



HAL
open science

Asservissement visuel d'un éclairage opératoire

Aurélien Gauvin

► **To cite this version:**

Aurélien Gauvin. Asservissement visuel d'un éclairage opératoire. Autre. Université d'Orléans, 2012. Français. NNT: 2012ORLE2019 . tel-01002339

HAL Id: tel-01002339

<https://theses.hal.science/tel-01002339>

Submitted on 6 Jun 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



UNIVERSITÉ D'ORLÉANS



ÉCOLE DOCTORALE SCIENCES ET TECHNOLOGIES

LABORATOIRE PRISME

THÈSE présentée par :

Aurélien GAUVIN

soutenue le : **5 juin 2012**

pour obtenir le grade de : **Docteur de l'université d'Orléans**

Discipline/ Spécialité : **Automatique et traitement du signal**

Asservissement visuel d'un éclairage opératoire

THÈSE dirigée par :

Youssef TOURE Professeur des Universités, Université d'Orléans
Estelle COURTIAL Maître de Conférences, Université d'Orléans

RAPPORTEURS :

Nicolas ANDREFF Professeur des Universités, Université de Franche-Comté
Jacques GANGLOFF Professeur des Universités, Université de Strasbourg

JURY :

Nicolas ANDREFF Professeur des Universités, Université de Franche-Comté
Jean-Pierre BREYSSE Directeur du service Recherche, MAQUET SA
Estelle COURTIAL Maître de Conférences, Université d'Orléans
Jacques GANGLOFF Professeur des Universités, Université de Strasbourg
Youssef TOURE Professeur des Universités, Université d'Orléans

Table des matières

Table des figures	vi
Remerciements	vii
Notations & Abréviations	ix
Introduction générale	1
1 Contexte et positionnement de l'étude	5
1.1 Le monde chirurgical d'aujourd'hui	6
1.2 Les systèmes d'éclairage pour la chirurgie	10
1.2.1 Les éclairages opératoires modernes	10
1.2.2 Les problèmes liés aux éclairages opératoires	14
1.2.3 Résumé des problèmes d'éclairage en chirurgie	18
1.3 Approches existantes pour la résolution des problèmes de mauvais placement	19
1.3.1 LightPilot	19
1.3.2 Eclairage autonome basé luminance	21
1.3.3 Voûte de Blin revisitée	22
1.4 Le concept d'éclairage opératoire asservi	24
1.4.1 Choix technologiques	24
1.4.2 Les objectifs de la thèse	26
1.5 Conclusion	27
2 Identification d'objets dans une scène chirurgicale	29
2.1 Introduction	30
2.1.1 Désignation "manuelle" d'une zone d'intérêt	30
2.1.1.1 Exemple 1 : Intervention superficielle localisée	30
2.1.1.2 Exemple 2 : Intervention superficielle étendue	32
2.1.1.3 Exemple 3 : Intervention en cavité	33
2.1.1.4 Exemple 4 : Intervention en cavité avec instruments multiples	34
2.1.2 Restriction du nombre d'objets à identifier	36
2.1.3 Processus de reconnaissance d'objet	39
2.2 Classification d'objets en milieu chirurgical	40
2.2.1 L'attribut "forme"	40
2.2.2 L'attribut "couleur"	49

2.2.2.1	Les espaces colorimétriques	50
2.2.2.2	Méthodes de reconnaissance basées couleur	52
2.2.3	L'attribut "mouvement"	57
2.3	Conclusion sur les attributs forme, couleur et mouvement	64
3	Fusion d'attributs et désignation de la zone d'intérêt	67
3.1	Fusion d'attributs	68
3.1.1	Définitions et généralités	68
3.1.2	Cadres théoriques de la fusion	72
3.1.2.1	Fusion par vote	72
3.1.2.2	Fusion dans le cadre probabiliste	73
3.1.2.3	Fusion floue	74
3.1.3	Fusion dans le cadre évidentiel	76
3.1.3.1	Etape 1 : l'étape crédale	77
3.1.3.2	Etape 2 : l'étape pignistique	85
3.2	Définition de la zone d'intérêt	88
3.2.1	Introduction à la théorie de la décision	88
3.2.1.1	Prise de décision en environnement incertain	89
3.2.1.2	Prise de décision en environnement risqué	93
3.2.2	Désignation de la zone d'intérêt	95
3.3	Conclusion	99
4	Asservissement visuel d'un éclairage opératoire	101
4.1	Généralités sur l'asservissement visuel	102
4.1.1	Classification des asservissements visuels	102
4.1.1.1	Position et nombre de caméra(s)	102
4.1.1.2	Asservissement direct ou indirect	103
4.1.1.3	Critère de Sanderson et Weiss	104
4.1.2	Approches de commandes	107
4.1.3	Application de l'asservissement visuel en milieu médical	108
4.1.3.1	Asservissement visuel sous imagerie X et IRM	108
4.1.3.2	Asservissement visuel échographique	108
4.1.3.3	Asservissement visuel endoscopique	110
4.1.3.4	Application de l'asservissement visuel à un éclairage opératoire	112
4.1.4	Les différentes étapes de mise en œuvre d'un AV2D	113
4.1.4.1	Etape 1 : le changement de repère	113
4.1.4.2	Etape 2 : le modèle de projection	114
4.1.4.3	Etape 3 : métrique du repère image	114
4.1.5	Calcul de la loi de commande	115
4.1.5.1	Notion de fonction de tâche	115
4.1.5.2	La matrice d'interaction L_s	115
4.1.5.3	Expression de la loi de commande	117
4.2	Problèmes de continuité de l'information	120
4.2.1	Les mesures de similarité	122

4.2.1.1	La mutuelle information	124
4.2.1.2	Comparaison des mesures de similarité	126
4.2.1.3	Perturbation du fragment	129
4.2.1.4	Essais de suivi d'une zone d'intérêt	131
4.2.2	Prédiction de la localisation du fragment	132
4.2.2.1	Filtre de Kalman	133
4.2.2.2	Filtre de Kalman étendu - EKF	134
4.2.2.3	Unscented Kalman Filter - UKF	135
4.2.2.4	Application du filtre de Kalman	137
4.3	Conclusion	139
5	Réalizations techniques & Résultats	141
5.1	Prototype d'un éclairage opératoire motorisé	142
5.1.1	Motorisation de l'arceau	142
5.1.2	Acquisition d'image	145
5.2	Mise en œuvre de la désignation de la zone d'intérêt	148
5.2.1	Reconnaissance de la classe instrument	148
5.2.2	Reconnaissance des autres classes	157
5.2.3	Désignation de la zone d'intérêt	166
5.3	Asservissement visuel de l'éclairage opératoire	173
5.3.1	Synthèse de la commande	173
5.3.2	Bilan des résultats expérimentaux	178
5.4	Conclusion	181
	Conclusion et perspectives	183
	Annexes	185
A	Evolution des éclairages opératoires	187
A.1	Caractéristiques d'un éclairage opératoire	189
B	Robots chirurgicaux hors étude	199
B.1	Les systèmes actifs	201
B.2	La télé-chirurgie mini-invasive	204
B.3	Les applications passives	209
B.4	Les systèmes semi-actifs et synergétiques en chirurgie	210
C	La chirurgie mini-invasive	213
	Bibliographie	241

Table des figures

1.1	Les catégories chirurgicales.	6
1.2	Evolution des techniques chirurgicales.	7
1.3	Différentes voies d'abords pour le traitement d'une fracture de l'humérus.	8
1.4	Application des champs stériles autour de la région à opérer.	9
1.5	Exemples illustrant le manque de lumière dans les interventions chirurgicales en superficie.	9
1.6	Exemples illustrant le manque de lumière dans les interventions chirurgicales en cavité.	9
1.7	Illustration du principe pour la réorientation et le repositionnement.	12
1.8	Principe d'un pontage coronarien par chirurgie à théâtre ouvert.	16
1.9	Principe d'une phlébectomie.	18
1.10	Module LightPilot développé par la société Berchtold.	20
1.11	Concept développé par [Spa07]	22
1.12	Système d'éclairage opératoire plafonnier [Pal11]	23
1.13	Processus du concept d'éclairage autonome et intelligent.	27
2.1	Isolation des objets dans le cas d'une chirurgie de l'épaule.	31
2.2	Suture de l'abdomen pratiqué sur un chien.	32
2.3	Partie d'une intervention chirurgicale se déroulant sans utilisation d'instruments chirurgicaux.	33
2.4	Difficulté de désigner la zone d'intérêt à l'aide uniquement des instruments chirurgicaux.	34
2.5	Variabilité des champs stériles (dimension, forme, ouverture, couleur, matériaux).	37
2.6	Evolution de la couleur d'une goutte de sang au cours du temps.	38
2.7	Exemple de descripteurs de forme (source : [PI97])	40
2.8	Perturbations appliquées à l'image d'origine.	42
2.9	Exemple de détection d'objets par segmentation.	43
2.10	Exemple de signature de contour (source : [GGMT10])	43
2.11	Exemple de détection de contours.	45
2.12	Exemple de détection de la classe instrument par transformée de Hough.	46
2.13	Exemple de contours actifs.	48
2.14	Représentation du sang dans différents espaces colorimétriques.	50
2.15	Exemple de classification par les séparateurs à vaste marge.	56
2.16	Exemple de flot optique à maillage régulier.	60
2.17	Descripteurs SURF.	63
3.1	Différents types de fusion au sens de Dasarathy (source : [Mar05]).	72

3.2	Fonction d'appartenance triangulaire des variables linguistiques "diamètre de cavité" et "luminance observée".	76
3.3	Représentation ensembliste des cadres mathématiques en fusion de données.	77
3.4	Exemple de classification par les séparateurs à vaste marge.	81
3.5	Exemple de surface représentative de l'intérêt que porte le chirurgien selon le temps opératoire et la situation opératoire.	89
3.6	Mise en évidence de la variabilité de l'intérêt de la classe sang selon le temps opératoire et la situation opératoire.	90
3.7	Décomposition de l'image courante en loterie.	91
3.8	Exemple de loterie sélectionnée selon différents critères ($ \mathcal{L} =5$ pixels, pas de recherche définie à 10pixels).	92
3.9	Exemple de loteries sélectionnées selon différents critères ($ \mathcal{L} =15$ pixels, pas de recherche définie à 20pixels).	92
3.10	Illustration du principe de sélection d'une loterie.	93
3.11	Sélection des loteries aux interfaces.	95
3.12	Sélection des loteries aux interfaces.	96
3.13	Illustration du principe de sélection de la zone d'intérêt.	98
3.14	Densité obtenue par estimateur à noyaux pour un jeu de données simulées.	99
4.1	Classification des différentes approches en asservissement visuel (source : [KC02])	102
4.2	Principe de l'asservissement visuel 3D	105
4.3	Principe de l'asservissement visuel 2D	105
4.4	Principe de l'asservissement visuel $2D^{1/2}$	106
4.5	Détermination de la pointe des instruments (source [DNdM07]).	111
4.6	Extraction d'un fragment d'image représentatif de la zone d'intérêt.	120
4.7	Illustration de transformations géométriques appliquées à un élément de référence.	121
4.8	Erreur de désignation pour une procédure d'extraction de lambeau de peau entraînant un mauvais suivi.	122
4.10	Erreur de similarité pour chacune des mesures.	128
4.11	Résumé des résultats pour le choix de la mesure de similarité.	129
4.12	Exemple de fragment bruité.	130
4.13	Influence des paramètres sur les mesures.	132
4.14	Résumé des résultats pour le choix de la mesure de similarité.	133
4.15	Prédiction de la position du fragment par filtrage de Kalman.	138
4.16	Variation de l'erreur entre position fournie par mesure de similarité et celle prédite par Kalman.	138
4.17	Variation de l'erreur entre position fournie par mesure de similarité et celle prédite par EKF.	139
5.1	Illustration du principe pour la réorientation et le repositionnement.	142
5.2	Vue globale des modifications apportées sur les articulations #5 et #6.	144
5.3	Prototype de l'étude.	144
5.4	Différentes techniques pour la diminution du contraste.	147
5.5	Reconnaissance de la classe instrument.	149

5.6	Histogramme représentatif de la classe instrument.	150
5.7	Reconnaissance de la classe instrument avec filtrage passe-haut.	151
5.8	Mouvement détecté lors de l'utilisation d'un instrument.	152
5.9	Pulsation d'une veine (aucun mouvement de la part du chirurgien).	152
5.10	Histogramme de l'attribut mouvement.	154
5.11	Isolation des mouvements de fortes amplitudes (retrait de la tête du chirurgien et mauvais appariement).	154
5.12	Comparaison des différentes règles de combinaison pour la classe instrument pour le cas d'une pose d'écarteur.	155
5.13	Comparaison des différentes règles de combinaison dans la classe instrument pour le cas d'une dissection.	156
5.14	Autres séquences issues d'essais en bloc opératoire.	157
5.15	Modélisation de l'histogramme de la classe sang pour le canal H.	158
5.16	Exemple de reconnaissance grâce à l'attribut couleur.	160
5.17	Modification du schéma de création des masses.	161
5.18	Exemple de reconnaissance de la classe sang pour différentes expositions.	161
5.19	Influence d'un ou deux éclairages opératoires supplémentaires sur la reconnaissance.	165
5.20	Détermination des coefficients α_1	167
5.21	Séquence opératoire déterminée par le système.	168
5.22	Conditionnement de la séquence opératoire au cas $\{S, C, S\}$	168
5.23	Densité obtenue au cours d'essais en clinique par estimateur à noyaux.	169
5.24	Sélection de la zone d'intérêt finale par moyenne glissante et par estimateur à noyaux.	170
5.25	Robustesse des méthodes face aux aléas.	171
5.26	Temps de retard en fonction de la taille de calcul.	171
5.27	Variation de l'erreur pour différentes tailles de fenêtre.	172
5.28	Résultats du processus de désignation pour une intervention chirurgicale.	172
5.29	Synopsis de la boucle de régulation.	174
5.30	Schéma mécanique de l'éclairage opératoire.	175
5.31	Comportement du système lorsque la cible est proche ou éloignée.	176
5.32	Insertion d'un bloc correcteur dans l'architecture de la loi de commande.	176
5.33	Erreur obtenue avec différents profils de correcteur.	178
5.34	Erreurs entre la zone d'intérêt désignée et réelle.	179
5.35	Comparaison entre l'activité du chirurgien et celle du système.	180
5.36	Mise en évidence des comportements "optimistes" et "prudents".	180
5.37	Résultats sur le déplacement de l'effecteur en conditions réelles.	181
A.1	Différents éclairages opératoires non manipulables par l'équipe chirurgicale.	189
A.2	Luminance pour différentes sources lumineuses	190
A.3	Modèle de Phong.	191
A.4	Diagramme CIExy 1931.	192
A.5	Différence d'aspect pour deux températures de couleur différentes.	192
A.6	I.R.C. pour différentes sources lumineuses	193
A.7	Participation de toutes les sources lumineuses à la création de la tache lumineuse.	195
A.8	Exemple de solutions pour l'augmentation du diamètre de tache.	196

A.9	Illustration de l'impact du spectre d'émission sur la vision qu'a le chirurgien de la plaie opératoire	196
A.10	Exemple de solutions pour l'augmentation du diamètre de tache.	197
B.1	Résumé sur les limites homme-machines ([LWCWN08])	209
C.1	Exemples de chirurgie mini-invasive.	215

Remerciements

J'exprime tout d'abord la plus grande gratitude à Nicolas Andreff de l'Université de Franche-Comté et Jacques Gangloff de l'Université de Strasbourg d'avoir bien voulu donner de leur temps et de me faire l'honneur de rapporter mon travail de thèse. Qu'ils trouvent ici l'expression de mes sentiments les plus chaleureux. Merci également à Jean-Pierre Breysse, Estelle Courtial, Serge Exshaw et Youssef Touré pour m'avoir donné l'opportunité d'étudier ce sujet passionnant tant au niveau scientifique qu'industriel. Plus particulièrement Estelle Courtial pour avoir tenté de m'apprendre ce qu'est la rigueur scientifique et Jean-Pierre Breysse pour son énergie et sa croyance inébranlable en la réussite de ce projet.

Je tiens à remercier les membres de l'équipe de l'INRA de Tours avec qui j'ai eu la chance de travailler et qui m'ont permis de tester le prototype dans des conditions réelles d'utilisation. Je tiens à remercier également l'ensemble des CHU, CHI et cliniques avec lesquels j'ai eu l'occasion de travailler. Ceci s'adresse notamment ceux de Villefranche-sur-Saône et de Bruxelles qui m'ont accueilli sur de longues périodes et fait découvrir le monde chirurgical.

Une thèse est une étape de vie, courte au final, où je n'ai cessé d'apprendre et où j'ai eu la chance de côtoyer des personnes formidables. Je tiens à adresser mes plus vifs remerciements à ceux qui ont de près ou de loin participé à ce projet.

Je remercie particulièrement l'équipe de recherche de l'entreprise MAQUET SAS au sein de laquelle j'ai pu travailler durant ces trois années notamment :

- Céline Delporte pour sa "rigueur scientifique" (inutile d'en dire plus j'entends d'ici son rire) ;
- Jérôme Désouches pour son tact légendaire et sa voiture (qu'il fallait ouvrir par le coffre pour accéder à la porte arrière pour enfin ouvrir la porte avant) ;
- Muriel Fidalgo pour sa bonne humeur permanente, ses délicieux gâteaux et son aptitude à rester zen et diplomate en toutes circonstances ;
- Bertrand Guillemillot pour sa capacité à générer une dizaine d'idées à la seconde et à simuler une intervention chirurgicale en découpant une calzone ;
- David Le ber pour son regard toujours neuf sur ce projet et pour m'avoir encouragé avec des "Ca ne marchera jamais !" ;
- Gabriel Peyras pour m'avoir répété un nombre incalculable de fois, avec une infinie patience, la différence entre luminance, réflectance et éclairement.

Je tiens à saluer également le BE mécanique, optique et électronique ainsi que les chefs de projet pour leurs réponses à mes questions parfois ésotériques mais aussi toute l'équipe de production pour m'avoir

accueilli aussi chaleureusement, ainsi que Messieurs Blanche et Salem pour leur soutien.

Je remercie tous ceux de l'Université d'Orléans notamment les gais lurons :

- Romain Agogué pour le sérieux dont il fait preuve lorsqu'il s'agit de science ;
- Jean-François Collumeau, Julien Moothoo, Dorra Nouri et Hamid Ouali a qui je souhaite bon courage pour leur thèse qui j'en suis sûr seront couronnées de succès ;
- Ali Damen pour sa contribution à ces travaux et la détermination dont il a fait preuve ;
- Aurélie Telmar pour son enthousiasme et sa joie de vivre toujours d'un grand réconfort pendant les moments de doute ;
- Romain Rieger pour avoir été un ami des plus fidèles avec qui j'ai pris plaisir à parler sciences et pour qui j'ai la plus grande estime ;
- Audrey Wendling pour avoir gardé son calme en toute occasion et m'avoir initié au monde mystique des matériaux composites.

Je tiens à remercier l'ensemble des stagiaires, personnel administratif, doctorants, post-doctorants, Maîtres de Conférence et Professeurs de l'Université d'Orléans, de l'IUT de Bourges et de l'ENSI de Bourges que je n'ai pas eu l'occasion de citer ici. Une attention spéciale pour Frédéric Kratz à qui je porte la plus grande des reconnaissances.

Mes remerciements vont également aux amis un peu perdus de vue durant ces années mais à qui j'ai souvent pensé entre autres Chen Jun (pour ces fous rires pour lesquels je ne t'ai jamais remercié), Lolo (LE super coloc' avec son célèbre "Peter Parker ..."), les Kronks ("bande de vautours", ps : "Amaury, j'ai caché les clé de ta voiture, c'est ton tour"), ...

Je terminerai cette parenthèse en remerciant ma famille sans qui il ne m'aurait été possible de terminer cette thèse. Merci à ma mère pour son courage et son soutien indéfectible ainsi qu'à ma femme qui m'a supporté et encouragé tout au long de cette thèse et qui a su mettre dans mon coeur 586 étoiles.

Notations & Abréviations

Chapitre 1

Notations	Désignation
$\# \cdot$	Numéro de l'articulation de l'éclairage opératoire
$\{\cdot, \cdot, \cdot\}$	Séquence opératoire correspondant à la situation opératoire en début, milieu et fin d'intervention chirurgicale
E .	Eclairement [lux] qui peut soit renvoyé (E_e) soit reçu (E_r)
ρ	Albédo [.]

Abréviations	Désignation
AIM	Automatic Illumination Management
IHM	Interface Homme Machine
IRC	Indice de rendu des couleurs [%]

Chapitre 2

Notations	Désignation
$\{\cdot\}$	Hypothèse intervenant dans un cadre de discernement
a	Scalaire
\vec{a}	Vecteur
\bar{a}	Elément contraire de a
$a * b$	Convolution de a et b
D	Espace occupé par un objet
\mathcal{D}	Décision retenue au terme d'une analyse
\mathcal{H}	Fonction d'Heavyside
h_o	Masque o utilisé pour la détection de contours
$I(x, y)$	Intensité du pixel de coordonnées (x, y)
l	Image
L_i	Norme i
$m_{p,q}$	Moment de Hu centré
$M_{p,q}$	Moment de Hu
$p(\vec{a} b)$	Probabilités conditionnelles de \vec{a} sachant la réalisation de b
s^-, s^+	Bornes inférieure et supérieure d'un intervalle

ΔI	Variation d'intensité entre deux images successives
$\eta_{i,j}$	Moment de Hu centré normalisé
$\vec{\nabla} I$	Gradient de l'image I
$ \vec{\nabla} I $	Module du gradient de l'image I
κ	Seuil appartenant à l'intervalle $[0, 1]$
μ	Moyenne d'une distribution
π	Pondération intervenant dans l'algorithme EM
ϕ_i	Moment dérivé de η à l'ordre i
ρ_i	Module du point i
σ	Ecart-type d'une distribution
Σ	Matrice de covariance
Θ	Cadre de discernement contenant les classes $\{\{a\}, \{b\}, \{c\}, \dots\}$

Abréviations

AIC	Akaike's Information Criterion
AV2D	Asservissement visuel 3D
AV3D	Asservissement visuel 3D
CGMM	Modèle à mélange de gaussiennes contraint
CIE-xyz	Espace colorimétrique
DoG	Difference of Gaussians
DWT	Transformée en ondelettes discrètes
EM	Espérance Maximisation
GFD	Descripteurs de Fourier génériques
GMM	Modèle à mélange de gaussiennes
GNG	Growing Neural Gas
HSV	Espace colorimétrique (hue, saturation, value)
HSI	Espace colorimétrique (hue, saturation, intensité)
i.i.d.	indépendamment et identiquement distribuée
$L^*_a*b^*$	Espace colorimétrique
LVQ	Learning Vector Quantization
nRGB	Espace colorimétrique (rouge, vert, bleu normalisé)
MHI	Motion History Image
MLPNN	Multi Layer Perceptron Neural Network
MSER	Maximally Stable Extremal Regions
PCA	Analyse en composante principale
RGB	Espace colorimétrique (rouge, vert, bleu)
ROC	Receiver Operating Characteristic
SF	Série de Fourier
SIFT	Scale Invariant Feature Transform
SOM	Self-Organizing Map
SURF	Speeded Up Robust Features
SVM	Support Vector Machine
Y_C, C_b	Espace colorimétrique

WPT Transformée par paquet d'ondes

Chapitre 3

Notations	Désignation
\oplus	Opérateur de combinaison conjonctive
2^Θ	Power set contenant les disjonctions du cadre de discernement <i>Theta</i>
$BetP(a)$	Probabilité pignistique de l'hypothèse <i>a</i>
D^{Theta}	Hyper-power set contenant les disjonctions \cap et \cup du cadre <i>Theta</i>
$\mathbb{E}(a)$	Espérance mathématique <i>a</i>
<i>g</i>	Gain potentiel
\mathbb{I}	Fonction indicatrice
\mathcal{L}	Loterie représentative d'un fragment de l'image
\mathcal{L}	Loterie représentative de la zone d'intérêt
\mathbb{L}	Ensemble de toutes les loteries
$m(\{a\})$	Masse représentative de la croyance en la classe <i>a</i>
$m^{\Theta_1}(\{a\})$	Masse représentative de la croyance en la classe <i>a</i> définie pour le cadre de discernement Θ_1
$\mathbb{P}(a)$	Probabilité de l'événement <i>a</i>
\mathcal{P}	Précision de la reconnaissance
\hat{p}	Distribution de probabilité estimée
\mathcal{R}	Rappel de la reconnaissance
S^Θ	Super-power set de loi $(\Theta, \cup, \cap, c(\cdot))$
$v_{i,k}$	Vote exprimé par l'expert (ou source) <i>i</i> concernant l'hypothèse <i>k</i>
\mathfrak{U}	Utilité du gain
β	Affaiblissement faisant appel à la précision et au rappel
\mathcal{E}	Image temporaire
$\mu_b(a)$	Fonction d'appartenance floue de l'objet <i>a</i> par rapport à l'objet <i>b</i>
θ_a	Classe d'objet <i>a</i>

Abréviations	Désignation
DFPM	Data Fusion Process Model
JDL	Joint Directors of Laboratories
OODA	Observation, Orientation Decision Action
SA	Situation Awareness Model

Chapitre 4

Notations	Désignation
\dagger	Pseudo inverse de Moore-Penrose
d_u, d_v	Taille d'un pixel suivant les axes \vec{u} et \vec{v}
<i>e</i>	Erreur exprimée dans l'image

f	Distance focale de la caméra
\mathbf{K}	Matrice des paramètres intrinsèques
k	Facteur d'échelle
L_s	Matrice d'interaction
${}^c\mathbf{M}_o$	Matrice homogène permettant le passage du repère R_0 au repère R_c
P_a	Point exprimé dans le repère a
R_a	Repère a
${}^c\mathbf{R}_o$	Ensemble des rotations permettant le passage du repère R_0 au repère R_c
r	Pose du système robotique
$ \mathcal{R} $	Cardinal d'un histogramme
s	Configuration des primitives dans l'image
s^*	Configuration désirée
${}^c t_o$	Ensemble des translations permettant le passage du repère R_0 au repère R_c
λ	Coefficient
τ	Torseur cinématique
θ	Gauchissement des pixels
\vec{u}, \vec{v}	Axes principaux de l'image

Abréviations

$2D^{\frac{1}{2}}$	Asservissement visuel hybride
d2D/dt	Asservissement visuel basé sur un champ de vitesse
EKF	Filtre de Kalman étendu
GPC	Generalized Predictive Controller
IBVS	Image Based Visual Servoing (équivalent à AV2D)
IRM	Imagerie par Résonance Magnétique
KF	Filtre de Kalman
LMI	Linear Matrix Inequalities
LQG	Linear Quadratic Gaussian
PBVS	Position Based Visual Servoing (équivalent à AV3D)
PRC	Prototype Repetitive Controller
PTZ	Pan Tilt Zoom
SDP	Semi-Definite Programming
UKF	Filtre de Kalman non parfumé

Introduction générale

Les travaux présentés dans ce manuscrit ont été réalisés dans le cadre d'une thèse CIFRE (Convention Industrielle de Formation par la REcherche), soutenue par l'entreprise MAQUET SAS et en collaboration avec le laboratoire PRISME de l'Université d'Orléans.

Au cœur des salles de chirurgie, les éclairages opératoires sont dévolus à une unique tâche : apporter au chirurgien les conditions d'éclairage nécessaires pour accomplir son geste. Cependant, bien que leurs performances répondent de plus en plus aux exigences des chirurgiens notamment en matière de qualité de tâche de lumière, il est encore aujourd'hui impératif de réorienter l'éclairage opératoire à chaque fois que la zone traitée par le chirurgien se déplace. Ce réajustement est effectué manuellement soit par le chirurgien, soit par une infirmière. Dans le cas où le chirurgien exécute cette manipulation, nous assistons alors à :

- des pertes de temps et d'énergie dues au déplacement de l'éclairage ;
- une interruption du geste opératoire impliquant une perte de concentration qui génère à son tour une perte de temps supplémentaire et peut altérer la qualité du geste opératoire ;
- un passage de gestes de faibles amplitudes à des mouvements de très fortes amplitudes puis retour à des mouvements de faibles amplitudes ;
- une augmentation des fautes d'asepsie qui peuvent être à l'origine de problèmes infectieux pour le patient ;
- une impossibilité d'atteindre la position voulue de l'éclairage compte-tenu de sa cinématique et de l'espace de travail du chirurgien.

Dans le cas où la manipulation est réalisée par l'infirmière, d'autres problèmes apparaissent :

- une réorientation de l'éclairage inadéquate car cette dernière ne peut visualiser la zone d'intérêt compte-tenu des règles d'hygiène ;
- des tensions entre le chirurgien et la panseuse en raison des mêmes règles ;
- une détérioration des équipements du fait des chocs répétés avec l'éclairage opératoire lors de son déplacement ;
- une augmentation des fautes d'asepsie.

Ceci explique pourquoi 88% du personnel chirurgical est insatisfait des éclairages opératoires ([KMJ11]). Il existe donc une réelle attente pour améliorer le déplacement de ce type d'équipement. De plus, du fait de ces difficultés, le réajustement de la position et de l'orientation de l'éclairage n'est faite que trop peu de fois comparativement à ce qui serait nécessaire pour garantir un éclairage suffisant pour le chirurgien.

Les approches développées par le passé consistent à désigner le lieu à éclairer grâce à l'utilisation d'une interface homme-machine ou d'une pastille stérile. De tels systèmes n'apportent en réalité qu'une réponse partielle aux besoins du corps chirurgical en matière d'éclairage de la scène dans la mesure où ils sont basés sur l'utilisation d'informations explicites. Celles-ci impliquent un arrêt du geste opératoire ce qui nuit à l'intervention (rallongement de l'anesthésie, faute d'asepsie, etc). Ce travail de recherche a pour objectif d'apporter une solution au problème de déplacement de l'éclairage opératoire durant une opération chirurgicale.

Pour atteindre cet objectif, nous proposons un éclairage opératoire "intelligent" et "autonome". "Intelligent" afin d'être capable de déterminer de lui-même la zone d'intérêt du chirurgien à éclairer. "Autonome" afin de se déplacer seul et de ne pas interférer avec les membres de l'équipe chirurgicale. Il s'agit donc de lever les trois verrous scientifiques suivants :

1. localiser l'endroit où travaille le chirurgien ;
2. déplacer le système jusqu'à cet endroit ;
3. assurer les performances du système malgré la variabilité des interventions chirurgicales.

Ne disposant pas d'un éclairage opératoire asservi pour la validation des approches proposées en conditions réelles (bloc opératoire) un prototype a été conçu en parallèle des travaux de recherches théoriques. Il a été développé dans le souci de satisfaire les préoccupations industrielles : limitation du temps de calcul, embarquabilité du système, gestion des erreurs, respect des normes, réglementations et éthiques propres au domaine chirurgical, etc.

Ce manuscrit s'articule suivant le plan ci-dessous :

- **Le premier chapitre est dédié à la compréhension du monde chirurgical, du besoin des chirurgiens en matière d'éclairage et des problèmes inhérents aux éclairages opératoires.** Après une brève présentation sur ces équipements, nous montrerons à travers deux exemples concrets que les conditions d'éclairage sont directement liées au placement de l'éclairage et qu'il est nécessaire, pour obtenir une visibilité optimale de la scène chirurgicale, de procéder à des réorientations fréquentes de l'éclairage. Nous exposerons les solutions existantes et mettrons en avant les problèmes dus aux interactions avec le corps chirurgical. Enfin, nous présenterons le concept d'éclairage opératoire asservi permettant de répondre à ce problème d'"optimalité".
- **Le chapitre 2 concerne l'identification des indices visuels nécessaires pour la désignation de la zone d'intérêt.** Nous illustrerons, à travers des exemples de désignation manuelle, que le processus de désignation de la zone d'intérêt basé sur une reconnaissance de certains objets dans la scène opératoire peut être automatisé. Consécutivement nous analyserons les techniques de reconnaissance d'objets existantes et mettrons en évidence qu'une étape de fusion et de décision sont nécessaires pour parvenir à localiser la zone d'intérêt du chirurgien.
- **Le chapitre 3 portera sur le processus de fusion et la désignation de la zone d'intérêt.** Nous présenterons les différents schémas de fusion et nous identifierons le plus adéquat pour notre

problématique. Dans un second temps, nous verrons comment à l'aide de cette étape de fusion allée à la théorie de la décision nous pouvons cibler la zone d'intérêt.

- **Le chapitre 4 sera consacré à la loi de commande permettant de traduire une désignation en déplacement interprétable pour le système.** Après un rappel des différents types d'asservissements et de leur application en milieu médical ou chirurgical, nous présenterons l'approche d'asservissement visuel utilisée. Les différentes étapes de calcul seront détaillées ainsi que les mécanismes spécifiques aux problèmes de continuité de la désignation permettant de résoudre le problème d'insuffisance d'information.
- **Le chapitre 5 est dédié aux réalisations techniques et aux résultats pratiques.** Nous commenterons les performances de la stratégie adoptée dans le cas d'essais effectués en bloc opératoire. Nous montrerons l'intérêt que représentent ces travaux ainsi que les efforts à fournir pour améliorer le comportement du système.

Chapitre 1

Contexte et positionnement de l'étude

Sommaire

1.1	Le monde chirurgical d'aujourd'hui	6
1.2	Les systèmes d'éclairage pour la chirurgie	10
1.2.1	Les éclairages opératoires modernes	10
1.2.2	Les problèmes liés aux éclairages opératoires	14
1.2.3	Résumé des problèmes d'éclairage en chirurgie	18
1.3	Approches existantes pour la résolution des problèmes de mauvais placement	19
1.3.1	LightPilot	19
1.3.2	Eclairage autonome basé luminance	21
1.3.3	Voûte de Blin revisitée	22
1.4	Le concept d'éclairage opératoire asservi	24
1.4.1	Choix technologiques	24
1.4.2	Les objectifs de la thèse	26
1.5	Conclusion	27

Afin de mieux comprendre les contraintes du milieu opératoire et le rôle des éclairages opératoires, nous allons proposer dans un premier temps un tour d'horizon des techniques chirurgicales. Il sera mis en exergue l'importance que revêtent les éclairages opératoires dans ce contexte, les moyens d'éclairages actuels et, au travers de scénarii opératoires, les difficultés liées à ce type de matériel notamment en termes de placement. Dans un second temps, nous présenterons les solutions existantes et montrerons qu'elles sont souvent inadaptées. Enfin, nous proposerons une solution à cette problématique : le concept d'un éclairage opératoire asservi visuellement. Une attention particulière sera portée sur le problème d'éclairage d'une cible dans le cadre chirurgical, le cahier des charges et les contraintes spécifiques à ce milieu.

Les informations concernant le déroulement des opérations, les différentes définitions relatives à l'éclairage et les tâches accomplies par le personnel chirurgical ont été recueillies au cours de visites d'établissements hospitaliers, pendant et après des interventions chirurgicales.

1.1 Le monde chirurgical d'aujourd'hui

En France, chaque année quatre millions d'interventions chirurgicales sont pratiquées ([IRD98]). Parmi ces opérations, il est possible de distinguer deux familles de chirurgie : la chirurgie majeure où le caractère traumatique pour le patient est important et la chirurgie mineure communément appelée chirurgie ambulatoire (figure 1.1).

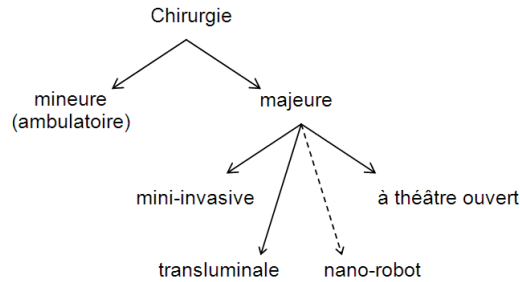


FIG. 1.1 – Les catégories chirurgicales.

Les interventions mineures ne nécessitent en général qu'un seul chirurgien et sont, dans la plupart des cas, dévolues à des actions superficielles, localisées et de courte durée. Compte tenu des dimensions et de la profondeur de la région à traiter, il n'est généralement pas nécessaire de déplacer l'éclairage pour ces types d'acte. Nous ne traiterons donc pas de ces cas.

La chirurgie majeure peut être subdivisée en quatre sous groupes tels que présentés en figure 1.2. Nous pouvons distinguer les opérations à théâtre ouvert qui seront à l'avenir de plus en plus supplantées par les interventions mini-invasives dont l'impact sur les tissus est moindre. La chirurgie transluminale consiste à opérer en passant par les voies naturelles sans dissection du derme. En raison de la difficulté technique de ce type d'intervention, la chirurgie transluminale reste à ce jour très peu pratiquée. Les nano-robots sont quant à eux en phase d'étude, encore non applicables cliniquement. La chirurgie transluminale et les nano-robots constituent les techniques chirurgicales de demain et ne seront pas concernées par ce travail de recherche.

Comme le montre la figure 1.2, la chirurgie mini-invasive ne nécessite que quatre incisions. Celles-ci servent aux passages des instruments, du tube d'insufflation CO₂ et de la caméra endoscopique. En l'occurrence, les conditions d'éclairage sont garanties par la source lumineuse intégrée à la caméra. Il n'y a donc pas besoin d'un éclairage opératoire pour ce type d'intervention. En conséquence, il ne sera pas traité dans cette étude des opérations mini-invasives. Le lecteur pourra cependant se référer à l'annexe C pour de plus amples informations sur ce sujet. Pour les mêmes raisons, il ne sera pas non plus discuté des opérations transluminales. Enfin, aucune information ne sera donnée dans ce manuscrit quant aux besoins d'éclairage pour la partie nano-robotique, aucun protocole clinique n'étant disponible.

Par la suite, nous nous focaliserons donc sur les interventions à théâtre ouvert qui représentent 82% des interventions totales réalisées en France (selon une étude marketing de TforG Euro Surgical Database réalisée en 2011) dont nous allons présenter le déroulement, les problèmes d'éclairage associés et les solutions apportées par les éclairages opératoires.

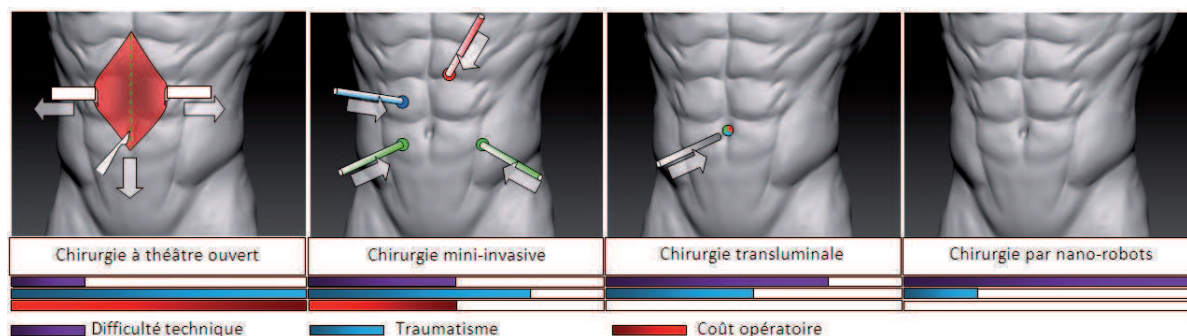


FIG. 1.2 – Evolution des techniques chirurgicales (les graphiques n'ont qu'une fonction indicative notamment pour le cas de la chirurgie transluminale qui est ici un cas particulier nécessitant une incision préliminaire; le coût pour la chirurgie transluminale et les nano-robots n'est pas encore connu).

Les interventions à théâtre ouvert

La chirurgie à théâtre ouvert, également appelée chirurgie conventionnelle, consiste à procéder à une large incision permettant d'accéder facilement à l'élément à traiter (figure 1.2). Le déroulement de ce type de chirurgie suit les cinq activités per-opératoires suivantes :

1. anesthésie du patient ;
2. sélection d'une voie d'abord ;
3. désinfection du patient par application d'un produit bétadiné et pose de champs stériles ;
4. début de l'intervention :
 - incision principale de la région à opérer ;
 - traitement de l'élément cible ;
 - suture des tissus et pose d'un drain si nécessaire.
5. fin de l'intervention, retrait des champs stériles.

Remarque : *Ces activités forment un cadre générique pour les interventions à théâtre ouvert. Dans la mesure où chaque opération suit un schéma propre, il n'est donc pas possible de détailler précisément le déroulement de ces étapes. Cependant, nous proposons par la suite une description qualitative globale afin d'illustrer les disparités pouvant survenir.*

Les voies d'abord

Avant de procéder à l'intervention chirurgicale proprement dite, le chirurgien sélectionne selon la pathologie à traiter la voie d'abord. Également appelées voies opératoires, les voies d'abord définissent le chemin qu'empruntera le chirurgien pour accéder à l'élément cible. Sa sélection dépend de nombreux facteurs dont :

- la facilité d'accès à la zone ;
- la restriction de l'espace de travail ;
- le risque de toucher des éléments sensibles durant l'opération (nerfs, vaisseaux principaux) ;
- le matériel mis à disposition ;
- les lésions potentielles ;

- l'entraînement ainsi que la préférence du chirurgien ;
- l'évolution des techniques chirurgicales.

La figure 1.3, illustre pour une fracture du bras les différentes voies d'abord possibles. Dans cet exemple, l'intervention consiste à poser une plaque métallique en vue de traiter une fracture simple ou multiple de l'humérus [LC03]. Le chirurgien vient donc choisir la voie d'abord la plus adéquate au vue de la situation. Comme les pointillés le montrent, il est possible d'accéder à la fracture par quatre chemins différents donc par quatre voies d'abords différentes (latérale, médiale, postéro-médiale, postérieure).

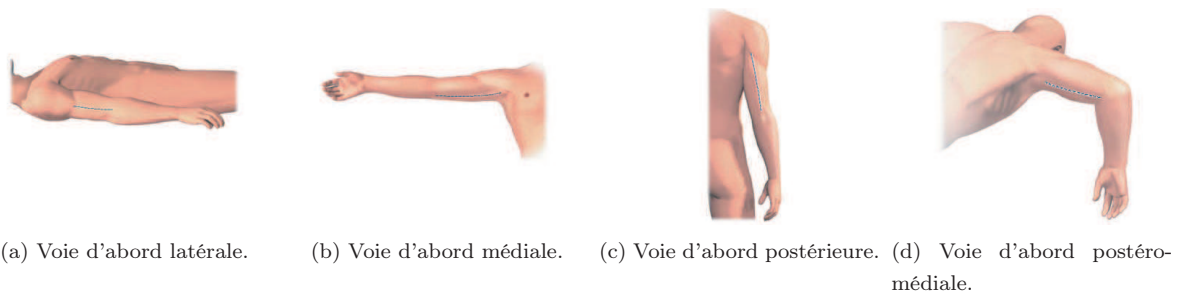


FIG. 1.3 – Différentes voies d'abords pour le traitement d'une fracture de l'humérus.

Remarque : *De cette diversité des voies d'abords possibles résulte une grande variabilité du type des tissus exposés. Nous verrons que cet élément revêtira une grande importance pour la suite du projet. En effet, en l'absence de connaissance sur le déroulement de l'intervention, il n'est pas possible de prédire sur le long terme quelles seront les caractéristiques des tissus observés (couleur, texture, forme, ...).*

Préparatifs de la scène opératoire

Une fois la voie d'abord sélectionnée, le chirurgien procède à la désinfection du patient grâce à de la bétadine dermique (povidone iodée). Ce produit permet de limiter temporairement les agressions bactériennes et ainsi diminuer la probabilité d'infection nosocomiale. Il est appliqué directement sur la peau (figure 1.4a). La couleur de la peau prend donc des teintes orangées plus ou moins prononcées selon l'origine du patient. Vient ensuite la pose des champs stériles. Ces tissus aux formes et aux couleurs variées isolent la région à opérer du reste du patient. Leur but est double : ils permettent d'une part d'éviter une contamination du site chirurgical et d'autre part de borner spatialement la région à opérer (figure 1.4b). En de rare cas, ils servent aussi à limiter les efforts appliqués à la peau pour éviter toutes déchirures. Le maintien des champs stériles est assuré par des instruments chirurgicaux (typiquement des ciseaux à clamper potentiellement en grand nombre (figure 1.6b)). Leur placement est fonction de la forme des champs stériles et des préférences du chirurgien.

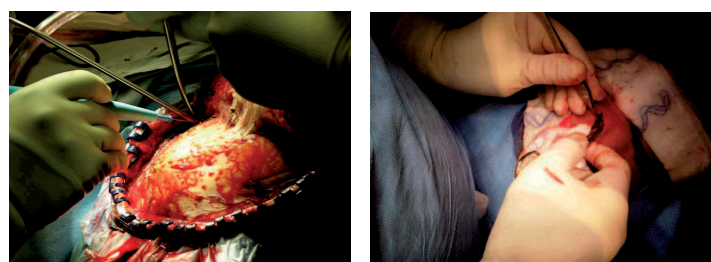
Une fois l'intégralité de ces étapes accomplies, le chirurgien incise la peau, puis par dissection des strates tissulaires successives, atteint l'élément à traiter. Le traitement appliqué est spécifique à la pathologie. De même les gestes chirurgicaux et le temps nécessaire pour les accomplir sont spécifiques au chirurgien en charge de l'opération. Ainsi à la difficulté technique de l'intervention s'ajoute celle liée à la lumière. En



(a) Application de bétadine iodée (source : site Dr Caputo).
 (b) Placement des champs stériles (source : site girlinscrubs).
 (c) Exemple de région stérile (source : site priestlakevet).

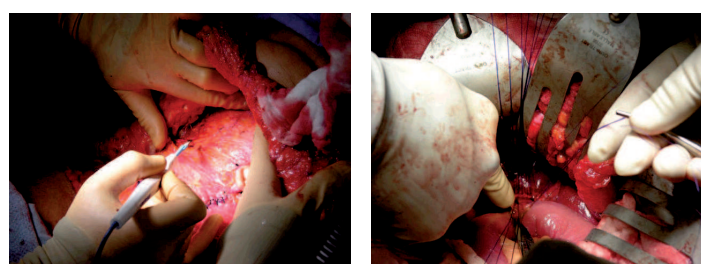
FIG. 1.4 – Application des champs stériles autour de la région à opérer.

effet, nous pouvons remarquer sur les figures 1.5 et 1.6 que, malgré l'utilisation d'un éclairage opératoire, les conditions d'éclairément ne sont pas parfaites tant pour des opérations en superficie qu'en cavité.



(a) Opération neurologique.
 (b) Suture superficielle.

FIG. 1.5 – Exemples illustrant le manque de lumière dans les interventions chirurgicales en superficie.



(a) Abdominoplastie.
 (b) Suture en cavité.

FIG. 1.6 – Exemples illustrant le manque de lumière dans les interventions chirurgicales en cavité.

En l'absence d'éclairage opératoire, il n'est pas possible pour le chirurgien d'accomplir l'opération malgré l'existence d'un éclairage ambiant. Ils sont donc indispensables pour la réalisation du geste opératoire. Ceci a conduit les industriels de ce milieu à imaginer des solutions permettant d'obtenir la tache lumineuse la plus puissante et évitant la formation d'ombres.

En résumé :

Nous avons vu qu'il existe différents types de chirurgie dont les contraintes et des besoins en éclairage différent. Dans le cadre de cette étude, nous nous limitons aux opérations majeures, et plus particulièrement à la chirurgie à théâtre ouvert. Nous avons présenté les activités per-opératoires, les voies d'abord ainsi que des exemples de problèmes d'éclairage rencontrés au cours d'une intervention pour différentes situations opératoires. Dans la partie suivante, nous nous attachons à décrire les grandes caractéristiques de l'éclairage opératoire et comment celui-ci doit être placé pour ne plus avoir ces problèmes de manque de lumière.

1.2 Les systèmes d'éclairage pour la chirurgie

1.2.1 Les éclairages opératoires modernes

Les anciens systèmes d'éclairage ne permettaient pas d'être intégrés directement dans les salles d'opération. Les projecteurs de trop grandes dimensions devaient être installés dans une salle indépendante située au-dessus du bloc opératoire, une verrière permettant la propagation du flux lumineux (cf. annexe A). Il faut attendre 1980 et l'avènement du "scialytique" (du grec skia/ombre et luein/détruire-dissoudre) pour voir un éclairage implanté au sein même d'un bloc opératoire. Bien que de nos jours obsolète, ce dispositif fut considéré comme étant l'une des avancées majeures du domaine. Son succès est dû à ses dimensions et ses performances notamment en termes d'émission infrarouge. En effet, l'élévation de la température plus raisonnable comparativement aux anciens dispositifs permettait aux chirurgiens de pouvoir travailler plus longtemps et ce, dans de meilleures conditions.

Ce type d'éclairage permet de mieux appréhender les besoins du chirurgien tout en orientant l'évolution des futurs éclairages. Aujourd'hui les éclairages opératoires tirent les enseignements du "scialytique" à savoir :

- être implantable au sein même du bloc ;
- être en mesure d'éclairer la zone opératoire tout au long de l'opération ;
- pouvoir se déplacer ;
- supprimer les émissions radiatives dans le domaine du proche infrarouge ;
- supprimer les ombres (notamment pour les actes en cavité) ;
- pouvoir modifier le diamètre de la tâche lumineuse.

Ces caractéristiques fondamentales sont séparables en deux catégories. La première concerne l'aspect "optique" où interviennent principalement l'éclairage, la luminance et le diamètre de la tâche lumineuse. La seconde porte sur la "mécanique" de l'éclairage opératoire notamment la cinématique et la maniabilité. D'autres grandeurs participant aux performances d'un éclairage opératoire ne seront pas présentées par la suite car trop éloignées du sujet d'étude. Néanmoins le lecteur pourra se reporter à l'annexe A.1 pour de plus amples informations concernant notamment les éléments se rapportant à la colorimétrie d'un éclairage opératoire (notion correspondant à la perception de la couleur et faisant intervenir l'Indice des Rendus des Couleurs et la température de couleur).

Eclairage et luminance

L'éclairage d'un plan est exprimé en lux et est caractérisé par le flux lumineux incident émis par l'éclairage opératoire rapporté à une unité de surface. Pour un éclairage opératoire l'éclairage peut atteindre 200klux. A titre de comparaison, celui du soleil est de 100klux loin devant celui des ampoules à incandescence qui n'est que de 200lux. Ce besoin de puissance lumineuse est motivé du fait de l'absorption de la lumière par les tissus. Il faut donc maximiser l'éclairage afin de compenser l'absorption sans pour autant éblouir le chirurgien. En conséquence, la puissance est limitée par réglementation à 160klux selon la norme IEC60601-1. Il y a donc un décalage entre la quantité de lumière émise par l'éclairage (E_e) et celle vue par le chirurgien (E_r). Cette fraction parvenant aux yeux du chirurgien s'appelle la luminance. Une définition plus rigoureuse est que la luminance au point de coordonnées (x, y) d'une surface réceptrice orientée dans la direction (m, p) est l'intensité par unité de surface apparente dans cette direction.

Ainsi plus le tissu est réfléchissant et plus la luminance est élevée. Cette propriété optique adimensionnelle du tissu s'appelle l'albédo (ρ). Elle se traduit par : $\rho = E_e \cdot E_r^{-1}$. Ainsi, pour un miroir, l'albédo sera de 1 (réflexion spéculaire) alors que pour un matériau très diffusant, l'albédo sera proche de 0. Dans ce dernier cas on parle de matériaux Lambertiens ou orthotropes. Ils ont pour propriété de réémettre la lumière dans toutes les directions de manière uniforme. Dans la nature, rares sont les matériaux parfaitement Lambertiens, la plupart possédant un comportement intermédiaire.

Diamètre de tache

Caractéristique la plus visible d'un éclairage opératoire, le diamètre de la tache lumineuse doit être suffisant pour permettre au chirurgien de voir la zone d'intérêt dans sa globalité. On entend par zone d'intérêt l'élément que le chirurgien est en train de traiter. Eclairer l'intégralité de la scène opératoire permettrait de conserver la zone d'intérêt visible malgré le déplacement de cette dernière dû à l'activité du chirurgien. Ceci n'est cependant pas souhaitable car exposer la scène à un niveau proche de 160klux sur une si grande région provoquerait un éblouissement qui empêcherait le chirurgien de travailler dans de bonnes conditions. De plus, il s'avère techniquement difficile de générer une tache lumineuse à ces dimensions (certains axes de recherche sont présentés en annexe A.1 figure A.8). En outre, les diamètres utilisés actuellement sont compris entre 150mm (Powerled300) et 220mm. Ils ne garantissent donc pas la couverture de la zone d'intérêt pour l'ensemble de l'intervention. Il faut donc se résoudre à déplacer l'éclairage lorsque celle-ci se trouve trop éloignée de la tache lumineuse. Ceci est d'autant plus valable que l'opération est superficielle et étendue. Les opérations en cavité sont une exception à cette règle. Plus profondes que larges, une fois l'éclairage bien positionné et sous condition que l'axe optique coïncide avec celui de la cavité, il n'est généralement plus nécessaire d'en rectifier la position ainsi que l'orientation. Par la suite nous utiliserons le terme "pose" pour faire référence au placement de l'éclairage opératoire.

Le point suivant aborde les moyens mis en œuvre pour modifier la pose ainsi que les problèmes associés qui sont au cœur de cette étude.

Cinématique d'un éclairage opératoire

De nos jours, la solution communément adoptée par les fabricants consiste en une cinématique organisée autour de six degrés de liberté (RRR RRR) (figure 1.7a). Une version dégradée basée sur cinq degrés de liberté est utilisée lorsque la hauteur sous plafond n'est pas assez grande. Celle-ci permet d'accéder à l'ensemble des poses désirées et ne nécessite que "peu" d'efforts pour son déplacement comparativement aux anciens produits d'éclairage. Ces efforts sont évalués (à l'aide d'un dynamomètre) à 22.5N et 21.6N pour les deux axes horizontaux au niveau de la poignée stérile du chirurgien et 38.2N pour la montée descente de la coupole.

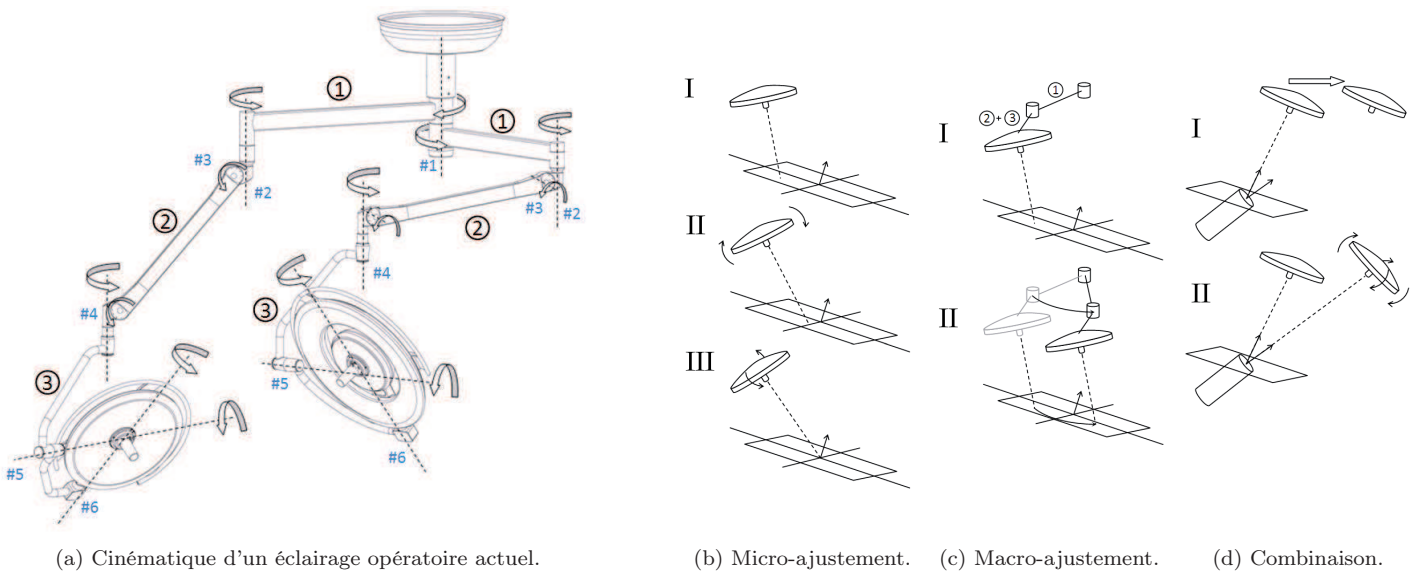


FIG. 1.7 – Illustration du principe pour la réorientation et le repositionnement.

Un éclairage opératoire (1.7a) se compose d'un arceau (①), d'un bras élévateur (②) et d'un bras horizontal (③). L'arceau est lui-même composé de deux fourches qui permettent la réorientation de la coupole. Ces deux axes non concourants (perpendiculaires l'un à l'autre) impliquent que les mouvements engendrés sont restreints à une sphère. En agissant sur l'arceau il est possible de recentrer la tache lumineuse sur la zone d'intérêt, on parle alors de micro-ajustement (figure 1.7b). La tache lumineuse n'étant optimale que pour une distance coupole-plaie opératoire comprise entre 1m et 1.20m (cf. volume de lumière en annexe A.1), il s'avère parfois nécessaire d'abaisser ou d'éloigner l'éclairage. Ce type de mouvement est permis grâce au bras élévateur qui maintient l'ensemble coupole-arceau à une altitude définie manuellement par l'utilisateur généralement en début d'intervention. Enfin, le bras horizontal permet des mouvements de grandes amplitudes, on parle de macro-ajustement (figure 1.7c). Assimilé à un repositionnement, il est utilisé lorsqu'une réorientation n'est plus suffisante. Ces micro et macro-ajustements peuvent donc être combinés (figure 1.7d). Cette utilisation conjointe des deux mouvements survient notamment pour les phases de prépositionnement en début d'intervention (le chirurgien n'ayant plus qu'à réajuster la pose si nécessaire en cours d'opération) ainsi que pour les interventions en cavité où il est nécessaire de rendre colinéaire l'axe optique de l'éclairage et celui de la cavité.

Selon l'activité du chirurgien, les déplacements nécessaires sont donc des micro-ajustements, macro-

ajustements ou bien une composition des deux. Suite aux visites en bloc opératoire effectuées dans le cadre de cette étude, il apparaît que les micro-ajustements sont les déplacements les plus usités. Nous nous focaliserons donc sur ces derniers.

Maniabilité

La manipulation d'un éclairage opératoire se fait grâce à des poignées se trouvant sur la tête d'éclairage (coupole). Bien que celles-ci soient accessibles à l'ensemble du personnel chirurgical, leur utilisation est régie par des règles d'hygiène strictes. Celles-ci diffèrent selon que la personne est stérile ou non. Le chirurgien étant stérile, il lui est interdit de toucher la coupole qui ne l'est pas. Afin de déplacer lui-même la coupole (micro-ajustement par exemple), le chirurgien vient clipser une poignée stérile sur un support placé au centre de la coupole orienté selon son axe optique. Grâce à ce dispositif, le chirurgien peut déplacer la coupole là où il le désire. Il n'est toutefois pas habilité à toucher le panneau de contrôle gérant le niveau d'éclairage et les fonctionnalités propres à la coupole (éclairage d'ambiance, focalisation, ...), généralement situé proche de l'articulation (#6).

Dans le cas où le chirurgien serait dans l'incapacité de déplacer l'éclairage par ce biais, il revient à la panseuse (ou à défaut un autre membre de l'équipe chirurgicale) d'y procéder. Tout comme le chirurgien, celle-ci doit respecter des consignes précises. N'étant pas stérile, elle ne peut donc pas toucher la poignée stérile au centre de la coupole. Pour accomplir cette tâche, elle ne peut utiliser que les poignées situées sur le pourtour de la coupole ; on parle de poignées auxiliaires. L'accès à ces poignées nécessite de se rapprocher des éléments stériles, chirurgien compris. Dès lors, et malgré les précautions, le risque de faire une faute d'asepsie existe et peut amener à des contaminations de la scène. Le tableau 1.1 condense les règles d'usage en matière de manipulation d'éclairage. L'une des difficultés pour déplacer l'éclairage opératoire réside donc dans le respect de la distance séparant la panseuse du chirurgien qui implique très souvent la non visibilité de la région en cours de traitement.

Utilisateur considéré	Stérile	Poignée à utiliser
Chirurgien	Oui	Poignée stérile au centre de la coupole
Assistant	Oui	Poignée stérile au centre de la coupole
Panseuse	Non	Poignée(s) auxiliaire(s) (à l'extérieur de la coupole)

TAB. 1.1 – Consignes d'utilisation d'un éclairage opératoire

En résumé :

Nous avons présenté les caractéristiques d'un éclairage opératoire. Nous avons montré les notions d'éclairage, de luminance, et comment en raison de l'absorption de la lumière par les tissus, il est nécessaire de générer une puissance lumineuse élevée. L'éclairage opératoire apparaît donc indispensable pour les interventions à théâtre ouvert. De plus, sur le plan de la manipulation, il est impératif de modifier la pose de l'éclairage lorsque la zone d'intérêt se trouve trop éloignée de la tache lumineuse. Ce déplacement est effectué manuellement par le biais de poignées stériles ou non grâce à six articulations. Généralement limitée aux micro-ajustements, la modification de la pose ne nécessite que deux degrés de liberté.

1.2.2 Les problèmes liés aux éclairages opératoires

Dans une étude portant sur les activités chirurgicales, [KMJ11] montre que l'éclairage opératoire est déplacé, en moyenne, toutes les 7.5 minutes. Deux tiers des déplacements interviennent durant le traitement de la zone d'intérêt et un tiers au cours de l'incision principale et de sa suture. Les micro-ajustements (réorientations de la coupole) constituent 30% des déplacements et les macro-ajustements (repositionnements de la coupole) interviennent à hauteur de 66%, le restant étant des élévations ou abaissements de l'éclairage. Dans 45% des cas, le chirurgien effectue lui-même le déplacement ce qui entraîne, 97% du temps, une interruption de l'intervention chirurgicale (respectivement 25% et 73% lorsqu'il s'agit de l'assistant chirurgien). Notons qu'il n'y a pas d'interruption lorsque le déplacement de l'éclairage est effectué par une panseuse ou un infirmier. En conséquence, sur une journée d'intervention (comptant 8 heures ouvrées), nous pouvons estimer que 64 déplacements de l'éclairage seront nécessaires dont environ 40 entraîneront un arrêt du geste opératoire. Ceci explique pourquoi l'utilisation d'un éclairage opératoire est de manière générale mal perçue par le personnel chirurgical. En effet, 88% des personnes interrogées sont insatisfaites par ce type de matériel. Pour en illustrer les problèmes nous proposons dans cette partie de présenter les conditions réelles d'utilisation d'un éclairage ainsi que les raisons pour lesquelles il ne donne pas entière satisfaction, pouvant engendrer des problèmes allant jusqu'à la faute d'asepsie par exemple.

Remarque : Par la suite, nous prendrons comme convention d'écriture la notation $\{S, C, S\}$ pour décrire une opération commençant en superficie ($\{\underline{S}, C, S\}$), évoluant vers une cavité ($\{S, \underline{C}, S\}$) et revenant à la fin de l'opération en superficie ($\{S, C, \underline{S}\}$). Une opération chirurgicale intégralement en superficie sera notée $\{S, S, S\}$ et $\{C, C, C\}$ lorsqu'elle se déroule intégralement en cavité .

Afin d'illustrer ces propos nous analyserons dans un premier temps une opération cardiaque de type $\{S, C, S\}$, puis une phlébectomie (excision de varices) $\{S, S, S\}$.

Scenario 1 : pontage coronarien - séquence de type $\{S, C, S\}$

L'opération considérée est un pontage coronarien. Il est possible suivant la gravité des obstructions de procéder à plusieurs pontages, dans ce cas, seul le processus d'anastomose est réitéré (fait de connecter deux éléments anatomiques, ici des vaisseaux sanguins).

Remarque : La voie d'abord sélectionnée s'est portée sur une thoracotomie médiane (figure 1.8b) pour des raisons de facilité de lecture.

Une fois le patient anesthésié, la panseuse prépositionne l'éclairage opératoire de manière à ce que la cage thoracique soit éclairée, sans aucune considération de pose optimale à lui donner. Une action complémentaire serait d'ajuster l'éclairage de la coupole afin de limiter l'éblouissement et ainsi la fatigue oculaire du chirurgien (en raison de l'albédo élevé de la peau). Or, lors de sa mise sous tension, il est courant de voir le personnel chirurgical mettre l'éclairage à son niveau d'éclairage maximum et ne plus y toucher durant le reste de l'intervention. Le chirurgien travaille donc avec un éclairage supérieur à celui nécessaire; la modulation de la luminance apparaît comme un asservissement possible pour les futures générations d'éclairages opératoires.

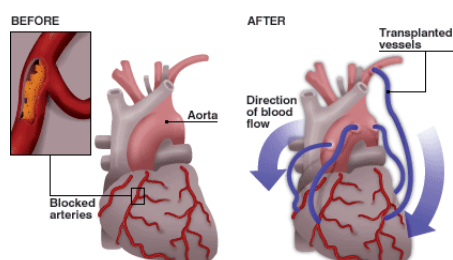
L'éclairage prépositionné, les champs opératoires stériles sont ensuite installés par le chirurgien et servent

à délimiter la zone opératoire. Une fois l'équipe en place et tous les indicateurs acceptables (pression artérielle, rythme cardiaque, ...), le chirurgien procède à une large incision qui servira d'ouverture principale pour la suite de l'opération. A ce moment, la tache lumineuse est focalisée sur la zone d'intérêt par l'instrumentiste grâce aux poignées stériles ou par la panseuse grâce aux poignées auxiliaires. L'opération pour le moment correspond donc à un état superficiel. Vue de la caméra, la plaie forme une ouverture de quelques centimètres de large (en raison de la rétractation naturelle de la peau), laissant apparaître du sang en grande quantité. Etant donné la situation superficielle de l'intervention et compte tenu du recouvrement de la zone d'intérêt par la tache lumineuse, l'éclairage n'a donc pas lieu d'être redéployé pour le moment.

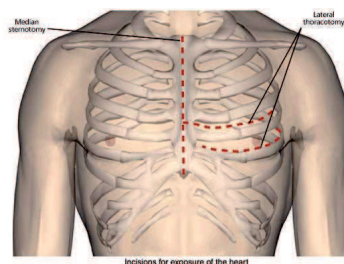
Le sternum assure une fonction protectrice pour les organes mais joue le rôle de barrière pour le chirurgien. Celui-ci doit donc l'inciser pour accéder au cœur. La découpe se fait à l'aide d'une scie le long du sternum (potentiellement suivant un "Z", figure 1.8b). Durant cette phase, les objets observés au-delà des gants et des instruments, sont les champs stériles, la peau, le sang, les os (qui sont recouverts de sang) et pour les personnes de forte corpulence, de la graisse. La fumée (poussière d'os) provoquée par la découpe altère momentanément la vue de la scène opératoire. Une fois l'incision terminée, des écarteurs sont ensuite insérés afin de dégager complètement la voie. A cette étape, le cœur est exposé, mais les contractions rendent difficile toute approche. Dans le but d'en limiter les effets, le chirurgien utilise un stabilisateur tel que le système Octopus ou STAR FISH qui, par compression ou succion, en limite les mouvements (stabilisation myocardique). Ce type d'équipement permet de ne pas recourir à un arrêt du cœur couplé à une circulation extracorporelle ce qui évite les phénomènes de nécrose. A ce moment de l'opération, la séquence opératoire évolue donc vers $\{S, C, \cdot\}$.

Jusqu'à présent le placement de l'éclairage était sommaire. La situation superficielle de l'intervention permettait de se satisfaire de l'éclairement de la zone d'intérêt sans positionnement particulier. Il s'agit maintenant de repositionner l'éclairage afin que la tache corresponde de nouveau à la zone d'intérêt. En effet, même si le cœur est éclairé, la région où sera effectuée la dérivation de l'influx sanguin (anastomose figure 1.8a) ne l'est pas forcément. La situation en cavité apparaît plus exigeante qu'en superficie et nécessite un déplacement combinant généralement micro et macro-ajustements. Pour ce faire, le chirurgien, l'instrumentiste ou la panseuse se saisissent d'une des poignées. Il leur faut faire coïncider l'axe optique de la coupole avec celui de la cavité, ce qui correspond également à la ligne de visée du chirurgien (axe virtuel reliant les yeux du chirurgien à la zone d'intérêt). Si cette manœuvre semble aisée, elle s'avère ardue à réaliser notamment pour le chirurgien pour les raisons suivantes :

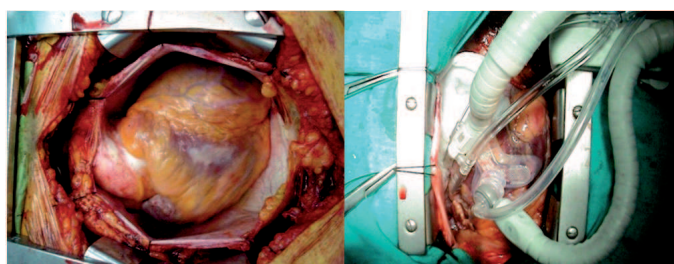
- impossibilité d'utiliser l'une de ses mains pour déplacer la coupole compte tenue de la criticité du geste en cours ;
- selon certaines écoles de pensée, il n'est pas à la charge du chirurgien de s'occuper de la tache de repositionnement, cette activité est donc laissée aux soins de la panseuse ;
- le chirurgien ne sachant pas où, ni comment déplacer l'éclairage opératoire, il peut solliciter la panseuse pour le faire considérant que son point de vue extérieur lui permettra de trouver la position correcte ;
- la configuration de l'éclairage ne lui permet pas d'être déplacé à cause d'une singularité, d'un croisement des suspensions, ou bien d'un effort à fournir trop important ce qui est le cas si le mouvement à effectuer nécessite par exemple une torsion du dos (cas où la poignée est au niveau



(a) Déviation d'un flux sanguin (source : BBC News).



(b) Voies d'abords possibles (sternotomie).



(c) A gauche sans stabilisateur ([YSO+10]), à droite avec le système Octopus.

FIG. 1.8 – Principe d'un pontage coronarien par chirurgie à théâtre ouvert.

de la tête du chirurgien et légèrement en retrait) ;

- le déplacement de la coupole impliquerait un déplacement du chirurgien autour de la table d'opération, chose qu'il ne peut pas faire (ou n'a pas envie de faire) ;
- une faute d'asepsie, due à la recherche à tâtons de la poignée stérile, l'activité du chirurgien l'ayant empêché de détourner les yeux pour savoir où était la coupole.

Lorsque le chirurgien est dans l'incapacité de déplacer l'éclairage, il revient à la panseuse d'accomplir cette tâche. Toutefois, tout comme le chirurgien, celle-ci peut se heurter à certaines difficultés lors de sa manipulation :

- l'éclairage opératoire est hors de portée. La panseuse, malgré l'ajout d'un marche-pied, ne peut atteindre l'une des poignées auxiliaires ;
- le temps alloué à un repositionnement est trop long face à la criticité de la situation, le chirurgien préfère se satisfaire des conditions d'illumination actuelles ;
- la panseuse n'a aucune idée de la position optimale de la coupole ;
- l'absence de visibilité de la scène chirurgicale l'oblige à suivre les directives du chirurgien qui ne lui permettent pas d'atteindre la position souhaitée ;
- le déplacement de la coupole l'oblige à s'approcher trop près du chirurgien ou de la table où reposent les instruments ce qui peut amener à une faute d'asepsie. Cette situation est très fréquente dans les opérations cardiaques qui font souvent appel à plusieurs chirurgiens ;
- la zone d'intérêt n'est visible que par le chirurgien (par exemple sur la paroi interne de la cage thoracique).

Les différentes causes listées ci-dessus rendant impossible le déplacement de l'éclairage opératoire ont été focalisées sur les acteurs (le chirurgien ou la panseuse) et non sur le matériel. Pour autant, les éclairages au même titre que les autres équipements peuvent subir des endommagements et dysfonctionner. Des problèmes de blocage d'axe, des accrochages avec d'autres appareils (comme les bras de distributions), la défaillance des freins induisant une dérive de la position, . . . , peuvent survenir. Cependant, l'erreur ou le mauvais positionnement dus à de tels dysfonctionnements restent rares.

Remarque : *Nous ne mentionnons pas les conséquences de ces problèmes telles que le stress, les disputes, la perte de concentration, l'arrêt de l'opération, la perte d'énergie et la variabilité de l'effort pouvant altérer la qualité du geste chirurgical. De plus, l'obligation réglementaire d'utiliser deux coupoles engendre des difficultés supplémentaires que nous n'évoquerons pas dans la mesure où cette étude se restreint à une seule coupole.*

Dans la majorité des cas il n'est pas possible pour l'équipe d'atteindre une position satisfaisante pour les raisons évoquées plus haut. Le chirurgien a donc recours à une source de lumière froide sous forme d'une lampe frontale afin de compléter l'éclairage opératoire aux endroits où aurait dû être la tache lumineuse. Grâce à l'emploi de cette source auxiliaire il ne lui est plus nécessaire de réajuster la pose de l'éclairage dans la mesure où cette tache lumineuse est centrée sur l'endroit où travaille le chirurgien. L'objectif de ce travail est donc d'affranchir le chirurgien du port d'une lampe frontale.

Une fois l'anastomose accomplie, il y a un retour vers une situation en superficie ($\{S, C, S\}$). Le sternum est suturé par du fil métallique pour des raisons de tenue mécanique, la peau quant à elle l'est par du fil simple. Pour cette dernière partie de l'intervention, les contraintes en termes de positionnement de l'éclairage opératoire sont relaxées et seul compte le recouvrement de la zone d'intérêt par la tache lumineuse. Ainsi, comme celui-ci se trouvait en bonne position pour la partie $\{S, C, \cdot\}$, il se trouve l'être d'autant plus pour $\{S, C, S\}$.

Scenario 2 : phlébectomie - séquence de type $\{S, S, S\}$

La phlébectomie consiste à supprimer les varices présentes sur les membres inférieurs. Celles-ci sont le résultat d'une insuffisance veineuse impliquant une dilatation des veines superficielles (c'est-à-dire entre le muscle et la peau). Bénignes, elles ne contribuent qu'à un préjudice esthétique, cependant, à un stade plus avancé, elles peuvent provoquer douleurs, œdème voir un ulcère variqueux.

Le cas considéré ici concerne un patient pour lequel il est nécessaire de retirer les varices en commençant par le haut de la cuisse et allant jusqu'à la cheville. Une fois le patient anesthésié, la panseuse prépositionne l'éclairage opératoire sur le haut de la cuisse. Ce début d'opération correspond donc à $\{S, \cdot, \cdot\}$.

A ce stade, il serait nécessaire, dans la mesure où la peau bétadinée renvoie une grande partie des rayons émis par la coupole, de diminuer le niveau d'éclairement afin que la fraction de lumière renvoyée ne soit pas trop agressive pour le chirurgien.

Grâce à une aiguille dont le bout est crocheté (crochet de Müller), le chirurgien vient tirer l'une des extrémités de la veine qui vient former un bourrelet à la surface de la peau. Le parcours des veines à retirer est connu du chirurgien grâce aux marques faites lors de la visite pré-opératoire (figure 1.9a). La suite de l'intervention est résumée sur l'illustration 1.9b et consiste à couper la veine en son milieu puis

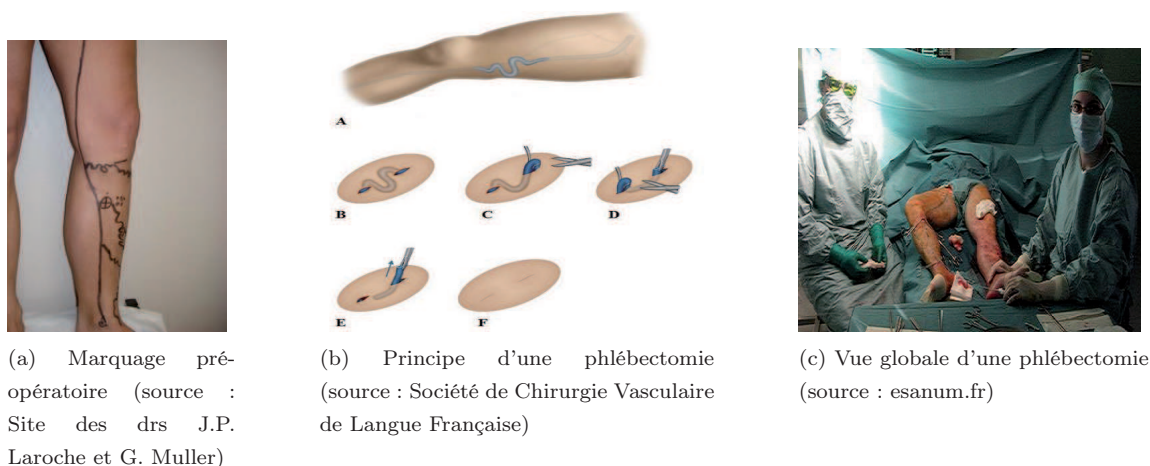


FIG. 1.9 – Principe d'une phlébectomie.

à réitérer ces gestes pour l'autre extrémité. Une fois le segment libre, il lui suffit de l'extraire ($\{S, S, \cdot\}$). Cette procédure est reproduite pour toutes les varices à traiter.

En principe, il serait nécessaire de procéder à des micro-ajustements à chaque changement de varice, de telle sorte à ce que la tache lumineuse puisse correspondre à la zone d'intérêt. Or, en pratique il n'en est rien, le chirurgien se contente de déplacer l'éclairage que lorsque la zone d'intérêt est complètement hors de la tache lumineuse. La raison réside principalement, pour ce type d'intervention, dans la perte de temps et d'énergie nécessaire au déplacement. De plus si le chirurgien est le seul à pouvoir le faire, il lui faudrait alors arrêter son geste un trop grand nombre de fois ce qui impliquerait une augmentation de la durée de l'intervention et un état de stress plus élevé. La répétition de ce type d'action peut induire un positionnement sans vigilance, pouvant mener à des fautes d'asepsies. Le chirurgien dans ces circonstances, préfère généralement se satisfaire de conditions d'éclairage insuffisantes. Sur la base des visites en bloc opératoire, nous estimons que ces modifications de pose de l'éclairage ne se font en moyenne qu'une fois sur quatre.

1.2.3 Résumé des problèmes d'éclairage en chirurgie

Nous avons vu qu'il existe deux types de déplacements possibles. Le premier, correspondant aux interventions en superficie, ne requiert qu'une réorientation de la coupole et ne mobilise que deux degrés de liberté situés au niveau de l'arceau. Le second, correspondant aux interventions en cavités nécessite de faire coïncider l'axe optique de l'éclairage avec la ligne de visée du chirurgien. Pour ce faire, les deux degrés de liberté du bras horizontal sont employés pour exécuter le repositionnement en plus de ceux utilisés pour la réorientation. Entre ces deux situations franches, il existe une infinité de cas mêlant superficie et cavité et qui par conséquent requièrent l'usage de ces quatre degrés de liberté.

Nous avons pu voir que les actions entreprises sur un éclairage opératoire s'appuient sur une vision directe de la scène chirurgicale, une connaissance a priori de l'opération ainsi que sur les instructions

explicités du chirurgien (si ce n'est pas lui qui manipule l'éclairage). Nous avons pu voir que :

1. le manque de maniabilité de l'éclairage entraîne des fautes d'asepsies, des conflits au sein de l'équipe chirurgicale, et des collisions avec les autres équipements présents dans la salle ;
2. la fréquence d'utilisation d'un éclairage opératoire n'est pas suffisante ce qui implique un éclairage non optimal pour le chirurgien.

Il existe donc une réelle attente du personnel chirurgical pour améliorer le déplacement des éclairages opératoires. Il s'agit maintenant d'évaluer les approches existantes pour permettre, in fine, de proposer une solution plus adéquate.

1.3 Approches existantes pour la résolution des problèmes de mauvais placement

Afin d'améliorer les conditions de travail de l'équipe chirurgicale, de nombreux systèmes robotiques ont été développés au cours des vingt dernières années. Nous allons montrer que ces dispositifs ne résolvent pas les problèmes évoqués précédemment. Dans un second temps, nous présenterons la stratégie envisagée pour résoudre ces problèmes ainsi que le cahier des charges imposé par l'industriel. Nous verrons que le concept d'éclairage opératoire capable de désigner de lui-même la zone d'intérêt du chirurgien et de s'y rendre de façon autonome permet d'être optimalement placé. Nous nous proposons dans un premier temps de faire un bref tour d'horizon des systèmes existants pour améliorer les conditions d'illumination.

1.3.1 LightPilot

Le système LightPilot (ou C950G) développé par la société Berchtold, est un éclairage opératoire équipé de moteurs au niveau des liaisons #5 et #6 (figure 1.10a). Par leurs actions combinées, ces moteurs permettent de réorienter la coupole et ainsi centrer la tache lumineuse sur la zone d'intérêt. Il s'agit donc de micro-ajustements. Pour ce faire, une personne dédiée place une pastille réfléchissante ou catadioptré (stérilisée afin d'éviter toute contamination du site opératoire) à l'endroit où se trouve la zone d'intérêt (celle-ci étant connue du chirurgien dans la mesure où il travaille à cette endroit). Un faisceau infrarouge émis en continu par la coupole en direction de la scène chirurgicale se reflète sur le catadioptré.

Remarque : Notons que ce faisceau provoque, par la même occasion, un dessèchement des tissus superficiels.

Le faisceau renvoyé est ensuite réceptionné par une caméra placée dans la coupole. Il s'agit donc de bien orienter la pastille, de manière à ce que le faisceau soit visible par la caméra (en prenant garde de ne pas envoyer le faisceau sur un équipement sensible aux infrarouges). La détermination de la position de la pastille dans le repère image s'effectue grâce à un traitement d'images basé sur un seuillage des valeurs des pixels permettant d'isoler l'empreinte lumineuse et sur un calcul de position du centre de gravité de celle-ci. Ceci revient donc par extension à localiser la zone d'intérêt. Il y a donc un lien direct entre le placement de la pastille et la réorientation de la coupole, le défaut de l'un entraînant l'échec de l'autre. Une fois cette information connue, le système calcule la loi de commande à appliquer aux moteurs pour réaliser le centrage de la tache lumineuse sur la zone d'intérêt (figure 1.10b). L'objectif atteint, il

faut procéder au retrait rapide de la pastille afin que celle-ci ne soit pas détectée et redéfinie comme la position de la zone d'intérêt ; le système étant dans l'incapacité de savoir si le déplacement en cours est adéquat ou non. Le cas échéant, il est nécessaire de placer la pastille de telle sorte à ramener la coupole dans une orientation convenable ce qui implique, outre un temps additionnel à l'accomplissement de cette manœuvre, une augmentation des risques d'asepsie en raison des gestes nécessaires à cette opération.

Ce système robotisé permet donc de résoudre les difficultés rencontrées lors d'opérations dont au moins l'une des parties est en superficie ($\{S, \cdot, \cdot\}$, $\{\cdot, S, \cdot\}$ ou encore $\{\cdot, \cdot S\}$) dans la mesure où seules les articulations #5 et #6 sont nécessaires. Il permet de conserver des conditions d'illumination optimales ce qui octroie un gain indirect en termes de qualité du geste.



(a) Agrandissement sur la motorisation du LightPilot (source : Service-manual Chromophare C950G)



(b) Pastille stérile utilisée par le module LightPilot (source : profimed.com)

FIG. 1.10 – Module LightPilot développé par la société Berchtold.

Si ces bénéfices sont attractifs, ils s'avèrent en revanche occultés par les conditions d'utilisation du LightPilot. Mis en cause, le type d'information utilisé pour la réorientation de l'éclairage et son mode de récupération. En effet, dans la mesure où le système est incapable de comprendre les actions entreprises ou les objets en présence dans la scène chirurgicale, il est impératif de le renseigner sur la localisation de la zone d'intérêt. Ceci impose deux constats. Le premier est qu'il est nécessaire d'interrompre le geste chirurgical pour placer le catadioptré et attendre le déplacement complet de la coupole. Le second est qu'il est impératif de redonner cette information à chaque fois qu'il y a déplacement de la zone d'intérêt. Or, réitérer la désignation manuelle de la zone d'intérêt n'est pas envisageable pour certaines chirurgies (phlébectomie par exemple) dans la mesure où celle-ci est amenée à se déplacer rapidement et de nombreuses fois. Conséquence immédiate, le chirurgien ne fait l'effort de renseigner le système qu'en de très rares occasions.

A la vue de ces éléments, ce système apparaît donc inapproprié à un usage en bloc principalement en raison du temps nécessaire qui s'avère parfois plus long qu'une utilisation manuelle. A ceci s'ajoutent l'arrêt obligatoire du geste chirurgical, le recours systématique à un objet spécifique pour désigner la zone d'intérêt (via catadioptré) et la possibilité de déplacement inapproprié de l'éclairage dans les cas où la pastille n'est pas retirée assez rapidement. Ceci explique donc l'échec commercial de cet éclairage.

Cependant, il est possible de retenir certains enseignements qui nous seront utiles pour cette étude :

- la désignation explicite (faite par une personne) de la zone d'intérêt par gestuelle est envisageable mais ne doit pas nécessiter d'arrêt du geste chirurgical majeur. L'idéal consisterait à désigner la zone d'intérêt de manière indépendante ;
- la motorisation des axes #5 et #6 permet de répondre aux opérations en superficie ;
- pour être utile, le déplacement du système doit, une fois les coordonnées de la zone d'intérêt connues, se faire de manière autonome.

1.3.2 Eclairage autonome basé luminance

Développé par [Spa07], ce système permet le centrage du flux lumineux sur la zone d'intérêt par l'intermédiaire d'une structure robotisée constituée d'un seul degré de liberté. Cependant, à la différence du système LightPilot, il n'est pas nécessaire de désigner manuellement la position de la zone d'intérêt. Cette absence d'interaction homme/machine permet d'éviter les problèmes inhérents à la phase de renseignement du système. L'idée adoptée consiste à désigner la zone d'intérêt à l'endroit de plus faible luminance. L'hypothèse est qu'il y a une plus forte absorption de la lumière (et donc moins de luminance) aux endroits où le chirurgien travaille du fait de son activité (les tissus internes absorbant plus de lumière que les tissus externes).

Pour ce faire, une photodiode donne l'"image" de la luminance locale de la scène chirurgicale. Positionnée en front d'éclairage, celle-ci n'est pas perturbée par l'éclairage des sources lumineuses placées derrière elle. Une fois l'observation enregistrée pour la position courante, il y a déplacement de l'éclairage suivant une trajectoire qui s'apparente à un "8". A intervalles réguliers, un nouveau jeu de mesures est enregistré. Le parcours terminé, le système compare les mesures de luminance obtenues et ne retient que la plus faible. Dès lors, le système revient à la position correspondant à ce minimum, ce qui par hypothèse correspond à la zone d'intérêt et donc à la bonne orientation de l'éclairage. Afin de conserver la zone d'intérêt au centre de la tache lumineuse, le système réitère à temps régulier l'intégralité de cette procédure.

Cet éclairage se révèle donc être très simple d'utilisation, aucune intervention humaine n'étant nécessaire. Si sa capacité à désigner la zone d'intérêt n'est pas représentative d'une intelligence apte à comprendre son environnement, son aptitude à s'y rendre par lui-même traduit une autonomie réelle. Autre avantage, l'absence de puissance de calcul importante réduit considérablement la complexité du système, ce qui lui assure un taux de fiabilité appréciable pour ce domaine d'application.

L'inconvénient majeur réside dans la validité de l'hypothèse utilisée pour la désignation de la zone d'intérêt. En effet, considérer le minimum de luminance comme critère pour la prise de décision revient à désigner la zone d'intérêt sur l'objet le plus sombre. Or, le système n'ayant pas les moyens de reconnaître les objets en présence, cette hypothèse peut se révéler inadéquate. Ceci a donc un impact sur l'orientation finale de la coupole.

Pour illustrer ces propos, considérons deux objets notés 1 et 2 de couleurs très différentes, mais d'albédo égaux, placés dans la scène chirurgicale (figure 1.11). Au début de l'opération, la coupole est préposi-

tionnée de telle sorte que la tache lumineuse soit équidistante des deux objets. Supposons que l'activité du chirurgien l'amène à travailler sur l'objet n°1 puis, après un certain temps sur l'objet n°2. Le système est activé et procède aux mesures de luminance à intervalles réguliers (figure 1.11a). La coupole s'oriente sur la mesure de luminance la plus basse c'est-à-dire sur l'objet n°2. Quelques temps après, le système réitère la procédure et n'effectue aucun mouvement, l'objet n°2 correspondant toujours au minimum de luminance.

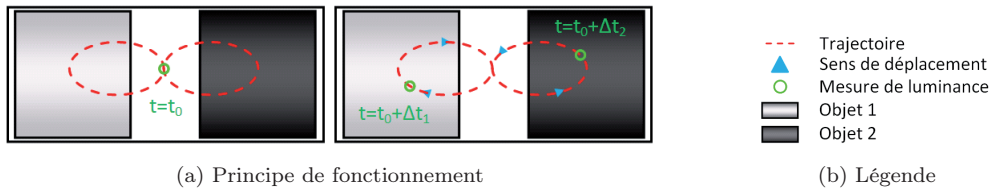


FIG. 1.11 – Concept développé par [Spa07]

Malgré le changement d'activité du chirurgien, amené à travailler sur l'objet n°1 puis l'objet n°2, le système n'a pas été en mesure de réajuster son orientation. L'hypothèse utilisée ne convient tout le temps et conduit à une désignation erronée de la zone d'intérêt, et donc à un comportement inadéquat de l'éclairage. Cet exemple permet d'illustrer les limites de ce type de système et montre combien la compréhension de la scène chirurgicale est primordiale pour l'exécution de micro-ajustements (ou de macro-ajustements). Bien que l'hypothèse de travail soit discutable, l'idée sous-jacente consistant à définir la zone d'intérêt à l'endroit où le chirurgien opère nous semble fondamentale.

1.3.3 Voûte de Blin revisitée

Ce système développé par [Pal11] consiste en une adaptation des voutes de Blin (voir annexe A) avec la technologie à led. Dans cette application, l'éclairage se compose de sept sources lumineuses indépendantes intégrées au plafond. Chacune d'entre elles est montée sur une station motorisée à deux degrés de liberté (pan et tilt) permettant au projecteur d'être dirigé en direction de la zone d'intérêt. La désignation de la zone d'intérêt se fait manuellement via une interface graphique, l'opérateur pouvant visualiser la scène opératoire grâce à une caméra intégrée dans le plafond. Outre la localisation de la zone d'intérêt, d'autres fonctionnalités sont proposées : le réglage du niveau d'éclairage, et celui des paramètres de la caméra (focus, zoom). Les coordonnées de la cible sont ensuite véhiculées par wifi jusqu'à un ordinateur chargé de la détermination des mouvements à appliquer pour chaque éclairage. Les ordres de déplacement sont donnés par liaison DMX512 jusqu'au contrôle-commande des projecteurs qui activent les moteurs en fonction. La réorientation s'effectue grâce à une boucle ouverte où seule est prise en compte l'information des codeurs, la vidéo n'est donc pas exploitée durant cette phase.

Les deux avantages de ce type de système sont d'une part l'absence de suspension qui permet de s'affranchir des problèmes de collisions et de maniabilité évoqués dans la partie consacrée aux éclairages modernes. D'autre part, il y a une diminution des risques de fautes d'asepsies en lien avec l'éclairage et l'exemption de perturbation du flux laminaire du fait de l'espace laissé entre la table d'opération et le chirurgien.

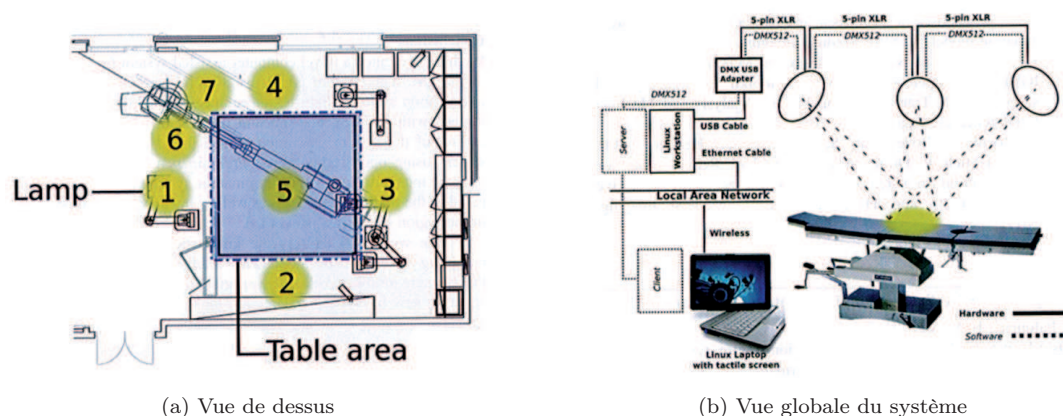


FIG. 1.12 – Système d'éclairage opératoire plafonnier [Pal11]

Cette distance séparant la plaie opératoire de l'éclairage constitue paradoxalement l'une des faiblesses de ce système. Le déport des sources impose l'augmentation de l'éclairage comparativement aux éclairages opératoires montés sur bras articulés. Ceci contribue à l'accroissement de la température en raison de l'empiètement du spectre d'émission dans l'infrarouge ce qui amène au dessèchement des tissus superficiels. Autre désavantage, la désignation manuelle de la zone d'intérêt nécessite d'avoir une personne dédiée au renseignement du système. L'incapacité du système à suivre la zone d'intérêt, implique sa dépendance vis-à-vis de l'opérateur. En effet, la position de la tache lumineuse est directement fonction de la performance de l'opérateur à localiser la zone d'intérêt et à la saisir sur l'interface homme-machine. D'autre part, cette obligation de passer par une interface homme-machine peut induire une lassitude du personnel, une consommation de temps importante voire devenir source de conflit.

En résumé :

L'ensemble de ces dispositifs met en valeur l'absence d'adaptabilité des éclairages opératoires actuels. Nous avons vu que pour réorienter la coupole il est nécessaire de renseigner le système sur la position de la zone d'intérêt par le biais d'une interface. Ceci limite considérablement les bénéfices apportés. De plus, pour conserver la zone d'intérêt centrée sur la tache lumineuse, il est nécessaire de réitérer fréquemment la désignation, d'où une perte de temps et une imprécision dues à une mauvaise transmission de l'information du fait de la lassitude. Rendre capable le système de s'orienter de lui-même sur la zone d'intérêt apparaît donc comme la solution optimale. Il faut cependant prêter la plus grande attention aux hypothèses conditionnant la désignation, car de leur validité dépend le bon comportement du système et donc son utilité.

En conséquence, le concept imaginé consiste en un éclairage opératoire asservi se déplaçant de manière autonome et sans interaction avec l'équipe chirurgicale. Celui-ci doit être capable de conserver la zone d'intérêt du chirurgien au centre de la tache lumineuse tout au long de l'opération. L'ambition de cette étude est d'autant plus grande, que le système devra être apte à évoluer seul (c'est-à-dire sans système additionnel) et être opérationnel quel que soit le type d'intervention. Ceci représente donc un défi de taille en raison de la variabilité des opérations (exhaustivité des instruments, des voies d'abord, des pathologies, des techniques chirurgicales, ...).

La réussite de ce programme permettrait d'utiliser l'éclairage opératoire de manière optimale et éviterait par la même occasion les problèmes d'illumination hors champ ou de manque d'éclairage. Dès lors, n'ayant plus besoin de se soucier de son placement, chacun des membres de l'équipe chirurgicale pourrait se concentrer sur des activités plus essentielles, ce qui répondrait collatéralement aux griefs mentionnés dans la section consacrée aux éclairages modernes. Outre cette optimisation des conditions d'illumination de la scène opératoire, celui-ci permettrait d'autre part un gain de temps et d'énergie ainsi qu'une élimination des variations importantes d'effort notamment dans le cas des chirurgies neurologiques (passage d'un micro-effort à un macro-effort dû au déplacement de la coupole, puis retour à un micro-effort).

La pluralité des interventions possibles couplée à ce souci de transparence sous-tend la nécessité d'avoir un système combinant **intelligence** et **autonomie**, c'est-à-dire un système capable de désigner de lui-même la zone d'intérêt du chirurgien et de s'y rendre. Toutefois ce concept soulève de nombreuses difficultés techniques pour lesquelles nous allons proposer des solutions.

1.4 Le concept d'éclairage opératoire asservi

Nous venons de voir que les concepts développés au cours de ces dernières années ne sont pas adaptés à la problématique de positionnement d'un éclairage opératoire. Toutefois, ils offrent des pistes de réflexion notamment sur le choix du capteur servant à renseigner le système. En l'occurrence, l'utilisation d'une caméra permet une mesure non-intrusive et non-destructive appropriée à notre environnement. La richesse d'information est un autre atout de ce type de capteur qui permet l'accès à des éléments de différents niveaux cognitifs.

Nous nous intéressons maintenant à exploiter ces enseignements pour élaborer le concept d'éclairage opératoire asservi permettant de centrer le flux lumineux sur la zone d'intérêt du chirurgien.

1.4.1 Choix technologiques

La problématique d'éclairage opératoire asservi soulève de nombreuses questions comme, en premier lieu, la cinématique à donner au système pour effectuer un déplacement adéquat. En s'appuyant sur les applications robotiques existantes, nous nous sommes orientés vers une motorisation des axes #5 et #6 à l'image du LightPilot. Cette solution est apparue comme la plus naturelle et la moins contraignante compte tenu de la cinématique existante, car il s'agit du point commun entre micro et macro-ajustements. Nous verrons que l'intégration des moteurs au niveau de ces articulations se prête tout particulièrement à la situation mais que sa mise en œuvre s'est avérée délicate.

La partie logicielle est nettement plus complexe dans la mesure où il n'existe pas d'antécédent, à notre connaissance, parmi les systèmes robotiques appliqués au domaine chirurgical (voir annexe B). En effet, les robots en chirurgie sont dédiés à un type d'acte particulier et utilisent des renseignements explicites d'origine humaine (désignation via écran tactile, joystick ou conditionnent leur prise de décision par l'observation de marqueurs placés dans la scène). Or, cette étude vise à déplacer l'éclairage sans indice visuel rajouté ou d'information explicite provenant de l'équipe chirurgicale. Facteur aggravant, ces

applications reposent généralement sur une connaissance a priori synthétisées sous la forme de modèle 3D (ce qui implique une étape de recalage). Ce prétraitement des informations relatives au patient est généralement complété par une connaissance précise du protocole opératoire. Ces informations pouvant être inaccessibles (pas d'imagerie pré-opératoire) et/ou inexploitable (pas de protocole opératoire défini et/ou standardisé), cela justifie donc notre choix de ne pas prendre en compte ces travaux dans le cadre de cette étude.

Une des approches envisagées consiste à s'intéresser à l'axe de regard (ligne de visée) du chirurgien, sa fatigue visuelle, l'état de son stress, etc. Or, certaines de ces informations relèvent plus du domaine qualitatif, quant aux autres, elles nécessiteraient des capteurs placés directement sur le corps du chirurgien. Dans la mesure où notre objectif consiste à ne pas interagir avec l'équipe, ne pas utiliser des informations relatives au patient, d'indices visuels placés dans la scène, ou des capteurs dans la scène chirurgicale, il a été décidé (au-delà des considérations de stérilisation, d'autonomie et de transmission de données) de ne pas considérer l'axe de regard pour le guidage de l'éclairage.

La solution retenue consiste finalement à extraire les informations à distance par le biais de l'observation. L'utilisation d'une caméra est donc le moyen adéquat. De plus en plus utilisée dans des systèmes robotiques chirurgicaux, riche en information, non-destructive et non-intrusive, la caméra répond à notre problématique de désignation de la zone d'intérêt sous condition de pouvoir interpréter les éléments observés. De plus, l'intégration de la caméra dans l'éclairage opératoire s'est révélée être la solution la plus simple dans la mesure où les éclairages opératoires modernes sont équipés d'interfaces mécaniques permettant l'intégration d'une caméra au centre de la coupole. Un second élément confortant ce choix est la distance coupole-zone d'intérêt gardée constamment à 1.20m tout au long de l'intervention. Enfin, l'axe optique de la coupole et celui de la caméra étant colinéaires, centrer la zone d'intérêt dans l'espace image de la caméra revient à focaliser la tache lumineuse sur celle-ci, ce qui facilitera l'expression de la loi de commande.

Remarque : *Une alternative à l'observation de la scène opératoire pourrait consister à observer le chirurgien. Or, la posture courbée du chirurgien ainsi que le port de masque, voire de lunettes grossissantes comme dans le cas des interventions sur le cœur ne permettent pas de distinguer les traits du visage porteurs d'informations utiles : ligne de visée, crispation du visage, sudation, etc. De plus, celui-ci peut être amené à porter son regard sur une partie extérieure à l'intervention, ce qui pourrait avoir une incidence sur le déplacement de l'éclairage opératoire.*

Pour pouvoir désigner la zone d'intérêt, il faut comprendre où travaille le chirurgien. Pour accomplir son acte, celui-ci se base sur un ensemble d'éléments objectifs (connaissance du corps humain, signaux patients) et subjectifs (communication avec l'équipe, capacité à gérer la tension), ce qui a été résumé par [Vor06] dans une étude portant sur la commande d'un robot-porte-optique pour les opérations mini-invasives :

"[...] , si les capacités cognitives d'un être humain sont de loin supérieures à celles d'une machine (aptitude à prendre des décisions dans un environnement incertain, à s'adapter à des situations imprévues, à apprendre par l'expérience), son comportement est beaucoup plus incertain (variabilité de ses réactions en fonction du stress, de la fatigue, de la charge cognitive) et imprédictible (inattention, possibilité de comportements irrationnels). Par ailleurs,

les prises de décision d'un être humain s'appuient sur une multitude de facteurs théoriques, empiriques, intuitifs, et sont de ce fait difficilement quantifiables. L'être humain a également des capacités sensorielles et motrices très perfectionnées (très bonne vision, grande dextérité, excellente coordination main-œil, sensations tactiles très fines), mais exécute des gestes avec une répétabilité faible, est sujet aux tremblements (accentués par la fatigue) et a des gestes dont la précision est difficilement quantifiable."

Il ressort de cet extrait que l'analyse du comportement du chirurgien basée uniquement sur l'observation est ardue. Ainsi, baser la détermination de la zone d'intérêt uniquement sur l'analyse de la scène chirurgicale nécessite de sélectionner des indices visuels pertinents.

Pour atteindre cet objectif, ces indices doivent être suffisamment récurrents et fiables malgré l'importante variabilité des actes, des éléments opérés (nature, localisation, aspect et forme différente), des équipements et l'exhaustivité des scénarii opératoires. En outre, il n'a pas été fait le choix de construire un modèle spécifique de la séquence opératoire pour chaque type d'opération en raison de l'explosion combinatoire des cas envisageables. Au-delà de la difficulté liée aux aspects humains (imprévisibilité, incertitudes, etc) et factoriels (multiplicité des éléments rencontrés par essence anormale), s'ajoute celle liée à la voie d'abord sélectionnée par le chirurgien qui peut avoir un impact sur les indices utilisés si ceux-ci sont en lien avec les tissus observés (couleur, texture, forme, etc).

1.4.2 Les objectifs de la thèse

Pour résoudre les problèmes de maniabilité, nous proposons de robotiser un éclairage opératoire et de le rendre intelligent et autonome.

La première étape du travail consiste à extraire de la scène chirurgicale, les informations nécessaires à la désignation de la zone d'intérêt en l'absence d'ordre explicite de l'équipe chirurgicale. Le processus de décision permettant, à partir d'indices visuels, de désigner la zone d'intérêt devra être défini.

La seconde étape s'attache à la mise en mouvement de l'éclairage opératoire afin d'éclairer de manière optimale la zone d'intérêt désignée. La loi de commande utilisée doit tenir compte de l'aspect interventionnel de la chirurgie (vitesse de déplacement de la tache lumineuse) mais aussi des spécificités dues à la cinématique de l'éclairage opératoire lui-même (espace de travail du chirurgien, zone stérile). Il faut garantir la sécurité du chirurgien (distances à la coupole raisonnables pour ne pas perturber le chirurgien, vitesse de déplacement peu élevée) ainsi que celle du patient (chute de l'appareil, compatibilité électromagnétique).

Le principe de fonctionnement du concept proposé (figure 1.13) met en avant ces deux étapes nécessaires pour atteindre les objectifs fixés, à savoir l'intelligence et l'autonomie.

Dans la partie suivante, nous montrerons que l'utilisation d'une caméra s'avère être un moyen intéressant pour cette tâche et nous mettrons en lumière les verrous scientifiques que cela impose de lever pour atteindre les buts fixés.

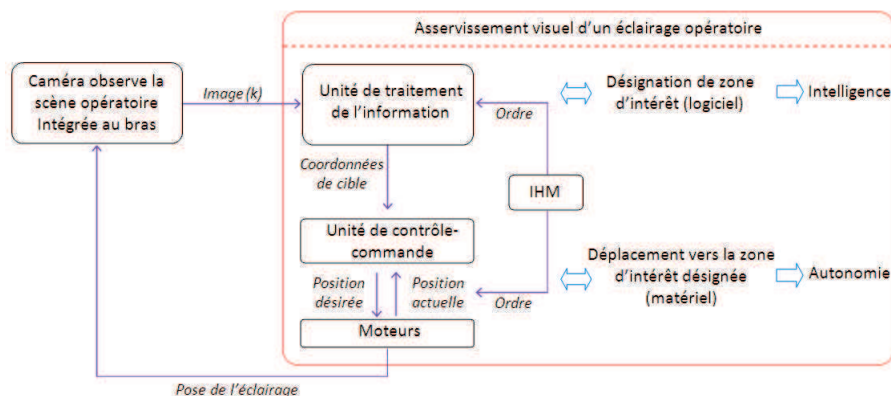


FIG. 1.13 – Processus du concept d'éclairage autonome et intelligent.

1.5 Conclusion

Dans ce chapitre, nous avons présenté les éclairages opératoires utilisés dans le cadre d'interventions à théâtre ouvert. Nous avons montré que pour obtenir des conditions d'éclairage idéales, il est nécessaire de réajuster constamment la pose de l'éclairage. Il s'agit pour les opérations superficielles d'une réorientation mobilisant deux degrés de liberté (micro-ajustement). Pour les opérations en cavité, il faut, en plus de cette réorientation, traduire l'éclairage (macro-ajustement) ce qui mobilise au total quatre degrés de liberté. En raison de la maniabilité des éclairages opératoires, ces deux types de déplacement ne sont généralement pas exécutés. Le chirurgien se contente alors d'un niveau d'éclairage anormalement faible.

Les approches existantes pour pallier ce point reposent soit sur une désignation manuelle de la zone d'intérêt, soit sur l'utilisation de marqueurs spécifiques placés dans la scène chirurgicale. Dans les deux cas, elles ne permettent de satisfaire les attentes de l'équipe chirurgicale. Notre objectif est donc de positionner l'éclairage opératoire de manière optimale quel que soit le type d'intervention sans interaction avec l'équipe chirurgicale, ni marqueurs ajoutés dans la scène.

Il n'existe pas, à notre connaissance, d'éclairage opératoire satisfaisant ces exigences, c'est pourquoi nous proposons de développer un éclairage opératoire robotisé autonome et intelligent apte à se déplacer sans ordre explicite provenant de l'équipe chirurgicale. Il a été identifié que l'utilisation d'une caméra est un moyen optimal pour renseigner le système sur la position de la zone d'intérêt du chirurgien et s'affranchir de toute interaction avec le corps chirurgical. La difficulté réside dans le choix des indices visuels permettant au système de désigner la zone d'intérêt malgré l'environnement qui se trouve être varié (diverses voies d'abord possibles, patients différents), changeant (accès au cours du temps à différentes strates tissulaires) et inconnu par avance (pas d'information pré-opératoire disponible).

Cette thèse a donc pour but d'apporter une solution au travers de deux objectifs distincts :

1. la désignation de la zone d'intérêt du chirurgien en l'absence de données explicites pour une scène inconnue, changeante et non contrôlée ;
2. l'asservissement visuel de l'éclairage opératoire en vue du centrage de la tâche lumineuse sur la zone d'intérêt désignée.

Chapitre 2

Identification d'objets dans une scène chirurgicale

Sommaire

2.1 Introduction	30
2.1.1 Désignation "manuelle" d'une zone d'intérêt	30
2.1.2 Restriction du nombre d'objets à identifier	36
2.1.3 Processus de reconnaissance d'objet	39
2.2 Classification d'objets en milieu chirurgical	40
2.2.1 L'attribut "forme"	40
2.2.2 L'attribut "couleur"	49
2.2.3 L'attribut "mouvement"	57
2.3 Conclusion sur les attributs forme, couleur et mouvement	64

L'objet de ce chapitre est de définir les indices visuels permettant à l'éclairage opératoire de désigner la zone d'intérêt du chirurgien de manière intelligente.

Dans un premier temps, nous montrerons à travers des exemples tirés de situations réelles comment désigner manuellement la zone d'intérêt du chirurgien. Il sera mis en évidence qu'il n'existe pas, dans une scène chirurgicale, d'information unique, récurrente, suffisamment robuste et représentative de la zone d'intérêt. Pour contourner cette difficulté inhérente à la variabilité des actes chirurgicaux, nous présenterons, dans un second temps, une démarche basée sur la reconnaissance de certains objets reposant sur l'extraction de plusieurs attributs.

2.1 Introduction

Le processus de réorientation de l'éclairage opératoire comprend deux étapes. La première, dite de désignation (partie intelligente), consiste à extraire, à partir d'une image provenant d'une caméra implantée dans la coupole, la région la plus probable d'être celle où travaille le chirurgien. L'utilisation d'une caméra permet d'obtenir des informations en grand nombre (sous un point de vue privilégié) et de qualité (riche de sens) tout en s'abstenant de capteurs additionnels placés dans la scène opératoire ou sur le chirurgien. Cette phase doit en outre permettre de pallier les problèmes liés à l'exhaustivité des interventions et le caractère non prévisible de la zone d'intérêt. Une fois celle-ci déterminée, l'étape suivante consiste à introduire les coordonnées de la zone d'intérêt dans la boucle d'asservissement visuel afin que l'éclairage opératoire s'y rende (partie autonome). L'asservissement visuel et par conséquent les performances du système d'éclairage opératoire proposé sont donc conditionnés par l'étape de désignation.

2.1.1 Désignation "manuelle" d'une zone d'intérêt

Nous allons identifier la zone d'intérêt du chirurgien en basant l'analyse uniquement sur une image. Les méthodes permettant d'extraire les différents éléments de la scène opératoire ne sont pas explicitées pour l'instant. Le but est de montrer à travers plusieurs exemples le principe de la désignation.

Remarque : *Etant donné la grande variabilité des cas de désignation d'une zone d'intérêt, la liste des exemples que nous allons présenter se veut représentative sans prétention d'exhaustivité. Parmi les cas qui ne seront pas détaillés ici, nous pouvons citer par exemple :*

- zone d'intérêt en dehors du champ de la caméra (évolution de l'acte au-delà des limites du champ de la caméra, mauvais prépositionnement de l'éclairage, ...). Il est alors impossible de procéder à une désignation correcte ;
- caméra occultée totalement (passage de la tête du chirurgien devant la caméra, main du chirurgien recouvrant la zone d'intérêt, ...). Il n'est alors pas possible de se prononcer sur la localisation de la zone d'intérêt. Cependant, si l'occultation est partielle, une approximation acceptable est de désigner la zone équivalente à proximité ;
- scène opératoire non délimitée par des champs opératoires (absents ou complètement imprégnés de sang, ...) à l'image de [Bal11]. Ce cas est récurrent pour les opérations du genou où, vu de la caméra, aucun champ n'est aperçu. La désignation ne doit par conséquent pas se baser uniquement sur ce type d'information.

2.1.1.1 Exemple 1 : Intervention superficielle localisée

En observant l'image 2.1a, il est possible de désigner la zone d'intérêt au niveau de l'extrémité du scalpel que tient le chirurgien dans sa main droite. Un test consistant à désigner la zone d'intérêt du chirurgien sur une image a été proposé à 18 personnes n'ayant que très peu de connaissances du domaine chirurgical. Sur l'ensemble des personnes interrogées, deux ont donné un résultat erroné en raison de supposition trop forte sur la nature de l'intervention. Seize personnes (89%) ont correctement identifié la zone d'intérêt. Comment expliquer ce résultat ?

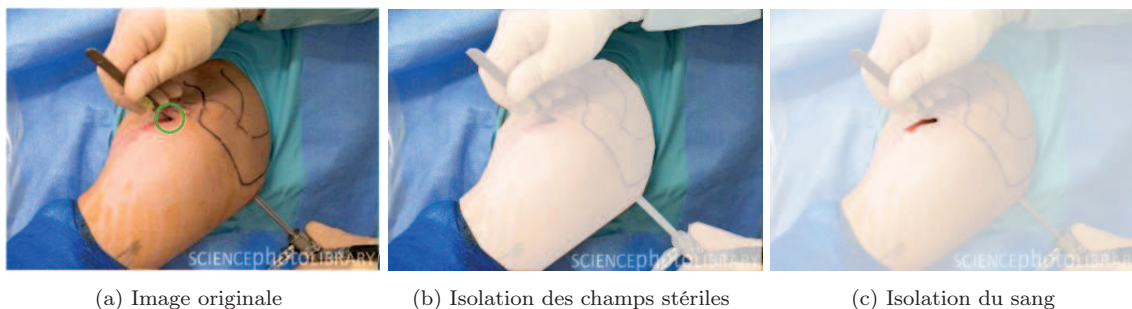


FIG. 2.1 – Isolation des objets dans la scène opératoire, dans le cas d’une chirurgie visant à remédier à une dislocation répétée de l’épaule, le geste en cours consiste à stabiliser (mécaniquement) l’épaule afin de ne pas endommager les ligaments (source : Science Photo Library - Dr. P. Marazzi)

Bien que les personnes interrogées n’ont qu’une expérience très limitée du domaine chirurgical, elles partagent une connaissance a priori sur les éléments qui se trouvent dans la scène :

- les champs opératoires, qui isolent la zone stérile et doivent être à bonne distance de la zone d’intérêt ;
- le chirurgien, qui utilise ses mains munies de gants pour opérer ;
- la peau, visible à l’endroit où il faudra opérer ;
- la couleur rouge, représentative du sang ou d’un élément anatomique (muscle, veine, ...).

Les champs opératoires placés au début de l’intervention (visibles sur la figure 2.1b), fournissent une localisation approximative du lieu où se trouve la zone d’intérêt. Enfin, la petite tache de sang due à la dissection par le scalpel parachève la désignation (figure 2.1c). En l’occurrence, l’utilisation de cette seule information aurait été suffisante pour localiser la zone d’intérêt. Si la petitesse de l’incision permet facilement de désigner la zone d’intérêt, ce type de situation chirurgicale ne représente pour autant qu’un pourcentage marginal des opérations pratiquées.

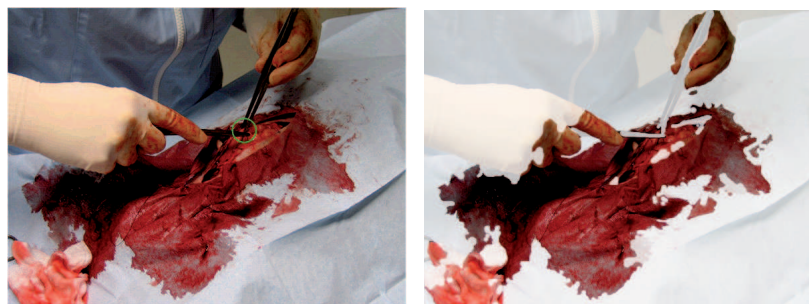
L’association de toutes ces informations explique donc la capacité des personnes interrogées à désigner correctement la zone d’intérêt. Cette mise en commun des informations permet de ne pas être perturbé par :

- l’application localement inhomogène de bétadine sur la peau ;
- les marques pré-opératoires (repères servant au chirurgien mais n’étant pas forcément les lieux des futures incisions) ;
- les différentes couleurs des champs opératoires (la chaussette stérile isolant le bras et l’avant-bras a été volontairement assimilée à un champ opératoire dans cet exemple) ;
- la forme inconnue des instruments manipulés par le chirurgien ;
- l’incapacité à reconnaître la région opérée.

Retenons de cet exemple que le sang permet de désigner la zone d’intérêt du chirurgien et que celle-ci se trouve à distance des champs opératoires.

2.1.1.2 Exemple 2 : Intervention superficielle étendue

L'intervention consiste à refermer une large incision au niveau de l'abdomen. Cette opération est pratiquée sur un chien mais peut s'apparenter à celles effectuées sur l'homme.



(a) Image originale

(b) Isolation manuelle du sang

FIG. 2.2 – Suture de l'abdomen pratiqué sur un chien (source : Azurs point net 2010).

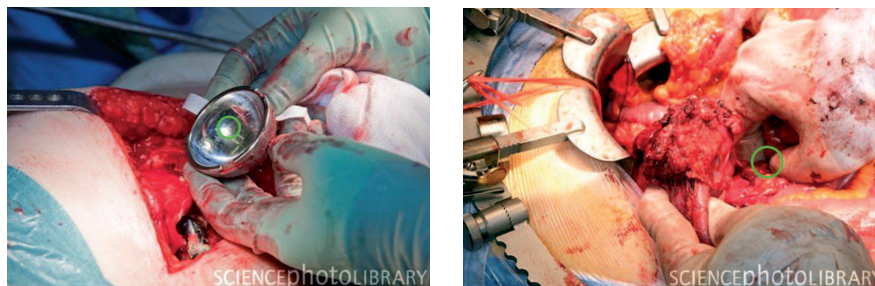
La désignation est plus ardue (comparativement à 2.1) du fait de l'étendue de sang qui ne permet pas de restreindre la localisation à une région de faible dimension. Les champs stériles qui servent à border la plaie opératoire ne sont pas reconnaissables dans leur intégralité en raison de l'imprégnation du sang. Notons que ces derniers sont de couleur différentes de ceux présentés dans l'exemple précédent et se différencient mal de la blouse du chirurgien.

Par conséquent, la désignation de la zone d'intérêt à l'aide de ces deux seules informations paraît compromise. L'identification des gants, facilement reconnaissables, malgré une partielle dénaturation de leur couleur à cause du sang recouvrant les extrémités, ne permet qu'une localisation approximative.

Les instruments sont déterminants et permettent la localisation de la zone d'intérêt qui se trouve au concours de leurs extrémités. A l'instant de la prise de vue, le chirurgien est amené à tirer le tissu adipeux avec sa main gauche et suturer avec sa main droite, ce qui confirme que la zone d'intérêt se trouve à cet endroit.

Le sang reste donc une source d'information intéressante mais insuffisante dans cet exemple. Les champs opératoires permettent de limiter l'espace de recherche sans pour autant donner des indices sur la localisation précise de la zone d'intérêt. Ils ne sont pas nécessairement verts, et peuvent être confondus avec d'autres éléments de la scène (blouse notamment). Enfin, les instruments permettent en l'occurrence de bien cibler la zone d'intérêt.

2.1.1.3 Exemple 3 : Intervention en cavité



(a) Remplacement d'une prothèse de hanche. La zone d'intérêt est la prothèse elle-même (source : Science Photo Library)

(b) Duodéno pancréatectomie. La zone d'intérêt se trouve entre l'index et le pouce droit (source : Science Photo Library)

FIG. 2.3 – Partie d'une intervention chirurgicale se déroulant sans utilisation d'instruments chirurgicaux.

Dans l'illustration 2.3a, aucun instrument n'est présent dans la scène chirurgicale. Nous notons la présence d'un écarteur (considéré comme un outil chirurgical et non comme un instrument) sur le côté gauche de l'image. Ce dernier ne nous renseigne pas sur la localisation de la zone d'intérêt. De même les champs stériles verts (au premier plan) et bleus (en arrière plan) n'offrent pas plus d'indication sur la position de la zone d'intérêt. Amalgamés aux gants du chirurgien, ils conduisent dans cet exemple à une erreur. En effet, la zone d'intérêt est la prothèse tenue au bout des doigts du chirurgien. L'une des hypothèses est que la zone d'intérêt doit se trouver à bonne distance des champs opératoires. Or, confondre les gants du chirurgien avec les champs stériles peut amener à ne pas désigner la prothèse comme zone d'intérêt. Différencier les gants du chirurgien des champs stériles apparaît donc indispensable pour ne pas prendre en compte ce type d'épiphénomène.

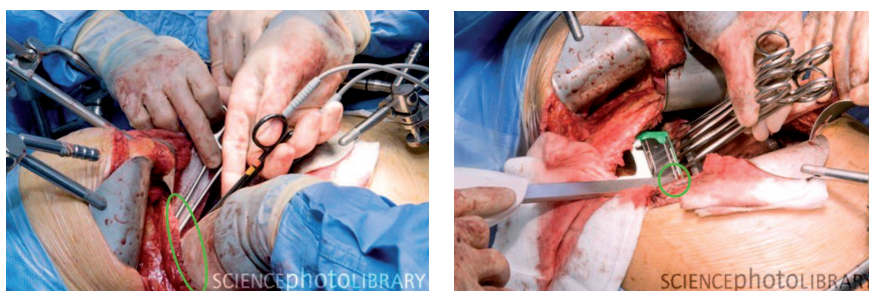
Le sang demeure une source d'information pertinente bien qu'approximative, comme mentionné pour les exemples 2.1 et 2.2. Compte-tenu des éléments identifiés, il n'est pas possible de désigner de manière précise la zone d'intérêt du chirurgien. Identifier la prothèse revient à reconnaître les éléments d'apparence métallique dans la scène. Ceci peut nous conduire à désigner la zone d'intérêt sur un écarteur voire sur un ciseau à clamper tenant les champs stériles, qui sont des outils chirurgicaux et qui n'ont par définition pas vocation à intervenir sur la zone d'intérêt.

La même analyse peut être faite sur l'image 2.3b qui concerne un cancer du pancréas où seuls des écarteurs sont présents à la périphérie de l'image (notons la différence de forme et de couleur de ces écarteurs par rapport à 2.3a). Il est manifeste que les conditions d'illumination ne sont pas optimales, la zone d'intérêt n'étant que très faiblement éclairée. Les gants sont cette fois d'une couleur s'apparentant à celle d'un os (blanc crème) et permettent difficilement la localisation de la zone d'intérêt. Le sang permet là encore d'approcher la position de la zone d'intérêt. La peau et les champs stériles ne permettent pas de désigner avec précision.

Dans ces deux exemples, nous pouvons constater qu'une zone d'intérêt peut exister malgré l'absence d'instrument dans la scène chirurgicale. Ceux-ci ne doivent pas être confondus avec les écarteurs sous

peine de procéder à des désignations erronées. La peau et les champs stériles permettent une localisation imprécise de la zone d'intérêt. Le sang ne permet pas plus d'affiner cette désignation. Les mains du chirurgien, notamment durant les phases de palpation et de manipulation de prothèse semblent d'intérêt. Cependant, leur utilisation comme indices visuels soulève la question de leur variabilité.

2.1.1.4 Exemple 4 : Intervention en cavité avec instruments multiples



(a) Opération de l'œsophage. La zone d'intérêt n'est pas visible de ce point de vue (source : Science Photo Library)

(b) Opération de l'œsophage. L'utilisation de nombreux instruments ne permet pas la localisation précise de la zone d'intérêt (source : Science Photo Library)

FIG. 2.4 – Difficulté de désigner la zone d'intérêt à l'aide uniquement des instruments chirurgicaux.

Les figures 2.4a et 2.4b illustrent la même intervention chirurgicale. On peut noter que le point de vue influence fortement le résultat de la désignation. En effet, la zone d'intérêt dans 2.4a n'est pas visible ; elle est seulement localisée à l'aide du sang. En l'occurrence, celle-ci se trouve à proximité de la main gauche du chirurgien au premier plan. Ici, il aurait été nécessaire d'extrapoler la pointe des instruments.

Cependant, déterminer où se trouve la pointe de l'instrument ne facilite pas toujours la désignation comme le montre la seconde prise de vue (2.4b) où de nombreux instruments sont présents. En effet, de prime abord, la zone d'intérêt semble se trouver au niveau des ciseaux que tient le chirurgien dans sa main gauche, alors qu'en réalité elle se situe à la pointe de la tige qu'il manipule avec sa main droite. Bien que l'erreur de désignation soit ici relativement faible en raison de la proximité de ces éléments, elle aurait constituée une faute notable si ceux-ci s'étaient trouvés distants.

En résumé :

Ces exemples montrent qu'il n'existe pas, dans la scène chirurgicale, un objet facilement reconnaissable et caractéristique de la zone d'intérêt du chirurgien. Identifier simultanément plusieurs objets semble donc un moyen efficace pour d'une part procéder à une désignation précise et d'autre part, pallier la diversité des actes opératoires comme le résume le tableau 2.1.

Type d'objet	Intérêt	Inconvénient
Champ stérile	Limite la localisation de la zone d'intérêt à une portion de l'espace	Variabilité des matériaux et des couleurs utilisés
Gant	Permet de localiser la zone d'intérêt dans les cas où le chirurgien travaille directement avec les mains	Le chirurgien utilise souvent des instruments
Instrument	Le repérage de l'extrémité de l'instrument est révélateur d'une zone d'intérêt	Présence parfois multiple d'instruments induisant une localisation erronée
Ecarteur	Permet potentiellement de localiser les cavités opératoires	Très forte diversité
Peau	Limite la localisation de la zone d'intérêt à une portion de l'espace	La dégradation de la peau (inhomogénéité de la bétadine, dépôt de sang, . . .) ainsi que la variabilité de la couleur ne permettent pas un repérage facile
Sang	Limite la localisation de la zone d'intérêt à une portion de l'espace	Détérioration de l'aspect du sang (coagulation), recouvrement d'autres objets (gants, champs stériles)
Muscle	Indicateur des lieux d'incisions	Impossible à différencier du sang
Organe	Indique que la phase opératoire est une cavité	Aspect soumis à une très forte variabilité
Graisse	Indique les endroits synonymes de voie d'abord	Similarité de la graisse avec la peau bétadinée

TAB. 2.1 – Qualité des objets présents dans la scène chirurgicale pour la désignation de la zone d'intérêt.

2.1.2 Restriction du nombre d'objets à identifier

Dans un cas idéal, la reconnaissance de tous les objets assurerait la robustesse de l'application face à la variabilité des actes chirurgicaux. Ceci impose en contrepartie une grande capacité de calcul et pose par conséquent le problème d'embarquabilité et de réactivité du système. Il est donc raisonnable de rechercher un compromis n'utilisant qu'une partie de la liste d'objets reconnaissables présentée plus haut. Le but est donc d'éliminer ou de regrouper certains objets en classes d'objets afin d'alléger les futurs traitements d'images.

Classes d'objets

Des classes d'objets sont des regroupements d'objets partageant une même caractéristique. Pour [RPR05] il s'agit d'activités d'un athlète comme "courir", "sauter", "chuter", "se relever" (séquence d'états); pour [YK06] ce sont des erreurs systèmes telles que : "condition normale", "déséquilibre mécanique", "casse de l'arbre moteur", "désalignement angulaire", etc.

Dans notre cas, nous pouvons distinguer les classes : champ stérile, gant, instrument, écarteur, peau, sang, muscle, organe et graisse. Les muscles, les organes et le sang, qui ont des apparences proches, peuvent être regroupés en une seule classe que nous appelons "méta-classe sang" ou par abus de langage simplement classe sang. Ce regroupement évite le recours à des algorithmes de traitements d'images particuliers pour distinguer spécifiquement les muscles des organes et du sang, ce qui permet de gagner du temps de calcul. Il est donc intéressant d'étendre cette réflexion aux autres classes.

Classes non-biologiques

Parmi les neuf objets les plus visibles lors d'une opération, nous évinçons d'emblée les écarteurs dans la mesure où ils ne permettent ni de désigner la zone d'intérêt, ni d'en limiter la localisation. Ils contribuent uniquement à nous renseigner sur la présence de cavités dans la scène opératoire. De plus, aux vues du très grand nombre de produits actuellement sur le marché (dont les formes, les couleurs et l'emploi sont très variables), une trop grande part de cette étude serait nécessaire pour élaborer un traitement permettant leur reconnaissance.

Cette dernière objection est également valable pour les instruments. Cependant, à la différence des écarteurs, les instruments participent directement aux gestes de dissection. Ce sont donc de très bons indices pour la localisation de la zone d'intérêt. Compte-tenu de la diversité des instruments, la solution retenue consiste à limiter la classe instrument aux scalpels et aux objets s'y apparentant tels que des ciseaux. Il faut cependant utiliser ce type d'information avec précaution dans la mesure où dans certaines situations, les instruments peuvent être temporairement posés loin de la zone d'intérêt. Ceci est fréquemment vérifié durant les phases de palpation où il est plus facile pour le chirurgien de placer l'instrument sur le côté de la scène chirurgicale plutôt que de le remettre sur la table dédiée. Dès lors, le traitement d'image devra être capable de différencier les instruments "actifs", utilisés par le chirurgien de ceux "passifs" posés dans la scène chirurgicale.

Les champs stériles permettent de restreindre l'espace de recherche de la zone d'intérêt. Il est donc intéressant de conserver ce type d'information pour garantir la fiabilité de la désignation. En effet, l'un des risques est la désignation en dehors de la scène opératoire, qui entraînerait un mouvement inadéquat

du système. Dès lors la limitation de la localisation à une portion d'espace rend l'algorithme plus fiable. Même s'il existe un grand nombre de tailles et de textures (figure 2.5a), la plupart des champs stériles reste dans une palette de couleur allant du vert au bleu (figure 2.5d). Nous ne retiendrons que ces derniers pour la classe champ stérile.

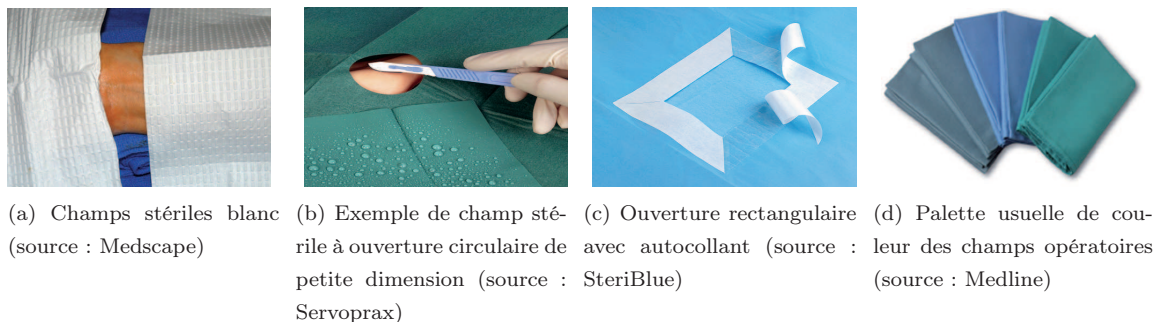


FIG. 2.5 – Variabilité des champs stériles (dimension, forme, ouverture, couleur, matériaux).

Dernier objet non biologique du cadre de discernement, les gants sont à la fois source d'information intéressante et source d'erreur. Outre la difficulté liée à leur reconnaissance compte tenu de leur dégradation par le sang l'utilisation des informations relatives à ces objets doit être considérée avec réserve (une solution possible consiste à imposer une couleur spécifique de sorte à faciliter leur reconnaissance, toutefois ceci apparaît difficilement acceptable pour le corps chirurgical ; cette solution est donc écartée). Comme illustré en figure 2.4b, le chirurgien est parfois amené, au cours de l'opération, à exécuter des actions spécifiques avec ses mains comme palper, éponger le sang avec une compresse, mettre en place une prothèse, etc. Dans ces cas précis, identifier l'extrémité des doigts peut permettre de localiser très précisément la zone d'intérêt. Paradoxalement, leur localisation peut conduire, en un temps très court, à une mauvaise désignation. En effet, la majorité des actions se déroule à l'aide d'instruments (scalpel entre autres), localiser l'extrémité des doigts n'a, dans ces circonstances, plus aucune signification. L'utilisation de ce type de renseignement étant trop sujette à discordance, il a donc été décidé d'éliminer cette classe.

Classes biologiques

La graisse représente un indicateur fiable des situations en cavité tout comme les écarteurs mais ne contribue pas à la désignation. Contrairement à ces derniers, son aspect est unique et ce quelle que soit l'origine du patient. Cependant, son utilisation est sujette à caution. Dans la section 1.2.2, il a été mentionné que la bétadine appliquée sur une peau claire la rend jaune-orangée. Cette teinte pose problème car elle s'apparente à celle de la graisse, ce qui faute d'information supplémentaire ne permet pas de les différencier. Dès lors, des erreurs d'appréciation de la situation opératoire peuvent survenir et entraîner des fausses décisions. Il faut prêter la plus grande vigilance quant à son utilisation. En conséquence, la classe graisse n'a pas non plus été retenue.

La peau permet de restreindre l'espace de recherche de la zone d'intérêt au même titre que les champs opératoires. Or, comme nous venons de le voir, il est possible de la confondre avec la graisse, sauf dans le cas d'un patient à peau sombre. Dans le cadre de cette étude, nous procéderons à des tests animaliers notamment sur des béliers et des cochons. Leur peau étant de couleur claire, la classe peau sera donc

restreinte à cette teinte uniquement. Cependant, il serait nécessaire pour une application réelle de ces recherches, d'augmenter la granularité de cette classe.

Dernier de la liste, le sang figure parmi les objets les plus intéressants pour la désignation d'une zone d'intérêt. Sa présence est synonyme de dissection, action récurrente lors d'une opération qui correspond justement à la zone d'intérêt à ce moment précis. Le sang est donc un indice probant mais a pour inconvénient d'imprégner les autres éléments de la scène notamment les gants, les champs stériles, la peau ainsi que les muscles et les organes. Ceci nuit donc à la précision de la désignation et peut dans certains cas amener à des désignations contre-intuitives. D'autre part, le sang a pour particularité de s'assécher au cours du temps, passant d'une couleur rouge vive à marron (figure 5.21a). Discerner le sang "frais" permet de s'affranchir de ce type de problème. En conséquence, la classe sang est réduite au sang frais. Comme il a été fait mention dans l'introduction aux classes d'objets, la classe sang intégrera les muscles et les organes afin de faciliter le traitement d'image.

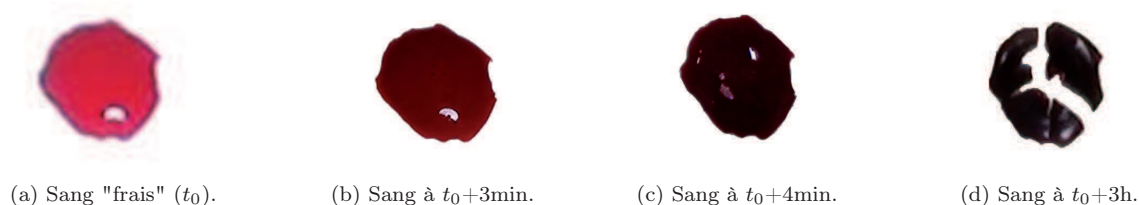


FIG. 2.6 – Evolution de la couleur d'une goutte de sang au cours du temps.

En résumé :

Les classes considérées pour cette étude sont :

1. *les champs stériles (gamme de couleurs allant du bleu au vert conformément à la figure 2.5) qui permettent de restreindre l'espace de recherche ;*
2. *les instruments chirurgicaux similaires aux scalpels afin de repérer l'activité du chirurgien ;*
3. *la peau dont la couleur est proche de celle d'une "peau claire" pour faciliter les essais en bloc opératoire et restreindre l'espace de recherche ;*
4. *le sang incluant le sang non coagulé synonyme de nouvelle dissection ainsi que les muscles et organes.*

2.1.3 Processus de reconnaissance d'objet

Il s'agit maintenant de définir les moyens permettant de reconnaître ces objets. Pour trouver des éléments de réponse, nous allons imiter le processus de raisonnement humain pour la conception de la partie intelligente du système [Sim57] :

Etape 0 : Initialisation

Création d'une base de données contenant les indices visuels de référence pour chaque classe d'objet.

Etape 1 : Renseignement

Acquisition de l'image de la scène chirurgicale.

Etape 2 : Dissociation des informations

Quantifier des informations "couleur", "mouvement",...

Etape 3 : Fusion

Recomposition sous une forme unique des informations.

Etape 4 : Désignation

Définition de la zone d'intérêt la plus probable pour le chirurgien.

L'étape 0 rassemble l'ensemble de la connaissance a priori sur la scène opératoire et l'environnement. Cette étape précède l'acquisition d'images assurée par la caméra qui observe la scène chirurgicale (étape 1). Ce flux est traité en étape 2 à la recherche de chaque classe d'objets. Chaque traitement d'images lié à la "couleur", au "mouvement", etc, nous donne une image. Il y a donc autant d'images que de types d'information traités. Ces images sont ensuite fusionnées en étape 3 pour n'en former qu'une, plus précise : nous parlons alors de fusion d'attributs. A partir de cette image, nous procédons à une prise de décision (étape 4), basée sur un ensemble de critères que nous détaillerons ultérieurement.

Le processus utilise pour la désignation de la zone d'intérêt une reconnaissance d'objets basée sur plusieurs types d'informations (forme, couleur, mouvement, etc). Il est toutefois possible d'y parvenir en ne se basant que sur un seul type. Cependant, les performances obtenues sont généralement moindres notamment dans les cas où :

1. les informations sont de même nature mais proviennent de plusieurs sources dont les performances et les domaines de validité sont différents ;
2. l'objet ne peut pas être facilement discriminé par un seul type d'information (couleur, mouvement, ...);
3. les informations ne sont pas persistantes (indisponibilité de l'information récurrente et représentative des objets).

Nous nous intéressons maintenant à l'étape 2 et plus particulièrement à la sélection des informations qui seront utilisées pour reconnaître les objets. L'étape 1 sera détaillée dans le chapitre 4 présentant le prototype développé durant ce travail. Nous restreindrons notre approche à la reconnaissance d'objets sans prise en compte de l'aspect temporel afin de prendre en compte l'aspect erratique de l'intervention chirurgicale (décorrélation du geste opératoire et de la désignation de la zone d'intérêt).

2.2 Classification d'objets en milieu chirurgical

Les classes d'objets à reconnaître étant définies, il est maintenant nécessaire d'extraire les attributs (forme, couleur, mouvement) permettant de les identifier. L'idée est de déterminer le jeu minimal d'attributs nécessaire pour identifier les objets tout en limitant le temps de calcul.

Remarque : Comme nous pourrions le voir sur les images qui suivent, l'apparence des objets est très variable en raison des changements d'éclaircement, des modifications des propriétés matériaux (assèchement, coagulation naturelle, etc), et de l'activité du chirurgien (coagulation forcée, dissection). En conséquence, nous ne retiendrons pas dans cette étude l'attribut texture au sens d'Haralick ([TSS⁺08]).

2.2.1 L'attribut "forme"

La forme est une information riche mais non structurée qui permet dans certaines situations d'identifier des objets dans une scène. Point pénalisant, ce type d'attribut est non redondant et affecté par le bruit (de la caméra, les projections, la fumée, etc), les distorsions de la caméra ainsi que les transformations perspectives (point de vue de la caméra par rapport à la scène opératoire inconnue) et les occultations (tête du chirurgien notamment). Présent dans des domaines aussi diversifiés que la reconnaissance d'écriture ([MB10]), de symbole ([NTZH10]), gestuelle ([HHW05]), la biométrie ([Cho07]), le médical ([SDT09]), le satellitaire ([CY09]), l'attribut forme s'appuie sur des techniques tout aussi variées.

Le point commun à tous ces travaux est l'utilisation des caractéristiques quantifiables que l'on appelle descripteurs dont [PI97] propose un résumé.

Descripteurs simples

Il existe des descripteurs simples capables de décrire la forme d'un objet (figure 2.7). Dans la littérature, nous rencontrons :

- la divergence rectangulaire (rectangular discrepancy, appelée aussi rectangularité) qui est donnée par le rapport de l'aire occupée par l'objet sur celle de sa boîte englobante (plus petit rectangle entourant l'objet), dont [Ros03] en propose une version améliorée ;
- la convexité qui est une description plus fine que la rectangularité et qui se base sur le rapport du périmètre de l'objet sur celui de l'enveloppe (convexe) ;
- la compacité (compactness) qui correspond au périmètre de l'objet au carré sur son aire [SK78] ;
- l'excentricité qui est définie comme étant le rapport de la longueur de l'axe principal sur celle de l'axe mineur (parfois appelé secondaire).

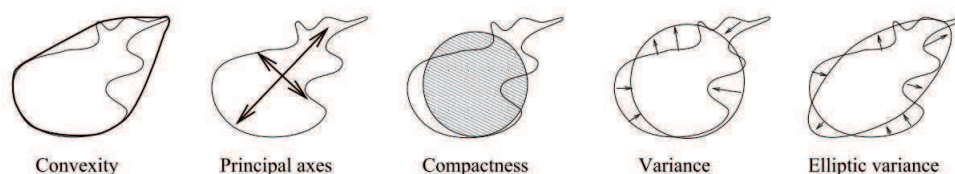


FIG. 2.7 – Exemple de descripteurs de forme (source : [PI97])

Ces descripteurs sont invariants en translation, rotation (autour de l'axe optique) et peu gourmands en temps de calcul. Cependant, ils ne sont capables de discriminer que des objets rigides et peu soumis aux changements de points de vue, ils ne sont donc pas adaptés pour cette application.

Moments invariants

Suite aux travaux de Hu ([Hu62]), les moments permettent de décrire les objets 2D (voir [NK04] pour une extension vers le 3D) sans spécificité particulière (exemple d'application en asservissement visuel d'objet [TC03]). L'expression du moment M_{pq} d'une fonction $f(x, y)$ (définie dans \mathbb{R}) à l'ordre $p + q$ ($p, q \in \mathbb{Z}$) s'écrit :

$$M_{p,q} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^p y^q f(x, y) dx dy \quad (2.1)$$

L'expression (2.1) appliquée à une image I se reformule par :

$$M_{pq} = \sum \sum_D x^p y^q I(x, y) \quad (2.2)$$

- D étant l'espace occupé par l'objet ;
- $I \in [0, 255]$ l'intensité du pixel de coordonnées (x, y) .

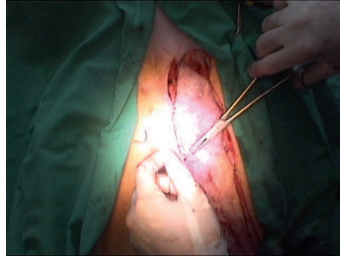
L'ordre 0 permet la détermination de l'aire de l'objet ($M_{0,0}$). Les ordres supérieurs nous permettent d'accéder à des informations plus précises telles que la position du centre de gravité si l'on parle de l'ordre 1 ($x_g = \frac{M_{1,0}}{M_{0,0}}$ et $y_g = \frac{M_{0,1}}{M_{0,0}}$) ou son orientation si l'on parle de l'ordre 2 ($M_{1,1}$, $M_{0,2}$ et $M_{2,0}$). Les ordres plus élevés sont plus difficilement appréhendables, citons néanmoins l'ordre 3 qui correspond à l'asymétrie de l'objet et l'ordre 4 qui quantifie l'aplatissement de la distribution de l'objet (kurtosis).

Pour obtenir les invariants en translation, le calcul des moments est effectué au niveau du centre de gravité noté (x_g, y_g) de l'objet. On parle alors de moments centrés, l'expression (2.2) s'écrit alors :

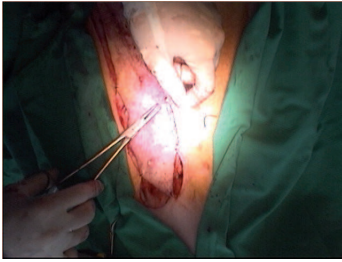
$$m_{p,q} = \sum \sum_D (x - x_g)^p (y - y_g)^q I(x, y) \quad (2.3)$$

L'invariance à l'échelle est obtenue en divisant les termes m_{ij} par $m_{0,0}^{1+\frac{(p+q)}{2}}$ noté alors η_{ij} . On parle alors de moments centrés normalisés. Ces invariances (translation, changement d'échelle) sont cependant mises en échec lors de la rotation de l'objet. Afin d'y pallier, des moments dits invariants ont été développés par [Hu62] qui permettent d'être, en plus, invariants en réflexion (uniquement jusqu'au sixième ordre). Nous présentons dans le tableau 2.2 les valeurs obtenues pour chaque ordre pour différentes perturbations appliquées à l'image (rotation, réflexion, dézoom), conformément aux équations (2.4). Nous pouvons constater que malgré ces modifications de l'image d'origine (figure 2.8), les moments décrivant chacune des images restent constants.

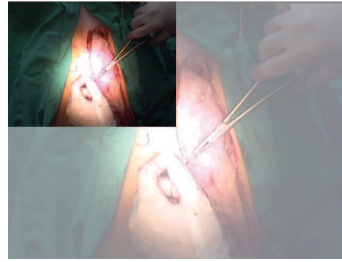
Transformation	ϕ'_1	ϕ'_2	ϕ'_3	ϕ'_4	ϕ'_5	ϕ'_6	ϕ'_7
image d'origine	2,8588	7,2608	9,9155	10,115	-22,772	-14,37	20,131
dézoom (50%)	2,8589	7,2601	9,9142	10,112	-21,975	-14,391	20,125
rotation (13°)	2,8588	7,2606	9,9155	10,116	-22,733	-14,37	20,131
réflexion	2,8588	7,2608	9,9155	10,115	-22,772	-14,37	20,131

TAB. 2.2 – Moments calculés ($\phi'_i = -\text{signe}(\phi_i) \log_{10}(|\phi_i|)$)

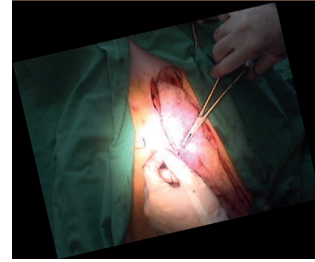
(a) Image d'origine.



(b) Réflexion.



(c) Réduction de 50%.



(d) Rotation de 13°.

FIG. 2.8 – Perturbations appliquées à l'image d'origine.

$$\begin{aligned}
\phi_1 &= \eta_{2,0} + \eta_{0,2} \\
\phi_2 &= (\eta_{2,0} - \eta_{0,2})^2 + 4\eta_{1,1}^2 \\
\phi_3 &= (\eta_{3,0} - 3\eta_{1,2})^2 + (3\eta_{2,1} - \eta_{0,3})^2 \\
\phi_4 &= (\eta_{3,0} + \eta_{1,2})^2 + (\eta_{2,1} + \eta_{0,3})^2 \\
\phi_5 &= (\eta_{3,0} - 3\eta_{1,2})(\eta_{3,0} + \eta_{1,2}) [(\eta_{3,0} + \eta_{1,2})^2 - 3(\eta_{2,1} + \eta_{0,3})^2] + (3\eta_{2,1} - \eta_{0,3})(\eta_{2,1} + \eta_{0,3}) \\
&\quad [3(\eta_{3,0} + \eta_{1,2})^2 - (\eta_{2,1} + \eta_{0,3})^2] \\
\phi_6 &= (\eta_{2,0} - \eta_{0,2}) [(\eta_{3,0} + \eta_{1,2})^2 - (\eta_{2,1} + \eta_{0,3})^2] + 4\eta_{1,1}(\eta_{3,0} + \eta_{1,2})(\eta_{2,1} + \eta_{0,3}) \\
\phi_7 &= (3\eta_{2,1} - \eta_{0,3})(\eta_{3,0} + \eta_{1,2}) [(\eta_{3,0} + \eta_{1,2})^2 - 3(\eta_{2,1} + \eta_{0,3})^2] + (3\eta_{2,1} - \eta_{0,3})(\eta_{2,1} + \eta_{0,3}) \\
&\quad [3(\eta_{3,0} + \eta_{1,2})^2 - (\eta_{2,1} + \eta_{0,3})^2]
\end{aligned} \tag{2.4}$$

Le lecteur pourra se rapporter aux travaux sur les polynômes de Legendre ([CRM04]), de Teague ([Tea80]), de Zernike ([RVS11]) et de Fourier-Mellin ([GZ10]) pour d'autres modalités de calcul des moments et qui sont appliqués à des domaines aussi divers que la reconnaissance de caractère, l'authentification de signature, la détection de masses malignes ou encore la reconnaissance de cellules. Notons que certains travaux s'intéressent à l'utilisation couplée de différents types de moments à l'image de [KS02].

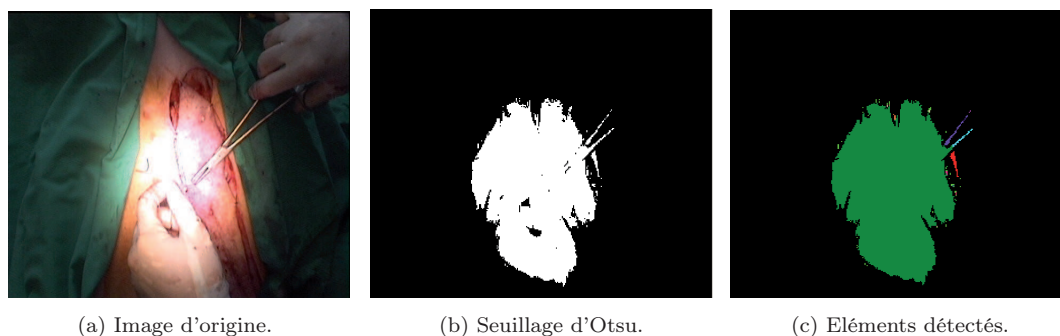


FIG. 2.9 – Exemple de détection d'objets par segmentation.

Dans l'exemple (figure 2.8), le calcul des moments est effectué dans l'image entière et non pour chaque élément dans la scène. Cette démarche reste donc tributaire du traitement d'image permettant d'isoler les objets. Un moyen simple consiste à seuiller l'image. Le seuillage d'Otsu est une méthode éprouvée mais qui dans notre cas ne donne pas satisfaction comme le montre la figure 2.9. La possibilité d'utiliser des techniques plus élaborées n'a pas été retenue compte-tenu du coût calculatoire que cela impose. En conséquence, l'utilisation des moments dans cette étude est écartée.

Signature de contour

La signature de contour est la compilation des distances séparant le centre de gravité (trouvé grâce au moment d'ordre 1 par exemple) des bords de l'objet. A titre d'exemple, nous proposons des exemples de signature de contours pour des objets à géométrie simple (figure 2.10). La signature obtenue est comparée à une base de données contenant celles d'autres objets. On associe à l'objet observé l'objet de la base de données dont le profil est le plus proche. Ce type d'approche a été utilisé dans [NIC⁺07] (comptage de cellules embryonnaires), plus récemment dans [GGMT10] (localisation de la pupille d'un individu) et [GJGMAR⁺11] (reconnaissance de panneau de signalisation), où sont mises en avant les propriétés d'invariance en translation et rotation (dans le plan image) ainsi que la capacité à montrer les imperfections des objets. Au final, la signature apparaît comme une fonction continue où la rotation de l'objet et le changement d'échelle engendrent respectivement le déphasage et la dilatation de sa signature en supposant que l'objet est convexe, afin d'éviter de voir apparaître plusieurs points de contour pour un même angle. Ceci implique de multiplier les signatures pour un même objet, la base de données est donc nécessairement de taille importante. Nous écartons donc ce type de technique de cette étude même pour les objets les plus courants.

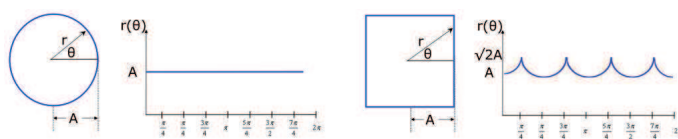


FIG. 2.10 – Exemple de signature de contour (source : [GGMT10])

Transformée de Hough

Utilisée pour la reconnaissance de droites ([JCS11], voir [HHM⁺07] pour une extension vers la reconnaissance de cercle), la transformée de Hough repose sur une détection de contours. Celle-ci repose sur l'idée que pour une image I (supposée continue), le contour correspond à une variation brutale de l'intensité des pixels (ce qui a des répercussions sur la dérivée première et la dérivée seconde de l'intensité dans l'image). Ces discontinuités sont générées par les bords de l'objet. Cependant, les conditions d'illumination (ombres, surexposition) et l'inhomogénéité de l'objet lui-même (texture, parasite) peuvent induire également des discontinuités et gêner le processus de reconnaissance de l'objet. Le lecteur pourra trouver dans [RSS11], [SRLA11] et [YTS11] des exemples récents d'applications où la détection de contours est utilisée.

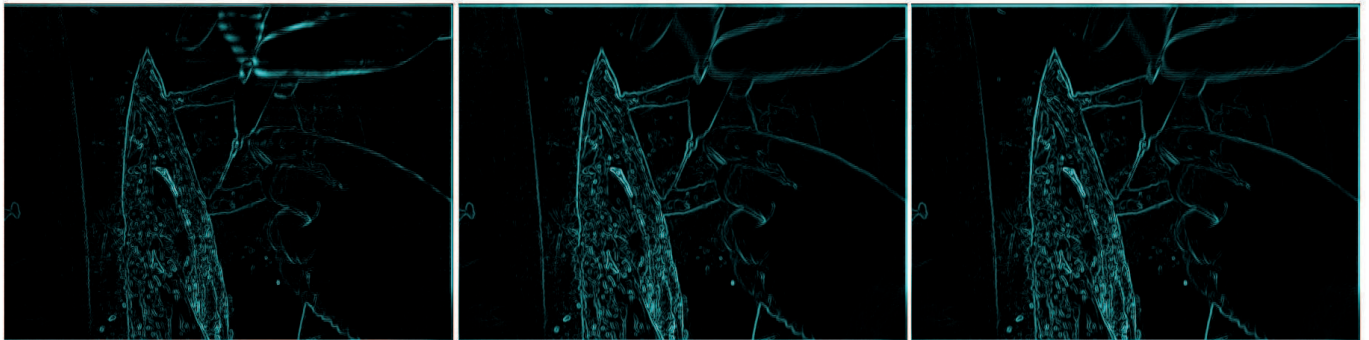
De la bonne détection de ces variations dépend donc le résultat de la localisation des contours. Pour ce faire, nous cherchons les endroits où le gradient est maximal dans la direction de \vec{g} définie par : $\vec{g} = \frac{\vec{\nabla}I}{|\vec{\nabla}I|}$, où $\vec{\nabla}I$ et $|\vec{\nabla}I|$ correspondent respectivement au gradient et au module du gradient de l'image ($\vec{\nabla}I = \left[\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right]^T$ et $|\vec{\nabla}I| = \left[\left(\frac{\partial I}{\partial x} \right)^2 + \left(\frac{\partial I}{\partial y} \right)^2 \right]^{\frac{1}{2}}$).

Rechercher les points du contour équivaut à trouver les extrema de la dérivée première de l'intensité des pixels ainsi que les passages à zéro de la dérivée seconde. Une manière simple permettant de contourner les difficultés d'une résolution analytique consiste à convoluer l'image par un masque h_o dans les directions \vec{x} et \vec{y} : $\frac{\partial I}{\partial o}(x, y) \approx (I * h_o)(x, y)$. Les masques les plus connus sont ceux de Roberts ([Rob63]), Sobel ([Sob78]), Prewitt ([Pre70]), Kirsh (1^{er} masque, [Kir71]) et Robinson ([Rob77]).

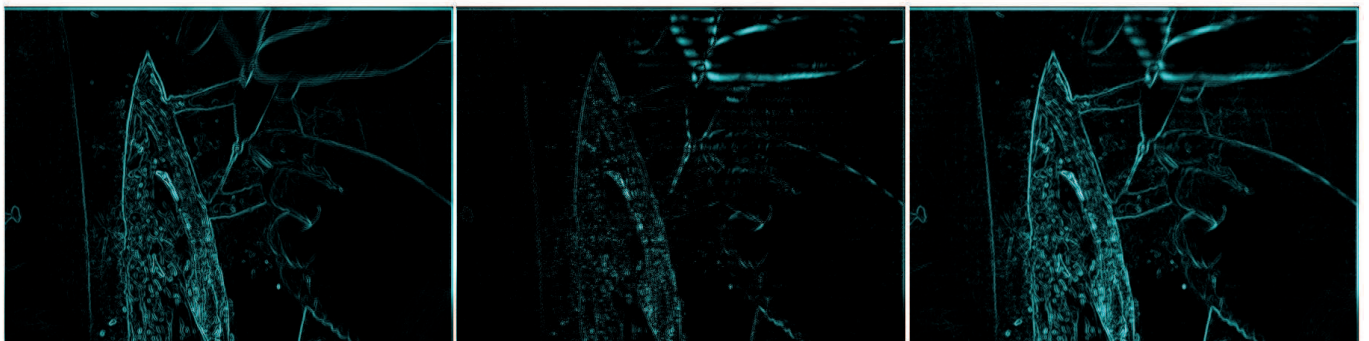
Les masques utilisés sont en général de taille 2×2 ou 3×3 mais peuvent être de plus grande dimension. Ceci est généralement valable pour les cas d'images fortement bruitées, mais requiert un temps de calcul plus conséquent. Une fois les calculs effectués, il est nécessaire d'approximer le calcul de $|\vec{\nabla}I|$ (module du gradient) par la norme L_1 ou L_2 et ainsi localiser les contours de l'objet : $|\widehat{\nabla}I| = |(I * h_x)(x, y)| + |(I * h_y)(x, y)|$ (norme L_1).

Dans le cas d'une image de scène opératoire, la détection de contours ne donne pas satisfaction (figure 2.11). En effet, il y a fusion de certains contours d'objets pourtant distincts. Les champs stériles facilement isolables dans l'image d'origine sont difficiles à distinguer pour l'ensemble des filtres. Seuls les contours de la plaie, des mains et de l'instrument sont potentiellement isolables malgré l'entrelacement de l'image.

La transformée de Hough repose sur le passage de l'espace image à l'espace dit des paramètres. L'idée sous-jacente est qu'une droite formée par les points du contour de coordonnées $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ est représentée dans l'espace des paramètres par un point unique (a_1, b_1) tel que : $y = a_1x + b_1$. Inversement, une droite dans l'espace des paramètres correspond un point unique dans l'espace image. Pour trouver (a_1, b_1) , nous déterminons donc la droite représentative de chaque point de contour (x_i, y_i) . Le croisement des droites dans l'espace des paramètres correspond au couple (a_1, b_1) recherché.



(a) De gauche à droite : masque de Roberts, Sobel et Prewitt.



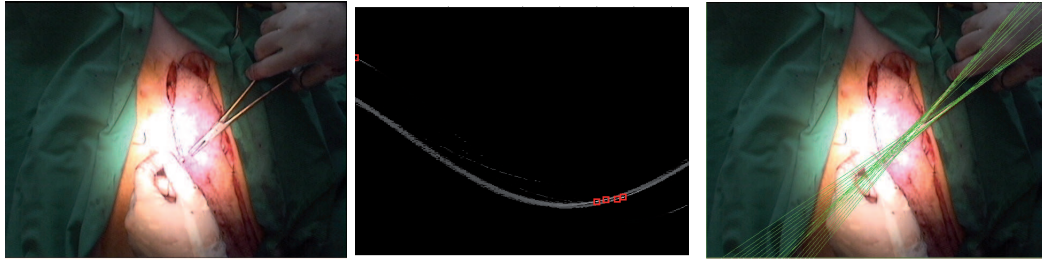
(b) De gauche à droite : gradient, Laplacien et gradient morphologique.



(c) Image originale.

FIG. 2.11 – Exemple de détection de contours.

L'utilisation de droites comme mode de représentation est imparfait car les droites verticales ne peuvent être représentées ($a \rightarrow \infty$). Le passage en coordonnées polaires permet de passer outre ce problème. L'écriture $y = a_1x + b_1$ devient $\rho_1 = x \cos(\theta_1) + y \sin(\theta_1)$ avec (ρ_1, θ_1) les deux coefficients représentatifs de la droite passant par $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$.



(a) Image originale.

(b) Exemple d'espace des paramètres.

(c) Détection des droites.

FIG. 2.12 – Exemple de détection de la classe instrument par transformée de Hough.

Ce mode de représentation implique l'apparition de sinusoides dans l'espace des paramètres (figure 2.12b). Les maxima symbolisés par des rectangles sur cette image correspondent aux paramètres des droites visibles (figure 2.12c). Nous pouvons voir que malgré des discontinuités de contours (figure 2.11) et le bruit dû à l'entrelacement, la détection est précise. Compte-tenu de ces résultats, il a été décidé de conserver la transformée de Hough pour la détection des instruments dans la scène chirurgicale. Cependant, notons qu'il est possible que des éléments droits soient considérés comme des instruments potentiels comme par exemple le bord d'une main. Il apparaît donc nécessaire de combiner cette méthode avec d'autres afin de restreindre la détection aux instruments seuls afin d'éliminer ce type d'erreur.

Signature spectrale

Les descripteurs de Fourier permettent de passer outre les obstacles liés à une mauvaise détection du contour ou à la non convexité de l'objet. L'idée repose sur une décomposition en séries de Fourier de la signature (fonction périodique) précédente et sur une identification de ses coefficients a_j donnés par : $SF(\psi) = \sum_{j=0}^{\infty} a_j e^{-ij\theta}$ (ψ est une variable réduite à valeurs dans $[0, 2\pi]$).

Une autre manière de procéder est de diviser le contour en N parties égales, puis de calculer la distance $z(p)$ ($p \in [0, N - 1], p \in \mathbb{N}^2$) séparant chacune des extrémités des segments, du centre de gravité de l'objet $((x_g, y_g))$ dans le plan complexe tel que : $z(p) = (x - x_g) + i(y - y_g)$. Le descripteur de Fourier s'obtient par identification des coefficients de la transformée de Fourier appliquée à $z(p)$, tel que : $Z_n = \frac{1}{N} \sum_{p=0}^{N-1} z(p) e^{(-\frac{2\pi i n p}{N})}$ [ZL02]. Ainsi, il est possible de simplifier un contour en agissant sur p . Si p est faible l'objet s'apparente à un cercle et prend, à mesure que p augmente, une forme de plus en plus fidèle au contour originel.

Principalement appliquée pour de la reconnaissance d'objets compacts ([YNGR07]), la transformée de Fourier est également utilisée sous l'aspect spectral (voir la transformée de Radon pour une représentation spatiale de l'image [BN07]). Appelés descripteurs de Fourier génériques (GFD [MSJM10]), ils s'avèrent moins sensibles au bruit (supprimé par un filtre passe-bas). Leur détermination nécessite le passage de l'image à traiter d'une base cartésienne $(I(x, y))$ à polaire $(I(r, \theta))$, ce qui garantit l'invariance en rotation. La transformée de Fourier est appliquée sur une image convertie, la normalisation achevant le calcul.

La transformée de Fourier permet une analyse à la fois temporelle et fréquentielle. D'autres techniques comme la transformée en ondelettes permettent en plus une analyse multi-résolution. Les modes de calcul les plus utilisés sont la transformée en ondelettes discrètes (DWT) et la transformée par paquet d'ondes (WPT), respectivement employées dans [KAS04] (utilisation conjointe avec un réseau de neurones pour la surveillance des phases d'éveil) et [CDBR10] (application à l'analyse de signaux issus d'un phonocardiogramme).

Si les ondelettes et les descripteurs de Fourier permettent d'analyser localement et globalement l'image, les performances de ces techniques sont cependant conditionnées par les résultats de la détection de contours. Or, nous avons vu dans le point précédent que cette étape ne donne pas satisfaction. La présence de contours non fermés, ajoutés à la présence répétée de contours uniques pour plusieurs objets distincts, ne motive pas le choix de ce type d'approche.

Approche par modèle

D'autres types d'approches reposent sur l'utilisation d'un modèle prédéfinissant les relations entre les attributs locaux de l'objet à l'image des travaux de [Sca98]. Dans les cas où ce type de modèle n'est pas accessible ou lorsque qu'il y a modification importante de l'apparence des objets (dans notre cas : dégradation de l'état de surface, activité du chirurgien notamment, etc), il est alors possible d'utiliser des modèles dits déformables plus connus sous la dénomination "contours actifs". L'idée consiste à trouver une fonction englobant au plus proche l'objet afin de pallier les difficultés liées à la faiblesse du gradient. Ceci revient à minimiser une fonctionnelle caractérisant l'énergie du contour. Cette énergie est composée en partie d'une énergie interne décrivant les caractéristiques du contour. Le contour est alors assimilé à une membrane mince dotée d'un coefficient de rigidité et d'élasticité (deux premiers termes de la régularisation de Tikhonov [CE95]). La seconde partie concerne l'énergie externe qui est fonction des propriétés locales de l'image. L'énergie contextuelle ou de contrainte peut apparaître dans la formulation de la fonctionnelle. Elle permet localement de relaxer certaines contraintes ([BPS+05]). La contribution de ces trois énergies permet l'attraction des points de contours vers les régions où le gradient est élevé. Ainsi, la forme du contour se modifie et vient épouser celle de l'objet.

L'application de cette technique à une scène opératoire est proposée en figure 2.13 (reprise des travaux de [LKGD08]).

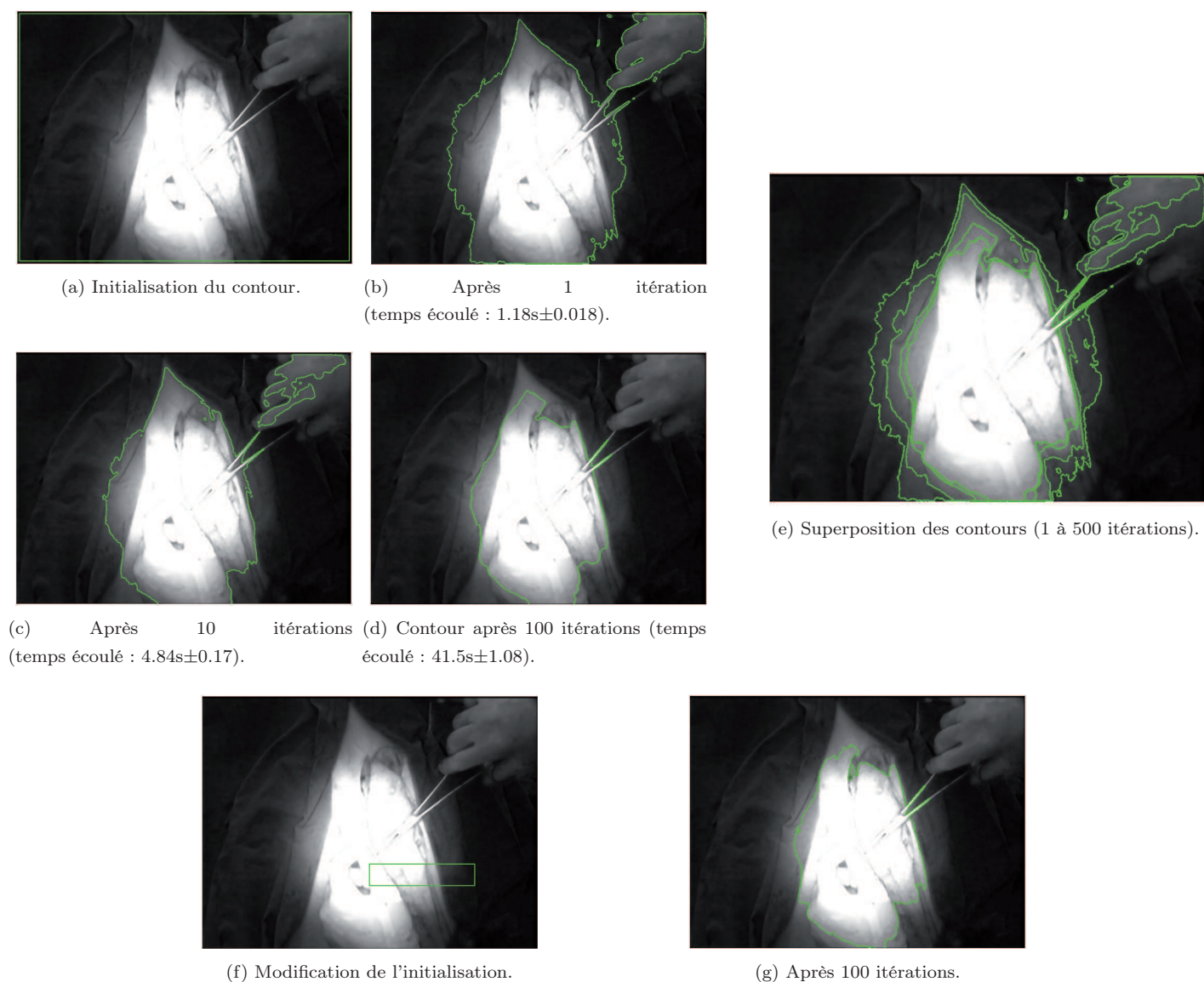


FIG. 2.13 – Exemple de contours actifs.

Nous pouvons constater sur cet exemple, que la forme du contour se modifie, par itérations successives, et tend vers la région où le gradient est le plus élevé. Comme nous pouvons le constater, le contour se concentre sur l'endroit le plus fortement éclairé (correspondant au contraste le plus grand) autrement dit aux abords de la tache lumineuse. Ceci ne nous renseigne pas sur les objets en présence. De plus, le temps imparti pour ce traitement d'image est conséquent ce qui ne permet pas d'être utilisé dans notre cas. Notons qu'une modification du contour initial n'influence que très légèrement le résultat (figures 2.13f et 2.13g). En conséquence, aux vues de ces résultats, les contours actifs sont exclus de cette étude.

En résumé :

Nous avons vu deux types d'approches pour la reconnaissance d'objets basée sur la forme. Celles utilisant l'objet dans son intégralité et celles prenant en entrée un contour. Les premières méthodes supposent que l'objet est isolable du reste de l'image. Or, cette étape n'est pas aisée pour des images de scène opératoire. A ceci s'ajoute le problème de la base de données nécessaire pour la comparaison des objets. Ces méthodes n'ont donc pas été retenues pour notre étude. Pour celles reposant sur une détection de contours, la difficulté réside dans l'obtention d'un contour net et précis. Dans la mesure où les techniques usuelles (laplacien, gradient morphologique, etc) ne donnent pas satisfaction, ce type d'approche n'est pas considéré dans le cadre de notre étude. Les méthodes fréquentielles reposent sur l'identification de coefficients caractéristiques de l'objet des N premières harmoniques à partir du contour. Ces techniques échouent donc également dans notre cas, la mauvaise délimitation du contour ne permettant pas l'établissement d'une signature robuste. Seule la transformée de Hough permet de passer outre ces problèmes et se révèle pertinente pour la classe instrument, les autres classes ne possédant pas de contours suffisamment "droits" pour permettre d'appliquer cette approche. Toutefois, cette technique doit être combinée à d'autres attributs afin d'éviter une mauvaise détection des instruments et permettre la reconnaissance des autres classes.

2.2.2 L'attribut "couleur"

Privilegié pour des tâches simples de reconnaissance, l'attribut couleur est facile d'emploi, nécessite peu de calculs, et est invariant en rotation, en translation et aux changements d'échelles.

Les techniques basées couleur nécessitent de définir au préalable un espace colorimétrique de travail. Les performances de ces techniques dépendent de la capacité de l'espace choisi à coder l'information (figure 2.14) et de sa mise en valeur de certains indices visuels même face aux perturbations extérieures telles que les variations d'éclairage ambiant ou de contraste. Ce cas se présente notamment dans cette étude car l'éclairage opératoire, qui offre une intensité lumineuse très élevée, induit un fort contraste au niveau de la caméra.

Le recours à des espaces colorimétriques permettant d'intégrer l'information d'intensité lumineuse et/ou d'en pallier les effets négatifs est dès lors de première importance pour notre application. Cependant, ce choix n'est pas figé. Certaines transformations permettent de passer d'un espace à un autre mais induisent potentiellement une perte d'informations qui dénature la qualité de l'image d'origine (en raison de la diminution du gamut).

Nous allons présenter les principaux espaces colorimétriques et déterminer le plus adéquat pour cette étude.

2.2.2.1 Les espaces colorimétriques

RGB

Acronyme pour rouge (Red), vert (Green) et bleu (Blue), cet espace trichromatique a été initialement utilisé en 1931 pour les moniteurs à tube cathodique. Il est aujourd'hui décrié en raison de la corrélation des différents canaux et de l'absence de la luminance comme grandeur indépendante, comme le montre [KMB07] au travers d'une revue des applications dans les domaines agroalimentaire, de la reconnaissance de visages et gestuelle. Dans une version plus actuelle, appelée espace RGB normalisé (normalized RGB - nRGB), il permet de désencapsuler l'information de luminance contenue dans RGB et d'en réduire la dimensionnalité. Grâce à (2.5) et dans la mesure où la somme des composantes r , g et b de nRGB vaut 1, il est possible de coder une couleur uniquement grâce à deux composantes. Bien que sensible au bruit, cet espace a gagné en popularité en raison de la rapidité de passage $RGB \rightarrow nRGB$. Une alternative à cet espace, permettant d'être moins sensible aux conditions d'éclairage, consiste à privilégier les rapports des composantes R/G et G/B comme dans [SAP10].

$$r = \frac{R}{R+G+B} \quad g = \frac{G}{R+G+B} \quad b = \frac{B}{R+G+B} \quad (2.5)$$

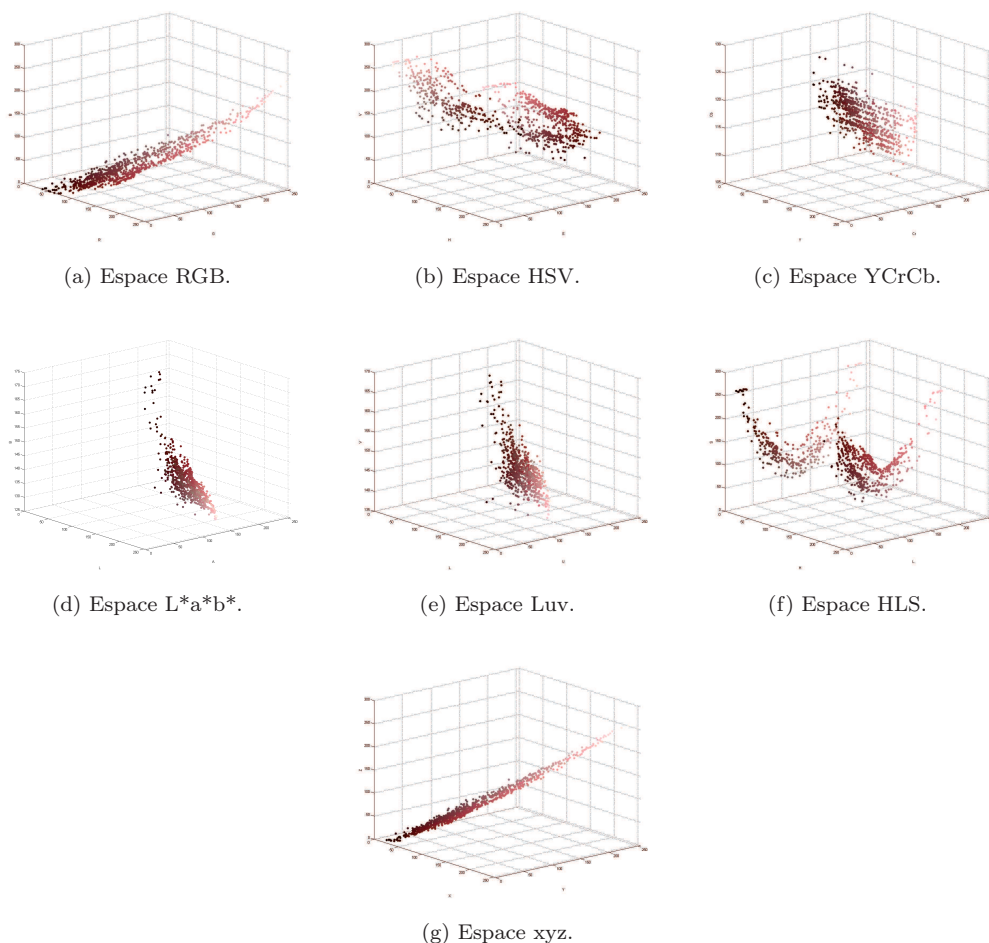


FIG. 2.14 – Représentation du sang dans différents espaces colorimétriques (taille du fragment : 38×24).

HS

Les espaces basés HS (Hue-Saturation / Teinte-Saturation) consistent en une transformation non linéaire mais cependant réversible de l'espace RGB.

Ils quantifient les notions de teinte et de saturation non présentes explicitement dans RGB et nRGB. L'expression de la teinte - également appelée couleur dominante ou tonalité telle que le rouge, jaune, vert, etc, sera codée dans le canal H. La saturation quant à elle contiendra la "pureté" du pixel considéré. Un objet rouge vif aura une haute valeur de pureté (et une teinte rouge), alors qu'à contrario, un rouge estompé (donc plus grisé) aura une faible valeur de pureté.

$$H = \arccos \frac{\frac{1}{2}((R - G) + (R - B))}{\sqrt{(R - G)^2 + (R - B)(G - B)}} \quad S = 1 - 3 \frac{\min(R, G, B)}{R + G + B} \quad (2.6)$$

Tout comme les espaces RGB et nRGB, les espace HS peuvent représenter n'importe quelle couleur par un mélange "additif" de trois canaux. Deux déclinaisons sont possibles pour ces espaces : HS-Value et HS-Intensity (on parle moins souvent de l'espace Hue Chrominance Intensity). La différenciation se fait sur le troisième canal qui est relatif à la perception qu'a notre œil de l'intensité lumineuse. Pour autant, il n'est pas possible de corrélérer cette grandeur à la luminance dans la mesure où ce troisième canal ne prend pas en compte dans son expression la sensibilité de l'œil comme c'est le cas pour l'espace HS-Lightness.

Toutefois, sans pour autant être représentatifs de la luminance les espace HSV et HSI permettent d'être robuste face aux variations d'illumination comme le montre [CJH01] qui utilise le premier espace pour extraire les pixels appartenant à la peau pour des conditions d'illumination variées. Toutefois, [EB11] mentionne que la détection à l'aide de cet espace n'est pas satisfaisante en présence d'ombre. Le lecteur pourra se référer, pour d'autres exemples d'applications, à [GSKM10] (fleur), [ICST11] (mélanome), [CGVRGPSP10] (visage).

YCrCb

De la même manière que HSV, la définition de l'espace $YCrCb$ se fait par transformation non-linéaire de RGB. Cette transformation s'avère plus aisée que le passage $RGB \rightarrow HSV$. L'information codée dans le canal "Y" est cette fois représentative de la luminance. Ceci est rendu possible grâce à la somme pondérée des coordonnées RGB dont les coefficients sont représentatifs de la perception des couleurs chez l'homme, à la différence de l'espace HSV où le calcul de l'entité "V" est effectué par une somme équi-pondérée (de coefficient 1/3). Les autres grandeurs de l'espace, C_r et C_b , forment la chrominance et sont obtenues par soustraction de la luminance Y et des composantes rouge et bleu (d'où l'annotation C_r et C_b) par :

$$Y = 0.299R + 0.587G + 0.114B \quad C_r = R - Y \quad C_b = B - Y \quad (2.7)$$

En raison de la décorrélation luminance-chrominance, $YCrCb$ est l'espace de référence pour l'extraction de la peau dans une séquence d'images ([BPR09]) ou lorsqu'il s'agit d'identifier des lésions cutanées [BBJP10]. Cet espace a donné naissance à d'autres concepts tels que YES et YIQ eux aussi dédiés à la reconnaissance d'objets de couleur chair ([NL09]). Dans la mesure où l'un des objets à reconnaître dans notre application est la peau, il est logique d'utiliser l'espace $YCrCb$ comme référence dans notre étude.

Remarque : Il n'a pas été fait mention des espaces $L^*a^*b^*$ ([LSR08]), CIE-xyz ([YLZ10]) ou encore Luv ([IW11]), dans la mesure où cette introduction aux espaces colorimétriques ne se veut pas exhaustive.

En résumé :

Nous avons vu que l'espace RGB permet un codage simple de l'information couleur mais s'avère peu robuste face à des variations de l'intensité lumineuse ambiante. L'utilisation d'un codage nRGB n'offre toujours pas de réponse à ce problème. Une solution partielle est apportée par les espace HS où la couche "V", "I" ou encore "L" n'est pas représentative de la luminance mais permet une certaine robustesse face aux changements d'éclairage. L'espace YC_rC_b permet de se rapprocher de la courbe de sensibilité de l'œil et par voie de conséquence de la notion de luminance. Ces deux derniers espaces seront donc privilégiés dans cette étude. Toutefois, compte-tenu que l'espace HSV est majoritaire, et afin de limiter les étapes de conversion pouvant pénaliser la charge processeur, nous n'utiliserons pas l'espace YC_rC_b .

Objet à reconnaître	Espace colorimétrique optimal	Espace colorimétrique utilisé
Peau	YC_rC_b	HSV
Sang	HSV	HSV
Graisse	HSV	HSV
Champs stérile	HSV	HSV
Instrument chirurgicaux	HSV	HSV

TAB. 2.3 – Récapitulatif des espaces colorimétriques utilisés pour l'attribut couleur

2.2.2.2 Méthodes de reconnaissance basées couleur

Maintenant que les espaces colorimétriques ont été introduits, nous allons aborder les méthodes de reconnaissance basées couleur et voir quelles sont leurs spécificités.

Définition explicite de la teinte d'une classe

L'une des premières méthodes utilisées pour la détection d'objets consiste à borner chaque composante de l'espace d'étude de sorte que la valeur d'un pixel appartenant à la classe souhaitée soit comprise dans cet intervalle. Les bornes de ces intervalles notées s_i sont définies expérimentalement en accord avec l'application visée et la classe à distinguer. Ainsi, un pixel observé dans l'espace E de base $\{\{e_1\}, \{e_2\}, \{e_3\}\}$ appartiendra à la classe $\{c\}$ si ses composantes respectives p_{e1} , p_{e2} et p_{e3} sont comprises dans des intervalles définis. Dans le cas contraire il appartiendra à $\{\bar{c}\}$. Mathématiquement, cela se traduit par :

$$\begin{aligned}
 \mathcal{D} &= \{c\} && \text{si } p_{e1} \in [s_1^-, s_1^+], p_{e2} \in [s_2^-, s_2^+] \text{ et } p_{e3} \in [s_3^-, s_3^+] \\
 \mathcal{D} &= \{\bar{c}\} && \text{sinon} \\
 \text{avec : } & && s_i^- = \mu_{e_i} - \alpha_i \sigma_{e_i} \\
 & && s_i^+ = \mu_{e_i} + \alpha_i \sigma_{e_i}
 \end{aligned} \tag{2.8}$$

où : \mathcal{D} correspond à la classe choisie à l'issue de la prise de décision ;
 μ et σ sont la moyenne et l'écart-type d'un modèle gaussien ;
 α est un paramètre de pondération.

Cette utilisation rigide de bornes ne convient pas dans les cas où les conditions d'éclairage sont variables. Il est alors possible d'utiliser un seuil κ pour délimiter la classe d'appartenance comme le montre [SAP10] avec $p_{e1}/p_{e2} > \kappa$. Il est également possible de réactualiser les intervalles de sorte à s'adapter aux nouvelles contraintes, comme dans [KPS03] qui distingue le cas d'une exploitation avec une lumière naturelle d'une exploitation avec lumière artificielle dans l'espace RGB :

$$\left[\begin{array}{l} \mathcal{D} = \{c\} \\ \mathcal{D} = \{\bar{c}\} \\ \text{Condition : } \end{array} \right. \text{ si } \left\{ \begin{array}{l} R > 95 \text{ et } G > 40 \text{ et } B > 20 \\ \max\{R, G, B\} - \min\{R, G, B\} > 15 \\ |R - G| > 15 \text{ et } R > G \text{ et } R > B \end{array} \right. \left[\begin{array}{l} \mathcal{D} = \{c\} \\ \mathcal{D} = \{\bar{c}\} \\ \text{Condition : } \end{array} \right. \text{ si } \left\{ \begin{array}{l} R > 220 \text{ et } G > 210 \text{ et } B > 170 \\ |R - G| \leq 15 \text{ et } G > B \text{ et } R > B \end{array} \right. \left. \begin{array}{l} \text{lumière naturelle} \\ \text{lumière artificielle} \end{array} \right.$$

L'idée est donc de tirer partie des avantages liés à l'espace colorimétrique d'étude, comme par exemple HSV qui est un espace peu influencé par les variations d'éclairage. La mise en application rapide de ce type de méthode est donc un atout mais reste dépendante du choix des espaces colorimétriques associés. Par contre, cette approche est inefficace dans les cas où les objets à distinguer possèdent des couleurs très proches. Dans la mesure où, pour cette étude, il est important d'identifier le sang "frais" du sang coagulé, cette technique ne peut pas être utilisée car les couleurs observées sont trop proches les unes des autres.

Approche par modèle gaussien

Le principe est de modéliser la distribution de l'histogramme représentatif de la classe par une loi normale (approche paramétrique multidimensionnelle). Des exemples sont donnés dans [WSLV08]. Dans le cas euclidien (voir [LY02] pour une utilisation de la métrique de Mahalanobis), les probabilités conditionnelles $p(\vec{a}|c)$ et $p(\vec{a}|\bar{c})$ correspondant à l'appartenance d'un pixel aux classes $\{c\}$ et $\{\bar{c}\}$, rapportées à un modèle gaussien et à la formule de Bayes, sont données par :

$$p(\vec{a}|k) = (2\pi)^{-\frac{d}{2}} (\det |\Sigma_k|)^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(\vec{a} - \vec{\mu}_k)^T \Sigma_k^{-1} (\vec{a} - \vec{\mu}_k)\right] \quad (2.9)$$

$$\vec{\mu}_k = \frac{1}{N_k} \sum_{i=1}^{N_k} \vec{a}_i \quad \Sigma = \frac{1}{N_k} \sum_{i=1}^{N_k} (\vec{a}_i - \vec{\mu}_k)(\vec{a}_i - \vec{\mu}_k)^T \quad (2.10)$$

avec : d la dimension de l'échantillon étudié (typiquement $d = 2$) ;
 \vec{a} le vecteur défini pour l'attribut a par $\vec{a} = [\vec{a}_1, \vec{a}_2, \dots, \vec{a}_N]^T$;
 k la classe d'appartenance sélectionnée parmi le cadre de discernement $\Theta = \{\{c\}, \{\bar{c}\}\}$;
 N_k le nombre total d'échantillons de la classe k ;
 $\vec{\mu}_k$ l'espérance de la classe k ;
 $|\Sigma_k|$ le déterminant de la matrice de covariance (définie positive).

Une fois $p(\vec{a}|c)$ et $p(\vec{a}|\bar{c})$ obtenus, il est possible de se prononcer sur l'appartenance d'un échantillon

à une classe grâce au résultat du rapport r de ces deux densités de probabilités conditionnelles :

$$r = p(\bar{a}|c) \cdot p(\bar{a}|\bar{c})^{-1} \quad (2.11)$$

La prise de décision \mathcal{D} concernant l'échantillon s'effectue par rapport à ce ratio telle que :

$$\mathcal{D} = \begin{cases} \{c\} & \text{si } r \in [0, \kappa[\\ \{\bar{c}\} & \text{si } r \in]\kappa, 1] \\ \{\emptyset\} & \text{si } r = \kappa \end{cases} \quad (2.12)$$

Remarque : L'affectation $\mathcal{D} = \{\emptyset\}$ dans l'expression (2.12) correspond à un cas d'indécision. Le système au vue de sa connaissance ne peut se décider sur l'appartenance d'un échantillon à une classe du cadre de discernement (Θ) .

Cette méthode repose donc sur la définition d'un seuil κ ($\kappa \in [0, 1]$) permettant la prise de décision \mathcal{D} . Le plus souvent, ce seuil est défini empiriquement à partir de la courbe de "bonne" détection en fonction des fausses détections (ROC - Receiver Operating Characteristic). En utilisant ce type de courbe, les auteurs ont pu définir que la valeur 0.4 est, pour le seuil κ , la plus adéquate pour les applications de détection de peau.

Modéliser l'histogramme couleur de l'objet à l'aide d'une seule gaussienne ne permet pas d'avoir une reconnaissance complète de l'objet notamment lorsque la distribution est multimodale. Utiliser un modèle plus complexe permet de restituer la distribution de façon fidèle et donc de pallier ce problème. Un moyen simple consiste à mélanger plusieurs gaussiennes.

Mélange de gaussiennes

Les approches reposant sur une distribution modélisée par une seule gaussienne souffrent de plusieurs lacunes. La première est qu'elles ne permettent pas de discriminer deux objets dont les distributions sont proches. La seconde est qu'en élargissant le champ d'action de la gaussienne (modification de la matrice de covariance Σ_k) et donc en améliorant la détection, il y aura une plus grande prise en compte des parasites. L'idée est donc de modéliser plus finement une distribution (un histogramme par exemple) grâce à un mélange de gaussiennes (Gaussian Mixture Models ou GMM). [COB03] dans une étude portant sur la détection de peau basée sur une modélisation mettant en jeu la couleur comme source d'information, montre que ce type de technique possède un taux de reconnaissance supérieur à celui d'un modèle gaussien unique grâce à la diminution de l'erreur de modélisation. Cependant, les auteurs précisent qu'une généralisation à n_G gaussiennes, avec n_G grand (>8), n'est que peu efficace. Ainsi, la difficulté réside dans le choix d'un nombre judicieux de gaussiennes pour la modélisation. [PFJ06] conforte cette analyse et précise que les meilleures performances sont obtenues pour 5 gaussiennes.

Remarque : La complexité des images médicales se traduit généralement par un nombre important d'éléments en présence. Dès lors, pour permettre la reconnaissance des objets présents, il est nécessaire de réitérer cette méthode pour chaque zone homogène dans l'image comme le montre [GRG06] avec un algorithme appelé CGMM (Constrained Gaussian Mixture Model) pour une application de classification de tissus dans une image scanner fortement bruitée.

Cette modélisation multi-gaussiennes est une addition de plusieurs gaussiennes dont on module l'influence. La densité de probabilités conditionnelles qu'un échantillon \vec{a} appartienne à la classe k ($k = \{c\}, \{\bar{c}\}$) devient en considérant que la variable a est indépendamment et identiquement distribuée (i.i.d.) :

$$p(\vec{a}|k) = \sum_{i=1}^{n_G} \pi_i \cdot p_i(\vec{a}|k) \quad (2.13)$$

L'équation (2.13) fait intervenir les scalaires notés π_i appelés paramètres de mélange (ou proportions) permettant la pondération des densités de probabilités conditionnelles individuelles sous contrainte : $\pi_i > 0$ et $\sum_{i=1}^{n_G} \pi_i = 1$. Ainsi, chaque gaussienne possédant son propre jeu $\mu_{k,i}, \Sigma_{k,i}$ tend si le nombre de gaussiennes est suffisamment grand vers la distribution réelle notée \mathbf{p} tel que :

$$\lim_{n_G \rightarrow \infty} \sum_{i=1}^{n_G} \pi_i \cdot p_i(\vec{a}|k) \rightarrow \mathbf{p}(\vec{a}|k) \quad (2.14)$$

Afin d'identifier ce jeu de paramètres, il est courant d'utiliser l'algorithme d'optimisation "Espérance Maximisation" (EM) ([EZ10]). L'estimation de $\mu_{k,i}, \Sigma_{k,i}$ et π_i implique un temps de calcul conséquent, entraînant l'utilisation d'un faible nombre de gaussiennes. De plus, il est possible que EM converge vers un maximum local sans que celui-ci ne donne un jeu de paramètres satisfaisant. A ceci s'ajoute pour l'utilisateur la nécessité de définir explicitement le nombre de gaussiennes n_G entrant dans la modélisation, ce qui impose de tester l'algorithme EM pour différentes valeurs de n_G et d'en déduire celle qui est la plus adaptée. Pour éviter ces essais-tests, il est possible de recourir à des versions plus adaptées de EM telles que AIC (Akaike's Information Criterion), qui propose une pénalisation lorsque n_G devient trop grand [Aka73].

Compte-tenu du temps de calcul requis, cette méthode n'a pas été retenue. Par contre elle peut être utilisée pour modéliser hors ligne, lors de la phase d'apprentissage, la distribution de l'histogramme couleur d'un objet, et être ensuite utilisée en ligne comme connaissance a priori.

Réseaux de neurones

Un réseau de neurones permet la construction d'une modélisation statistique de données (approche non-paramétrique) reposant sur une architecture composée de strates d'éléments [MPR⁺11] (comparaison aux algorithmes génétiques). Ces neurones sont composés de synapses établissant des connections plus ou moins fortes avec les neurones mitoyens. Une synapse peut donc être vue comme une pondération affectant les impulsions d'entrée d'un neurone. Le cœur du neurone est quant à lui chargé de sommer ces impulsions et d'utiliser une fonction d'activation générant une impulsion de sortie. La fonction d'activation peut être sous forme de seuil (neurone binaire), d'une identité (neurone linéaire), d'une saturation, d'une gaussienne ou plus fréquemment d'une sigmoïde. La structure la plus simple est le perceptron utilisé en reconnaissance de forme ([GR05]). Cette structure aujourd'hui moins prisée a conduit à l'élaboration d'autres réseaux comme le perceptron multi-couches (MLPNN - Multi Layer Perceptron Neural Network) dont un exemple d'application médicale est donné dans [RSMR10].

L'un des principaux griefs de ces approches est la nécessité d'ajuster manuellement l'ensemble de la structure (apprentissage supervisé). Réduisant pour la phase de classification dans notre cas, ce point peut néanmoins être contourné par une variante proposée par [VdM73] et reprise par [Koh82], appelée

système auto-organisateur (SOM - Self-Organizing Map plus connu sous le nom de carte de Kohonen), permettant un apprentissage non-supervisé. Bien que les résultats soient encourageants, le temps de calcul nécessaire nous oblige à ne pas considérer cette technique, la puissance de traitement d'un réseau étant fonction du nombre de neurones mis en jeu. Par conséquent les techniques plus évoluées telles que les GNG (Growing Neural Gas) qui consistent à ajouter de la matière neuronale là où il y a un manque de représentation, ou encore les LVQ (Learning Vector Quantization) qui créent des classes à part entière pour répartir les données du vecteur d'entrée [KHK⁺95], ne sont pas étudiées.

Les séparateurs à vaste marge

Résultant des travaux de [Vap95] sur la théorie statistique de l'apprentissage, les séparateurs à vaste marge (SVM : Support Vector Machine) permettent de classer un ensemble de données d'entrée par la génération d'un hyperplan séparant les échantillons [Bur98]. Cette définition permet la construction d'une infinité d'hyperplans possibles ce qui impose d'utiliser un critère en vue d'une séparation optimale. La forme de l'hyperplan reste un problème ouvert. Ce type d'approche a été utilisé pour l'identification rétinienne alliée aux filtres de Gabor ([PP07]), ainsi qu'en reconnaissance d'activité humaine [QMXW10].

Dans le cas le plus simple, l'hyperplan est un séparateur linéaire, scindant en deux un ensemble de données formant ainsi deux catégories : d'une part la classe $\{c\}$ et de l'autre la classe $\{\bar{c}\}$. L'optimalité est alors obtenue lorsque la distance (d) séparant les échantillons est maximale (voir figure 2.15b). Bien qu'il soit plus simple d'utiliser un séparateur linéaire, il est toutefois possible de recourir à des formes non linéaires. En la matière nous retiendrons les séparateurs polynômiaux (ordre défini par l'expert), à base radiale ou sigmoïdaux (figures 2.15c et 2.15d).

Cette méthode nécessite une phase d'apprentissage. Il est nécessaire de remanier la base de données associée, chaque fois que l'on souhaite ajouter ou supprimer un objet de la reconnaissance. De plus, dans notre cas, il y a altération de l'apparence de l'objet à mesure que celui-ci se rapproche de la tache lumineuse. Ceci implique de devoir inclure dans cette base de données des images de l'objet sous différents niveaux d'éclairément ce qui accroît sa taille. En conséquence, il n'a donc pas été choisi de poursuivre dans cette voie compte-tenu du manque de flexibilité de cette méthode.

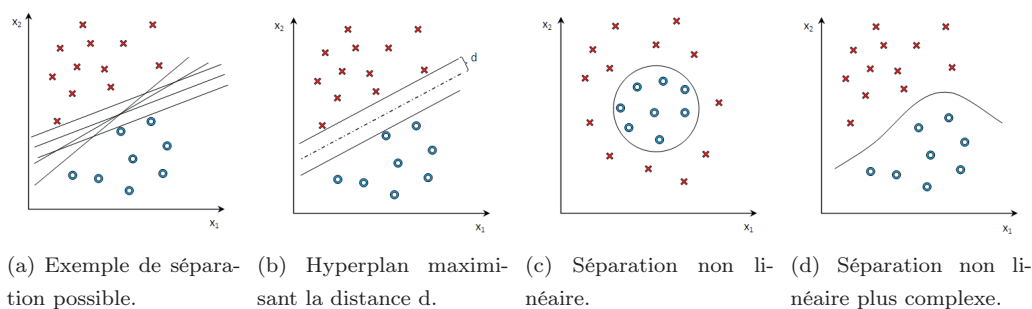


FIG. 2.15 – Exemple de classification par les séparateurs à vaste marge.

En résumé :

Parmi les principales méthodes de classification d'objet basées sur la couleur présentées, certaines ont été écartées pour des raisons de stockage de l'information (base de données de taille trop importante), comme les réseaux neuronaux et les séparateurs à vastes marges, et d'autres à cause du manque d'adaptabilité qui ne leur permettrait pas d'être utilisées pour un cadre chirurgical (méthode basée sur la définition explicite de la teinte d'une classe). L'approche par mélange de gaussiennes (approche paramétrique) offre de bonnes perspectives d'étude compte-tenu du peu de connaissances a priori nécessaires qui est synthétisé dans des histogrammes représentatifs de l'objet. Si cette flexibilité est prometteuse, elle ne résout cependant pas les problèmes liés à l'utilisation passive ou active des objets. Il est donc nécessaire d'incorporer l'information mouvement dans le processus de classification, ce qui fait l'objet de la section suivante.

2.2.3 L'attribut "mouvement"

L'attribut mouvement est utilisé pour des applications de classification d'objets non statiques, mais soulève le problème des mouvements apparents. En effet, la quantification du mouvement de l'objet ne peut être absolue car influencée par le point de vue de la caméra et le mouvement entre la caméra et l'objet observé. Ainsi un objet se déplaçant à vitesse donnée, observé par une caméra statique, génèrera le même mouvement apparent qu'un objet statique vu par une caméra se déplaçant à cette même vitesse (ce problème est partiellement résolu dans notre cas dans la mesure où seule la vitesse de la caméra est connue et non celle du patient et de la table le supportant qui peuvent être amenés à évoluer). Un autre exemple typique est le problème d'ouverture décrit dans [BBF94]. Il faut donc traiter avec réserve cette information et connaître les conditions de prise de vue afin que l'analyse ne soit pas biaisée.

Parmi les approches possibles, nous distinguons :

- la première vise à mettre en évidence les régions où le mouvement a eu lieu. Il s'agit des méthodes de variation temporelle d'intensité et d'historique du mouvement ;
- la seconde vise à quantifier le mouvement, c'est-à-dire à affecter un vecteur à chaque pixel qui s'est déplacé : on trouve les techniques de flot optique (par approche différentielle, fréquentielle) et de mise en correspondance (par point d'intérêt et région).

Variation temporelle d'intensité

Cette technique est basée sur la variation de l'intensité I d'un pixel de coordonnées $\mathbf{x} = \{x, y\}$ entre deux instants t et $t + 1$ [VR04]. Ainsi, entre deux images prises successivement, la variation d'intensité $\Delta I(\mathbf{x}, \Delta t)$ s'écrit :

$$\Delta I(\mathbf{x}, \Delta t) = \mathcal{H}(|I(\mathbf{x}, t + 1) - I(\mathbf{x}, t)| - \Gamma) \quad (2.15)$$

Le paramètre $\Gamma \in \mathbb{N}$ est un entier constant permettant de pallier certaines fluctuations comme par exemple le bruit de la caméra. \mathcal{H} est la fonction d'Heavyside telle que :

$$\mathcal{H}(u) = \begin{cases} 0 & \text{si } u \leq 0 \\ 1 & \text{sinon} \end{cases}$$

Ainsi, les pixels dont la valeur est à 1 seront représentatifs des régions où la variation d'intensité est suffisante, autrement dit là où il y a eu un mouvement. A l'inverse, les pixels mis à 0 seront les zones considérées comme étant immobiles.

Dans le cas d'un fond statique et connu (contenu dans l'image I_{ref}), la détection de mouvement par variation temporelle d'intensité initialement définie par (2.15) devient :

$$\Delta I(\mathbf{x}, \Delta t) = \mathcal{H}(|I(\mathbf{x}, t + 1) - I_{ref}(\mathbf{x}, t)| - \Gamma) \quad (2.16)$$

La technique dite d'historique du mouvement (MHI : Motion History Image) repose sur l'affectation d'une intensité du pixel en fonction de l'antériorité du mouvement ([ZBA09]). Les pixels ayant été traités avant $t - \Delta t$, où Δt correspond à la taille de la fenêtre d'observation, sont considérés comme obsolètes :

$$I_{MHI}(\mathbf{x}, t) = \begin{cases} t & \text{si } I(\mathbf{x}, t) - I_{ref}(\mathbf{x}, t) \neq 0 \\ I_{MHI}(\mathbf{x}, t) & \text{si } I(\mathbf{x}, t) - I_{ref}(\mathbf{x}, t) = 0 \text{ et si } I_{MHI}(\mathbf{x}, t) \geq t - \Delta t \\ 0 & \text{si } t > \Delta t \end{cases} \quad (2.17)$$

Le choix de la fenêtre Δt est laissé à l'utilisateur : une valeur trop haute impliquera une surabondance de pixels correspondant à des mouvements dans l'image alors qu'une valeur trop faible ne permettra pas son analyse.

Ces types de détecteurs de mouvement, simples et rapides à mettre en place, sont généralement utilisés dans les applications de vidéosurveillance où l'environnement est connu et statique (ce qui n'est pas notre cas). En revanche, ces techniques sont sujettes aux fausses détections lorsqu'il y a modification des conditions de prise de vue : zoom de la caméra, variation de l'éclairage ambiant, soustraction d'un des éléments appartenant au fond, etc. De plus, ces techniques ne permettent pas de différencier les mouvements physiologiques des gestes du chirurgien. Facteur aggravant, les éléments en mouvement peuvent avoir des connexions avec le bord de l'image ce qui rend plus difficile l'interprétation des résultats. Il apparaît ardu d'utiliser ces techniques pour de la reconnaissance d'objets dans le cadre d'intervention chirurgicale.

Une possibilité pour pallier ces problèmes de fausses détections consiste à construire de manière adaptative une image représentative du fond, on parle alors de "background subtraction" ([MKK11]). L'estimation par moyenne glissante sur une fenêtre de n images est l'un de ces mécanismes possibles, l'image de référence devient donc :

$$I_{ref}(\mathbf{x}, t) = \frac{1}{n} \sum_{i=1}^{i=t} I(\mathbf{x}, t - i) \quad (2.18)$$

Grâce à (2.18), les perturbations ont moins d'impact mais peuvent tout de même engendrer des mouvements erronés si le nombre d'images servant à la construction n'est pas assez grand. Conjointement, le temps mort nécessaire à la construction de l'image de fond sera d'autant plus important que n est grand, ce qui retardera d'autant l'algorithme de détection de mouvement. Comme le rappelle [BMM07], il existe d'autres techniques pour l'extraction de l'arrière-plan autre que la construction par moyenne glissante : les mélanges de gaussiennes [ZvdH06], les méthodes d'estimation de densité par noyaux [EHD00], le filtre de Kalman [ZS03], par géométrie [?], etc.

Les méthodes basées sur la soustraction, offrent une plus grande robustesse face au bruit et une meilleure performance pour l'isolation de l'objet en mouvement. Cependant, dans notre cas, il apparaît difficile de construire une image représentative du fond, et ce même de manière itérative compte-tenu de la vitesse d'exécution des gestes du chirurgien. Ces techniques n'ont donc pas été retenues pour cette étude.

Flot optique - approche différentielle

Cette technique est basée sur l'idée que l'intensité lumineuse d'un objet reste constante entre deux images successives, ce qui se traduit par :

$$I(x, y, t) = I(x + dx, y + dy, t + dt) \quad (2.19)$$

Si l'objet s'est déplacé d'une quantité (dx, dy) entre deux prises de vue séparées de $dt = 1/f_c$ (f_c étant la fréquence d'acquisition de la caméra supposée suffisamment grande) celui-ci gardera ses propriétés photométriques. Autrement dit, les variations d'intensité observées ne seront dues qu'aux déplacements de l'objet. Cette hypothèse s'avère très contraignante et tend à ne pas être vérifiée lors de longues séquences. En effet, l'éclairage doit être constant quel que soit le point de vue, ceci suppose que l'objet observé est de nature Lambertienne.

Remarque : *On rappelle qu'un objet est dit Lambertien (parfaitement mat) si tous les faisceaux lumineux provenant d'une source lumineuse sont réémis dans toutes les directions. A contrario, un objet possédant une forte composante spéculaire tendra vers le comportement d'un miroir ; les faisceaux lumineux sont redirigés suivant une direction privilégiée.*

Ceci implique que pour un objet du quotidien (comportement hybride à mi-chemin entre le Lambertien et le spéculaire), une modification de la position de l'éclairage entraînera une modification de la luminance perçue et sera donc considérée comme un mouvement à part entière. Une expression plus générale de (2.19) formulée par [PBB⁺06] permet de pallier ce type de phénomène et de prendre en considération les occlusions. La version, appelée contrainte du flot optique, a pour expression :

$$(\nabla I)^T v + I_t = 0 \quad (2.20)$$

avec : ∇I , le gradient spatial de I défini par : $\nabla I = \left(\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right)$;

v , le vecteur vitesse du point considéré tel que $v = v(\mathbf{x}, t) = \left(\frac{dx}{dt}, \frac{dy}{dt} \right) = (v_x, v_y)$;

I_t , le gradient temporel correspondant à $\frac{\partial I}{\partial t}$.

Le passage de (2.19) à (2.20) se fait sur la base de l'absence de variation d'intensité entre deux instants d'où : $\frac{\partial I}{\partial x} dx + \frac{\partial I}{\partial y} dy + \frac{\partial I}{\partial t} dt = 0$ ce qui revient à : $\frac{\partial I}{\partial x} v_x + \frac{\partial I}{\partial y} v_y + \frac{\partial I}{\partial t} = 0$, équivaut à (2.20). Dans cette expression, il n'est possible de déterminer la composante du mouvement que dans la direction du gradient (autrement dit normale au contour de l'objet), les deux composantes v_x et v_y ne pouvant être quantifiées. En ajoutant d'autres contraintes, [HS81] (Horn et Schunck) et [LK81] (Lucas et Kanade), ont permis l'estimation du flot optique en considérant respectivement l'ensemble de l'image et un fragment.

Ainsi la méthode développée par Horn et Schunck repose sur la minimisation de la fonctionnelle dérivée de (2.20) :

$$J_f = \int \int (\nabla I \cdot v + I_t)^2 dx dy \quad (2.21)$$

Ils ajoutent à (2.21) un terme de régularisation permettant d'avoir un champ de vecteurs plus ou moins régulier par l'intermédiaire d'un paramètre k . On parle alors de contrainte de régularisation.

$$J_{HS} = \int \int [(\nabla I \cdot v + I_t)^2 + k^2 \|\nabla v_x\| + k^2 \|\nabla v_y\|] dx dy \quad (2.22)$$

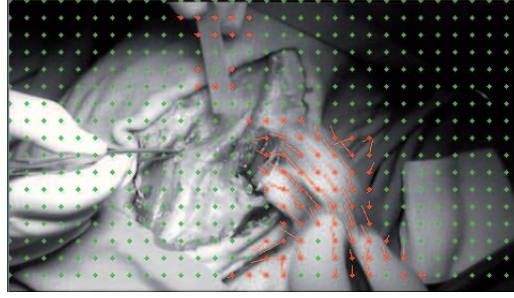


FIG. 2.16 – Exemple de flot optique à maillage régulier.

Bien que cette technique fournisse, après résolution, un champ dense (association pour chaque pixel de l'image d'un vecteur mouvement), elle s'avère sujette au bruit et ne met en avant que les mouvements de grande amplitude en raison de la contrainte de lissage. Une alternative à l'exploitation de l'image dans sa globalité a été proposée par [LK81] sous la forme d'une minimisation d'une fonction d'appariement :

$$J_{LK} = \sum_{\mathbf{x} \in D_x^y} W^2 \left(\frac{\partial I}{\partial x} dx + \frac{\partial I}{\partial y} dy + \frac{\partial I}{\partial t} \right)^2 \quad (2.23)$$

où le voisinage D_x^y correspond aux pixels se trouvant autour du pixel central \mathbf{x} de coordonnées (x, y) , l'usage courant voulant D_x^y inférieur à la dizaine. Le terme W renforce l'importance du pixel central par rapport aux pixels se trouvant aux abords. Il est vu comme une pondération pour les moindres carrés.

Ce type de technique se base donc sur une étude locale du flot optique. L'hypothèse de constance de l'intensité est moins restrictive, assurant ainsi une meilleure robustesse au bruit et une meilleure description des discontinuités. Cependant, lorsque le mouvement est de grande amplitude, cette méthode tend à fournir des résultats erronés. Dans le but de garder les bénéfices de la méthode de Lucas et Kanade, l'idée est de modifier artificiellement le facteur d'échelle pour que le mouvement, à l'origine de grande amplitude, soit considéré de faible amplitude. Par interpolation, le champ de vecteurs est propagé à toutes les images sous-échantillonnées pour arriver à une définition globale du flot optique. A titre d'exemple, nous présentons en figure 2.16 l'application de cette méthode sur une image tirée d'essais en bloc opératoire.

De nombreux travaux ont poursuivi ces études et ont permis une amélioration des calculs de flots optiques par méthodes différentielles, synthétisés dans [Mém03] où l'approche consiste à paramétrer le mouvement de l'objet suivant un modèle ou plus récemment [NBK08] qui propose une méthode reposant sur un cadre variationnel sur-paramétré.

Flot optique - approche fréquentielle

Initiée par [AB85], l'approche fréquentielle consiste en une analyse du spectre spatio-temporel local d'une séquence d'images. Considérons un objet se déplaçant à vitesse constante $v = (v_x, v_y)$ dans un plan parallèle à celui de la caméra. Notons I_0 l'intensité du pixel à l'instant 0, on trouve donc :

$$I(x, y, t) = I_0(x + v_x t, y + v_y t) \quad (2.24)$$

En notant $\widehat{I}(f_x, f_y, f_t)$, la transformée de Fourier de $I(\mathbf{x}, t)$ avec f_x et f_y les fréquences spatiales et f_t la fréquence temporelle, on obtient grâce à l'emploi d'une distribution de Dirac notée δ :

$$\widehat{I}(f_x, f_y, f_t) = \widehat{I}_0(f_x, f_y) \delta(-f_x v_x - f_y v_y + f_t) \quad (2.25)$$

Dans l'hypothèse où il y a conservation de l'intensité lumineuse entre les deux prises de vue, nous avons alors :

$$f_x v_x + f_y v_y = f_t \quad (2.26)$$

L'équation (2.26) étant équivalente à un plan, l'idée est d'en estimer les paramètres pour retrouver v [Hee88]. L'une des approches les plus utilisées est l'emploi de filtres orientés tels que les filtres de Gabor. Ces filtres sont le résultat du produit d'une fonction trigonométrique et d'une gaussienne (l'orientation est obtenue lors du passage vers le domaine fréquentiel). En analysant les réponses des filtres placés dans l'espace de base $\{f_x, f_y, f_t\}$, il est donc possible de retrouver les paramètres du plan de vitesse et d'estimer le mouvement apparent de l'objet. Lorsque le mouvement diffère d'une translation mais s'apparente à une similitude, il est possible de recourir aux coordonnées log-polaires grâce à la transformée de Fourier-Mellin. Ceci permet de lever l'indétermination du facteur d'échelle et de rotation ([JLS04]).

Outre sa sensibilité au bruit, l'approche différentielle implique une charge processeur élevée en raison de la création d'un champ de vecteurs dense. L'application visée étant très sujette au problème de temps de calcul, ce type de méthode n'a pas été retenu pour la quantification du mouvement.

Mise en correspondance - approche par point d'intérêt

Alors que les techniques de flot optique se concentrent au niveau pixel, les méthodes de mise en correspondance se trouvent à une échelle supérieure, discernable en deux catégories : les approches par point d'intérêt (feature-based) et les approches région (region-based). Ces deux approches n'ont nullement besoin d'avoir en mémoire l'estimation antérieure du mouvement.

Dans la première approche, on suppose que des points particuliers appelés points d'intérêts, solidaires de l'objet, sont extrayables tout au long de la séquence. En mesurant leur déplacement entre deux instants successifs, il est possible, en supposant la fréquence d'acquisition de la caméra connue, d'évaluer le mouvement de l'objet. On parle d'appariement. Ces points particuliers possèdent des propriétés d'invariance permettant d'être robustes à l'échelle ainsi qu'aux rotations et translations pouvant affecter l'objet.

Dès lors, commettre une erreur dans la position de ces points revient donc à mal estimer le déplacement de l'objet. De même, dans la mesure où il y a redéfinition des points à intervalle régulier, il est possible qu'à un nombre n de points à l'instant t il ne reste que p points à $t + 1$ (avec $p < n$), il faut donc veiller à affecter le vecteur déplacement à la bonne paire de descripteurs. Parmi l'ensemble des techniques de localisation de points d'intérêt, nous retiendrons les plus connues à savoir : le détecteur de Harris [GHT11], les méthodes SIFT [MEF11] et SURF [QGC⁺11]. Un comparatif des performances des méthodes de calculs des descripteurs peut être trouvé dans [MS05] et [CELR08].

Les détecteurs de Harris-Stephen ([GHT11]) se basent sur les variations brutales de niveaux de gris suivant les deux axes, ce qui correspond à un coin de l'objet. La démarche proposée considère que le coin est plus significatif qu'une texture homogène, ce qui se traduit par une matrice d'auto-corrélation multipliée

a posteriori par un terme de lissage défini par (2.27) :

$$M_{HS}(\mathbf{x}) = \begin{bmatrix} \sum_{\mathbf{x} \in D_x^y} \left(\frac{\partial I}{\partial x}(\mathbf{x}) \right)^2 & \sum_{\mathbf{x} \in D_x^y} \left(\frac{\partial I}{\partial x}(\mathbf{x}) \frac{\partial I}{\partial y}(\mathbf{x}) \right) \\ \sum_{\mathbf{x} \in D_x^y} \left(\frac{\partial I}{\partial x}(\mathbf{x}) \frac{\partial I}{\partial y}(\mathbf{x}) \right) & \sum_{\mathbf{x} \in D_x^y} \left(\frac{\partial I}{\partial y}(\mathbf{x}) \right)^2 \end{bmatrix} \quad (2.27)$$

Cette matrice d'auto-corrélation synthétise les variations d'intensité I suivant les deux axes au point \mathbf{x} de coordonnées (x, y) pris dans une fenêtre D_x^y . En analysant les deux valeurs propres de la matrice M_{HS} il est possible d'identifier trois cas :

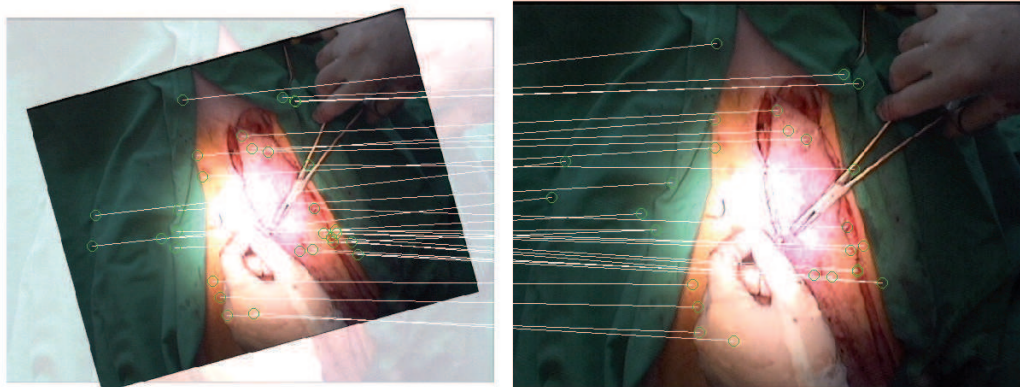
- les deux valeurs propres sont faibles : ceci traduit une texture homogène car il n'y pas de variation d'intensité suivant les deux axes ;
- l'une des valeurs propres est forte : la zone D_x^y se trouve donc sur l'un des contours de l'objet ;
- les deux valeurs propres sont fortes : un point anguleux a été trouvé, ce point sera dit d'intérêt.

La définition des points d'intérêt s'effectue par la recherche des maxima de la fonction $f_{HS} = \det(M_{HS}) - k \text{trace}(M_{HS})^2$ avec k de l'ordre du centième ; le calcul des valeurs propres par la fonction f_{HS} étant plus pratique que par la méthode classique. Tolérante aux changements d'éclairage ambiant, cette technique est cependant mise en échec lors de changements d'échelle (ce qui dans notre cas peut survenir si le chirurgien décide d'abaisser ou éloigner l'éclairage opératoire). Une variante des détecteurs de coins de Harris-Stephen, appelée Harris-Laplacien, offre une réponse partielle à ce problème par adjonction du calcul du Laplacien pour différentes échelles. Bien que le problème du changement d'échelle ait été résolu par [MS01], cette dernière version s'avère encore insuffisante lors de brusques changements d'illumination.

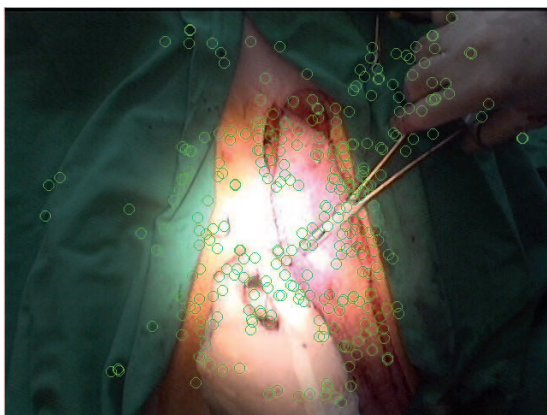
Pour répondre à ce problème [Low99] propose un nouvel algorithme appelé Scale Invariant Feature Transform (SIFT). Le principe repose sur une convolution successive de l'image d'entrée par une gaussienne (approche multi-échelle). Chacune des ces images sera soustraite avec l'image adjacente (communément appelé DoG - Difference of Gaussians), l'image résultante correspondant à une échelle particulière. Les points d'intérêt sont ensuite définis comme étant les extrema locaux (minima et maxima) communs à ces différentes échelles. A chacun de ces points d'intérêt sont associés les gradients des pixels à leur voisinage par construction des histogrammes des orientations suivant des directions privilégiées, pondérées par l'intensité du pixel et l'éloignement au point d'intérêt (inversement proportionnel à la distance au pixel d'intérêt) dont on a calculé l'orientation principale.

Les descripteurs SIFT possèdent donc une dimension importante car dans le cas où l'on choisit un voisinage au point d'intérêt de $n \times n$ et p degrés de liberté pour la construction de l'histogramme, alors ceci constitue un vecteur de $n \times n \times p$ éléments. L'usage conventionnel des descripteurs SIFT étant un voisinage 4×4 et 8 degrés de liberté, nous avons donc un vecteur de 128 composantes pour 1 descripteur (un point SIFT est donc la congruence d'une position, d'une résolution, d'une direction principale et d'un descripteur de 128 composantes). Ces descripteurs améliorent les caractéristiques des détecteurs de Harris-Laplace, car invariants aux translations, rotations, au facteur d'échelle et partiellement au changement d'éclairage ainsi qu'à certaines distorsions (notamment affines). Le lecteur trouvera dans [FL07] une combinaison des SIFT avec les MSER (Maximally Stable Extremal Regions) dont [MS⁺04] présente une étude comparative. A titre d'exemple la figure 2.17c présente une application des MSER à notre cas (issu de routine OpenCV et exécuté en approximativement 300ms). Suite aux travaux sur les descripteurs SIFT, d'autres variantes telles que les PCA-SIFT ont été proposées afin de réduire la di-

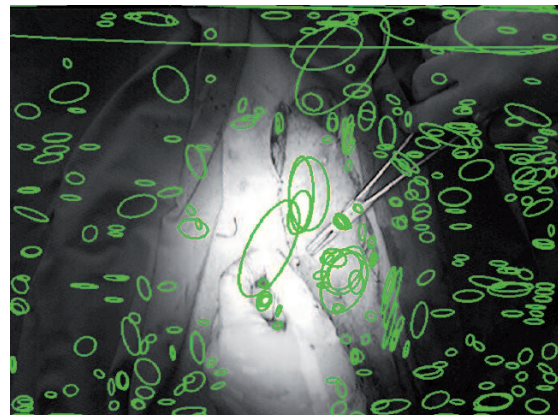
mensionnalité des descripteurs et ainsi réduire le temps de calcul nécessaire [KS04]. Notons que le temps nécessaire pour l'intégralité du processus est de l'ordre de 8.6s pour une image de dimensions 369×287 et 2200 points d'intérêt. Ceci est donc un argument rédhitoire pour notre application malgré les progrès effectués comme l'approche GLOH (basée sur un renforcement du contenu descriptif du point d'intérêt) ou bien SURF (Speeded Up Robust Features - qui s'efforce d'être plus rapide grâce à l'estimation de la matrice hessienne M_{HS} [BTVG06]). La figure 2.17 présente les résultats obtenus à l'aide de cette dernière technique pour notre cas (temps d'exécution : 6.87s).



(a) Appariement (rotation de l'image d'origine : 15° , dézoom : 20%).



(b) Points SIFT (316 points détectés - résolution : 480×640).



(c) Exemple d'extraction de point d'intérêt par méthode MSER.

FIG. 2.17 – Descripteurs SURF.

Remarque : Dans l'approche par "mise en correspondance par région" (*matching region based*), l'image ou un fragment d'image, est considéré dans sa globalité : il n'y a pas de recherche d'élément spécifique comme des coins. Le principe repose sur la mesure d'un critère de dissimilarité permettant de mettre en concurrence des zones de l'image à tester. Ce type d'approche ne peut pas, par définition, permettre l'obtention d'un vecteur déplacement pour de petits objets. Ces approches ont donc été écartées pour notre étude.

En résumé :

Bien que le maillage obtenu par le flot optique basé par point d'intérêt est moins dense par rapport aux deux approches présentées, il constitue néanmoins une bonne alternative pour la quantification du mouvement. Les approches de type SIFT trop gourmandes en temps de calcul sont écartées. Ainsi pour cette étude, les descripteurs de Harris seront retenus en raison de la charge processeur modéré nécessaire. Bien que le défaut principal de cette méthode soit sa mise en échec lors de changement d'échelle, il ne constitue pas un problème majeur. En effet, nous supposons que la distance séparant la coupole et la scène opératoire sera constante tout au long de l'intervention (l'élévation n'étant pas asservie), il n'est donc pas à craindre de voir ce type de problème apparaître.

2.3 Conclusion sur les attributs forme, couleur et mouvement

L'attribut forme s'avère être une information trop délicate à manipuler et sujette à caution dans la mesure où le point de vue influence fortement le résultat de la reconnaissance. L'objet à reconnaître doit posséder une forme spécifique invariante tout au long de la séquence (une connaissance a priori qui ne fait pas partie du cahier des charges). Or, les opérations chirurgicales traitent des tissus anormaux qui peuvent avoir une forme unique. Ceci explique pourquoi cet attribut n'a pas été pris en considération.

Autre attribut étudié, la couleur permet une discrimination simple et rapide des objets. Elle sera donc utilisée dans cette étude comme principal vecteur d'information. Cependant, elle ne permet pas de discerner les objets statiques, non utilisés par le chirurgien (par exemple : clampes, ciseaux, greffon, ...), des objets non statiques ce qui nous oblige à considérer le mouvement comme un attribut supplémentaire pour l'application.

Nous venons de voir les principaux attributs permettant de caractériser un objet. Il a été montré qu'il n'est pas possible de procéder à la reconnaissance des objets en ne se basant que sur un seul attribut. Cependant, certains semblent prévaloir sur les autres, nous retiendrons :

- la transformée de Hough pour mettre en évidence les éléments s'apparentant aux scalpels ;
- la modélisation par mélange de gaussiennes combinée à l'espace HSV pour la détection du sang "frais", des champs opératoires, de la peau dans la scène opératoire ;
- l'approche basée sur les points d'intérêts, permet de quantifier les mouvements apparents. Il est donc possible d'éliminer les mouvements physiologiques et ainsi mettre en avant les régions de l'image "active".

Il s'agit donc maintenant de mettre en commun ces informations pour obtenir une représentation globale des objets en présence permettant à terme de désigner la zone d'intérêt

du chirurgien de manière robuste et précise.

Chapitre 3

Fusion d'attributs et désignation de la zone d'intérêt

Sommaire

3.1 Fusion d'attributs	68
3.1.1 Définitions et généralités	68
3.1.2 Cadres théoriques de la fusion	72
3.1.3 Fusion dans le cadre évidentiel	76
3.2 Définition de la zone d'intérêt	88
3.2.1 Introduction à la théorie de la décision	88
3.2.2 Désignation de la zone d'intérêt	95
3.3 Conclusion	99

L'objet de ce chapitre est de présenter les schémas de fusion permettant de combiner les attributs forme, couleur et mouvement. Le résultat de cette fusion doit permettre la reconnaissance des classes d'objets : sang, peau, champ stérile et instrument. Nous cherchons à l'issue de ce processus de fusion à attribuer à chaque pixel de l'image d'origine une classe d'objet. Nous présenterons ensuite la théorie de la décision qui offre le cadre nécessaire pour désigner la zone d'intérêt à partir de cette reconnaissance. Nous montrerons que cette étape est délicate et conditionne l'ensemble des performances du système en termes de réactivité et de précision.

3.1 Fusion d'attributs

3.1.1 Définitions et généralités

Il n'existe pas à ce jour une définition unique de la fusion de données. Ceci s'explique par la forte hétérogénéité des applications et des buts recherchés [Val01]. Ainsi, des auteurs comme [Wal99] proposent d'harmoniser la terminologie, notamment la définition même de la fusion de données :

"La fusion de données constitue un cadre formel dans lequel s'expriment les moyens et techniques permettant l'alliance des données provenant de sources diverses. Elle vise à l'obtention d'informations de plus grande qualité; la définition exacte de "plus grande qualité" dépendra de l'application."

Cette citation met en évidence la difficulté de couvrir l'exhaustivité des travaux sous une forme unique. Cependant, les méthodes de fusion visent selon [BBC02] à augmenter les performances suivant l'un au moins des critères suivants :

- la granularité : l'information en sortie du processus de fusion est plus riche que celle en entrée ;
- la confiance : elle est matérialisée par une plus grande probabilité a priori de l'information ;
- la précision : l'écart type de l'information se trouve diminué à l'issue du traitement ;
- la complétude : elle permet au système de mieux entrevoir son environnement, d'en avoir une meilleure vision.

Il est donc possible de retrouver la fusion de données dans des domaines aussi variés que :

- le domaine médical [Das12], [MZW11], [ZRL⁺11] ;
- la surveillance [NvLSZ12], [RGH11], [SVF11] ;
- la défense [SRMD11], [VMGBS11], [MN09] ;
- la robotique d'exploratoire [SE11], [FLK11], [VBA11] ;
- ...

Que la fusion soit multi-source, multi-temporelle ou multi-échelle, il est important de comprendre ce que représentent les informations en entrée de ce processus, d'où elles proviennent et quelles en sont les imperfections. Ces dernières sont catégorisées par [Blo05] en cinq familles (une autre typologie de l'incertitude peut être trouvée dans [KY95] ou plus récemment dans [Liu04]) :

1. **L'incertitude.** Elle caractérise le degré de discordance d'une information par rapport à la réalité (notion de fidélité). Ainsi des termes comme "sûr", "certain", "plausible", "probable" ou "vraisemblable" font référence à cette notion de doute qui entache l'avis sur la véracité de l'information (vrai/faux) ;
2. **L'imprécision.** Si l'incertitude porte sur la manière dont la donnée a été définie, l'imprécision quant à elle, caractérise l'information même. L'origine de cette imprécision peut être liée à l'information elle-même ou bien aux modes d'acquisition l'ayant générée par exemple. Soit le cadre de discernement $\Theta = \{\{\text{peau}\}, \{\text{sang}\}, \{\text{instrument}\}\}$ et l'affirmation suivante : "il est incontestable que le pixel appartient au sang ou à la peau". Cette information est donc certaine ("incontestable") mais imprécise ("sang", "peau"). Autre exemple : "il est probable que le pixel appartienne à l'instrument" ; cette fois l'information est incertaine mais précise. Il est immédiat de

trouver une affirmation précise et certaine de même qu'imprécise et incertaine. Dès lors il apparaît qu'une information possède à la fois une incertitude et de l'imprécision, ces deux entités étant couplées ;

3. **L'incomplétude.** Elle caractérise l'absence de vision globale de l'environnement. Ainsi, une information fournie par une caméra est incomplète du fait des limitations physiques liées au capteur : projection perspective, restriction à un domaine spectral spécifique, etc, auxquelles s'ajoutent le bruit propre de la caméra et celui lié à l'environnement ;
4. **L'ambiguïté.** Une information est dite ambiguë si elle mène à deux interprétations différentes. Cette imperfection peut être le résultat d'une imprécision couplée à l'incomplétude de la source. Dès lors il n'est pas toujours possible de se prononcer sur l'appartenance d'un objet à une classe en ne se basant que sur ce type d'informations. La fusion a donc pour but de supprimer l'indécision par la mise en commun d'informations ;
5. **Le conflit.** Fréquent dans les processus de fusion, le conflit apparaît lorsque les informations mènent à plusieurs résultats contradictoires. Plusieurs solutions sont possibles : la première consiste à éliminer les sources d'informations les moins fiables, ce qui suppose d'avoir une connaissance a priori sur celles-ci. La seconde repose sur l'absence de décision, ce qui suppose que la source du conflit provient d'un phénomène transitoire. Enfin, une troisième solution consiste à intégrer des informations supplémentaires permettant de supprimer le conflit.

Afin de rendre l'information plus précise, plus complète ou moins conflictuelle, de nombreuses architectures de fusion sont possibles. L'une des plus connues est le DFPM (Data Fusion Process Model) proposé par le Joint Directors of Laboratories (JDL) en 1986, sous la tutelle du département de la défense. Ce modèle générique permet de codifier et standardiser les aspects de la fusion de données et de définir une terminologie. Le tableau 3.1 propose une synthèse des travaux de [LBR⁺04] visant à réactualiser ce modèle.

Description du DFPM	Interprétation
<p>Niveau 0 (Pre-Processing) : Pré-traitement des données (filtrage, normalisation, recalage spatial et temporel)</p>	<p>Estimation de l'état des signaux Vise notamment à diminuer, par traitement d'images dans le cadre de cette étude, le bruit de l'information ainsi qu'à extraire les informations utiles de la donnée (attributs).</p>
<p>Niveau 1 (Object refinement) : Association de données, estimation et prédiction, identification de l'objet</p>	<p>Estimation des caractéristiques Etape permettant la reconnaissance de l'objet par le biais de l'identification de ses attributs.</p>
<p>Niveau 2 (Situation refinement) : Reconnaissance de modèles, détermination des relations entre les entités</p>	<p>Evaluation de la situation Compréhension de l'environnement par aggrégation des connaissances sur les objets et les liens les unissant.</p>
<p>Niveau 3 (Threat refinement/Impact assessment) : Evaluation du comportement futur (connaissant les événements passés), détermination des menaces, des alliés, des ennemis</p>	<p>Evaluation des impacts Donne l'univers des possibilités pouvant affecter la situation actuelle ainsi que le cours des événements.</p>
<p>Niveau 4 (Process refinement) : Surveillance, évaluation et amélioration des performances des sous-niveaux</p>	<p>Amélioration de la procédure [SBW99] analyse cette étape d'auto-contrôle des performances comme n'étant utile que si les données d'entrée sont significativement exogènes.</p>

Eléments additionnels	
<p>Interface Homme-machine : Visualisation des données, interprétation des données, contrôle global du système</p>	<p>Gestion du système de base de données : Base de données de soutien, connaissance a priori servant à la fusion (base de données)</p>

TAB. 3.1 – Modèle du processus de fusion proposé par le JDL

Cette architecture de fusion proposée par le JDL doit être vue non pas comme une structure rigide (au sens de framework) mais plus comme une base de réflexion. Ainsi, en fonction du problème abordé, ce modèle doit être adapté, ce qui conduit parfois à inverser, paralléliser ou ne pas exécuter certains niveaux du modèle.

D'autres travaux ont porté sur l'élaboration d'autres modèles. Leur comparaison est délicate dans la mesure où certains sont dédiés à des situations précises. Dès lors, critiquer ces schémas de fusion dans un cadre général n'a que peu de sens. Néanmoins, citons ici les modèles :

- Waterfall ([MOK⁺05]), qui consiste en une hiérarchisation des tâches à la manière du JDL sans pour autant inclure un raisonnement cognitif;
- Observation, Orientation Decision, Action (OODA) de Boyd ([Boy87]), qui peut être associé au DFPM selon [SBL01] qui montre que l'étape d'observation correspond aux niveaux 0 et 1, l'orientation aux niveaux 2 et 3. Quant au niveau 4, il est partagé par la décision et l'action. Les auteurs mettent en avant son aspect cyclique ainsi que la diminution de la quantité d'informations lors du passage d'une étape à une autre;
- Omnibus ([BO00]), extension du modèle OODA grâce à l'intégration de schémas plus précis des flux de données le rendant plus rigide pour des applications;
- Situation Awareness Model ([End95]) qui est un modèle de fusion calqué sur l'homme. Couramment appelé SA, ce modèle de conscience de la situation repose sur "la perception des éléments dans un environnement borné en temps et en espace, la compréhension de leur signification, et la projection de leur état dans un futur proche". Ce type de modèle possède donc un cœur correspondant aux trois questions (3Q) suivantes : Qui/Quand?, Que font-ils?, Que vont-ils faire? Même si ces travaux prennent mieux en compte le facteur humain ils ne précisent pas, là non plus, les flux de données internes.

Cette présentation des schémas de fusion se voulant brève, nous ne présenterons pas les modèles de Luo-Kay ([LK]), Pau ([Pau88]), Thomopoulos ([Tho90]) et CECA ([NvLSZ12]). Cependant, le lecteur pourra se référer à [Elm07] pour une liste plus exhaustive d'architectures.

Nous retiendrons pour cette étude l'architecture DFPM en raison de la similitude des étapes 0 à 2 avec le raisonnement suivi dans cette étude : recherche d'attributs/reconnaissance d'objets/prise de décision. Cette architecture est dans notre cas à plus proprement parler, une fusion d'attribut au sens de Dasarathy ([Das97]) qui distingue trois types de schémas de fusion selon que les informations sont de bas niveau (au sens cognitif), d'un niveau intermédiaire ou haut niveau, comme le montre la figure 3.1 ([Mar05]). On parle alors de fusion de données (abusivement utilisée pour désigner un processus de fusion en général), d'attributs ou de décisions. La première correspond à la manipulation d'informations telles que celles issues de capteurs. La seconde, à mi-chemin entre la fusion de données et la fusion de décisions (fusion de haut niveau), la fusion d'attribut s'attache à fusionner les caractéristiques de l'information comme la couleur, le mouvement, etc. Le lecteur pourra se rapporter à [DW87] et [BF04] pour d'autres classifications des schémas de fusion ou à [Mit07] pour une vue plus complète.

En résumé :

Dans cette étude, la problématique réside dans la reconnaissance des objets présents dans la scène chirurgicale. Nous avons montré en section 2.3 qu'il est nécessaire de fusionner différentes caractéristiques telles que la couleur, la forme et le mouvement. Nous avons présenté dans cette partie les imperfections de ce type d'information ainsi que les schémas de fusion permettant de la valoriser au regard de l'application. Dans notre cas, nous utiliserons une architecture DFPM car adaptée à la problématique : recherche d'attributs/reconnaissance d'objets/prise de décision. Cette dernière doit permettre d'émettre un avis sur l'appartenance de pixels à une classe d'objets.

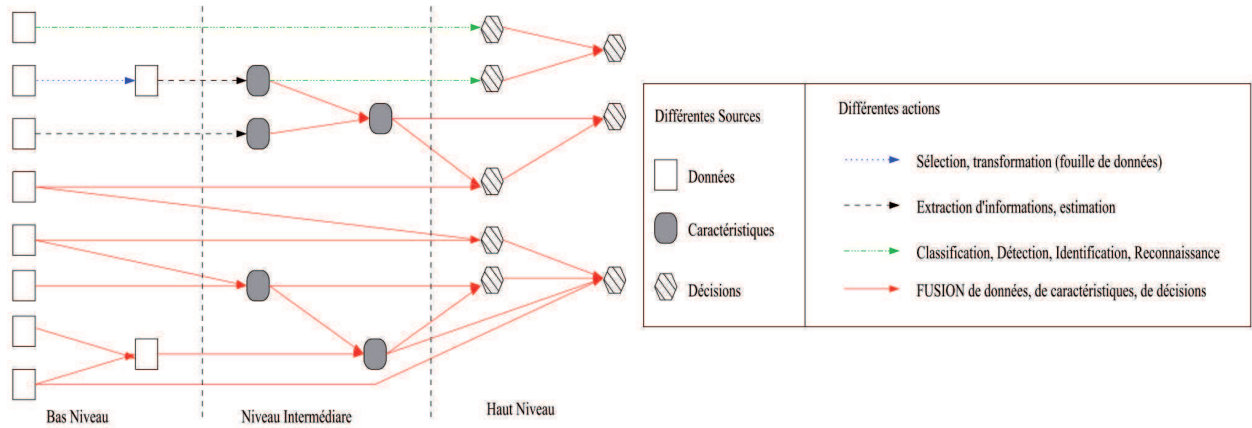


FIG. 3.1 – Différents types de fusion au sens de Dasarthy (source : [Mar05]).

3.1.2 Cadres théoriques de la fusion

Il existe différents cadres mathématiques pour la fusion d'informations : le vote, les probabilités, les possibilités et l'évidentiel. Nous allons voir que ces familles ne traitent pas les mêmes imperfections de l'information. Par la suite nous prenons le parti d'appeler "fusion" la mise en commun d'informations extraites d'une image. Il s'agit en fait d'un abus de langage car ce domaine spécifique de la fusion est à proprement parlé de la "cue integration" ([TM01]).

Remarque : Les termes "théorie de l'évidentiel", "cadre évidentiel" ou "cadre crédibiliste" seront utilisés par la suite pour traduire le mot anglais "evidence" qui signifie "preuve".

3.1.2.1 Fusion par vote

L'idée de cette méthode est que des experts (estimateurs-classifieurs) votent pour ou contre l'appartenance d'une observation à une classe. Soit n le nombre d'avis à évaluer et θ les classes pour lesquelles ces experts peuvent voter. Les avis apportés à chaque classe (supposée exclusive) sont notés $v_{i,k}$, l'indice i faisant référence au numéro de l'expert et k à celui de la classe. Ces votes ne peuvent prendre que deux valeurs possibles (vote binaire) telles que $v_{i,k} : \Theta \rightarrow \llbracket 0; 1 \rrbracket$. La prise de décision quant à elle repose sur le dépassement d'un seuil ([Par94]). La décision \mathcal{D} sera prise par le dépassement (ou non dépassement) d'un seuil prédéfini.

Pour illustrer la fusion par vote, considérons trois experts devant se prononcer sur l'appartenance d'un pixel au cadre de discernement $\Theta = \{\{\theta_1\}, \{\theta_2\}, \{\theta_3\}\}$, $\{\theta_3\}$ étant une classe fictive dédiée aux situations de conflit (non accessible explicitement par les experts mais seulement comme résultat de décision). Après évaluation les experts donnent les avis suivants :

- Expert n°1 : $v_{1,1} = 1$ et $v_{1,2} = 0$
- Expert n°2 : $v_{2,1} = 1$ et $v_{2,2} = 0$
- Expert n°3 : $v_{3,1} = 0$ et $v_{3,2} = 1$

On obtient alors : $v_{.,1} = \sum_{i=1}^n v_{i,1} = 2$ et $v_{.,2} = \sum_{i=1}^n v_{i,2} = 1$. L'utilisateur a le choix parmi plusieurs seuils possibles, dont voici les quatre plus utilisés :

- Vote à l'unanimité : $D = \{\theta_k\}$ si $\max_{k \in \Theta / \{\theta_3\}} \sum_{i=1}^n v_{i,k} = n$, $\{\theta_3\}$ sinon
- Vote majoritaire : $D = \{\theta_k\}$ si $\max_{k \in \Theta / \{\theta_3\}} \sum_{i=1}^n v_{i,k} > \frac{n}{2}$, $\{\theta_3\}$ sinon
- Vote byzantin : $D = \{\theta_k\}$ si $\max_{k \in \Theta / \{\theta_3\}} \sum_{i=1}^n v_{i,k} > \frac{2}{3}n$, $\{\theta_3\}$ sinon
- Vote m sur n : $D = \{\theta_k\}$ si $\max_{k \in \Theta / \{\theta_3\}} \sum_{i=1}^n v_{i,k} \geq m$, $\{\theta_3\}$ sinon

Notons qu'il existe une variante du vote m sur n dite vote t sur V (où $V = \sum_k v_{.,k}(\theta)$), où une classe est choisie si $\sum_k v_{.,k}(\theta) > t$. Ce type de vote mène à une solution non unique dans les cas où : $t \leq \frac{1}{2} \sum_k v_{.,k}(\theta)$.

Dans cet exemple, le vote majoritaire conduit à sélectionner θ_3 ; les votes majoritaire et byzantin à θ_1 et le vote m sur n à θ_1 si $m \in \{0, 1, 2\}$ et θ_3 si $m = 3$. On voit donc ici l'intérêt d'avoir une classe symbolisant les conflits entre les experts.

Pour ce type de processus, les experts qui émettent les avis (votes) pour élire une classe, sont supposés identiques. Dans le cas contraire, il est possible de pondérer les votes en fonction de leur fiabilité et de leur précision, permettant ainsi à l'algorithme d'être plus robuste. Cependant, l'estimation de l'erreur des experts peut se révéler ardue voire inaccessible. Ces votes peuvent faire l'objet d'une possible réactualisation en cours de processus afin de prendre en compte la dynamique du capteur. Il sera montré dans le chapitre 4 comment définir un jeu de pondération basé sur l'analyse des performances des experts (matrice de confusion). Bien que cette pondération résolve certains conflits, elle nécessite une connaissance a priori de la fiabilité des experts. Dans un cas général, la prise de décision s'effectue sur la base de $\max_{k \in \Theta / \{\theta_3\}} \sum_{i=1}^n \delta_{i,k} v_{i,k} > \kappa$ avec $\sum_{i=1}^n \delta_{i,k} = 1$ et κ le critère fonction du type de vote choisi.

Adaptée pour la prise de décision symbolique, la fusion par vote a pour avantage, outre sa facilité de compréhension et sa rapidité de mise en place, d'être très peu coûteuse en temps de calcul. Bien que ce dernier point soit un élément critique de cette étude, ce type de fusion a été écarté dans la mesure où il n'y a pas de prise en compte de la fiabilité des sources émettant l'information. Or, comme nous le verrons plus en détail dans le chapitre 5, les méthodes d'extraction des attributs ont des performances inhomogènes.

3.1.2.2 Fusion dans le cadre probabiliste

Incontestablement l'approche la plus utilisée, la fusion probabiliste repose sur des bases mathématiques solides (depuis les travaux de T. Bayes justifié par Cox [Cox46]). Son usage est préconisé lorsqu'il s'agit de traiter des informations incertaines.

Considérons un cadre de discernement Θ contenant n classes d'objets θ_i (supposées discrètes, finies et dénombrables). On appelle distribution de probabilités $p : \Theta \rightarrow [0, 1]$ telle que $\sum_{i=1}^n p(\theta_i) = 1$ et $\mathbb{P}(\theta_i)$ la probabilité de l'événement θ_i (information a priori). $\mathbb{P}(\theta_i)$ représente le degré de croyance de l'expert face à l'apparition (ou la réalisation) de l'événement θ_i . Du point de vue fréquentiste, cette probabilité est le rapport des cas favorables sur ceux défavorables.

Ce degré de croyance $\mathbb{P}(\theta_i)$ est donc inclus entre d'une part l'événement impossible $\Leftrightarrow \mathbb{P}(\emptyset) = 0$ et

d'autre part l'événement certain correspondant à $\Leftrightarrow \mathbb{P}(\Theta) = 1$, la réalisation de Θ étant par définition sûre car incluant tous les événements possibles. L'axiome de dualité est une seconde justification dans la mesure où $\forall A \subseteq \Theta, \mathbb{P}(A) = 1 - \mathbb{P}(\bar{A})$.

Ces informations a priori renseignent sur le passé des événements et permettent d'en déduire l'occurrence future. Ainsi, la survenue d'un événement tend à réviser cette connaissance. Par application du théorème de Bayes, la probabilité conditionnelle $\mathbb{P}(B|A)$ (probabilité a posteriori) correspondant à la probabilité d'observer l'événement B sachant que A s'est produit, est donnée par :

$$\mathbb{P}(B|A) = (\mathbb{P}(A|B) \times \mathbb{P}(B)) \times (\mathbb{P}(A))^{-1} \quad (3.1)$$

avec $\mathbb{P}(A|B) = \mathbb{P}(A, B) \times \mathbb{P}(B)^{-1}$.

Remarque : Dans les cas où B comprend un ensemble d'événements $\{B_j\}_{1 \leq j \leq N}$, la formule de Bayes devient :

$$\mathbb{P}(B_j|A) = (\mathbb{P}(A|B_j) \times \mathbb{P}(B_j)) \times \left(\sum \mathbb{P}(A|B_j) \times \mathbb{P}(B_j) \right)^{-1} \quad (3.2)$$

Une extension vers des variables continues x et y est possible, on utilise alors des densités de probabilité données par :

$$f(x|y) = (f(y|x) \times f(x)) \times \left(\int f(y|x) \times f(x) dx \right)^{-1} \quad (3.3)$$

avec $f : \mathbb{R}^n \rightarrow [0, \infty[$.

La fusion probabiliste est un cadre mathématiquement cohérent notamment sous le point de vue fréquentiste (notons qu'il existe d'autres types d'approches telles que les probabilités usuelles et les probabilités subjectives). Son emploi est conseillé lorsque l'ensemble des données est accessible, mais s'avère inapproprié lorsque les données sont peu ou mal connues ([Kos90]). Une autre limitation des probabilités concerne la restriction de la décision aux hypothèses (singleton) constituant le cadre de discernement (celle-ci s'effectuant grâce au maximum de probabilité a posteriori (MAP), maximum de vraisemblance, etc), l'impossibilité de traiter les systèmes non additifs (paradoxe d'Ellsberg) ainsi que les événements non indépendants. De plus, l'impossibilité d'intégrer le conflit (l'ignorance quant à elle est traduite par équi-probabilité) nous pousse à envisager d'autres cadres mathématiques pour la fusion d'attributs dans notre étude.

3.1.2.3 Fusion floue

La théorie des sous-ensembles flous fut développée initialement par [Zad65] pour traiter des cas où les états ne sont pas binaires, généralisant ainsi la théorie des ensembles classiques. Ce type de logique a donné naissance à la théorie des possibilités ([Zad99]) pour gérer des problèmes où les informations sont à la fois imprécises et incertaines (l'incertitude n'étant pas abordée de manière probabiliste). Ainsi des qualificatifs tels que "petit", "grand", "éloigné" ou "proche" sont modélisables à l'aide de fonctions d'appartenance floues - FMF (Fuzzy Membership Function) - notées μ (fonctions d'appartenance floues) telles que $\mu : \Theta \rightarrow [0, 1]$. Ainsi plus $\mu_{\theta_i}(x)$ tend vers 1, plus l'objet x (dans notre cas le pixel x) appartient à la classe θ_i . A contrario, à mesure que $\mu_{\theta_i}(x)$ s'en éloigne les chances pour que x relève de θ_i diminuent.

Chaque classe appartenant à un cadre de discours Θ possède donc sa propre fonction d'appartenance.

Il faut donc définir l'allure de ces fonctions, sans laquelle il est impossible d'attribuer une valeur représentative de l'appartenance de l'objet considéré à une classe. Cette étape permettant le passage de linguistique en numérique s'appelle la fuzzification. Elle repose sur un partitionnement de l'univers de discours qui nécessite une connaissance experte pour le choix de la forme des fonctions (triangulaire, gaussienne, etc). Une fois l'ensemble des fonctions d'appartenance définies, on utilise une série de règles représentatives du fonctionnement du système telles que : Si $x \in \theta_1$ Et $y \in \theta_2$ Alors $\mathcal{D} \in \theta_3$. Chaque règle met en commun des prémisses ($x \in \theta_1, y \in \theta_2$), des conjonctions (Et), une implication (Alors) et une conclusion ($\mathcal{D} \in \theta_3$) correspondant à ce que doit faire le système si les prémisses sont accomplies.

Pour illustrer ce fonctionnement considérons un système dont l'objectif est d'asservir visuellement la luminance dans une cavité chirurgicale. La variable de sortie (floue) est l'éclairement de la coupole. Celle-ci est conditionnée par deux variables d'entrée : le diamètre de la cavité et la luminance observée qui peuvent prendre respectivement les valeurs "élevée", "moyenne" et "faible" pour le diamètre de la cavité contre "forte" et "faible" pour la luminance observée.

Les règles floues auxquelles le système doit se conformer sont présentées dans le tableau 3.2.

SI	Diamètre de la cavité	ET SI	Luminance dans la cavité	ALORS	Action
	grande		faible		niveau d'éclairement actuel conservé
	moyenne		faible		éclairage augmenté
	petite		faible		éclairage augmenté
	grande		forte		éclairage diminué
	moyenne		forte		éclairage diminué
	petite		forte		niveau d'éclairement actuel conservé

TAB. 3.2 – Exemple de règles utilisées en logique floue

A partir de ces éléments et des valeurs courantes du diamètre de la cavité et de la luminance observée, il est possible de fournir en sortie du processus (figure 3.2) une valeur représentative de l'action à entreprendre. Ceci n'est rendu possible que grâce à la méthode dite de combinaison (ou d'agrégation) qui attribue à la variable de sortie une valeur unique.

Moins restrictif, le cadre possibiliste permet la modélisation des mécanismes humains mettant en jeu l'imprécis et l'incertain mais nécessite pour ce faire une connaissance experte (choix de l'opérateur de combinaison et création des règles) pour la création d'une base de règles (heuristiques). Compte-tenu du type d'information traité la fusion possibiliste peut convenir pour cette étude.

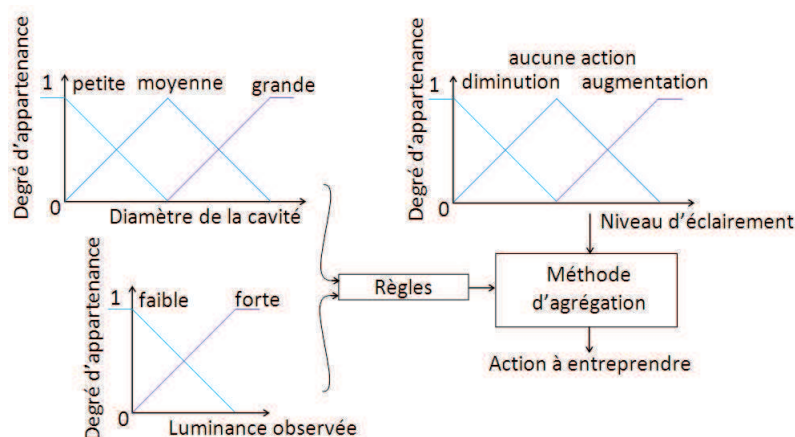


FIG. 3.2 – Fonction d'appartenance triangulaire des variables linguistiques "diamètre de cavité" et "luminance observée".

3.1.3 Fusion dans le cadre évidentiel

La fusion évidentielle (ou crédibiliste) permet à la fois de gérer l'incertitude et l'imprécision mais aussi le conflit. Ceci n'est rendu possible que grâce au modèle des croyances transférables de Smets. Ce modèle permet la représentation et l'agrégation des informations pour permettre en sortie de processus de se prononcer, dans notre cas, sur l'appartenance d'un pixel à une classe d'objets. Les travaux dans ce domaine portent sur des applications aussi variées que :

- le diagnostic : défaut logiciel ([PG11]), défaillance de pales de rotor ([YHH⁺11]), détection de fautes ([LYHH11]), détection de bruit ([TC11]);
- le médical : classification des patients atteints d'osteoarthrite ([JBHR06]), diagnostic médical ([Str06]), estimation de la viabilité du myocarde ([MRJ01]);
- la segmentation d'images : [CCFM04], [HMS04], [BSF⁺11], [AMMM06];
- l'environnement : intrusion de contaminants dans un réseau d'eau [YWR11], détection de feux de forêt [ZMA⁺11], cartographie [MFF11], évaluation d'impacts environnementaux [Awa11].

Compte-tenu de la supériorité de l'évidentiel sur les probabilités et les possibilités, nous choisissons pour cette étude d'utiliser ce cadre pour le processus de reconnaissance des objets en présence dans la scène chirurgicale. La mise en œuvre de ce type de fusion passe par les deux étapes suivantes :

1. l'étape crédale (ou étape de représentation) correspond au processus de création/agrégation des informations. Celle-ci consiste à :
 - (a) définir le cadre de discernement servant à définir le niveau de précision du système ;
 - (b) procéder à la création de fonctions de croyance que l'on appelle masses, représentatives de la "confiance" que l'on a pour une proposition donnée ;
 - (c) sélectionner la règle de combinaison servant à agréger ces masses.
2. l'étape pignistique consiste à choisir le critère de décision permettant de statuer sur la classe à laquelle appartient le pixel.

Remarque : Même si aujourd'hui des ponts existent entre les différents cadres mathématiques comme l'illustre la figure 3.3 (le lecteur pourra se référer à [Nor88] pour un passage du cadre crédibiliste aux chaînes de Markov), il est préférable de choisir celui ayant le meilleur rapport simplicité de mise en œuvre/précision du résultat. En effet, utiliser le cadre le plus général implique une complexité parfois difficile à gérer notamment lorsque des données manquent ou lors d'une perte d'information du fait de la limitation de la représentation des données. La figure 3.3 présente les cadres mathématiques les plus répandus : la théorie des probabilités (XVIIème siècle), des sous-ensembles flous (1965), et évidentielle (1967) ainsi que les ponts permettant de passer des uns aux autres. Sont présentes également la théorie des ensembles aléatoires (1974) et la théorie des ensembles approchés (1982). La première, développée par [Ken74] (random sets) appliquée à l'origine pour des transformations morphologiques, a été rapprochée de la théorie de l'évidence par les travaux de [Ngu08] (la référence fait état d'une revisite de la version datant de 1978) et des ensembles flous par [Orl78]. Le lecteur pourra se référer à [FJGB02] pour plus de détails à ce propos.

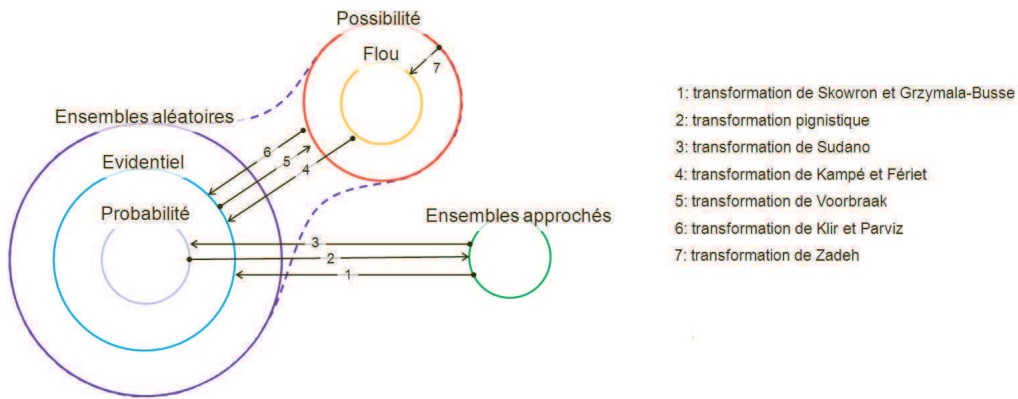


FIG. 3.3 – Représentation ensembliste des cadres mathématiques en fusion de données.

3.1.3.1 Etape 1 : l'étape crédale

Définition du cadre de discernement

Soit l'ensemble Θ le cadre de discernement (ou monde) ayant pour expression (3.4) dans le cas de n classes :

$$\Theta = \{\{\theta_1\}, \{\theta_2\}, \dots, \{\theta_n\}\} \quad (3.4)$$

L'adaptation de l'équation (3.4) à notre étude s'écrit :

$$\Theta = \{\{\text{peau}\}, \{\text{sang}\}, \{\text{graisse}\}, \{\text{champs stériles}\}, \{\text{instruments}\}\} \quad (3.5)$$

A la différence des probabilités, le résultat de la fusion évidentielle n'est pas restreint à cet ensemble mais à une extension de Θ intégrant les unions de classes, appelée power set.

Ce power set noté $2^\Theta \triangleq (\Theta, \cup)$ représente l'ensemble exclusif ($\theta_i \cap \theta_j = \emptyset, \forall i \neq j$) des 2^n disjonctions d'hypothèses formées par :

1. $\emptyset, A_1, A_2, \dots, A_n \subseteq 2^\Theta$, A_i étant une proposition.
2. si A_i et $A_j \subseteq 2^\Theta$ alors $A_i \cup A_j \subseteq 2^\Theta$.

Dans ces conditions, l'équation (3.4) donne le power set 2^Θ suivant :

$$2^\Theta = \{\{\emptyset\}, \{\theta_1\}, \dots, \{\theta_n\}, \{\theta_1 \cup \theta_2\}, \dots, \{\theta_{n-1} \cup \theta_n\}, \dots, \{\theta_i \cup \theta_j \cup \theta_n\}, \dots, \{\theta_1 \cup \dots \cup \theta_n\}\} \quad (3.6)$$

L'ensemble 2^Θ contient donc des classes uniques (singletons) telles que $\{\theta_i\}$ et des combinaisons de classes suivant l'opérateur \cup à l'image de $\{\theta_i \cup \theta_n\}$. Notons la présence de deux autres éléments particuliers notés $\{\emptyset\}$ et $\{\theta_1 \cup \dots \cup \theta_n\}$. Ces deux propositions permettent explicitement de modéliser le conflit et l'ignorance (ou le doute), ce qui n'est pas permis dans les cadres probabilistes et possibilistes. La notion d'ignorance étant dans le cas probabiliste modélisée par une équi-probabilité des propositions et non contenue dans une proposition unique. Le cadre crédibiliste offre donc des mécanismes supplémentaires comparativement aux autres cadres.

Les combinaisons d'hypothèses ne se résument pas uniquement à l'opérateur union. Ainsi d'autres opérateurs comme l'intersection servent à enrichir le power set et par voie de conséquence offrent plus de possibilités en termes de décisions. Comme l'illustrent les équations (3.9), il existe des sur-ensembles du power set dénommés hyper-power sets notés $D^\Theta \triangleq (\Theta, \cup, \cap)$ (tel que $2^\Theta \subseteq D^\Theta$) et le super-power set noté $S^\Theta \triangleq (\Theta, \cup, \cap, c(.))$. Ce dernier reprend l'hyper-power set complété de l'opérateur complément $c(.)$ (tel que $2^\Theta \subseteq D^\Theta \subseteq S^\Theta$).

$$D^\Theta = \{\{\emptyset\}, \{\theta_1\}, \{\theta_2\}, \{\theta_1 \cup \theta_2\}, \{\theta_1 \cap \theta_2\}\} \quad (3.7)$$

$$S^\Theta = \{\{\emptyset\}, \{c(\emptyset)\}, \{\theta_1\}, \{c(\theta_1)\}, \{\theta_2\}, \{c(\theta_2)\}, \{\theta_1 \cup \theta_2\}, \dots, \{c(\theta_1 \cup \theta_2)\}, \{\theta_1 \cap \theta_2\}, \{c(\theta_1 \cap \theta_2)\}\} \quad (3.8)$$

Il appartient donc à l'utilisateur de définir le cadre de discernement approprié compte-tenu de l'application et de la puissance de calcul disponible. En effet, passer d'un power set à un hyper-power set ou à un super-power set implique un accroissement des hypothèses à traiter comme le montre le tableau 3.3. [Abb09] estime à $1,7 \cdot 10^{15} Go$ l'espace mémoire nécessaire dans le cas où $n = 8$ dans l'hypothèse où chaque élément de D^Θ occupe $2^n - 1$ bits.

$ \Theta = n$	$ 2^\Theta = 2^n$	$ D^\Theta $	$ S^\Theta \sim 2^{2^n - 1}$
2	4	5	8
3	8	19	128
4	16	167	32768
5	32	7580	$2,1 \cdot 10^9$
6	64	7828353	$9,2 \cdot 10^{18}$
7	128	$\sim 2,4 \cdot 10^{12}$	$\sim 1,7 \cdot 10^{38}$

TAB. 3.3 – Nombre d'hypothèses à évaluer fonction du nombre de classes et types de cadres utilisés.

L'énumération des hypothèses pour $|S^\Theta|$ n'est pas triviale, et est communément appelée problème de

Dedekind. Le lecteur pourra se rapporter à [KT10] et [DS09] pour de plus amples détails.

Remarque : *Nous avons fait l'hypothèse dans ces descriptions que le cadre de discernement est suffisamment exhaustif pour contenir le résultat approprié. On parle alors de monde "clos" (ou "fermé"). Dans le cas contraire il s'agira de monde "ouvert" et le résultat peut alors être en dehors de Θ . Usuellement, l'insertion d'une classe fictive θ_{n+1} permet de combler ce manque d'exhaustivité sans pour autant permettre d'appréhender sa signification.*

Utiliser des sur-ensembles de Θ permet théoriquement des résultats plus précis mais a pour inconvénient de nécessiter des capacités de traitement importantes. Nous nous limiterons donc dans cette étude au power set 2^Θ contenant les disjonctions de Θ . Il est toutefois possible de passer ultérieurement à D^Θ voire à S^Θ , moyennant peu de frais ce qui est un avantage du cadre crédibiliste par rapport aux autres cadres de discernement. Nous nous intéressons maintenant à la création des masses permettant d'évaluer la croyance que l'on a en une information.

Création du jeu de masses

Les experts, ou agents rationnels selon [Den97], émettent un avis pour chacune des propositions de 2^Θ (identiquement pour D^Θ et S^Θ). Ces avis ou masses, notés m , prennent la forme de fonctions définies par :

$$m : 2^\Theta \rightarrow [0, 1] \quad (3.9)$$

La somme de ces masses est, pour un expert donné, égale à 1 (i étant le numéro de l'expert et A la proposition considérée) :

$$\sum_{A \in 2^\Theta} m_i(A) = 1 \quad (3.10)$$

Chacune des masses est représentative de la croyance accordée. Plus le degré de croyance est élevé et plus la masse associée est proche de 1. Ainsi la certitude totale en une proposition A se traduit par $m(A) = 1$ et l'ignorance totale par $m(\theta_1 \cup \dots \cup \theta_n) = 1$ plus couramment notée $m(\Theta) = 1$.

Pour affecter des valeurs aux masses, deux méthodes sont possibles :

1. La méthode dite de Dencœux ([Den00]) qui repose sur l'évaluation d'une distance euclidienne notée d_i séparant le vecteur x correspondant aux caractéristiques de l'objet à évaluer et x_i le $i^{\text{ème}}$ le terme correspondant au centroïde du vecteur d'apprentissage \mathcal{L} (notons qu'il est possible de remplacer x_i correspondant au centroïde intervenant dans d_i par les k plus proches voisins comme utilisé dans [LMK⁺06]). Ce type d'approche s'appuie sur une fonction strictement décroissante ϕ_i (vérifiant $\phi_i(0) = 1$ et $\lim_{d \rightarrow 0} \phi_i(d) = 0$) qui prend généralement la forme : $\phi_i(d_i) = \exp(-\gamma_i(d_i)^2)$. Le paramètre $\gamma_i > 0$ (pouvant être identique pour l'ensemble des classes), influence la volatilité des classes ainsi que les performances de la fusion. [Mar05] propose d'utiliser l'hypothèse suivante $\forall i \in \mathbb{N}, \gamma_i = \gamma = 1/d^2$, d étant la distance moyenne entre les vecteurs d'apprentissage vérifiant $\{\theta_i\}$. Ainsi la fonction de croyance $m_{ij}(\cdot)$ (premier indice pour le numéro de l'hypothèse et le second pour celui de la source) est définie par :

$$\begin{cases} m_{ij}(\{\theta_i\}) = \alpha_j \phi_i(d_i) \\ m_{ij}(\Theta) = 1 - \alpha_j \phi_i(d_i) \end{cases} \quad (3.11)$$

Le coefficient α_j ($\in [0, 1]$) traduit la représentativité de l'information, ce qui s'avère différent de la notion de fiabilité faisant intervenir les mécanismes d'affaiblissement qui seront explicités plus en détail par la suite ;

2. La seconde est celle d'Appriou ([App98]) qui s'oriente plus vers les techniques probabilistes dont voici les deux variantes :

$$\text{modèle 1} \begin{cases} m_{ij}(\{\theta_i\}) = \alpha_j (R_j p(s_j | \{\theta_i\})) \\ m_{ij}(c\{\theta_i\}) = \alpha_j (1 - R_j p(s_j | \{\theta_i\})) \\ m_{ij}(\Theta) = 1 - \alpha_j \end{cases} \quad \text{modèle 2} \begin{cases} m_{ij}(\{\theta_i\}) = 0 \\ m_{ij}(c\{\theta_i\}) = \alpha_j (1 - R_j p(s_j | \{\theta_i\})) \\ m_{ij}(\Theta) = 1 - m_{ij}(c\{\theta_i\}) \end{cases} \quad (3.12)$$

Le coefficient R_j sert à normaliser les fonctions de masse et est compris entre 0 et $\frac{1}{\max_i p(s_j | \{\theta_i\})}$, où $p(s_j | \{\theta_i\})$ représente la densité de probabilités de s_j considérant $\{\theta_i\}$. L'estimation de $p(s_j | \{\theta_i\})$ peut se faire à l'aide d'une distribution de probabilités modélisée ou bien approchée expérimentalement.

Application du cadre évidentiel dans notre étude

Pour cette étude, la création des masses reposera sur l'approche d'Appriou (modèle 1) pour sa simplicité de mise en œuvre. Chaque expert va statuer sur la "vérité" de chaque proposition de 2^Θ . L'attribut mouvement n'est constitué que d'un seul expert. Ceci est également l'attribut forme. Cependant, pour l'attribut couleur, trois experts sont utilisés : un pour chaque couche de l'image (H/S/V). Effectuer l'allocation sur 2^Θ pour cinq experts demande un temps de calcul conséquent. Pour ne pas altérer la réactivité du système, nous segmentons le cadre de discernement Θ en quatre cadres distincts Θ_i (cf. remarque ci-dessous). Chaque Θ_i est composé des singleton $\{\theta_i\}$, $c\{\theta_i\}$, $\{\emptyset\}$ et $\{\Theta_i\}$. Cette fragmentation limite la puissance de la fusion évidentielle, la règle de combinaison n'ayant plus "accès" à l'intégralité du cadre de discernement d'origine. Cependant, l'adoption du formalisme crédibiliste nous permet de réutiliser rapidement l'ensemble des disjonctions 2^Θ si le système de calcul le permet.

Remarque : Les cadres de discernement peuvent être amenés à évoluer par des opérations de raffinement et de grossissement. Ainsi pour deux cadres de discernement distincts Θ_1 et Θ_2 , une masse $m^{\Theta_1}(X)$ peut s'écrire comme le produit résultant de ces deux espaces induisant une redéfinition de la fonction de croyance : $m^{\Theta_1 \uparrow \Theta_1 \times \Theta_2}(X)$. L'expression de cette nouvelle fonction de croyance au sein d'une règle de combinaison est proposée ici sous sa forme générique :

$$(m_1^{\Theta_1} \odot m_2^{\Theta_2})(X) = (m_1^{\Theta_1 \uparrow \Theta_1 \times \Theta_2} \odot m_2^{\Theta_2 \uparrow \Theta_1 \times \Theta_2})(X) \quad (3.13)$$

où l'opérateur \odot correspond à toute opération de combinaison de fonction de croyance.

Ce cas particulier n'est pas d'actualité ici car nous relaxons le cadre de discernement (génération de quatre cadres n'interagissant pas ensemble). Il ne faut donc pas confondre grossissement (ou raffinement) d'un cadre et décomposition.

	Couleur			Ligne	Mouvement
	s1	s2	s3	s4	s5
\emptyset					
Sang					
Peau					
Ch. stériles					
Instrument					
Sang \cup Peau					
Sang \cup Ch. stériles					
Sang \cup Instrument					
\vdots					
Θ					

(a) Cadre de discernement avant fractionnement.

	Couleur			Ligne	Mouvement
	s1	s2	s3	s4	s5
\emptyset	m_{01}	m_{02}	m_{03}	m_{04}	m_{05}
θ_i	m_{11}	m_{12}	m_{13}	m_{14}	m_{15}
$c\{\theta_i\}$	m_{21}	m_{22}	m_{23}	m_{24}	m_{25}
Θ_i	m_{31}	m_{32}	m_{33}	m_{34}	m_{35}

(b) Obtention du cadre Θ_i .

	Couleur			Ligne	Mouvement
	s1	s2	s3	s4	s5
\emptyset	m_{01}	m_{02}	m_{03}	m_{04}	m_{05}
Sang	m_{11}	m_{12}	m_{13}	m_{14}	m_{15}
$c\{\text{Sang}\}$	m_{21}	m_{22}	m_{23}	m_{24}	m_{25}
Θ_1	m_{31}	m_{32}	m_{33}	m_{34}	m_{35}

(c) Déconditionnement pour la classe sang.

	Couleur			Ligne	Mouvement
	s1	s2	s3	s4	s5
\emptyset	m_{01}	m_{02}	m_{03}	m_{04}	m_{05}
Peau	m_{11}	m_{12}	m_{13}	m_{14}	m_{15}
$c\{\text{Peau}\}$	m_{21}	m_{22}	m_{23}	m_{24}	m_{25}
Θ_2	m_{31}	m_{32}	m_{33}	m_{34}	m_{35}

(d) Déconditionnement pour la classe peau.

	Couleur			Ligne	Mouvement
	s1	s2	s3	s4	s5
\emptyset	m_{01}	m_{02}	m_{03}	m_{04}	m_{05}
Champ stérile	m_{11}	m_{12}	m_{13}	m_{14}	m_{15}
$c\{\text{Champ stérile}\}$	m_{21}	m_{22}	m_{23}	m_{24}	m_{25}
Θ_3	m_{31}	m_{32}	m_{33}	m_{34}	m_{35}

(e) Déconditionnement pour la classe champs stériles.

	Couleur			Ligne	Mouvement
	s1	s2	s3	s4	s5
\emptyset	m_{01}	m_{02}	m_{03}	m_{04}	m_{05}
Instrument	m_{11}	m_{12}	m_{13}	m_{14}	m_{15}
$c\{\text{Instrument}\}$	m_{21}	m_{22}	m_{23}	m_{24}	m_{25}
Θ_4	m_{31}	m_{32}	m_{33}	m_{34}	m_{35}

(f) Déconditionnement pour la classe instrument.

FIG. 3.4 – Exemple de classification par les séparateurs à vaste marge.

Ainsi, la classe "sang", tout comme les autres classes d'objets, forme un cadre de discernement à part entière, composé des propositions "sang", "non sang", "ignorance" et "conflit". Par convention nous noterons par Θ_1 le cadre de discernement de la classe sang, Θ_2 la classe peau, Θ_3 la classe champ stérile et Θ_4 les instruments.

Pour affecter les masses à chaque proposition des cadres Θ_1 , Θ_2 , Θ_3 et Θ_4 , nous utilisons la démarche proposée par [FL05]. Celle-ci consiste à parcourir l'image acquise par la caméra et à récupérer les coordonnées chromatiques du premier pixel. Au préalable, des histogrammes représentatifs des classes d'objets pour chacun des experts (couleur, mouvement et forme) sont stockés en mémoire. Chaque classe possède donc cinq histogrammes.

Pour l'attribut couleur, la valeur du premier histogramme pour la première composante colorimétrique du pixel correspond à $p(s_1 | \{\theta_i\})$. Il est alors possible de calculer m_{01} , m_{11} , m_{21} et m_{31} . Cette procédure est répétée pour les deux autres composantes. On obtient donc $p(s_2 | \{\theta_i\})$ et $p(s_3 | \{\theta_i\})$ ce qui nous donne au final neuf masses représentatives de la classe d'objet pour l'attribut couleur. Vient ensuite l'attribut forme qui nous permet de calculer les masses m_{04} , m_{14} , m_{24} et m_{34} . Nous effectuons la même chose pour l'attribut mouvement, qui grâce à l'estimation du flot optique est capable de nous restituer une image des pixels en mouvement. Le niveau de gris des pixels est représentatif de l'amplitude du mouvement. Ainsi, en comparant le niveau de gris du pixel considéré à l'histogramme de la classe pour cet attribut, nous obtenons de nouvelles masses : m_{05} , m_{15} , m_{25} et m_{35} . Au final, chacune des propositions de Θ_1 possède cinq masses issues de cinq experts (sources) différents.

Afin de gagner encore en temps de calcul nous prenons le parti d'éliminer certains experts superflus :

- les experts forme et mouvement pour les classes sang, peau et champ stérile dans la mesure où la reconnaissance de ce type d'objet est principalement conditionnée par la couleur (tableaux 3.4c, 3.4d et 3.4e) ;
- l'expert couleur pour la classe instrument car nous ne possédons pas les histogrammes représentatifs de $p(s_i | \{\text{instrument}\})$, $i \in \{1, 2, 3\}$ (tableau 3.4f).

Nous avons considéré, jusqu'à présent, que ces sources ont des performances équivalentes. Or, ceci n'est pas toujours le cas notamment lorsqu'il s'agit de sources hétérogènes. Pour intégrer ces disparités, les masses vont être affaiblies.

Affaiblissement des masses

L'affaiblissement de masse permet de prendre en compte la fiabilité des sources d'informations. Ceci revient à pondérer la masse m_{ij} suivant les modalités suivantes :

$$\begin{cases} m_{ij}^\beta(A) = \beta_j m_{ij}(A), \forall A \in 2^\Theta \\ m_{ij}^\beta(\Theta) = 1 - \beta_j (1 - m_{ij}(\Theta)) \end{cases}$$

L'affaiblissement vérifie donc $\beta_{ij} \in [0, 1]$, les extrema représentant respectivement la non-fiabilité (entraînant $m_{ij}^\beta(\Theta) = 1$) et la fiabilité totale de la source pour la proposition considérée. Cependant, quantifier la fiabilité d'un capteur n'est pas toujours aisé. De plus, dans notre cas, la reconnaissance des objets est influencée par l'éclairage de la scène. Dans la mesure où le chirurgien peut potentiellement en moduler le niveau, il est intéressant de s'attarder sur les performances de la méthode pour différentes configurations d'éclairage.

Nous proposons d'utiliser une matrice de confusion qui condense dans une matrice, l'aptitude des méthodes à bien classer les pixels mais aussi les erreurs d'appariement. Il est alors possible d'attribuer β_{ij} aux valeurs contenues dans la matrice. Dans un cadre plus courant, la définition de β_{ij} se fait en calculant la moyenne de la diagonale de la matrice de confusion ($\beta_{ij} \sim \beta_j$). Cette matrice de confusion est une synthèse simple des résultats de classification où chaque cellule donne le taux de classification de la proposition A_i en A_j . Si l'on considère un problème à deux classes $\Theta = \{\{\text{sang}\}, \{\text{peau}\}\}$ où un expert doit se prononcer sur l'appartenance des pixels d'une image au cadre de discernement Θ , il serait possible de trouver la matrice de confusion suivante :

	sang ¹	non sang ¹
sang ²	VP	FP
non sang ²	FN	VN

TAB. 3.4 – Exemple de matrice de confusion ("1" correspond à la vérité terrain et "2" à celle du système).

- VP : nombre de pixels reconnus correctement comme étant du sang ;
- VN : identification correcte des pixels non-sang ;
- FP : pixels associés à du sang et qui en réalité ne l'étaient pas (fausse détection) ;
- FN : pixels manqués par le système.

A partir des valeurs contenues dans les cellules, il est possible de calculer certaines grandeurs plus représentatives de l'efficacité du système de reconnaissance telles que la précision \mathcal{P} et le rappel \mathcal{R} . La première définie par $\mathcal{P} = \frac{VP}{VP+FP}$, correspond à la probabilité que l'avis soit correct lorsqu'il s'agit de sang. Inversement, la seconde rend compte des bonnes détections de la classe non-sang, $\mathcal{R} = \frac{VP}{VP+FN}$.

Dans le cadre de ces travaux, nous utiliserons comme coefficient de pondération des masses, la distance définie par :

$$D = \sqrt{(1 - \text{Précision})^2 + (1 - \text{Rappel})^2}$$

$$\beta_j = 1 - \frac{1}{\sqrt{2}}D \quad (3.14)$$

Ainsi une reconnaissance parfaite se traduit par $\beta_j = 1$ (a contrario $\beta_j = 0$). Une fois chacune des masses affaiblies, il faut procéder à leur agrégation pour n'obtenir qu'une seule masse représentative de chaque proposition. Les règles de combinaison permettant la fusion de ces avis sont explicitées dans la section suivante.

Les règles de combinaison en fusion évidentielle

Le but des règles de combinaison consiste en une mutualisation des connaissances permettant, suivant différentes modalités, de fournir en sortie de processus une masse unique pour chacune des propositions composant le power set.

Règle de combinaison conjonctive et disjonctive

La première règle conçue est celle de Dempster en 1967, étendue en 1976 par Shafer, dite de combinaison conjonctive normalisée (ou encore de combinaison de Dempster et de somme orthogonale). Pour une proposition donnée, les masses sont "additionnées" les unes aux autres et le résultat est noté m_{\oplus} ([LDB01]) :

$$m_{\oplus} = m_1 \oplus m_2 \oplus \dots \oplus m_n \quad (3.15)$$

avec $\sum_{A \in 2^{\Theta}} m_{\oplus}(A) = 1$.

Pour deux experts 1 et 2 fournissant respectivement m_1 et m_2 , nous obtenons (sous hypothèse de totale

fiabilité des sources et A étant un élément focal appartenant à Θ :

$$\text{règle de Dempster} \quad \left\{ \begin{array}{l} m_{\oplus}^{\Theta}(A) = (m_1^{\Theta} \oplus m_2^{\Theta})(A), \quad \forall A \subseteq \Theta, A \neq \emptyset \\ \quad = \frac{1}{1-k} m_{1 \cap 2}^{\Theta}(A) \\ m_{\oplus}^{\Theta}(\emptyset) = 0 \end{array} \right. \quad (3.16)$$

(règle conjonctive à 2 experts)

Le terme $m_{1 \cap 2}$ de l'équation (3.16), connu sous le nom de règle conjonctive de Smets, est présenté en (3.17). Une interprétation intuitive de cette règle conjonctive correspond à l'opération "ET" de la théorie ensembliste.

$$m_{1 \cap 2}(A) = \sum_{A_1 \cap A_2 = A} m_1^{\Theta}(A_1) m_2^{\Theta}(A_2) \quad (3.17)$$

Le réel k (appartenant à $[0,1]$), défini par $k = m_{\oplus}^{\Theta}(\emptyset) = \sum_{A_1 \cap A_2 = \emptyset} m_1^{\Theta}(A_1) m_2^{\Theta}(A_2)$ dans (3.16), est révélateur de l'inconsistance de la fusion ([JF96]) et traduit le conflit global (k attribué à $\{\emptyset\}$). La règle de combinaison de Dempster peut donc être appliquée aux cadres de discernement fermés ; sa généralisation à N experts s'écrit donc :

$$\text{règle de Dempster} \quad \left\{ \begin{array}{l} m_{\cap}^{\Theta}(A) = (\oplus_{j=1}^N m_j^{\Theta})(A), \quad \forall A \subseteq \Theta \\ \quad = \frac{1}{1-k} \sum_{A_1 \cap \dots \cap A_N = A} \prod_{j=1}^N m_j^{\Theta}(A_j) \end{array} \right. \quad (3.18)$$

(règle conjonctive à N experts)

L'expression généralisée du conflit devient donc : $k = \sum_{A_1 \cap \dots \cap A_N = \emptyset} \prod_{j=1}^N m_j^{\Theta}(A_j)$.

Une variante de la règle conjonctive, dite règle disjonctive proposée par [DP88] permet de réaliser l'opération logique "OU" entre les masses de croyance. On obtient pour N sources :

$$\text{règle disjonctive} \quad \left\{ \begin{array}{l} m_{\cup}^{\Theta}(A) = \sum_{A_1 \cup \dots \cup A_N = A} \prod_{j=1}^N m_j^{\Theta}(A_j), \quad \forall A \subseteq \Theta \\ m_{\cup}^{\Theta}(\emptyset) = 0 \end{array} \right. \quad (3.19)$$

(N sources)

Les règles de combinaison conjonctive et disjonctive ne sont pas les seules permettant d'agréger des masses de croyance. D'autres auteurs ont proposé des solutions notamment pour répartir plus judicieusement le conflit. Par la suite, nous utiliserons les règles conjonctives et disjonctives. Cependant de nombreux travaux décrivent ces règles du fait de leur non-idempotence. Dès lors, ce processus de fusion sera étendu aux règles de [Yag87], [DP88] et [Sma05]. Nous nous limitons ici aux règles les plus connues. Le lecteur pourra compléter cette description des règles de combinaison en se rapportant aux travaux de [LCV02], [Ina91] et [FDV⁺06] sur les fonctions génériques.

Règle de combinaison de Yager

Cette règle permet de résoudre le problème de non-idempotence de la règle conjonctive. [Yag87] proposa de répartir le conflit sur l'ignorance totale à savoir Θ telle que pour $\forall A \in 2^{\Theta}$, $A \neq \Theta$, $A \neq \emptyset$:

$$\text{règle de Yager} \quad \left\{ \begin{array}{l} m_Y^{\Theta}(A) = m_{1 \cap 2}^{\Theta}(X) \\ m_Y^{\Theta}(\Theta) = m_{1 \cap 2}^{\Theta}(\Theta) + m_{1 \cap 2}^{\Theta}(\emptyset) \\ m_Y^{\Theta}(\emptyset) = 0 \end{array} \right. \quad (3.20)$$

Règle de combinaison de Dubois et Prade

Proposé par Dubois et Prade, cette règle distribue le conflit partiel sur l'union des sources le générant (l'intersection de ces deux propositions étant vide) :

$$\text{règle de Dubois et Prade} \begin{cases} m_{\text{DP}}^{\ominus}(A) &= m_{\cap}^{\ominus}(A) + \sum_{\substack{A_1 \cup A_2 = A \\ A_1 \cap A_2 = \emptyset}} m_1^{\ominus}(A_1)m_2^{\ominus}(A_2) \\ m_{\text{DP}}^{\ominus}(\emptyset) &= 0 \end{cases} \quad (3.21)$$

Le résultat tendra donc vers une conjonction en l'absence de conflit et vers une disjonction lorsque celui-ci sera proche de 1.

Règle de combinaison PCR5

Actuellement, l'une des techniques les plus abouties pour la gestion du conflit, la règle dite Proportional Conflict Redistribution développée par [Sma05] est opérante même dans D^{\ominus} . Sa cinquième version (PCR5) est plus performante que les précédentes mais ne s'applique qu'à deux sources (3.22). Ce dernier point a été pallié par [MO06] qui propose une extension à N experts sous le nom de règle PCR6.

$$m_{\text{PCR5}}^{\ominus}(X) = m_{\cap 2}^{\ominus}(X) + \sum_{A \cap B = \emptyset} \left(\frac{m_1^{\ominus}(A)^2 \cdot m_2^{\ominus}(B)}{m_1^{\ominus}(A) + m_2^{\ominus}(B)} + \frac{m_1^{\ominus}(B) \cdot m_2^{\ominus}(A)^2}{m_1^{\ominus}(A) + m_2^{\ominus}(B)} \right) \quad (3.22)$$

Nous venons de voir les règles de combinaison de masse les plus utilisées permettant d'avoir un avis unique sur chaque proposition de 2^{\ominus} . Il s'agit donc maintenant de décider de la classe d'objets la plus représentative du pixel considéré. En l'absence d'information complémentaire les règles conjonctives, disjonctives, de Yager, Dubois et Prade et PCR5 seront utilisées. L'analyse des performances statuera sur la règle la plus adaptée au problème de reconnaissance d'objets dans une scène chirurgicale.

3.1.3.2 Etape 2 : l'étape pignistique

Cette étape consiste, à partir des masses résultantes issues de l'étape crédale, à en extraire une sortie unique représentative du problème posé. Pour cela, plusieurs critères sont possibles :

- le maximum de croyance ;
- le maximum de plausibilité ;
- le maximum de probabilité pignistique.

Les deux premiers critères permettent un comportement respectivement pessimiste et optimiste du système. Dans la mesure où le maximum de probabilité pignistique est un compromis entre ces deux critères, nous privilégions son usage. Basé sur le principe de la raison insuffisante, la probabilité pignistique notée $BetP$ est définie par :

$$BetP(A) = \sum_{\substack{A \in \Theta, B \neq \emptyset \\ B \in 2^{\ominus}}} \frac{|A \cap B|}{|B|} \frac{m(B)}{1 - m(\emptyset)} \quad (3.23)$$

Ainsi, pour chaque singleton du cadre de discernement, la mesure $BetP$ est calculée. La proposition correspondant au maximum de $BetP$ est définie comme étant la réponse du système au problème de reconnaissance d'objets. Cette prise de décision peut être combinée à la notion de choix en environnement risqué. Résumé dans [Den97], ce type de méthode s'appuie, à l'image du cadre probabiliste, sur une

décision minimisant l'espérance du coût f (à valeur dans \mathbb{R}). Ainsi, dans le cas d'une distribution de probabilités $BetP$, l'espérance associée s'écrit :

$$E_{BetP}(f) = \sum_{A \in \Theta} f(A)BetP(A) \quad (3.24)$$

Il est possible d'utiliser d'autres mécanismes que celui de la probabilité pignistique pour le calcul de (3.24), et atteindre des comportements pessimistes (minimisation de l'espérance supérieure) ou optimistes (minimisation de l'espérance inférieure).

Dans le cadre de cette étude, nous utiliserons la prise de décision définie en (3.23). Quatre propositions composant le cadre de discernement Θ_i sont calculés. Pour composer une image unique représentative des objets en présence dans l'image, nous sélectionnons le maxima des probabilités pignistiques des premières propositions. L'utilisation de la notion de risque pour l'établissement des probabilités pignistiques constitue une perspective de ces travaux.

Nous avons vu l'ensemble des mécanismes permettant de reconnaître des objets présents dans la scène chirurgicale. Nous proposons en table 1 un synopsis de la procédure de désignation de la zone d'intérêt.

Entrée du programme :	Image I couleur de taille $M \times N$ de la scène opératoire.
Sortie du programme :	Image H en niveaux de gris représentative de la présence de la classe dans la scène opératoire de taille $M \times N$ ($\times 4$).

```

I=Acquisition de l'image;
IM=Estimation du flot optique (I);
IL=Transformée de Hough (I);
foreach  $i : 1 \rightarrow M$  do
  |
  | foreach  $j : 1 \rightarrow N$  do
  | |
  | | foreach  $c : 1 \rightarrow 4$  do
  | | |
  | | | ( $p_1, p_2, p_3$ )=Récupération des coordonnées chromatiques du pixel ( $I(i, j)$ );
  | | |  $p_4$ =Récupération du niveau de gris du pixel ( $I_M(i, j)$ );
  | | |  $p_5$ =Récupération du niveau de gris du pixel ( $I_L(i, j)$ );
  | | | foreach  $k : 1 \rightarrow 5$  do
  | | | |
  | | | | foreach  $q : 1 \rightarrow 3$  do
  | | | | |
  | | | | | ( $m_{qk}, \dots, m_{qk}$ )=Création des fonctions de croyances ( $\text{hist}_k(\text{classe}), p_k$ );
  | | | | | end
  | | | | |  $V_q = (m_{qk}, \dots, m_{qk})$ ;
  | | | | end
  | | | | ( $m_1, m_2, m_3$ )=Combinaison des masses ( $V_1, V_2, V_3$ );
  | | | | ( $\text{Bet}P_1, \text{Bet}P_2, \text{Bet}P_3$ )=Calcul des probabilités pignistiques ( $m_1, m_2, m_3$ );
  | | | |  $\tilde{H}_{\text{classe}}(i, j) = (\text{Bet}P_1)$ ;
  | | | end
  | | |  $H(i, j) = \max_{\text{classe}} \tilde{H}_{\text{classe}}(i, j)$ ;
  | | end
  | end
end

```

Algorithme 1 : Processus pour la réorientation de la coupole.

Notations :

I	image acquise par la caméra ;	I_M	image représentative du mouvement dans I ;
I_L	image représentative des lignes droites dans I ;	(i, j)	coordonnée du pixel dans l'image I ;
M, N	taille de l'image I ;	classe	classe d'objet considéré (sang, instrument, etc) ;
p_\circ	coordonnées chromatiques du pixel $[0, 255]$ pour la couche \circ ;	$hist_\circ$	histogramme représentatif de la classe dans la couche \circ ;
$m_{\circ 1}$	masse de la couche \circ pour la proposition 1 ;	V_\circ	vecteur contenant l'ensemble des masses d'une même couche ;
m_\circ	masse représentative de la proposition \circ ;	$BetP_\circ$	probabilité pignistique de la proposition \circ ;
\tilde{H}_\circ	image pour la couche \circ contenant l'ensemble des probabilités pignistiques de la proposition 1 ;	H_\circ	image finale représentative des objets en présence dans la scène chirurgicale ;

3.2 Définition de la zone d'intérêt

Pour accéder à la position de la zone d'intérêt notre approche repose sur la reconnaissance des objets en présence dans la scène opératoire. Il faut s'interroger sur le comportement que doit adopter le système face à ces objets pour ainsi identifier la zone d'intérêt. Ceci revient donc à comprendre l'importance qu'ont ces objets pour le chirurgien et de là, privilégier certains plutôt que d'autres.

Nous nous attachons donc dans le point suivant à aborder la théorie de la décision qui offre une structure pour ce type de prise de décision, ainsi que les principaux critères utilisés en la matière. Nous verrons que les choix d'un individu (chirurgien inclus) ne sont pas nécessairement objectifs (ou rationnels) ; et qu'il est possible d'opposer à cette vision normative (prise de décision optimale) des choix descriptifs (ce que les agents font réellement).

3.2.1 Introduction à la théorie de la décision

Issue des activités économiques notamment pour les assurances et les banques, la théorie de la décision traite des problèmes de choix en environnement risqué (au sens de [Kni21] où les probabilités associées aux issues sont connues) ou incertain (où les probabilités ne sont pas connues). Cette théorie permet de modéliser le comportement d'un individu devant se prononcer entre différentes loteries. Le terme "loterie" est utilisé dans le sens de proposition. En l'occurrence une loterie \mathcal{L} est un vecteur dont les composantes sont des couples associant probabilité d'occurrence p et gain potentiel g :

$$\mathcal{L} = \{(g_1, p_1), \dots, (g_m, p_m)\}$$

L'ensemble des loteries forme le vecteur $L = \{\mathcal{L}_1, \mathcal{L}_2, \dots, \mathcal{L}_N\}$ (avec N le nombre maximal de loteries). La traduction de la préférence d'une loterie \mathcal{L}_q par rapport à \mathcal{L}_k ($k \neq q$) sera notée \succeq , telle que : $\mathcal{L}_q \succeq \mathcal{L}_k$. Cette convention d'écriture permet de matérialiser les préférences d'un individu, mais doit répondre à certaines propriétés qui sont :

- la complétude : $\forall \mathcal{L}_q \in L$ et $\forall \mathcal{L}_k \in L$, $\mathcal{L}_q \succeq \mathcal{L}_k$;
- l'indifférence : $\forall \mathcal{L}_q \in L$ et $\forall \mathcal{L}_k \in L$, $\mathcal{L}_q \sim \mathcal{L}_k \Leftrightarrow \mathcal{L}_q \succeq \mathcal{L}_k$ et $\mathcal{L}_k \succeq \mathcal{L}_q$;

- la transitivité : $\forall \mathcal{L}_q \in L, \forall \mathcal{L}_k \in L \text{ et } \forall \mathcal{L}_m \in L, \text{ si } \mathcal{L}_q \succeq \mathcal{L}_k \text{ et } \mathcal{L}_k \succeq \mathcal{L}_m \text{ alors } \mathcal{L}_q \succeq \mathcal{L}_m$.

Au-delà des aspects normatifs ou descriptifs précédemment évoqués, il n'est pas obligatoire d'impliquer des probabilités dans le processus de décision. Dès lors, ceci conduit à deux alternatives distinctes :

- Environnement incertain : $\mathcal{L} = \{(g_1), \dots, (g_m)\}$.
- Environnement risqué : $\mathcal{L} = \{(g_1, p_1), \dots, (g_m, p_m)\}$.

Les grandeurs g_i et p_i correspondent aux notions de gain et de probabilité. Dans notre application, désigner la zone d'intérêt du chirurgien revient à sélectionner la loterie ayant la plus haute valeur. Cette loterie optimale notée \mathcal{L}^* sera injectée à l'entrée de la boucle de régulation pour être in fine centrée dans le repère image et donc éclairer optimalement.

Ceci soulève la question de l'environnement à utiliser ainsi que la signification des g_i et p_i dans notre application.

3.2.1.1 Prise de décision en environnement incertain

La notion de probabilité est absente pour ce type d'environnement. Il y a donc une équi-probabilité des propositions. La prise de décision se base uniquement sur le gain. Dans notre étude, le gain peut être vu comme une matérialisation de l'importance accordée par le chirurgien à chacune des classes. En conséquence, plus l'importance est grande, plus la classe d'objets est représentative de la zone d'intérêt. Or, cette importance varie au cours de l'opération. Compte tenu des discussions menées en chapitre 2 et sur la base des visites en bloc opératoire, nous estimons que l'instant de la prise de décision et la situation opératoire sont les deux principaux paramètres influençant l'importance de l'objet. La figure 3.5 illustre ces variations pour la classe sang.

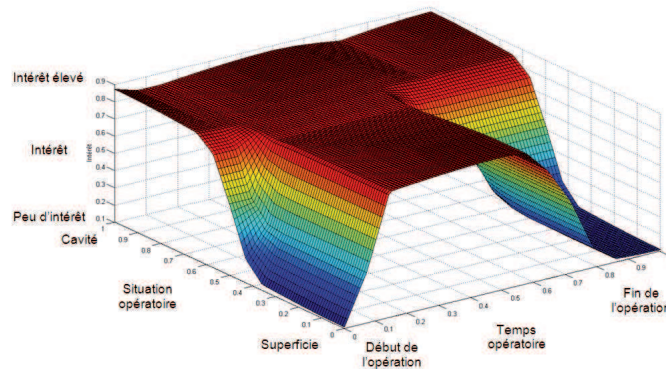
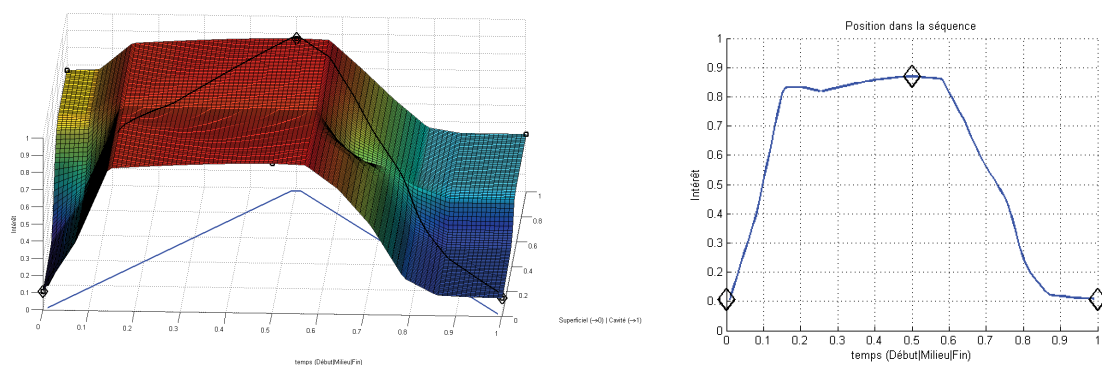


FIG. 3.5 – Exemple de surface représentative de l'intérêt que porte le chirurgien selon le temps opératoire et la situation opératoire.

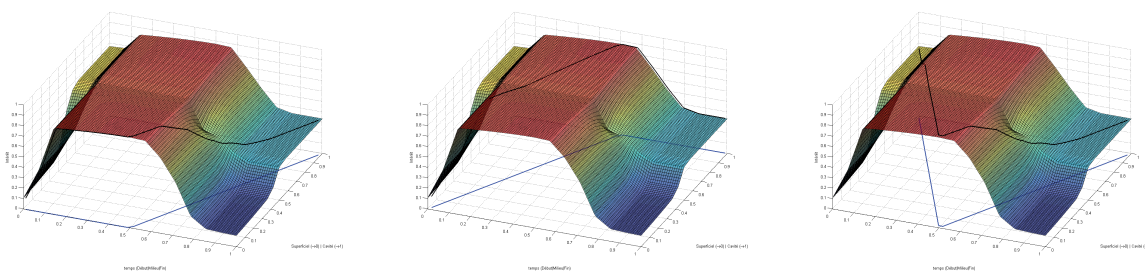
Sur cette figure, si l'on considère une opération de type $\{S, C, S\}$ (autrement dit qui commence en superficie tend en milieu d'intervention vers une cavité et revient en superficie) le sang, en début de séquence, n'est pas très représentatif de la zone d'intérêt car l'importance accordée à cette classe est faible. Cependant, à mesure que l'opération se déroule, l'intérêt de cette classe augmente, puis devient mineur lorsque la situation opératoire revient en superficie. Pour mieux comprendre l'évolution de l'importance

d'une classe, nous présentons figure 3.6a la même surface qu'en figure 3.5 à laquelle nous ajoutons la trajectoire correspondant à la séquence opératoire $\{S, C, S\}$ (courbe bleue) et la projection d'une séquence sur la surface dite d'intérêt (en noir). La courbe bleue montre le cheminement de l'opération passant de la superficie en début d'opération, à une situation en cavité au milieu puis de nouveau en superficie en fin d'intervention. La courbe noire visible au-dessus de la surface d'intérêt illustre la variabilité de l'importance que prend la classe sang pour cette opération dans le processus de désignation, celle-ci étant présentée sous un autre point de vue (figure 3.6b). Ainsi, pour une classe considérée, les huit séquences possibles conduisent à huit modèles d'évolution de l'importance (figures 3.6c, 3.6d et 3.6e).

Pour déterminer l'importance g_{classe} d'une classe, il suffit donc d'estimer conjointement quels sont la situation opératoire (S_{op}) et l'instant opératoire (t_{op}) à l'instant courant. En réitérant cette opération pour chacun des pixels constituant l'image, nous formons une nouvelle image cartographiant la répartition d'importance des objets présents dans l'image courante.



(a) Représentation 3D de l'évolution de l'intérêt de la classe sang (courbe bleue) : projection de la séquence sur la surface d'intérêt. (b) Trajet noir présent en (a) en fonction du temps pour la même séquence opératoire $\{S, C, S\}$.



(c) Exemple de scénario $\{S, S, C\}$. (d) Exemple de scénario $\{S, C, C\}$. (e) Exemple de scénario $\{C, S, C\}$.

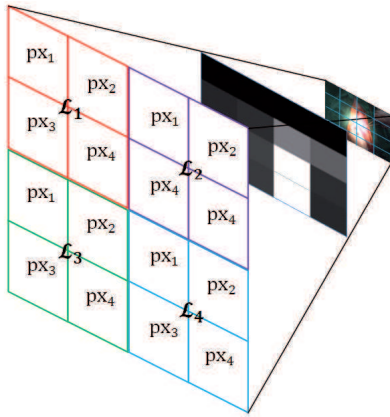
FIG. 3.6 – Mise en évidence de la variabilité de l'intérêt de la classe sang selon le temps opératoire et la situation opératoire.

Il reste maintenant à expliciter par quel critère il convient de discriminer une loterie, plutôt qu'une autre. La première possibilité est de considérer le critère de la raison insuffisante, appelé critère de Laplace. Compte tenu de l'équi-probabilité des loteries, le choix se base sur la moyenne des gains la plus forte,

autrement dit :

$$\mathcal{L}^* = \arg \max_{\mathcal{L}_q \in \mathcal{L}} \frac{1}{|\mathcal{L}|} \sum_{i=1}^{|\mathcal{L}|} g_i \quad (3.25)$$

Le tableau 3.5 illustre ce type de critère et fait référence à la figure 3.7 qui présente la décomposition de l'image courante en quatre loteries (supposées homogènes pour l'exercice) composées de quatre pixels chacune.



Loterie	px1	px2	px3	px4	Laplace
\mathcal{L}_1	0.1	0.5	0.4	0.9	0.475
\mathcal{L}_2	0.5	0.4	0.3	0.2	0.35
\mathcal{L}_3	0.1	0.6	0.2	0.7	0.4
\mathcal{L}_4	0.8	0.4	0.1	0.5	0.45

TAB. 3.5 – Jeu de valeurs pour différentes loteries en environnement incertain ($|\mathcal{L}| = 4$).

FIG. 3.7 – Décomposition de l'image courante en loterie.

Compte tenu des valeurs affectées aux pixels, la loterie sélectionnée est \mathcal{L}_1 ($\mathcal{L}_1 \succeq \mathcal{L}_4 \succeq \mathcal{L}_3 \succeq \mathcal{L}_2$). Cependant, maximiser le gain moyen n'est pas le seul mécanisme de décision possible en environnement incertain, citons ici les critères :

- **Maximax** : $\mathcal{L}^* = \arg \max_{\mathcal{L}_q \in \mathcal{L}} \sup \mathcal{L}_i$

Le système est à la recherche du gain maximal qu'il peut obtenir sans prise en compte des risques inhérents (comportement optimiste du système) ; le gain maximal pouvant être issu d'une erreur de reconnaissance de l'objet, de l'évaluation du temps opératoire ou encore de la situation opératoire ;

- de **Wald** : $\mathcal{L}^* = \arg \max_{\mathcal{L}_q \in \mathcal{L}} \inf \mathcal{L}_i$

Correspond à un comportement pessimiste. Le système va chercher à s'assurer un gain minimal ;

- d'**Hurwicz** : $\mathcal{L}^* = \arg \max_{\mathcal{L}_q \in \mathcal{L}} \alpha \inf \mathcal{L}_i + (1 - \alpha) \sup \mathcal{L}_i$

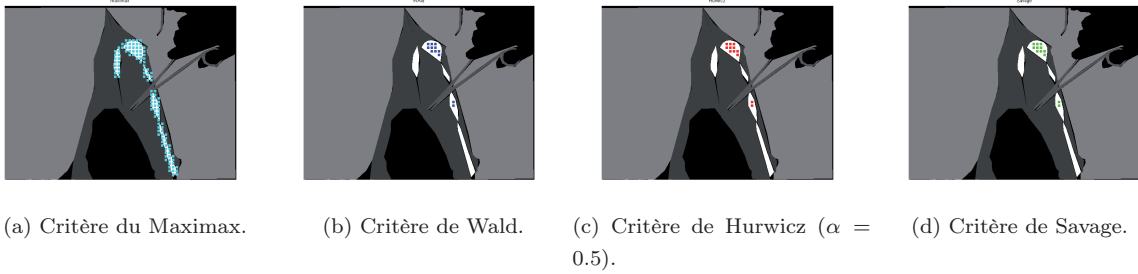
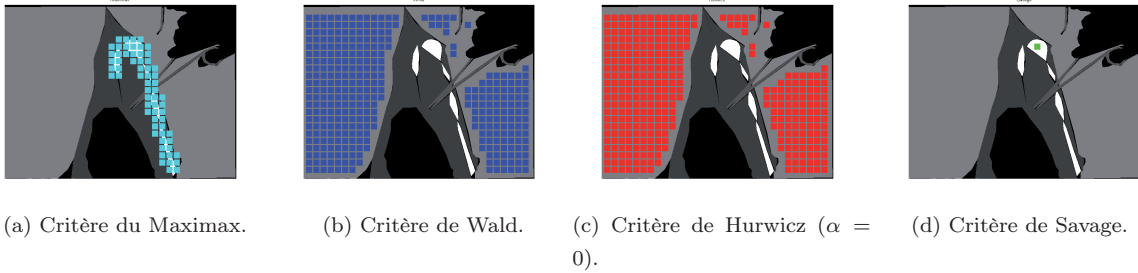
Suivant la valeur de α , le système adopte un comportement plus ou moins pessimiste ; ce qui suppose de connaître à l'avance α ;

- de **Savage** : $\mathcal{L}^* = \arg \min_{\mathcal{L}_q \in \mathcal{L}} \sum_{i=1}^{|\mathcal{L}|} (\sup \mathcal{L}_i) - \mathcal{L}_i$

Correspond au regret le plus faible parmi toutes les loteries.

Critère	Loterie sélectionnée	Détail du calcul du critère pour \mathcal{L}^*
Laplace	$\mathcal{L}_1(0.475) \succeq \mathcal{L}_4(0.45) \succeq \mathcal{L}_3(0.4) \succeq \mathcal{L}_2(0.35)$	$\mathcal{L}^* = (0.1 + 0.5 + 0.4 + 0.9)/4$
Maximax	$\mathcal{L}_1(0.9) \succeq \mathcal{L}_4(0.8) \succeq \mathcal{L}_3(0.7) \succeq \mathcal{L}_2(0.5)$	$\mathcal{L}^* = \sup(0.1, 0.5, 0.4, 0.9)$
Wald	$\mathcal{L}_2(0.2) \succeq \mathcal{L}_1(0.1) \sim \mathcal{L}_3(0.1) \sim \mathcal{L}_4(0.1)$	$\mathcal{L}^* = \inf(0.5, 0.4, 0.3, 0.2)$
Hurwicz ($\alpha = 0$)	$\mathcal{L}_1(0.9) \succeq \mathcal{L}_4(0.8) \succeq \mathcal{L}_3(0.7) \succeq \mathcal{L}_2(0.5)$	$\mathcal{L}^* = 0 \times 0.1 + (1 - 0) \times 0.9$
Hurwicz ($\alpha = 1$)	$\mathcal{L}_2(0.2) \succeq \mathcal{L}_1(0.1) \sim \mathcal{L}_3(0.1) \sim \mathcal{L}_4(0.1)$	$\mathcal{L}^* = 1 \times 0.1 + (1 - 1) \times 0.9$
Hurwicz ($\alpha = 0.5$)	$\mathcal{L}_4(0.6) \succeq \mathcal{L}_1(0.5) \succeq \mathcal{L}_3(0.4) \succeq \mathcal{L}_2(0.35)$	$\mathcal{L}^* = 0.5 \times 0.4 + (1 - 0.5) \times 0.8$
Savage	$\mathcal{L}_1(0.8) \succeq \mathcal{L}_1(0.9) \succeq \mathcal{L}_3(1.1) \succeq \mathcal{L}_2(1.3)$	$\mathcal{L}^* = 0.8 - 0.1 + 0.6 - 0.5 + 0.4 - 0.4 + 0.9 - 0.9$

TAB. 3.6 – Exemple de critères de décision en environnement incertain.

FIG. 3.8 – Exemple de loterie sélectionnée selon différents critères ($|\mathcal{L}|=5$ pixels, pas de recherche définie à 10pixels).FIG. 3.9 – Exemple de loteries sélectionnées selon différents critères ($|\mathcal{L}|=15$ pixels, pas de recherche définie à 20pixels).

Afin d'illustrer ces critères, nous proposons en tableau 3.6, le même jeu de valeurs que celui exposé en 3.5. Nous pouvons constater que le choix du critère a une influence sur le résultat final. Ces disparités de comportement sont d'autant plus visibles pour des exemples non numériques. Considérons la figure 3.7, où l'importance des objets est définie manuellement (le blanc étant représentatif des régions de grande importance). Nous pouvons constater que toutes les loteries sélectionnées se sont focalisées sur les régions de plus grande importance (ici le sang) avec un comportement plus ou moins laxte (figure 3.8). Cependant, les résultats sont fonctions du cardinal de la loterie ainsi que du pas de recherche (figure 3.9). Cette modification du comportement affecte plus particulièrement les critères de Wald et d'Hurwicz en raison de l'intégration dans le critère du terme $\inf \mathcal{L}_i$. Seul les critères de Savage et Maximax sont insensibles

à ce type de modifications. Ils restent cependant discernables par leur comportement face à l'exclusivité de \mathcal{L}^* par rapport à la classe sang.

Ces exemples montrent que le choix d'un critère de décision n'est pas simple. Ce processus de décision se base uniquement sur des informations décorréelées de l'étape de reconnaissance et de ce fait des erreurs peuvent être engendrées. C'est pour cette raison qu'il est intéressant d'intégrer la notion d'incertitude dans la prise de décision ; on parle alors de choix en environnement risqué.

3.2.1.2 Prise de décision en environnement risqué

Nous avons vu qu'une loterie est composée de gain fictif mais aussi, dans le cas d'un environnement risqué, des probabilités notées p_i . Dans le cadre de notre étude, nous pouvons associer à ces p_i , entrant dans la composition de la loterie \mathcal{L}_q , les probabilités pignistiques (*BetP*) traduisant la croyance en l'identification de l'objet. Il reste cependant encore à définir comment combiner ces probabilités aux gains et quel critère utiliser pour procéder à un choix de la zone d'intérêt. Un moyen simple consiste à calculer l'espérance mathématique de la loterie \mathcal{L}_q telle que :

$$\mathbb{E}_{\mathcal{L}_q} = \sum_{i=1}^{|\mathcal{L}|} p(i)g(i) \quad (3.26)$$

La décision s'écrit alors :

$$\mathcal{L}^* = \arg \max_{\mathcal{L}_q \in L} \mathbb{E}_{\mathcal{L}_q} \quad (3.27)$$

Un exemple illustrant (3.26) est proposé (figure 3.10a et 3.10b) où sont modélisées respectivement la probabilité d'appartenance et l'importance des classes (la transformation des images en surface tridimensionnelle n'est qu'à but illustratif). De ces deux images, il est possible de construire une nouvelle image correspondant à l'espérance mathématique (figure 3.10c).

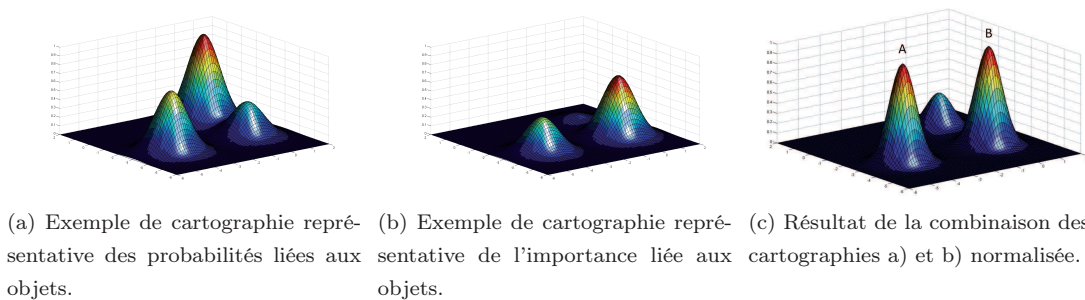


FIG. 3.10 – Illustration du principe de sélection d'une loterie.

Dès lors, en fragmentant l'image (figure 3.10c) en L morceaux contenant $|\mathcal{L}|$ pixels (comme en 3.7) et en évaluant chacune de ces loteries, il est possible de choisir la plus représentative de la zone d'intérêt du chirurgical.

Ce type de raisonnement ne permet pas de garantir l'unicité de \mathcal{L}^* ce qui peut mener, pour un système artificiel, à une impossibilité décisionnelle comme dans le cas présenté en figure 3.10c, où A et B ne sont pas discriminables et sont discordants par rapport au choix réel d'un utilisateur. Le lecteur pourra se rapporter à [AH79] ainsi qu'au paradoxe de "Saint-Pétersbourg" pour deux exemples de choix irrationnel.

En reprenant l'exemple (figure 3.10c), ce type de raisonnement aurait donc conduit à choisir le pic A plutôt que le B . En effet, bien que le gain associé à B s'avère plus grand que celui accordé à A , le degré plus important de certitude en A aurait rassuré l'individu qui l'aurait donc privilégié.

Pour combler cet écart entre choix optimal et choix réel, divers auteurs modifièrent (3.26) pour inclure cet aspect irrationnel. Ainsi Bernouilli introduisit le critère d'espérance morale qui fut la première apparition de la notion de risque tel que plus l'individu est fortuné et plus sa prise de risque est importante. Réciproquement moins l'individu a de biens et plus il sera tenté par un gain sûr. Ainsi (3.26) devient :

$$\mathbb{E}_{\mathcal{L}_q}^m = \sum_{i=1}^{|\mathcal{L}|} p(i) \log(g(i)) \quad (3.28)$$

Von Neuman et Morgenstern ([VNM53]) poursuivirent ces travaux et introduisirent la fonctionnelle dite notion d'utilité \mathfrak{U} (de classe \mathcal{C}^2) remplaçant les critères (3.26) et (3.28). Ce passage permet d'inclure la notion d'aversion au risque comme exprimée par Jensen sous la forme de l'inégalité suivante, avec X une variable aléatoire réelle ([Cai92]) :

$$\mathfrak{U}(\mathbb{E}(X)) > \mathbb{E}(\mathfrak{U}(X)) \quad (3.29)$$

Dès lors il est possible de poser certaines hypothèses sur le profil de la fonction d'utilité :

1. \mathfrak{U} est une fonction croissante, ce qui traduit la préférence d'un individu à choisir un gain élevé plutôt qu'un gain moindre ;
2. \mathfrak{U} est strictement concave ; les individus sont averses au risque.

Si l'hypothèse n°2 implique un comportement averse au risque qui se traduit par $\mathfrak{U}(x)'' < 0, \forall x \in X$, elle peut néanmoins être remplacée par l'hypothèse de neutralité face au risque ($\mathfrak{U}(x)'' = 0, \forall x \in X$) ou bien privilégier la prise de risque ($\mathfrak{U}(x)'' > 0, \forall x \in X$). Il peut donc exister une infinité de fonctions d'utilité, dont certaines trahissent des comportements spécifiques selon que le gain déjà acquis est élevé ou non (telles que les fonctions logarithmique, exponentielle, puissance et quadratique). Nous retiendrons seulement l'expression générale définie par :

$$f_{\mathcal{L}_q} = \sum_{i=1}^{|\mathcal{L}|} p(i) \mathfrak{U}(g(i)) = \sum_{i=1}^{|\mathcal{L}|} \text{BetP}(i) \mathfrak{U}(g(i)) \quad (3.30)$$

telle que la loterie optimale est définie par :

$$\mathcal{L}^* = \arg \max_{\mathcal{L}_q \in \mathcal{L}} f_{\mathcal{L}_q} \quad (3.31)$$

La prise de décision en environnement risqué apparaît donc plus aboutie que celle en environnement incertain grâce à la prise en compte de l'incertitude matérialisée par BetP .

3.2.2 Désignation de la zone d'intérêt

Nous venons de voir comment déterminer la loterie la plus représentative de la zone d'intérêt du chirurgien. Cependant, le processus de désignation mis en place ne permet pas de distinguer les interactions qui existent entre plusieurs classes mais seulement des loteries appartenant aux classes ayant le meilleur ratio importance-probabilité de bonne reconnaissance. Dès lors, il est possible de s'interroger sur la précision de la désignation compte-tenu que l'intégralité de certains objets est considérée comme la zone d'intérêt (figure 3.9a). Ramener alors les coordonnées de la zone d'intérêt à celles du centre de gravité de l'objet permet de se focaliser sur un point unique. Si cette technique permet de pallier la multiplicité des loteries sélectionnées, elle engendre un problème d'erreur de localisation de la zone d'intérêt. En effet, plus l'objet est grand et plus il y a de chances que la zone d'intérêt réelle ne soit pas aux coordonnées du centre de gravité de l'objet.

Une solution consiste à s'intéresser dans un premier temps aux interfaces entre les classes d'objets notamment celles entre la classe sang et la classe instrument. L'illustration 3.11b reprend l'exemple précédent où les loteries ont été limitées à celles satisfaisant ce type d'interaction. Puis dans un second temps, définir la zone d'intérêt au niveau du "centre de gravité" de ces loteries, comme le montre la figure 3.11c.

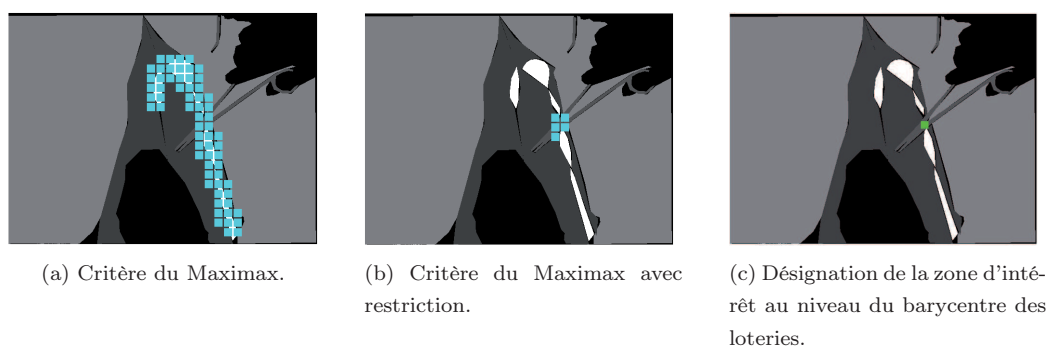
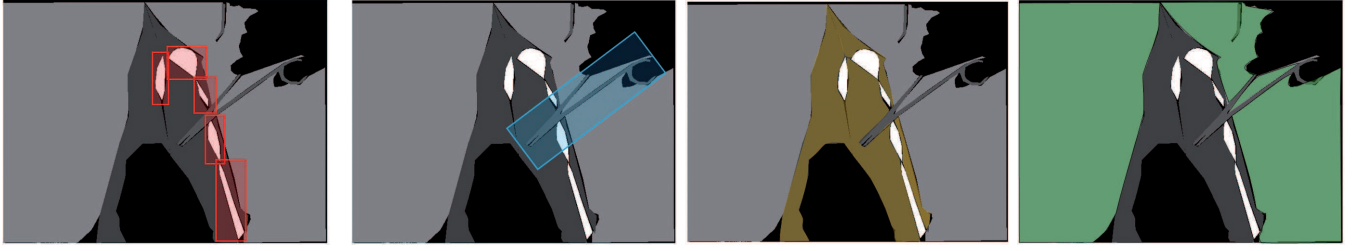


FIG. 3.11 – Sélection des loteries aux interfaces.

Pour mettre en pratique cette technique il est nécessaire de localiser les interfaces entre les classes d'objets. Cette étape pouvant être longue nous privilégions l'approche suivante :

1. recherche de la fenêtre englobante pour chaque classe d'objets (figures 3.12a et 3.12b) à l'exception de la classe champ stérile ;
2. superposition dans une image unique de l'ensemble des fenêtres ;
3. conservation des régions communes ;
4. élimination des régions communes à la classe champ stérile (régions en jaune figure 3.12e).

Les classes peau et champ stérile servent à limiter les régions où il est possible de désigner. Utiliser des fenêtres englobantes pour ce type de classe entraînerait une absence de recouvrement des classes instrument et sang. En conséquence, nous conservons les classes champ stérile et peau dans leur globalité (figure 3.12c et 3.12d). Cette technique de désignation de la zone d'intérêt s'avère très peu gourmande en temps de calcul, de nombreuses fonctions déjà optimisées pour ce type de traitement étant accessibles

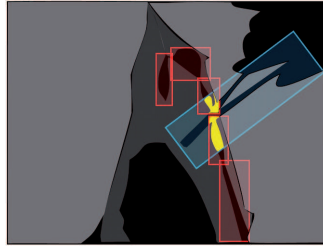


(a) Fenêtre englobante de la classe sang.

(b) Fenêtre englobante de la classe instrument.

(c) Classe peau.

(d) Classe champ opératoire



(e) Superposition de a) et b) avec élimination des éléments appartenant à c) et d).

FIG. 3.12 – Sélection des loteries aux interfaces.

dans la plupart des langages de programmation. Ainsi, pour obtenir les régions jaunes représentatives de la zone d'intérêt il suffit d'appliquer :

$$\mathcal{E}(i, j) = \left[\alpha_1 \frac{1}{2} I_{\text{sang}}(i, j) + \alpha_2 \frac{1}{2} I_{\text{instrument}}(i, j) \right] \mathbb{I}_{\text{champ opératoire}}(i, j) \cdot \mathbb{I}_{\text{peau}}(i, j)$$

$$I(i, j) = \begin{cases} 255 & \text{si } \mathcal{E}(i, j) = 255 \\ 0 & \text{sinon} \end{cases} \quad (3.32)$$

avec (i, j) les coordonnées du pixel considéré, I_{sang} et $I_{\text{instrument}}$ les images contenant les pixels reconnus respectivement comme du sang et des instruments tel que le niveau de gris est représentatif de $BetP$. Les termes α_i correspondent à l'utilité du gain de la classe $\mathcal{U}(g(i))$. La multiplication pixel à pixel de l'image avec α_i revient donc à $BetP(i)\mathcal{U}(g(i))$ intervenant dans (3.30). Enfin, \mathbb{I}_{θ_k} intervenant dans (3.32) correspond à l'indicatrice de la classe θ_k telle que :

$$\mathbb{I}_{\theta_k} = \begin{cases} 1 & \text{si } I_{\theta_k}(i, j) > 0 \\ 0 & \text{sinon} \end{cases}, \theta_k = \{\{\text{champ opératoire}\}, \{\text{peau}\}\} \quad (3.33)$$

Dans l'expression (3.32) nous nous servons des coefficients α_1 et α_2 pour moduler dynamiquement l'apport des classes d'objet dans la désignation de la zone d'intérêt afin d'augmenter la précision de la désignation de la zone d'intérêt. En effet, lors d'intervention en cavité le sang est localisé. Maximiser l'importance de cette information dans le processus de désignation est donc intéressante. En superficie, le sang est peu présent et souvent répandu à la surface de la plaie. Dans ces conditions, il s'agit de minimiser l'impact de la classe sang et d'augmenter celui de la classe instrument. Dès lors nous pouvons

considérer que $\alpha_2 = 1 - \alpha_1$. Nous verrons dans le chapitre 5, quel indicateur permet dans un cas concret d'évaluer ces pondérations et quelle est leur influence sur la précision de la désignation.

Dès lors, connaissant l'image $I(i, j)$ représentative des recouvrements des différentes classes, il suffit de calculer le "centre de gravité" commun à ces régions, pour désigner la zone d'intérêt du chirurgien.

Sélection d'un loterie unique

La démarche précédemment exposée, suppose qu'à chaque instant il n'existe qu'un seul lieu de recouvrement. Or, il est possible, en raison d'une mauvaise analyse de la scène, que plusieurs recouvrements distants existent. Comment dans ces conditions sélectionner une zone d'intérêt ?

Pour y répondre, nous ne gardons de ces recouvrements que leur centre de gravité que nous assimilons à des "mesures" de la position de la zone d'intérêt. Ces mesures ont une densité de probabilité p inconnue qui est représentative de l'activité du chirurgien. Déterminer p permet donc de désigner la zone d'intérêt. Une approche possible consiste à utiliser un estimateur de Rosenblatt dont voici le principe de fonctionnement : placés à des coordonnées aléatoires dans l'espace nous cherchons par estimation successive les coordonnées maximisant la densité de points pondérés par un terme K_0 .

$$\hat{p}_R = \frac{1}{nh} \sum_{i=1}^n K_0 \left(\frac{X_i - x}{h} \right) \quad (3.34)$$

avec :

- h , le paramètre de fenêtrage de l'estimateur ;
- K_0 , une fonction définie par : $K_0(u) = \frac{1}{2}\mathbb{I}(-1 < u \leq 1)$;
- \mathbb{I} , une fonction indicatrice qui prend la valeur 1 quand la condition est respectée et 0 dans le cas contraire ;
- n , le nombre de données.

Parzen étendit ces résultats, et introduisit la notion de noyau remplaçant K_0 d'origine, on parle alors d'estimateur à noyaux ou d'estimateur Parzen-Rosenblatt ([Par62]). Ces noyaux sont des fonctions devant satisfaire $\int_{-\infty}^{+\infty} K(u)du = 1$. Dès lors, il existe une infinité de formes possibles dont certaines sont présentées dans le tableau 3.7.

Les estimateurs à noyaux se sont vus principalement utilisés pour des techniques de suivi d'objet basées sur les histogrammes couleur. Bien qu'entachées par leur manque de discernement face à des objets aux couleurs proches (ce qui justifiera l'utilisation des mesures de similarité pour la phase de suivi de fragment), de nombreuses études tentent de combler cet aspect en se focalisant soit sur le noyau utilisé [Art08], soit sur l'intégration de plusieurs histogrammes afin d'être plus robustes aux variations brutales et rapides de couleur [LLR10].

L'application de cette technique pour cette étude quel que soit le noyau utilisé est présentée en figure 3.13.

Noyau	$K(u)$	
-Rectangulaire	$K(u) = \frac{1}{2}\mathbb{I}(u \leq 1)$	[WW10]
-Triangulaire	$K(u) = (1 - u)\mathbb{I}(u \leq 1)$	[Hyn08]
-Epanechnikov(parabolique)	$K(u) = \frac{3}{4}(1 - u^2)\mathbb{I}(u \leq 1)$	[Yam11]
-Biweight	$K(u) = \frac{15}{16}(1 - u^2)^2\mathbb{I}(u \leq 1)$	[GP11]
-Gaussien	$K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right)$	[RW10]

TAB. 3.7 – Liste des principaux noyaux utilisés.

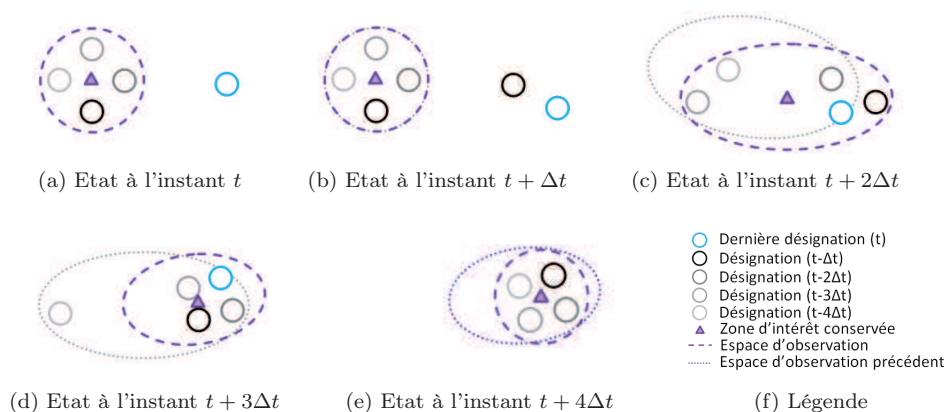


FIG. 3.13 – Illustration du principe de sélection de la zone d'intérêt.

Dans cet exemple, la désignation courante est en bleu et les anciennes en gris (noir pour la plus récente et gris pâle pour la plus ancienne). Cette nouvelle désignation est éloignée de ses prédécesseurs, il peut donc s'agir d'une désignation erronée ou bien d'un changement d'activité du chirurgien. En l'état actuel des connaissances il n'est pas possible de trancher. Sur ce schéma, le triangle violet symbolise la zone d'intérêt conservée qui servira de référence pour le déplacement. Cependant, dans le cas d'un changement d'activité du chirurgien, il est nécessaire d'intégrer au plus vite cette désignation et ne plus la considérer comme source d'erreur (réactivité du système). Ainsi, au fur et à mesure des désignations ce triangle va être amené à se déplacer pour finalement aller à l'endroit où les observations sont les plus nombreuses. Bien que non mentionnée dans cette explication, cette technique suppose la "mort" des observations les plus anciennes ceci afin de ne pas considérer des activités obsolètes. L'utilisation d'un estimateur à noyaux permet donc un comportement plus sûr, réfractaire aux variations brusques de position de la zone d'intérêt.

Remarque : Nous verrons dans la partie consacrée aux résultats que l'adjonction de cette étape influence la réactivité du système et parfois même, pose problème dans les cas où le chirurgien passe rapidement d'une zone d'intérêt à une autre. Cependant, le parti pris de cette étude a été de privilégier une absence ou un retard de déplacement à une erreur de déplacement. L'un des axes de recherche postérieur à cette étude serait donc d'améliorer ce point en intégrant l'aspect dynamique et non un estimateur purement statistique.

Une densité de probabilités obtenue grâce à un estimateur à noyaux (utilisation d'un noyau gaussien) est présentée (figure 3.14) pour un jeu de 55 lieux de recouvrements assimilés à 55 désignations (chaque lieu de recouvrement donnant un point unique considéré comme une désignation). Sur cette figure nous pouvons voir que les lieux de désignation isolés n'ont qu'une faible influence sur l'évaluation de la densité de probabilités (en bleu sur la figure). En effet, par définition, plus les désignations sont éparpillées et moins elles sont représentatives de la zone d'intérêt du chirurgien. En conséquence, une désignation isolée est considérée comme une erreur, alors qu'un groupement compact d'échantillons est fortement caractéristique de la zone d'intérêt réelle du chirurgien (en rouge sur la figure). Dès lors, la zone d'intérêt finale est obtenue par simple localisation du maxima de la distribution sur une période de temps défini; le fenêtrage des échantillons permet de ne plus considérer des désignations obsolètes.

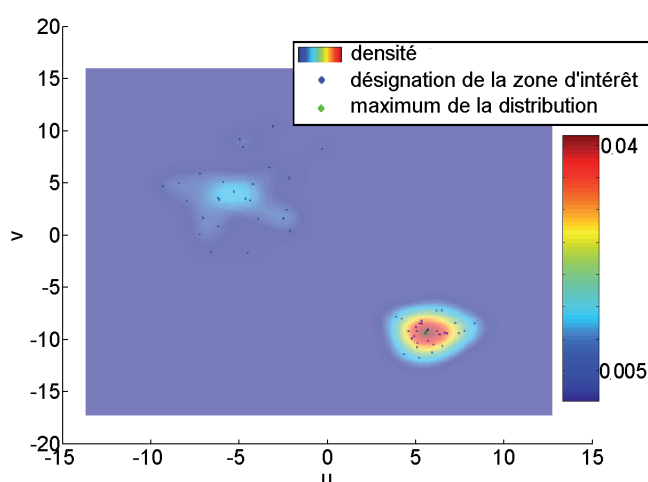


FIG. 3.14 – Densité obtenue par estimateur à noyaux pour un jeu de données simulées.

3.3 Conclusion

Nous avons présenté dans ce chapitre un processus de désignation de la zone d'intérêt du chirurgien basé sur une reconnaissance des objets en présence dans la scène chirurgicale à partir d'images provenant d'une caméra intégrée dans l'éclairage opératoire.

La première partie a porté sur des exemples de désignation manuelle. Nous avons pu voir que seule la reconnaissance simultanée du sang, de la peau, des champs stériles et des instruments permet une désignation fiable quel que soit le type d'intervention.

Pour pouvoir reconnaître ces objets, nous avons étudié les attributs forme, couleur et mouvement. Au terme de cette analyse, nous pouvons conclure qu'il apparaît nécessaire d'utiliser **conjointement** plusieurs attributs pour permettre une reconnaissance robuste. Cette mise en commun d'informations se fait grâce à la fusion de données.

Nous avons montré qu'il existe plusieurs cadres mathématiques. Nous retiendrons :

- le cadre probabiliste qui permet de manipuler les informations incertaines. Il s'avère inefficace face à des informations imprécises ;
- le cadre possibiliste qui traite des informations imprécises mais non conflictuelles ;
- le cadre évidentiel qui répond aux contraintes de l'information utilisée tout en permettant de futures évolutions.

Dans la mesure où les informations traitées sont incertaines, imprécises et conflictuelles, le support de fusion s'est orienté vers le cadre évidentiel. Chacune des règles de combinaison évoquées dans ce chapitre (étape 1 du processus de fusion, appelée étape crédale) seront employées lors de tests numériques (exposées en chapitre 4) qui nous permettront in fine de porter un choix sur l'une d'entre elles. Dans la suite de l'étude, nous prenons le parti d'utiliser les probabilités pignistiques comme critère de décision (étape 2 du processus de fusion, appelée étape pignistique) pour attribuer au pixel considéré une proposition du cadre de discernement. Afin de gagner en temps de calcul, nous choisissons de fragmenter ce cadre en quatre cadres distincts correspondant aux quatre classes d'objets sélectionnés.

Pour permettre de désigner la zone d'intérêt à partir de cette reconnaissance d'objets nous avons vu qu'il est possible de recourir à la théorie de la décision. En la matière, nous avons proposé deux environnements pour la prise de décision : l'environnement incertain et l'environnement risqué. Pour être opérants, ces deux environnements nécessitent d'attribuer aux classes d'objets un "gain" représentatif de leur intérêt pour le chirurgien. Nous choisissons de prendre l'instant opératoire et la situation opératoire ainsi qu'une connaissance a priori pour la définition de ce gain fictif. La différence entre ces deux environnements est que le choix en environnement incertain n'intègre pas de probabilité. La zone d'intérêt est donc désignée sur la classe d'objets ayant le plus haut degré d'importance, ce qui expose le système à de mauvaises désignations. L'intégration de la notion de risque par l'intermédiaire des probabilités permet de pallier ce point, on parle alors de choix en environnement risqué. En modulant l'importance des classes d'objets en fonction du gain déjà acquis, il est possible de faire ressortir des comportements spécifiques tels que l'aversion aux risques. Dans le cadre de cette étude, nous retiendrons ce dernier aspect pour le processus de désignation.

Nous avons également vu que pour permettre une désignation précise il est nécessaire de ne s'intéresser qu'aux interfaces d'objets. Nous avons montré comment sélectionner des loteries aux interfaces de manière simplifiée et comment éviter tout déplacement abusif du système à l'aide d'un estimateur à noyaux. Ce module finalise la partie consacrée à l'intelligence du système.

Il s'agit maintenant d'étudier la partie autonome chargée de placer la zone d'intérêt désignée au centre de la tache lumineuse.

Chapitre 4

Asservissement visuel d'un éclairage opératoire

Sommaire

4.1 Généralités sur l'asservissement visuel	102
4.1.1 Classification des asservissements visuels	102
4.1.2 Approches de commandes	107
4.1.3 Application de l'asservissement visuel en milieu médical	108
4.1.4 Les différentes étapes de mise en œuvre d'un AV2D	113
4.1.5 Calcul de la loi de commande	115
4.2 Problèmes de continuité de l'information	120
4.2.1 Les mesures de similarité	122
4.2.2 Prédiction de la localisation du fragment	132
4.3 Conclusion	139

Ce chapitre est consacré à l'autonomie du système. Après un rappel sur l'asservissement visuel en général, nous aborderons en détail l'architecture de l'asservissement visuel 2D considéré dans cette étude. La loi de commande sera exposée faisant le lien entre la localisation de la zone d'intérêt et le déplacement à effectuer pour centrer la tache lumineuse sur celle-ci. Nous verrons que pour avoir un mouvement continu du système, il est nécessaire de procéder au suivi de la zone d'intérêt durant le déplacement de la coupole.

4.1 Généralités sur l'asservissement visuel

L'asservissement visuel est né des travaux de [SI73] portant sur la prise en compte d'une information visuelle dans une boucle de commande. Il faut attendre les travaux de [HP79] pour voir apparaître l'appellation "*asservissement visuel*" (visual feedback). Initialement dévoué à des tâches simples de positionnements de bras robotisés, l'asservissement visuel est utilisé pour des applications dans les domaines aussi variés que l'aéronautique ([HS10] vol non habité), le médical (chirurgie du genou [CHC⁺10], alignement d'instruments suivant une trajectoire préétablie [SW10]), ou encore en robotique mobile ([CCO08]).

L'asservissement visuel consiste à contrôler à partir d'informations extraites d'une ou plusieurs caméra(s) un système articulé motorisé. Un exemple d'application consiste à déplacer l'effecteur d'un robot d'un point A à un point B défini au préalable. Pour ce faire, le système utilise exclusivement des informations provenant d'une ou plusieurs caméra(s). Ainsi, de la position courante (A) à la position désirée (B), le bras manipulateur doit se déplacer uniquement grâce aux informations extraites de l'image prise par une caméra.

4.1.1 Classification des asservissements visuels

4.1.1.1 Position et nombre de caméra(s)

Pour réaliser ce déplacement, différentes approches ont été proposées. Elles diffèrent selon le nombre de caméra(s) utilisée(s), leur position, le type d'informations visuelles à disposition, la connaissance a priori sur l'objet et son environnement. [KC02] propose de hiérarchiser ces approches selon le nombre de caméra(s) en action ainsi que leur position. On distingue alors des macro-groupes (figure 4.1).

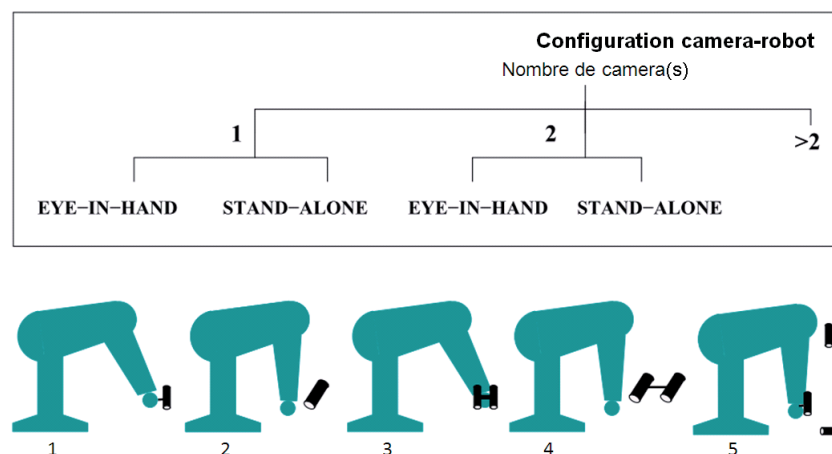


FIG. 4.1 – Classification des différentes approches en asservissement visuel (source : [KC02])

1. la caméra se trouve au bout de l'effecteur (Eye-in-hand) : on parle alors de caméra embarquée ;
2. la caméra observe l'effecteur (Stand-alone parfois Eye-to-hand) : il s'agit de caméra déportée ;
3. plusieurs caméras se trouvent au bout de l'effecteur ;
4. plusieurs caméras sont déportées ;

5. utilisation combinée de caméras embarquées et déportées (nous ajoutons à la figure 4.1 l'association de deux caméras, l'une embarquée l'autre déportée à l'image de [FCM00]).

Remarque : *Pièce centrale dans une tâche d'asservissement visuel, la caméra est un capteur constitué d'une optique et d'une matrice photosensible. Elle permet l'observation d'une scène à distance dans un domaine spectral donné, dans la plupart des cas le visible. Toutefois il est possible de trouver des caméras dans d'autres domaines tels que l'infrarouge ([CHYQNS04]), les rayons X ([OIA⁺11]) et γ ([XSK⁺10]), utilisées notamment pour la surveillance, le médical et le nucléaire. Le domaine spectral n'est pas le seul critère de différenciation, sont également présents la résolution, l'encombrement mais aussi la qualité de leur optique. Si l'optique permet d'atténuer l'astigmatisme (dû au mauvais accolement des lentilles), la courbure de champ, les aberrations chromatiques ainsi que les distorsions, elle joue également un rôle dans la capture de la scène. La caméra peut alors être omnidirectionnelle (caméra catadioptrique permettant une vue à 360°) grâce à l'utilisation de miroir plan ([HA01]), hyperbolique ([YYY93]) ou encore parabolique ([GN98]). Plus classiquement il s'agit de caméras perspectives (caméras dioptriques). Enfin elle peut être motorisée; il s'agit notamment des caméras dites PTZ (Pan Tilt Zoom [DJ08]) qui possèdent deux degrés de liberté.*

4.1.1.2 Asservissement direct ou indirect

Une autre manière de classifier les asservissements visuels consiste à regarder la distance "logicielle" séparant la loi de commande des moteurs : on parle alors d'asservissement visuel indirect ou direct.

L'asservissement visuel indirect repose sur l'emploi de contrôleurs bas-niveau chargés de convertir la consigne provenant d'une loi de commande définie en vitesse. Il est donc nécessaire d'avoir un cadencement des contrôleurs bas-niveau plus élevé que celui du calculateur chargé de la loi de commande et par extension plus élevé que la fréquence d'acquisition de la caméra. Ceci conditionne donc le choix des caméras à utiliser dans ce type d'asservissement, les caméras rapides étant à proscrire. Bien que ce point puisse constituer un élément rédhitoire selon les applications, l'indépendance de l'organe chargé du traitement d'images et de la loi de commande en fait une technique pertinente lorsque la définition du robot est sujette à débat.

A contrario, dans un asservissement visuel direct, l'état du robot est estimé au niveau du contrôleur vision. Charge à cet organe de gérer les moteurs afin de réaliser le mouvement souhaité. Cet aspect rend donc plus difficile le changement du robot dédié à la manipulation (passage par exemple d'un robot cylindrique RPP à un robot anthropomorphe 6R). Cependant, l'avantage majeur de cette approche est qu'elle permet l'emploi de caméras vingt à quarante fois plus rapides ([GDM00], [NITM00]) qu'une caméra classique (cadencée à 25Hz) grâce à l'augmentation des capacités de calcul des ordinateurs (loi de Moore).

4.1.1.3 Critère de Sanderson et Weiss

Dernier critère de classification, celui de Sanderson et Weiss ([SW80]), repose sur le type d'information utilisé dans la boucle de commande. On distingue trois approches : 3D, 2D et hybride. Ces différents asservissements se différencient par leur mode de calcul de l'erreur qui s'effectue respectivement dans le repère cartésien, image ou mixte (bien que certains travaux proposent le passage d'un mode à un autre [GH08]).

Asservissement visuel 3D - AV3D

L'asservissement visuel 3D (Position Based Visual Servoing - PBVS) englobe les techniques utilisant dans leur boucle de commande la pose du robot. Celle-ci correspond à la position et à l'orientation de l'effecteur. L'objectif est donc de faire tendre la pose courante \hat{p} vers la pose désirée p^* connue au préalable (figure 4.2). Le calcul de l'erreur, intervenant dans la boucle de régulation, s'effectue dans le repère cartésien. L'estimation de \hat{p} est non triviale et donne lieu à deux principales approches :

- La première consiste à prendre plusieurs images simultanées sous différents points de vue (cas 3, 4 et 5 figure 4.1) et à appliquer le principe de géométrie épipolaire (stéréovision). Ceci permet une estimation de \hat{p} même en l'absence de connaissance sur la géométrie de la scène ;
- La seconde s'applique aux cas d'applications monoculaires. La détermination de la pose requiert alors le modèle géométrique (CAO) d'un ou plusieurs objet(s) observé(s) que ce soit par apprentissage ou par hypothèse. Cette technique repose sur l'extraction à chaque itération d'indices visuels que l'on appelle primitives visuelles. Ces primitives peuvent être "naturelles" (primitives intrinsèques à l'objet comme les coins) ou "rajoutées" (amers visuels collés sur un cœur par exemple). Il peut s'agir de points, de droites ([ECR93]), d'ellipses particuliers de l'objet d'intérêt ([Vin01]) ou encore les moments [TC05] ; le lecteur pourra se reporter à [MC05] pour une synthèse en la matière. Cependant, plus le nombre de paramètres nécessaires pour décrire la primitive est élevé et plus le temps de calcul pour identifier ces paramètres sera grand. Il faut donc mettre en regard la complexité des indices visuels identifiables dans la scène observée et la puissance de calcul accessible. Ceci explique donc pourquoi les points sont généralement privilégiés ([Cha90]). Il faut également s'assurer qu'au cours du déplacement un nombre suffisant de primitives soit vu par la caméra, afin de se prémunir des problèmes de non unicité de pose dus à l'insuffisance du nombre de points détectés.

L'effort de reconstruction à faire limite les applications à des objets de géométrie simple (le temps de calcul pour l'estimation de la pose augmentant de façon significative avec la complexité de l'objet). De plus, l'estimation de la pose reposant exclusivement sur la connaissance de la géométrie de l'objet d'intérêt, cette technique est donc vulnérable aux erreurs de modélisation de l'objet lui-même ainsi qu'aux bruits de mesure. Bien que l'asservissement visuel 3D ait pour avantage de déplacer l'effecteur selon des trajectoires réalisables, il ne garantit pas la non disparition d'une ou plusieurs primitives.

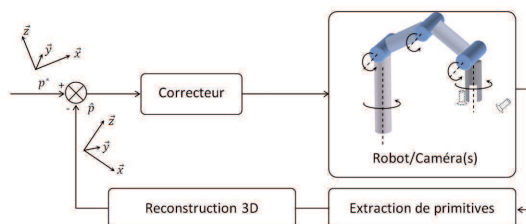


FIG. 4.2 – Principe de l'asservissement visuel 3D

Asservissement visuel 2D - AV2D

L'asservissement visuel 2D (Image Based Visual Servoing - IBVS) consiste à faire tendre une configuration de primitives courantes, $s(t)$ vers une configuration désirée s^* préalablement connue (figure 4.3). L'avantage d'une telle approche est l'absence de reconstruction géométrique même dans le cas d'applications monoculaires où il y a une perte de profondeur.

Cependant, en se basant sur un contrôle uniquement dans l'image sans regard de la position de l'effecteur dans l'espace, il est parfois possible que le robot tente d'atteindre des positions inappropriées (hors de son espace de travail par exemple). Mis en évidence par [Cha98a] le problème "d'avance/retrait" est un exemple de ces mauvais comportements. Lors d'une tâche d'asservissement, où quatre primitives de type point servent au guidage, il est demandé à un robot à six degrés de liberté (de type free-flying camera) d'exécuter une rotation de 180° autour de l'axe optique de la caméra. Le phénomène observé est un retrait vers "l'infini" suivant l'axe optique. Ceci s'explique par le fait que le déplacement des points dans le repère image de la caméra est restreint à des droites (en raison de la forme imposée de la dérivée de l'erreur).

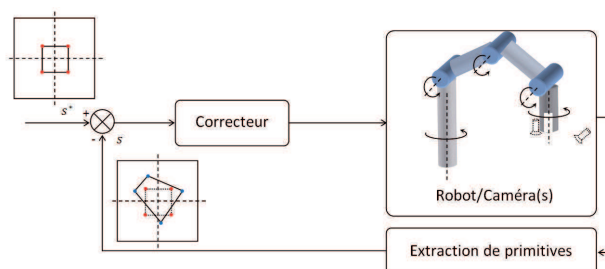


FIG. 4.3 – Principe de l'asservissement visuel 2D

Asservissement visuel 2D^{1/2} - AV2D^{1/2}

L'approche hybride, développée par [MCB98], utilise une consigne définie à la fois dans l'espace image de la caméra et dans l'espace cartésien, ce qui permet de pallier au problème "d'avance/retrait" possible dans l'AV2D grâce au découplage entre la rotation et la translation du robot. Ainsi, le contrôle en rotation est directement issu du calcul de l'homographie mettant en jeu l'appariement des points dans l'image, les translations sont quant à elles régies par l'erreur de configuration dans le plan image. En

s'affranchissant d'un modèle géométrique de la cible mais en conservant un contrôle dans l'espace image, il est possible d'éviter la fuite des primitives hors du champ de vision de la caméra possible en AV2D. Cette technique suppose la mise à disposition des informations à la fois 3D et 2D. Dans le cas où aucune

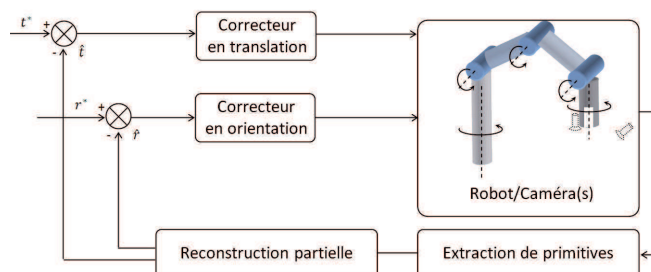


FIG. 4.4 – Principe de l'asservissement visuel $2D^{1/2}$

de ces informations n'est accessible, le recours à l'asservissement $d2D/dt$ est une dernière option.

Asservissement visuel AVd2D/dt

La technique d'asservissement visuel $d2D/dt$ utilise comme consigne un champ de vitesse désiré. Bien que la définition de la consigne se fasse dans le plan image, aucune extraction de primitives n'est nécessaire. Dans [SW86] et [CC01], les auteurs utilisent un champ de vitesse basé sur un modèle du mouvement apparent de l'objet limité à un plan pour leur loi de commande (4.1). Le modèle développé prend la forme d'une expression quadratique telle que les variations de position d'un point $(x, y)^T$ de la scène observée (appartenant au plan image π) sont exprimées par :

$$\begin{cases} \dot{x} = c_1 + a_1x + a_2y + a_3x^2 + a_4xy \\ \dot{y} = c_2 + a_5x + a_6y + a_7y^2 + a_8xy \end{cases} \quad (4.1)$$

La démarche consiste donc à identifier le jeu de paramètres a_i et c_i nécessaire à l'accomplissement de la tâche demandée. Cependant, pour être effective cette approche requiert un temps de calcul important (notamment si le modèle du mouvement est quadratique) et s'avère très sensible aux bruits de mesures.

En résumé :

Nous venons de voir les différents types d'asservissements visuels. Nous retiendrons que :

- *l'asservissement 3D effectue la régulation dans l'espace cartésien. Ceci passe par la planification et le suivi de trajectoire pour le déplacement du système robotique, ce qui permet d'assurer le succès de la tâche d'asservissement. En contrepartie, l'AV3D nécessite un modèle tridimensionnel de l'objet d'intérêt. La puissance de calcul nécessaire est proportionnelle à la complexité de ce modèle ce qui en limite les applications ;*
- *l'asservissement 2D effectue la régulation dans l'espace image grâce à des indices visuels appelés primitives. Ceci permet de ne pas avoir recours nécessairement à un modèle de l'objet. Le traitement d'images est donc crucial pour ce type d'asservissement. La planification d'une trajectoire n'est donc pas possible ;*

- l'asservissement $2D^{\frac{1}{2}}$, utilise à la fois les informations 3D et 2D. L'idée est conserver les avantages des deux méthodes tout en palliant leurs inconvénients. Ainsi translation et rotation sont découplées ce qui permet entre autres d'être robuste au problème "d'avance/retrait";
- l'asservissement $d2D/dt$ utilise la notion de champ de vitesse. Aucun modèle ou primitive n'est utilisé ce qui simplifie le traitement d'image. Toutefois, la difficulté d'accéder au modèle de mouvement de l'objet limite son application. Comme nous le verrons dans le chapitre dédié aux réalisations techniques, l'emploi d'une caméra unique nous limite à l'AV2D.

4.1.2 Approches de commandes

Nous venons d'explicitier précédemment les différents schémas de régulation dans le cadre d'asservissement visuel. Toutefois, il est possible de trouver au sein même de ces schémas différents types de commande permettant de prendre en compte les contraintes et la dynamique du robot.

Les contraintes sont généralement de deux natures : elles peuvent être articulaires (limite des actionneurs) auquel cas le robot ne peut atteindre certaines positions, ou bien dans l'image (contrainte de visibilité [Cha98b]). [MC00] propose une génération de trajectoire pour résoudre ce problème basée sur des fonctions de potentiels où le robot atteint la configuration désirée par le jeu de force entre un potentiel attractif le guidant vers la cible et un potentiel répulsif forçant les primitives à rester dans l'image. Grâce à l'intégration de ce terme lors de la génération de trajectoire (somme de toutes les forces), la méthode est robuste et la cible visible tout au long du déplacement. Une extension de ces travaux prenant en compte la saturation des actionneurs est proposée dans [MC02]. D'autres auteurs comme [DB06] (étendant les travaux de [EGS96]), utilisent une analyse et une synthèse multicritères de lois de commande. Celle-ci prend la forme d'une optimisation sous contraintes résolue grâce aux LMI (Linear Matrix Inequalities) ou programmation semie-définie (Semi-Definite Programming-SDP). Cette approche permet de compléter le bassin de convergence (où toutes les contraintes sont respectées) mais repose sur une linéarisation globale qui en limite les applications. Cette formulation sous forme d'un problème d'optimisation est commune à d'autres auteurs comme [MGdM02] qui n'utilise pour sa tâche d'asservissement aucun modèle a priori que ce soit pour le robot ou l'objet (l'interface homme-machine permet de spécifier la localisation des marqueurs à suivre). L'optimisation est réalisée grâce au simplex mais reste pénalisée par un temps de convergence trop important.

D'autres approches comme la commande prédictive ont été utilisées à des fins de suivi pour compenser les mouvements dus aux battements de cœur ([GDM00] poursuivi par [Gin03] et [Sau06]). Cette approche utilise une commande GPC (Generalized Predictive Controller) qui permet après apprentissage, de prédire la position de marqueurs posés sur à la surface du cœur (quatre primitives de type point) et d'anticiper leur mouvement dans l'image. Cette commande prédictive soulève cependant le problème de temps de calcul nécessaire à sa résolution. Toutefois, [ACT07] a montré qu'en utilisant la notion de platitude différentielle, le temps de calcul peut être divisé par deux. L'approche GPC introduite par [CMT87] peut être considérée comme une simplification de la commande LQG (Linear Quadratic Gaussian) qui cherche à minimiser une fonctionnelle. Celle-ci est représentative d'un coût de la forme : $x^T(t)Qx(t) + u^T(t)Ru(t)$, avec $u(t) = -Kx(t)$, $\dot{x}(t) = Ax(t) + Bu(t) + v(t)$ et $v(t)$ le bruit de mesure.

Le gain K est calculé, considérant l'horizon infini, à partir de l'équation algébrique de Riccati.

4.1.3 Application de l'asservissement visuel en milieu médical

Cette partie offre un tour d'horizon des différents asservissements visuels vus précédemment dans le domaine médical. Nous distinguerons entre autres les travaux sous imagerie X et IRM, les échographies interventionnelles et enfin ceux dédiés aux interventions mini-invasives (le lecteur pourra en trouver une très bonne synthèse dans [Kru09]). Le but est d'une part de montrer le champ applicatif de l'asservissement visuel dans ce milieu et d'autre part mettre en évidence les écarts avec notre étude et celles développées notamment en termes de connaissance a priori de l'objet d'intérêt.

4.1.3.1 Asservissement visuel sous imagerie X et IRM

Nous devons les premiers travaux d'asservissement visuel dans le domaine des rayons X à [LN00] et [NBL+00] pour le guidage autonome d'une aiguille jusqu'à un élément cible en vue d'une biopsie. Leur approche repose sur un renseignement explicite du système effectué par le chirurgien qui spécifie le point d'insertion ainsi que la localisation de la cible sur deux image prises arbitrairement. Dans ces travaux il n'est pas nécessaire de calibrer le système (utilisation d'un phantom ou de marqueurs) ni d'utiliser des capteurs supplémentaires autres que l'imageur X. L'asservissement visuel de type AV3D est réalisé à l'aide d'un robot à six degrés de liberté distant du patient. Le placement de l'aiguille suivant trois plans de coupes orthogonales conduit au bon positionnement du système. Une application similaire est proposée dans [MPB+04] où le robot est cette fois-ci posé directement sur le patient (éliminant ainsi les mouvements relatifs) et contrôlé à distance afin de limiter l'exposition du praticien aux radiations. Le robot (CT-bot) permet grâce à deux degrés de liberté d'orienter et d'insérer l'aiguille. L'asservissement visuel de ce robot est étudié dans [DMBdM06], où il s'agit d'éliminer les mouvements résiduels dus à la respiration. La mise en œuvre de cet asservissement repose sur la régulation dans l'image (image de coupe obtenue à l'aide du scanner), d'un ensemble de marqueurs placés sur le robot (AV2D).

Dans [BM08], le but est de positionner une table où se trouve un patient atteint d'une tumeur au cerveau, de telle sorte qu'il soit dans l'alignement d'un faisceau collimaté de proton. Pour placer correctement la table, deux images de coupes réalisées par rayons X sont comparées à une image de référence. Des marqueurs radio-opaques implantés sur la tête du patient permettent sur ces images de connaître l'erreur de placement. La loi de commande calculée permet de déterminer le mouvement à effectuer. D'autres auteurs encore comme [TGC+08] et [FACM09] permettent le guidage d'une capsule à travers le réseau veineux jusqu'à l'atteinte d'un élément cible. Munie d'une partie ferromagnétique, sa propulsion est générée par les variations de champ magnétique produit par un imageur IRM. La loi de commande à l'origine de ces fluctuations est basée sur le suivi d'une trajectoire de référence bidimensionnelle décomposée en segment. Les corrections de trajectoire résultent de la différence entre la position courante de la capsule et celle définie par la trajectoire de référence.

4.1.3.2 Asservissement visuel échographique

Les développements en matière de robotique télé-opérée et les propriétés de non intrusivité et de non destructivité de l'échographe sont à l'origine de l'essor de l'asservissement visuel dans ce domaine. Dans [ASZ+02], le praticien procède à un examen de la carotide, et manipule à distance la base d'un

robot de telle sorte à déplacer la sonde sur le cou du patient. L'asservissement visuel consiste à contrôler la tête supportant la sonde de manière à centrer la carotide dans l'image. La primitive utilisée est de type point et correspond au centre de la carotide. Toutefois, il est nécessaire pour contrôler les trois degrés de liberté de la sonde, de localiser le centre de deux artères. Ceux-ci sont détectés à l'aide d'une mesure de similarité qui recherche dans l'image les éléments s'apparentant à la carotide (connaissance a priori sous forme de fragment d'image). Les techniques utilisées sont de type "corrélation croisée" et SSD.

Poursuivant les travaux de [GGdM⁺04] sur la compensation des mouvements physiologiques de faibles amplitudes intégrant un modèle de contrôle prédictif couplé à un observateur adaptatif, [LKO⁺07] propose un asservissement visuel d'une sonde échographique à forte intensité dans le cadre d'une élimination d'un calcul rénal. L'asservissement est de type AV3D et se base sur une reconstruction de la pose du calcul rénal à l'aide de deux coupes échographiques perpendiculaires. Dans chacune de ces coupes, le calcul est détecté grâce à un algorithme de segmentation. La pose ainsi calculée est envoyée en consigne à l'entrée de la boucle de régulation afin d'agir sur les trois degrés de liberté du robot. L'élimination des mouvements relatifs est donc tributaire de la phase de calibration intervenant entre les deux sondes auxiliaires et la sonde principale, et nécessite un modèle représentatif du calcul ce qui pénalise la robustesse de l'approche.

Dans [VMO05], le but est le placement d'un instrument chirurgical dans le cadre d'une réparation de valve mitrale effectuée en mini-invasif. La technique développée repose sur une détection des extrémités de l'instrument (mords d'un forceps) dans l'image fournie par une sonde échographique trans-œsophagienne. La récupération du centre de gravité s'effectue par seuillage. Les primitives utilisées sont donc de type point, cependant pour augmenter la robustesse du système, les auteurs utilisent également le barycentre des extrémités ainsi que l'angle formé par ce point et le plan horizontal de l'image. Dès lors, connaissant l'image échographique désirée et celle courante, l'asservissement permet le déplacement de l'instrument par action sur un robot à quatre degrés de liberté. L'application de cette détection de la pointe des instruments par seuillage est également utilisée dans les travaux de [SPDC06] qui utilise cette information pour compenser les mouvements cardiaques. Cette étude se base sur une commande prédictive prenant en compte les contraintes de visibilité de l'instrument ainsi que les mécaniques du robot à six degrés de liberté : saturation des actionneurs, butées articulaires.

Poursuivant le même objectif que [VMO05], [NSDH07] s'appuie sur une détection de droite dans l'espace cartésien réexprimé dans le repère du robot, rendue possible grâce à l'ajout de marqueurs spécifiques sur l'instrument (notons que la transposition dans l'AV2D s'effectue à l'aide de la transformée de Hough ce qui ne nécessite aucun marqueur [HDH⁺04]). La pose de la cible est là aussi obtenue à l'aide d'un marqueur (en forme de croix), ce qui nuit à l'utilisation de cette méthode mais pose aussi le problème de calibration. La régulation se résume à un correcteur proportionnel dérivé prenant en entrée l'erreur entre la pose courante et celle désirée.

De manière générale, les asservissements visuels se basent soit sur des marqueurs rajoutés à la scène soit sur des caractéristiques de la cible comme la forme. L'approche développée par [KFH09] pour des activités de suivi d'organe, s'appuie sur le speckle obtenu dans l'image (interférence obtenue dans l'image s'apparentant à du bruit). Résultant des multiples interfaces tissulaires, des échos forment ce qui s'apparente à du bruit dans l'image. La sonde échographique n'effectuant pas une coupe parfaite mais légèrement

conique, il y a donc corrélation des plans parallèles. Retrouver des fragments speckle sur différents plans de coupe permet donc d'estimer le déplacement de l'organe et d'atteindre, connaissant l'image désirée, la position correcte (validée expérimentalement par des essais expérimentaux réalisés sur fantôme). Cette idée de travail en dehors du plan de l'effecteur, a été reprise dans [NK11]. L'asservissement repose sur trois coupes orthogonales de l'objet ainsi que sur un calcul des moments d'où sont extraits six paramètres géométriques. L'application de ces travaux est le suivi d'une coupe désirée, une extension intégrant la compensation des mouvements physiologiques périodiques est étudiée à l'aide d'un contrôleur prédictif (GPC) [NKG11].

Dans [MKC10], l'objectif est de déplacer la sonde échographique dans une configuration désirée préalablement définie. L'asservissement visuel est basé sur un calcul des moments sur la section observée de l'objet d'intérêt, permettant ainsi d'accéder à son aire, aux coordonnées de son centre de gravité, son orientation, ainsi qu'à la longueur de son axe principal. L'hypothèse utilisée dans ces travaux est que la forme non symétrique de l'objet est connue (modèle approché effectué avant l'intervention).

4.1.3.3 Asservissement visuel endoscopique

Comme nous l'avons mentionné dans le chapitre 1, les interventions chirurgicales à théâtre ouvert tendent à être supplantées par les opérations mini-invasives moins traumatisantes pour les tissus. En annexe B, nous présentons certains problèmes liés à ce type d'interventions et notamment le besoin d'avoir un assistant pour tenir la caméra endoscopique. Pour éviter ce recours systématique à un assistant, nombre de travaux proposent d'orienter la caméra sur la zone d'intérêt du chirurgien.

Pour [CM04], la technique employée (à l'image des travaux qui vont être présentés) est basée sur un asservissement visuel 2D reposant sur une recherche des éléments rectilignes dans l'image (monochromatique) préalablement filtrée (filtre gaussien). Il s'ensuit une sélection parmi les droites candidates. La discrimination s'effectue suivant certains critères tels que la longueur de la droite (qui doit être supérieure à un seuil prédéfini), l'intersection de l'instrument avec le cadre de l'image, l'orientation de l'instrument qui doit être dirigé vers le centre de l'image. Les essais sur une séquence de 128 images issues d'opérations montrent que le taux de bonne identification est de 77%. Ce taux avoisine les 99% lorsqu'un mécanisme de prédiction (linéaire) est utilisé pour estimer la position de l'instrument (en supposant la détection à l'instant précédent correcte et l'amplitude du mouvement faible). Ces résultats sont satisfaisants mais souffrent de plusieurs défauts dont notamment l'impossibilité de différencier plusieurs instruments simultanément, ce qui limite l'exploitation de cette méthode en conditions réelles.

Pour pallier ce point [Rou02] propose une approche basée sur le mouvement des instruments. En faisant la différence de deux images successives il est possible d'accéder uniquement aux éléments qui se sont déplacés. De là, une détection de contours permet de récupérer l'enveloppe de ces éléments que l'on modélise par des droites. Leur intersection définit la zone d'intérêt à centrer dans l'image. Les hypothèses selon lesquelles les instruments sont facilement différenciables (norme du gradient élevé lors de la détection de contours) et rectilignes apparaissent raisonnables. Cependant, bien que facilement implémentable, cette technique suppose qu'en dehors des instruments, aucun élément dans la scène n'est en mouvement. Pour éviter cette contrainte forte rarement respectée du fait de la proximité des organes de la caméra endoscopique, [WUW98], [MCF05] et plus récemment [AKN+11] utilisent l'information couleur pour

localiser les instruments (les auteurs se focalisent sur le suivi des instruments et non la désignation de la région à centrer dans l'image). Cette technique utilise un classifieur bayésien pour discriminer les régions de couleur (espace couleur RGB) semblables à celle d'un instrument dont on a au préalable construit l'histogramme. En la matière, quatre fragments d'image contenant l'instrument servent à l'élaboration de l'histogramme afin d'avoir la vue la plus complète de l'instrument. Ceci suppose donc qu'il n'y a pas de dénaturation de la couleur de l'instrument tout au long de l'intervention (comme des projections de sang par exemple) ainsi qu'une absence de modification des conditions d'éclairage. Il est possible de trouver des travaux mêlant ce type d'approche à ceux de [CM04] comme par exemple dans [SSS⁺08].

Pour [KGD⁺03], [VLC06] et [DNdM07] la technique employée consiste dans un premier temps à segmenter l'image (figure 4.5b). Puis ils adaptent un cylindre dégénéré aux contours des régions précédemment trouvées (figures 4.5c et 4.5d). En parcourant l'axe de l'instrument, il est possible de retrouver la pointe de l'instrument (figure 4.5e). Même si ce type d'approche permet la localisation de plusieurs instruments simultanément, elle repose sur des hypothèses fortes qui empêchent toute transposition vers les interventions à théâtre ouvert : absence de dégradation de l'apparence des instruments (projection de sang par exemple), formes des instruments constantes et non variées, présence d'éléments exogènes rares (compresse, prothèse), aucune occlusion de la zone d'intérêt, etc.

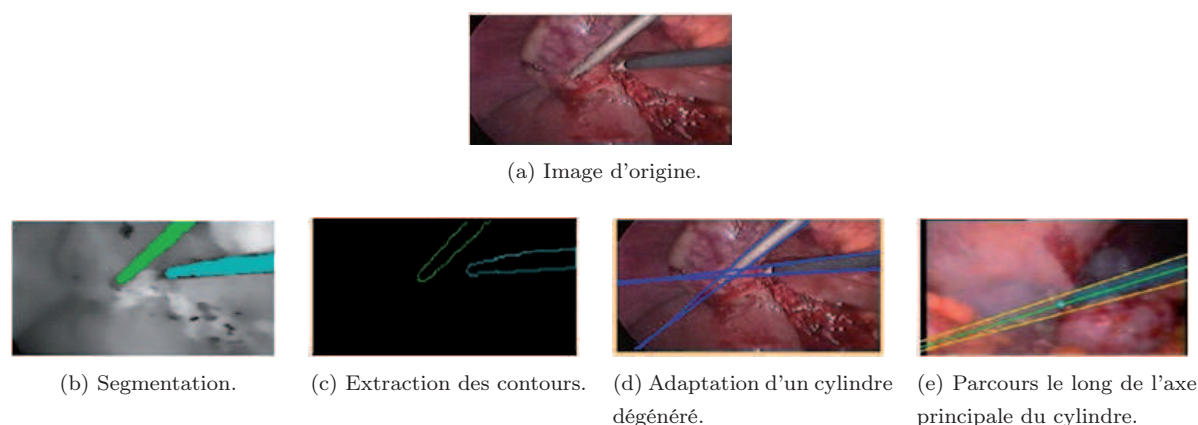


FIG. 4.5 – Détermination de la pointe des instruments (source [DNdM07]).

Une autre contrainte des interventions mini-invasives concerne les mouvements physiologiques qui vus de la caméra endoscopiques sont de très grandes amplitudes et peuvent conduire à la sortie de la zone d'intérêt de l'image. Pour y remédier, [BRL⁺07] propose une compensation par stabilisation active du myocarde. La structure élaborée est à liaisons compliantes ce qui permet une absence de jeu mécanique. Celle-ci est montée sur un robot sphérique à deux degrés de liberté, contrôlé par une loi de commande à rejet de perturbation (H_∞). Plusieurs schémas de commande ont été développés, basés sur différentes connaissances a priori du mouvement cardiaque à compenser (nous retiendrons notamment celle utilisant la commande prédictive). Dans la même thématique, [BBM10] utilise une commande basée sur un apprentissage des mouvements respiratoires, dans le cas d'une découpe du thorax. Un laser servant à la découpe est placé en bout d'effecteur, sa position est repérée à l'aide de marqueur. De même le thorax est rigidement lié à des marqueurs de même type. La position relative du robot est donc connue. La

répétitivité du mouvement respiratoire est prédite et compensée par action sur le robot.

D'autres auteurs comme [OZN⁺08] proposent une aide pour le cas d'interventions transluminales (qui diffèrent des chirurgies mini-invasives dans la mesure où celles-ci utilisent les voies naturelles). Dans ces travaux, il s'agit d'asservir visuellement la tête d'un endoscope flexible de telle sorte à compenser les mouvements physiologiques. En l'occurrence la technique utilisée repose sur un suivi d'une zone d'intérêt préalablement définie. Celle-ci est ensuite envoyée à l'entrée d'une boucle de régulation de type AV2D. La compensation du mouvement (périodique dont l'harmonique est à 3Hz) s'effectue à partir d'un modèle ARIMAX et à l'aide des correcteurs R-GPC (commande prédictive généralisée répétitive) et PRC (Prototype Repetitive Controller). L'adjonction d'un terme commutatif permet de rejeter les perturbations non périodiques.

En résumé :

Ces quelques exemples montrent les trois grandes thématiques de l'asservissement visuel dans le domaine médical. L'ensemble de ces travaux utilise soit des informations rajoutées, soit une connaissance forte de la forme ou l'apparence de l'objet d'intérêt. Notons néanmoins les travaux de [DNdM07] qui s'approchent de notre objectif par la localisation de la pointe des instruments. Toutefois ceux-ci ne peuvent être transposés à notre cas compte-tenu des écarts en matière de contraintes : la distance par rapport à la zone d'intérêt, les mouvements apparents dans l'image, les occultations, le type d'instrument utilisé, etc.

4.1.3.4 Application de l'asservissement visuel à un éclairage opératoire

Dans notre étude :

- le système de vision est une caméra unique. L'utilisation de plusieurs caméras est rejetée dans la mesure où aucune information supplémentaire facilitant la reconnaissance de la zone d'intérêt ne peut être extraite. En effet, la distance séparant les objets ne permet pas de localiser la zone d'intérêt et peut même dans certains cas amener à des résultats contre-intuitifs et de là conduire à effectuer des mouvements erronés (notamment dans le cas de la chirurgie du genou) ;
- la caméra est située au centre de la coupole. Ceci facilite la tâche d'asservissement et s'avère aisément intégrable. Il s'agit donc de la configuration "eye-in-hand" au sens de [KC02] ;
- la caméra choisie est une caméra perspective afin de ne voir que la scène opératoire sans être influencé par les éléments se situant autour de la table d'opération (activité de l'infirmière de bloc par exemple) comme cela serait le cas avec une caméra omnidirectionnelle ;
- la bande spectrale est dans le visible afin de faciliter la reconnaissance des objets en présence. Toutefois, le recours à une caméra infrarouge est envisageable mais complexifie grandement cette tâche car les matériaux composant les objets peuvent différer d'un bloc à l'autre ;
- un asservissement visuel indirect est considéré. Le souci d'évolutivité combiné à la rapidité de mise en application nous ont poussés à privilégier ce type d'asservissement. Ainsi l'unité chargée du traitement d'image est complètement dissociable de la partie commande, ce qui a pour conséquence

de restreindre le choix à des caméras basses fréquences (mais permettant d'apprécier les mouvements du chirurgien) ; cela n'est pas incompatible avec le processus de désignation de zone d'intérêt ;

- l'asservissement 2D est le choix le plus rationnel par rapport à notre application. Nous écartons l'asservissement 3D car il nécessite un modèle des objets en présence dans la scène. Un apprentissage (à réitérer à chaque itération) est toutefois possible mais s'avère très coûteux en temps de calcul. L'asservissement d2D/dt basé sur le champ de vitesse observé depuis la caméra est lui aussi écarté. A l'origine de ce rejet, la nécessité durant la phase de déplacement d'avoir une scène immobile. Or, cela n'est pas envisageable dans le cadre de notre application : il devient dès lors très complexe de calculer la bonne direction à prendre pour le système ;
- la primitive utilisée est de type point. Celle-ci correspond au maximum de la densité de probabilité définie à l'aide de l'estimateur à noyaux. Il ne s'agit donc pas d'un élément spécifique de la zone d'intérêt car par définition inconnue par avance ce qui implique qu'il n'est pas possible de déterminer les indices visuels devant être extraits par le traitement d'image.

4.1.4 Les différentes étapes de mise en œuvre d'un AV2D

Pour traduire une variation dans l'image en déplacement à effectuer, il faut comprendre comment est perçu un point de l'espace cartésien par la caméra. Pour ce faire, nous appliquerons les transformations suivantes :

1. passage du repère cartésien au repère lié à la caméra ;
2. définition du modèle de projection correspondant à la caméra utilisée et projection du point exprimé dans le repère de la caméra ;
3. définition de la métrique liée au plan de projection.

4.1.4.1 Etape 1 : le changement de repère

Soit P_o un point exprimé dans l'espace cartésien R_o et P_c le même point exprimé dans le repère caméra R_c . Ces deux points ont pour coordonnées homogènes :

$$\begin{cases} P_o = (X_o, Y_o, Z_o, 1)^T \\ P_c = (X_c, Y_c, Z_c, 1)^T \end{cases} \quad (4.2)$$

Ce passage de R_o à R_c n'est possible que grâce à la matrice de transformation homogène ${}^c\mathbf{M}_o$ telle que :

$$[X_c, Y_c, Z_c, 1]^T = {}^c\mathbf{M}_o [X_o, Y_o, Z_o, 1]^T \quad (4.3)$$

Pour des raisons de lisibilité il est courant de partitionner la matrice de transformation homogène en deux sous-matrices correspondant respectivement aux rotations ${}^c\mathbf{R}_o$ (matrice 3x3) et aux translations ${}^c\mathbf{t}_o$ (matrice 1x3) nécessaires pour passer de R_o à R_c :

$${}^c\mathbf{M}_o = \begin{pmatrix} {}^c\mathbf{R}_o & {}^c\mathbf{t}_o \\ \mathbf{0}_{1*3} & 1 \end{pmatrix} \quad (4.4)$$

L'équation (4.3) devient donc :

$$[X_c, Y_c, Z_c, 1]^T = {}^c\mathbf{R}_o [X_o, Y_o, Z_o, 1]^T + {}^c\mathbf{t}_o \quad (4.5)$$

4.1.4.2 Etape 2 : le modèle de projection

Il existe plusieurs types de projections possibles : orthographique (perspective axonométrique), paraperspective (approximation de la projection perspective), perspective, etc. Le modèle de projection choisi dans notre étude est la projection perspective car, outre le fait qu'elle soit largement utilisée pour les tâches d'asservissement visuel, elle est une généralisation des autres types de projections.

Nous souhaitons projeter le point P_c exprimé dans le repère R_c , dans le plan image π de la caméra se situant à une distance f appelée distance focale. Le point P_c aura pour projection le point P ayant pour coordonnées $(x, y, z)^T$.

$$P = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} f \frac{X_c}{Z_c} \\ f \frac{Y_c}{Z_c} \\ f \end{pmatrix} \quad (4.6)$$

4.1.4.3 Etape 3 : métrique du repère image

Estimer les coordonnées de P à partir de l'image n'est pas encore possible. Comme l'illustre l'équation (4.7), il existe une relation de passage entre le point p vu dans l'image et P :

$$P = \mathbf{K}p \Leftrightarrow \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \mathbf{K} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (4.7)$$

où le terme \mathbf{K} compile les paramètres intrinsèques ([Fau93]) :

$$\mathbf{K} \triangleq \begin{pmatrix} k_u f & k_u f \cos \theta & u_0 \\ 0 & k_v f / \sin \theta & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (4.8)$$

avec :

- (u_0, v_0) : pixel central du plan image (ou point principal) ;
- f : distance focale de la caméra ;
- k_u, k_v : facteur d'échelle suivant respectivement les axes \vec{u} et \vec{v} tel que $k_u = 1/d_u$ et $k_v = 1/d_v$ avec (d_u, d_v) les dimensions d'un pixel ;
- θ : angle que forment les vecteurs \vec{u} et \vec{v} du repère image correspondant au gauchissement des cellules photosensibles de la caméra.

Il est possible de simplifier l'équation (4.8). Dans le cas où nous supposons que la matrice CCD est composée de pixels parfaitement rectangulaires, c'est-à-dire sans gauchissement, nous obtenons :

$$\mathbf{K} \xrightarrow{\theta \rightarrow \frac{\pi}{2}} \begin{pmatrix} k_u f & 0 & u_0 \\ 0 & k_v f & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (4.9)$$

Cette nouvelle forme de la matrice \mathbf{K} peut être encore simplifiée si les pixels sont de taille $d_u = d_v = 1$:

$$\mathbf{K} \xrightarrow{k_i \rightarrow 1} \begin{pmatrix} f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (4.10)$$

Nous choisirons cette dernière expression comme matrice des paramètres intrinsèques de la caméra.

Nous venons de voir comment remonter au point d'origine à partir d'une mesure dans l'image. Sur ces bases nous allons maintenant aborder les techniques de guidage de systèmes robotiques à l'aide d'une mesure d'erreur dans l'image.

4.1.5 Calcul de la loi de commande

4.1.5.1 Notion de fonction de tâche

Pour une tâche d'asservissement visuel 2D, il est nécessaire de connaître la configuration courante des primitives $s(t)$ en fonction de $r(t)$, la pose de l'éclairage telle que :

$$s(t) = g(r(t)) \quad (4.11)$$

L'objectif est donc de faire tendre $s(t)$ vers une configuration désirée $s^*(t)$ connue au préalable. La réalisation de cet objectif conduit donc à la minimisation de l'erreur (ou fonction de tâche) définie dans l'image, notée $e(t)$, telle que :

$$e(t) = s(t) - s^*(t) \quad (4.12)$$

La caméra est donc bien dans la pose désirée si $e(t) = 0$. L'objectif de l'asservissement visuel est donc de trouver l'enchaînement des poses permettant d'atteindre s^* . Cependant résoudre $r^* = g^{-1}(s^*)$ pour définir le chemin optimal -méthode globale- pour le robot s'avère généralement coûteux en temps de calcul, ce qui exclut son utilisation en ligne. Il est donc plus simple de recourir à la méthode locale consistant en une approximation de la fonction g :

$$\dot{s}(t) = \begin{pmatrix} \dot{u}(t) \\ \dot{v}(t) \end{pmatrix} = \frac{\partial s(t)}{\partial r(t)} \frac{dr(t)}{dt} = L_s \tau(t) \quad (4.13)$$

où L_s est la matrice d'interaction de dimension 2×6 liant l'espace cartésien à celui de l'image ([HHC96]). Le second membre τ est le torseur cinématique quantifiant le mouvement à effectuer pour atteindre la pose désirée, exprimé dans le repère cartésien. τ est une composition de translations \mathbf{t} et de rotations telle que :

$$\tau = [\mathbf{t}, \mathbf{R}]^T, \text{ avec } \mathbf{t} = (t_x, t_y, t_z)^T \text{ et } \mathbf{R} = (R_x, R_y, R_z)^T \quad (4.14)$$

4.1.5.2 La matrice d'interaction L_s

Considérant des primitives est de type point et une caméra perspective, la matrice d'interaction L_s a pour expression (4.15). Le lecteur pourra se reporter aux travaux de [Cha90] pour une extension vers d'autres types de matrices d'interaction.

$$L_s(s, Z) = \begin{bmatrix} \frac{-1}{Z} & 0 & \frac{u}{Z} & uv & -(1+u^2) & v \\ 0 & \frac{-1}{Z} & \frac{v}{Z} & (1+v^2) & -uv & -u \end{bmatrix} \quad (4.15)$$

Il faut noter que cette forme n'est valable que dans le cas d'une seule primitive. Or, comme l'a montré [MR93] il est nécessaire, dans le cas d'un robot à six degrés de liberté, d'en avoir au minimum trois pour assurer l'unicité de la pose. En conséquence, nous concaténons k fois la matrice (4.15) de manière

à obtenir une matrice d'interaction (de dimension $(2k) \times 6$) englobant l'ensemble des primitives visuelles (4.16).

$$L_{s_{total}}(s, Z) = \begin{bmatrix} L_{s_1} \\ L_{s_2} \\ \vdots \\ L_{s_k} \end{bmatrix} \quad (4.16)$$

Dans cette expression L_{s_i} correspond à la matrice d'interaction de la $i^{\text{ème}}$ primitive. Par abus de langage nous utiliserons la notation $L_s(s, Z)$ pour parler de $L_{s_{total}}(s, Z)$.

Ce calcul de la matrice d'interaction est donc fonction de s et de Z . Si la détermination de s est aisée car elle revient à connaître les coordonnées $(u, v)^T$ de la zone d'intérêt, il n'en est pas de même pour la profondeur Z séparant la caméra du point dans l'espace cartésien. En effet, la projection dans l'espace de la caméra dénature la qualité des informations, ne permettant pas de revenir à l'appréciation de la profondeur. L'ajout de capteurs extéroceptifs est dans notre cas interdit, ce qui implique de devoir estimer Z_c et par conséquent L_s . L'évaluation de L_s noté \hat{L}_s peut s'effectuer suivant plusieurs modalités :

1. $\hat{L}_s \triangleq L_s(s, \hat{Z}_c)$. Cette approche consiste à estimer à chaque itération l'altitude Z_c permettant de se rapprocher au plus près de $L_s(s, Z_c)$. Les techniques employées en la matière sont basées sur le traitement d'images permettant la reconstruction 3D de la scène, ce qui suppose la mise à disposition de plusieurs caméras ou d'un modèle géométrique de la cible. Ce type d'estimation implique le déplacement rectiligne des primitives dans l'espace image, ce qui peut engendrer des mouvements inappropriés dans l'espace 3D ;
2. $\hat{L}_s \triangleq L_s(s^*, Z^*)$. La matrice d'interaction est choisie constante, égale à celle évaluée à la position désirée c'est-à-dire pour $s \equiv s^*$ et $Z \equiv Z^*$. L'avantage est le contrôle des mouvements dans l'espace 3D, à l'instar de la méthode précédente, mais cela invalide le caractère rectiligne du déplacement des primitives ;
3. $\hat{L}_s \triangleq \frac{1}{2}(L_s(s, \hat{Z}_c) + L_s(s^*, Z^*))$. Proposé par [Mal04], le calcul de la série de Taylor de la matrice d'interaction arrêté au second ordre permet le contrôle commun du robot dans les espaces 3D et image. Cependant, cette technique requiert comme pour $L_s(s^*, Z^*)$ la connaissance de la distance séparant la primitive de la caméra ;
4. D'autres auteurs proposent un calcul différent de la matrice d'interaction : citons par exemple un développement suivant une méthode dynamique quasi-Newtonienne [PML04] et une forme généralisée [MC08] permettant le passage de $L_s(s^*, Z^*)$ à $L_s(s, \hat{Z}_c)$ par ajustement d'un scalaire pondérant leur importance, ainsi qu'un apprentissage non pas de la matrice d'interaction comme dans [HA94] mais de son inverse [LJDC04].

Dans un cas général, au voisinage de la pose de référence les différents jacobiens image seront similaires en raison du rapprochement de l'altitude courante et de l'altitude désirée. Le choix de la technique d'estimation sera donc négligeable si la tâche du robot reste dévolue aux abords de la position de référence (car l'erreur d'estimation est moindre). A contrario, l'éloignement de la pose courante par rapport à la

pose désirée occasionnera de nettes disparités de comportement. Il est donc nécessaire de choisir judicieusement le type d'approximation en fonction des contraintes de l'application. Dans notre cas, l'élévation et l'abaissement ne sont pas motorisés ; nous estimerons donc constante la distance séparant la caméra de la scène opératoire (et par extension de la zone d'intérêt) ce qui correspond au cas n°1 c'est-à-dire où $\hat{L}_s \triangleq L_s(s, \hat{Z}_c)$. L'éclairage opératoire étant conçu pour une distance comprise entre 1m et 1.2m nous considérons $\hat{Z}_c=1.1\text{m}$.

4.1.5.3 Expression de la loi de commande

Nous cherchons à établir le lien entre les variations dans l'image et le déplacement à effectuer pour atteindre s^* . Pour ce faire, nous nous basons sur la fonction de tache vue précédemment qui s'écrit $e(r, t) = (s(r, t) - s^*)$ (équation 4.12). Pour simplifier l'écriture nous ne ferons plus mention des variables internes t et r correspondant respectivement au temps et à la pose de la caméra. Dès lors, nous avons sous condition que la cible est fixe :

$$\dot{e} = \dot{s} \quad (4.17)$$

Afin de garantir la rapidité ainsi que la convergence de l'erreur de poursuite, il est courant d'utiliser une décroissance exponentielle de l'erreur :

$$\dot{e} = -\lambda e \quad (4.18)$$

où λ est un scalaire permettant la paramétrisation de la vitesse de convergence ($\lambda \in \mathbb{R}^+$). La dérivation du terme \dot{e} permet d'écrire :

$$\dot{e} = \frac{de}{dt} = \frac{\partial e}{\partial r} \cdot \frac{\partial r}{\partial t} + \frac{\partial e}{\partial t} \Leftrightarrow \dot{e} = \frac{de}{dt} = L_s \cdot \tau(t) + \frac{\partial e}{\partial t} \quad (4.19)$$

En combinant (4.19) et (4.18), le torseur cinématique s'exprime :

$$\tau(t) = -\lambda \hat{L}_s^+ e - \hat{L}_s^+ \frac{\partial e}{\partial t} \quad (4.20)$$

Dans cette expression, \hat{L}_s^+ correspond à l'inverse de la matrice d'interaction. Cette opération n'est pas aisée. On distingue trois cas possibles :

1. $d = n$: la matrice d'interaction est carrée, l'inversion revient à faire $L_s^\dagger = L_s^{-1}$;
2. $d < n$: la matrice d'interaction n'est pas carrée, l'inversion se fait grâce à l'inverse généralisée. Il est d'usage en asservissement visuel de recourir à l'opérateur pseudo-inverse de Moore-Penrose noté \dagger . Son emploi est discuté par [LJDC04]. Lorsque L_s est de rang plein, la pseudo-inverse a pour expression : $L_s^\dagger = L_s^T (L_s L_s^T)^{-1}$;
3. $d > n$: la matrice n'est toujours pas carrée, l'usage de la pseudo-inverse est encore possible mais avec un léger ajustement (sous condition que L_s soit de rang plein) : $L_s^\dagger = (L_s^T L_s)^{-1} L_s^T$. Ce cas peut induire d'autres problématiques telles que des minima locaux.

Le terme $\frac{\partial e}{\partial t}$, intervenant dans l'équation 4.20, correspond au mouvement propre de l'objet dans l'image. Dans la mesure où il n'est pas possible d'accéder à sa valeur exacte, il est nécessaire de procéder à son estimation. Nous reformulons donc 4.20 comme suit :

$$\tau(t) = -\lambda \hat{L}_s^+ e - \hat{L}_s^+ \frac{\hat{\partial} e}{\partial t} \quad (4.21)$$

La non prise en compte du mouvement propre de la cible induit des erreurs de traînage. Pour y pallier il est nécessaire d'estimer $\frac{\partial e}{\partial t}$. L'équation (4.22) est un exemple d'estimation proposé par [BC95] qui consiste à soustraire la contribution du mouvement effectué par le système de l'erreur e obtenue dans l'image :

$$\left(\frac{\partial e}{\partial t}\right)_t = \frac{e_t - e_{(t-1)}}{\Delta t} - (\hat{L}_s)_t(\tau)_{t-1} \quad (4.22)$$

Dans la mesure où, dans notre cas, la zone d'intérêt n'est pas statique (le chirurgien n'arrête pas son geste) il apparaît intéressant de procéder à ce type d'estimation. Toutefois, compte-tenu que la zone d'intérêt évolue dans un espace limité et à vitesse faible, nous faisons l'hypothèse que $\frac{\partial e}{\partial t} = 0$. Ceci implique une réécriture de (4.21) :

$$\tau(t) = -\lambda \hat{L}_s^+ e \quad (4.23)$$

L'étude de la stabilité de ce système revient à considérer une fonction de Lyapunov V telle que :

- V doit être définie positive ;
 - \dot{V} doit être semi-définie négative (cas stable) ou définie négative (cas asymptotiquement stable).
- L'interprétation physique de cette propriété est que la dissipation continue d'énergie doit amener le système vers un point de stabilité.

Dans le cas d'une tâche de positionnement, la fonction candidate généralement considérée est ([DHKHC04]) :

$$V(t) = \frac{1}{2} \|e(t)\|^2 = \frac{1}{2} e(t)^T e(t) \quad (4.24)$$

La dérivée de V par rapport au temps s'écrit :

$$\dot{V}(t) = e(t)^T \dot{e}(t) \quad (4.25)$$

En remplaçant $\dot{e}(t)$ par $\dot{s}(t)$ (4.17) nous obtenons :

$$\dot{V}(t) = e(t)^T \dot{s}(t) = e(t)^T L_s \tau \quad (4.26)$$

Or, nous avons vu précédemment que τ peut s'exprimer sous la forme : $\tau = -\lambda \hat{L}_s^+ e(t)$, ce qui conduit à :

$$\dot{V}(t) = -\lambda e(t)^T L_s \hat{L}_s^+ e(t) \quad (4.27)$$

On cherche à montrer que le système adopte un comportement asymptotiquement stable et que l'erreur $e(t)$ est nulle uniquement lorsque $s(t) = s^*$. Il faut donc prouver que $\dot{V}(t) < 0$; la condition sur $V(t) > 0, \forall e(t) \neq 0$ étant déjà obtenue (4.24). Conformément à l'équation (4.27), ceci revient à vérifier $-\lambda e(t)^T L_s \hat{L}_s^+ e(t) < 0$ ($\forall e(t)$) autrement dit que $L_s \hat{L}_s^+ > 0$.

Nous pouvons isoler les trois cas suivants :

1. le nombre de primitives visuelles est inférieur à 3 : l'atteinte de la position désirée n'est pas possible ;
2. le nombre de primitives visuelles est égal à 3 : La matrice L_s est carrée, et inversible ($\dim(L_s) = (2 \times 3) \times 6$). Le système est globalement asymptotiquement stable car stable et globalement attractif. Le calcul de τ (4.20) conduit à l'atteinte de $e(t) = 0$ telle que $s = s^*$;

3. le nombre de primitives visuelles est supérieur à 3 : il y a existence de plusieurs minima locaux. Le système n'est pas globalement asymptotiquement stable mais uniquement stable localement. L'utilisation de quatre primitives pour un robot à six degrés de liberté est préconisée ([CA07]).

Remarque : Ce problème de minima locaux rencontré dans ce dernier cas n'affecte pas l'AV3D et l'AV2D^{1/2} qui combinent les deux approches. D'autres stratégies le permettent aussi comme celle dite hybride (ou à commutation) qui consiste à passer de l'AV2D à l'AV3D en fonction de leur erreur respective. Ceci permet d'augmenter le domaine de stabilité du système et d'éviter la sortie de primitives de l'image [GH07] (voir [KB08] pour une adaptation "matérielle" de cette stratégie). Une autre technique repose sur la notion de trajectoire du mouvement à exécuter (path planning). En la matière, il est possible d'utiliser des points de passage dans l'espace 3D constituant des étapes locales que le robot devra franchir pour atteindre la pose demandée. Ceci pose le problème de la traduction des informations tridimensionnelles en position articulaire car l'usage du modèle géométrique inverse permettant ce type de passage induit potentiellement des anomalies (conditions aux limites). Afin de ne pas passer par un modèle inverse, le recours au modèle géométrique direct est envisageable. L'avantage immédiat est l'évitement des singularités de ce modèle ; en contrepartie, l'aisance de travail apportée par la définition de la trajectoire dans l'espace 3D n'est plus de mise dans l'espace articulaire. Cette technique s'avère donc plus difficile à exploiter. Une dernière option repose sur la génération de la trajectoire dans l'espace image, on parle alors de trajectoire visuelle ([BMG97]).

Nous avons considéré jusqu'à présent que le robot servant à placer la zone d'intérêt au centre de la tâche lumineuse possédait six degrés de liberté. Or, comme nous le verrons dans le chapitre 5 consacré aux réalisations techniques, le robot que nous utilisons dans cette étude n'en possède que deux. Dans ces conditions $L_{s_{total}}$ n'est composé que de L_{s_1} correspondant à la zone d'intérêt qui a pour expression (4.28) ([CC00]).

$$L_s(s, Z) = \begin{bmatrix} uv & -(1+u^2) \\ (1+v^2) & -uv \end{bmatrix} \quad (4.28)$$

Compte-tenu des dimensions de L_s , une seule primitive suffit. L'atteinte de s^* est assurée sous condition qu'à chaque instant les coordonnées de $s(t)$ soient connues. Ceci suppose que le processus de la désignation puisse fournir à chaque itération la position de la zone d'intérêt.

En résumé :

Nous avons présenté dans cette partie nos hypothèses de travail ainsi que les différentes étapes de mise en œuvre d'un asservissement visuel 2D. Nous avons montré que la loi de commande utilisée permet la décroissance exponentielle de l'erreur et assure la stabilité du système. Nous avons mis en avant que la détermination du torseur cinématique correspondant au déplacement à effectuer, nécessite l'inversion de la matrice d'interaction qui peut s'avérer délicate. Cependant dans notre cas elle est simplifiée du fait de l'altitude constante de l'éclairage et de l'utilisation d'une seule primitive visuelle. Toutefois pour garantir le déplacement continu de l'éclairage il est nécessaire d'avoir à chaque itération la position de la zone d'intérêt.

4.2 Problèmes de continuité de l'information

Problématique : Le traitement utilisé pour la désignation de la zone d'intérêt repose sur la reconnaissance simultanée de plusieurs classes d'objet dans une même région de l'image. En l'absence d'une de ces informations (due aux erreurs de reconnaissance du traitement d'images par exemple), aucune zone d'intérêt ne peut être désignée. Ceci implique qu'aucune coordonnée ne peut être envoyée dans la boucle de régulation. Par conséquent, aucun torseur ne peut être calculé ce qui entraîne un arrêt du déplacement. S'il est aisé de combler ce manque d'information pour une itération (utilisation d'une mémoire : $\tau(t) = \tau(t - 1)$, $t \geq 1$), ce problème devient très délicat lorsque ce manque se prolonge (sortie potentielle de la tache lumineuse de la scène chirurgicale). De plus, l'arrêt répété du système ne permet pas de garantir au chirurgien des conditions d'éclairage optimales.

L'objectif de cette partie est la recherche de moyens permettant le déplacement continu de la coupole. Autrement dit, fournir à la boucle de régulation les coordonnées de la zone d'intérêt entre deux désignations.

Pour illustrer ce problème considérons l'exemple, présenté en figure 4.6, où pour un intervalle de temps donné aucune désignation de la zone d'intérêt n'est possible. Il s'agit donc de retrouver la zone d'intérêt entre les instants t_1 et t_2 . Une solution immédiate consiste à récupérer un fragment d'image centré sur les coordonnées de la zone d'intérêt définie à l'instant t_1 et le rechercher sur l'intervalle $[t_1, t_2]$. Retrouver ce fragment permet donc de retrouver la zone d'intérêt, combler l'absence de désignation et ainsi permettre au système d'effectuer un déplacement continu. Cette procédure est réitérée lors de toute cessation de désignation (comme par exemple entre les instants t_3 et t_4).

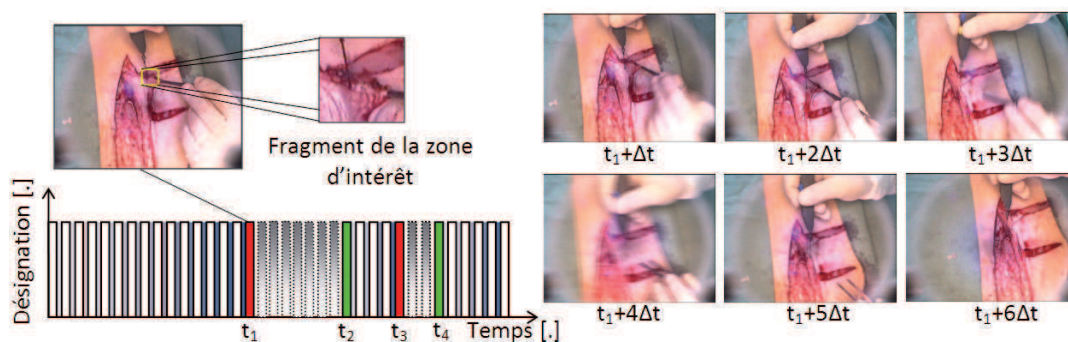


FIG. 4.6 – Extraction d'un fragment d'image représentatif de la zone d'intérêt.

Cette recherche de la zone d'intérêt soulève plusieurs problèmes dont notamment :

1. l'apparence du fragment qui ne peut être connue à l'avance. La zone d'intérêt ne correspond pas à un objet spécifique (ni même à une partie) mais résulte de la maximisation d'un critère, le fragment contient donc une proportion inconnue d'éléments biologiques et non-biologiques. La technique chargée de retrouver la zone d'intérêt ne doit donc pas s'attacher à retrouver un élément particulier ;

2. l'évolution de l'apparence de la zone d'intérêt en raison :
 - (a) du déplacement de la coupole. Lors de l'extraction du fragment (à l'instant t_1), la zone d'intérêt se trouve à distance du centre de la tache lumineuse (induisant une erreur $e(t_1) \neq 0$) donc dans une région sombre (comparativement au centre de la tache qui est au maximum d'éclairement). Le fragment est donc lui aussi sombre. A mesure que la tache lumineuse se rapproche de la zone d'intérêt, celle-ci va pâlir accentuant d'avantage la différence entre le fragment pris à l'instant t_1 (représentatif de la zone d'intérêt à cet instant) et la zone d'intérêt à l'instant $t > t_1$;
 - (b) du changement de point de vue sous lequel est observée la zone d'intérêt qui est amenée à se modifier à mesure que l'éclairage se rapproche ;
3. la modification de l'apparence de la zone d'intérêt du fait de l'activité du chirurgien qui poursuit son geste durant les phases de réorientation. Un exemple est présenté en figure 4.6 où le fragment extrait contient l'extrémité d'une pince (que tient le chirurgien dans sa main gauche) qui tend à sortir du champ de vue de la caméra dès l'instant $t_1 + 3\Delta t$. Dans un cas extrême, l'activité du chirurgien peut conduire à la perte de la zone d'intérêt telle que définie à l'instant t_1 ;
4. l'altération de l'image due aux mouvements de la coupole (comme par exemple sur l'image prise à l'instant $t_1 + 4\Delta t$ de la figure 4.6).

A cette liste il faut ajouter les difficultés dues à l'occultation partielle ou totale de la scène chirurgicale (par la tête du chirurgien par exemple), le déplacement de la zone d'intérêt du fait des mouvements physiologiques, le bruitage de la caméra qui perturbe les traitements d'images, etc.

Remarque : *Le changement de réorientation de la coupole va nécessairement entraîner une modification géométrique du fragment. L'illustration (figure 4.7) met en évidence que les transformations successives). Dans notre cas, le type de transformation appliquée au fragment est une homographie. Cependant, compte-tenu de la faible amplitude des réorientations nous supposons par la suite que celle-ci n'affecte pas le processus de localisation de la zone d'intérêt.*

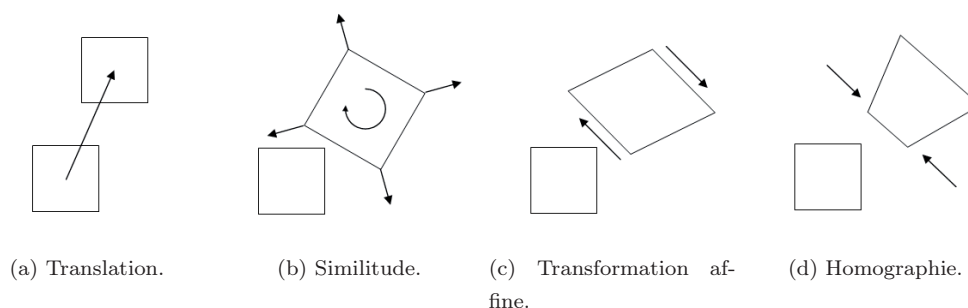


FIG. 4.7 – Illustration de transformations géométriques appliquées à un élément de référence.

Nous avons déjà vu en chapitre 2 certaines approches permettant de retrouver un fragment dans l'image. Les techniques basées sur les histogrammes comme SIFT ou SURF ne sont pas ici à utiliser en raison du

stockage d'informations requis généralement important. Celles utilisant la détection de points d'intérêt telle que la technique de Harris sont écartées compte-tenu de la perte potentielle d'une partie de la zone d'intérêt (troisième point de la liste). Enfin, celles reposant sur la reconnaissance d'un élément particulier sont elles aussi écartées car il n'est pas possible de garantir la pérennité de cette information (congruence du premier et troisième point de la liste).

Notre choix s'est porté sur une mesure de similarité qui permet de ne pas s'attacher à des éléments spécifiques du fragment, la globalité du fragment devenant un indice visuel à part entière.

4.2.1 Les mesures de similarité

Le principe de similarité repose sur la quantification des ressemblances entre un échantillon courant (pouvant être un histogramme H_e ou bien une image I_e) et celui de référence (H_{ref} ou I_{ref}). Parmi les mesures de similarité, on trouve la notion d'indice et de distance. Leur but commun est de fournir un scalaire proportionnel aux disparités entre les deux entités à tester. En réitérant ce calcul pour l'ensemble de l'image, il est possible de créer une carte de disparité. L'idée est alors de définir la zone d'intérêt (non désignée mais retrouvée) à l'endroit où la disparité est la moins grande.

Du fait de son incapacité à détecter les erreurs de localisation, le choix de la mesure n'est pas anodin et conditionne le bon comportement du système (figure 4.8). Sur cet exemple, une petite erreur de localisation conduit à moyen terme à un écart important avec la zone d'intérêt réelle. Il s'agit donc de sélectionner la mesure la plus adéquate. Nous nous baserons sur le tableau 4.1 regroupant les mesures les plus courantes. Le lecteur pourra se référer à [CC11] pour de plus amples détails concernant les familles de mesures (les croisées, les classiques, les dérivées, les non-paramétriques et les robustes). Compte-tenu de la criticité de l'application, l'intervalle de temps pendant lequel la coupole se déplace grâce à une mesure de similarité est limité à cinq images. Cette sécurité doit permettre de se préserver des dérives potentielles du système.

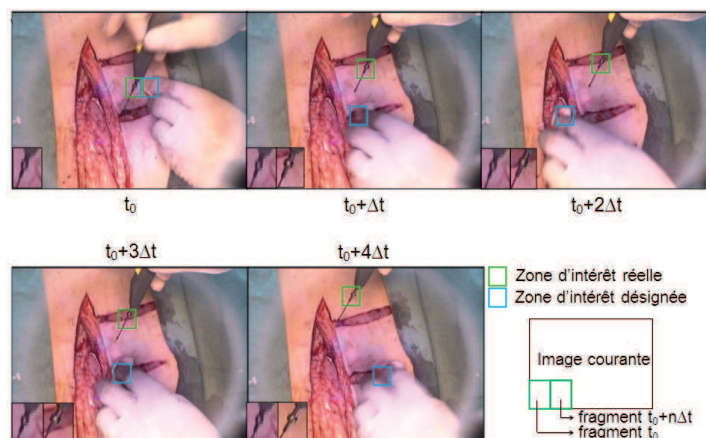


FIG. 4.8 – Erreur de désignation pour une procédure d'extraction de lambeau de peau entraînant un mauvais suivi.

Mesure	Expression	Référence
Chebychev ⁽¹⁾	$\max_{1 \leq i \leq \mathcal{R} } H_e(i) - H_{ref}(i) $	[Klo11]
χ^2	$\sum_i^{ \mathcal{R} } \frac{(H_{ref}(i) - H_e(i))^2}{H_{ref}(i) + H_e(i)}$ test statistique non-paramétrique	[DM11]
Euclidienne ⁽¹⁾	$\sqrt{\sum_i^{ \mathcal{R} } (H_e(i) - H_{ref}(i))^2}$	[LL09]
Jeffreys-Matusita	$\sqrt{\sum_i^{ \mathcal{R} } (\sqrt{\tilde{H}_e(i)} - \sqrt{\tilde{H}_{ref}(i)})^2}$	[HC03]
Jeffreys (divergence de)	$\sum_i^{ \mathcal{R} } (H_e(i) \log \frac{H_e(i)}{H_{ref}(i)} + H_{ref}(i) \log \frac{H_{ref}(i)}{H_e(i)})$	[YXY10]
Manhattan ⁽¹⁾	$\sum_i^{ \mathcal{R} } H_e(i) - H_{ref}(i) $ approche efficace lorsque la résolution est basse.	[WWL10] [WC04]
Minkowski	$\left(\sum_i^{ \mathcal{R} } (H_e(i) - H_{ref}(i))^p \right)^{\frac{1}{p}}$	[FWV05]
SQDIFF	$\sum_i^{\mathcal{D}_x} \sum_j^{\mathcal{D}_y} (I_{ref}(i, j) - I_e(i, j))^2$	[LYY+10]
SQDIFFN	$\frac{\sum_i^{\mathcal{D}_x} \sum_j^{\mathcal{D}_y} (I_{ref}(i, j) - I_e(i, j))^2}{\sqrt{\sum_i^{\mathcal{D}_x} \sum_j^{\mathcal{D}_y} I_{ref}(i, j)^2 \sum_i^{\mathcal{D}_x} \sum_j^{\mathcal{D}_y} I_e(i, j)^2}}$	[LYY+10]
CCORR	$\sum_i^{\mathcal{D}_x} \sum_j^{\mathcal{D}_y} I_{ref}(i, j) I_e(i, j)$	[LYY+10]

CCORRN	$\frac{\sum_i^{\mathcal{D}_x} \sum_j^{\mathcal{D}_y} I_{ref}(i, j) I_e(i, j)}{\sqrt{\sum_i^{\mathcal{D}_x} \sum_j^{\mathcal{D}_y} I_{ref}(i, j)^2 \sum_i^{\mathcal{D}_x} \sum_j^{\mathcal{D}_y} I_e(i, j)^2}}$	[Cor01]
CCOEF	$\sum_i^{\mathcal{D}_x} \sum_j^{\mathcal{D}_y} (I_{ref}(i, j) - \overline{I_{ref}}(i, j))(I_e(i, j) - \overline{I_e}(i, j))$	[LYY+10]
ZNCC	$\frac{\sum_i^{\mathcal{R}} (H_e(i) - \overline{H_e})(H_{ref}(i) - \overline{H_{ref}})}{\sqrt{\sum_i^{ \mathcal{L} } (H_e(i) - \overline{H_e})^2 (H_{ref}(i) - \overline{H_{ref}})^2}}$ (zero-mean normalized cross correlation coefficient)	[DKC11]

Notations :

- (1) : distance déclinée de la distance de Minkowski
 $p = 1$ Manhattan
 $p = 2$ Euclidienne
 $p \rightarrow \infty$ Chebychev

- H_e : Histogramme échantillon
 H_{ref} : Histogramme de référence
 \tilde{H}_i : Histogramme normalisé de H_i
 $|\mathcal{R}|$ cardinal de l'histogramme

TAB. 4.1 – Liste des principales mesures de similarité

Remarque : La divergence de Kullback-Leibler, sensible à la résolution des histogrammes (binning), ne correspond pas tout à fait à une mesure mais à une différence d'entropie entre deux distributions (divergence basée sur la théorie de l'information) [AGXX10].

Comme il a été fait mention, cette liste n'est pas exhaustive. Nous ne ferons que citer les distances de Kolmogorov-Smirnov (proche de la distance de Chebychev [FBFSB+98]), indice de Czekanowski ([TWT+96]), Canberra (complémentaire de l'indice de Bray-Curtis [Blo81]), la dissimilarité de Pearson et celle de Spearman ([YLA02]), la fonction de similarité cosinusale [BMS07].

Parmi les mesures de similarité que nous n'avons pas cité figure la mutuelle information. Son utilisation dans des travaux récents d'asservissements visuels nous incite à expliciter plus en détail cette technique, car sa robustesse aux changements d'éclairage importants est un point critique de notre application.

4.2.1.1 La mutuelle information

La mutuelle information découle des travaux de Shannon et a déjà été exploitée dans des applications d'asservissement visuel ([DM10]). Elle utilise comme outil principal la notion d'entropie telle que définie par l'équation (4.29) où $p(x)$ est la probabilité marginale liée à une variable aléatoire X .

$$H(X) = \sum_{x \in X} p(x) \log \frac{1}{p(x)} = - \sum_i p(x) \log p(x) \quad (4.29)$$

Ainsi l'équation (4.29) devient pour deux variables aléatoires X et Y :

$$H(X, Y) = - \sum_{x \in X} \sum_{y \in Y} p(x, y) \log p(x, y) \quad (4.30)$$

Deux variables aléatoires réelles X et Y seront dites semblables si la mesure $MI(X, Y)$ est forte, ce qui se traduit :

$$MI(X, Y) = H(X) + H(Y) - H(X, Y) \quad (4.31)$$

Il est possible de réécrire cette équation à l'aide de (4.29) et (4.30) sous la forme :

$$MI(X, Y) = \sum_{x \in X} \sum_{y \in Y} \log \frac{p(x, y)}{p(x)p(y)} \quad (4.32)$$

ce qui n'est pas sans rappeler la divergence de Kullback-Leibler (4.1). Cependant, déterminer les probabilités marginales ou jointes n'est pas direct lorsqu'il s'agit d'histogrammes. [DM10] propose un calcul en trois étapes :

1. normalisation des valeurs du fragment ;
2. utilisation de tricubiques B-splines pour calculer les histogrammes ;
3. somme sur l'ensemble du fragment des différences entre l'histogramme estimé et courant.

Si minimiser l'entropie jointe revient à maximiser la MI, donc à chercher le lieu où le fragment se trouve, ce type de technique souffre de faiblesses liées aux hypothèses utilisées qui sont : la constance de l'éclairage tout au long du mouvement, et l'obligation d'avoir une scène observée statique proche d'un comportement Lambertien (ce qui n'est pas notre cas dans la mesure où il n'y a pas cessation d'activité). D'autre part, la MI ne différencie pas les cas où les fragments ont une partie en commun, des cas où les deux fragments sont totalement superposés mais non identiques. Afin d'éviter ce type d'ambiguïté, [Stu97] introduisit la mutuelle information normalisée (NMI) telle que :

$$NMI(X, Y) = \frac{H(X) + H(Y)}{H(X, Y)} \quad (4.33)$$

Même si ces travaux montrent une amélioration notable de ce dernier point, l'"Entropy Correlation Coefficient" ([MCV⁺97]) et plus récemment la "Symmetric Uncertainty" ([MHA09] étudiée dans [Cah10]) lui sont souvent préférées. Le lecteur pourra se reporter à [PMV03] pour une revue des approches reposant sur l'information mutuelle.

La mutuelle information et ses dérivées semblent être appropriées pour cette étude. Toutefois, n'ayant pas de critère pour discriminer la mesure de similarité à privilégier, nous procéderons à ce choix lors des essais en bloc opératoire au regard des erreurs de localisation de la zone d'intérêt.

En résumé :

Dans certaines circonstances, la désignation de la zone d'intérêt ne peut être effectuée à chaque itération. Ceci altère la capacité du système à fournir une tache de lumière optimalement positionnée. Dès lors, il est nécessaire de pouvoir la localiser durant ces instants, pour permettre le déplacement continu de la coupole. Pour ce faire, nous nous basons sur une mesure de similarité prenant en compte un fragment d'image centré sur la zone d'intérêt. L'intégralité de ce fragment est alors considérée comme une primitive visuelle. Nous avons également vu qu'il existe un certain nombre de ces mesures. La mutuelle information déjà utilisée dans des tâches d'asservissement visuel et robuste aux changements d'éclairage est donc privilégiée. Cependant, ne pouvant discriminer ces mesures, nous nous appuyerons sur les résultats des essais en bloc opératoire pour juger laquelle se trouve être la plus adéquate.

4.2.1.2 Comparaison des mesures de similarité

Nous cherchons dans cette partie la mesure de similarité la plus robuste, dans la mesure où le suivi conditionne le comportement du système durant les phases où aucune désignation n'est possible. Leur comparaison nécessitant des espaces de définition communs, il a été privilégié de caractériser ces mesures sur un intervalle $[0, 1]$. De même, certaines de ces mesures étant caractéristiques de la dissimilarité, certaines conversions ont été appliquées afin de permettre leur comparaison.

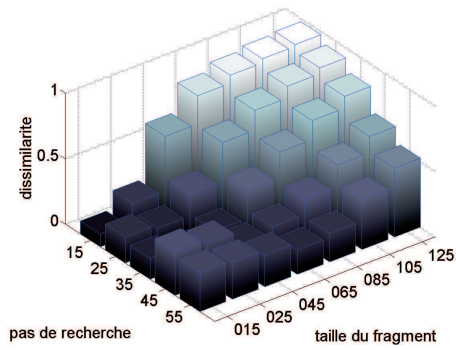
Le tableau 4.2 présente les résultats obtenus pour un fragment quelconque de taille 85×85 pour une image initiale de 600×450 . Le pas de recherche est de 125 pixels. L'évaluation de la performance d'une mesure s'effectue par multiplication de l'erreur moyenne par le temps d'exécution nécessaire pour cette mesure. La sélection de la mesure se base donc sur le minimum obtenu par ce calcul.

Mesure	Temps de calcul [ms]	$\frac{\text{Erreur moyenne}}{\min(\text{Erreur moyenne})}$	Score(★)
Euclidienne	47	2.793	16.23
Manhattan	63	3.311	25.79
Chebychev	47	1.256	7.30
χ^2	78	1.778	17.15
Jeffreys-Matusita	62	2.099	16.09
ZNCC	63	1.333	10.39
SQDIFF	47	1	5.81
SQDIFF NORMED	47	2.883	16.75
CCORR	62	1.968	15.09
CCORR NORMED	47	1.446	8.40
CCOEFF	47	1.973	11.46
Mutuelle information	243	1.163	34.96
Distance la plus adéquate :	SQDIFF		

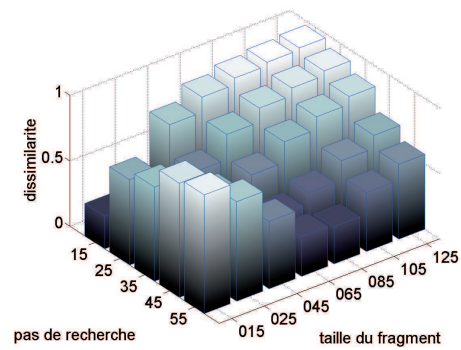
TAB. 4.2 – Simulation de mesure de similarité (★ : temps de calcul \times erreur moyenne).

A partir du tableau 4.2, nous pouvons déduire que la mesure la plus appropriée pour ces paramètres, est la mesure SQDIFF. Cette mesure se révèle à la fois la plus rapide à réaliser et celle qui donne le meilleur résultat.

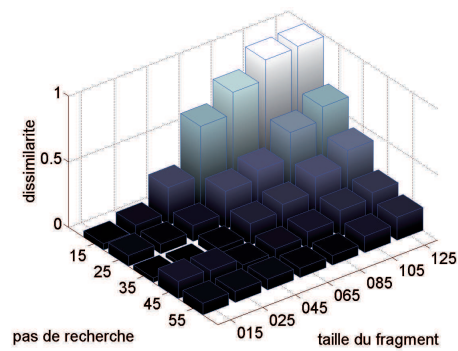
Toutefois, cette évaluation n'est pas objective car il s'agit d'un jeu de paramètres particulier. Nous nous intéressons donc à faire varier la taille du fragment et le pas de recherche puis à sélectionner la mesure qui présente l'erreur de localisation la plus basse (figures 4.9a à 4.10l).



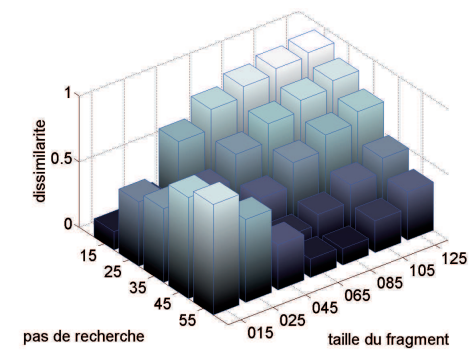
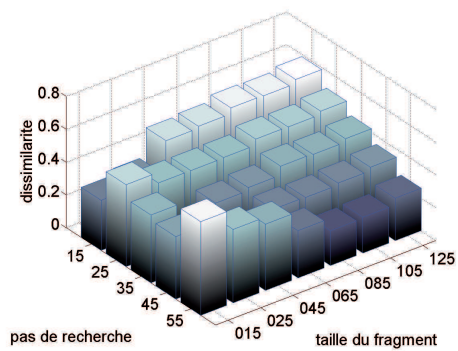
(a) Mesure euclidienne.



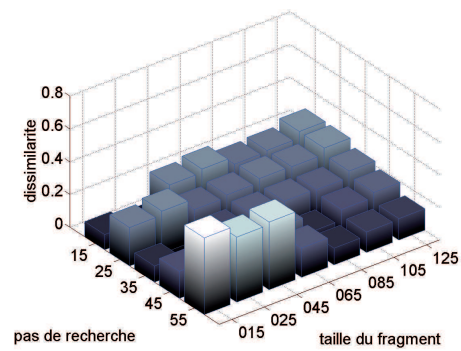
(b) Mesure de Manhattan.



(c) Mesure de Chebychev

(d) Mesure du χ^2 

(e) Mesure de Jeffreys-Matusita.



(f) Mesure ZNCC.

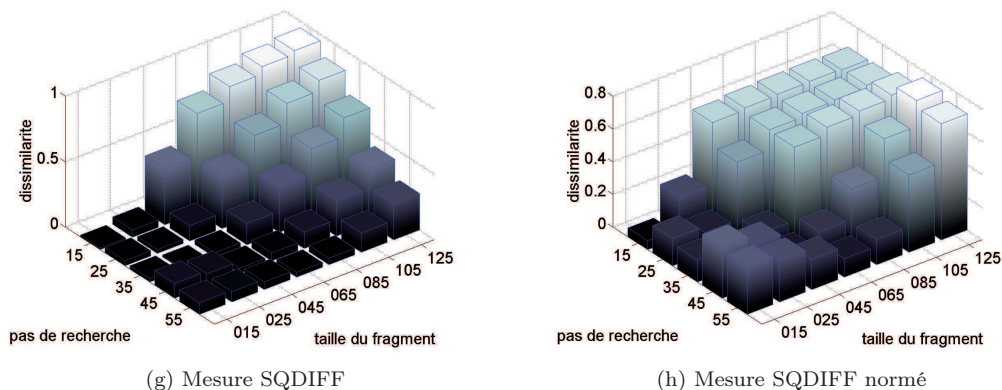


FIG. 4.9 – Erreur de similarité pour chacune des mesures.

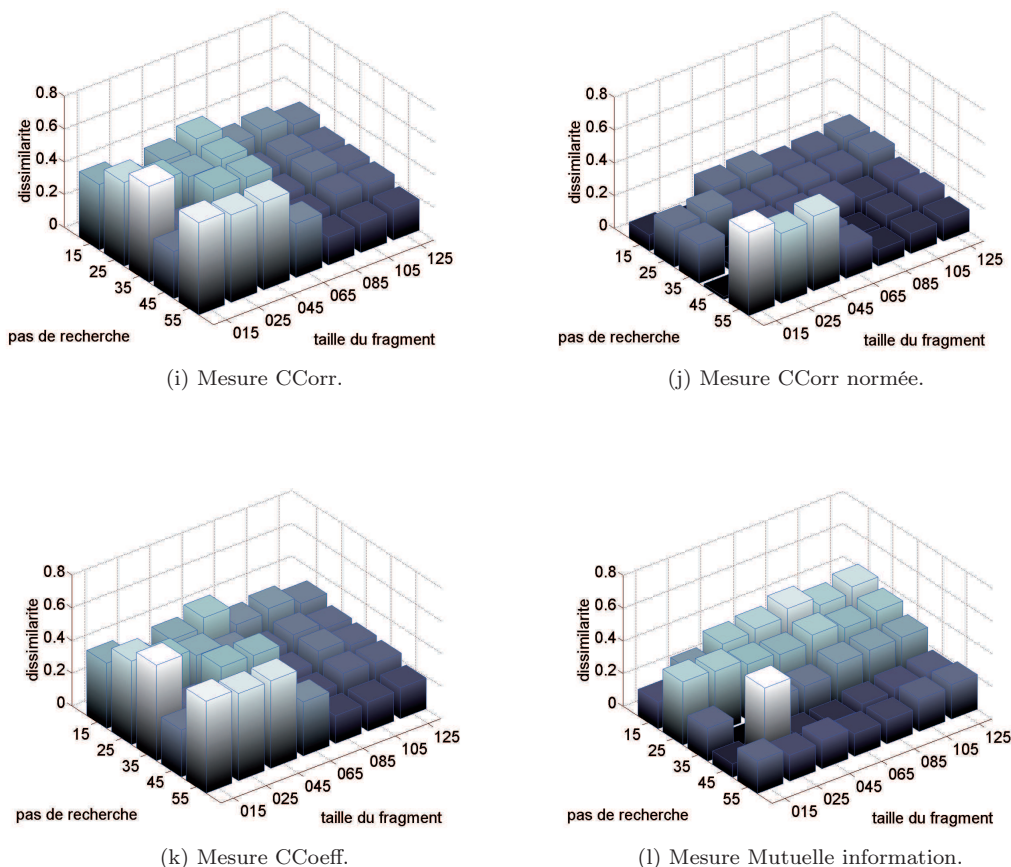


FIG. 4.10 – Erreur de similarité pour chacune des mesures.

Il en ressort de cette étude, que les mesures SQDIFF et CCorr normée sont les plus performantes. Ceci est visible sur la figure 4.11a qui présente pour chaque couple (taille de fragment, pas de recherche) la mesure qui a obtenu le meilleur résultat. Ces deux mesures à elles seules occupent 43% et 49% des couples. L'erreur minimale globale est obtenue à l'aide la mesure SQDIFF pour un jeu de paramètres de (35,25) comme le montre la figure 4.11 qui présente l'inverse de l'erreur (afin de faciliter la lecture).

Mesure	Sang	Peau	Champs stériles	Instruments	Ecart-type
Euclidienne	0.089	0.101	0.168	0.067	0.039
Manhattan	0.301	0.372	0.331	0.313	0.028
Chebychev	0.057	0.032	0.085	0.015	0.028
χ^2	0.179	0.243	0.193	0.18	0.026
Jeffreys	0.358	0.422	0.338	0.356	0.032
ZNCC	0.242	0.149	0.117	0.245	0.06
SQDIFF	0.018	0.012	0.033	0.005	0.011
SQDIFF normé	0.071	0.072	0.105	0.047	0.022
CCORR	0.39	0.176	0.15	0.395	0.146
CCORR normé	0.259	0.154	0.119	0.255	0.063
CCoeff	0.391	0.176	0.15	0.395	0.146
MI	0.24	0.196	0.484	0.114	0.138
Ecart-type max : 0.146		Ecart-type moyen : 0.061			

TAB. 4.3 – Influence de la classe observée sur l'erreur de similarité moyenne.

Ce jeu de paramètre permet un temps de calcul raisonnable (<45ms) tout en maximisant les chances d'obtenir un bon suivi de la localisation de la zone d'intérêt.

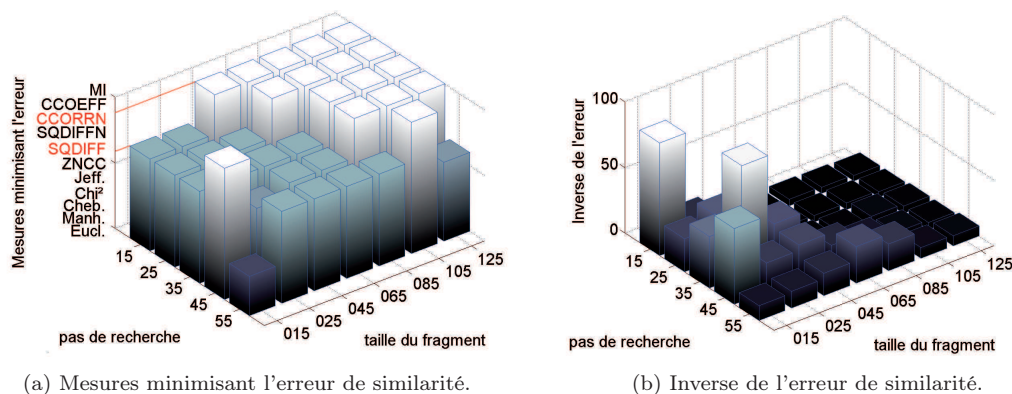


FIG. 4.11 – Résumé des résultats pour le choix de la mesure de similarité.

Nous cherchons maintenant à voir le comportement de la mesure SQDIFF lorsque la classe d'objet observé varie. Autrement dit l'objet contenu dans le fragment perturbe-t-il cette mesure de similarité ?

4.2.1.3 Perturbation du fragment

Pour répondre à cette question, considérons les cinq classes considérées dans cette étude, à savoir le sang, la peau, les champs stériles et les instruments. Le tableau 4.3 synthétise les écart-type obtenus pour différentes classes d'objets. A titre informatif les autres mesures de similarité sont également présentes dans ce tableau. Nous pouvons constater que de manière générale, ce type de perturbation n'est pas gênante, seules les mesures CCorr, CCoeff et la mutuelle information sont influencées.

Mesure	Aucun bruit	Faiblement bruité	Modérément bruité	Fortement bruité	Ecart-type
Euclidienne	0.073	0.077	0.083	0.08	0.04
Manhattan	0.349	0.382	0.442	0.417	0.041
Chebyshev	0.022	0.023	0.023	0.022	2.10^{-4}
χ^2	0.202	0.235	0.291	0.269	0.039
Jeffreys	0.382	0.421	0.47	0.454	0.039
ZNCC	0.171	0.168	0.088	0.141	0.038
SQDIFF	0.005	0.006	0.008	0.007	0.001
SQDIFF normé	0.049	0.051	0.06	0.055	0.005
CCORR	0.342	0.317	0.244	0.268	0.045
CCORR normé	0.23	0.224	0.159	0.197	0.032
CCoeff	0.342	0.317	0.244	0.268	0.045
MI	0.248	0.243	0.298	0.286	0.03
Ecart-type max : 0.045		Ecart-type moyen : 0.026			

TAB. 4.4 – Influence du bruit sur les mesures de similarité.

Comparer les méthodes pour des images sans défaut n'est pas représentatif des conditions d'exploitations rencontrées durant cette étude. Pour prendre en compte les perturbations d'origine matérielle (bruitage de l'image, entrelacement, résolution, ...) et/ou humain (occlusion partielle ou totale, dénaturation et/ou inclinaison de la zone d'intérêt, ...) pouvant survenir, nous proposons de bruitez artificiellement le fragment servant à la comparaison. Bruiter le fragment ou l'image courant est équivalent en termes d'erreur car la similarité est commutative. Dès lors, nous proposons trois niveaux de bruit basés sur une affectation aléatoire des valeurs d'un jeu de pixels sélectionnés aléatoirement parmi ceux du fragment. Nous distinguons trois cas, le premier où $\frac{1}{4}$ des pixels sont perturbés, puis $\frac{1}{2}$. Bruiter complètement le fragment revenant à désigner aléatoirement les zones d'intérêt, nous nous sommes limités à modifier $\frac{3}{4}$ des pixels du fragment. Des exemples de fragments sont proposés en figure 4.12.

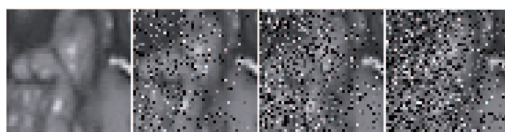


FIG. 4.12 – Exemple de fragment bruité.

Le tableau 4.4 montre les résultats les écart-type obtenus pour ce type de dégradation du fragment. Nous pouvons constater que le bruit est un paramètre peu perturbateur pour les mesures de similarité. Celui-ci peut donc être négligé. Nous savons que la taille du fragment, le pas de recherche, la classe observée ainsi que le bruit perturbent dans des proportions plus ou moins fortes la mesure de similarité. Si dans une application conventionnelle cette étude serait suffisante, elle s'avère ici incomplète. En effet, il nous faut considérer les perturbations liées au niveau d'éclairage. Le tableau 4.5 présente les résultats obtenus pour différents perturbations typiques de ces travaux : une exposition correcte de la scène chirurgicale, une sur- et une sous-exposition de la scène, la présence de halo (reflet dû filtre apodisant) et l'entrela-

Mesure	Halo	Sous-exposition	Sur-exposition	Entrelacement
Euclidienne	0.038	0.016	0.075	0.057
Manhattan	0.004	0.013	0.005	0.003
Chebyshev	0.039	0.013	0.071	0.026
χ^2	0.011	$> 10^{-4}$	0.005	0.012
Jeffreys	0.003	0.018	0.005	0.026
ZNCC	0.014	0.02	0.011	0.07
SQDIFF	0.026	0.017	0.038	0.014
SQDIFFNORM	0.044	0.035	0.07	0.052
CCORR	0.031	0.239	0.144	0.079
CCORRNORM	0.016	0.018	0.009	0.074
CCoEFF	0.031	0.239	0.144	0.063
MI	0.07	0.048	0.081	0.061

TAB. 4.5 – Influence des conditions d'éclairément sur l'erreur de similarité

cement de l'image en raison du type de transmission des données. Nous remarquons que les conditions d'éclairément n'impactent pas les mesures de similarité. La majorité conserve un comportement constant malgré les fortes variations apportées à l'échantillon.

En résumé :

Nous proposons en figure 4.13 un résumé de cette étude. Notre choix s'est porté, pour le suivi de la zone d'intérêt, sur la mesure SQDIFF, mesure la plus performante dans nos conditions d'essais. Nous avons testé cette mesure pour différentes perturbations : variation du contenu du fragment, parasitage du fragment et différents niveaux d'exposition du fragment. Cette mesure, outre le fait d'obtenir l'erreur de suivi la plus basse, s'avère être la plus robuste. Ceci valide donc notre sélection et justifie l'emploi d'une mesure de similarité pour le processus de suivi dans la mesure où cette technique est robuste face aux aléas.

4.2.1.4 Essais de suivi d'une zone d'intérêt

La mise en application de la mesure SQDIFF pour le suivi d'une zone d'intérêt, dans des conditions réelles, donne satisfaction comme le montrent les figures 4.14a et 4.14b. Celles-ci mettent en évidence la concordance entre la position réelle de la zone d'intérêt et celle estimée par la mesure de similarité. Malgré la durée importante pendant laquelle le suivi est exécuté (30 images), l'erreur d'appréciation reste faible comme le montre la figure 4.14. Quantitativement, l'erreur obtenue au sens de la norme L_2 est en moyenne $1.03\text{pixel} \pm 0.69$, nous pouvons donc en conclure que ce processus est fiable et adéquat pour l'étape de suivi. Sur ce graphique, nous pouvons voir un accroissement de l'erreur pour les images 5 et 6. Ce pic correspond dans cette séquence au phénomène d'entrelacement combiné à un mouvement rapide effectué par le chirurgien. Malgré cette difficulté, le comportement de la mesure reste fiable, l'erreur n'excédant pas 3pixels.

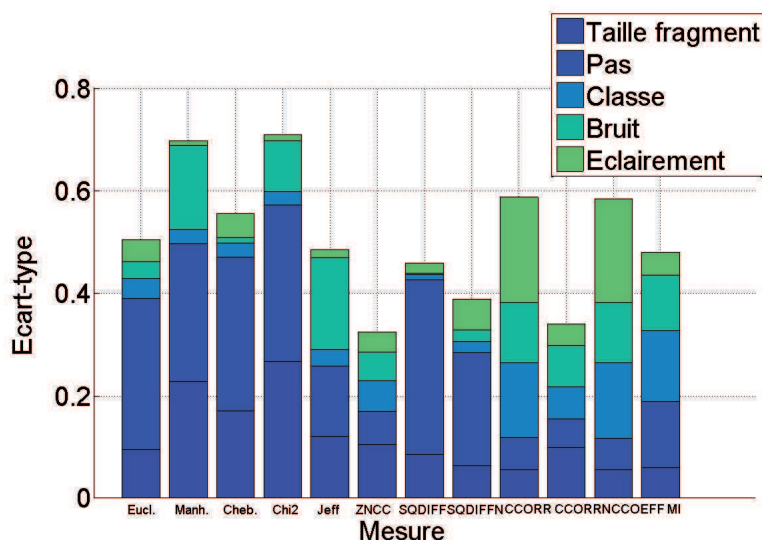
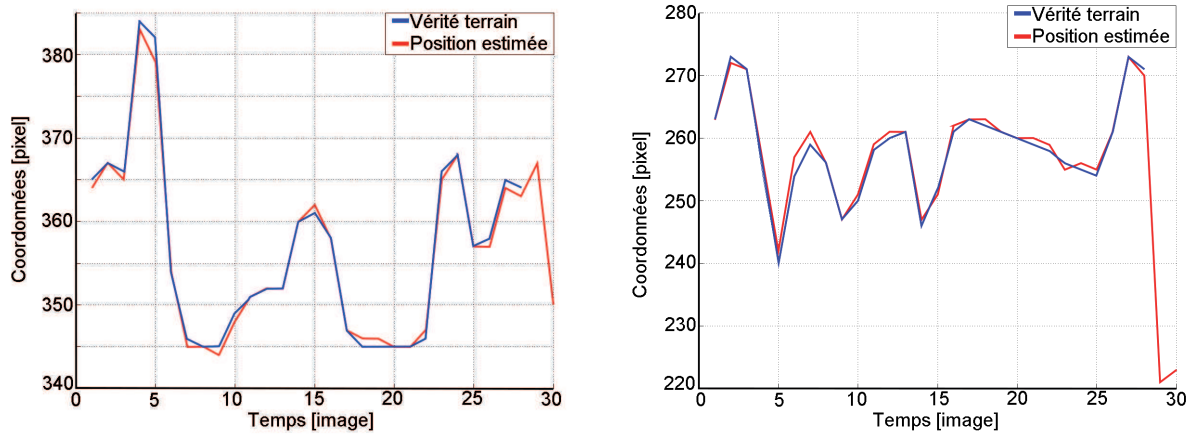


FIG. 4.13 – Influence des paramètres sur les mesures.

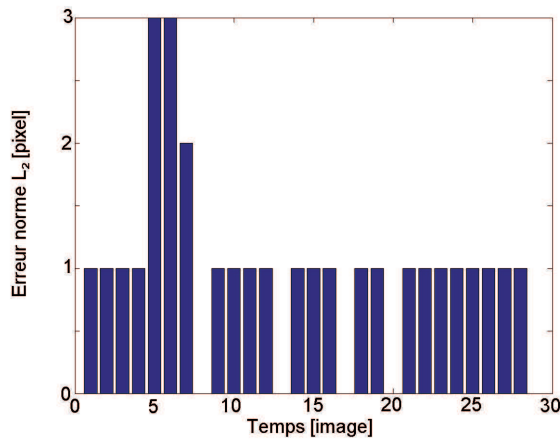
Cependant, cet exemple met en avant le problème de mauvaise traque de la zone d'intérêt lorsque celle-ci sort du champ de vue de la caméra. En effet, pour les images #28 et #29 aucune zone d'intérêt n'est visible ; ce qui explique l'arrêt à l'image #27 de la courbe bleue correspondant à la vérité terrain pour les figures 4.14a et 4.14b. Or, pour les deux images suivantes le processus de suivi fournit des coordonnées de la région s'apparentant le plus à la zone d'intérêt. Le système n'ayant aucun moyen de savoir si cette localisation est correcte ou non, celui-ci va se déplacer vers ces coordonnées et donc s'éloigner de la zone d'intérêt réelle. Il est donc de première importance de limiter le temps alloué à la phase de suivi. Ceci ne garantit cependant pas la conservation de la zone d'intérêt dans l'espace image. Il est donc intéressant de s'attacher à surveiller les coordonnées fournies à l'issue de la mesure de similarité. Dans la section suivante nous nous intéressons aux méthodes permettant de surveiller ce processus de suivi. Nous montrerons que cette étape supplémentaire permet d'éviter les mauvaises localisations et le cas échéant les problèmes de déplacements erronés.

4.2.2 Prédiction de la localisation du fragment

Réitérer cette recherche de la zone d'intérêt, définie lors de la dernière désignation, dans l'image courante ne garantit pas la robustesse du système. En effet, quelle que soit la mesure utilisée, la localisation de la zone d'intérêt correspond au maxima de la carte de similarité, il n'y a donc aucune relation entre cette position et celles définies aux instants précédents. Dans ces conditions, il est intéressant d'utiliser une stratégie couplée reposant sur deux briques indépendantes. La première repose sur l'estimation de la position du fragment retrouvé (entre l'instant courant et l'instant futur), ce type de démarche faisant l'objet de la section suivante. Quant à la seconde, elle consiste à utiliser un jeu de règles définissant les variations admises pour ce suivi (entre l'instant courant et le précédent).



(a) Ecart entre position réelle et position estimée suivant l'axe des x. (b) Ecart entre position réelle et position estimée suivant l'axe des y.



(c) Erreur de suivi.

FIG. 4.14 – Résumé des résultats pour le choix de la mesure de similarité.

4.2.2.1 Filtre de Kalman

Le filtre de Kalman-Bucy ([Kal60]) utilise une modélisation du système dynamique qui permet à partir de mesures ($y \in \mathbb{R}^m$) d'estimer tout ou partie des états du système ($x \in \mathbb{R}^p$) décrit par l'équation (4.34) en temps discret. Les matrices A , B sont déterministes et modélisent la dynamique du système, la commande du système est symbolisée par u , w et v matérialisent respectivement les bruits blancs non corrélés du système (de covariance Q avec $w \sim \mathcal{N}(0, Q)$) et de mesure (de covariance R avec $v \sim \mathcal{N}(0, R)$).

$$\begin{cases} x_k = Ax_{k-1} + Bu_{k-1} + Mw_{k-1} \\ y_k = Cx_k + v_k \end{cases} \quad (4.34)$$

En s'appuyant sur cette dernière expression, le filtre de Kalman-Bucy (KF) permet l'estimation de l'état courant x_k , notée \hat{x}_k . La récursivité proprement dite consiste à prédire l'estimation pour l'état courant \hat{x}_k dans un premier temps puis à la rectifier lorsque l'observation y_k est disponible. Ainsi la première étape correspondant à une prédiction de l'état peut se résumer à l'aide du système d'équation suivant

(avec $k \in [1, +\infty[$) :

$$\begin{cases} \hat{x}_k^- = Ax_{k-1} + Bu_{k-1} \\ P_k^- = AP_{k-1}A^T + MQM^T \end{cases} \quad (4.35)$$

La première équation du système (4.35) présente l'estimation a priori de l'état x_k (donc avant réactualisation). La seconde fait appel à la matrice P (covariance de l'erreur), qui dans sa version a priori est notée P_k^- . En l'occurrence, l'erreur est définie par : $e_k = x_k - \hat{x}_k$ (ce qui devient $e_k^- = x_k - \hat{x}_k^-$). Dès lors, $P_k^- = \mathbb{E}[(x_k - \hat{x}_k^-][x_k - \hat{x}_k^-]^T)$ autrement dit $P_k^- = \mathbb{E}(e_k^- e_k^{-T})$. L'idée sous-jacente du filtre de Kalman-Bucy est de minimiser cette covariance de l'erreur, ce qui est équivalent à $\hat{x}_k \rightarrow x_k$ donc à $P_k^- \rightarrow 0$.

Une fois cette prédiction connue, il est nécessaire de procéder à une étape de mise à jour (on parle parfois de recalage ou encore de réactualisation). Connaissant \hat{x}_k^- et P_k^- , elle consiste à les corriger en fonction des observations courantes. Dès lors, il ne s'agit plus d'état et de matrice a priori mais a posteriori, notés \hat{x}_k^+ et P_k^+ . Pour cela, nous utilisons un gain optimal K , appelé gain de Kalman (ou innovation), défini par l'équation (4.36) :

$$\begin{cases} \hat{x}_k^+ = \hat{x}_k^- + K_k(y_k - C\hat{x}_k^-) \\ K_k = P_k^- C^T (C P_k^- C^T + R)^{-1} \\ P_k^+ = (I_n - K_k C) P_k^- \end{cases} \quad (4.36)$$

On dispose donc d'une estimation sur l'état et d'un indice de confiance sur cette estimation (P_k^+). Bien qu'optimal pour des filtres linéaires, ce type de filtre ne peut être opérant si les bruits ne sont pas blancs gaussiens et si le système n'est pas linéaire. Dans ces cas il est possible de procéder à une linéarisation locale, on parle alors de filtre de Kalman étendu (EKF).

4.2.2.2 Filtre de Kalman étendu - EKF

Considérons un système non-linéaire tel que l'équation (4.34) s'écrive (avec $w_k \sim \mathcal{N}(0, Q)$, $v_k \sim \mathcal{N}(0, R)$ et $k \in [1, +\infty[$) :

$$\begin{cases} x_k = f_{k-1}(x_{k-1}, u_{k-1}, w_{k-1}) \\ y_k = g_k(x_k, v_k) \end{cases} \quad (4.37)$$

Comme précisé précédemment, l'EKF utilise une linéarisation du système, ce qui suppose de connaître le point de fonctionnement où il faut appliquer la linéarisation au point \hat{x}_{k-1}^+ . Dès lors linéariser le système revient à calculer F et G définis par l'équation (4.38) (f et g fonctions non-linéaires de \mathbb{R}^i dans \mathbb{R}^i supposés dérivables) :

$$\begin{aligned} F_{k-1} &\sim \left(\frac{\partial f_{k-1}}{\partial x} \right)_{\hat{x}_{k-1}^+} \\ G_k &\sim \left(\frac{\partial g_k}{\partial x} \right)_{\hat{x}_{k-1}^+} \end{aligned} \quad (4.38)$$

Ainsi l'étape de prédiction (4.35) devient :

$$\begin{cases} \hat{x}_k^- = f_{k-1}(\hat{x}_{k-1}^+, u_{k-1}, 0) \\ P_k^- = F_k P_{k-1}^+ F_k^T + MQM^T \end{cases} \quad (4.39)$$

L'équation (4.39) conduit au gain optimal $K_k = P_k^- G_k^T (G_k P_k^- G_k^T + R)^{-1}$, ainsi qu'à la matrice de covariance $P_k^+ = (I_n - K_k G_k) P_k^-$ et donc à l'estimation $\hat{x}_k^+ = \hat{x}_k^- + K_k(y_k - G_k \hat{x}_k^-)$. Tout comme pour le filtre de Kalman, l'initialisation de l'EKF est délicate ($\hat{x}_0^- = \bar{x}_0 = \mathbb{E}[x_0]$ et $P_0^- = Q$) et la définition de

Q et R peut poser problème, notamment dans les cas où l'humain est une partie du processus. Bien que cette version du filtre de Kalman moins restrictive permette un accroissement des domaines d'application, elle se trouve limitée là encore par l'hypothèse gaussienne.

4.2.2.3 Unscented Kalman Filter - UKF

Développé par [JU97], l'UKF (pour Unscented Kalman Filter traduit en français par filtre de Kalman sans biais) se caractérise par une absence de linéarisation des modèles. L'idée est d'approximer les distributions gaussiennes par un ensemble minimal de points (au nombre de $2n + 1$, n étant la dimension du vecteur d'état) appelés "sigma points" selon les auteurs, le calcul des Jacobiennes (4.38) étant non trivial. Ils modélisent la moyenne ainsi que la covariance de la distribution, qui viennent alimenter le système non-linéaire et par conséquent permettent la prédiction de la moyenne et la covariance de l'état quelles que soient les non-linéarités (preuve jusqu'au 2^{ème} ordre de la série de Taylor de G sans calcul de Jacobienne).

Le choix de ces "sigma points" est toutefois régi par des règles strictes dont les étapes clés sont les suivantes : considérons la matrice χ contenant l'ensemble de ces points pondérés par W tels que :

$$\begin{cases} \chi_{k-1}^0 = \bar{x}_{k-1}, & W_0 = \frac{\kappa}{n+\kappa} \\ \chi_{k-1}^i = \bar{x}_{k-1} + [(n+\kappa)P_{k-1}]_i^{\frac{1}{2}}, & W_i = \frac{1}{2(n+\kappa)} \quad \forall i \in [1, n] \\ \chi_{k-1}^i = \bar{x}_{k-1} - [(n+\kappa)P_{k-1}]_i^{\frac{1}{2}}, & W_i = \frac{1}{2(n+\kappa)} \quad \forall i \in [n+1, 2n] \end{cases}$$

Par convention d'écriture, les auteurs utilisent l'indice i pour indiquer la $i^{\text{ème}}$ ligne ou colonne de $[(n+\kappa)P_{k-1}]^{\frac{1}{2}}$, le terme κ conditionne quant à lui la répartition des sigma points. Nous procédons alors à l'étape de prédiction telle que :

$$\hat{x}_k^- = \sum_{i=0}^{2n} W_i f((\chi_k)_i)$$

Ce calcul n'est rendu possible que grâce à la connaissance de χ_k qui est une propagation des sigma points dans le modèle non-linéaire et traduite par : $\chi_k = f(\chi_{k-1}, u_{k-1}, k-1)$. Dès lors possédant χ_k et \hat{x}_k^- , il est possible de calculer la matrice de covariance a priori telle que :

$$P_k^- = \sum_{i=0}^{2n} [\chi_k^i - \hat{x}_k^-][\chi_k^i - \hat{x}_k^-]^T$$

Similaire à EKF, UKF utilise une étape de mise à jour visant à déterminer P_k^+ , K_k et \hat{x}_k^+ :

$$\begin{cases} K_k = P_k^x (P_k^y)^{-1} \\ \hat{x}_k^+ = \hat{x}_k^- + K_k (y_k - \sum_{i=0}^{2n} W_i g(\chi_k^i)) \\ P_k^+ = P_k^- - K_k P_k^y K_k^T \end{cases}$$

avec :

$$P_k^y = \sum_{i=0}^{2n} W_i [g(\chi_k^i) - \sum_{i=0}^{2n} W_i g(\chi_k^i)] [g(\chi_k^i) - \sum_{i=0}^{2n} W_i g(\chi_k^i)]^T + R_k \text{ et,}$$

$$P_k^x = \sum_{i=0}^{2n} W_i [\chi_k^i - \hat{x}_k^-] [g(\chi_k^i) - \sum_{i=0}^{2n} W_i g(\chi_k^i)]^T$$

Ce type de filtre s'avère plus simple à implémenter que la version EKF mais reste cependant comparable en termes de performances [CJ10]. Le lecteur pourra se reporter à [RC04], [JJ06], [XJ07] et [QH09] pour quelques exemples d'application.

Remarque : Ces trois déclinaisons du filtre de Kalman (KF, EKF, UKF) ne sont pas les seules existantes à ce jour; nous nous sommes focalisés sur ceux-ci car ils condensent les contributions majeures en la matière. Dans la mesure où notre besoin est d'anticiper les potentielles défaillances ou erreurs liées à la phase de suivi, la complexité du filtre nécessaire est donc faible c'est pourquoi nous nous concentrerons uniquement sur les filtres KF, EKF et UKF. Néanmoins notons l'existence des quelques travaux suivants : AEKF pour la réactualisation dynamique de la matrice de covariance liée à la mesure ([WGD09]), EnKF prémisses de l'UKF pour sa vision stochastique plutôt que déterministe de KF ([ZJGHHFL11]) amélioré en sa version DEnKF ([SB12]), SEEK et SEIK visant à réduire la complexité calculatoire de l'EKF ([PVR98]), rUKF permettant la réduction de l'ordre du filtre UKF ([XLL⁺11]), SRUKF pour l'augmentation de la robustesse et de la stabilité de l'UKF ([HLY10]).

Filtrage particulaire Nous avons vu dans la partie consacrée à l'UKF que pour modéliser une distribution, des points particuliers en nombre restreint sont utilisés (sigma points). Cependant, ce type de filtre se trouve limité lorsque la distribution est informée (éloignée d'une distribution gaussienne) ou lorsque la variance du bruit est forte. Le filtre particulaire -ou méthode de Monte Carlo séquentielle- décrit dans [HM69] se différencie de l'UKF par l'utilisation d'un filtre Bayésien récursif (et non grâce à la méthode des moindres carrés). La représentation de la distribution est effectuée par des points en grand nombre que l'on appelle particules. Ces particules vont être amenées à s'agréger les unes aux autres de manière à prendre "la forme" de la distribution. Ceci revient donc à approximer numériquement (dans l'étape de prédiction) la distribution définie par Chapman-Kolmogorov (équation (4.41)). Par convention $y_{1:k}$ désigne la collection (y_1, \dots, y_k) , δ la fonction de Dirac, $p(\cdot)$ la densité de probabilité et ξ_k^i les particules ($i \in [1, N]$, N étant le nombre de particules mises en jeu) pondérées par ω_k^i ([LS10]).

$$\begin{aligned} \hat{x}_k &= \int x_k p(x_k | y_{1:k}) dx_k \approx \int x_k \sum_{i=1}^N \omega_k^i \delta(x_k - \xi_k^i) dx_k \\ &\approx \sum_{i=1}^N \omega_k^i \xi_k^i \end{aligned} \quad (4.40)$$

Ce type de méthode repose sur les étapes suivantes :

1. création d'un jeu de particules aléatoirement positionnées (échantillonnage) pour la première itération ;
2. pondération des particules selon leur vraisemblance conditionnelle d'appartenance à la distribution réelle (affectation pour chaque particule d'un poids proportionnel à sa distance à la mesure) ;
3. normalisation des poids ;
4. conservation des particules les plus proches de la mesure, les autres sont éliminées ;
5. duplication des particules survivantes (génération d'un nouveau jeu de particules \Leftrightarrow rééchantillonnage).

Utilisée principalement pour des applications de suivi de cible ([XZ10], [MdROM11], à noter les travaux de [Rei11] rapprochant les filtres particuliers de la théorie de Dempster-Shafer), cette méthode nécessite un nombre de particules proportionnel à la complexité de la distribution. Dès lors, bien que cette méthode soit une solution élégante pour l'estimation de distributions, le coût calculatoire devient un handicap sérieux pour des applications telles que celles étudiées ici. En effet, nous justifions son éviction

pour ce motif, la méthode de suivi devant cohabiter avec la loi de commande chargée du déplacement ainsi qu'avec la strate logicielle dédiée à la surveillance du système.

En résumé :

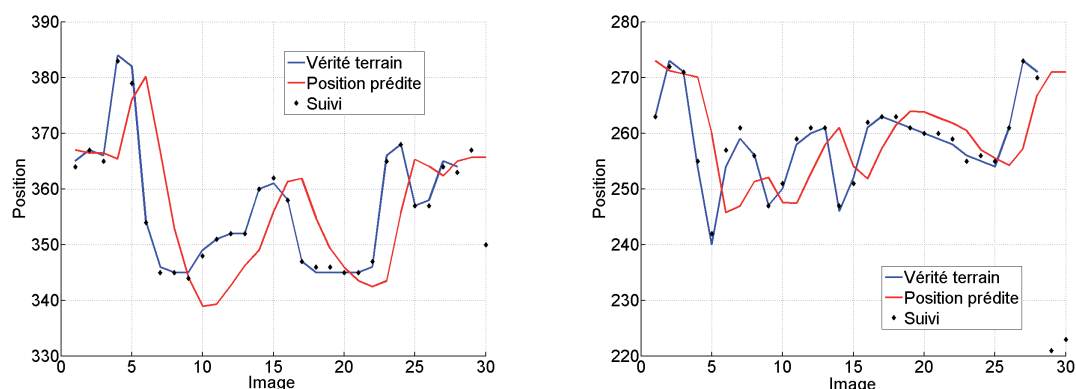
Nous venons de voir différentes approches pour la prédiction de la position du fragment représentatif de la zone d'intérêt. Le filtre de Kalman octroie une solution optimale dans le cas de systèmes linéaires. Dans sa version EKF, ce filtre permet d'être utilisé pour des systèmes non-linéaires mais s'avère limité par des hypothèses de fonctionnement restrictives. Les hypothèses plus laxées dans sa variante UKF lui permettent d'être utilisé pour des densités de probabilité éloignées d'une distribution gaussienne. Enfin, le filtrage particulaire ne souffre quant à lui d'aucune restriction et peut être utilisé pour des cas d'applications très variées. Cependant, le temps de traitement nécessaire limite en soi son usage pour cette étude.

4.2.2.4 Application du filtre de Kalman

Afin d'éviter toute erreur de localisation de la zone d'intérêt, une étape supplémentaire a été insérée entre la mesure de similarité et la loi de commande. Celle-ci consiste à surveiller si les coordonnées du fragment retrouvé sont cohérentes. Pour ce faire on regarde les variations entre la position du fragment prédite par un filtrage et celle obtenue par le processus de suivi. Dans le cas où les variations pour l'un des axes dépassent un certain seuil une alerte apparaît. Celle-ci traduit une erreur de suivi et vient déclencher l'arrêt du déplacement de la coupole. Le problème est que le chirurgien poursuit son geste durant toute la phase de suivi. Des erreurs peuvent donc survenir même dans les cas où le fragment est bien retrouvé. Le choix du seuil doit donc être fait avec précaution : pour un seuil trop faible, l'occurrence des arrêts empêcherait un déplacement fluide, et s'il est trop haut aucune alerte ne serait donnée, ce qui pourrait amener la coupole à se diriger vers une région non souhaitée. Le choix s'est porté sur un seuil de 20pixels/image, ce qui permet expérimentalement d'obtenir un bon compromis sécurité/efficacité.

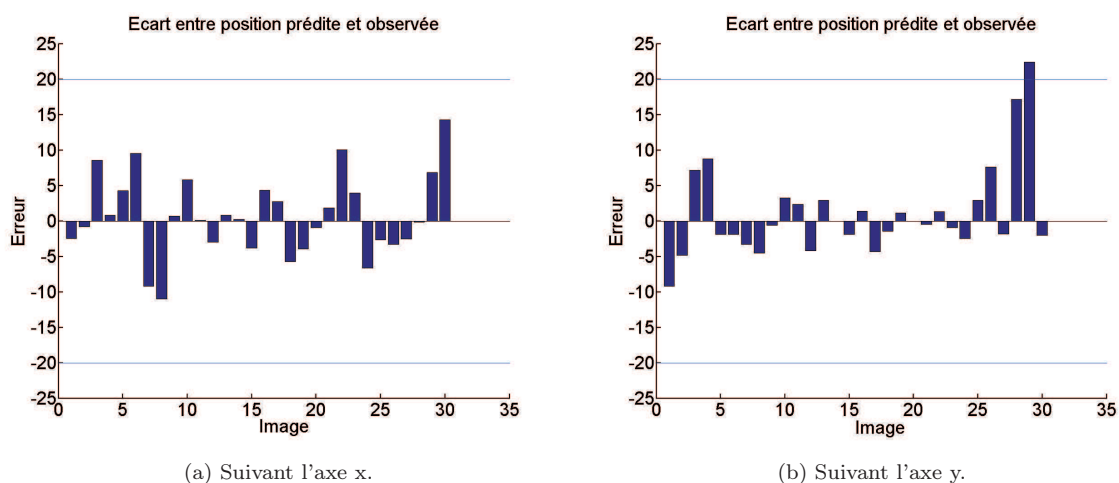
Dans un premier temps, nous prédisons à l'aide d'un filtre de Kalman la position du fragment. Les figures 4.15a et 4.15b exposent les résultats obtenus pour l'exemple présenté dans la section précédente pour les axes x et y.

Comme présenté précédemment, le filtrage de Kalman se déroule en deux étapes. La première s'attache à prédire la position du fragment pour l'instant $t + 1$. Cet avis est révisé dans un second temps par l'étape dite de réactualisation suivant l'observation prise à l'instant t . L'hypothèse suivant laquelle le système est linéaire est ici erronée, le chirurgien ne suivant aucun schéma d'action, les gestes s'apparentant plus à des mouvements erratiques. Cependant comme le montre la figure 4.17, cette approximation est toutefois acceptable. On peut remarquer qu'il y a un prémisses d'alerte pour l'image #29 correspondant à la sortie du chirurgien, mais que du fait du non dépassement du seuil l'alerte n'est donnée qu'en image #30.



(a) Prédiction de la position du fragment suivant l'axe x. (b) Prédiction de la position du fragment suivant l'axe y.

FIG. 4.15 – Prédiction de la position du fragment par filtrage de Kalman.



(a) Suivant l'axe x.

(b) Suivant l'axe y.

FIG. 4.16 – Variation de l'erreur entre position fournie par mesure de similarité et celle prédite par Kalman.

Cette technique semble donc effective mais repose sur un modèle inexact basé sur l'hypothèse de linéarité du système. Nous nous intéressons donc à trouver un modèle proche du comportement observé. Ce modèle est injecté dans un filtre de Kalman étendu. Le principe de fonctionnement est semblable au précédent à ceci près que le modèle est cette fois non linéaire. Sur les figures 4.17a et 4.17b il est immédiat de remarquer que le modèle choisi n'est pas très adéquat. Celui-ci ne permet pas de générer des alertes compte-tenu du seuil utilisé, ce qui conduit à des déplacements incorrects. Il aurait été toutefois possible de réviser la valeur du seuil utilisé. Cependant, l'abaisser entraîne un arrêt prématuré du système dès la première image.

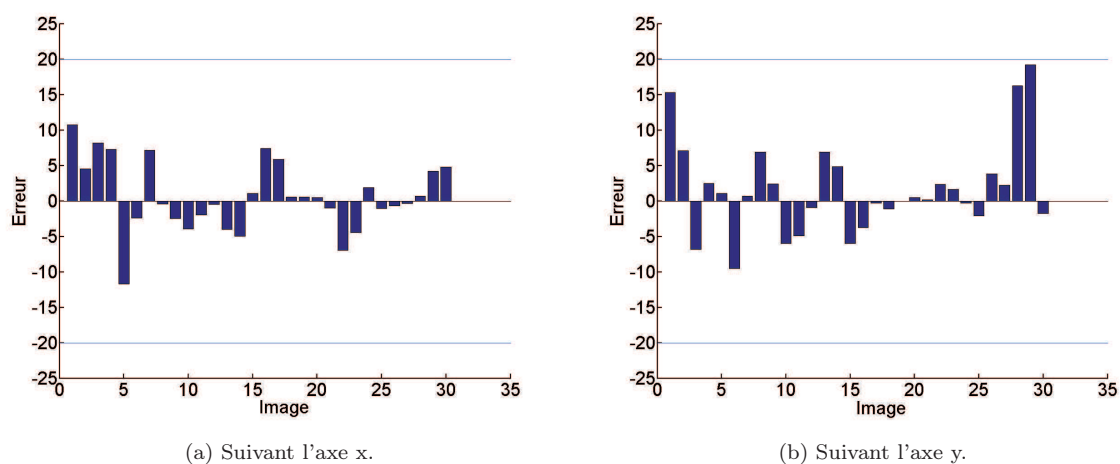


FIG. 4.17 – Variation de l'erreur entre position fournie par mesure de similarité et celle prédite par EKF.

De nombreuses modifications du modèle représentatif des mouvements du chirurgien ont été élaborées. Le modèle sinusoïdal est celui donnant le plus satisfaction du fait des mouvements effectués lors de sutures ou de dissections. Compte tenu des résultats obtenus avec le filtrage de Kalman, du fait de sa mise en place simple et du coût calculatoire faible ($\sim 16\text{ms}$), ce type de surveillance a donc été conservé. Toutefois nous limiterons à 5 images la durée pendant laquelle la phase de suivi est active même en l'absence d'alerte par le filtrage de Kalman. Ce délai dépassé la coupole est automatiquement arrêtée, ceci afin d'éviter tout mauvais comportement. Ce type d'arrêt dû à l'expiration du délai n'est jamais atteint, la fréquence de désignation permettant une redéfinition de la zone d'intérêt dans un temps inférieur à 5 images.

4.3 Conclusion

Dans ce chapitre, nous avons montré qu'il est possible de gérer le déplacement d'un éclairage opératoire à partir d'un retour vidéo.

Parmi les différents asservissements visuels possibles, l'asservissement visuel 2D se révèle être le plus adéquat. Les asservissements 3D, $2D^{1/2}$, et $d2D/dt$ ont été écartés compte-tenu que nous n'avons pas accès à un modèle (d'apparence ou de champ de vitesse) représentatif des objets en présence dans la scène chirurgicale. Dans notre cas, l'asservissement visuel est indirect et l'éclairage opératoire est assimilé à un système robotique à deux degrés de liberté doté d'une caméra perspective en bout d'effecteur observant dans le visible. Les principales difficultés de mise en œuvre de cette régulation dans l'image sont relatives à l'inversion de la matrice d'interaction qui pose généralement problème mais qui s'avère aisée ici compte-tenu de l'altitude constante du système.

La primitive visuelle utilisée pour cette étude est de type point. Elle ne correspond pas une caractéristique de la zone d'intérêt mais résulte de la maximisation d'un critère. Le déplacement continu de

l'éclairage est donc conditionné par la capacité du traitement d'image à désigner la zone d'intérêt à chaque instant. Or, cette hypothèse n'est pas toujours vérifiée ce qui implique l'arrêt du système. Pour combler ce manque d'information et ainsi retrouver la zone d'intérêt dans l'image entre deux désignations, la technique retenue consiste en une mesure de similarité, capable de retrouver dans l'image courante le fragment correspondant à la dernière zone d'intérêt désignée. Nous avons vu que cette technique n'est pas garante de la fiabilité du déplacement, et qu'il est possible pour y pallier d'ajouter un filtre de type Kalman.

Ceci conclut la partie théorique du sujet d'asservissement visuel d'un éclairage opératoire.

Chapitre 5

Réalisations techniques & Résultats

Sommaire

5.1	Prototype d'un éclairage opératoire motorisé	142
5.1.1	Motorisation de l'arceau	142
5.1.2	Acquisition d'image	145
5.2	Mise en œuvre de la désignation de la zone d'intérêt	148
5.2.1	Reconnaissance de la classe instrument	148
5.2.2	Reconnaissance des autres classes	157
5.2.3	Désignation de la zone d'intérêt	166
5.3	Asservissement visuel de l'éclairage opératoire	173
5.3.1	Synthèse de la commande	173
5.3.2	Bilan des résultats expérimentaux	178
5.4	Conclusion	181

Dans les chapitres précédents, nous avons présenté le concept d'éclairage opératoire asservi visuellement. Nous avons montré comment, à partir d'une unique caméra placée dans l'effecteur du robot, il est possible d'identifier la zone d'intérêt du chirurgien et de s'y rendre.

Dans ce chapitre, nous nous intéressons à la mise en œuvre de ce concept. Nous détaillerons les étapes de désignation de la zone d'intérêt et d'asservissement visuel.

Des expérimentations sur le prototype illustrent les performances de la solution proposée en termes d'"autonomie" et d'"intelligence".

5.1 Prototype d'un éclairage opératoire motorisé

5.1.1 Motorisation de l'arceau

La maquette utilisée pour cette étude est une modification d'un produit existant. Ceci permet d'une part de se placer dans des conditions réelles (niveau d'éclairage entre autres) et d'autre part d'être conforme à la norme IEC60601 concernant l'utilisation d'éclairages opératoires dans les blocs chirurgicaux et ainsi éviter les contraintes de conception concernant les risques de chute d'éléments dans la scène, la perturbation des équipements environnants, les précautions d'isolation électrique, etc. Le modèle choisi est un Powerled700 de la société MAQUET (voir figure 5.1a) dans la mesure où cet éclairage opératoire est le plus difficile à manipuler et donc le plus intéressant à robotiser.

Remarque : *Ce choix n'est pas sans conséquence et a rendu la tâche de robotisation de l'arceau difficile dans la mesure où certaines contraintes inhérentes à la robotisation n'ont pas été prises en compte lors de la conception originale de ce produit (structure non équilibrée).*

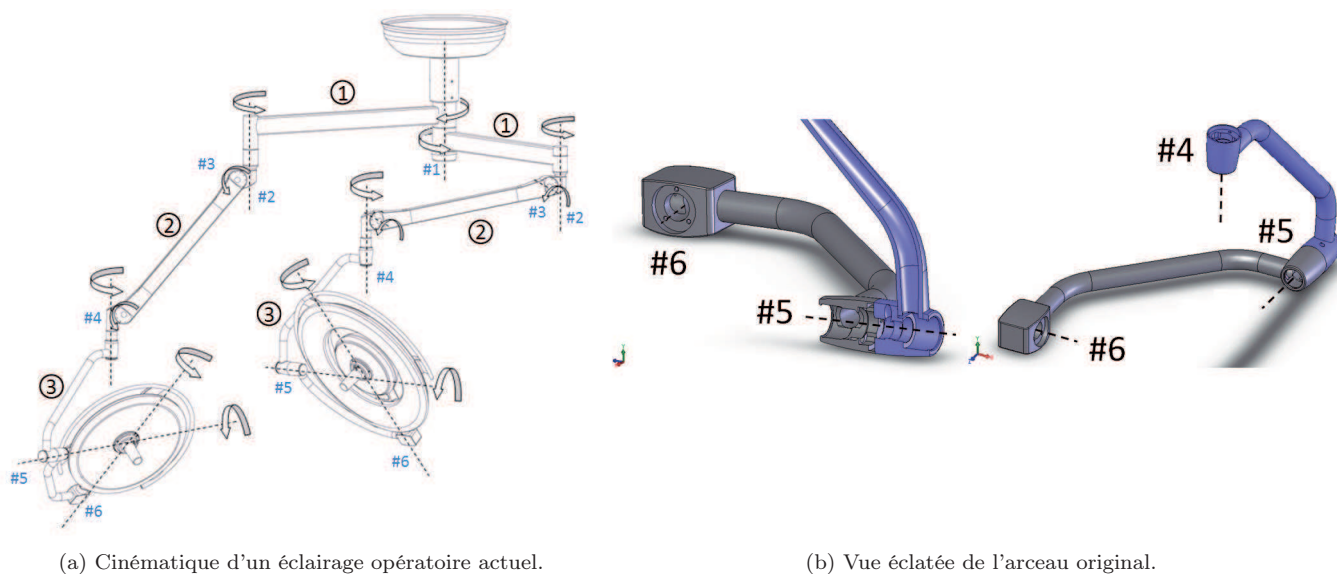


FIG. 5.1 – Illustration du principe pour la réorientation et le repositionnement.

Choix et intégration des moteurs

Pour pouvoir dimensionner les moteurs, nous avons étudié les efforts générés au niveau des axes #5 et #6 par une méthode analytique et une méthode par éléments finis, respectivement mises en œuvre à l'aide des logiciels Matlab et ADAMS.

Afin de faciliter le choix des moteurs, nous considérons les conditions suivantes :

- le coefficient de sécurité est supposé égal à un ;
- le poids total de la coupole est de 18.5kg ;
- la vitesse de déplacement de la coupole est de $6\text{tr}\cdot\text{min}^{-1}$.

Remarque : *La vitesse de déplacement est le résultat d'un compromis entre la nécessité de réactivité du système et le confort du chirurgien. Une vitesse trop grande induirait une variation de l'éclairage brutale, éblouissant le chirurgien. Le choix de la vitesse de déplacement a été ultérieurement approuvé par les différents chirurgiens, lors d'essais expérimentaux.*

Type de sollicitation	Axe considéré [.]	Couple calculé par méthode analytique [Nm]	Couple calculé par méthode par éléments finis [Nm]
Statique	#5	6.454	7.301
	#6	0	0
Dynamique	#5	7.269	7.988
	#6	0.196	0.198

TAB. 5.1 – Evaluation du couple nécessaire pour mettre en mouvement la coupole.

Les résultats présentés dans le tableau 5.1 montrent que les moteurs devront développer un couple de plus de 8Nm pour permettre la mise en mouvement de la structure.

La prospection des moteurs s'est donc basée principalement sur le couple à développer et sur des critères tels que le poids (les liaisons étant en série, le poids du moteur de l'articulation #6 affecte la masse à déplacer pour le moteur #5), les dimensions (encombrement restreint dans les articulations de l'arceau), la tension de fonctionnement (réglementairement limitée à 24V) et la capacité du moteur à pouvoir effectuer un mouvement réversible (dans les deux sens de rotation).

Remarque : *Le nombre de cycles (dépendant dans notre cas de l'activité du chirurgien), le couple natif (à prendre en compte pour une version future), ainsi que la thermique des moteurs (constante de temps thermique, température de fonctionnement) n'ont pas été considérés pour cette activité de maquetage, ils constituent une piste d'amélioration du prototype.*

En croisant les références disponibles avec les contraintes de l'application, nous avons sélectionné le moteur Amax22CLL (société MDP) couplé à un réducteur GP022C. Lors des premiers essais, cette combinaison s'est avérée sous-dimensionnée par rapport aux sollicitations réelles. Nous n'avions pas pris en compte, dans les simulations, la présence des câbles dans la structure qui génèrent un couple résistant. L'intégration de moteurs plus puissants (Amax32G de couple maximum 36.7Nm fonctionnant sous 24V, avec réducteur PLG32 à rapport de réduction 1 :411) a permis de résoudre ce problème mais s'est révélée inadéquate avec l'utilisation d'une courroie lisse dans les dimensions imposées. Les illustrations 5.2c et 5.2d présentent l'intégration des moteurs pour ce cas. La courroie lisse a donc été remplacée par une courroie crantée donnant meilleure satisfaction. Malheureusement, les points durs dus à la mauvaise répartition de charge de l'éclairage ont provoqué leurs casses répétées. Il a donc été décidé d'une part, d'interchanger la coupole avec celle d'un Powerled300 plus petite (450mm de diamètre contre 740mm) et plus légère (10kg contre 18.5kg) et d'autre part, de retravailler les arceaux (par allègement de la structure et raccourcissement de la fourche entre les axes #5 et #6). De plus, afin de garantir la robustesse mécanique de la structure lors d'opérations réelles, il a été décidé de remplacer la transmission par des

engrenages droits à rapport de réduction 1:1. Dans la mesure où le prototype se trouve à bonne distance du chirurgien, il n'intègre aucun système de débrayage permettant, par désaccouplement du moteur, de limiter la force de l'impact en cas de choc.

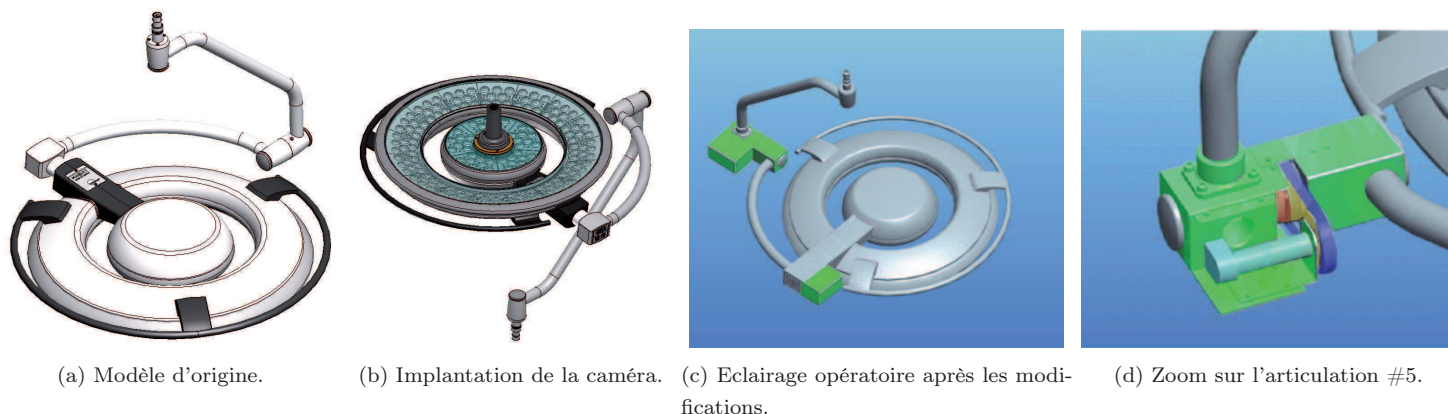


FIG. 5.2 – Vue globale des modifications apportées sur les articulations #5 et #6.

Cette maquette est montée sur une base mobile ("Rolite" de la société MAQUET voir figure 5.3a) permettant les tests en bloc opératoire sans nécessité de désinstaller l'éclairage déjà implanté. Cette base permet de placer l'éclairage opératoire dans des conditions d'emploi identiques à celles rencontrées en conditions réelles : distance coupole-plaie de l'ordre de 1.1m, réorientation de la coupole possible par le chirurgien ou le personnel chirurgical, etc. Seul le repositionnement diffère et s'effectue dans notre cas par déplacement de la base mobile et non par rotation autour de l'axe #1.

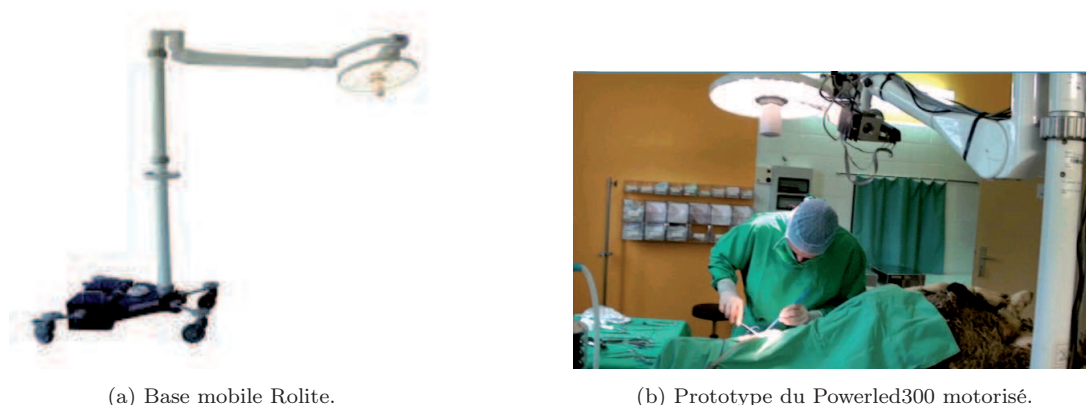


FIG. 5.3 – Prototype de l'étude.

La figure 5.3b présente la dernière version du prototype intégrant l'ensemble de ces modifications. Nous utiliserons ce prototype lors des différents essais en bloc opératoire dont les résultats sont présentés par la suite.

L'asservissement visuel de l'éclairage opératoire se fait de manière indirecte. Le bloc d'asservissement visuel fournit, à partir des informations visuelles et de leurs traitements, le torseur cinématique à appliquer à la caméra pour atteindre la zone d'intérêt. Le calcul de la loi de commande s'effectue selon les modalités vues en chapitre 4; la régulation étant assurée par des modules indépendants placés en cascade. Cette architecture est flexible et nous permet de ne pas être limités à deux moteurs. De plus, les traitements nécessaires à la régulation s'effectuent au niveau des modules de contrôle, ce qui permet d'alléger les calculs au niveau de l'ordinateur. Il est donc possible de mettre ainsi en cascade plusieurs autres modules afin de permettre un mouvement complet dans \mathbb{R}^3 . Cependant dans notre cas, l'absence d'information sur la pose désirée de l'éclairage opératoire nous oblige à ne considérer que les réorientations.

De ce torseur sont extraites les deux vitesses angulaires appliquées en consigne aux deux moteurs. L'ordre combinant ces deux vitesses est envoyé à un module maître par liaison USB (EPOS2 cadencé à 60MHz). Ce module récupère la partie liée à l'articulation #5 et la renvoie par liaison CANopen à un second module dit esclave correspondant à l'articulation #6. Maître et esclave modulent le courant envoyé à leur moteur respectif par rapport à cet ordre mais aussi relativement aux calibrations définies lors de pré-essais (profil d'accélération-décélération, appel de courant, gestion des erreurs et des pannes, etc). Le retour d'information au niveau moteur est assuré par des codeurs optiques 500 points (HEDL 5540 société Maxxon). Ceux-ci permettent d'informer, à intervalles de temps réguliers, les modules sur la position des arbres moteurs. L'intégralité de ce processus est régie par une couche prioritaire (au niveau de l'ordonnanceur) permettant d'arrêter tout déplacement en cours et de passer à un mode manuel où les ordres sont donnés explicitement par l'équipe chirurgicale via une interface logicielle (mode dégradé).

5.1.2 Acquisition d'image

La réorientation autonome de la coupole nécessite de récupérer le signal vidéo de la caméra se situant dans la poignée centrale de la coupole (figure 5.2b) et de le traiter. La prise d'images est assurée par un bloc SONY FCB-EX1000P (résolution de 0.44Mpixels) qui permet l'observation de la scène chirurgicale malgré le haut niveau d'éclairement. Ce choix est justifié également par ses caractéristiques qui facilite son intégration dans la coupole (poids, compacité, connectique, ...). Le signal (composite) créé par cette caméra est envoyé à un boîtier de numérisation Dazzle (DVC-100). Celui-ci renvoie le flux par liaison USB à l'ordinateur portable Asus G51J (système d'exploitation Windows 7, 1GB de mémoire vive, processeur Intel i7 famille 7200M), chargé du traitement d'images. Ce choix simule la puissance de calcul finale post étude. Les programmes du traitement d'images sont développés en C++ et utilisent des bibliothèques OpenCV permettant une récupération très simple des images de la scène opératoire.

Les conditions d'exploitation sont très spécifiques et diffèrent de celles habituellement rencontrées pour des applications de reconnaissance d'objets. En effet, la luminance perçue par la caméra est dans notre cas démultipliée en raison de la puissance du projecteur pouvant atteindre 125klux. Ce problème de fort éclairage est l'une des principales difficultés de cette étude. Compte-tenu des dimensions de la tache lumineuse (180mm pour un Powerled300 placé à une distance de 1m de la scène), l'image obtenue est fortement contrastée malgré notre choix de caméra, comme le montre la figure 5.4a. Il est alors difficile de reconnaître des objets au-delà de la région couverte par la tache.

Ceci entraîne trois cas de figures :

1. les gestes du chirurgien sont effectués dans la tache lumineuse. Il est possible d'accomplir une réorientation de la coupole sans trop de difficulté, l'éclairage ayant le temps de centrer le faisceau lumineux sur la zone d'intérêt du chirurgien ;
2. les gestes du chirurgien sont effectués en dehors de la tache lumineuse. Aucune reconnaissance n'étant possible en dehors de la tache lumineuse il n'y a donc pas de réorientation de la coupole. Pour amener le système à éclairer une zone d'intérêt placée en dehors de la tache lumineuse, le chirurgien peut effectuer une série de "zones d'intérêt temporaires" permettant au système de converger vers la zone d'intérêt réelle. Le temps nécessaire à cette opération est trop important par rapport à une réorientation manuelle, le chirurgien privilégie donc cette dernière option ;
3. les gestes du chirurgien sont à la frontière de la tache lumineuse. La mise en mouvement de la coupole n'est pas assurée car la zone d'intérêt peut ne pas être "reconnue" ce qui empêche toute réorientation du système.

Il apparaît donc primordial de trouver une solution au problème de saturation de l'image qui limite fortement l'intérêt du système si nous maintenons les caractéristiques du projecteur (puissance, diamètre de tache, ...). Plusieurs solutions ont été envisagées :

1. l'adjonction d'un matériau diffusant ;
2. l'ajout de projecteurs supplémentaires éclairant à la périphérie de la tache principale ;
3. l'application d'un filtre à densité neutre ;
4. l'application d'un filtre à transmission variable.

Remarque : *Une possibilité aurait été de modifier les paramètres internes à la caméra, notamment diminuer le temps d'exposition de la caméra. Cette solution s'avère ici impossible à mettre en œuvre en raison de l'inaccessibilité de ce paramètre pour le matériel choisi.*

La première solution consiste à accoler un matériau diffusant (verre dépoli par exemple) à la source de lumière pour la rendre moins directive. Comme le montre la figure 5.4b, cette technique permet une exploitation plus importante de l'image. Cependant, elle reste insuffisante compte-tenu de l'assombrissement des bords de l'image et de l'inhomogénéité de la couleur sur le carton témoin.

Une seconde technique consiste à ajouter des projecteurs supplémentaires éclairant la périphérie de la tache principale, ceci toujours dans le but de diminuer le contraste dans l'image. A l'éclairage est ajouté quatre projecteurs basse puissance de type LUCEA (de la société MAQUET) d'un éclairage de 10klux chacun. Il n'a malheureusement pas été possible d'intégrer les tests effectués pour cette solution à ce manuscrit.

Troisième technique éprouvée au cours de cette étude, l'ajout d'un filtre à densité neutre au niveau de la pupille d'entrée de la caméra, jouant le rôle d'atténuateur de lumière afin d'obtenir un contraste moins grand. Le problème est que modifier uniformément la luminance n'affecte pas le contraste. Dans notre cas, il s'agit de diminuer fortement l'éclairage au centre sans affecter le reste de l'image. Coller une surface opaque sur la caméra, nous permet d'obtenir une image exploitable dans sa globalité (figure

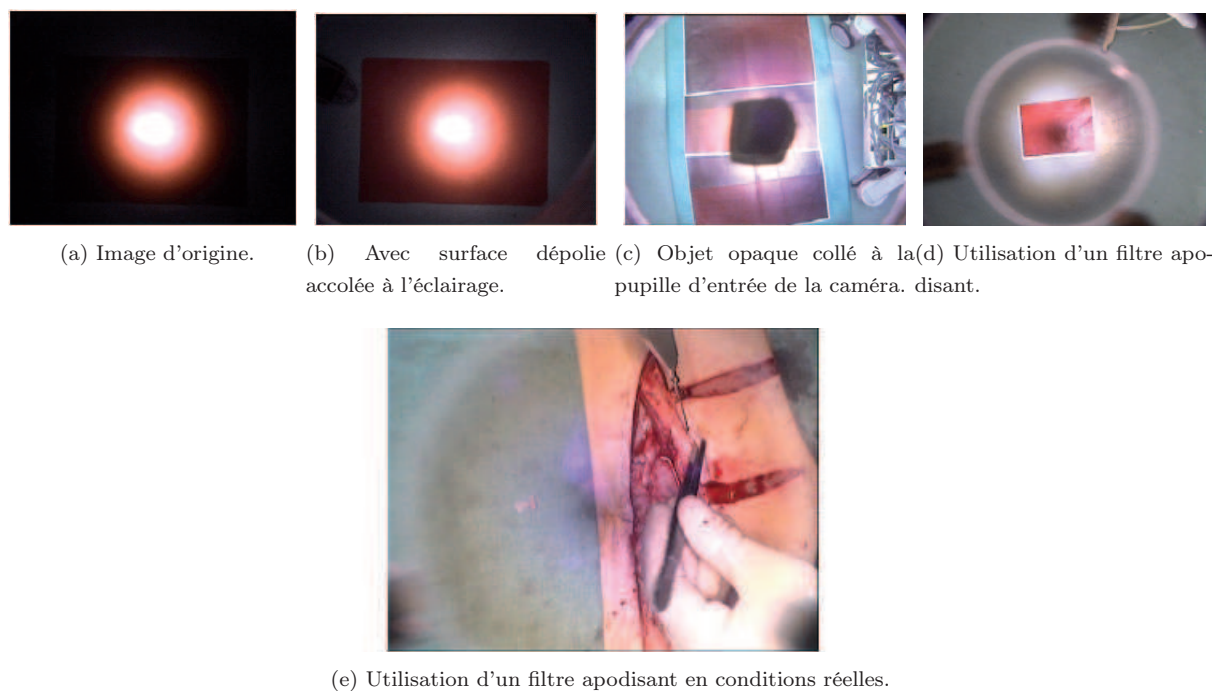


FIG. 5.4 – Différentes techniques pour la diminution du contraste.

5.4c). Cependant, le centre de l'image est non visible et empêche toute désignation au centre. Une erreur statique (distance entre la zone d'intérêt réelle et celle désignée) apparaît et dépend du diamètre de l'objet utilisé. Cette erreur empêche a fortiori tout réajustement optimal de la coupole. Bien qu'inadaptée, cette solution donne une idée de l'approche à suivre.

La solution retenue consiste à utiliser un filtre à transmission variable. Ces filtres dits apodisants se distinguent des filtres neutres par leur capacité à uniformiser la luminance malgré une répartition d'éclairage inhomogène. Le profil de transmission utilisé s'approche d'une loi gaussienne; l'image obtenue donne satisfaction (figure 5.4e). Sur cette image, nous pouvons remarquer tout de même certaines variations de luminance dues à l'écart entre le profil de la coupole et celui du filtre (formation de halos sombres concentriques et de réflexions de certaines parties de l'image). A défaut d'autres solutions plus adaptées et malgré l'existence d'anneaux opaques dans l'image, ce dispositif sera conservé pour l'intégralité des tests à venir.

Remarque : Dans le cas d'interventions en chirurgie majeure, cadre de notre étude, la réglementation oblige d'utiliser deux coupoles afin de prévenir tout défaut. Il est donc légitime de se demander si les problèmes évoqués pour une coupole ne sont pas démultipliés avec l'utilisation d'une seconde coupole. Comme nous le verrons par la suite, l'emploi d'une source supplémentaire améliore le rendu dans l'image. Ceci est dû à la différence entre les distributions du filtre et de la coupole. En effet, l'augmentation de l'éclairage contrebalance les régions sombres là où la transmission était trop élevée. L'utilisation du prototype avec des éclairages opératoires supplémentaires ne pose donc pas de problème pour l'exploitation de l'image dans sa totalité.

En résumé :

Un prototype d'un éclairage opératoire motorisé a été développé. Les normes très draconiennes du milieu chirurgical ont été respectées et nous permettent d'utiliser le prototype en bloc opératoire. Cette activité a nécessité de nombreuses modifications très consommatrices en temps. La caméra, soumise au fort éclairage de la coupole, a été munie d'un filtre apodisant afin de rendre les images exploitables. Le profil de transmission de ce filtre ne correspond pas exactement à celui de la distribution de lumière fournie par la coupole. Ceci entraîne une légère inhomogénéité dans l'image acceptable dans cette étude. L'ensemble des traitements (désignation de la zone d'intérêt, calcul de la loi de commande) sera assuré par un ordinateur portable.

5.2 Mise en œuvre de la désignation de la zone d'intérêt

Nous présentons maintenant l'extraction des informations forme, couleur et mouvement de manière pratique pour la reconnaissance des classes instrument, sang, peau et champs stériles présents dans la scène opératoire.

5.2.1 Reconnaissance de la classe instrument

Attribut forme

Par hypothèse les instruments chirurgicaux possèdent un grand nombre d'éléments rectilignes. Ceci nous permet de nous affranchir d'une quelconque connaissance a priori sur les dimensions, la couleur, la texture ainsi que sur la forme exacte des instruments. La classe instrument se base donc principalement sur l'attribut forme (correspondant à la quatrième source d'information utilisée dans la reconnaissance d'objet).

Pour trouver les éléments rectilignes dans l'image nous utilisons la transformée de Hough (cf. chapitre 2) dont nous rappelons brièvement les étapes clés :

- détection de contour grâce un masque de type Prewitt ;
- recherche de tous les couples (ρ, θ) , paramètres de la droite passant par le point de coordonnées (i,j) (figure 5.5b) ;
- répétition de la recherche pour chaque point de l'image ;
- recherche des maxima dans l'espace des paramètres (en rouge sur la figure 5.5b) ;
- création des droites correspondant aux couples trouvés (figure 5.5c).

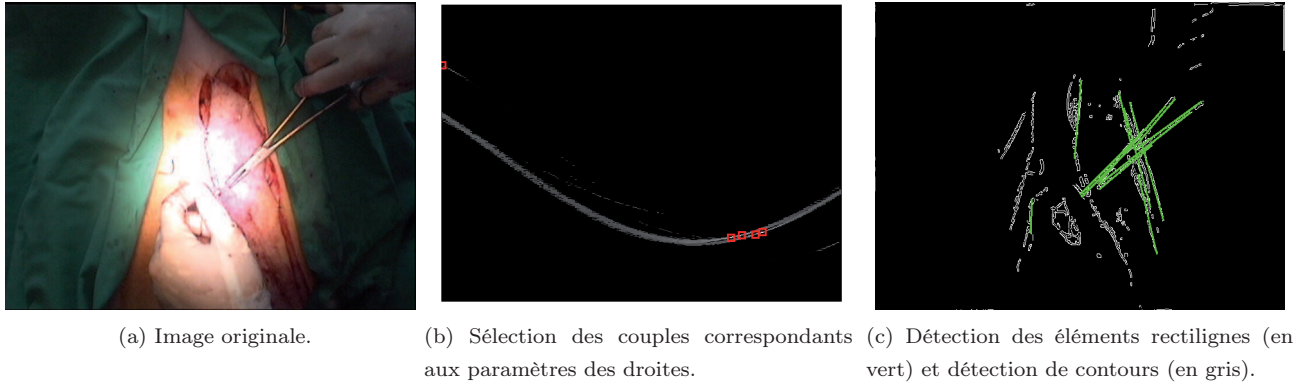


FIG. 5.5 – Reconnaissance de la classe instrument.

A partir de l'image (figure 5.5c) que nous appellerons I_L (la valeur d'un pixel est bornée entre 0 et 255), nous créons les masses m_{i4} représentatives de la classe instrument vue de la source forme. Cette étape s'effectue selon le modèle d'Appriou :

$$\begin{cases} m_{14}(\{\text{instruments}\}) = \beta_4(R_4 p(s_4 = N | \{\text{instruments}\})) \\ m_{24}(c\{\text{instruments}\}) = \beta_4(1 - R_4 p(s_4 = N | \{\text{instruments}\})) \\ m_{34}(\Theta) = 1 - \beta_4 \end{cases}$$

avec :

- $p(s_4 = N | \{\text{instruments}\})$, la valeur de la densité de probabilité conditionnelle de la source 4 (correspondant à l'attribut forme) lorsque le niveau de gris du pixel de coordonnées (k, l) de l'image I_L est égal à N , sachant l'hypothèse $\{\text{instruments}\}$;
- R_4 , le coefficient permettant la normalisation de la masse. Considéré égal à 1 dans la mesure où la densité de probabilité conditionnelle $p(s_4 | \{\text{instruments}\})$ est normalisée. L'injection du gain de la classe non pas au niveau de la prise de décision mais dans R_4 constitue une perspective possible ;
- β_4 , le coefficient représentatif de la fiabilité de l'affectation (pris par défaut égal à 1) ;
- m_{14} , la croyance que le pixel appartient à un instrument ;
- m_{24} , la croyance que le pixel n'appartient pas à un instrument ;
- m_{34} , la masse représentative de l'ignorance.

La création des masses nécessite une connaissance a priori. Celle-ci est contenue dans la densité de probabilité conditionnelle $p(s_4 | \{\text{instruments}\})$, synthétisée sous la forme d'un histogramme comme présenté en figure 5.6. La détermination de $p(s_4 = N | \{\text{instruments}\})$ s'effectue par lecture de l'histogramme au niveau de gris du pixel considéré. Compte-tenu de l'allure de cet histogramme, seuls les pixels appartenant aux lignes droites détectées auront une masse $m_{14}(\{\text{instruments}\})$ non nulle :

$$\begin{cases} m_{14}(\{\text{instruments}\}) = \beta_4(R_4) & \text{si } I_L(k, l) = N = 255; \quad \beta_4(1 - R_4) \text{ sinon;} \\ m_{24}(c\{\text{instruments}\}) = \beta_4(1 - R_4) & \text{si } I_L(k, l) = N = 255; \quad \beta_4(R_4) \text{ sinon;} \\ m_{34}(\Theta) = 1 - \beta_4 \end{cases}$$

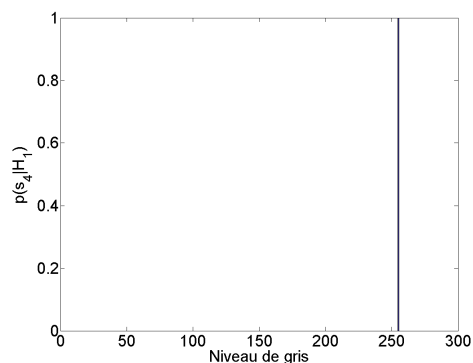


FIG. 5.6 – Histogramme représentatif de la classe instrument.

Dans la mesure où il est possible que la transformée de Hough assimile une interface champs opératoires-peau (ou toutes autres délimitations franches et rectilignes trouvées lors de la détection de contours) à un instrument, nous affaiblissons les masses liées à cette source à l'aide d'une matrice de confusion. Ceci revient à moduler l'importance des masses selon que la transformée de Hough permet de localiser fidèlement les instruments ou non. Il existe donc un lien direct entre les masses m_{i4} et l'aptitude de la méthode à détecter les lignes droites. Pour établir les différentes composantes de la matrice de confusion, nous utilisons un masque isolant l'objet (ici un instrument) défini manuellement. Le tableau 5.2 présente les taux obtenus pour ce type d'attribut.

	instrument ¹	non instrument ¹
instrument ²	46.4%	1.4%
non instrument ²	53.6%	98.6%

TAB. 5.2 – Matrice de confusion pour la classe instrument (éclairage au niveau maximal-avec filtre apodisant). "1" correspond à la vérité terrain, "2" à celle du système.

Nous pouvons déduire du tableau 5.2 que les faux négatifs (correspondant aux pixels appartenant à la classe d'objets manqués par le système) sont à un taux anormalement élevé comparativement aux vrais positifs (correspondant aux bonnes reconnaissances). Ceci signifie que notre méthode ne reconnaît pas un grand nombre de pixels appartenant à l'instrument. A contrario, le taux de vrai négatif indique une bonne capacité de la méthode à identifier les pixels appartenant à la proposition non instrument.

La reconnaissance est donc insuffisante mais fiable. Cependant, notre objectif principal est de ne détecter que les instruments, ceci afin d'éviter au maximum les mauvaises désignations (l'objectif secondaire étant de détecter plus de pixels appartenant à la classe instrument.). Il s'agit donc de faire tendre la précision vers 1.

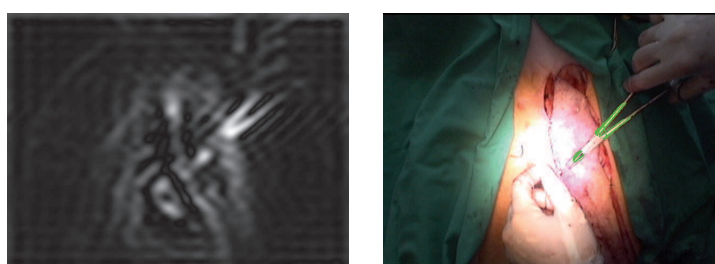
Pour rendre plus robuste la méthode, nous renforçons la détection de contours grâce à un filtre passe-haut

afin d'éliminer les gradients de faible amplitude. Nous obtenons les résultats présentés en tableau 5.3.

	instrument ¹	non instrument ¹
instrument ²	19.7%	$< 10^{-2}\%$
non instrument ²	80.3%	99.9%

TAB. 5.3 – Matrice de confusion pour la classe instrument après ajout d'un filtre passe-haut. "1" correspond à la vérité terrain, "2" à celle du système.

La précision est, comme souhaité, augmentée de manière significative passant de 0.9707 à 0.9999, au détriment du rappel qui chute à 0.1971. Ceci s'explique par le bruit généré par l'étape de filtrage qui pénalise la quantité de pixels identifiés, mais a pour conséquence d'augmenter la distance qui est alors égale à 0.8029.



(a) Résultat du filtrage passe-haut. (b) Surimpression du résultat dans l'image courante.

FIG. 5.7 – Reconnaissance de la classe instrument avec filtrage passe-haut.

L'attribut forme n'est pas suffisant pour affecter un pixel à la classe instrument. En effet, il faut éviter de détecter les instruments "inactifs" posés aux abords de la plaie opératoire et pouvant être source de mauvaises désignations. Dès lors, pour reconnaître cette classe d'objet nous utilisons les masses m_{i5} correspondant au mouvement.

Remarque : *Comme nous l'avons vu en chapitre 2, la couleur ne participe pas à la discrimination des instruments. Dès lors, les masses $m_{ij}, j \in \{1, 2, 3\}$ ne sont pas utilisées et sont exclues de la fusion afin de gagner en temps de calcul.*

Attribut mouvement

Pour permettre la quantification du mouvement, l'évaluation du flot optique grâce aux détecteurs de Harris apparaît comme le moyen le plus simple pour créer un champ de vecteurs représentatif des objets en mouvement dans la scène. Comme le montrent les figures 5.8a et 5.9a (l'intensité des pixels a été augmentée pour faciliter la lecture), le champ de vecteurs obtenu par cette méthode n'est pas sous forme de maillage régulier, mais est conditionné par les points d'intérêt détectés. Grâce à l'emploi d'un filtre

gaussien permettant de répartir localement les vecteurs mouvements, la quantification du mouvement est cohérente et met même en avant les mouvements physiologiques (figure 5.9b).

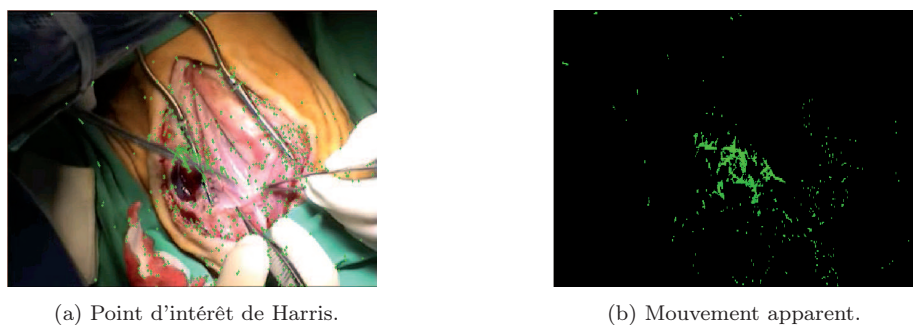


FIG. 5.8 – Mouvement détecté lors de l'utilisation d'un instrument.

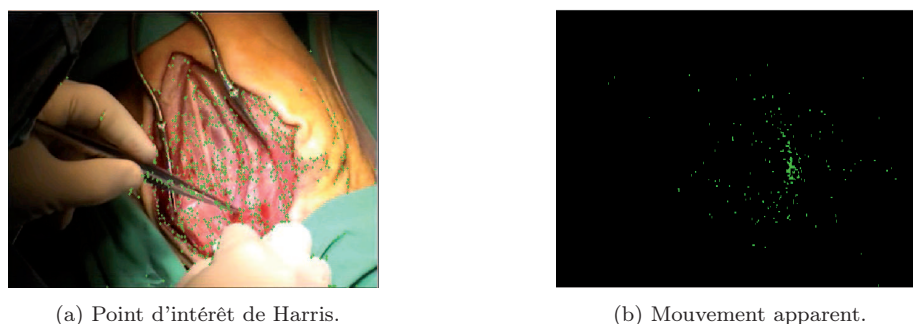


FIG. 5.9 – Pulsation d'une veine (aucun mouvement de la part du chirurgien).

Pour garantir un nombre suffisant de vecteurs, les points d'intérêt doivent être régénérés à intervalles de temps réguliers pour éviter la disparition progressive de l'ensemble des points en raison de la sortie de l'objet ou bien de leur non appariement entre deux images consécutives. Les tests ont montré qu'une régénération toutes les 10 images semble être la valeur la plus adéquate.

Comme précédemment, l'attribution des masses m_{i5} s'effectue en utilisant le modèle d'Appriou :

$$\begin{cases} m_{15}(\{\text{instrument}\}) = \beta_5(R_5 p(s_5 = N | \{\text{instruments}\})) \\ m_{25}(c \{\text{instrument}\}) = \beta_5(1 - R_5 p(s_5 = N | \{\text{instruments}\})) \\ m_{35}(\Theta) = 1 - \beta_5 \end{cases}$$

- $p(s_5 = N | \{\text{instruments}\})$, la valeur de la densité de probabilité conditionnelle de la source 5 (correspondant à l'attribut mouvement) lorsque le niveau de gris du pixel de coordonnées (k, l) de l'image I_M égal à N , sachant l'hypothèse $\{\text{instruments}\}$;
- R_5 , le coefficient permettant la normalisation de la masse. Considéré égal à 1 dans la mesure où la densité de probabilité conditionnelle $p(s_5 | \{\text{instruments}\})$ est normalisée;
- β_5 , le coefficient représentatif de la fiabilité de l'affectation (pris par défaut égal à 1);

- m_{15} , la croyance que le pixel appartient à un instrument ;
- m_{25} , la croyance que le pixel n'appartient pas à un instrument ;
- m_{35} , la masse représentative de l'ignorance.

La détermination de $p(s_5 = N | \{\text{instruments}\})$ se base sur un histogramme représentatif du mouvement. Pour ce faire, nous conditionnons au préalable le niveau de gris de chaque vecteur par rapport à sa norme. Ainsi un niveau de gris de 255 correspond à un déplacement de $\sqrt{n^2 + m^2}$ pixels (n et m étant les dimensions de l'image) et 0 correspond à un mouvement nul. Dès lors, par simple lecture des coordonnées chromatiques du pixel, il est possible d'accéder à la vitesse apparente de l'objet en ce point.

Nous cherchons par l'intermédiaire de l'histogramme mouvement à mettre à zéro les masses m_{i5} lorsque la norme du vecteur est soit faible soit forte, ceci afin de ne pas prendre en compte les deux types de mouvements spécifiques :

1. les mouvements de retrait exécutés par le chirurgien. Ces mouvements rapides correspondent généralement aux cas où le chirurgien se tourne et prend un autre instrument sur une table mitoyenne. Désigner la zone d'intérêt sur un élément en cours de retrait amènerait l'éclairage opératoire à suivre cet élément et donc à s'éloigner de la zone d'intérêt réelle, voire à sortir de la scène opératoire si les champs opératoires servant à la délimiter sont mal détectés. Ces mouvements rapides apparaissent également lors de mauvais appariements ;
2. les mouvements dits physiologiques générés par exemple par la respiration. Ce type de mouvement lent fausse l'analyse du système et ce dernier peut alors désigner une région inappropriée.

Dans la mesure où il n'est pas possible de discerner les gestes de très petites amplitudes des mouvements physiologiques, nous ne seuillons pas brutalement le début de l'histogramme. Nous intégrons progressivement les masses de m_{i5} afin de laisser la possibilité au système de désigner ces endroits s'il n'y a pas d'activité dans les autres régions de l'image (la probabilité pignistique étant plus forte dans ce cas), d'où la forme en "S" de la courbe (figure 5.10) qui s'apparente à une sigmoïde d'expression (forme de Verhulst) :

$$p(s_5 | \{\text{instruments}\}) = (1 + 2.10^4 \exp^{-0.1N})^{-1} \quad (5.1)$$

avec N , le niveau de gris du pixel considéré.

Cet histogramme permet de se restreindre aux mouvements ayant une vitesse supérieure à 1cm.s^{-1} et inférieure à 6cm.s^{-1} .

Ainsi comme le montre la figure 5.11, il y a élimination d'une partie des vecteurs pour ne conserver que les mouvements utiles. Nous pouvons également voir que les vecteurs assimilés pour des gestes rapides sont dus à de mauvais appariements. La preuve de la reconnaissance des gestes rapides est donnée en figure 5.11d.

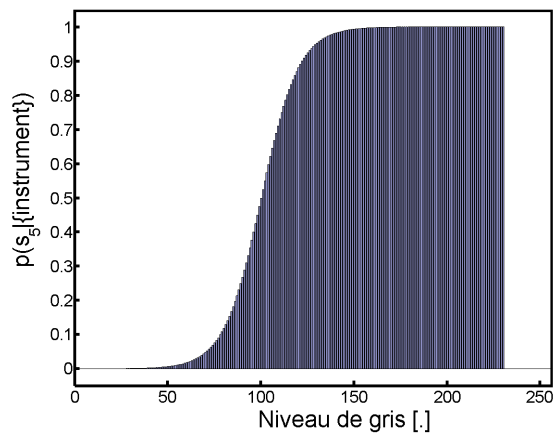
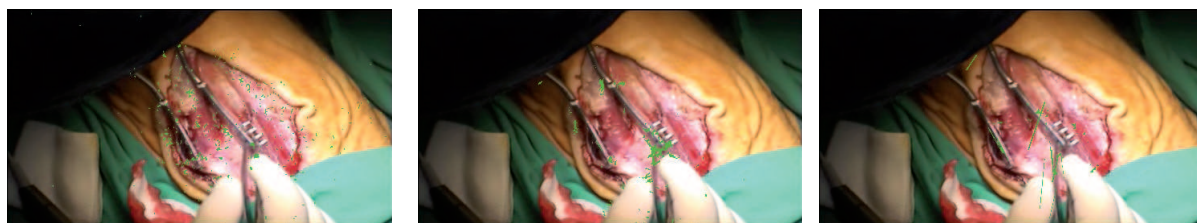
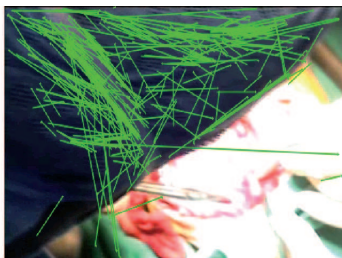


FIG. 5.10 – Histogramme de l'attribut mouvement.



(a) Vecteurs de norme inférieure à 1cm.s^{-1} .
 (b) Vecteurs de norme comprise entre 1cm.s^{-1} et 6cm.s^{-1}
 (c) Vecteurs de norme supérieure à 6cm.s^{-1} .



(d) Vecteurs de norme supérieure à 6cm.s^{-1} dans le cas d'une occultation.

FIG. 5.11 – Isolation des mouvements de fortes amplitudes (retrait de la tête du chirurgien et mauvais appariement).

Fusion des attributs

Maintenant que les masses correspondant à la forme et au mouvement sont créées, il faut les fusionner à l'aide d'une règle de combinaison telle que celles présentées en chapitre 3. Afin de faciliter la comparaison entre les performances des règles conjonctive, disjonctive, de Smets et PCR#5, nous utilisons l'analyse de la probabilité pignistique moyenne de la proposition {instrument} calculée pour l'ensemble de l'image. Nous ne nous attachons qu'aux fluctuations (dont le chirurgien est responsable) et non à la valeur même de $\overline{BetP}(\{\text{instrument}\})$ (correspondant à la probabilité pignistique de la proposition {instrument}) qui est nécessairement plus faible que $BetP(\{\text{instrument}\})$ calculée pour un pixel appartenant à la classe instrument. Nous présentons en figure 5.12 les résultats obtenus lors d'une séquence d'essais en conditions réelles.

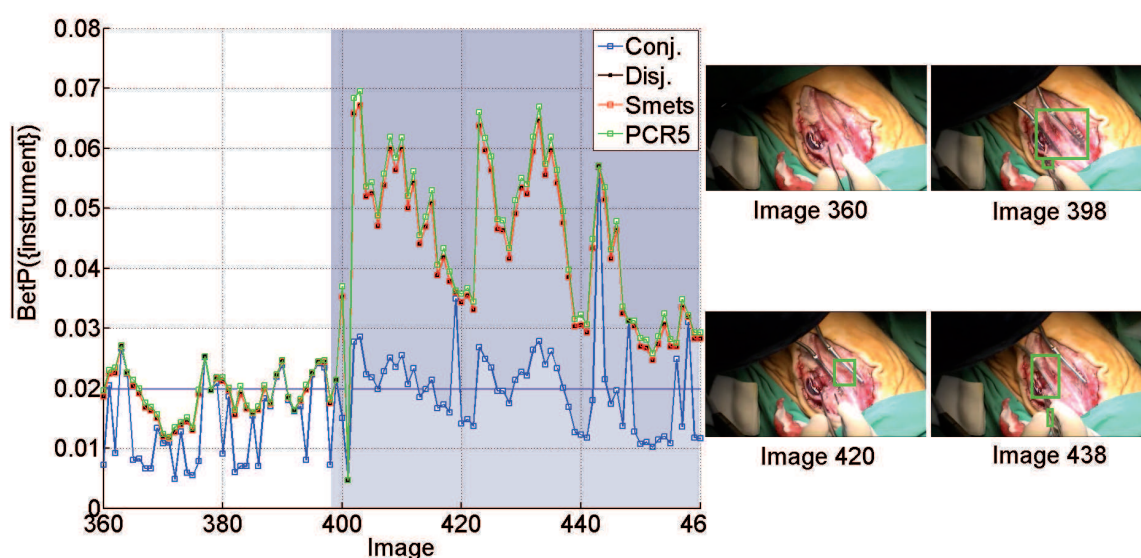


FIG. 5.12 – Comparaison des différentes règles de combinaison pour la classe instrument pour le cas d'une pose d'écarteur.

La règle conjonctive adopte un comportement pessimiste (obligation d'avoir un grand nombre d'informations se recoupant) ce qui se retranscrit par une moyenne basse de $\overline{BetP}(\{\text{instrument}\})$ comparativement aux autres règles plus optimistes. Nous constatons également sur cette séquence une variation brutale de \overline{BetP} pour les règles disjonctives, de Smets et PCR#5. Jusqu'à l'image 397, les mains du chirurgien sont visibles mais ne participent à aucune activité (examen de la plaie). Il n'y a donc pas d'instrument "actif", la réponse du système à cette sollicitation est donc adéquate car aucun rectangle symbolisant la présence d'instrument n'est visible (ce sont ces rectangles qui nous seront ultérieurement utiles pour la désignation de la zone d'intérêt). La probabilité pignistique moyenne est non nulle du fait de l'intégration des amplitudes faibles dans l'histogramme représentatif du mouvement. A partir de l'image 398 commence la pose d'un écarteur qui est interprétée correctement par le système comme une activité à part entière du chirurgien. Le nombre de pixels assimilés à des instruments actifs augmente et avec lui les rectangles représentatifs des lieux où se trouvent les instruments "actifs". Les fluctuations observées pour cette période découlent des variations de vitesse apparente de sorte que plus le chirurgien exécute son

geste rapidement (sans toutefois dépasser la limite des $6\text{cm}\cdot\text{s}^{-1}$), plus grandes sont les chances d'avoir $p(s_5|\text{instrument})$ proche de 1 et donc d'avoir $\overline{BetP}(\{\text{instrument}\})$ élevée.

La figure 5.13 expose une deuxième séquence tirée de la même opération. Outre une période stable sur l'intervalle $[940, 1000]$ correspondant à une période d'inactivité du chirurgien, nous observons deux pics. Ces augmentations de $\overline{BetP}(\{\text{instrument}\})$ nous montrent qu'il y a accroissement des pixels considérés comme appartenant à la classe instrument. En effet, sur les intervalles $[1001, 1016]$ et $[1030, 1040]$ le chirurgien procède à une dissection (ouverture et fermeture des ciseaux). Nous pouvons voir que sur ces deux intervalles, les boîtes englobantes cernant la région de l'image où les ciseaux sont en action sont correctement localisées malgré la présence de la main du chirurgien assistant.

Pour compléter cette analyse des performances de la reconnaissance de la classe instrument, nous

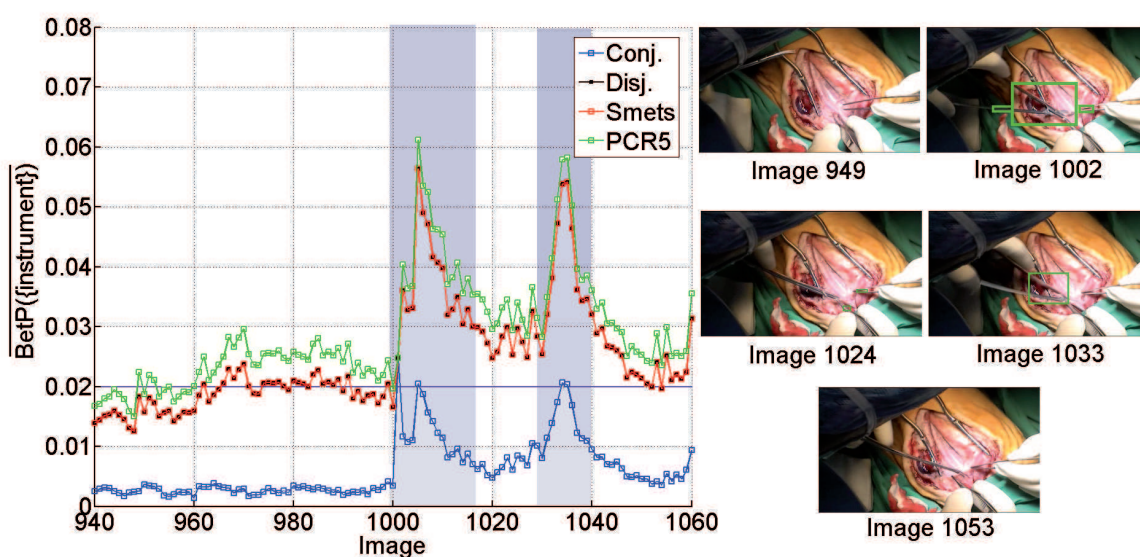


FIG. 5.13 – Comparaison des différentes règles de combinaison dans la classe instrument pour le cas d'une dissection.

proposons en figures 5.14a et 5.14b, deux autres séquences. La première expose le cas où il n'y a aucun mouvement en dehors des mouvements physiologiques (pulsation de l'artère, respiration). Nous assistons alors à une valeur stable de la probabilité pignistique en dessous de la barre symbolique des 0.02. Le second cas proposé illustre une activité du chirurgien qui est interrompue par une palpation sur la période $[2461, 2469]$ où le chirurgien principal arrête son geste et où la main du chirurgien assistant vient tirer la peau du sujet. Le système interprète correctement cette action et n'intègre pas la main comme instrument potentiel ceci afin d'éviter ultérieurement une mauvaise désignation. La valeur de $\overline{BetP}(\{\text{instrument}\})$ chute donc à un niveau assimilable à de l'inactivité.

Dans la mesure où la fusion conjonctive ne met pas en valeur les phases d'activité du chirurgien, nous prenons le parti de privilégier les autres règles de combinaison. Le choix de la règle utilisée pour la fusion des masses pour la classe instrument ainsi que pour les classes sang, peau et champs stériles se base sur

la gestion du conflit qui est une perspective de ces travaux. En conséquence, la règle PCR#5 sera la règle utilisée pour l'ensemble des classes d'objets.

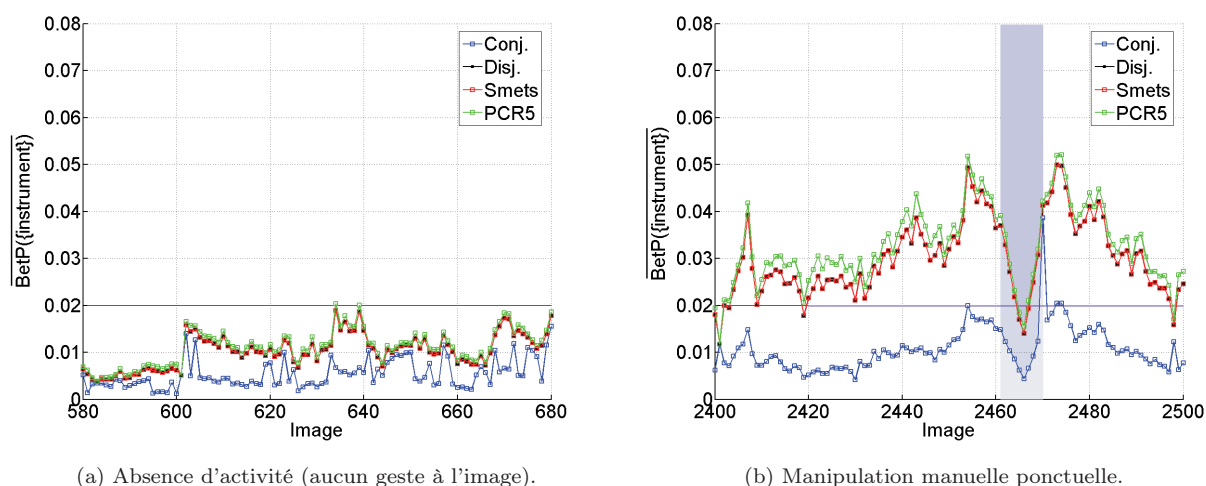


FIG. 5.14 – Autres séquences issues d'essais en bloc opératoire.

En résumé :

La classe instrument est correctement reconnue grâce à la fusion des attributs de forme et de mouvement. L'attribut forme s'appuie sur une transformée de Hough, nécessitant peu de calculs. Un filtre passe-haut a été ajouté afin de rendre plus robuste la détection et améliorer la précision. La combinaison des attributs forme et mouvement permet d'éliminer les instruments "inactifs". Grâce à un choix adéquat des paramètres de création des masses m_{i5} (attribut mouvement), le système est capable d'éliminer les mouvements rapides et les mouvements trop lents. La règle de fusion (PCR#5) se révèle être la règle la mieux adaptée à notre étude. Les performances obtenues dans des situations particulières (passage d'une période d'inactivité à une période d'activité, intermittence rapide de périodes inactives et actives, absence totale d'activité avec présence de mouvements physiologiques, interruption brutale d'activité) illustrent l'efficacité de la reconnaissance de la classe instrument malgré les perturbations et les contraintes.

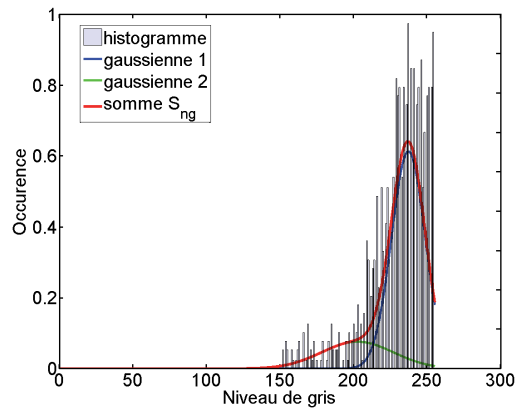
5.2.2 Reconnaissance des autres classes

Contrairement à la classe instrument, les classes sang, peau et champs opératoires peuvent être reconnues grâce à l'attribut couleur. Les attributs forme et mouvement ne sont pas utilisés pour ces classes.

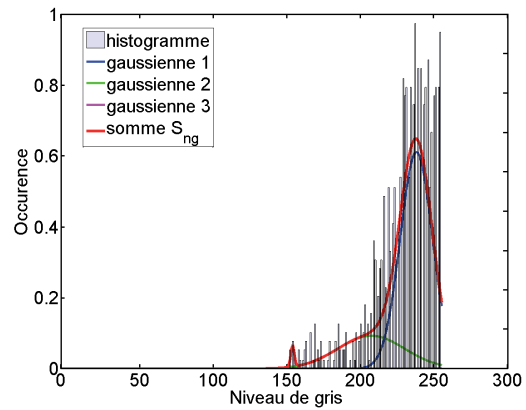
Attribut couleur

L'attribut couleur permet l'allocation des masses m_{ij} , $j \in 1, 2, 3$. Tout comme pour les autres attributs, nous utilisons une connaissance a priori sous forme de trois histogrammes. Chacun d'entre eux est représentatif de la classe pour un canal considéré (H, S ou V). A la différence de l'attribut forme, la création de ces histogrammes est plus délicate car il ne s'agit pas de relation binaire. Pour ce faire, nous utilisons un fragment d'image représentatif de chacune des classes pris au cours d'essais précliniques. Nous en

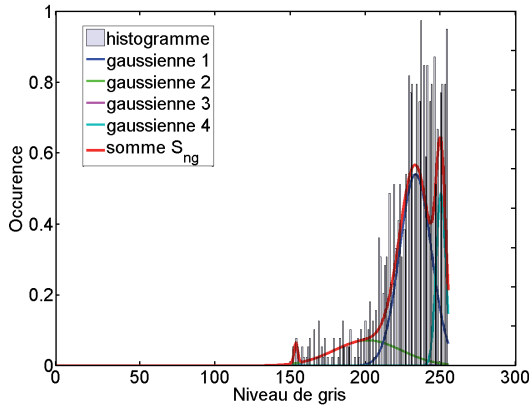
extrayons l'histogramme pour chacune des couches puis nous modélisons ceux-ci par deux gaussiennes (sauf dans le cas de la graisse et de la peau qui sont pour la couche H modélisables facilement par une seule gaussienne) à l'aide de l'algorithme EM afin d'être laxe et ainsi lisser les perturbations locales (bruit, réflexions spéculaires, etc). Un autre avantage de la modélisation par mélange de gaussiennes est la facilité d'exportation de ces modèles ; un modèle d'une classe étant synthétisé par un vecteur de taille $2 \times 2 \times 3$ composantes contre $255 \times 1 \times 3$ dans le cas d'histogramme complet. Comme le montrent



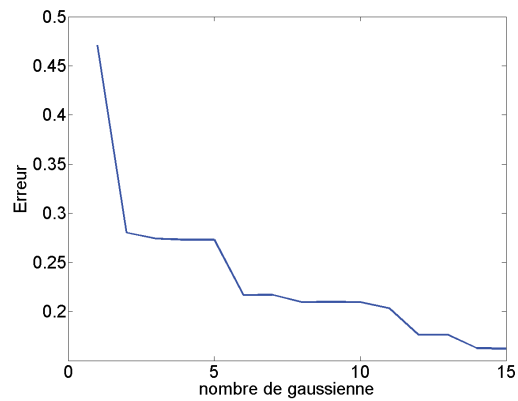
(a) Modélisation par deux gaussiennes.



(b) Modélisation par trois gaussiennes.



(c) Modélisation par quatre gaussiennes.



(d) Influence du nombre de gaussiennes dans l'erreur de modélisation.

FIG. 5.15 – Modélisation de l'histogramme de la classe sang pour le canal H.

les figures 5.15a, 5.15b et 5.15c, plus le nombre de gaussiennes est important et plus la modélisation (courbe en rouge) est proche de la distribution réelle. Cependant, au-delà d'un certain point, la contribution de chaque gaussienne devient mineure comme le montre la figure 5.15d qui présente le résultat de $\sum_{i=1}^{255} |S_{n_G}(i) - \mathbf{p}(i)|$, avec S_{n_G} la somme des n_G gaussiennes et \mathbf{p} la distribution réelle. Ainsi, les paramètres estimés grâce à l'algorithme EM sont la moyenne (μ_i) et l'écart-type (σ_i) (les coefficients π_i n'étant utilisés que dans la construction du modèle) synthétisés dans le tableau 5.4 pour les différentes classes d'objets (nous utilisons la convention d'écriture $\begin{pmatrix} \mu_1 & \sigma_1 \\ \mu_2 & \sigma_2 \end{pmatrix}$).

Classe d'objet	Modélisation de la classe d'objets pour les couches H, S et V					
Sang	$\begin{pmatrix} 168.81 & 3.9338 \\ 154.54 & 4.9723 \end{pmatrix}$		$\begin{pmatrix} 121.91 & 23.905 \\ 196.1 & 30.338 \end{pmatrix}$		$\begin{pmatrix} 111.04 & 12.406 \\ 193.63 & 37.202 \end{pmatrix}$	
Ch. stériles	$\begin{pmatrix} 64.762 & 1.6851 \\ 53.924 & 8.9259 \end{pmatrix}$		$\begin{pmatrix} 99.961 & 4.4614 \\ 85.517 & 10.665 \end{pmatrix}$		$\begin{pmatrix} 46.15 & 9.205 \\ 53.1 & 0.81545 \end{pmatrix}$	
Peau	$\begin{pmatrix} 21.838 & 0.94987 \end{pmatrix}$		$\begin{pmatrix} 246.53 & 4.8953 \\ 230.81 & 7.4255 \end{pmatrix}$		$\begin{pmatrix} 236.34 & 12.03 \\ 196.63 & 23.398 \end{pmatrix}$	

TAB. 5.4 – Résumé des modélisations des classes d'objets pour les couche H, S et V.

L'allocation des fonctions de croyance se base donc sur ces modélisations. Ainsi pour le premier canal (H), la valeur affectée à la masse m_{i1} pour le premier pixel est définie pour la classe θ_k ($k \in \{\text{sang}\}, \{\text{peau}\}, \{\text{champ stérile}\}$) par :

$$\begin{cases} m_{11}(\{\theta_k\}) = \beta_1(R_1 p(s_1 | \{\theta_k\})) \\ m_{21}(c \{\theta_k\}) = \beta_1(1 - R_1 p(s_1 | \{\theta_k\})) \\ m_{31}(\Theta) = 1 - \beta_1 \end{cases}$$

- $p(s_1 | \{\theta_k\})$, densité de probabilité de la source 1 (correspondant à l'attribut couleur) pour l'hypothèse $\{\theta_k\}$;
- R_1 : coefficient permettant la normalisation de la masse. Considéré égal à 1 dans la mesure où la densité de probabilité $p(s_1 | \{\theta_k\})$ est normalisée;
- β_1 , coefficient représentatif de la fiabilité de l'affectation (pris par défaut égal à 1);
- m_{11} , croyance que le pixel appartient à la classe $\{\theta_k\}$;
- m_{21} , croyance que le pixel n'appartient pas à la classe $\{\theta_k\}$;
- m_{31} , représentative de l'ignorance.

En répétant cette procédure pour l'ensemble des pixels, il est possible d'obtenir une image représentative de chaque proposition pour ce canal (figure 5.16). Nous pouvons voir que la reconnaissance du sang n'est pas à la hauteur de nos attentes dans la mesure où peu de pixels appartenant à cette classe sont détectés. La matrice de confusion présentée en table 5.5 illustre les performances liées au fragment choisi. Cette dépendance au fragment sélectionné est confirmée par la figure 5.16c où le fragment choisi pour la classe peau permet cette fois une reconnaissance correcte.

	sang ¹	non sang ¹
sang ²	62.9%	0.6%
non sang ²	37.1%	99.4%

TAB. 5.5 – Matrice de confusion pour la classe sang avec le fragment sélectionné. "1" correspond à la vérité terrain, "2" à celle du système.

Nous pouvons voir que l'inhomogénéité d'exposition affecte la reconnaissance. La qualité du fragment

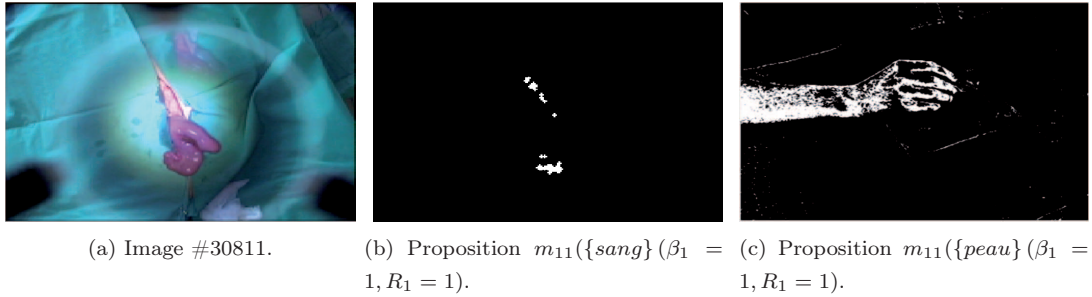


FIG. 5.16 – Exemple de reconnaissance grâce à l’attribut couleur.

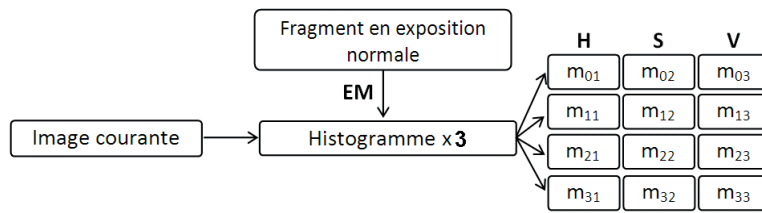
sélectionné pour la modélisation de la classe a un très fort impact sur les masses. En conséquence, nous remplaçons la reconnaissance basée sur un fragment (trois histogrammes) par une reconnaissance formée par trois fragments (neuf histogrammes). Ceux-ci sont représentatifs des régions sombres, en sur-exposition et normalement exposées de sorte à couvrir l’ensemble des différentes expositions.

Nous passons donc, pour une classe donnée, d’un système à neuf masses à un système à vingt-sept masses, que nous appellerons masses temporaires notées $m'_{i1}(\{\theta_k\})$. Il s’agit ensuite de les combiner pour n’en obtenir que trois en sortie de processus et ainsi permettre d’exécuter une fusion par règle PCR#5 pour attribuer au pixel la valeur de $BetP(\{\theta_k\})$.

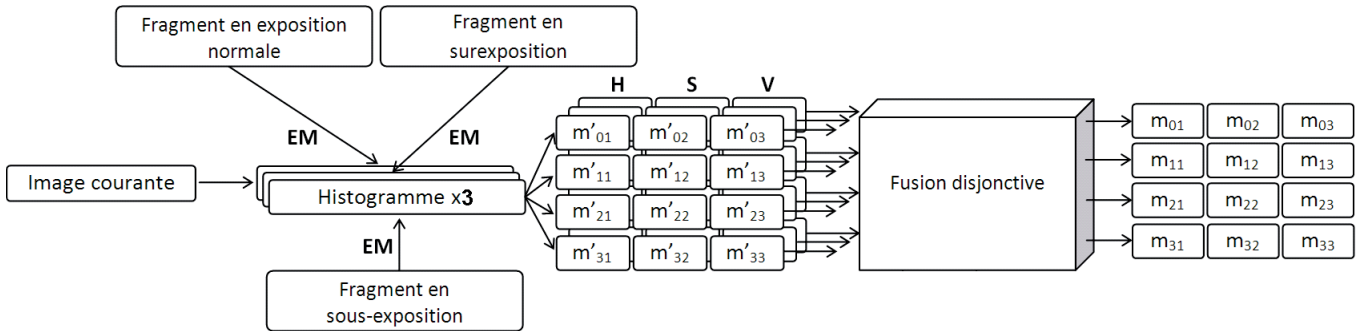
Pour intégrer ces masses temporaires m'_{i1} , nous utilisons une fusion disjonctive telle que définie en (3.19). Celle-ci agit à la manière d’un OU logique entre les images représentatives de la classe pour chaque type d’exposition. L’illustration 5.17 présente une vue globale de la méthode.

Nous proposons en figure 5.18 les résultats de cette méthode (en référence à l’image 5.16a). Nous pouvons y voir que la contribution du fragment normalement exposé est majoritaire dans la reconnaissance de la classe sang. Il serait donc judicieux de ne garder que ce type d’exposition. Cependant, compte-tenu de la variabilité des interventions chirurgicales, il est souhaitable de conserver les expositions faibles notamment pour pallier les effets de l’absorption par les tissus lors d’intervention en cavité. Pour cette étude nous prenons le parti de garder également la reconnaissance des classes en sous-exposition forte bien que son apport soit mineur.

La matrice de confusion présentée en tableau 5.6 montre que cette prise en compte des différentes expositions améliore la reconnaissance notamment en termes de faux négatifs (correspondant aux pixels appartenant à la classe d’objets manqués par le système) et de vrais positifs (correspondant aux bonnes reconnaissances). Nous constatons que parallèlement à cette amélioration des performances de la reconnaissance (augmentation par 8 du taux de vrais positifs et diminution par 48 de celui des faux négatifs), il y a augmentation des faux positifs (détections erronées). Ceci s’explique par l’intégration des bruits de chaque reconnaissance pour les expositions faibles, fortes et normales. La précision obtenue de 0.9504 montre le peu d’erreurs commises pour cette reconnaissance. Le rappel, synonyme d’erreur de désignation (et potentiellement de déplacements erronés), est quant à lui de 0.9985.

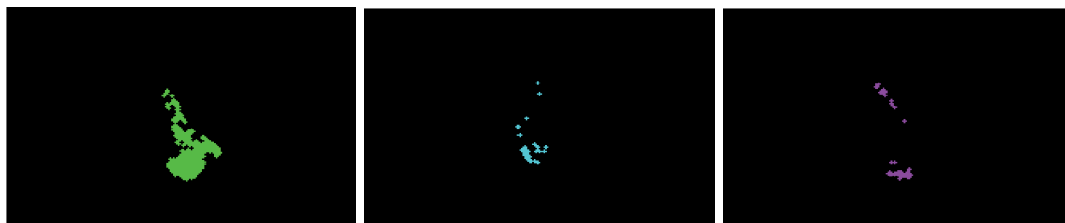


(a) Schéma de création des masses.

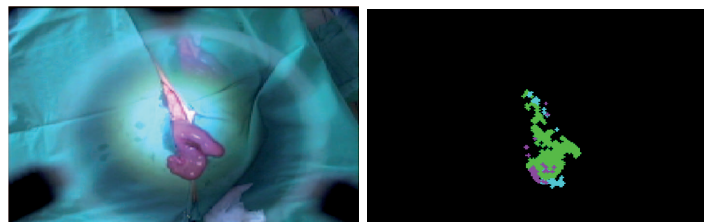


(b) Nouveau schéma intégrant une fusion disjointive.

FIG. 5.17 – Modification du schéma de création des masses.



(a) Classe sang normalement expo- (b) Classe sang en sur-exposition. (c) Classe sang en sous-exposition. sée.



(d) Image #30811.

(e) Résultat en sortie de fusion disjointive.

FIG. 5.18 – Exemple de reconnaissance de la classe sang pour différentes expositions.

sans fusion	sang ¹	non sang ¹	avec fusion	sang ¹	non sang ¹
sang ²	62.9%	0.6%	sang ²	95.0%	< 10 ⁻² %
non sang ²	37.1%	99.4%	non sang ²	5%	99.9%

TAB. 5.6 – Comparaison des matrices de confusion pour la classe sang (éclairage au niveau maximal-avec filtre apodisant) sans et avec fusion disjonctive.

Nous venons de voir que l'exposition a une influence sur la reconnaissance de la classe sang et qu'il faut pour y pallier ajouter au processus une étape de fusion disjonctive. Cependant, ces tests ne sont représentatifs que d'un point de fonctionnement du système. Il faut maintenant vérifier que ces performances sont également valables pour l'ensemble de la plage d'éclairage ; l'équipe chirurgicale pouvant modifier son niveau en cours d'intervention. Cette procédure a donc été menée pour trois niveaux d'éclairage, le tableau (5.7) présente les résultats obtenus.

Remarque : *Nous ne menons une étude de sensibilité que pour l'attribut couleur et non pour la forme ou le mouvement car ces derniers sont basés sur une étude soit du gradient soit sur une analyse multi-échelle. Dans les deux cas ils s'avèrent peu impactés par les variations de l'éclairage.*

Sans filtre	Sans éclairage	Eclairage à 36klux	Eclairage à 110klux
Rappel	$0.069 \pm 3.63 \cdot 10^{-2}$	$0.750 \pm 3.7 \cdot 10^{-2}$	$0.757 \pm 2.06 \cdot 10^{-2}$
Précision	$0.994 \pm 2.6 \cdot 10^{-3}$	$0.982 \pm 6.4 \cdot 10^{-3}$	$0.977 \pm 1.34 \cdot 10^{-2}$

Avec filtre	Sans éclairage	Eclairage à 36klux	Eclairage à 110klux
Rappel	$0.0922 \pm 1.15 \cdot 10^{-2}$	$0.994 \pm 3.4 \cdot 10^{-3}$	$0.998 \pm 1.8 \cdot 10^{-3}$
Précision	$0.995 \pm 1.3 \cdot 10^{-3}$	$0.965 \pm 9.6 \cdot 10^{-3}$	$0.957 \pm 5.3 \cdot 10^{-3}$

TAB. 5.7 – Calcul du rappel et de la précision pour différentes configurations d'éclairage pour la classe sang.

Remarque : *Il est cohérent de trouver que l'ajout du filtre n'a aucune incidence sur la précision et le rappel dans le cas où l'éclairage est éteint car il n'y a pas d'inhomogénéité d'éclairage à compenser.*

Il ressort de ce tableau que, pour la classe sang, la qualité de la reconnaissance est peu influencée par l'éclairage. Ceci est rendu possible grâce à l'utilisation conjointe d'un mélange de gaussiennes et de l'espace HSV qui permet d'absorber les variations de coordonnées chromatiques. Nous pouvons également constater que l'utilisation d'un filtre apodisant permet une diminution significative des taux de faux négatifs (intervenant dans l'expression du rappel), ce qui se retrouve dans les distances calculées pour ces deux configurations de test (tableau 5.8).

Sang				
	Sans filtre	Sans écl.	Ecl. à 36klux	Ecl. à 110klux
	Distance	$0.930 \pm 3.6.10^{-2}$	$0.251 \pm 3.6.10^{-2}$	$0.244 \pm 2.1.10^{-2}$
	Avec filtre	Sans écl.	Ecl. à 36klux	Ecl. à 110klux
	Rappel	$0.0922 \pm 1.15.10^{-2}$	$0.994 \pm 3.4.10^{-3}$	$0.998 \pm 1.8.10^{-3}$
	Précision	$0.995 \pm 1.3.10^{-3}$	$0.965 \pm 9.6.10^{-3}$	$0.957 \pm 5.3.10^{-3}$
	Distance	$0.908 \pm 1.15.10^{-2}$	$0.035 \pm 9.0.10^{-3}$	$0.042 \pm 5.3.10^{-3}$
Accroissement des performances à l'aide d'un filtre apodisant		2.47%	85.94%	82.60%

TAB. 5.8 – Grandeurs caractéristiques pour la classe sang en fonction de l'éclairement.

Champs opératoires				
	Avec filtre	Sans écl.	Ecl. à 36klux	Ecl. à 110klux
	Rappel	$0.680 \pm 3.1.10^{-3}$	$0.506 \pm 4.7.10^{-3}$	$0.377 \pm 3.4.10^{-3}$
	Précision	$0.649 \pm 3.2.10^{-3}$	$0.606 \pm 5.3.10^{-3}$	$0.527 \pm 9.6.10^{-3}$
	Distance	$0.475 \pm 1.6.10^{-3}$	$0.632 \pm 1.8.10^{-3}$	$0.783 \pm 5.18.10^{-2}$

TAB. 5.9 – Grandeurs caractéristiques pour la classe champs opératoires en fonction de l'éclairement.

Peau				
	Avec filtre	Sans écl.	Ecl. à 36klux	Ecl. à 110klux
	Rappel	$0.387 \pm 1.25.10^{-2}$	0.629 ± 9.10^{-3}	$0.726 \pm 1.59.10^{-2}$
	Précision	$0.865 \pm 1.6.10^{-3}$	$0.588 \pm 5.2.10^{-3}$	$0.752 \pm 1.65.10^{-2}$
	Distance	$0.628 \pm 1.2.10^{-3}$	$0.555 \pm 9.4.10^{-3}$	$0.370 \pm 1.53.10^{-2}$

TAB. 5.10 – Grandeurs caractéristiques pour la classe peau en fonction de l'éclairement.

A la lecture du tableau 5.8, nous concluons que les meilleurs résultats sont obtenus pour un éclairage minimal avec l'utilisation d'un filtre apodisant. Il faut maintenant généraliser ces résultats aux autres classes d'objets : peau et champs stériles.

Le tableau 5.9 montre que les performances pour la classe champs opératoires augmentent lorsque l'éclairage diminue. Ces résultats sont à l'encontre de ceux obtenus pour la classe peau, où les performances maximum (autrement dit la distance minimum) sont pour un éclairage de 110klux. Compte-tenu que pour les classes sang, peau et champs stériles, les distances minimales ne sont pas trouvées pour les mêmes éclairages, il faut donc définir quel éclairage, aux vues des tableaux 5.8, 5.9 et 5.10, est le meilleur compromis.

Au regard du tableau 5.11 contenant les distances moyennes (obtenues avec un filtre apodisant) et compte-tenu que l'éclairage ne peut être éteint, les meilleures performances sont obtenues avec le maximum d'éclairage. Nous nous placerons donc dans cette configuration pour la suite des essais.

	Sans éclairage	Eclairage à 36klux	Eclairage à 110klux
Distance moyenne	0.318 ± 0.51	0.63 ± 0.15	0.518 ± 0.13

TAB. 5.11 – Distance fonction de l'éclairage.

Les performances entre les différentes reconnaissances n'étant pas égales, nous intégrons ces disparités lors de la création des masses. Ainsi pour chaque source (forme, mouvement, couleur), nous attribuerons au coefficient β_j (affaiblissement des masses représentatives de la confiance en la reconnaissance) la valeur de sa distance obtenue pour l'éclairage maximum. Nous obtenons donc :

- pour la classe instrument : $\beta_4 = \beta_5 = 1 - \frac{1}{\sqrt{2}}0.803$;
- pour la classe sang : $\beta_1 = \beta_2 = \beta_3 = 1 - \frac{1}{\sqrt{2}}0.042$;
- pour la classe champs opératoires : $\beta_1 = \beta_2 = \beta_3 = 1 - \frac{1}{\sqrt{2}}0.783$;
- pour la classe peau : $\beta_1 = \beta_2 = \beta_3 = 1 - \frac{1}{\sqrt{2}}0.370$.

En résumé :

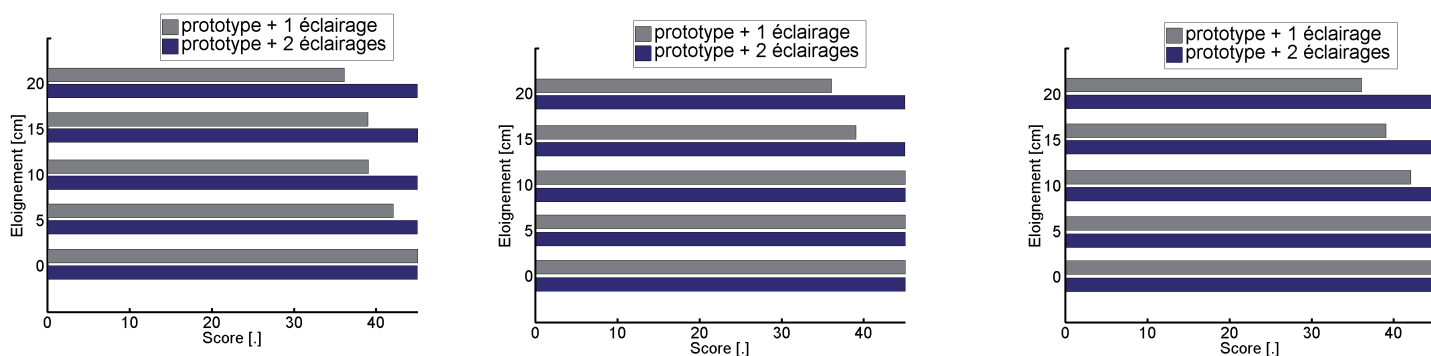
Nous venons de voir comment reconnaître les objets en présence dans la scène opératoire ainsi que les performances associées. L'éclairage joue un rôle non négligeable sur les résultats et affecte non uniformément les classes d'objets. Suite au comparatif des performances de chaque classe pour différents niveaux d'éclairage, nous pouvons conclure que le niveau maximal est préférable. Ce choix s'est basé sur la mesure de distance moyenne qui est un bon indicateur du comportement des méthodes de reconnaissance. Nous prendrons en compte les disparités de reconnaissance dans les affaiblissements des masses.

Ajout de coupoles supplémentaires

Afin de compléter cette étude, nous nous intéressons au comportement de la reconnaissance en présence d'éclairages opératoires supplémentaires pour recréer l'éclairage réel rencontré lors d'interventions chirurgicales (deux éclairages opératoires obligatoires au minimum pour les chirurgies majeures). Pour ce faire nous utilisons un test qui consiste à discerner neuf cibles réparties dans le champ de vision de la caméra. Celui-ci s'articule en deux étapes. Dans un premier temps, les taches lumineuses de toutes les coupoles sont confondues, puis dans un second temps nous éloignons volontairement la tache de la coupole additionnelle par pas de 5cm jusqu'à une limite de 20cm afin d'évaluer le comportement de la méthode face à de mauvais prépositionnements de l'éclairage. Nous recommençons ce test en modifiant le niveau d'éclairage de la maquette ainsi que celui de la coupole additionnelle. Trois niveaux d'éclairage sont utilisés pour chacune des coupoles.

Pour chacune des cibles, nous attribuons une note comprise entre 0 et 5 : 0 étant l'absence de dé-

tection et 5 la reconnaissance totale de la cible.



(a) Score obtenu pour un éclairage de 30klux. (b) Score obtenu pour un éclairage de 70klux. (c) Score obtenu pour un éclairage de 110klux.

FIG. 5.19 – Influence d'un ou deux éclairages opératoires supplémentaires sur la reconnaissance.

Les résultats de ces tests (figure 5.19) montrent que l'éclairage de la coupole additionnelle affecte la reconnaissance. Les performances sont directement proportionnelles à la distance séparant les deux taches lumineuses. Ainsi, à mesure que l'écart se réduit, le taux de détection s'améliore. Il est donc intéressant pour les futurs tests animaliers d'utiliser deux éclairages opératoires dont les taches lumineuses sont les plus proches possibles quel que soit le niveau d'éclairage de la coupole additionnelle. Cette conclusion est également valable lorsque l'on ajoute une troisième coupole aux deux précédemment installées (résultats en bleu sur la figure 5.19).

En résumé :

Nous avons présenté dans cette partie la reconnaissance des objets dans la scène chirurgicale et montré, à travers des tests sur des images prises dans des conditions réelles, l'utilité du filtre apodisant. Nous pouvons conclure d'après ces résultats que la méthode utilisée est suffisamment fiable et précise pour la classe sang mais que des efforts sont encore à fournir pour les classes champs opératoires et peau. La robustesse du traitement a été testée dans des conditions d'éclairage variées voire extrêmes. Son bon comportement quelles que soient les conditions nous permet donc d'être confiants quant à la réussite de la désignation. Nous avons vu dans cette partie les performances de la stratégie de fusion PCR#5 pour la reconnaissance des objets en présence dans la scène chirurgicale.

Remarque : *La reconnaissance de l'ensemble des classes d'objets dans l'image (peau, sang, champs stériles et instrument) n'englobe pas l'intégralité des éléments présents dans la scène chirurgicale car n'appartenant pas au cadre de discernement. Comme souligné dans le chapitre 2, il est possible de créer une classe supplémentaire correspondant à ces éléments (et ainsi répartir le conflit sur cette classe on parle alors de hedging). Ceci peut être particulièrement profitable si la région correspond à un élément d'intérêt pour le chirurgien comme une prothèse. L'importance à donner à cette nouvelle classe dans le processus de décision reste une question ouverte et n'a pas été traitée dans cette étude.*

5.2.3 Désignation de la zone d'intérêt

Nous nous intéressons maintenant à partir de la reconnaissance des différentes classes d'objets à créer dans un premier temps des boîtes englobantes qui nous serviront à déterminer les zones d'intérêt potentielles. Nous avons vu dans le chapitre 3 que ce recouvrement s'effectue à l'aide d'une somme pondérée des images représentatives des classes d'objets dont on rappelle ici l'expression :

$$\mathcal{E}(i, j) = \left[\alpha_1 \frac{1}{2} I_{\text{sang}}(i, j) + (1 - \alpha_1) \frac{1}{2} I_{\text{instrument}}(i, j) \right] \mathbb{I}_{\text{champ opératoire}}(i, j) \cdot \mathbb{I}_{\text{peau}}(i, j)$$

$$I(i, j) = \begin{cases} 1 & \text{si } \mathcal{E}(i, j) = 255 \\ 0 & \text{sinon} \end{cases} \quad (5.2)$$

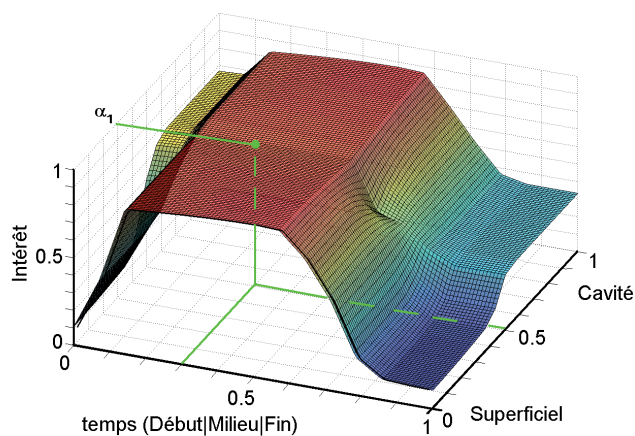
La quantification du gain α_1 passe par la détermination de la situation opératoire (superficielle ou cavité) et de l'instant opératoire (début, milieu ou fin d'intervention). Pour ce faire, nous utilisons le nombre de pixels reconnus comme du sang pour évaluer la situation opératoire. De manière schématique, nous supposons que plus il y a de sang dans l'image, plus la situation opératoire tend vers une cavité. A contrario, moins il y a de sang et plus l'intervention est superficielle. Pour le temps opératoire, aucun indicateur pertinent n'a été trouvé. Nous avons donc considéré que la durée de l'intervention est fixe et n'excède pas 90 minutes. Ceci suppose donc que l'intervention se déroule sans encombre et qu'aucun retour en arrière n'est effectué.

Nous injectons ces deux paramètres dans le graphe représentatif du gain (ou plus à proprement parler de l'utilité du gain) défini par règles floues (figure 5.20a). L'altitude du point nous donne le gain α_1 . Plus celui-ci est élevé et plus cette classe sera révélatrice de la zone d'intérêt. Pour réduire l'incertitude sur le gain, nous choisissons de n'étudier que certaines trajectoires que peut prendre l'intervention chirurgicale. Nous distinguons ainsi les actes $\{S, S, S\}$, $\{S, S, C\}$, $\{S, C, S\}$, ..., $\{C, C, C\}$ comme le montre les figures 5.20b et 5.20c.

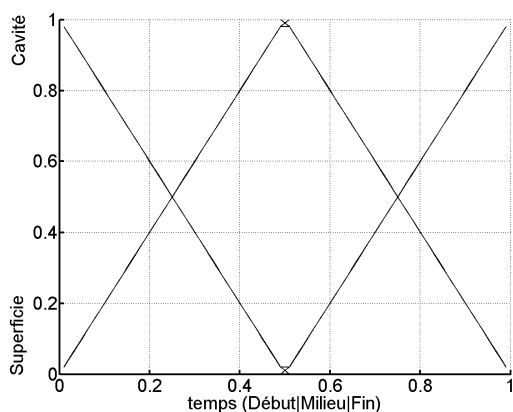
Ainsi comme le montre la figure 5.21a, le nombre de pixels évolue aux grès de l'activité du chirurgien. En outre dans cette figure, le temps a été normalisé. De même, connaissant le nombre maximum de pixels pouvant être détectés, la quantité de pixels reconnus comme du sang a été également normalisée. Ces étapes permettent d'accéder directement par simple lecture sur la figure 5.20a au coefficient α_1 .

Sur la figure 5.21b est représentée en bleue la séquence suivie par le chirurgien. Les cercles rouges indiquent la séquence déterminée par le système à partir de la détection du sang (figure 5.21a). Nous pouvons voir que le système considère à tort en début d'opération que l'intervention se déroule en superficie, puis corrige son erreur et vient suivre la bonne trajectoire. La fin de l'intervention se partage à part égale entre situation en superficie et en cavité ce qui explique les sauts brusques entre situation superficielle et cavité.

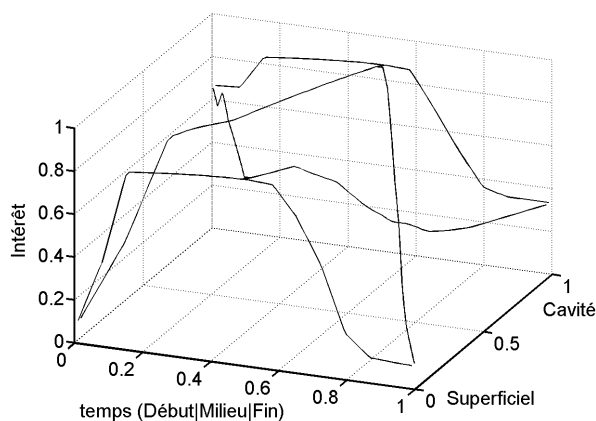
Dans le cas où la séquence opératoire est connue, il est alors possible de contraindre le système à suivre une trajectoire précise. Nous présentons en figure 5.22 le cas d'une séquence $\{S, C, S\}$. Toutefois, dans cette étude nous ne disposons pas de cette connaissance en conséquence nous supposons pour le reste de ce manuscrit que toutes les trajectoires sont accessibles.



(a) Détermination de α_i pour une situation opératoire et un instant opératoire quelconque.



(b) Trajectoires les plus probables (vue de dessus).

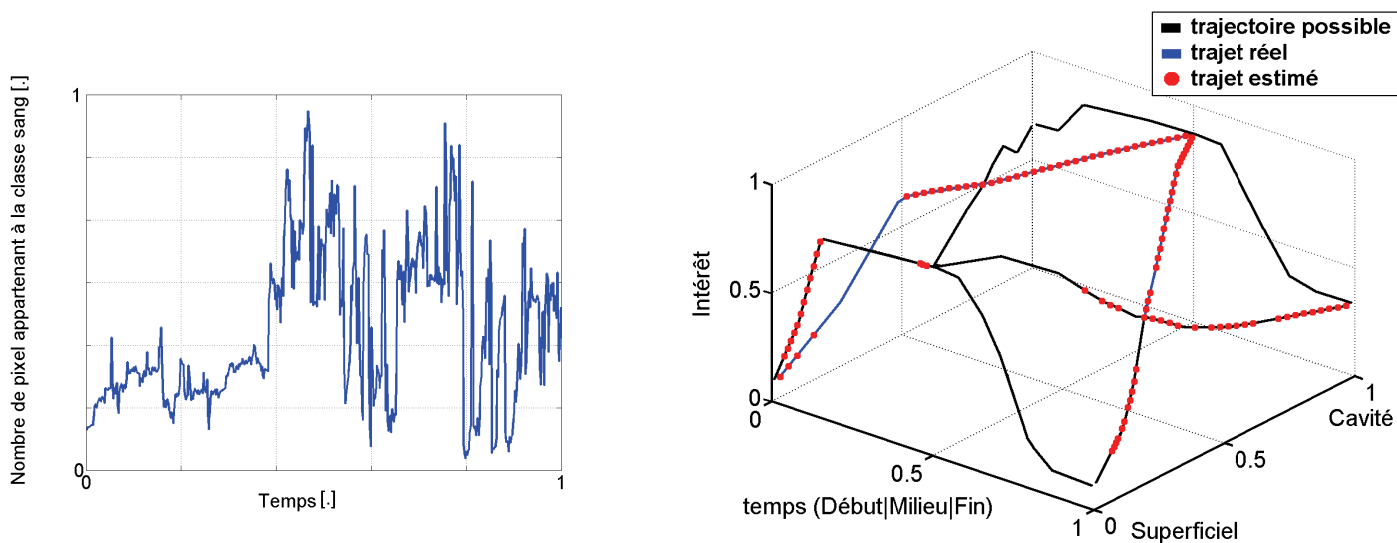


(c) Trajectoires les plus probables.

FIG. 5.20 – Détermination des coefficients α_1 .

Dès lors, le recouvrement des objets dans la scène conduit à deux cas de figure :

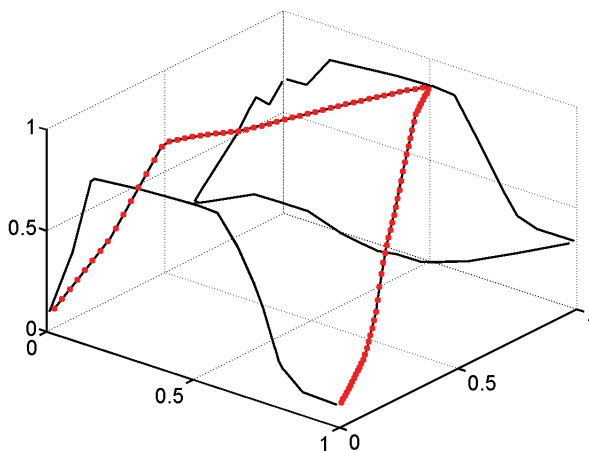
1. le recouvrement donne un pixel unique. Nous utilisons directement ces coordonnées dans la boucle de régulation pour calculer le torseur cinématique ;
2. le recouvrement donne plusieurs recouvrements à différents endroits de l'image. Il y a donc indétermination de la zone d'intérêt. L'origine de cette multiplicité est due principalement à une mauvaise reconnaissance des objets, plusieurs zones d'intérêt simultanées ou à un cas chirurgical non pris en compte. Dans ces conditions, le calcul du torseur cinématique n'est pas possible.



(a) Nombre de pixels reconnus comme appartenant à la classe sang.

(b) Trajectoire calculée.

FIG. 5.21 – Séquence opératoire déterminée par le système.

FIG. 5.22 – Conditionnement de la séquence opératoire au cas $\{S, C, S\}$.

Au cours de nos essais nous avons pu constater que le premier cas n'est jamais rencontré. Le recouvrement des zones même localisées ne permet pas d'isoler un pixel unique. Dans ces circonstances, il s'agit de procéder à une sélection de la zone d'intérêt parmi ces multiples recouvrements. Pour ce faire, nous utilisons un estimateur à noyaux car cette technique permet de ne garder que les recouvrements les plus récurrents représentatifs de zone d'intérêt. L'idée est que plus l'accumulation de recouvrements dans une région de l'image est importante, plus grandes sont les chances d'avoir une zone d'intérêt en ce lieu. Ceci équivaut à la sélection du maximum d'une densité de probabilité.

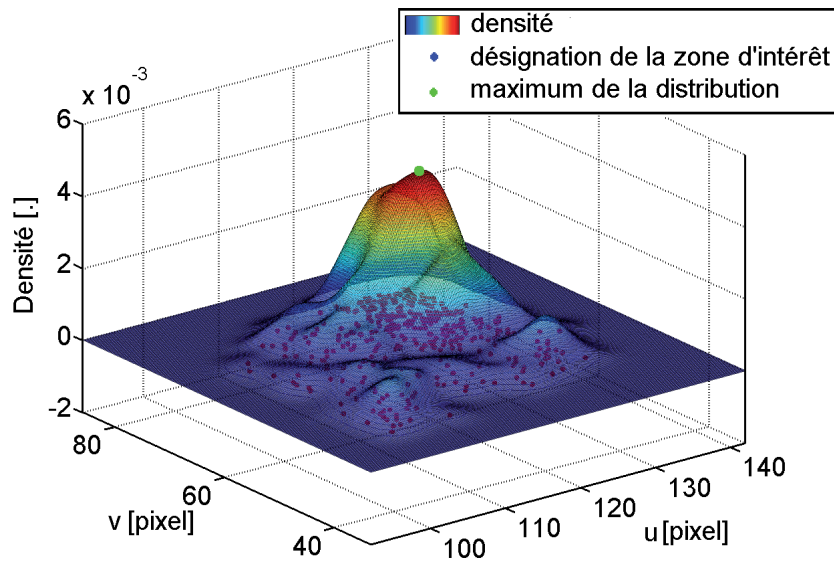


FIG. 5.23 – Densité obtenue au cours d'essais en clinique par estimateur à noyaux.

Nous présentons en figure 5.23, un exemple de densité obtenue pour un jeu de données réelles modélisées à l'aide de l'équation (5.3) où nous n'avons retenu que le centre de gravité de chaque recouvrement pour simplifier les calculs. Dans cet exemple, nous avons privilégié un noyau gaussien ($K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right)$) bien qu'il eut été possible d'utiliser un autre modèle. Nous pouvons constater que pour cette séquence chirurgicale les lieux de recouvrements sont concentrés dans une portion de l'image (40×40 pixels). Nous pouvons également observer que les recouvrements regroupés conduisent à une augmentation importante de la densité, le point le plus haut étant considéré comme étant la zone d'intérêt "finale". Ses coordonnées (u, v) serviront au calcul du torseur cinématique permettant, in fine, de centrer la zone d'intérêt ("finale") dans le repère image.

$$\hat{p}_R = \frac{1}{nh} \sum_{i=1}^n K_0\left(\frac{X_i - x}{h}\right) \quad (5.3)$$

Il n'est pas possible d'appréhender sur la figure 5.23 l'aspect dynamique de la sélection de la zone d'intérêt. Pour ce faire, nous présentons en figure 5.24 les performances de ce type d'estimation. Nous utilisons pour cet exemple un jeu de simulation afin de faciliter la lecture. A titre de comparaison, nous présentons en figure 5.24b les résultats d'une estimation qui prend en compte l'ensemble des données (de type moyenne glissante).

Sur ces figures sont représentées en bleu les désignations (correspondant au centre de gravité des lieux de recouvrement) où nous avons à intervalles de temps réguliers inséré des désignations erronées afin de tester le comportement des deux méthodes. Les endroits où le chirurgien travaille réellement sont les régions 1, 3, 5 et 7. Les régions 2, 4 et 6 sont quant à elles représentatives des erreurs de désignation pouvant être générées par une mauvaise reconnaissance de classe, un bruitage important de l'image, etc. Il y a donc alternance de régions "correctement" désignées et "faussement" désignées ; la proportion des

désignations de ces dernières étant de 1 pour 10. Les points verts (figure 5.24) illustrent les zones d'intérêt retenues par le système.

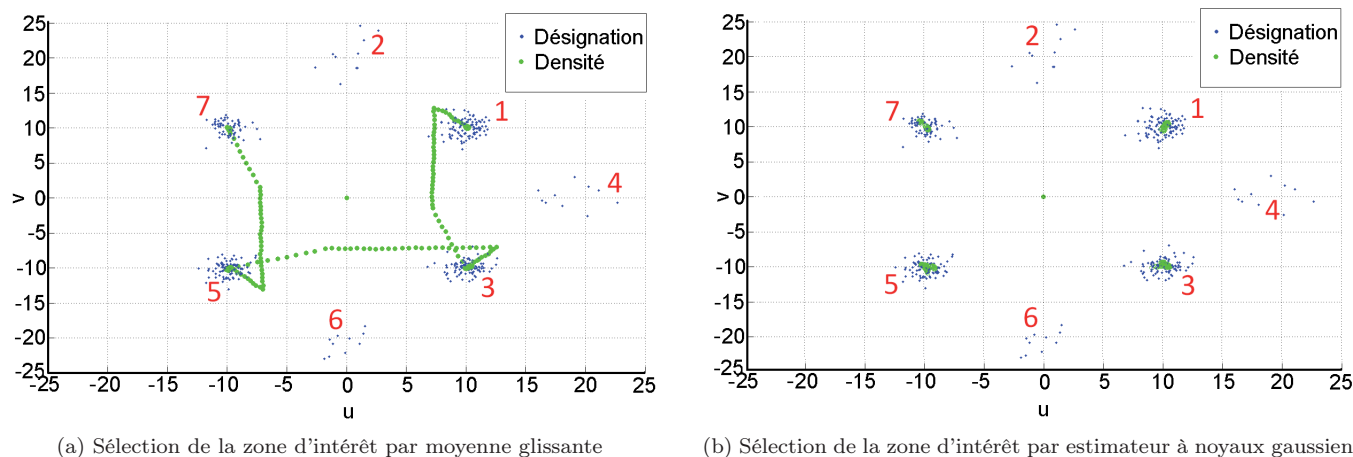
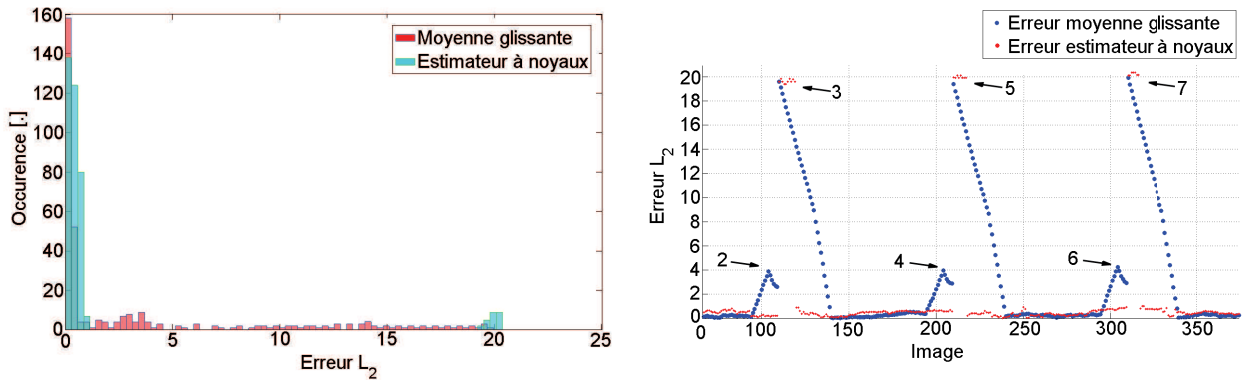


FIG. 5.24 – Sélection de la zone d'intérêt finale par moyenne glissante et par estimateur à noyaux.

La méthode par moyenne glissante intègre à parts égales chacune des désignations qu'elles soient erronées ou non (figure 5.24a). Dès lors la localisation de la zone d'intérêt s'en trouve affectée. Cette influence des désignations "erronées" sur la zone d'intérêt finale peut se voir sur la figure 5.25b représentant l'erreur en fonction du temps (erreur au sens de la norme L_2). La transition entre deux régions ne se fait donc pas sous la forme de saut mais de manière "continue", ce qui se retrouve dans l'histogramme des erreurs (5.25a), où les erreurs sont uniformément réparties.

A contrario, nous pouvons voir sur la figure 5.24b que l'estimateur à noyaux offre de bien meilleurs résultats que la méthode par moyenne glissante. En effet, l'absence de marquage vert en dehors des régions 1, 3, 5 et 7 montre que cette méthode permet de ne pas retenir des désignations incorrectes, la zone d'intérêt finale étant tour à tour au centre de ces régions "correctes". L'estimateur n'est donc pas influencé par les désignations erronées et permet donc une meilleure sélection de la zone d'intérêt. Cependant, cette technique implique un retard dans la sélection de la zone d'intérêt finale (figure 5.25b). En effet, aux premiers instants de l'émergence de la région 3, l'estimateur à noyaux ne considère pas ces échantillons comme étant représentatifs d'une nouvelle zone d'intérêt. Durant ce court moment, ces échantillons sont considérés comme étant des parasites au même titre que ceux de la région 2, leur contribution étant mineure face à ceux de la région 1. Il faut attendre l'accumulation de désignations dans la région 3 pour voir le maximum de la densité de probabilités correspondre au centre de cette nouvelle région. Un retard apparaît entre la première désignation de la région 3 et l'instant où l'estimateur considère que cette région est d'intérêt. Ce retard inhérent à la méthode pose donc problème dans la mesure où il détériore la réactivité du système. Ainsi, plus ce retard est important, plus le système mettra du temps avant de se réorienter. Il s'agit de trouver un moyen permettant de diminuer ce temps de retard.



(a) Histogramme des erreurs pour l'estimation par moyenne glissante et par noyaux.

(b) Evolution de l'erreur en fonction du temps.

FIG. 5.25 – Robustesse des méthodes face aux aléas.

Diminution du temps de retard

Jouer sur la taille de la fenêtre sur laquelle les estimations sont faites permet de modifier ce retard. En effet, comme le montre la figure 5.26 un minimum de 2 images est atteint pour une fenêtre de 55 désignations. Cependant, modifier ce paramètre modifie également la robustesse des résultats comme le montre la figure 5.27a qui présente l'erreur de sélection sous forme de boxplots calculés pour différents horizons d'estimation (nous exposons également à titre de comparaison les résultats obtenus pour la méthode moyenne glissante). Les données non représentatives sont symbolisées en rouge. Elles correspondent justement aux instants où l'estimateur à noyaux a adopté une position prudente.

Nous pouvons constater que le minimum d'erreur est obtenu également pour un intervalle de 55 désignations. L'ensemble de ces résultats est basé sur l'hypothèse qu'il n'y a pas congruence des échantillons erronés. Dans le cas contraire, le retard est proportionnel au volume des échantillons composant les zones erronées (loi polynomiale d'ordre 2, coefficient de corrélation $R^2 \sim 0.98$).

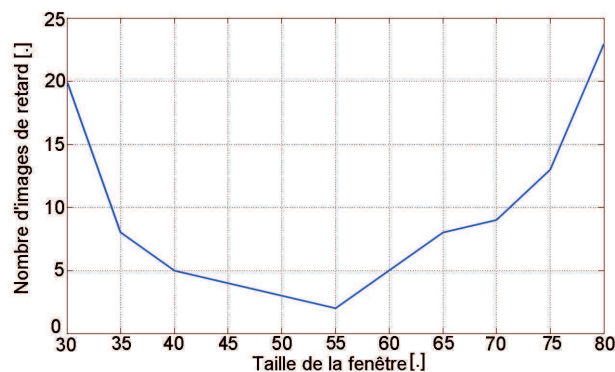


FIG. 5.26 – Temps de retard en fonction de la taille de calcul.

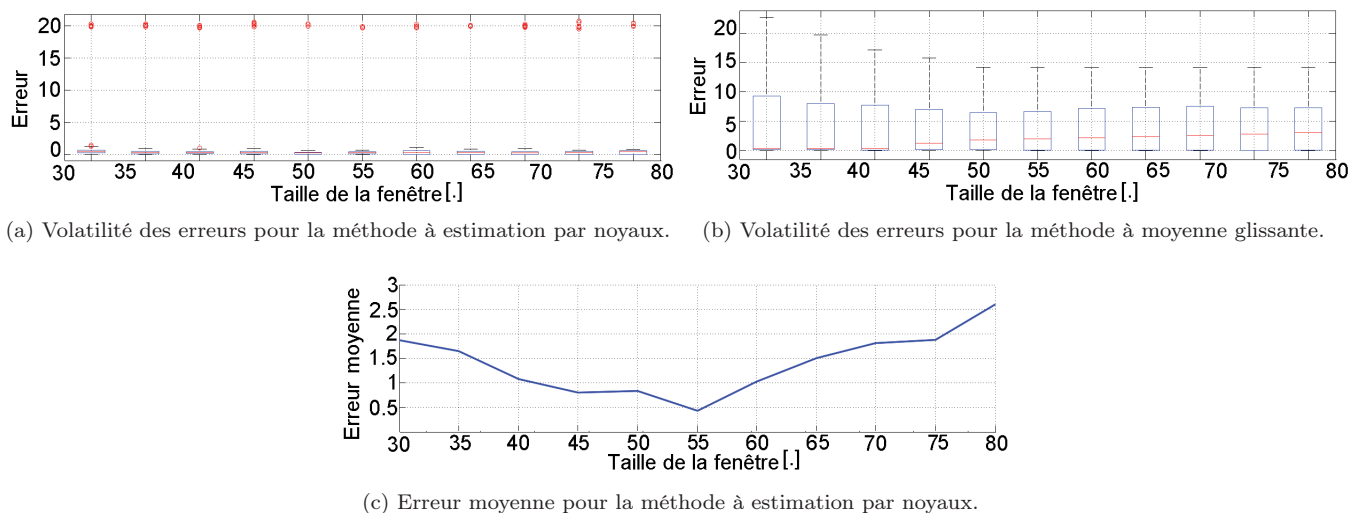


FIG. 5.27 – Variation de l'erreur pour différentes tailles de fenêtre.

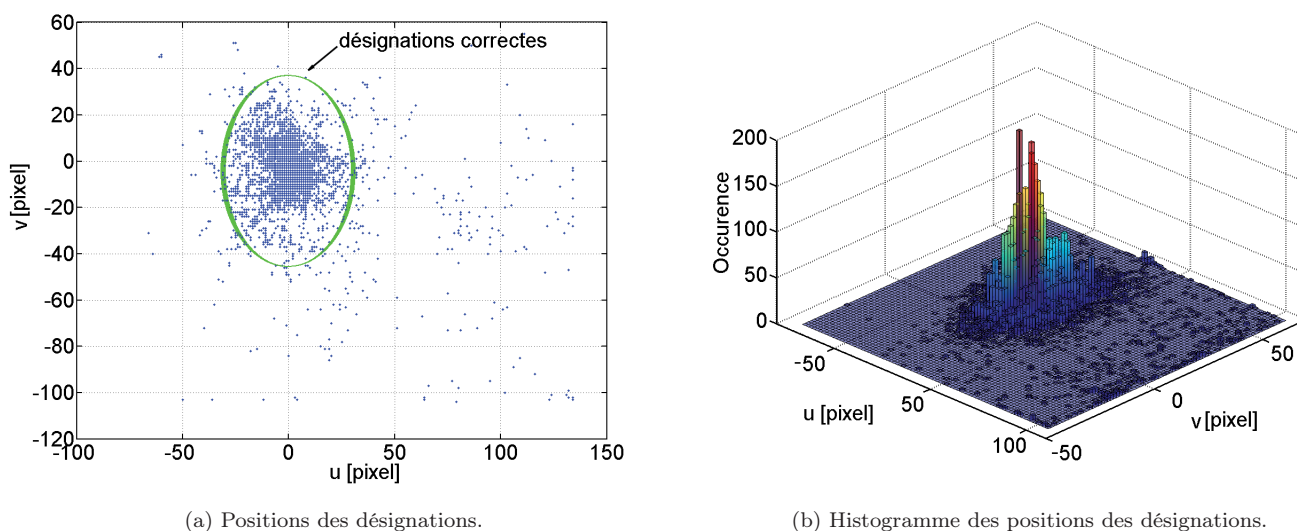


FIG. 5.28 – Résultats du processus de désignation pour une intervention chirurgicale.

L'exploitation de l'estimateur à noyaux dans des conditions réelles a déjà été présentée en figure 5.23 qui illustre la zone d'intérêt retenue pour un instant donné. Nous proposons en figure 5.28a, l'ensemble de ces zones d'intérêt retenues pour toute une séquence opératoire. Nous pouvons constater que la plupart des désignations sont à proximité de l'origine ce qui traduit un bon prépositionnement de l'éclairage qui n'a pas besoin de se déplacer grandement. Nous pouvons voir également qu'une partie de ces désignations se trouve éparpillée du fait des erreurs de reconnaissance. Il est intéressant de noter que celles-ci sont majoritairement situées d'un côté de l'image. Ce côté correspond à l'emplacement du chirurgien lors de l'intervention (figure 5.28b). Ces erreurs de reconnaissance sont en réalité dues à une mauvaise interprétation des masses de la classe instruments qui assimile les mains du chirurgien à des instruments "actifs"

(concordance d'un mouvement et d'éléments rectilignes). Sur le plan statistique, le délai minimum entre deux désignations lors de cette séquence est de deux images dans 91% des cas. Ceci est en accord avec les résultats de simulation précédents. Autre donnée importante, le taux d'erreur sur cette séquence de 20000 images est de 9.71%.

En résumé :

L'estimateur à noyaux permet de définir la zone d'intérêt du chirurgien à retenir. Les essais montrent qu'un horizon d'estimation de 55 désignations permet d'obtenir les meilleures performances. Toutefois, l'expérience montre que ce type de méthode n'est capable de sélectionner une zone d'intérêt qu'une image sur deux. Ceci diminue donc de moitié la réactivité du système mais permet de différencier un échantillon correct d'une désignation erronée et donc de garantir la robustesse du système. Le taux d'erreur obtenu est inférieur à 10% ce qui est perfectible mais toutefois acceptable pour cette étude.

5.3 Asservissement visuel de l'éclairage opératoire

5.3.1 Synthèse de la commande

A chaque itération, les coordonnées de la zone d'intérêt retenue sont transmises à un module de commande dont le rôle est de déterminer le mouvement à appliquer à la coupole pour confondre la tache lumineuse et la zone d'intérêt. Pour élaborer cette boucle de régulation nous nous sommes appuyés sur un certain nombre de choix :

- la robotisation de l'éclairage se limite aux axes #5 et #6 ;
- la caméra est placée au centre de la coupole. Centrer la zone d'intérêt dans l'espace image revient à placer l'éclairage de façon optimale ;
- à chaque instant $s(t)$ et s^* sont connus ;
- obligation pour l'erreur de décroître suivant une loi exponentielle ;
- le calcul de la matrice d'interaction se base sur l'hypothèse que l'altitude du système est constante ($\hat{L}_s = L_s(s, Z^*)$).

Cette boucle de régulation comporte trois étapes distinctes. Dans un premier temps, le traitement d'image désigne la zone d'intérêt du chirurgien. Dans un second temps, la loi de commande détermine le mouvement à effectuer dans le repère monde. Cette étape se décompose comme suit :

- accès aux informations $s(t)$ et s^* ;
- calcul de l'erreur $e(t)$;
- calcul du torseur cinématique $\tau(t)$;
- réécriture du torseur cinématique dans le repère R_0 (τ_0).

Enfin, dans une troisième étape, ce torseur cinématique est converti en consigne de vitesse pour les moteurs. Nous distinguons les activités suivantes :

- détermination des vitesses articulaires ($\dot{q} = (Q_1, Q_2)$ avec $\dim(q) = 1 \times 2$) à l'aide de la pseudo-inverse du Jacobien J du robot ;

- action sur les moteurs compte-tenu des profils de vitesse et d'accélération définis lors d'essais préliminaires, permettant de ne pas endommager prématurément les moteurs.

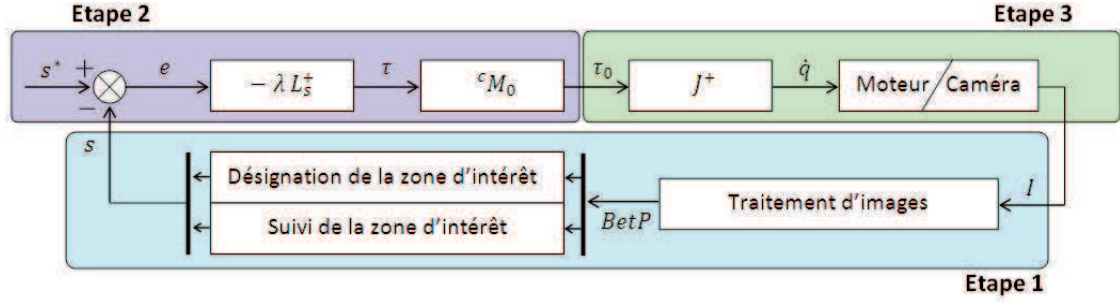


FIG. 5.29 – Synopsis de la boucle de régulation.

Etape 1 : Détermination de la zone d'intérêt

Nous avons vu précédemment comment désigner la zone d'intérêt du chirurgien. L'expression de la configuration courante s'écrit $s(t) = [u_{zi}, v_{zi}]^T$, où u_{zi} et v_{zi} correspondent aux coordonnées dans l'image du maximum de la densité définie par l'estimateur à noyaux.

Etape 2 : Calcul du torseur cinématique

Le calcul de l'erreur $e(t)$ se base sur un accès aux valeurs de $s(t)$ mais aussi de s^* . Si la configuration courante est connue, celle de référence n'a pas encore été définie. Dans la mesure où nous utilisons une seule primitive pour cette tâche d'asservissement, l'expression de la configuration de référence est donc $s^* = [u_0, v_0]$, avec u_0 et v_0 les coordonnées du pixel central de l'image.

Connaissant $e(t)$, l'accès au torseur cinématique est immédiat (en posant le scalaire λ égal à 0.1 pour éviter les problèmes de divergences et Z^* , l'altitude de l'éclairage opératoire, fixée à 1.1m). Celui-ci s'effectue dans notre cas à l'aide de la matrice d'interaction (4.28).

$$L_s(s, Z^*) = \begin{bmatrix} uv & -(1+u^2) \\ (1+v^2) & -uv \end{bmatrix} \quad (5.4)$$

Une fois le torseur cinématique τ calculé, nous procédons à sa réécriture dans le repère R_0 à l'aide de la matrice de passage cM_o qui relie le repère R_0 au centre de l'articulation #5 et le repère R_c correspondant au repère caméra situé au centre de la coupole. Nous utilisons pour cela l'expression (5.5). Ce calcul du torseur τ_0 achève l'étape 2.

$${}^cM_o = \begin{pmatrix} {}^cR_o & {}^c t_o \\ \mathbf{0}_{1 \times 3} & 1 \end{pmatrix} = \begin{pmatrix} \cos(\theta_1) \cos(\theta_2) & -\cos(\theta_1) \sin(\theta_2) & \sin(\theta_1) & -D_2 \sin(\theta_1) \\ \sin(\theta_1) \cos(\theta_2) & -\sin(\theta_1) \sin(\theta_2) & -\cos(\theta_1) & D_2 \cos(\theta_1) \\ \sin(\theta_2) & \cos(\theta_2) & 0 & D_1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (5.5)$$

Etape 3 : Calcul des vitesses articulaires

Pour calculer les vitesses articulaires, il est nécessaire de connaître le Jacobien du robot. Pour ce faire nous nous appuyons sur le modèle géométrique direct basé sur la représentation illustrée figure 5.30.

$$\begin{aligned} T_x &= (D_3 - D_2) \sin(\theta_1) & \omega_x &= \arctan\left(\frac{\tan \theta_2}{\sin \theta_1}\right) \\ T_y &= (D_2 - D_3) \cos(\theta_1) & \omega_y &= \arctan\left(\frac{\tan \theta_2}{\cos \theta_1}\right) \\ T_z &= D_1 & \omega_z &= \theta_1 \end{aligned} \quad (5.6)$$

Ceci nous permet d'écrire le Jacobien du robot (5.7) et de là, sa pseudo-inverse définie par : $L_s^\dagger = L_s^T(L_s L_s^T)^{-1}$.

$$J = \begin{pmatrix} \frac{\partial T_x}{\partial \theta_1} & \frac{\partial T_x}{\partial \theta_2} \\ \frac{\partial T_y}{\partial \theta_1} & \frac{\partial T_y}{\partial \theta_2} \\ \frac{\partial T_z}{\partial \theta_1} & \frac{\partial T_z}{\partial \theta_2} \\ \frac{\partial \omega_x}{\partial \theta_1} & \frac{\partial \omega_x}{\partial \theta_2} \\ \frac{\partial \omega_y}{\partial \theta_1} & \frac{\partial \omega_y}{\partial \theta_2} \\ \frac{\partial \omega_z}{\partial \theta_1} & \frac{\partial \omega_z}{\partial \theta_2} \end{pmatrix} = \begin{pmatrix} (D_3 - D_2) \cos(\theta_1) & 0 \\ (D_3 - D_2) \sin(\theta_1) & 0 \\ 0 & 0 \\ \frac{-\cos \theta_1 \tan \theta_2}{\sin^2 \theta_1 + \tan^2 \theta_2} & \frac{\sin \theta_1}{\sin^2 \theta_1 \cos^2 \theta_2 + \sin^2 \theta_2} \\ \frac{-\sin \theta_1 \tan \theta_2}{\cos^2 \theta_1 + \tan^2 \theta_2} & \frac{\cos \theta_1}{\cos^2 \theta_1 \cos^2 \theta_2 + \sin^2 \theta_2} \\ 1 & 0 \end{pmatrix} \quad (5.7)$$

En multipliant ce terme par le torseur cinématique obtenu précédemment, nous pouvons déterminer les vitesses articulaires des axes #5 et #6. Par leur mise en action, l'éclairage se déplace et agit donc sur $s(t)$. La réitération de l'ensemble de cette procédure permet d'atteindre s^* .

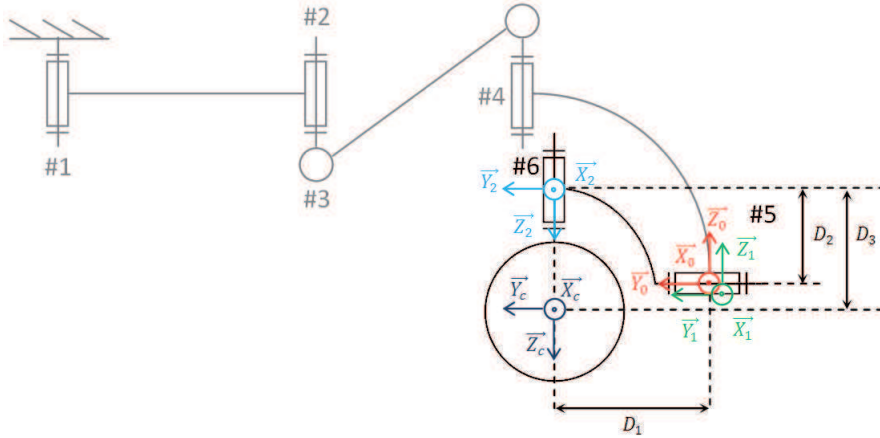


FIG. 5.30 – Schéma mécanique de l'éclairage opératoire.

Les premiers essais pratiques montrent deux particularités (figure 5.31). La première concerne le retard au démarrage qui affecte inégalement les deux moteurs (figure 5.31a). Ce retard de 14s et 23s, respectivement pour les axes #5 et #6, peut s'expliquer par un couple au démarrage important, une mauvaise répartition de charge et un jeu mécanique au niveau de la transmission du couple.

Le second point plus problématique concerne la vitesse de convergence qui n'est pas la même selon que la cible est éloignée ou proche du centre de l'image. Dans le cas d'une cible proche (figure 5.31a), la

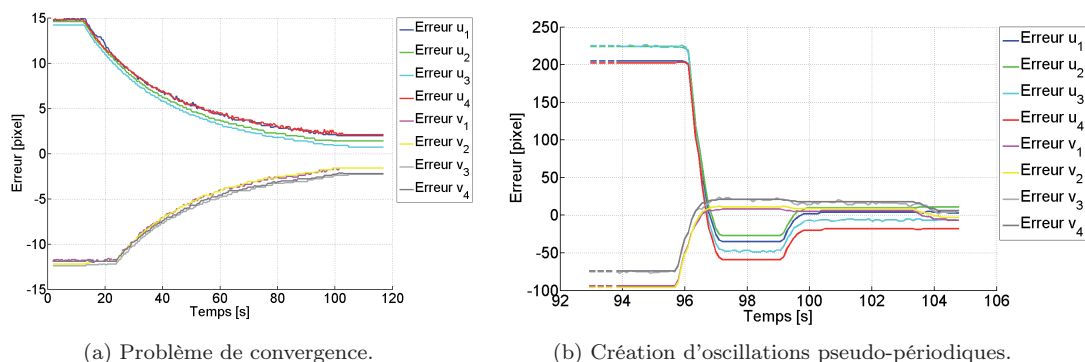


FIG. 5.31 – Comportement du système lorsque la cible est proche (à gauche) ou éloignée (à droite).

convergence est lente et laisse apparaître une erreur statique importante ($s \neq s^*$) malgré $\tau \neq \vec{0}$. Là encore la réalisation de la maquette est à l'origine de ce phénomène. L'absence de roulement au niveau des articulations rend plus difficile la rotation des moteurs à basse vitesse en raison du coefficient d'adhérence statique (proche de 0.3 dans notre cas). Malgré l'ajout d'un graissage, il n'est pas possible avec cette forme de loi de commande d'obtenir un placement précis de la coupole.

Ce phénomène est également présent lorsque la cible est éloignée (figure 5.31b). Cependant, dans ce cas, le problème de dépassement de la cible se substitue à la lenteur de convergence. Il y a apparition d'oscillations pseudo-périodiques générées par la norme élevée des composantes du torseur cinématique, qui lorsque la cible est loin sollicite grandement les moteurs se déplaçant alors au-delà de l'angle souhaité (le module interne chargé de la régulation bas niveau compensant le frottement vu plus haut) (figure 5.31b). On note un dépassement d'environ 20% de l'erreur maximale qui occasionne un retard supplémentaire, temps nécessaire aux moteurs pour appliquer les contre-actions.

La difficulté de modifier de façon importante le prototype, la présence systématique d'un retard au démarrage, le manque de précision à basse vitesse et le temps de convergence de l'ordre de la minute pour un déplacement n'excédant pas 15 pixels nous ont donc incités à revoir l'architecture sur laquelle est basée la loi de commande. Dans la mesure où il s'agit d'obstacles propres au prototype, il est préférable que les modifications soient les plus proches de l'étage moteur. Dès lors, la stratégie adoptée a consisté à intégrer un module correcteur entre l'étape du calcul du Jacobien et l'envoi de la consigne aux moteurs (figure 5.32). Ce correcteur de type proportionnel permet de diminuer le temps de convergence

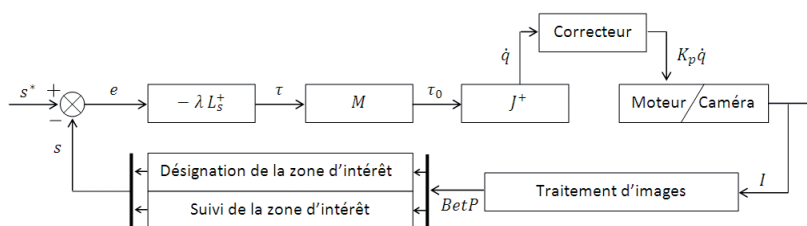


FIG. 5.32 – Insertion d'un bloc correcteur dans l'architecture de la loi de commande.

de près de 87% et l'erreur statique à moins d'un pixel. Cependant, face aux problèmes de dépassement, la correction a eu un effet inverse avec une augmentation de 80% de l'amplitude des oscillations pour un temps de convergence semblable à celui avant correction. De plus, bien que le temps de convergence ait été diminué grandement, le temps de retard au démarrage est resté constant.

La solution retenue consiste donc à modifier le gain du correcteur en fonction de l'éloignement de la cible. Ainsi, à basse vitesse, le gain est identique à celui précédemment utilisé afin de conserver les mêmes performances alors qu'à haute vitesse (et donc pour des éloignements plus importants de la zone d'intérêt), le gain est abaissé pour éviter les dépassements. Plusieurs profils de gain ont été testés : le tableau 5.12 propose un résumé qualitatif des performances.

Type de profil (définition sur $[0, +\infty[$)	Retard au démarrage	Phénomènes oscillatoires
Constante (a)	++	--
Linéaire ($-ax + b$)	++	++
Parabolique ($-x^2$)	++	+
Exponentielle ($-e^{ax}$)	+	-
Exponentielle tronquée ($-e^{ax}$)	+	-
Echelon	++	-

TAB. 5.12 – Evaluation de différents profils de gain (échelle de notation : ++ : comportement très bon, +, -, -- : mauvais comportement).

Le profil linéaire est la solution retenue, offrant un comportement adéquat à haute ou basse vitesse et garantissant, quelle que soit la consigne, un temps de convergence inférieur à 5s (figure 5.33a).

Certains tests ont pu montrer que le jeu mécanique au niveau de la transmission du couple est la principale cause du retard résiduel. Évalué à $3.75^\circ(\pm 0.4)$, il est à parts égales entre l'étage motoréducteur (2°) et celui de la transmission mécanique (1.75°). La réduction du jeu mécanique est donc un enjeu important en vue d'une commercialisation.

Remarque : *Toutes les stratégies mises en œuvre pour compenser ce retard se sont révélées inefficaces. L'utilisation d'une consigne compensatrice chargée de solliciter constamment les moteurs afin d'être à la limite du couple de démarrage n'a donné que des résultats mitigés. L'utilisation d'une consigne compensatrice chargée de solliciter constamment les moteurs afin d'être à la limite du couple de démarrage n'a donné que des résultats mitigés. L'une des meilleures techniques proposées consiste à injecter au niveau du correcteur la différence entre la consigne q calculée à l'instant t par le logiciel et l'estimation de la vitesse articulaire pour l'instant $t - 1$. Cependant, cette modification introduit des latences supplémentaires déformant la décroissance exponentielle.*

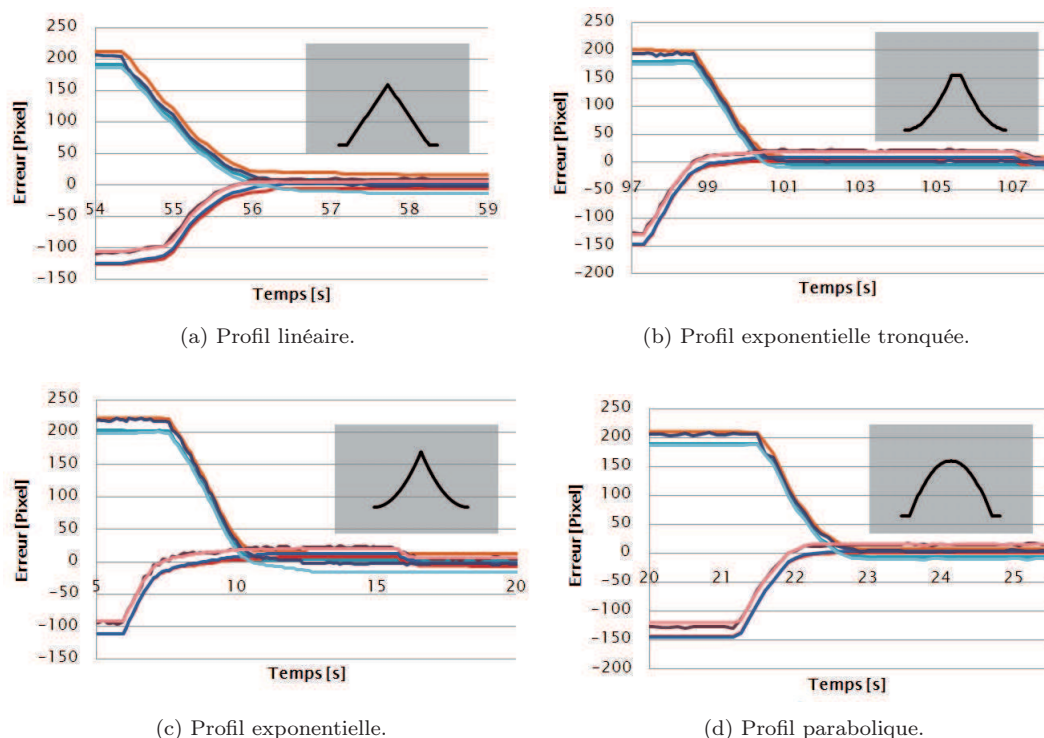


FIG. 5.33 – Erreur obtenue avec différents profils de correcteur.

5.3.2 Bilan des résultats expérimentaux

L'évaluation en conditions réelles du système est simple et ne repose que sur la quantification de trois grandeurs :

1. la désignation de la zone d'intérêt est-elle précise (erreur de désignation) ?
2. le système désigne-t-il souvent (fréquence de désignation) ?
3. le système se déplace-t-il jusqu'aux zones d'intérêt désignées (trajectoire suivie) ?

Les résultats présentés par la suite sont issus d'une intervention sur un porc pratiquée par un seul chirurgien sans assistant ni instrumentiste. L'éclairage opératoire asservi est le seul système d'éclairage utilisé. Celui-ci est à son niveau d'éclairage maximum, positionné à proximité de la table d'opération et orienté au milieu de la cage thoracique. La durée de l'opération est d'environ 90 minutes, ce qui est représentatif des interventions sur l'homme. Le chirurgien exécute son intervention sans considération de l'éclairage, le but étant d'obtenir des gestes similaires à ceux rencontrés en bloc. Ceci inclut les occultations partielles ou totales de la caméra, les phases d'inactivité (examen de la surface tissulaire par exemple), etc.

Précision du système

Pour juger de la précision de notre méthode de désignation, nous regardons la distance entre la zone d'intérêt définie par le système et celle du chirurgien. Pour ce faire, nous déterminons image par image la zone d'intérêt réelle puis calculons l'erreur sur les axes \vec{u} et \vec{v} de l'image (figure 5.34). Nous constatons que sur cette séquence opératoire, l'évolution des erreurs suit une décroissance exponentielle mais se trouve être fortement bruitée. Ceci est dû à la réactualisation permanente de la position de la zone d'intérêt

du chirurgien. En effet, la décroissance exponentielle n'est possible que si le chirurgien travaille sur une région très précise. Or, les activités de suture et de dissection ne sont pas par définition locales. Dès lors, il ne peut y avoir que des fragments de décroissance exponentielle. De plus, dans près de 10% des cas, une erreur de désignation est commise, ce qui a pour effet d'altérer encore plus la décroissance.

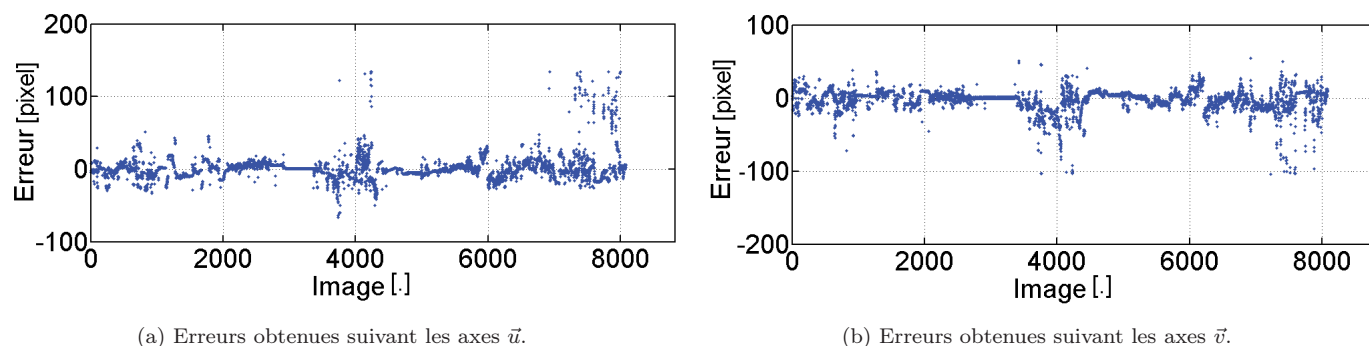


FIG. 5.34 – Erreurs entre la zone d'intérêt désignée et réelle.

Pour cette séquence l'erreur moyenne est de 13 pixels avec 9 pixels pour un intervalle [0 pixel ; 134 pixels] pour l'axe \vec{u} et 9 pixels pour un intervalle de [0 pixel, 104 pixels] pour l'axe \vec{v} . Compte-tenu du diamètre de la tache lumineuse nous pouvons conclure que la désignation est suffisante et conforme aux attentes.

Régularité de la désignation

Pour évaluer la régularité de la désignation, nous regardons les instants pour lesquels il y a désignation comparativement à l'activité du chirurgien. Les figures 5.35a et 5.35b présentent les résultats obtenus pour cette séquence. Nous pouvons voir que globalement le système réagit plus que nécessaire. Ceci est dû notamment à une mauvaise élimination des mouvements physiologiques leurrant le système qui croit reconnaître des zones d'intérêt ce qui entraîne, comme nous l'avons vu précédemment, une erreur inférieure à 10%. La superposition de l'activité du chirurgien et des instants où il y a eu désignation (figure 5.36), met en évidence des "surplus" de désignation, moment où le système a été trop optimiste. Nous pouvons voir également qu'il y a, par instants, absences de désignation, où cette fois le système a été trop prudent. Ceci s'explique par une superposition insuffisante des informations représentatives de la zone d'intérêt qui a entraîné une mise en sécurité du système, celui-ci préférant ne pas désigner.

En termes de réactivité, l'intervalle de temps le plus long entre deux désignations a été de 509 images (durant les images 16100 et 16609). Ce délai n'est pas représentatif de l'efficacité de la désignation car il correspond à l'une des phases d'inactivité du chirurgien. Nous quantifions la réactivité par rapport à l'intervalle de temps durant lequel le chirurgien a travaillé mais où il n'y a eu aucune désignation, il s'agit donc des instants où le système a adopté un comportement "prudent". Ainsi, 36 images sont passées sans avoir de désignation, ce qui a entraîné un arrêt du déplacement de coupole. Pour le chirurgien, cet arrêt n'a pas été gênant pour deux raisons :

- la zone d'intérêt était déjà bien éclairée avant l'interruption ;
- durant l'interruption la zone d'intérêt s'est peu déplacée restant du même coup à l'intérieur de la tache lumineuse.

Éliminer ce type d'interruption constitue une voie d'amélioration du système, ce qui serait profitable pour les chirurgies où la zone d'intérêt évolue rapidement. Toutefois, rendre le système plus réactif implique de relaxer certaines conditions (ou du moins de les réajuster). Les futures versions devront donc faire l'objet de la plus grande attention pour ne pas augmenter l'erreur de désignation.

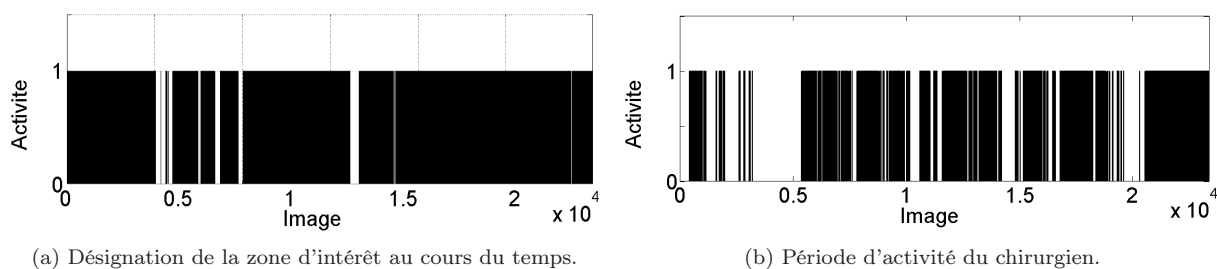


FIG. 5.35 – Comparaison entre l'activité du chirurgien et celle du système.

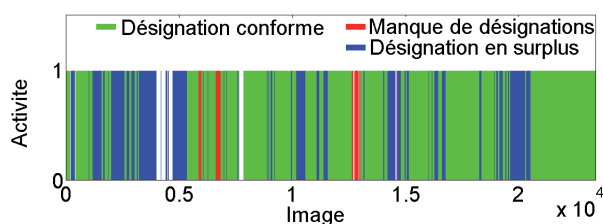


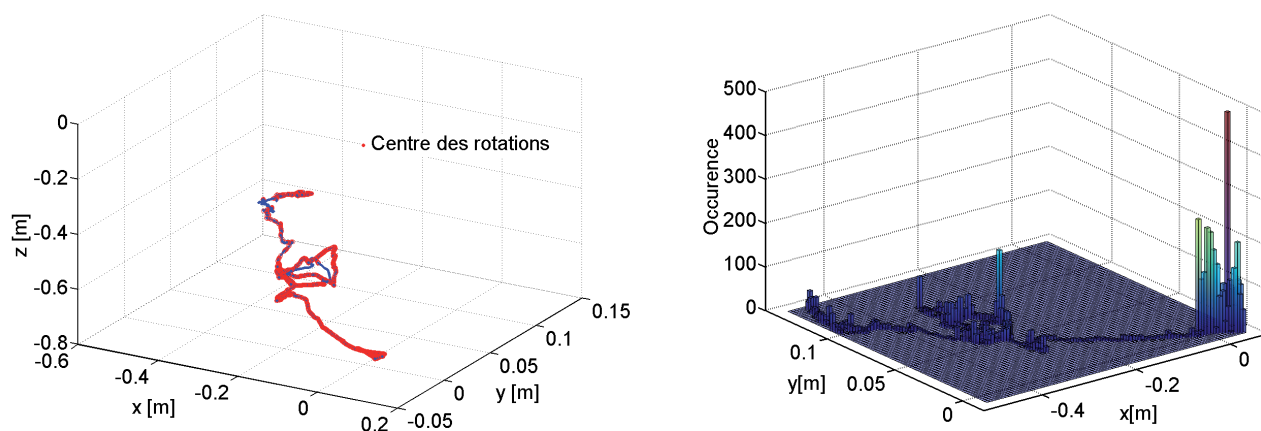
FIG. 5.36 – Mise en évidence des comportements "optimistes" et "prudents".

Réorientation de l'éclairage

Pour illustrer les déplacements effectués durant ces essais cliniques, nous présentons la trajectoire de l'effecteur du robot au cours du temps et par extension celle de la tache lumineuse (figure 5.37a).

Nous pouvons voir sur la figure 5.37b représentant les positions les plus souvent atteintes qu'il existe un axe de déplacement privilégié correspondant pour cette opération à l'incision principale. Nous remarquons que la position la plus fréquemment éclairée se trouve proche des coordonnées d'origine, ce qui traduit un bon prépositionnement de l'appareil en début d'intervention (point qui avait déjà été mis en avant concernant les lieux de désignation). Le déplacement effectué se base dans une très grande majorité des cas sur la désignation de la zone d'intérêt. Les points de la trajectoire obtenus à l'aide du processus de désignation sont en rouge sur la figure 5.37a. Nous constatons que le suivi, en bleu sur cette figure, permettant de déplacer l'éclairage en l'absence de zone d'intérêt désignée n'intervient qu'en de très rare occasion.

Ainsi le système, pour cette intervention, a réalisé 12786 réajustements allant du millimètre à une dizaine de centimètre (le cumul sur l'ensemble de l'intervention des réorientations de la coupole équivaut à un



(a) Déplacement de l'effecteur au cours d'un test réalisé en bloc opératoire. (b) Histogramme des lieux de déplacements de la tache lumineuse.

FIG. 5.37 – Résultats sur le déplacement de l'effecteur en conditions réelles.

déplacement au niveau de l'effecteur à 2.17m). Ceci met en évidence l'intérêt du système par rapport à une utilisation manuelle. En effet, il n'est pas concevable de demander à la panseuse ou au chirurgien d'actualiser l'orientation de la coupole à hauteur de ce nombre.

En résumé :

Les essais cliniques montrent que la loi de commande est effective malgré certains problèmes mécaniques du prototype. Des mesures palliatives ont permis de restreindre la portée de ces difficultés. Les essais en bloc ont montré que l'erreur de centrage obtenue est faible ce qui indique que la réorientation du système est adéquate. Le cumul des déplacements effectués indique que ce type d'asservissement permet un réel bénéfice pour l'équipe chirurgicale. Le nombre de réorientations au cours de l'opération montre qu'il n'est pas possible pour l'équipe chirurgicale d'accomplir cette tâche, ce qui justifie la plus value de ce système.

5.4 Conclusion

Dans ce chapitre, nous avons présenté le prototype développé dans cette étude. Nous avons vu les difficultés rencontrées en termes d'intégration moteur et d'exploitation de l'image ainsi que les mesures permettant de contourner ces problèmes (utilisation d'un filtre apodisant et d'un correcteur l'élimination des imperfections de conception).

Dans un deuxième temps, la mise en œuvre de la méthode de reconnaissance des objets dans la scène a été développée. La détection des instruments "actifs" grâce à l'intégration d'une étape de fusion d'attribut utilisant la forme et le mouvement se montre très satisfaisantes. Des essais en conditions réelles prouvent le bon comportement de la méthode et sa capacité à discriminer les phases où le chirurgien utilise un instrument ou non. Dans le cas de la peau, du sang et des champs stériles, seul l'attribut couleur est utilisé faute de puissance de calcul. Nous avons vu que l'éclairage localisé induit trois types d'exposi-

tions. Ces classes d'objets peuvent être normalement exposées, surexposées ou bien sous-exposées. Nous utilisons pour intégrer ces variations d'aspect un processus interne à l'attribut couleur qui se compose d'une fusion disjonctive. Les résultats sont suffisants pour cette application, mais toutefois perfectibles. Des tests exhaustifs ont montré que la reconnaissance basée couleur est robuste aux très forts éclairagements.

L'étape de désignation reposant sur une superposition des classes d'objets et d'un estimateur à noyaux montre que cette stratégie est appropriée au problème de localisation de la zone d'intérêt du chirurgie. L'exploitation en bloc opératoire de ces deux strates (reconnaissance d'objets, désignation) en entrée d'une boucle d'asservissement visuel prouve le bénéfice apporté par ce système et s'articule autour de deux points :

1. la précision du système qui permet d'obtenir un placement de la tache lumineuse inférieure à 13pixels ;
2. l'éclairage opératoire asservi visuellement permet des réorientations continues.

Ces deux résultats montrent la plus value du système par rapport à une utilisation manuelle.

Conclusion et perspectives

Dans cette étude, nous avons proposé une solution pour le centrage de la tache lumineuse d'un éclairage opératoire sur la zone d'intérêt du chirurgien dans le cas d'interventions à théâtre ouvert. Les principales difficultés rencontrées sont liées à la pluridisciplinarité des travaux mêlant à la fois automatique, traitement du signal et fusion de données prenant en compte des contraintes industrielles fortes.

Les solutions proposées actuellement dans la littérature pour remédier à ce problème reposent sur une interaction avec l'équipe chirurgicale soit par placement dans la scène opératoire d'un élément spécifique à l'endroit de la zone d'intérêt, soit par désignation manuelle par l'intermédiaire d'une interface homme-machine. Ces solutions nécessitent l'interruption répétée de l'intervention, ce qui est coûteux en temps, en énergie, et peut être source de conflit entre les membres de l'équipe chirurgicale.

L'approche envisagée dans ces travaux a consisté à asservir visuellement l'éclairage grâce à une caméra embarquée placée au centre de la coupole pour permettre son déplacement intelligent et autonome. La problématique a été de désigner la zone d'intérêt compte-tenu du milieu complexe, changeant et sans information explicite. L'absence de capteur additionnel a été un challenge supplémentaire, garant de la transparence du système pour l'équipe.

Le concept développé permet de s'affranchir de la réorientation manuelle de l'éclairage et garantit un éclairage optimal de la zone d'intérêt pour l'ensemble de l'intervention.

Pour cela, il a fallu reconnaître les objets composant la scène opératoire pour permettre une désignation robuste de la zone d'intérêt malgré la variabilité des opérations. Cependant, identifier tous les éléments composant la scène nécessite une grande puissance de calcul ce qui n'est pas acceptable dans cette étude compte-tenu des contraintes d'embarquabilité. Pour en limiter l'ampleur, nous avons mis en évidence qu'un jeu de cinq classes d'objets suffit.

Nous avons montré que pour accomplir cette différenciation, l'utilisation d'un seul attribut n'est pas suffisante. L'emploi combiné de la couleur, de la forme et du mouvement est seule garante d'une reconnaissance fiable des classes d'objets sélectionnées. L'étude des mécanismes de fusion nous a orientés vers le cadre crédibiliste plus versatile et aux meilleures performances lorsqu'il s'agit de gérer des informations conflictuelles. Cette étape est couplée à une prise de décision basée sur la théorie de la décision permettant à partir de ces objets de désigner la zone d'intérêt du chirurgien. En outre, nous avons mis en avant le fait qu'intégrer la notion d'importance dans ce processus permet d'éviter certains phénomènes contre-intuitifs.

Une fois la zone d'intérêt désignée dans le repère image, il restait à assurer le déplacement autonome de l'éclairage. Notre choix s'est porté sur un asservissement visuel 2D seul capable de réguler l'erreur uniquement à partir de l'image. Pour des raisons de sécurité, la désignation n'est effectuée que lorsque suffisamment d'informations sont accessibles, et donc par intermittence. Pour combler les discontinuités et ainsi éviter les à-coups lors du déplacement de la coupole pouvant gêner le chirurgien, nous avons utilisé une mesure de similarité permettant de résoudre le problème de lien entre la zone d'intérêt et les objets en présence.

Le prototype développé est aujourd'hui opérationnel et a été testé en conditions réelles. Les essais en bloc opératoire ont montré que la désignation de la zone d'intérêt et le déplacement de la coupole sont globalement satisfaisants.

L'asservissement visuel d'un éclairage opératoire reste un problème ouvert. Les perspectives de ces travaux devront notamment passer par :

- l'optimisation du code de calcul permettant une synthèse des processus de reconnaissance de l'ensemble des classes d'objets ;
- l'utilisation de la notion de risque dans le calcul de la probabilité pignistique ceci afin de limiter les erreurs d'assignation des pixels à une classe d'objets ;
- l'utilisation d'une caméra plus adaptée permettant la vision homogène de l'ensemble de la scène chirurgicale sans dispositif de filtration ;
- l'équilibrage de la structure de la maquette afin d'éviter l'étape de correction au niveau de la boucle de commande ;
- la comparaison des résultats de reconnaissance par rapport à un filtrage particulière ;
- l'amélioration de la désignation de la zone d'intérêt en superficie.

A ces points d'amélioration, il faut ajouter également les essais cliniques visant à étudier de manière plus exhaustive la précision ainsi que la stabilité du système pour d'autres types d'opérations telles que les interventions neurologiques (matière grise en grande proportion) et du genou (pas de champs stériles conventionnels).

Annexes

Annexe A

Evolution des éclairages opératoires

Dans les premiers temps de la chirurgie, l'inexistence de matériel d'éclairage adéquat rendait obligatoire l'implantation des blocs opératoires au dernier étage des hôpitaux. L'utilisation de grandes baies vitrées dépolies pour obtenir une diffusion maximale des faisceaux lumineux était, à l'époque, la seule méthode pour opérer dans des conditions plus favorables. Soulignons que l'emploi de la lumière naturelle en tant que source lumineuse principale ne permettait pas la vision en cavité, la mauvaise dilution des ombres ainsi que le faible éclairage gênant fortement ce type d'opération. Ce n'est qu'en 1854 avec l'arrivée du gaz distribué que ce type d'installation fut mis de côté pour privilégier d'autres formes de disposition, cependant, il fallut attendre 1889 pour voir apparaître l'une des premières percées notables de l'éclairage opératoire dans le monde chirurgical. En France celle-ci est due au chirurgien-physiopathologiste Antonin Poncet (1849-1913) qui introduisit une installation au gaz permettant une nette amélioration des conditions de vision pour un acte chirurgical (ce système composé d'une lampe à gaz fut conçue par Francis Wenham (1824-1908) améliorée plus tard par l'adjonction d'un bec de Punch renversé). Bien que l'utilisation de gaz inflammable en environnement confiné présentait certains risques, cette invention permit un réel progrès des conditions de travail pour le chirurgien qui n'était plus tributaire des cycles jour/nuit (circadien), de la météorologie et des saisons.

Une dizaine d'années plus tard, avec l'arrivée de l'électricité, les éclairages opératoires connurent des progrès significatifs. Ce saut technologique permit de diminuer les risques inhérents à l'utilisation du gaz dans les salles d'opérations et donc de contribuer à la sécurité dans les milieux médicaux. Les premiers éclairages opératoires électriques étaient constitués d'éclairages multi-hublots de fort éclairage dont le domaine spectral empiétait largement dans l'infrarouge (en réalité la désignation multi-hublot est, dans le cas considéré, impropre dans la mesure où il était plus question d'éclairages individuels rassemblés que d'un système unique composé de plusieurs ouvertures optiques). Ce point est une des caractéristiques majeures d'un éclairage opératoire car il contribue directement à l'élévation de température au niveau de la tête du chirurgien et ainsi qu'à celle des tissus en surface. On rapporte à cette occasion des cas graves de nécrose dus uniquement à l'empiètement du spectre de l'éclairage opératoire dans le proche infrarouge (dû à la conception de l'éclairage, soit par oubli de réinstallation des filtres, ...).

A ce point s'ajoute d'autres éléments entachant l'agrément d'utilisation de ce type d'éclairage opératoire tels que la difficulté de nettoyage et le fait que ce dispositif une fois implanté n'était plus déplaçable. Bien que les désavantages soient nombreux, ils permirent de mieux appréhender les besoins du chirurgien

tout en donnant un axe directeur de recherche et de développement. En effet, les besoins ont pu être clairement dégagés à savoir :

- être en mesure d'éclairer principalement la zone opératoire ;
- la possibilité de déplacer l'éclairage ;
- la suppression des émissions radiatives dans le domaine du proche infrarouge ;
- la gestion des ombres (notamment pour les actes en cavité) ;
- la focalisation du faisceau lumineux (modification du diamètre de tâche lumineuse).

Les années qui suivirent virent l'apparition de plusieurs concepts; on ne s'attachera ici qu'aux plus importants tels que la voûte réfléchissante d'André Walter, la voûte de Blin et le superscialytique.

La voûte de Blin

Cette voûte éclairante, conçue en 1981 par Jean Blin et Jean Hubert pour la société Air Liquide Médical France, est constituée de projecteurs répartis sur l'ensemble du plafond de la salle d'opération d'où le terme de voûte éclairante (figure A.1a). Cette voûte ajourée pour permettre aux étudiants de pouvoir assister aux opérations confirme la tendance d'avoir de plus en plus un lieu clos, au profit de l'hygiène. De nos jours de nouvelles évolutions sont apparues, où les seules modifications sont les sources lumineuses.

La voûte de Walter

Suivant le même principe mais avec cependant quelques modifications, la voûte de Walter n'utilisait qu'un seul projecteur implanté dans une salle mitoyenne à la celle de l'opération (figure A.1b). La voûte à proprement parler était composée non pas de projecteurs mais de panneaux réfléchissants répartis sur une surface hémisphérique ou encore héli-ellipsoïdale. Les changements de position de la source lumineuse permettait d'ajuster la tâche lumineuse au grès du chirurgien. En effet, en modifiant l'emplacement de la source de lumière, les faisceaux lumineux venaient se réfléchir sur différents panneaux, permettant ainsi de pouvoir ajuster la position de la tâche lumineuse dans le champ. A l'image du dispositif de Blin, cette voûte était pourvue de hublots permettant l'observation directe de l'opération par des étudiants. D'autres types de voûtes ont été conçus citons ici celle de Nelson.

Le superscialytique

L'agrément visuel apporté par la voûte de Walter a permis d'entrouvrir la voie à de nouveaux concepts en matière d'éclairage chirurgical notamment au niveau de leur cinématique. En effet, outre l'évolution des sources lumineuses en elles-mêmes, les systèmes développés ont offert de plus en plus de degrés de liberté afin d'améliorer le confort du chirurgien qui ne possédait jusqu'alors qu'un équipement statique auquel il devait s'adapter. Un bon exemple de cette volonté est le superscialytique inventé en 1919 par le professeur Louis Véraïn de la faculté des Sciences d'Alger (figure A.1c). Ce dispositif consistait en une source lumineuse collimatée et orientable montée sur une crémaillère en arc de cercle et entourée d'une couronne de miroirs qui seront plus tard remplacés par des lentilles de Fresnel. Cette source lumineuse commercialisée par la société Barbier, Bernard et Turenne (BBT), était située à l'extérieur d'une voûte vitrée permettant l'obtention d'un lieu chirurgical clos.

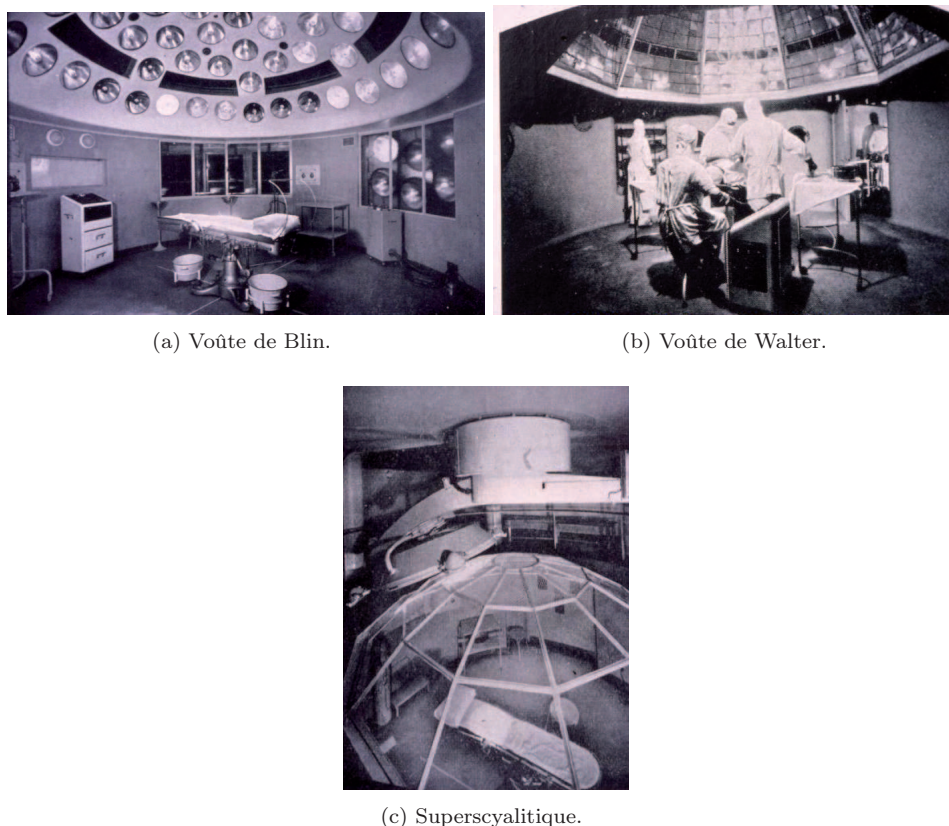


FIG. A.1 – Différents éclairages opératoires non manipulable par l'équipe chirurgicale (source : CHU de Besançon).

A.1 Caractéristiques d'un éclairage opératoire

Afin de simplifier la lecture de ce manuscrit, nous nous sommes limités aux caractéristiques utiles à ces travaux. Nous présentons ici brièvement les autres.

Flux lumineux (F_L) [lm]

Le flux lumineux exprimé en lumens, permet de quantifier l'impact de la lumière émise en termes de stimuli sur l'œil humain "standard". Une formulation pourrait se définir ainsi : le flux lumineux correspond au nombre de photons émis par une source de lumière rapporté à la sensibilité de l'œil humain standard à ces photons. Ainsi, une source de lumière émettant une quantité importante de lumens sera perçue par l'œil humain de manière très intense.

Intensité lumineuse (I) [$W.sr^{-1}$]

L'intensité lumineuse d'une source émettrice correspond au flux qu'elle émet par unité d'angle solide dans la direction considérée. L'intensité peut donc s'écrire ainsi :

$$I(m, p) = \frac{dF_S(m, p, d\Omega_S)}{d\Omega_S}$$

où I est l'intensité lumineuse de la source considérée émise dans la direction (m, p) et $dF_S(m, p, d\Omega_S)$ le flux lumineux de cette même source considérée dans un angle solide $d\Omega_S$.

Luminance (L) [$W.m^{-2}.sr^{-1}$]

La luminance au point (x,y) d'une surface réceptrice regardant dans la direction (m,p) est l'intensité par unité de surface apparente dans cette direction :

$$L(x, y, m, p) = \frac{dF_S^2(x, y, m, p, d\Omega_S)}{dA_S \cdot \cos\theta_S \cdot d\Omega_S} \quad (\text{A.1})$$

Une autre interprétation est de considérer le cas suivant : un éclairage opératoire émet de la lumière sur la scène chirurgicale, la luminance correspond donc à l'intensité, par unité de surface, observée par l'œil du chirurgien (par exemple) quand celui-ci observe les tissus opérés.

Source de lumière	Luminance [$Mcd.m^2$]
Soleil	1500
Lampe à incandescence	10-50
Lampe au xenon haute pression	200-2600
Lampe au mercure haute pression	300-1700
Lampe de voiture D_2 (Xe-Hg)	65
Lampe de métal halidé	30-400
Lampe céramic	300-1000
LED blanche	0.3-3.0

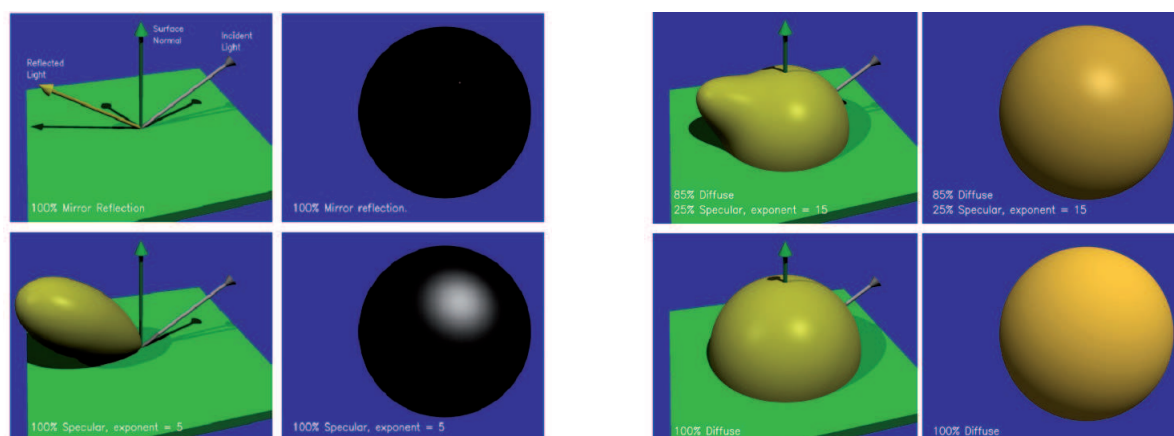
FIG. A.2 – Luminance pour différentes sources lumineuses

Albédo (ρ) [.]

Les matériaux possèdent, de par leurs caractéristiques physiques, une certaine capacité à réfléchir la lumière, cette notion (adimensionnelle) que nous appelons albédo traduit ce phénomène. L'albédo est donc la fraction de lumière que réémet la surface de l'objet considéré. Ainsi, pour un miroir l'albédo sera de 1 (réflexion spéculaire) alors que pour un solide Lambertien l'albédo sera proche de 0. Ces surfaces Lambertiennes ou orthotropes ont la particularité de réémettre la lumière dans toutes les directions. Une représentation plus implicite est présentée en figure A.3, où sont également présents des modèles intermédiaires dont le comportement est soit orienté vers un modèle Lambertien soit vers un celui d'un miroir.

Interprétation de la couleur

La couleur telle que nous la percevons n'est en fait qu'une interprétation de notre cerveau. Le mécanisme physiologique inhérent à cette interprétation est la résultante d'une interaction entre la lumière et les cônes tapissant la rétine. Ces cônes (au nombre de sept millions) ont pour unique fonction de réagir à la composante spectrale de la lumière. Il existe trois types de cônes : les L, M, S ; les cônes de type L sont sensible au rouge, ceux de type M au vert et enfin les cônes S réagissent au bleu. Ainsi, en recomposant les réponses de ces "capteurs" nous sommes capables d'analyser le spectre visible d'une source de lumière. Cependant, la sensibilité de l'oeil à distinguer les couleurs est une fonction non linéaire comme le montre le diagramme de Macadam ([Mac85]) présenté en figure A.4. Les couleurs proches du bleu seront facilement discernable de celles dans le vert, en témoigne la taille des ellipses représentatives de l'espace de discernement. Ajoutons que la capacité de l'œil à discerner les couleurs est dépendante



(a) En haut : un miroir parfait, en bas : une réflexion spéculaire

(b) En haut : un modèle d'illumination des matériaux courants, en bas : une diffusion Lambertienne

FIG. A.3 – Modèle de Phong.

de l'entraînement, ainsi les radioastronomes et les personnes travaillant dans le domaine du textile sont plus à même à différencier deux teintes proches.

Température de couleur [K]

La température de couleur d'une source de lumière est la température à laquelle il faut élever un corps noir pour obtenir un rayonnement semblable à celui observé (figure A.5a). La flamme d'un briquet en est un bon exemple. A sa base, la température est élevée, la couleur bleutée et à mesure que l'on s'en éloigne la température diminue et la couleur tend vers les teintes orangées. Il en va de même pour un éclairage opératoire ; une tache bleutée aura une température élevée et une tache orangée une température faible. Cet écart a pour origine les différences culturelles entre les pays nordiques et latins. Ainsi, dans les pays nordiques les produits auront une couleur de tâche lumineuse bleutée alors que dans les pays latins, la couleur sera plus rougeâtre (figures A.5c et A.5d). Adapter la température de couleur en jouant sur la quantité de "rouge" envoyée sur la scène chirurgicale peut être un bon moyen de faire ressortir certains détails. Cependant, il s'avère dangereux de modifier la température de couleur (de manière continue) pendant l'opération car certains tissus (comme par exemple les tissus nécrosés) peuvent ne plus être visibles et donc générer des erreurs cliniques.

Indice de rendu des couleurs (I.R.C) [%]

L'indice des rendus des couleurs comme son nom l'indique, définit l'aptitude d'un éclairage opératoire à restituer les mêmes couleurs d'un objet sous l'exposition du Soleil. La norme IEC 60601 définit cette caractéristique dans une échelle allant de 0% à 100%. Les produits contemporains en matière d'éclairage opératoire ont un I.R.C de l'ordre de 86% à 93% pour les meilleurs. Ainsi, plus l'I.R.C. sera élevé et plus le rendu sera meilleur à éclairage égal. L'I.R.C. d'une source lumineuse en salle d'opération est calculé à partir de la comparaison avec une source lumineuse standard assimilable au Soleil. Ainsi, plus le spectre tend à ressembler à celui du Soleil et plus l'aptitude de l'éclairage à restituer les couleurs sera élevée. De nos jours, l'I.R.C. se heurte à des limites dans la mesure où :

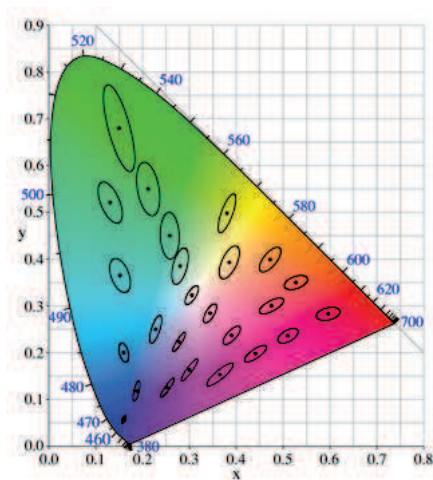
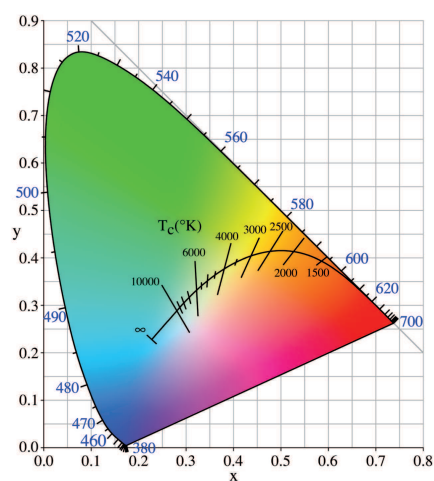
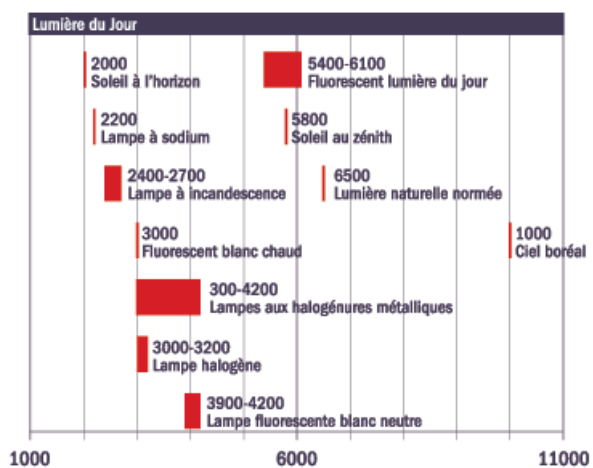


FIG. A.4 – Diagramme CIExy 1931.



(a) Température de couleur sur le diagramme de CIExy 1931.



(b) Plage de température de couleur selon le type de source lumineuse.



(c) Température de 3300K.



(d) Température de 4300K.

FIG. A.5 – Différence d'aspect pour deux températures de couleur différentes.

- pour faire ressortir les tissus vascularisés un éclairage opératoire se doit d'injecter une quantité supérieure de rouge (par rapport au Soleil) ;
- l'ajout potentiel de filtre infrarouge entache d'autant plus la comparaison ;
- l'I.R.C. est fortement dépendante de la technologie de la source lumineuse utilisée (les ampoules à décharge ont un mauvais I.R.C. car leur spectre est non lisse - présence de beaucoup de pics - alors que pour une ampoule halogène l'I.R.C. est très proche de 93%).

Pour ces raisons, il est donc impossible d'atteindre physiquement la barre des 100% pour un éclairage opératoire.

Source de lumière		I.R.C [%]
Soleil		100
Lampe à incandescence		98
Lampe au mercure à haute pression		40-60
Lampe au métal halide		80-90
Lampe au sodium	basse pression	0
	haute pression	20-40
Lampe à fluorescence	compacte	80-90
	blanche	80-95
	3 bandes	80-95
LED	RGB (faible qualité)	25
	RGB (haute qualité)	60
	blanche	80

FIG. A.6 – I.R.C. pour différentes sources lumineuses

Température dans le champ opératoire [K]

Pour des raisons d'objectivité de la mesure il est nécessaire de décorrélérer la mesure de la température dans le champ opératoire, de celle de la pièce. Ainsi, la température dans le champ opératoire correspond à la chaleur dégagée par les tissus superficiels du patient ; ces mêmes tissus étant très sensibles à une élévation de température il est donc impératif de maîtriser les émissions énergétiques de l'éclairage opératoire pour éviter les phénomènes de nécroses. Les sources lumineuses utilisées dans les produits d'ancienne génération se trouvent au cœur de ce problème, en effet, leur spectre empiétant largement dans l'infrarouge il a nécessité de filtrer ces émissions dangereuses (par exemple : l'halogène est plus une source de chaleur qu'une source de lumière) d'autant plus que les supprimer améliorera l'I.R.C. du produit. Les techniques usuelles en la matière sont au nombre de deux : la filtration athermique et la filtration dichroïque.

Filtration athermique :

Ce type de filtre (situé à la perpendiculaire du filament de la source lumineuse et la plupart du temps de

couleur bleu ou vert) implique la mise en contact avec une autre partie de la structure (pont thermique) afin d'éviter l'augmentation de la température de contact. Cette précaution est nécessaire dans la mesure où ce type de filtre fonctionne sur le principe d'absorption du spectre infrarouge.

Filtration dichroïque :

Ce type de technique utilise le principe de filtration du spectre infrarouge par dissipation vers la carrosserie (conduction). Une fois cette filtration effectuée il faut être en mesure d'évacuer la chaleur. L'utilisation de plus en plus fréquente de leds (pas d'UV et d'IR) se justifie par le fait que l'on cherche de plus en plus à limiter l'émission dans le domaine spectral ultraviolet et infrarouge pour ainsi éviter l'utilisation de filtres.

Température de contact

La norme IEC 60-601-2-413 impose la température maximale des pièces susceptibles d'être touchées à 333.15K à l'exception de la sous-face de la coupole qui n'est pas censée être touchée. Cette norme définit un seuil de brûlure supérieur au seuil de sensibilité d'un être humain au niveau de la main (323.15K).

Les solutions pour éviter l'accumulation de chaleur au sein de la coupole et l'évacuation de la chaleur pour éviter toute brûlure par contact direct :

- Distribution de la chaleur (convection) ;
- Pont thermique entre les composants « chauds » (conduction) ;
- Aération (cf. Prismatic) dans la structure de la coupole. Cette solution est désormais abandonnée pour des raisons d'hygiène ; il est préférable de conserver une structure fermée.

Contrôle des ombres

Le contrôle des ombres est l'une des nouvelles fonctionnalités d'un éclairage opératoire, il permet au chirurgien de pouvoir de se pencher sur le champ opératoire sans pour autant occulter le faisceau lumineux de la coupole. Cette aptitude de l'éclairage opératoire est possible grâce à une conception basée sur le principe de contribution totale de tâche. L'idée est que chaque source lumineuse concourt à l'ensemble de la tâche lumineuse ; ainsi, bien que l'utilisateur occulte certaines sources, la tâche reste inchangée. Cependant, l'éclairement baisse nécessairement, une compensation est possible comme sur les produits MAQUET utilisant la technologie A.I.M. (Automatic Illumination Management). D'autres méthodes sont possibles pour avoir la maîtrise de l'espace d'éclairement :

- augmenter la surface d'éclairement (implique que l'éclairage ne sera manœuvrable que par l'infirmière) ;
- éclairage de faible dimension pour augmenter la manœuvrabilité - destiné à être manipulé par le chirurgien ;
- coupole de forme compacte (et/ou associative) ;
- génération de la lumière plutôt en périphérie (meilleure vision du volume des objets).

Diamètre de tache

La première technique possible consiste à augmenter la taille de la coupole en conséquence (figure A.8a). Cette stratégie, outre le fait qu'elle induit une élévation du coût de production et une réadaptation de la suspension en raison de l'élévation du poids, pose le problème de la gestion de l'excédent de chaleur (émis en direction des tissus et au sein de la coupole). De plus, au niveau réglementaire aucun texte de loi



FIG. A.7 – Participation de toutes les sources lumineuses à la création de la tache lumineuse.

n'impose une puissance d'alimentation limite, mais la norme IEC 60601 contraint cependant à intégrer des éléments d'isolation (exigence d'avoir une double isolation lorsque la tension dépasse 48V) et de précautions proportionnelles à celle-ci. Bien que les points abordés soient techniquement surmontables, c'est avant tout l'encombrement et la maniabilité de l'éclairage qui a conduit à ne pas considérer cette solution ; la tendance actuelle étant en faveur des coupoles plus petites donc plus maniables.

Une autre solution possible consiste à réorienter les sources de lumière ponctuelles afin de couvrir une zone plus grande (figure A.8b). Facile à réaliser, cette solution n'a pour autant jamais été mise à l'épreuve. La raison réside dans la dilution des ombres qui permet à un éclairage, malgré la présence d'un obstacle (main, tête du chirurgien, organe masquant un autre organe, . . .), de continuer à éclairer un objet. Le principe pour réussir cette performance consiste à réaliser une tache lumineuse qui est la contribution de chaque source ponctuelle. Ainsi, en occultant une partie d'entre elles, la tache reste inchangée grâce à l'apport des sources restantes. Dès lors, en défocalisant trop les sources ponctuelles cette technique de dilution n'est plus efficace ce qui entraîne une visibilité nulle de l'objet lors des occultations. Cette seconde solution n'est donc pas à retenir.

Pour répondre au problème de diamètre de tache trop petit ne recouvrant pas suffisamment la zone d'intérêt, une troisième solution serait envisageable. Elle consiste à augmenter la taille des sources de lumières passant ainsi de ponctuelle à étendue (une solution analogue pourrait être d'augmenter leur cône d'émission) comme le montre la figure A.8c. Ainsi l'image de la source projetée sur l'objet serait plus grande ce qui augmenterait par voie de conséquence le diamètre de tache. Le problème réside dans le degré de complexité de l'optique qui est proportionnel à l'extension à apporter à la source lumineuse. Cette difficulté a pour origine les aberrations géométriques (déformation de la forme de la tache) et chromatiques (détérioration de la teinte de la tache ainsi que de la répartition des couleurs) liées à l'optique. Une source étendue telle que les leds pourrait répondre à cette problématique si l'efficacité lumineuse de cette technologie pouvait atteindre celles des leds actuelles. Cette solution est donc également à écarter.

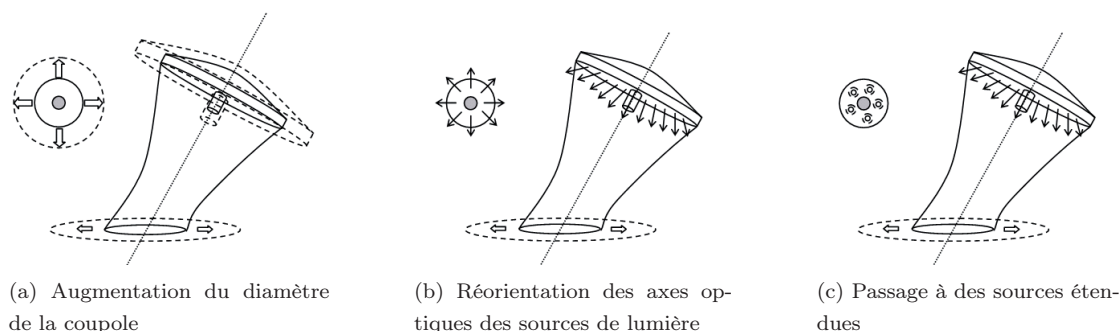


FIG. A.8 – Exemple de solutions pour l'augmentation du diamètre de tache.

Remarque : Dans le paragraphe précédent concernant le diamètre de tache il n'a pas été fait mention de l'augmentation de l'altitude comme solution potentielle. La raison est qu'il n'est pas possible d'assurer une tache uniforme au-delà d'une variation d'altitude d'une vingtaine de centimètres ; la multitude de taches créées ne permettant pas au chirurgien de pouvoir accomplir son geste dans des conditions suffisantes.

Eclairage et luminance

Ce que notre œil perçoit n'est en fait qu'une partie du spectre d'émission de la coupole, qui semble

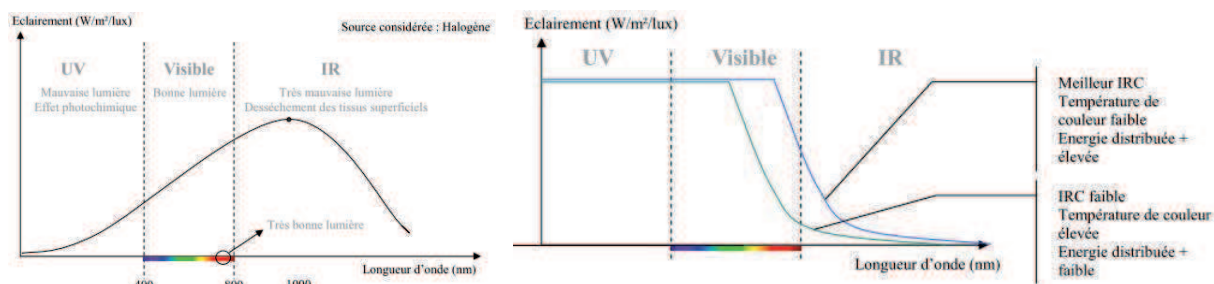


FIG. A.9 – Illustration de l'impact du spectre d'émission sur la vision qu'a le chirurgien de la plaie opératoire

homogène lorsque nous la regardons. Or, il n'en est rien, comme le montre la figure A.9a, où la répartition de l'éclairage couvre l'infrarouge et l'ultraviolet. Cette fraction du spectre a un impact sur les tissus mis à découvert durant l'opération. En effet, une exposition prolongée aux rayonnements ultraviolets peut conduire à des mutations dans l'ADN provoquant une faiblesse du système immunitaire (effet photochimique). Si ce phénomène peut mettre un certain temps à se déclencher il n'en est pas de même pour les infrarouges qui assèchent immédiatement les tissus superficiels. Des filtres sont donc généralement placés dans les coupoles d'éclairages afin d'éliminer l'infrarouge contenu dans le spectre d'émission. Or,

la filtration d'une partie du spectre a pour conséquence de détériorer l'IRC ainsi que la température de couleur (on parle de conversion de couleur). Il s'agit donc de réaliser un compromis entre un bon rendu des couleurs (figure A.9a) et une absence de chaleur (figure A.9b).

Volume de lumière

Le volume de lumière fait référence à la variation de distance zone d'intérêt-coupole pour laquelle les conditions d'éclairage sont considérées identiques (éclairage entre autres). Plus ce volume est important et plus le chirurgien peut évoluer dans les strates tissulaires différentes sans nécessité de déplacer l'éclairage. Ainsi, pour une altitude donnée en dehors de l'intervalle $[A_1, A_3]$ ($\sim [1\text{m}, 1.2\text{m}]$ pour un Powerled700), l'éclairage et l'homogénéité de la tache sont jugés insuffisants. Cependant, comme l'illustre la figure A.10a, en restant dans ces normes on peut s'apercevoir que le diamètre de la tache varie. Ce phénomène est une conséquence de l'architecture multi-sources fréquemment utilisée qui impose une divergence des faisceaux lumineux afin de résoudre le problème de dilution des ombres (voir figure A.10b).

Lors du prépositionnement de l'éclairage en début d'opération, l'infirmière de bloc vient placer la future zone d'intérêt entre A_1 et A_3 . Dans le cas où l'altitude se verrait inappropriée, l'un des membres de l'équipe chirurgicale peut la rectifier en rapprochant ou en éloignant la coupole.

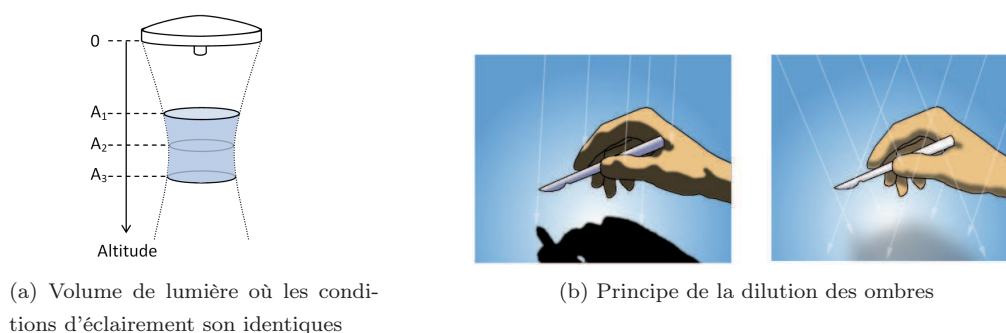


FIG. A.10 – Exemple de solutions pour l'augmentation du diamètre de tache.

Annexe B

Robots chirurgicaux hors étude

Nous avons vu en chapitre 1 que les concepts existants ne permettent pas au chirurgien d'obtenir des conditions d'éclairage satisfaisantes tout au long de l'intervention. Seule la robotisation de la structure, à l'image du C950 de la société Berchtold, semble être une solution appropriée mais s'avère limitée du fait de son incapacité à désigner par elle-même la zone d'intérêt du chirurgien. Nous nous intéressons donc dans cette partie à explorer les solutions robotiques développées pour les interventions chirurgicales. La majorité des applications robotiques étant dédiées aux opérations mini-invasives, nous ajouterons ces solutions à celles développées pour les interventions à théâtre ouvert.

Plusieurs critères existent pour classer les applications robotiques en chirurgie comme : le type de tâche allouée au robot, le domaine d'application ou encore "les contraintes imposées par les utilisateurs" ([PP05]). D'autres encore plus spécifiques portent sur les types de liaisons mécaniques permettant un effort nul au niveau du trocart (applications dédiées au mini-invasif). Nous retiendrons pour cette présentation celle de [CDL+01], qui se concentre sur le degré d'initiative laissé au chirurgien et structure simplement les applications en familles :

- **Les systèmes passifs** sont principalement dédiés à des tâches de navigation. Ils n'interviennent donc pas directement sur le patient mais uniquement par l'intermédiaire du chirurgien qu'ils renseignent. Un exemple typique concerne l'erreur d'insertion d'aiguille de ponction par rapport à une trajectoire prédéfinie. Ceci revient donc à mettre en correspondance la pose actuelle de l'instrument, estimée à l'aide de systèmes basés sur la stéréovision, et la pose prévue établie au préalable à l'aide d'un système d'imagerie radiatif (IRM). Ainsi plus ce recalage est précis, meilleure sera l'estimation de l'erreur et par voie de conséquence la précision du geste chirurgical.
- **Les systèmes actifs** sont conçus pour réaliser un geste chirurgical spécifique où une grande précision est requise. Leur succès repose sur l'accès aux modèles et informations pré-opératoires (connaissances a priori), un recalage précis (ce qui nécessite un système passif additionnel) et des conditions d'exploitation peu soumises à variation (mouvement du patient, variabilité de la zone à traiter, etc). Une fois le système installé et la pré-séquence opératoire réalisée par le chirurgien (comme la dissection des tissus permettant de mettre l'os à jour), le robot est activé. L'accomplissement du geste chirurgical se fait alors sans intervention du chirurgien dont l'action résiduelle se résume à l'arrêt du système. Une fois la tâche terminée et le robot retiré dans la scène opératoire,

le chirurgien reprend la main et termine l'intervention.

- **Les systèmes semi-actifs** sont à mi-chemin des deux précédents types de robots. Ils ont pour vocation de restreindre les gestes du chirurgien à une trajectoire définie dans un planning préopératoire. Le robot joue le rôle de guide et assure un geste sûr et précis. L'action est laissée au chirurgien qui a toute latitude pour procéder au mouvement, le robot n'étant présent que pour améliorer les conditions de réalisation de la séquence opératoire. Leur architecture dédiée les rend difficilement adaptables à d'autres types de chirurgies.
- **Les systèmes synergétiques** sont une extension des travaux menés pour la robotique semi-active. Ces dispositifs ont pour particularité de mettre en coopération le chirurgien et le robot qui est alors plus vu comme un guide intelligent. La réalisation du geste est laissée au chirurgien mais sous "surveillance" du robot. Généralement, trois modes sont accessibles au chirurgien. Le premier permet une totale liberté de mouvement : le robot ne sert que de localiseur. Les modes intermédiaires permettent d'atteindre une pose précise et de suivre une trajectoire prédéfinie. Le dernier mode est réservé à la restriction des mouvements à des zones admissibles. Par retour haptique le chirurgien sent une résistance à mesure qu'il se rapproche de ces régions. Dès lors, il devient impossible d'atteindre des zones dangereuses. A l'extrême les mouvements sont réduits à une dimension optimisant ainsi ses gestes (élimination des mouvements latéraux dus aux tremblements). L'avantage de ces systèmes est qu'ils offrent une architecture générique permettant une plus grande évolutivité contrairement aux systèmes semi-actifs plus "rigides". Cependant, à la décharge de ces derniers et malgré leur rapidité d'action, le manque de sécurité peut s'avérer être un frein pour leur mise en application. Le choix du type de technologie doit donc se faire en fonction des contraintes du geste à exécuter.
- **Les systèmes télé-opérés** offrent la possibilité de répercuter un ordre émis par le chirurgien et ce à distance. Le chirurgien joue le rôle de maître et le robot celui de l'esclave sur lequel un système de vision capture la scène. Le report de l'image sur un afficheur (3D ou non) permet au chirurgien par action sur des manettes (potentiellement à retour d'effort) de diriger les effecteurs du robot. L'avantage de ce type de système est qu'il permet l'élimination des mouvements parasites et un gain de précision. Ce type de dispositif est privilégié dans les cas où les tissus à traiter nécessitent une grande finesse comme pour les opérations de l'œil.
- **Les systèmes endocorporels** sont des dispositifs agissant directement dans le corps humain ("endo" signifiant à l'intérieur). Leur principe d'action est simple : une fois introduit par des voies naturelles ou artificielles (ouverture par le chirurgien de la cage thoracique par exemple), le robot doit atteindre une zone cible par ses propres moyens et la traiter suivant les ordres du chirurgien. Les modes de locomotion peuvent différer mais restent généralement "bio-inspiré", comme le montre [MD03]. L'avantage de cette approche réside dans sa faible invasivité qui permet de réduire au maximum les traumatismes occasionnés.

B.1 Les systèmes actifs

A l'opposé des systèmes passifs où l'action repose uniquement sur le chirurgien, les dispositifs actifs ont pour rôle d'accomplir le geste chirurgical sans intervention humaine. Ce type de robot permet l'obtention de meilleurs résultats en termes de précision, de rapidité et de robustesse.

L'EndoBot est un exemple de ce type de système ([KW01]). Développé afin d'exécuter lui-même les activités de suture pour des interventions mini-invasives, il permet un gain de temps par rapport à une approche conventionnelle qui s'avère ardue en raison des restrictions de l'espace de travail. L'origine de cette différence réside dans le nombre d'étapes nécessaires pour une activité de suture, décomposée pour le robot en quatre tâches élémentaires facilement appréhendables et plus rapidement exécutées (insertion de la fibre, création de la maille, développement et sécurisation du nœud). L'adjonction d'un contrôle de force permet d'éviter le déchirement des tissus et une interface homme-machine renseigne sur les défauts potentiels (coupure de courant, échec de l'algorithme, ...) ce qui garantit au chirurgien le bon fonctionnement de l'opération.

Autre application des systèmes actifs, les porte-endoscopes motorisés ont pour but de maintenir dans le champ de vision de la caméra la zone d'intérêt du chirurgien. Jusqu'alors dévolus aux assistants chirurgiens, ces robots permettent d'optimiser les conditions d'exploitation de l'image, à savoir :

1. une erreur de centrage nulle grâce au centrage systématique de la zone d'intérêt, ce qui permet d'améliorer l'agrément de travail (l'image affichée sur l'écran n'est plus vacillante) et évite les heurts liés aux mauvaises coordinations des actions entre le chirurgien et l'assistant ;
2. une suppression des tremblements. Le porte-endoscope n'est plus maintenu par la main de l'assistant chirurgien ;
3. un non recours à un assistant pour cette tâche. Celui-ci peut s'intéresser à d'autres activités de plus haut niveau. L'espace potentiellement gagné par son absence peut se traduire par un confort supplémentaire et supprimer les interférences ([LKK+03]).

Le cahier des charges de ce type de dispositif s'articule principalement sur les éléments suivants :

- le type d'information servant au guidage de la caméra ;
- le choix de la fréquence de la boucle de commande (influençant directement la réactivité du système) ;
- l'embarquabilité du système (minimisation des dimensions et du poids, ergonomie, choix de l'alimentation, bruit des moteurs généré) ;
- le mode de report des erreurs et de suivi de trajectoire ;
- la minimisation du coût qui implique le plus souvent l'utilisation de matériaux plastiques (ce qui influence la rigidité du système et donc sa précision) ;
- la compatibilité avec les appareils d'imagerie par émissions radiatives ;
- la garantie d'avoir un système fiable (au niveau mission) et sûr (au niveau patient).

Découlant des travaux de Jurong et Zemiti sur les manipulateurs compacts pour la chirurgie endo-

scopique, le système MC²E est un de ces porte-endoscopes motorisés ([KGD⁺03]). Ce robot permet de centrer l'image sur l'instrument manipulé par le chirurgien. Pour ce faire, trois leds sont rigidement fixées sur l'instrument. La localisation du centre de chacun de ces points lumineux s'effectue par détection de contour et estimation des paramètres des ellipses les modélisant. L'ensemble des calculs nécessaires au traitement d'images n'excède pas 20ms pour une image de taille 768×572. A ce stade, connaissant les coordonnées de la pointe de l'instrument par rapport au centre de l'image on peut en déduire le déplacement à effectuer, et dès lors centrer l'instrument dans le repère image de la caméra. Une alternative à l'utilisation de leds consiste à utiliser la couleur comme source d'information. Ce type de solution proposé par [KM11] permet de contourner les difficultés de détection de la pointe en raison du sang dans la scène opératoire.

Il n'y a donc pas à proprement parler de désignation de zone d'intérêt. Le déplacement du système est conditionné par la détection des leds fixées sur l'instrument. Dans la mesure où nous nous attachons dans cette étude à ne pas interagir avec les éléments de la scène chirurgicale (instruments compris), nous ne nous inspirerons pas de ces travaux. Cependant, il est intéressant de remarquer le véritable challenge qu'ont constitué les conditions d'éclairage dans ces travaux. En effet, la source de lumière embarquée induit une saturation au centre de l'image, ce qui nuit à la détection des marqueurs. Facteur aggravant, la "viscosité" des organes provoque des réflexions spéculaires qui ne permettent pas d'avoir des contours nets. De plus, les mouvements physiologiques génèrent à cette distance de grands déplacements qui perturbent d'autant plus l'algorithme de suivi. L'ensemble de ces éléments, ajouté à la faible structuration des tissus, rend donc le traitement d'images plus complexe. Une fois correctement détectés ces marqueurs permettent d'estimer la distance entre l'instrument et l'organe par retour vidéo. L'ensemble s'appuie sur un robot de type AESOP (système robotique présenté plus en avant dans la partie B.2) équipé d'une caméra à 50fps. Le temps nécessaire pour le centrage de la zone d'intérêt est de l'ordre de 10s (10.12s) mais avec une variabilité importante 5s (4.84s) au mieux et 20s (19.28s) dans le pire des cas.

Bien que ce système soit une avancée en matière de confort pour le chirurgien, il faut ajouter qu'il existe un écart entre la position de la zone d'intérêt réelle et celle définie par le système. En effet, considérer la pointe de l'instrument comme étant la zone d'intérêt du chirurgien induit une erreur qui rend cette technique efficace mais non optimale. De plus, cette application repose sur le fait que l'instrument n'est jamais posé sur le côté ou laissé inactif, ce qui est une particularité des chirurgies mini-invasives. Enfin, l'inexistence d'un point d'insertion unique, l'assurance de voir la zone à traiter, et la possibilité d'exécuter de grandes amplitudes sont encore d'autres points de divergences entre les applications mini-invasives et celles à théâtre ouvert. Nous verrons que ce type d'hypothèse n'est pas valide pour le cas de la chirurgie à théâtre ouvert (2.1.2).

Remarque : Dans la mesure où les porte-endoscopes ne participent pas directement au traitement des tissus cibles, un autre choix aurait été de les classer parmi les systèmes passifs. Cependant, dans le cadre de ce manuscrit l'aspect autonome a été privilégié, ce qui justifie leur intégration dans la famille des systèmes actifs. De plus, il ne sera pas fait mention des appareils d'auscultation échographique autonomes ([BBRW11]) ainsi que des dispositifs de radiothérapie tels que CyberKnife ([GW03]) car bien qu'il s'agisse de systèmes actifs dont le principe d'action est identique à l'EndoBot, ils ne relèvent pas à proprement parler du domaine chirurgical. Ici, l'effecteur est un accélérateur linéaire émettant des radiations ciblées sur une zone à traiter. Le guidage du système repose sur deux sources à rayons X qui émettent des faisceaux orthogonaux permettant de localiser des marqueurs radio-opaques implantés sur le patient. Un appareil complémentaire permet de synchroniser l'émission du flux avec la respiration du patient afin de réduire l'erreur de ciblage. L'avantage est qu'il n'est plus nécessaire d'attacher le patient sur la table, ce qui est une avancée en termes de bien-être. Les performances sont décrites dans [WMT09].

Autre exemple de système actif, le dispositif RobacKA permet la découpe autonome d'une partie de la calotte crâniale ([KEB+03]). La dissection est assurée par un robot Stäubli RX90 à six degrés de liberté. Un laser en amont de la trajectoire de dissection permet au chirurgien de voir les dérives potentielles. Là encore une étape de recalage est nécessaire. Celle-ci est rendue possible grâce à l'implantation au préalable de trois billes en titane radio-opaques dans le crâne du patient, permettant l'établissement d'une trajectoire pour la dissection ("pin-based registration"). Ce mode de détermination de la pose de l'os nécessite leur retrait après l'acte chirurgical ce qui induit des traumatismes supplémentaires. Il n'y a donc pas de désignation de la zone d'intérêt proprement dite. Pire encore, il est nécessaire de réitérer le recalage chaque fois que le patient est déplacé, ce qui implique un arrêt supplémentaire de l'intervention. Pour accélérer cette étape de recalage, l'approche envisagée consiste à utiliser à la fois des marqueurs radio-opaques et des leds infrarouges, permettant un recalage plus rapide mais moins précis. Cependant, comme le souligne [DKKS07], il est parfois impossible pour des raisons médicales de placer de tels marqueurs, notamment pour la chirurgie du rachis. Pour y remédier, Les auteurs proposent une solution alternative ("pinless registration") en se basant sur des algorithmes de point-to-surface matching. Il est alors possible de mettre en correspondance un modèle 3D et les régions observées. Ainsi par itérations successives, la pose de l'os est retrouvée. Cependant, pour des raisons de sécurité il est nécessaire d'obtenir une validation du recalage par le chirurgien ce qui nuit au concept de système autonome auquel s'ajoute un coût de 600k\$ bien que compensé par un gain potentiel de 50k\$/an ([BBB98]). Cette approche a été notamment utilisée pour le système RoboDoc employé principalement pour le fraisage crânial mais aussi pour le remplacement de l'articulation de la hanche [BBL97]. D'autres auteurs comme [FGHP03] ou [Bel11] proposent des applications comparables (implantation d'une prothèse auditive). Celles-ci se différencient de la précédente du point de vue du robot utilisé (RX130 contre RX90 précédemment), de la stratégie de découpe, de l'utilisation conjointe de la simulation de trajectoire et d'un capteur de force permettant d'éviter l'endommagement du crâne. Cependant, là encore aucune désignation de zone d'intérêt n'est effectuée.

Ce constat est également valable pour des systèmes autonomes utilisés dans des interventions autres que neurologiques, exemple en est du Dermarob qui sert au prélèvement cutané en vue de la chirurgie réparatrice [DDPP03]. Par appui d'un dermatome sur la peau, outil servant à découper une fine pellicule de peau, il est possible pour le praticien d'extraire un lambeau qui sera plus tard greffé sur la région à

traiter (brûlure, réparation). La faible occurrence de ce type d'intervention complique la tâche de l'opérateur (absence d'entraînement régulier). De plus, la difficulté d'obtenir une épaisseur de peau constante avec un dermatome, condition nécessaire pour la prise du greffon, a conduit à l'élaboration d'un robot (collaboration du LIRMM et de la société Sinters) placé le long de la table d'opération sur un chariot mobile possédant six degrés de liberté (bras de type SCARA PRR-RRR) dont le poignet a été modifié afin d'éviter les singularités (axes non concourants). Là encore aucune zone d'intérêt n'est désignée. Après une phase d'apprentissage où le chirurgien exécute les gestes à faire pour prélever le greffon, le robot est activé et reproduit par mimétisme ces mouvements. Le chirurgien n'a plus qu'à indiquer au robot, par le biais d'interface homme-machine, l'effort axial à fournir sur la peau régissant l'épaisseur de coupe. Le prélèvement s'effectue grâce à un capteur d'effort qui mesure la pression exercée sur la peau et un télémètre qui lui assure le contrôle de l'inclinaison de la tête de découpe.

En résumé :

Nous avons vu que les applications robotiques dédiées au domaine chirurgical se différencient en six grandes catégories :

- *passives : aide à la perception et/ou au guidage passif ;*
- *actives : l'action est laissée au robot ;*
- *semi-actives : l'action est déléguée momentanément au robot ;*
- *synergétiques : système coopératif homme-machine ;*
- *téléopérés : le robot retranscrit les mouvements exécutés à distance par le chirurgien ;*
- *endocorporels : insertion du robot à l'intérieur du corps.*

Les quelques exemples vus dans cette partie montrent l'incapacité des robots actuels à comprendre l'intervention chirurgicale dans sa globalité. Focalisés sur une partie d'un geste spécifique, ils ne permettent pas d'être adaptés à d'autres types d'opérations. De plus, leur performance se base sur une forte connaissance a priori ainsi que sur une étape de recalage conséquente en temps de calcul et non transposable à cette étude. Enfin la nécessité de préparer la zone d'intérêt avant la mise sous tension de l'appareil et le besoin d'achever l'opération une fois la séquence automatisée terminée, sont un dernier exemple illustrant le manque d'autonomie réelle de ces systèmes. Ainsi malgré les avantages procurés par ces dispositifs en termes de réduction de l'intervention (50min dans le cas du RoboDoc), de précision et de répétabilité, l'ensemble de ces avancées ne peut être utilisé dans cette étude.

B.2 La télé-chirurgie mini-invasive

La télé-chirurgie repose sur la reproduction des gestes effectués par un chirurgien (maître) par une entité robotique distante (esclave) dont les capacités sont supérieures. En 2002, une cholécystectomie a été effectuée grâce au système Tonatiuh par liaison satellitaire à 2400km de distance du chirurgien ([MVMMGG+02]). De telles interventions sont rares, plus couramment maître et esclave ne sont distants que de quelques mètres. Cette partie consacrée aux dispositifs télé-opérés n'a pas pour vocation d'être exhaustive. Nous nous sommes focalisés sur les systèmes les plus répandus dans ce domaine. Le lecteur pourra se référer à l'annexe B.2 pour une description de systèmes complémentaires tels que LAPROTEK, EndoXirob, Naviped, etc.

AESOP Acronyme anglais pour système endoscopique automatique pour le positionnement optimal (Automatic Endoscopic System for Optimal Positioning) fabriqué par Computer Motion, il fut l'un des premiers robots télé-opérés commercialisés (l'AESOP 1000 a reçu l'agrément de la Food Drug Administration en 1994 pour une utilisation en bloc opératoire). L'objectif est de maintenir et de positionner un endoscope grâce à une structure à six degrés de liberté dont la base est attachée à la table d'opération (rail latéral). L'avantage de ce système par rapport à une utilisation manuelle de l'endoscope est la stabilisation du champ de vue tout au long de la séquence opératoire (argument identique au MC²E). Les mouvements erratiques dus à la fatigue et aux tremblements de l'opérateur ne sont donc plus un obstacle pour le chirurgien. Malheureusement, malgré un temps d'installation relativement court son encombrement, pourtant faible au regard des dimensions du système Da Vinci, est une réelle gêne pour l'équipe chirurgicale. De plus, il s'avère incapable de désigner la zone d'intérêt, le contrôle du robot étant effectué par commande vocale dès la version AESOP3000 (1998). La signature vocale du chirurgien, sous forme de carte insérée en début d'opération, permet de ne pas prendre en considération les discussions ambiantes comme ordre potentiel. Le chirurgien se voit donc obligé de commander le robot grâce à des ordres exprimés à haute et intelligible voix. En l'occurrence, trois modes d'utilisation sont possibles. Le premier consiste à donner la séquence des mouvements à effectuer grâce à des mots-clés de type "gauche", "droite", "haut", "bas". Le second permet un mouvement "continu" (au sens de [LDS+06]), les mots "move" et "stop" conditionnant l'activation et l'arrêt du déplacement. Enfin, un troisième permet le déplacement de l'endoscope vers une position prédéfinie par le chirurgien au cours de l'opération.

EndoAssist Développé par la société Armstrong Healthcare pour réaliser la même fonction que l'AESOP, l'EndoAssist est actionné non pas vocalement mais par les mouvements de tête. Des marqueurs infrarouges placés sur un casque porté par le chirurgien permettent au système d'interpréter les ordres. Lorsque celui-ci souhaite déplacer la caméra endoscopique, il lui suffit de faire face au capteur placé à proximité de l'écran de contrôle, presser un commutateur au sol et effectuer un geste avec sa tête correspondant au déplacement que doit exécuter le robot. Conjointement un système additionnel vérifie que la tension exercée au niveau du trocart (point d'insertion de la caméra) n'est pas excessive et le cas échéant vient désactiver le robot.

Bien que facile à installer (moins de 10 minutes lors de la première utilisation) et à prendre en main, l'EndoAssist ne s'avère pas aussi efficace que l'AESOP ([LDS+06], [BJHM01]), malgré le gain de temps de 8min réalisé ([AGS+02]). Il ne s'avère pas non plus utile pour cette étude dans la mesure où il n'y a pas de désignation de zone d'intérêt.

Zeus Le système Zeus est dédié aux opérations cardiaques à cœur battant. Il est constitué de trois bras robotisés contrôlés par le chirurgien. Deux des bras servent à manipuler les instruments. Le troisième agit directement sur la caméra endoscopique insérée dans le patient. Dans sa version actuelle, la caméra est stéréoscopique, ce qui permet au chirurgien de voir en trois dimensions la scène opératoire, lui garantissant ainsi une immersion totale (grâce à des lunettes polarisantes comparables à celles utilisées dans les cinémas). Le chirurgien vient centrer l'image sur la zone d'intérêt grâce à une commande vocale, cette unique tâche pouvant aller jusqu'à 2.0 ± 0.8 min ([ZGY+06]). Malgré tous les bénéfices apportés par ce système (élimination des tremblements des gestes du chirurgien, mouvements plus précis, etc), il n'est

pas d'intérêt pour cette étude dans la mesure où il n'y a pas de processus de désignation.

Da Vinci Développé sur le même principe d'action que le système Zeus mais principalement pour les opérations thoracoscopiques et laparoscopiques, le Da Vinci (Intuitive Surgery Inc.) est le robot le plus cher actuellement en exploitation. La structure mécanique adoptée repose sur un mât principal où trois à quatre bras sont rattachés. Les gestes sont les répliques du chirurgien saisis par des manettes placées dans le poste de commande. Comme pour le système Zeus, il est possible de diminuer ou d'augmenter le facteur de multiplication des gestes (downscaling), celui-ci pouvant aller de 1 :1 à 5 :1, contre 2 :1 à 10 :1 pour Zeus [LDS⁺06]. L'élimination des tremblements par filtrage assure à l'esclave un comportement stable. [NK07] rapporte qu'en 2007 près de 400 Da Vinci ont été vendus à travers le monde (présent dans plus de 300 hôpitaux), que le nombre de publications s'y référant dépassait le millier et que le nombre d'opérations s'élevaient déjà à plus de 16000 en 2004. [HD11] rapporte que les procédures pratiquées par l'intermédiaire du Da Vinci sont plus sûres qu'une chirurgie conventionnelle et s'avèrent plus profitables pour l'hôpital (étude clinique pour la chirurgie de l'intestin antérieur sur un échantillon de cent vingt quatre patients). Une caméra stéréoscopique permet de voir la scène chirurgicale en 3D. Cependant, l'immersion se fait ici via deux écrans indépendants (chacun reproduisant le décalage de l'image nécessaire à une vision binoculaire) et non grâce à des lunettes polarisantes. Le chirurgien contrôle les déplacements de la caméra au moyen d'une pédale (servant aussi aux phases de dissection) ce qui indique l'absence de désignation de la zone d'intérêt.

En résumé :

A l'issue de cette brève présentation des applications robotiques, nous pouvons conclure que les systèmes développés sont dans l'incapacité d'analyser la scène opératoire et de désigner la zone d'intérêt du chirurgien. En l'occurrence, leur performance n'est rendue possible que par une préparation chirurgicale minutieuse peu soumise à aléa et sous condition d'accès à une connaissance a priori utilisant très majoritairement l'imagerie tridimensionnelle ; l'absence d'une de ces modalités entraînant l'impossibilité d'utiliser le robot. Ils ne prennent donc pas en compte toutes les difficultés et contraintes d'une opération dans leur globalité.

Robots compatible IRM

Cette revue des applications robotiques en chirurgie a fait état de robots passifs, semi-actifs, actifs et télé-opérés. Certaines contraintes ont été mises en avant, notamment la sécurité du patient ainsi que l'espace alloué au dispositif. Cependant, certains d'entre eux doivent évoluer dans des conditions encore plus sévères que celles déjà exposées. Ces robots sont conçus pour résister aux perturbations générées par les systèmes d'imageries permettant la vue interne du patient. L'absence de conception dédiée engendre un mauvais comportement du système comme le montre [TCM⁺11]. En l'occurrence les perturbations sont classiquement les champs électromagnétiques émis par une bobine supraconductrice ou des émissions radiatives de type X ou γ .

Le but de ces applications est double. Le robot doit, d'une part, permettre de contourner les difficultés liées à la restriction de l'espace imposé par le système d'imagerie tout en garantissant ou en améliorant la précision du geste. D'autre part, éviter au personnel chirurgical d'être exposé inutilement à des rayonnements.

Il est possible de citer des systèmes tels que le robot CT-Bot (projet ROBEA IRASIS), adapté pour les CT-scan, dont l'objectif est d'augmenter la précision liée à l'insertion d'une aiguille grâce à un dispositif monté et maintenu sur le patient à l'aide de bandes spéciales (le poids total n'excédant pas 3kg) [BBD⁺05]. L'adoption d'un système posé sur le patient, impose deux contraintes. La première est que le robot ne doit pas dépasser un poids raisonnable pour le patient, la seconde étant que toute l'électronique doit pouvoir supporter les rayonnements. L'opérateur peut choisir d'insérer lui-même l'aiguille en s'aidant du robot (sous-système à 5 degrés de liberté) ou bien, grâce à une interface, piloter cette insertion (activation d'un sous-système à 2 degrés de liberté ajoutés aux précédents). Bien que le mouvement relatif entre le patient et le robot soit nul, il se pose le problème de celui du robot par rapport aux organes (à la différence des systèmes type MARS ou Mbars monté sur l'os). Pour surmonter cette difficulté deux approches sont possibles : la compensation par méthode active ou bien l'arrêt du moteur pendant la respiration (la reprise étant lors phases d'apnées). La première peut être sous la forme d'un asservissement visuel conjoint à l'imagerie fluoroscopique. Cette technique est éliminée en raison des expositions inutiles aux radiations durant le tracking, ce qui a motivé l'adoption de l'approche reposant sur la seconde méthode. Ces robots actuels sont légers, d'un ratio poids/puissance bon, sûrs et stérilisables, comparativement aux premières versions qui étaient dérivées de robots industriels. Les nouvelles générations sont de plus en plus dédiées dès la conception à ce type d'application comme le montre [MBP⁺08].

Une alternative au robot monté, est de déporter la partie active du robot, électronique, contrôle-commande et les moteurs, en dehors du système d'imagerie. Un robot similaire au CT-Bot mais utilisant cette stratégie est proposé dans [Vir06]. Nommé KIMRO, il réalise l'insertion d'une aiguille ou d'un outil par le biais d'un bras mécanique de grande dimension. Extension de ces robots, le MR SAM permet le placement d'instruments (au sens large) sous imagerie IRM. Alliant les deux principaux avantages des architectures parallèles à savoir précision supérieure (comparativement aux structures sérielles) et espace de travail réduit, ce robot de type PSU (à six degrés de liberté) est placé directement dans la partie interne du scanner et sert à améliorer les performances du chirurgien [PM]. [WFG⁺03] propose une application similaire pour la réduction de fracture du fémur sous imagerie à rayons X et par l'injection d'un fluorophore. Le chirurgien pilote le robot dont l'effecteur est solidaire d'une partie de l'os fracturé, le but étant de remettre l'os dans la bonne position. Celui-ci peut se guider grâce à des marqueurs placés sur l'os. La manipulation se fait par joystick, le retour haptique étant disponible dans une version supérieure. Bien d'autres systèmes existent, le lecteur intéressé pourra se reporter à [GTSY10] pour une procédure prostatique transpérinéale, [SOA⁺10] concernant la brachythérapie de la prostate. L'activité chirurgicale combinée à l'imagerie IRM ou X ne se limite pas à ce type de gestes, d'autres comme celui développé par [LTCWN08] permettent le forage dans le cadre orthopédique. Conçu pour être rigide, stable et précis, il utilise la fluoroscopie comme système de navigation pour connaître la taille et la profondeur de perçage.

Concurrents aux Da Vinci et Zeus

Traitant des cas opératoires similaires mais avec un coût moindre (250keuros à l'achat), le robot LA-PROTEK (société EndoVia Medical) s'articule autour du concept d'évolution permanente où tout ou partie du robot peut être amené(e) à être remplacé(e) par un modèle plus récent sans impact sur le reste du robot. Les outils disponibles sont comparables en termes de nombre et de degrés de liberté à l'Endo-Wrist. D'un fonctionnement similaire au Da Vinci, il n'est pas de même de ses performances. Ceci est

dû aux multiples connections mécaniques, la déviation de la structure de support, l'étirement des câbles ainsi que la variabilité de la fabrication [Pei06]. Au-delà de l'encombrement qui reste toujours un point délicat pour les systèmes télé-opérés, une autre critique formulée dans [Dom02] concerne le manque de retour haptique, ce qui a été pris en compte dans le projet ROBEA MARGE. Ce projet propose un robot anthropomorphe PA-10 Mitsubishi à 7 degrés de liberté (EndoXirob) pour la suture sur cœur qui est dans le cadre mini-invasif une activité difficile. Les qualités de transparence du système allié à un retour haptique permettent une dextérité précieuse pour un pontage coronarien. Le Naviped quant à lui propose une solution pour la gestion des collisions sans changement de la position et de l'orientation de l'effecteur du robot [Kon04]. La difficulté réside dans l'évitement des configurations singulières qui peuvent être à l'origine de défaillances système, et la possibilité pour le chirurgien d'accomplir le mouvement désiré au niveau de l'instrument tout en conservant la précision constante. Ces auteurs réussissent cette tâche sur un robot à sept degrés de liberté à liaisons redondantes.

Les systèmes télé-opérés ne sont pas nécessairement volumineux et peuvent être comme le propose [BCB⁺03] avec un robot apposé directement sur l'abdomen du patient, le contrôle de l'orientation ainsi que de la profondeur se faisant par reconnaissance vocale. L'avantage principal repose sur la facilité d'installation comparé au système sur pied (AESOP et EndoAssist). Dans sa version plus récente, un système de traque externe POLARIS (Northern Digital Inc.), permet de suivre le bout de l'instrument afin que celui-ci reste dans le champ de vue de la caméra. Des delts jouent le rôle d'indice visuel et aident au calcul de la pose de l'endoscope et de l'instrument qu'il doit centrer. Des câbles actionnés par des muscles artificiels de McKibben (voir [KCH02] à ce propos pour les applications biorobotiques) sont gérés par un contrôleur PID (proportionnel, intégral, dérivé) mais se sont révélés difficiles de mise en œuvre et jugés insuffisamment fiables et précis. Pour ces raisons, un déplacement généré par des moteurs a été privilégié tendant à contrebalancer leurs inconvénients notamment de non-linéarité ainsi que leur délai de réponse (le prototype que nous avons développé tient compte de ces travaux et s'oriente vers un système de transmission par moteurs à l'image de ces auteurs).

A l'extrême, cette quête de réduction de la dimensionnalité a poussé l'émergence de systèmes évoluant à la surface même du tissu comme le HeartLander développé par [PZR04] dont les dimensions sont de 16mm de haut pour 13mm de diamètre. Dédié aux opérations intrapéricardiques à cœur battant, l'HeartLander se déplace suivant les ordres du chirurgien (via un joystick et un retour vidéo grâce à une fibre optique de 1.6mm de diamètre) sur la paroi épocardique. Ne possédant pas d'instrument, sa mission est dédiée au transport d'éléments thérapeutiques sur une zone cible, ce qui peut lui permettre d'être potentiellement dérivé au cadre ambulatoire. L'adhésion par succion est garantie par une pompe à vide (dépression de $-0.08\text{N}\cdot\text{mm}^2$) et les déplacements par le biais de câbles. Outre le gain de place procuré, un autre avantage est l'annulation de tout mouvement relatif entre la zone d'intérêt et le robot, ce qui est généralement le problème des systèmes sur pieds et partiellement de ceux posés sur le patient. A l'image de ce dernier robot, la réduction de l'espace occupé est encore un axe d'amélioration privilégié. Plusieurs projets témoignent de cette volonté et ont eu pour vocation de réduire la taille des dispositifs comme Active Trokar [KCW⁺02], composé d'un robot esclave tenant la base d'un trocart de manière fixe et d'un robot maître tenant l'autre extrémité permettant de déplacer le trocart. Pesant à peine 630g, les dimensions occupées par le système ne sont que de $88\times 131\times 197\text{mm}^3$.

Ceci explique l'augmentation de près de 46% de cette part de marché est prévue entre 2006 et 2011 (estimation de BCC research en date de juillet 2006). Si la tendance est aux robots, elle s'explique comme nous venons de le voir par l'apport au niveau clinique et financier mais aussi du fait même de nos différences physiologiques, comme le montre [LTWCWN08] :

Humain (avantages)	Robots (avantages)
Forte coordination main-œil Dextérité (à l'échelle macroscopique) Flexibilité et adaptabilité Intégration d'informations diversifiées et en grand nombre Appréciation d'information qualitative Bon jugement des situations Facilité à instruire et à débriefier	Bonne précision géométrique Stable et infatigable Extensibilité à d'autres ordres de grandeur Peut utiliser des capteurs divers dans sa boucle de commande Résistant aux radiations et aux infections Stérilisable
Humain (inconvénients)	Robots (inconvénients)
Dextérité limité en dehors de l'échelle macroscopique Sujet aux tremblements et à la fatigue Précision géométrique limité Capacité limitée à utiliser des données quantitatives Hygiène restreinte (non stérilisable) Sensibilité aux radiations et aux infections	Jugement médiocre Difficulté à concevoir et à debugger Restreint à des tâches simples Cher Limitation de la dextérité limitée ainsi que de la coordination main-œil

FIG. B.1 – Résumé sur les limites homme-machines ([LTWCWN08])

B.3 Les applications passives

Les systèmes passifs sont dans leur très grande majorité dédiés à la navigation (AXIOS Cambiar B1, Vicra, CT-Nav, Omnisight Excel, StealthStation, VectorVision, ...). Le bénéfice réel de ces dispositifs reste encore aujourd'hui controversé. Encore très récemment, un groupe de travail des Hautes Autorités de Santé constitué de professionnels du milieu, a dû être créé afin de mesurer leur utilité. Dans le cas de la chirurgie prothétique du genou, les résultats de cette étude (d'une durée de huit ans) montrent que la navigation permet une diminution de 14 à 25% du taux de mauvais alignement des pièces prothétiques lorsque l'exigence est inférieure à 3°, ce qui réduit du même coup les risques de déviation ([SP09]). Le tableau B.1 présente l'apport de la navigation confronté aux résultats en chirurgie conventionnelle.

Ces performances reposent sur l'étape dite de recalage qui implique une perte de temps estimée à 14 minutes ([SP09]). Celle-ci a pour origine l'installation d'un dispositif de repérage rigidement lié au patient ou à l'instrument, on parle d'ancillaire. Plusieurs moyens sont possibles pour sa localisation. Les systèmes de navigation basés sur de la stéréovision sont les plus répandus. Ils utilisent une caméra dans le domaine infrarouge afin de rendre l'ancillaire plus facilement visible et moins sensible aux variations

Auteurs	Echantillon patient	Déviatiion (avec navigation)	Déviatiion (sans navigation)
[SSY+07]	42	$0.7^{\circ} \pm 1.6^{\circ}$	$0.8^{\circ} \pm 2.5^{\circ}$
[JCK+05]	235	$0.0^{\circ} \pm 2.0^{\circ}$	$0.6^{\circ} \pm 3.4^{\circ}$
[PBP+05]	25	$0.78^{\circ} \pm 1.1^{\circ}$	$1.2^{\circ} \pm 1.7^{\circ}$

TAB. B.1 – Comparatif du gain de la navigation.

d'éclairage ambiant dans cette gamme spectrale. La précision du recalage obtenue aujourd'hui avoisine les 0.15mm (Vicra) voir 0.05mm dans le cas du 3D AXIOS Cambiar B1.

Comparer de manière équitable les performances des systèmes de navigation comme ceux présentées en tableau B.1 peut se révéler ardu. La raison, selon [LLR+11], réside dans le manque de standardisation des termes définissant les erreurs systèmes ainsi que la précision réelle obtenue. Pour ce faire, les auteurs décomposent le système de navigation en cinq strates représentatives des tâches accomplies par le système dans le but d'obtenir un langage commun pour les personnes impliquées dans le développement de type d'équipement. L'adoption de cette convention permettrait donc de mieux quantifier les besoins clients en termes de la précision nécessaire que l'application doit avoir :

1. l'imagerie pré-opérative ;
2. suivi chirurgical (tracking) ;
3. le recalage ;
4. la visualisation ;
5. la vérification.

Dans de plus rare cas, les systèmes passifs, à l'image du BlueDRAGON ([RBC+02]) ou du module "MiniBird" (développé dans le cadre du projet MARGE, [Dom02]), s'attachent à récupérer au niveau d'un effecteur des informations telles que la force et le couple appliqué par le chirurgien, ceci afin de déterminer l'action entreprise. Leur objectif secondaire est de donner un score représentatif de la performance accomplie et ainsi permettre l'amélioration du geste. Pour ce faire, le dispositif est solidairement fixé à un ciseau et couplé à un capteur d'effort à 6 composantes ce qui lui permet d'une part de récupérer la trajectoire empruntée par l'instrument et d'autre part de modéliser l'intégralité du geste de suture.

En résumé :

Les systèmes passifs reposent sur la mise à disposition d'informations a priori et d'un ancillaire solidaire de l'instrument ou de l'objet d'intérêt. Le manque d'adaptabilité, du fait de l'utilisation de cette connaissance exhaustive des éléments de la scène opératoire et de l'ajout d'information spécifique, est donc éloigné de la problématique visée dans cette étude. Ceci justifie donc l'exclusion de ce type de système des recherches.

B.4 Les systèmes semi-actifs et synergétiques en chirurgie

La difficulté de mise en œuvre et de sécurisation des systèmes actifs sont à l'origine de l'essor des robots semi-actifs qui participent au geste du chirurgien sans pour autant en être le donneur d'ordre. Le projet Mbars (Miniature Bone-attached robotic system) est un exemple de robot semi-actif dédié aux

arthroplasties [WJLD05]. Celui-ci se compose d'une plateforme Stewart-Gough à 6 degrés de liberté équipée d'une fraise servant à abraser l'os selon un schéma pré-plannifié. Apte à déterminer la surface de l'élément sur lequel il se trouve, il n'est pas capable de réaliser seul l'ensemble de l'intervention. Sur le même principe de fonctionnement, le robot miniature MARS (Miniature Robot for Surgical procedures) pour la neurochirurgie s'appose directement sur l'os du patient et permet d'orienter l'axe de forage ou d'insertion d'une aiguille, éliminant du même coup (comme pour MARS) le mouvement relatif entre la zone à traiter et le robot [SBZ+03]. Ses faibles dimensions ($5 \times 5 \times 7 \text{ cm}^3$), la précision obtenue (0.1mm), la possibilité de le stériliser et son poids ($\sim 200\text{g}$) lui permettent d'être implanté sans risques de dommages pour le patient ; il constitue du même coup une réelle alternative au guidage passif. Ces deux exemples présentent encore une fois l'absence de désignation de zone d'intérêt.

Un autre exemple est Surgicobot qui s'inscrit dans les systèmes collaboratifs appliqués à la chirurgie à théâtre ouvert ([RGT+11]). Conçu pour la chirurgie du rachis (notamment laminectomie correspondant à une libération de la moelle épinière), son but est de contraindre les gestes du chirurgien de telle sorte à ce qu'ils ne rentrent pas dans des zones prédéfinies comme étant dangereuses. L'instrument de fraisage est maintenu d'une part par le robot et d'autre part par le chirurgien qui contrôle les gestes à effectuer ainsi que l'effort à appliquer. Cette restriction de l'espace accessible au chirurgien est usuellement appelée contrainte active. Ceci est rendu possible grâce à un ancillaire solidaire de l'os et à une étape de recalage (précision obtenue de 0.2mm). Il n'y a donc pas de désignation de la zone d'intérêt ou encore de compréhension de la scène opératoire.

Trois systèmes d'évitement comparables au Surgicobot peuvent être cités ici. Il s'agit des robots Acrobot (pour Active Constraint robot), PADyC (Passive Arm with Dynamic Constraints) et MICRON. Le premier permet le forage de l'os pour les prothèses du genou sous le contrôle du robot afin d'éviter des zones prédéfinies [YD10]. Il utilise à la différence de Surgicobot une boucle de régulation hybride (coopérative) incluant un contrôle de la force exercée par le chirurgien, la position de l'effecteur, et un observateur permettant la compensation des perturbations. Ces développements permettent une erreur de suivi de l'ordre de $0.01\text{mm}\cdot\text{s}^{-1}$. Le second robot quant à lui permet d'obtenir des renseignements sur la pose de l'instrument, de le positionner à un endroit prédéfini, de lui imposer le suivi d'une trajectoire spécifique ou bien encore de restreindre l'espace admissible. Développé par l'équipe TIMC-IMAG [ST01], le PADyC développé pour les ponctions péricardiques permet de limiter à 15mm l'insertion maximale de l'aiguille et ainsi les risques d'endommagement du cœur.

Enfin, le dispositif MICRON ([ARK00], [BMHR11]) propose à l'image de [Vac11] une aide aux gestes en limitant les tremblements et les mouvements non désirés. Pour cela une estimation de la position 3D de l'effecteur, un réseau de neurones et un algorithme WFLC (weighted-frequency Fourier linear combiner) sont combinés. La précision obtenue est dans le cas unidirectionnel de $2.5\mu\text{m}$ et lors d'un mouvement libre de $11.2\mu\text{m}$. Les informations utilisées proviennent de trois accéléromètres et de trois gyroscopes. Même si des travaux supplémentaires comme ceux de [ISU+11] et [KBG+00] ont amélioré la partie robotique (utilisation d'un robot JHU à sept degrés de liberté), l'utilisation de capteurs au niveau de l'instrument ne rentre pas dans le cadre de cette étude.

En résumé :

Les systèmes coopératifs permettent de préserver la sécurité du patient et d'améliorer la précision du chirurgien tout en lui laissant l'initiative du geste. Dédiés principalement à des opérations orthopédiques et ophtalmologiques, ils nécessitent d'être guidés par l'opérateur durant toute la séquence. Ils ne sont donc pas adaptés à notre problématique, c'est pourquoi ils sont écartés de l'étude.

Le coût des systèmes robotisés

La quête de précision et de minimisation des traumatismes générés par l'acte chirurgical impose de plus en plus le passage à des systèmes robotisés. Cependant, les dispositifs actuels sont souvent très coûteux, ce qui explique pourquoi les recherches s'orientent vers des systèmes plus simples et ne nécessitant pas une infrastructure lourde. Pour ces raisons, la tendance dans ce domaine va en faveur des dispositifs directement implantables dans le corps humain utilisant, le plus souvent, les voies naturelles.

Parmi les systèmes passifs, les dispositifs de navigation sont un poste de dépense non négligeable pour l'hôpital. Composés d'une plateforme permettant la reconnaissance et/ou la localisation des tissus, d'une station logicielle où seront effectués les calculs (de type reconstruction) ainsi que la base de données (atlas en neurochirurgie) et enfin l'instrumentation (par exemple ancillaires), l'ensemble de ce matériel représente un coup estimé entre 95k€ et 115k€, dont les frais fixes par an s'élèvent à 15k€ pour la formation du personnel ainsi que la maintenance. Pour un système encore plus imposant comme le Da Vinci, le coût d'achat varie de 1.83M€ à 1.95M€ selon le nombre de bras et l'achat du système de vidéo HD et il faudra compter environ 10% du prix de vente pour le rendre opérationnel durant un an. Les consommables quant à eux s'élèvent à 2k€/instrument, leur durée d'opérabilité n'étant que de 10 interventions, à ceci s'ajoute d'autres frais annexes d'un montant moyen de 5k€. Dans ces conditions, on comprend bien que cette tendance vers une robotisation globale de la salle chirurgicale, évolue vers des systèmes parallèles moins chers et plus simples à l'usage. [LCA10] dans leur travaux proposent une comparaison des opérations à théâtre ouvert et mini-invasif pour le traitement des pathologies de l'aorte thoracique. Ils montrent que du strict point de vue de l'efficacité clinique il y a équivalence des taux de survie à un an (86.34% pour la chirurgie à théâtre ouvert contre 87.74% pour le mini-invasif) mais que les temps de séjour sont moitié moindres pour les opérations mini-invasives (15.88 jours contre 7.17 jours). Ceci justifie donc l'essor des robots impliqués dans ce type d'interventions. De plus, du point de vue financier, il y a une économie réelle faite par ce biais de l'ordre de 38% à 24.7% ce qui conforte d'autant plus cet avantage.

Annexe C

La chirurgie mini-invasive

Mode opératoire

De plus en plus populaire, la chirurgie mini-invasive (Minimally Invasive Surgery-MIS en anglais), diminue le temps de récupération ainsi que les traumatismes occasionnés à l'organisme. Par introduction continue de CO₂ dans l'abdomen du patient (tube en rouge dans la figure 1.2), il est possible de créer artificiellement une cavité (pneumopéritoine provoqué). L'orifice nécessaire, dont le diamètre ne dépasse pas 15mm, est effectué à l'aide d'un trocart qui a la double tâche d'assurer l'étanchéité et de limiter les sollicitations au niveau du point d'insertion. Une fois la cavité obtenue, les tissus cibles sont traités par deux instruments, symbolisés en vert, insérés par deux trous additionnels ce qui permet de ne pas recourir à une ouverture massive. L'absence de vision directe est compensée quant à elle, par une caméra dite endoscopique (en bleu).

La chirurgie mini-invasive se voit inappropriée pour des opérations effectuées dans l'urgence ou dans le cas de greffes par exemple. Cependant, elle s'avère préférable pour le patient pour les activités suivantes (exemples tirés du site chirurgie-mini-invasive.com) :

- **Gynécologie** (ablation d'un kyste ovarien, traitement d'une grossesse extra-utérine, reperméabilisation des trompes, hystérectomie, traitement d'un prolapsus) ;
- **Orthopédie** (prothèse totale de hanche, prothèse de genou, chirurgie rachidienne, disectomie, thoracoscopie, chirurgie du pied) ;
- **Chirurgie abdominale** (chirurgie de la vésicule biliaire, chirurgie de l'estomac, chirurgie de l'intestin, hernie inguinale, chirurgie de l'obésité) ;
- **Chirurgie cardiaque** (pontage coronarien, valvuloplastie) ;
- **Chirurgie urologique** (prostatectomie, adénomectomie, calculs urinaires, rétrécissement de la jonction pyélo-urétérale).

Remarque : *La chirurgie mini-invasive est généralement associée aux opérations où l'on insuffle un gaz pour créer une cavité dans le corps du patient, celle-ci pouvant être dans l'abdomen ou ailleurs. Cependant, il est possible de rencontrer ces termes lorsque que l'incision principale est de petites dimensions (et où aucun gaz n'est injecté).*

Avantages et inconvénients de la chirurgie mini-invasive

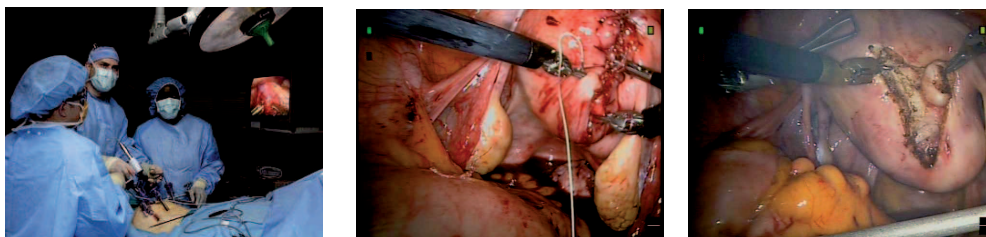
Les avantages de la chirurgie mini-invasive sont liés à la diminution du nombre de tissus mous disséqués, principale cause de douleurs chez le patient. De plus, outre une cicatrice plus esthétique grâce aux faibles dimensions de l'ouverture pratiquée, moins de vaisseaux sanguins sont coupés ce qui implique des saignements en proportion moindre et donc une diminution des douleurs post-opératoires. Le temps de récupération ainsi réduit, tend à populariser cette pratique au sein des hôpitaux qui réalisent là de substantielles économies. Ainsi pour un remplacement total de hanche, selon une communication du Centre Médico-Chirurgical du Mans (CMCM), la cicatrice par approche mini-invasive est de 5-8 cm contre 12-20 cm pour la chirurgie conventionnelle voire dans le cas d'un remplacement total de hanche de 6-8cm contre 25cm ; la durée d'hospitalisation quant à elle passe de 10-12 jours à 3-5 jours.

Cependant, si l'emploi de cette méthode est avantageuse pour le patient, elle se révèle l'être beaucoup moins pour l'équipe chirurgicale qui se heurte à des difficultés de mise en œuvre relatives à la technique elle-même, citons ici :

- une inversion des mouvements (effet fulcrum) due au passage de l'instrument dans le trocart ;
- un travail à distance (les instruments pouvant mesurer jusqu'à 50cm) engendrant une amplification des tremblements accentuée par la posture souvent inconfortable que prend le chirurgien ;
- une diminution de la précision des mouvements malgré les appuis possibles (trocart) ;
- une restriction de l'espace à un cône à base elliptique (le sommet des cônes étant le trocart), le point d'entrée de l'instrument étant supposé statique ;
- un réapprentissage des gestes à faire avec de surcroît une modification de la gestion de l'effort (en raison du type d'instrument utilisé) ;
- impossibilité de palper les organes ;
- création d'un volume faible en raison de la limite élastique de la peau ;
- l'absence de vision directe de la zone d'intérêt.

On entend par zone d'intérêt, la région que le chirurgien est en train de traiter. Bien qu'une caméra introduite dans la cavité lui en restitue l'image, plusieurs problèmes persistent. Le chirurgien ne pouvant dédier une main pour la manipulation de la caméra, il délègue donc cette activité à un assistant (instrumentiste, panseuse ou assistant chirurgien). Charge à lui de maintenir manuellement, la zone d'intérêt centrée dans l'image tout en gardant les instruments dans le champ de vision. Ceci occasionne des tremblements ainsi que des changements de point de vue fréquents ce qui complique la tâche du chirurgien. La perte de la profondeur du fait de l'utilisation d'une caméra, l'importance des saignements dans l'image, les fumées lors des cautérisations, la proximité des éléments, ainsi que la perte de coordination main-œil du fait de l'observation indirecte de la scène sont des gênes supplémentaires liées à l'emploi d'une caméra pour le chirurgien (l'écran ne se trouvant pas dans l'axe de regard naturel du chirurgien, figures C.1).

A ces difficultés s'ajoutent l'éclairage de la scène. Assuré par une source de lumière déportée raccordée à une fibre optique intégrée à l'endoscope, ce dispositif permet d'obtenir les conditions d'illumination rendant l'image exploitable par l'équipe. Cependant, cette source de lumière froide ponctuelle placée en bout d'endoscope (une lumière d'ambiance sur le patient étant appréciable pour l'équipe chirurgicale) génère des ombres marquées et limite la vision des cavités profondes (notamment pour l'intestin) comme le montre [DPDP03] qui parle d'"éclairage imparfait".



(a) Exemple d'opération mini-invasive (source : CFTC Santé-Sociaux Hôpitaux de Strasbourg)

(b) Vue qu'à le chirurgien sur l'écran (source : DaVinci)

(c) Génération de fumée lors d'une dissection (source : DaVinci)

FIG. C.1 – Exemples de chirurgie mini-invasive.

En résumé :

Les caméras ont donc un rôle déterminant pour les opérations mini-invasives. Cependant leur utilisation suppose des conditions d'éclairage suffisantes. Comme nous venons de le voir, ce couplage entraîne certaines difficultés propres à ce contexte. Nous verrons par la suite que résoudre le problème d'éclairage pour les opérations à théâtre ouvert pose des difficultés analogues.

Bibliographie

- [AB85] E.H. Adelson and J.R. Bergen. Spatiotemporal energy models for the perception of motion. *J. Opt. Soc. Am. A*, 2(2) :284–299, 1985.
- [Abb09] N. Abbas. *Développement de modèles de fusion et de classification contextuelle d’images satellitaires par la théorie de l’évidence et la théorie du raisonnement plausible et paradoxal*. PhD thesis, Faculté d’électronique et informatique, 2009.
- [ACT07] G. Allibert, E. Courtial, and Y. Touré. A flat model predictive controller for trajectory tracking in image based visual servoing. In *Nonlinear Control Systems*, volume 7, pages 993–998, 2007.
- [AGS⁺02] S. Aiono, J. Gilbert, B. Soin, P. Finlay, and A. Gordan. Controlled trial of the introduction of a robotic camera assistant (endo assist) for laparoscopic cholecystectomy. *Surgical Endoscopy*, 16 :1267–1270, 2002.
- [AGXX10] Y. Anbang, W. Guijin, L. Xinggang, and C. Xiujuan. An incremental bhattacharyya dissimilarity measure for particle filtering. *Pattern Recognition*, 43(4) :1244 – 1256, 2010.
- [AH79] M. Allais and O. Hagen. *Expected utility hypotheses and the Allais paradox : Contemporary discussions of decisions under uncertainty with Allais’ rejoinder*. Reidel, Dordrecht, The Netherlands, 1979.
- [Aka73] H. Akaike. Information theory and an extension of the maximum likelihood principle. In *Second international symposium on information theory*, volume 1, pages 267–281. Springer Verlag, 1973.
- [AKN⁺11] B.F. Allen, F. Kasper, G. Nataneli, E. Dutson, and P. Faloutsos. Visual tracking of laparoscopic instruments in standard training environments. *Studies in health technology and informatics*, 163 :11, 2011.
- [AMMM06] B. Al Momani, S. McClean, and P. Morrow. Using dempster-shafer to incorporate knowledge into satellite image classification. *Artificial Intelligence Review*, 25 :161–178, 2006.
- [App98] A. Appriou. Spatially ambiguous multisensor data processing. *Office National d’Etudes et de Recherches Aérospatiales*, 1998.
- [ARK00] W. Ang, C. Riviere, and P. Khosla. An active hand-held instrument for enhanced microsurgical accuracy. In *Medical Image Computing and Computer-Assisted Intervention*, pages 243–291. Springer, 2000.
- [Art08] N.M. Artner. A comparison of mean shift tracking methods. In *12th Central European Seminar on Computer Graphics*, pages 197–204, 2008.

- [ASZ⁺02] P. Abolmaesumi, S.E. Salcudean, W.H. Zhu, M.R. Siroospour, and S.P. DiMaio. Image-guided control of a robot for medical ultrasound. *IEEE International Conference on Robotics and Automation*, 18(1) :11–23, 2002.
- [Awa11] S.S. Awasthi, A. Chauhan. Using ahp and dempster-shafer theory for evaluating sustainable transport solutions. *Environmental Modelling & Software*, 26(6) :787 – 796, 2011.
- [Bal11] H. Baldwin. Open-sky (anterior approach) müller’s muscle resection for the correction of blepharoptosis. In Adam J. Cohen and David A. Weinberg, editors, *Evaluation and Management of Blepharoptosis*, pages 189–195. Springer New York, 2011.
- [BBB98] W.L. Bargar, A. Bauer, and M. Börner. Primary and revision total hip replacement using the robodoc (r) system. *Clinical orthopaedics and related research*, 354 :82, 1998.
- [BBC02] D. Bellot, A. Boyer, and F. Charpillat. A new definition of qualified gain in a data fusion process : application to telemedicine. In *Information Fusion*, volume 2, pages 865–872, 2002.
- [BBD⁺05] L. Barbé, B. Bayle, C. Doignon, J. Gangloff, M. de Mathelin, B. Maurin, P. Zanne, O. Piccin, P. Renaud, A. Gangi, et al. Irais : Insertion robotisée d’aiguille sous imagerie scanner robotized needle insertion with ct guidance. *3ème Journées du Programme interdisciplinaire de Recherche ROBEA*, pages 131–138, 2005.
- [BBF94] J.L. Barron, S.S. Beauchemin, and D.J. Fleet. On optical flow. In *6th Int. Conf. on Artificial Intelligence and Information-Control Systems of Robots*, pages 3–14, 1994.
- [BBJP10] K. Bunte, M. Biehl, M.F. Jonkman, and N. Petkov. Learning effective color features for content based image retrieval in dermatology. *Pattern Recognition*, 44 :1892–1902, 2010.
- [BBL97] A. Bauer, M. Börner, and A. Lahmer. Robodoc - animal experiment and clinical evaluation. In *CVRMed*, volume 1205 of *Lecture Notes in Computer Science*, pages 561–564. Springer, 1997.
- [BBM10] M. Busack, D. Bellot, and G. Morel. Breathing motion compensation for robot assisted laser osteotomy. In *IEEE International Conference on Robotics and Automation*, Anchorage, Alaska, 2010.
- [BBRW11] T. Beyl, T. Brennecke, J. Raczowsky, and H. Wörn. Ultrasound tomography using a light weight robot. *Computer Assisted Radiology*, 6 :355–356, 2011.
- [BC95] F. Bensalah and F. Chaumette. Compensation of abrupt motion changes in target tracking by visual servoing. In *IEEE International Conference on Intelligent Robots and Systems*, volume 1, pages 181–187, August 1995.
- [BCB⁺03] P. Berkelman, P. Cinquin, E. Boidard, J. Troccaz, C. Létoublon, and JM Ayoubi. Design, control and testing of a novel compact laparoscopic endoscope manipulator. *Proceedings of the Institution of Mechanical Engineers, Part I : Journal of Systems and Control Engineering*, 217(4) :329–341, 2003.
- [Bel11] B. Bell. A purpose-built robot for direct access cochlear implantation. *International Journal of Computer Assisted Radiology and Surgery*, 6 :129, 2011.

- [BF04] R. Boudjemaa and AB. Forbes. Parameter estimation methods for data fusion. *NPL Report CMSC*, 38(04), 2004.
- [BJHM01] P. Ballester, Y. Jain, K.R. Haylett, and R.F. McCloy. Comparison of task performance of robotic camera holders endoassist and aesop. *International Congress Series*, 1230(0) :1100 – 1103, 2001.
- [Blo81] S.A. Bloom. Similarity indices in community studies : potential pitfalls. *Mar. Ecol. Prog. Ser.*, 5(2) :125–128, 1981.
- [Blo05] I. Bloch. Fusion d'informations numériques : panorama méthodologique. *Journées Nationales de la Recherche en Robotique*, pages 79–88, 2005.
- [BM08] R. Belaroussi and G. Morel. Visual servoing for patient alignment in protontherapy. In *Advances in Visual Computing*, volume 5359 of *Lecture Notes in Computer Science*, pages 843–850. Springer Berlin / Heidelberg, 2008.
- [BMG97] F. Berry, P. Martinet, and J. Gallice. Trajectory generation by visual servoing. In *IEEE International Conference on Intelligent Robots and Systems*, volume 2, pages 1066–1072, 1997.
- [BMHR11] B.C. Becker, R.A. MacLachlan, G.D. Hager, and C.N. Riviere. Handheld micromanipulation with vision-based virtual fixtures. In *IEEE International Conference on Robotics and Automation*, pages 4127–4132, 2011.
- [BMM07] H. Bhaskar, L. Mihaylova, and S. Maskell. Background modeling using adaptive cluster density estimation for automatic human detection. *Lecture Notes in Informatics*, pages 130–134, 2007.
- [BMS07] R.J. Bayardo, Y. Ma, and R. Srikant. Scaling up all pairs similarity search, 2007.
- [BN07] L. Bocchi and J. Nori. Shape analysis of microcalcifications using radon transform. *Medical Engineering & Physics*, 29(6) :691 – 698, 2007.
- [BO00] M. Bedworth and J. O'Brien. The omnibus model : a new model of data fusion ? *IEEE Aerospace and Electronic Systems Magazine*, 15(4) :30–36, 2000.
- [Boy87] J.R. Boyd. Organic design for command and control. *A Discourse on Winning and Losing*, 1987.
- [BPR09] M. Balasubramanian, S. Palanivel, and V. Ramalingam. Real time face and mouth recognition using radial basis function neural networks. *Expert Systems with Applications*, 36(3) :6879–6888, 2009.
- [BPS⁺05] T. Bailloeuil, V. Prinet, B. Serra, P. Marthon, P. Chen, and H. Zhang. Digital building map refinement from knowledge-driven active contours and very high resolution optical imagery. *Photogrammetrie fernerkundung geoinformation*, 2005(6) :511, 2005.
- [BRL⁺07] W. Bachta, P. Renaud, E. Laroche, J. Gangloff, and A. Forgione. Cardiolock : An active cardiac stabilizer. *Medical Image Computing and Computer-Assisted Intervention*, pages 78–85, 2007.
- [BSF⁺11] S. Ben Chaabane, M. Sayadi, F. Fnaiech, E. Brassart, and F. Betin. A new method for the estimation of mass functions in the dempster-shafer's evidence theory : Application to colour image segmentation. *Circuits, Systems, and Signal Processing*, 30 :55–71, 2011.

- [BTVG06] H. Bay, T. Tuytelaars, and L. Van Gool. Speeded-up robust features. *Computer Vision*, pages 404–417, 2006.
- [Bur98] C.J.C. Burges. A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, 2(2) :121–167, 1998.
- [CA07] F. Chaumette and N. Andreff. Quoi de neuf en asservissement visuel depuis les jurr’03 ? In *Journées Nationales de la Recherche en Robotique, JNRR’07*, Obernai, October 2007.
- [Cah10] N. Cahill. Normalized measures of mutual information with general definitions of entropy for multimodal image registration. In *Biomedical Image Registration*, volume 6204 of *Lecture Notes in Computer Science*, pages 258–268. Springer Berlin / Heidelberg, 2010.
- [Cai92] R.D. Cairns. La recherche de rentes en situation d’incertitude avec ou sans opposition. *L’Actualité économique*, 68(3) :477–498, 1992.
- [CC00] A. Crétual and F. Chaumette. Dynamic stabilization of a pan and tilt camera for submarine image visualization. *Computer Vision and Image Understanding*, 79(1) :47–65, 2000.
- [CC01] A. Cretual and F. Chaumette. Visual servoing based on image motion. *Int. Journal of Robotics Research*, 20, 11 :857–877, 2001.
- [CC11] S. Chambon and A. Crouzil. Similarity measures for image matching despite occlusions in stereo vision. *Pattern Recognition*, 44(9) :2063 – 2075, 2011.
- [CCFM04] A.-S. Capelle, O. Colot, and C. Fernandez-Maloigne. Evidential segmentation scheme of multi-echo mr images for the detection of brain tumors using neighborhood information. *Information Fusion*, 5(3) :203 – 216, 2004.
- [CCO08] A. Cherubini, F. Chaumette, and G. Oriolo. A position-based visual servoing scheme for following paths with nonholonomic mobile robots. In *IEEE International Conference on Intelligent Robots and Systems*, pages 1648–1654, 2008.
- [CDBR10] L.H. Cherif, S.M. Debbal, and F. Berekci-Reguig. Choice of the wavelet analyzing in the phonocardiogram signal analysis using the discrete and the packet wavelet transform. *Expert Systems with Applications*, 37(2) :913 – 918, 2010.
- [CDL⁺01] P. Cinquin, Emmanuel Dubois, S. Lavallée, Laurence Nigay, and J. Troccaz. Surgétique. de l’intérêt de la coopération entre médecin et machine au service du malade dans la réalisation de gestes médico-chirurgicaux. *Les Cahiers du Numérique-L’information médicale numérique*, 2(2) :197–207, 2001.
- [CE95] J.M. Chassery and Y. Elomary. Coopération contour actif et multirésolution en segmentation d’images. In *15° Colloque sur le traitement du signal et des images*. GRETSI, Groupe d’Etudes du Traitement du Signal et des Images, 1995.
- [CELR08] A. Choksuriwong, B. Emile, H. Laurent, and C. Rosenberger. Comparative study of global invariant descriptors for object recognition. *Journal of Electronic Imaging*, 17 :023015, 2008.
- [CGVRGPSP10] J.M. Chaves-González, M.A. Vega-Rodríguez, J.A. Gómez-Pulido, and J.M. Sánchez-Pérez. Detecting skin in face recognition systems : A colour spaces study. *Digital Signal Processing*, 20(3) :806–823, 2010.

- [Cha90] F. Chaumette. *La relation vision-commande : théorie et application à des tâches robotiques*. PhD thesis, Université de Rennes 1, 1990.
- [Cha98a] F. Chaumette. *De la perception à l'action : l'asservissement visuel, de l'action à la perception : la vision active*. PhD thesis, Habilitation à diriger des recherches de l'Université de Rennes 1, 1998.
- [Cha98b] F. Chaumette. Potential problems of stability and convergence in image-based and position-based visual servoing. *The confluence of vision and control*, pages 66–78, 1998.
- [CHC⁺10] C.S. Chen, M.S. Hsieh, Y.W. Chiu, C.H. Tsai, S.M. Liu, C.C. Lu, and P.L. Yen. An unconstrained virtual bone clasper for a knee surgical robot using visual servoing technique. *Journal of the Chinese Institute of Engineers*, 33(3) :379–386, 2010.
- [Cho07] M. Choras. Human lips recognition. In *Computer Recognition Systems 2*, volume 45 of *Advances in Soft Computing*, pages 838–843. Springer Berlin / Heidelberg, 2007.
- [CHYQNS04] J. Chun-Hong, Y. Yong-Qiang, C. Nagashima, and S. Sato. Optical design of infrared camera cascaded. *Chinese Astronomy and Astrophysics*, 28(2) :238 – 252, 2004.
- [CJ10] G. Chowdhary and R. Jategaonkar. Aerodynamic parameter estimation from flight data applying extended and unscented kalman filter. *Aerospace Science and Technology*, 14(2) :106 – 117, 2010.
- [CJH01] K.M. Cho, J.H. Jang, and K.S. Hong. Adaptive skin-color filter. *Pattern Recognition*, 34(5) :1067–1073, 2001.
- [CM04] J. Climent and P. Marés. Automatic instrument localization in laparoscopic surgery. *ELCVIA*, 4(1) :21–31, 2004.
- [CMT87] D.W. Clarke, C. Mohtadi, and P.S. Tuffs. Generalized predictive control—part i. the basic algorithm. *Automatica*, 23(2) :137–148, 1987.
- [COB03] T.S. Caetano, S.D. Olabarriaga, and D.A.C. Barone. Do mixture models in chromaticity space improve skin detection? *Pattern Recognition*, 36(12) :3019–3021, 2003.
- [Cor01] Intel Corporation. Open source computer vision library. 2001.
- [Cox46] R.T. Cox. Probability, frequency and reasonable expectation. *American journal of physics*, 14(1) :1–13, 1946.
- [CRM04] C.W. Chong, P. Raveendran, and R. Mukundan. Translation and scale invariants of legendre moments. *Pattern Recognition*, 37(1) :119–129, 2004.
- [CY09] J.Y. Choi and Y.K. Yang. Vehicle detection from aerial images using local shape information. In *Advances in Image and Video Technology*, volume 5414 of *Lecture Notes in Computer Science*, pages 227–236. Springer Berlin / Heidelberg, 2009.
- [Das97] B.V. Dasarthy. Sensor fusion potential exploitation-innovative architectures and illustrative applications. *Proceedings of the IEEE*, 85(1) :24–38, 1997.
- [Das12] B. V. Dasarthy. Information fusion in the realm of medical applications - a bibliographic glimpse at its growing appeal. *Information Fusion*, 13(1) :1 – 9, 2012.
- [DB06] P. Danes and D. Bellot. Towards an lmi approach to multicriteria visual servoing in robotics. *European journal of control*, 12(1) :86, 2006.

- [DDPP03] E. Dombre, G. Duchemin, P. Poignet, and F. Pierrot. Dermarob : a safe robot for reconstructive surgery. *IEEE International Conference on Robotics and Automation*, 19(5) :876–884, 2003.
- [Den97] T. Denoeux. Application du modèle des croyances transférables en reconnaissance de formes. 14(5) :443–451, 1997.
- [Den00] T. Denoeux. A neural network classifier based on dempster-shafer theory. *Systems, Man and Cybernetics, Part A : Systems and Humans, IEEE Transactions on*, 30(2) :131–150, 2000.
- [DHKHC04] D.Y.K. Do Hyong Kim, H.S. Hong, and M.J. Chung. An image-based control scheme for an active stereo vision system. In *IEEE International Conference on Intelligent Robots and Systems*, 2004.
- [DJ08] W. Dingrui and Z. Jie. Stereo vision using two ptz cameras. *Computer Vision and Image Understanding*, 112(2) :184 – 194, 2008.
- [DKC11] A. Dutta, A. Kar, and B.N. Chatterji. A new approach to corner matching from image sequence using fuzzy similarity index. *Pattern Recognition Letters*, 32(5) :712 – 720, 2011.
- [DKKS07] D. Dimitrov, C. Knauer, K. Kriegel, and F. Stehn. Approximation algorithms for a point-to-surface registration problem in medical navigation. *Frontiers in Algorithmics*, pages 26–37, 2007.
- [DM10] A. Dame and E. Marchand. Improving mutual information-based visual servoing. In *IEEE International Conference on Robotics and Automation*, pages 5531–5536, 2010.
- [DM11] P. Dunne and B. Matuszewski. Choice of similarity measure, likelihood function and parameters for histogram based particle filter tracking in cctv grey scale video. *Image and Vision Computing*, 29(2-3) :178 – 189, 2011.
- [DMBdM06] C. Doignon, B. Maurin, B. Bayle, and M. de Mathelin. Vers un asservissement visuel en tomodesitométrie à rayon x. In *Actes de la conférence francophone Reconnaissance des Formes et Intelligence Artificielle (RFIA)*, 2006.
- [DNdM07] C. Doignon, F. Nageotte, and M. de Mathelin. Segmentation and guidance of multiple rigid objects for intra-operative endoscopic vision. *Dynamical Vision*, pages 314–327, 2007.
- [Dom02] E. Dombre. Projet marge : modélisation, apprentissage et reproduction du geste endochirurgical. In *Proc. Journées ROBEA*, 2002.
- [DP88] D. Dubois and H. Prade. Representation and combination of uncertainty with belief functions and possibility measures. *Computational Intelligence*, 4(3) :244–264, 1988.
- [DPDP03] E. Dombre, P. Poignet, G. Duchemin, and F. Pierrot. Quelques problemes de conception et commande de robots medicaux. *Mécanique et Industries*, 4(5) :559–563, 2003.
- [DS09] J. Dezert and F. Smarandache. An introduction to dsmt. *CoRR*, 2009.
- [DW87] H.F. Durrant-Whyte. Consistent integration and propagation of disparate sensor observations. *The International journal of robotics research*, 6(3) :3, 1987.
- [EB11] J. Egbert and R.W. Beard. Low-altitude road following using strap-down cameras on miniature air vehicles. *Mechatronics*, 21(5) :831 – 843, 2011.

- [ECR93] B. Espiau, F. Chaumette, and P. Rives. A new approach to visual servoing in robotics. *Geometric Reasoning for Perception and Action*, pages 106–136, 1993.
- [EGS96] L. El Ghaoui and G. Scorletti. Control of rational systems using linear-fractional representations and linear matrix inequalities. *Automatica*, 32(9) :1273–1284, 1996.
- [EHD00] A. Elgammal, D. Harwood, and L. Davis. Non-parametric model for background subtraction. *Computer Vision-ECCV 2000*, pages 751–767, 2000.
- [Elm07] W. Elmenreich. A review on system architectures for sensor fusion applications. *Software Technologies for Embedded and Ubiquitous Systems*, pages 547–559, 2007.
- [End95] M.R. Endsley. Toward a theory of situation awareness in dynamic systems. *Human Factors : The Journal of the Human Factors and Ergonomics Society*, 37(1) :32–64, 1995.
- [EZ10] A. El-Zaart. Expectation-maximization technique for fibro-glandular discs detection in mammography images. *Computers in Biology and Medicine*, 40(4) :392–401, 2010.
- [FACM09] O. Felfoul, E. Aboussouan, A. Chanu, and S. Martel. Real-time positioning and tracking technique for endovascular untethered microrobots propelled by mri gradients. In *IEEE International Conference on Robotics and Automation*, pages 2693–2698, 2009.
- [Fau93] O. Faugeras. *Three-dimensional computer vision : a geometric viewpoint*. the MIT Press, 1993.
- [FBFSB⁺98] J. Ferré-Borrull, X. Fernandez, S. Salvador Bosch, S. Vallmitjana, and E. Martín-Badosa. Post-processing of joint-transform correlator outputs with the kolmogorov-smirnov statistical test. *Optics Communications*, 147 :359 – 365, 1998.
- [FCM00] G. Flandin, F. Chaumette, and E. Marchand. Eye-in-hand/eye-to-hand cooperation for visual servoing. In *Robotics and Automation*, volume 3, pages 2741–2746, 2000.
- [FDV⁺06] M.C. Florea, J. Dezert, P. Valin, F. Smarandache, and A.L. Joussetme. Adaptive combination rule and proportional conflict redistribution rule for information fusion. *CoRR*, 2006.
- [FGHP03] P.A. Federspil, U.W. Geithoff, D. Henrich, and P.K. Plinkert. Development of the first force-controlled robot for otoneurosurgery. *The Laryngoscope*, 113(3) :465–471, 2003.
- [FJGB02] M.C. Florea, A.L. Joussetme, D. Grenier, and E. Bosse. Unified approach to the fusion of imperfect data? In *Proceedings of SPIE*, volume 4731, page 75, 2002.
- [FL05] F. Faux and F. Luthon. Modélisation de la teinte du visage par fusion d’informations couleur dans le cadre de la théorie de dempster-shafer. *CORESA ’05*, 2005.
- [FL07] P.E. Forssen and D.G. Lowe. Shape descriptors for maximally stable extremal regions. In *IEEE International Conference on Computer Vision*, pages 1–8, 2007.
- [FLK11] N.E.E. Faouzi, H. Leung, and A. Kurian. Data fusion in intelligent transportation systems : Progress and challenges—a survey. *Information Fusion*, 12(1) :4–10, 2011.
- [FWV05] D. Francois, V. Wertz, and M. Verleysen. Non-euclidean metrics for similarity search in noisy datasets. In *Proceedings of the European symposium on artificial neural networks*, pages 339–334, 2005.

- [GDM00] J. Gangloff and M. De Mathelin. High speed visual servoing of a 6 dof manipulator using mimo predictive control. In *IEEE International Conference on Robotics and Automation*, volume 4, pages 3751–3756, 2000.
- [GGdM⁺04] R. Ginhoux, JA Gangloff, M. de Mathelin, L. Soler, M.M.A. Sanchez, and J. Marescaux. Beating heart tracking in robotic surgery using 500 hz visual servoing, model predictive control and an adaptive observer. In *IEEE International Conference on Robotics and Automation*, volume 1, pages 274–279, 2004.
- [GGMT10] I. Guarneri, M. Guarnera, G. Messina, and V. Tomaselli. A signature analysis based method for elliptical shape. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 7537, page 20, 2010.
- [GH07] N.R. Gans and S.A. Hutchinson. Stable visual servoing through hybrid switched-system control. *Robotics, IEEE Transactions on*, 23(3) :530–540, 2007.
- [GH08] N.R. Gans and S.A. Hutchinson. Multi-attribute utility analysis in the choice of a vision-based robot controller. In *IEEE International Conference on Intelligent Robots and Systems*, volume 1, pages 355–362, 2008.
- [GHT11] S. Gauglitz, T. Höllerer, and M. Turk. Evaluation of interest point detectors and feature descriptors for visual tracking. *International Journal of Computer Vision*, 94 :335–360, 2011.
- [Gin03] R. Ginhoux. *Compensation des mouvements physiologiques en chirurgie robotisée par commande prédictive*. These, Université Louis Pasteur - Strasbourg I, December 2003.
- [GJGMAR⁺11] P. Gil-Jiménez, H. Gómez-Moreno, J. Acevedo-Rodríguez, R.J. López-Sastre, and S. Maldonado-Bascón. Evaluation of shape classification techniques based on the signature of the blob. *Signal Processing*, 2011.
- [GN98] J. Gluckman and S.K. Nayar. Ego-motion and omnidirectional cameras. In *Computer Vision, 1998. Sixth International Conference on*, pages 999–1005. IEEE, 1998.
- [GP11] E.L. Geist and T. Parsons. Assessing historical rate changes in global tsunami occurrence. *Geophysical Journal International*, 2011.
- [GR05] A. Goltsev and D. Rachkovskij. Combination of the assembly neural network with a perceptron for recognition of handwritten digits arranged in numeral strings. *Pattern Recognition*, 38(3) :315 – 322, 2005.
- [GRG06] H. Greenspan, A. Ruf, and J. Goldberger. Constrained gaussian mixture model framework for automatic segmentation of mr brain images. *Medical Imaging, IEEE Transactions on*, 25(9) :1233–1245, 2006.
- [GSKM10] D.S. Guru, Y.H. Sharath Kumar, and S. Manjunath. Textural features in flower classification. *Mathematical and Computer Modelling*, 2010.
- [GTSY10] A. Goldenberg, J. Trachtenberg, M. Sussman, and Y. Yi. Mri-guided robot-assisted prostatic interventions. In *International Congress and Exhibition, Computer Assisted Radiology and Surgery*, page S23, 2010.
- [GW03] P.C. Gerszten and W.C. Welch. Cyberknife radiosurgery for the spine. *Tech Neurosurg*, 9(3) :232–241, 2003.

- [GZ10] L.Q. Guo and M. Zhu. Quaternion fourier-mellin moments for color images. *Pattern Recognition*, 2010.
- [HA94] K. Hosoda and M. Asada. Versatile visual servoing without knowledge of true jacobian. In *IEEE International Conference on Intelligent Robots and Systems*, volume 1, pages 186–193, 1994.
- [HA01] H. Hua and N. Ahuja. A high-resolution panoramic camera. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages I–960, 2001.
- [HC03] M. Hansen and J. Carstensen. Color-based image retrieval from high-similarity image databases. In *Image Analysis*, volume 2749, pages 1098–1105. Springer Berlin / Heidelberg, 2003.
- [HD11] R. Doubet J. Ryan M. Huettner, F. Alley and D. Dynda. Robot assisted foregut surgery : one surgeon’s. *International Journal of Computer Assisted Radiology and Surgery*, pages 105–106, 2011.
- [HDH⁺04] J. Hong, T. Dohi, M. Hashizume, K. Konishi, and N. Hata. An ultrasound-driven needle-insertion robot for percutaneous cholecystostomy. *Physics in Medicine and Biology*, 49 :441, 2004.
- [Hee88] D.J. Heeger. Optical flow using spatiotemporal filters. *International Journal of Computer Vision*, 1(4) :279–302, 1988.
- [HHC96] S. Hutchinson, G.D. Hager, and P.I. Corke. A tutorial on visual servo control. *IEEE International Conference on Robotics and Automation*, 12(5) :651–670, 1996.
- [HHM⁺07] Y. Hafizal, Y. Haniza, H. Mohd, M. Shukri, A.A. Mohamed, M. Rizon, and S. Sayuti. Circular discontinuities detection in welded joints using circular hough transform. *NDT & E International*, 40(8) :594 – 601, 2007.
- [HHW05] Jung-Ning Huang, Pi-Fuei Hsieh, and Chung-Hsien Wu. Vision-based recognition of hand shapes in taiwanese sign language. In Jianhua Tao, Tieniu Tan, and Rosalind Picard, editors, *Affective Computing and Intelligent Interaction*, volume 3784 of *Lecture Notes in Computer Science*, pages 224–231. Springer Berlin / Heidelberg, 2005.
- [HLY10] M. Huang, W. Li, and W. Yan. Estimating parameters of synchronous generators using square-root unscented kalman filter. *Electric Power Systems Research*, 80(9) :1137 – 1144, 2010.
- [HM69] J.E. Handschin and D.Q. Mayne. Monte carlo techniques to estimate the conditional expectation in multi-stage non-linear filtering. *International journal of control*, 9(5) :547–559, 1969.
- [HMS04] S. Le Hégarat-Masclé and R. Seltz. Automatic change detection by evidential fusion of change indices. *Remote Sensing of Environment*, 91(3-4) :390 – 404, 2004.
- [HP79] J. Hill and W.T. Park. Real time control of a robot with a mobile camera. In *Proc. 9th ISIR*, pages 233–246. Washington, DC, 1979.
- [HS81] B.K.P. Horn and B.G. Schunck. Determining optical flow. *Artificial intelligence*, 17(1-3) :185–203, 1981.
- [HS10] S. Huh and D. H. Shim. A vision-based landing system for small unmanned aerial vehicles using an airbag. *Control Engineering Practice*, 18(7) :812 – 823, 2010.

- [Hu62] M.K. Hu. Visual pattern recognition by moment invariants. *Information Theory, IRE Transactions on*, 8(2) :179–187, 1962.
- [Hyn08] R.J. Hyndman. Etc5410 : Nonparametric smoothing methods. 2008.
- [ICST11] H. Iyatomi, M.E. Celebi, G. Schaefer, and M. Tanaka. Automated color calibration method for dermoscopy images. *Computerized Medical Imaging and Graphics*, 35(2) :89–98, 2011.
- [Ina91] T. Inagaki. Interdependence between safety-control policy and multiple-sensor schemes via dempster-shafer theory. *Reliability, IEEE Transactions on*, 40(2) :182–188, 1991.
- [IRD98] IRDES. La chirurgie ambulatoire en france bilan et perspectives. *Bulletin d'information en économie de la santé*, 1998.
- [ISU⁺11] Y. Ida, N. Sugita, T. Ueta, Y. Tamaki, K. Tanimoto, and M. Mitsuishi. Microsurgical robotic system for vitreoretinal surgery. *International Journal of Computer Assisted Radiology and Surgery*, pages 1–8, 11.
- [IW11] D.E. Ilea and P.F. Whelan. Image segmentation based on the integration of colour-texture descriptors-a review. *Pattern Recognition*, 2011.
- [JBHR06] L. Jones, M.J. Beynon, C.A. Holt, and S. Roy. An application of the dempster-shafer theory of evidence to the classification of knee function and detection of improvement due to total knee replacement surgery. *Journal of biomechanics*, 39(13) :2512–2520, 2006.
- [JCK⁺05] J.Y. Jenny, U. Clemens, S. Kohler, H. Kiefer, W. Konermann, and R.K. Miehle. Consistency of implantation of a total knee arthroplasty with a non-image-based navigation system : A case-control study of 235 cases compared with 235 conventionally implanted prostheses. *The Journal of arthroplasty*, 20(7) :832–839, 2005.
- [JCS11] J. Junhong, Guodong C., and Lining S. A novel hough transform method for line detection by enhancing accumulator array. *Pattern Recognition Letters*, 32(11) :1503 – 1510, 2011.
- [JF96] F. Janez and J.L. Ferrier. Fusion de sources d'information définies sur des référentiels non exhaustifs différents. solutions proposées sous le formalisme de la théorie de l'évidence. 1996.
- [JJ06] A. Jamshaid and F. Jiancheng. Sins/ans integration for augmented performance navigation solution using unscented kalman filtering. *Aerospace Science and Technology*, 10(3) :233 – 238, 2006.
- [JLS04] A. T. B. Jin, D.N.C. Ling, and O.T. Song. An efficient fingerprint verification system using integrated wavelet and fourier-mellin invariant transform. *Image and Vision Computing*, 22(6) :503 – 513, 2004.
- [JU97] S.J. Julier and J.K. Uhlmann. A new extension of the kalman filter to nonlinear systems. In *Int. Symp. Aerospace/Defense Sensing, Simul. and Controls*, volume 3, page 26, 1997.
- [Kal60] R.E. Kalman. A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82(1) :35–45, 1960.

- [KAS04] M.K. Kiyimik, M. Akin, and A. Subasi. Automatic recognition of alertness level by using wavelet transform and artificial neural network. *Journal of neuroscience methods*, 139(2) :231–240, 2004.
- [KB08] K. Kühnlenz and M. Buss. On sensor switching visual servoing. *International Journal of Optomechatronics*, 2(3) :233–256, 2008.
- [KBG⁺00] R. Kumar, P. Berkelman, P. Gupta, A. Barnes, P.S. Jensen, L.L. Whitcomb, and R.H. Taylor. Preliminary experiments in cooperative human/robot force control for robot assisted microsurgical manipulation. In *IEEE International Conference on Robotics and Automation*, volume 1, pages 610–617, 2000.
- [KC02] D. Kragic and H.I. Christensen. Survey on visual servoing for manipulation. *Computational Vision and Active Perception Laboratory*, 15, 2002.
- [KCH02] G.K. Klute, J.M. Czerniecki, and B. Hannaford. Artificial muscles : Actuators for biorobotic systems. *The International Journal of Robotics Research*, 21(4) :295–309, 2002.
- [KCW⁺02] Y. Kobayashi, S. Chiyoda, K. Watabe, M. Okada, and Y. Nakamura. Small occupancy robotic mechanisms for endoscopic surgery. *Medical Image Computing and Computer-Assisted Intervention*, pages 75–82, 2002.
- [KEB⁺03] W. Korb, D. Engel, R. Boesecke, G. Eggers, B. Kotrikova, R. Marmulla, J. Raczkowski, H. Wörn, J. Mühling, and S. Hassfeld. Development and first patient trial of a surgical robot for complex trajectory milling. *Computer Aided Surgery*, 8(5) :247–256, 2003.
- [Ken74] D. G. Kendall. Foundations of a theory of random sets. *Stochastic Geometry*, pages 322–376, 1974.
- [KFH09] A. Krupa, G. Fichtinger, and G.D. Hager. Real-time motion stabilization with b-mode ultrasound using image speckle information and visual servoing. *The International Journal of Robotics Research, IJRR*, 28(10) :1334–1354, 2009.
- [KGD⁺03] A. Krupa, J. Gangloff, C. Doignon, M. de Mathelin, G. Morel, J. Leroy, L. Soler, and J. Marescaux. Autonomous 3-d positioning of surgical instruments in robotized laparoscopic surgery using visual servoing. *IEEE International Conference on Robotics and Automation*, 19(5) :842–853, 2003.
- [KHK⁺95] T. Kohonen, J. Hynninen, J. Kangas, J. Laaksonen, and K. Torkkola. Lvq-pak. *The learning vector quantization program package. Version*, 3, 1995.
- [Kir71] R.A. Kirsch. Computer determination of the constituent structure of biological images. *Computers and biomedical research*, 4(3) :315–328, 1971.
- [Klo11] T. Klove. Lower bounds on the size of spheres of permutations under the chebychev distance. *Designs, Codes and Cryptography*, 59 :183–191, 2011.
- [KM11] S. Liao H. Kobayashi E. Kim, H. Joung and N. Motomura. Stereo endoscopic image-based slave detection method for handeye coordination of surgical robot. *International Journal of Computer Assisted Radiology and Surgery*, 6 :128, 2011.
- [KMB07] P. Kakumanu, S. Makrogiannis, and N. Bourbakis. A survey of skin-color modeling and detection methods. *Pattern Recognition*, 40(3) :1106–1122, 2007.

- [KMJ11] A. Knulst, R. Mooijweer, and F. W. Jansen. Indicating shortcomings in surgical lighting systems. *Minimally Invasive Therapy & Allied Technologies*, 20 :267–275, 2011.
- [Kni21] F. H. Knight. *Risk, Uncertainty, and Profit*. Houghton Mifflin Compagny, 1921.
- [Koh82] T. Kohonen. Self-organized formation of topologically correct feature maps. *Biological cybernetics*, 43(1) :59–69, 1982.
- [Kon04] H. Hirzinger G. Engelke R. Konietschke, R. Ortmaier T. Weiss. Manipulability and accuracy measures for a medical robot in minimally invasive surgery. In *9th International Symposium on Advances in Robot Kinematics (ARK)*, 2004.
- [Kos90] B. Kosko. Fuzziness vs. probability. *International Journal of General Systems*, 17(2-3) :211–240, 1990.
- [KPS03] J. Kovac, P. Peer, and F. Solina. *Human skin color clustering for face detection*, volume 2. IEEE, 2003.
- [Kru09] A. Krupa. Asservissement visuel par imagerie médicale. In *Journées Nationales de la Recherche en Robotique, JNRR'09*, Neuvy-sur-Barangeon, France, November 2009.
- [KS02] C. Kan and M.D. Srinath. Invariant character recognition with zernike and orthogonal fourier-mellin moments. *Pattern Recognition*, 35(1) :143–154, 2002.
- [KS04] Y. Ke and R. Sukthankar. Pca-sift : A more distinctive representation for local image descriptors. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2 :506, 2004.
- [KT10] S. Kurz and N. Tautenhahn. On dedekind's problem for complete simple games. 2010.
- [KW01] H. Kang and J.T. Wen. Endobot : a robotic assistant in minimally invasive surgeries. In *IEEE International Conference on Robotics and Automation*, volume 2, pages 2031–2036, 2001.
- [KY95] G.J. Klir and B. Yuan. *Fuzzy sets and fuzzy logic : theory and applications*. Prentice Hall PTR New Jersey, 1995.
- [LBR⁺04] J. Llinas, C. Bowman, G. Rogova, A. Steinberg, E. Waltz, F. White, Space, and Naval Warfare Systems Command. *Revisiting the JDL Data Fusion Model II*. 2004.
- [LC03] Jouve E. El Barnoussi A. Laporte C., Biette G. Les différentes voies d'abord pour l'ostéosynthèse par plaques des deux tiers distaux de l'humérus. *Maîtrise Orthopédique n°128*, 2003.
- [LCA10] B. Leroy, H. Constant, and G. Aulagner. Analyse "coût-efficacité" de la voie endovasculaire pour le traitement des pathologies de l'aorte thoracique. In *20ème Journées nationales sur les dispositifs médicaux*, 2010.
- [LCV02] E. Lefevre, O. Colot, and P. Vannoorenberghe. Belief function combination and conflict management. *Information fusion*, 3(2) :149–162, 2002.
- [LDB01] O. Vannoorenberghe P. Lefevre, E. Colot and D. De Brucq. Informations et combinaison : les liaisons conflictuelles. *Revue Traitement du Signal*, 18(3) :161–177, 2001.
- [LDS⁺06] J.A. Long, J.L. Descotes, O. Skowron, J. Troccaz, P. Cinquin, B. Boillot, N. Terrier, and J.J. Rambeaud. Utilisation de la robotique en chirurgie laparoscopique urologique : état de l'art. *Progrès en Urologie*, 16 :3–11, 2006.

- [Liu04] C. Liu. *A general measure of uncertainty-based information*. PhD thesis, Université Laval, 2004.
- [LJDC04] J.T. Lapresté, F. Jurie, M. Dhome, and F. Chaumette. An efficient method to compute the inverse jacobian matrix in visual servoing. In *IEEE International Conference on Robotics and Automation*, volume 1, pages 727–732, 2004.
- [LK] R.C. Luo and M.G. Kay. A tutorial on multisensor integration and fusion. In *16th Annual Conference of IEEE Industrial Electronics Society*, pages 707–722. IEEE.
- [LK81] B.D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. 1981. In *DARPA Image Understanding Workshop*, 1981.
- [LKGD08] C. Li, C.Y. Kao, J.C. Gore, and Z. Ding. Minimization of region-scalable fitting energy for image segmentation. *Image Processing, IEEE Transactions on*, 17(10) :1940–1949, 2008.
- [LKK⁺03] Y.J. Lee, J. Kim, S.Y. Ko, W.J. Lee, and D.S. Kwon. Design of a compact laparoscopic assistant robot : Kalar. *Proc. Int. Conf. on Control, Automation and Systems*, pages 2648–2653, 2003.
- [LKO⁺07] D. Lee, N. Koizumi, K. Ota, S. Yoshizawa, A. Ito, Y. Kaneko, Y. Matsumoto, and M. Mitsuishi. Ultrasound-based visual servoing system for lithotripsy. In *IEEE International Conference on Intelligent Robots and Systems*, pages 877–882, 2007.
- [LL09] J. Li and B.L Lu. An adaptive image euclidean distance. *Pattern Recognition*, 42(3) :349 – 357, 2009.
- [LLR10] I. Leichter, M. Lindenbaum, and E. Rivlin. Mean shift tracking with multiple reference color histograms. *Computer Vision and Image Understanding*, 114(3) :400–408, 2010.
- [LLR⁺11] C.A. Linte, P. Lang, M.E. Rettmann, D.S. Cho, D.R. Holmes, R.A. Robb, and T.M. Peters. Accuracy considerations in image-guided cardiac interventions : experience and lessons learned. *International Journal of Computer Assisted Radiology and Surgery*, page S161, 2011.
- [LMK⁺06] H. Laanaya, A. Martin, A. Khenchaf, D. Aboutajdine, et al. Classification par régression floue et crédibiliste à base de machines à vecteurs de support. 2006.
- [LN00] M. Loser and N. Navab. A new robotic system for visually controlled percutaneous interventions under ct fluoroscopy. In *Medical Image Computing and Computer-Assisted Intervention*, volume 1935 of *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg, 2000.
- [Low99] D.G. Lowe. Object recognition from local scale-invariant features. In *7th IEEE International Conference on Computer Vision*, volume 2, pages 1150–1157, 1999.
- [LS10] T. Li and W. Ser. Three dimensional acoustic source localization and tracking using statistically weighted hybrid particle filtering algorithm. *Signal Processing*, 90(5) :1700 – 1719, 2010.
- [LSR08] RE Larraín, DM Schaefer, and JD Reed. Use of digital images to estimate cie color coordinates of beef. *Food Research International*, 41(4) :380–385, 2008.

- [LTWCWN08] K.S. Leung, N. Tang, L. WH Cheung, and E. WK NG. Robotic arm in orthopaedic trauma surgery-early clinical experience and a review. *Pb Journal of Orthopaedics*, X :5–9, 2008.
- [LY02] J.Y. Lee and S.I. Yoo. An elliptical boundary model for skin color detection. In *Proc. of the 2002 International Conference on Imaging Science, Systems, and Technology*, 2002.
- [LYHH11] H. Luo, S.-L. Yang, X.-J. Hu, and X.-X. Hu. Agent oriented intelligent fault diagnosis system using evidence theory. *Expert Systems with Applications*, 3(0) :2524–2531, 2011.
- [LYY⁺10] S.K. Lam, C.Y. Yeong, C.T. Yew, W.S. Chai, and S.A. Suandi. A study on similarity computations in template matching technique for identity verification. *International Journal on Computer Science and Engineering*, 2(8) :2659–2665, 2010.
- [Mac85] D.L. MacAdam. *Color measurement : theme and variations*. Springer-Verlag, 1985.
- [Mal04] E. Malis. Improving vision-based control using efficient second-order minimization techniques. In *IEEE International Conference on Robotics and Automation*, volume 2, pages 1843–1848, 2004.
- [Mar05] A. Martin. Fusion d’informations haut niveau-application à la classification d’images sonar. *Atelier Fouille de données complexes-Extraction et gestion de connaissance (EGC), Paris, France*, 2005.
- [MB10] H. Mandana and A. Borji. Invariance analysis of modified c2 features : case study-handwritten digit recognition. *Machine Vision and Applications*, 21 :969–979, 2010.
- [MBP⁺08] B. Maurin, B. Bayle, O. Piccin, J. Gangloff, M. de Mathelin, C. Doignon, P. Zanne, and A. Gangi. A patient-mounted robotic platform for ct-scan guided procedures. *IEEE Transactions on Biomedical Engineering*, 55(10) :2417–2425, October 2008.
- [MC00] Y. Mezouar and F. Chaumette. Génération de trajectoire dans l’image : vers l’asservissement visuel robuste. In *12ème Journées des Jeunes Chercheurs en Robotique*, pages 131–137, February 2000.
- [MC02] Y. Mezouar and F. Chaumette. Path planning for robust image-based control. *IEEE International Conference on Robotics and Automation*, 18(4) :534–549, August 2002.
- [MC05] E. Marchand and F. Chaumette. Feature tracking for visual servoing purposes. *IEEE International Conference on Intelligent Robots and Systems*, 52(1) :53–70, 2005.
- [MC08] M. Marey and F. Chaumette. Analysis of classical and new visual servoing control laws. In *IEEE International Conference on Robotics and Automation*, pages 3244–3249, 2008.
- [MCB98] E. Malis, F. Chaumette, and S. Boudet. Positioning a coarse-calibrated camera with respect to an unknown object by 2d 1/2 visual servoing. In *IEEE International Conference on Robotics and Automation*, volume 2, pages 1352–1359, 1998.
- [MCF05] S.J. McKenna, H.N. Charif, and T. Frank. Towards video understanding of laparoscopic surgery : Instrument tracking. In *Image and Vision Computing New Zealand*, 2005.
- [MCV⁺97] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Trans. Med. Imaging*, 16(2) :187–198, 1997.

- [MD03] A. Menciassi and P. Dario. Bio-inspired solutions for locomotion in the gastrointestinal tract : background and perspectives. *Philosophical Transactions of the Royal Society of London. Series A : Mathematical, Physical and Engineering Sciences*, 361(1811) :2287–2298, 2003.
- [MdROM11] J. Martinez-del Rincon, C. Orrite, and C. Medrano. Rao-blackwellised particle filter for colour-based tracking. *Pattern Recognition Letters*, 32(2) :210 – 220, 2011.
- [MEF11] M. Mühling, R. Ewerth, and B. Freisleben. On the spatial extents of sift descriptors for visual concept detection. In *Computer Vision Systems*, volume 6962 of *Lecture Notes in Computer Science*, pages 71–80. Springer Berlin / Heidelberg, 2011.
- [MFF11] B. Mora, R.A. Fournier, and S. Foucher. Application of evidential reasoning to improve the mapping of regenerating forest stands. *International Journal of Applied Earth Observation and Geoinformation*, 13(3) :458 – 467, 2011.
- [MGdM02] K. Miura, J.A. Gangloff, and M. de Mathelin. Robust and uncalibrated visual servoing without jacobian using a simplex method. In *IEEE International Conference on Intelligent Robots and Systems*, volume 1, pages 311–316, 2002.
- [MHA09] A. Melbourne, D. Hawkes, and D. Atkinson. Image registration using uncertainty coefficients. In *Biomedical Imaging : From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on*, pages 951–954, 2009.
- [Mit07] H.B. Mitchell. *Multi-sensor data fusion : an introduction*. Springer, 2007.
- [MKC10] R. Mebarki, A. Krupa, and F. Chaumette. 2d ultrasound probe complete guidance by visual servoing using image moments. *IEEE Trans. on Robotics*, 26(2) :296–306, April 2010.
- [MKK11] N.A. Mandellos, I. Keramitsoglou, and C.T. Kiranoudis. A background subtraction algorithm for detecting and tracking vehicles. *Expert Systems with Applications*, 38(3) :1619 – 1631, 2011.
- [Mém03] E. Mémin. *Estimation du flot optique : contribution et panorama de différentes approches*. Habilitation à diriger les recherches, Université de Rennes 1, July 2003.
- [MN09] A.C. Muller and S. Narayanan. Cognitively-engineered multisensor image fusion for military applications. *Information Fusion*, 10(2) :137 – 149, 2009.
- [MO06] A. Martin and C. Osswald. A new generalization of the proportional conflict redistribution rule stable in terms of decision. *Advances and applications of DSMT for information fusion : collected works*, 2 :69, 2006.
- [MOK⁺05] D.P. Mandic, D. Obradovic, A. Kuh, T. Adali, U. Trutschell, M. Golz, P. De Wilde, J. Barria, A. Constantinides, and J. Chambers. Data fusion for modern engineering applications : An overview. *Artificial Neural Networks : Formal Models and Their Applications-ICANN 2005*, pages 715–721, 2005.
- [MPB⁺04] B. Maurin, O. Piccin, B. Bayle, J. Gangloff, M. De Mathelin, L. Soler, and A. Gangi. A new robotic system for ct-guided percutaneous procedures with haptic feedback. In *International Congress Series*, volume 1268, pages 515–520. Elsevier, 2004.
- [MPR⁺11] O. Marin, I. Pérez, D. Ruiz, A. Soriano, and J. Garcia. Neural networks versus genetic algorithms as medical classifiers. In *Foundations on Natural and Artificial Computation*, volume 6686, pages 393–400. Springer Berlin / Heidelberg, 2011.

- [MR93] H. Michel and P. Rives. Singularités dans la détermination de l'attitude de l'effecteur d'un robot à partir de la vue perspective de 3 points. Technical Report 1850, INRIA, 1993.
- [MRJ01] C. Muller, M. Rombaut, and M. Janier. Dempster shafer approach for high level data fusion applied to the assessment of myocardial viability. In *Functional Imaging and Modeling of the Heart*, volume 2230, pages 104–112. Springer Berlin / Heidelberg, 2001.
- [MS01] K. Mikolajczyk and C. Schmid. Indexing based on scale invariant interest points. *ICCV*, pages 525–531, 2001.
- [MS⁺04] K. Mikolajczyk, C. Schmid, et al. Comparison of affine-invariant local detectors and descriptors. In *Proc. European Signal Processing Conf*, 2004.
- [MS05] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. *IEEE transactions on pattern analysis and machine intelligence*, pages 1615–1630, 2005.
- [MSJM10] J. Mennesson, C. Saint-Jean, and L. Mascarilla. De nouveaux descripteurs de fourier géométriques pour l'analyse d'images couleur. 2010.
- [MVMMGG⁺02] J.L. Mosso-Vázquez, A. Minor-Martínez, A. Gómez-González, L. Padilla-Sánchez, A. de la Torre-Gutiérrez, and V. Lara-Vaca. Colecistectomías laparoscópicas asistidas por un robot y teleguiadas vía satélite en méxico. *Cirugía y Cirujanos*, 70(6) :449–454, 2002.
- [MZW11] Y. Ma, K. Zhan, and Z. Wang. Image fusion. In *Applications of Pulse-Coupled Neural Networks*, pages 83–109. Springer Berlin Heidelberg, 2011.
- [NBK08] T. Nir, A.M. Bruckstein, and R. Kimmel. Over-parameterized variational optical flow. *International Journal of Computer Vision*, 76(2) :205–216, 2008.
- [NBL⁺00] N. Navab, B. Bascle, M. Loser, B. Geiger, and R. Taylor. Visual servoing for automatic and uncalibrated needle placement for percutaneous procedures. In *Proc. of Int. Conf. On Computer Vision and Pattern Recognition (CVPR'2000)*, pages 2327–2334, 2000.
- [Ngu08] H. Nguyen. On random sets and belief functions. *Classic Works of the Dempster-Shafer Theory of Belief Functions*, pages 105–116, 2008.
- [NIC⁺07] J.A. Newmark, W.C.W. II, C.C. Chang, G.E. Herrera, D.H. Brooks, C.A. DiMarzio, and C.M. Warner. Determination of the number of cells in preimplantation embryos by using noninvasive optical quadrature microscopy in conjunction with differential interference contrast microscopy. *Microscopy and Microanalysis*, 13(02) :118–127, 2007.
- [NITM00] Y. Nakabo, M. Ishikawa, H. Toyoda, and S. Mizuno. 1 ms column parallel vision system and its application of high speed target tracking. In *IEEE International Conference on Robotics and Automation*, volume 1, pages 650–655, 2000.
- [NK04] M. Novotni and R. Klein. Shape retrieval using 3d zernike descriptors. *Computer-Aided Design*, 36(11) :1047–1062, 2004.
- [NK07] Z. Nawrat and P. Kostka. The robin heart vision telemanipulator for camera holding - preliminary test results. *Journal of Automation, Mobile Robotics & Intelligent Systems*, 1 :48–53, 2007.

- [NK11] C. Nadeau and A. Krupa. Intensity-based direct visual servoing of an ultrasound probe. In *IEEE International Conference on Robotics and Automation*, pages 5677–5682, Shanghai, China, May 2011.
- [NKG11] C. Nadeau, A. Krupa, and J. Gangloff. Automatic tracking of an organ section with an ultrasound probe : Compensation of respiratory motion. In *Medical Image Computing and Computer-Assisted Intervention*, Toronto, Canada, September 2011.
- [NL09] L. Nanni and A. Lumini. Fusion of color spaces for ear authentication. *Pattern Recognition*, 42(9) :1906–1913, 2009.
- [Nor88] J. Norton. Limit theorems for dempster’s rule of combination. *Theory and Decision*, 25 :287–313, 1988.
- [NSDH07] P.M. Novotny, J.A. Stoll, P.E. Dupont, and R.D. Howe. Real-time visual servoing of a robot using three-dimensional ultrasound. In *IEEE International Conference on Robotics and Automation*, pages 2655–2660, 2007.
- [NTZH10] N. Nacereddine, S. Tabbone, D. Ziou, and L. Hamami. Un descripteur efficace pour la reconnaissance des symboles graphiques basé sur la transformée de radon. 2010.
- [NvLSZ12] M. Nilsson, J. van Laere, T. Susi, and T. Ziemke. Information fusion in practice : A distributed cognition perspective on the active role of users. *Information Fusion*, 13(1) :60 – 78, 2012.
- [OIA⁺11] I. Ordavo, S. Ihle, V. Arkadiev, O. Scharf, H. Soltau, A. Bjeoumikhov, S. Bjeoumikhova, G. Buzanich, R. Gubzhokov, A. Günther, R. Hartmann, P. Holl, N. Kimmel, M. Kühbacher, M. Lang, N. Langhoff, A. Liebel, M. Radtke, U. Reinholz, H. Riese-meier, G. Schaller, F. Schopper, L. Strüder, C. Thamm, and R. Wedell. A new pnccd-based color x-ray camera for fast spatial and energy-resolved measurements. *Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment*, 654(1) :250 – 257, 2011.
- [Orl78] A. Orlov. Fuzzy and random sets. *Prikladnoi Mnogomerni Statisticheskii Analys*, 262 :280, 1978.
- [OZN⁺08] L. Ott, P. Zanne, F. Nageotte, M. de Mathelin, and J. Gangloff. Physiological motion rejection in flexible endoscopy using visual servoing. In *IEEE International Conference on Robotics and Automation*, pages 2928–2933, 2008.
- [Pal11] E. Elle O.J. Palomar, R. Samset. Computer-assisted fixed ceiling illumination system in the or. *International Journal of Computer Assisted Radiology and Surgery*, pages 153–154, 2011.
- [Par62] E. Parzen. On estimation of a probability density function and mode. *The annals of mathematical statistics*, 33(3) :1065–1076, 1962.
- [Par94] B. Parhami. Voting algorithms. *Reliability, IEEE Transactions on*, 43(4) :617–629, 1994.
- [Pau88] L.F. Pau. Sensor data fusion. *Journal of Intelligent & Robotic Systems*, 1(2) :103–116, 1988.
- [PBB⁺06] N. Papenberg, A. Bruhn, T. Brox, S. Didas, and J. Weickert. Highly accurate optic flow computation with theoretically justified warping. *International Journal of Computer Vision*, 67(2) :141–158, 2006.

- [PBP⁺05] L. Perlick, H. Bächis, C. Perlick, C. Lüring, M. Tingart, and J. Grifka. Revision total knee arthroplasty : a comparison of postoperative leg alignment after computer-assisted implantation versus the conventional technique. *Knee Surgery, Sports Traumatology, Arthroscopy*, 13(3) :167–173, 2005.
- [Pei06] W. J. Peine. *Standard and Metrology Needs for Surgical Robotics*. PhD thesis, Purdue University, 2006.
- [PFJ06] H. Permuter, J. Francos, and I. Jermyn. A study of gaussian mixture models of color and texture features for image classification and segmentation. *Pattern Recognition*, 39(4) :695–706, 2006.
- [PG11] A. Paksoy and M. Göktürk. Information fusion with dempster-shafer evidence theory for software defect prediction. *Procedia Computer Science*, 3 :600–605, 2011.
- [PI97] M. Peura and J. Iivarinen. Efficiency of simple shape descriptors. *Advances in Visual Form Analysis, World Scientific, Singapore*, pages 443–451, 1997.
- [PM] J. Petitt and K. Miller. Mr sam : Magnetic resonance compatible surgical manipulator.
- [PML04] J.A. Piepmeier, G.V. McMurray, and H. Lipkin. Uncalibrated dynamic visual servoing. *IEEE International Conference on Robotics and Automation*, 20(1) :143–147, 2004.
- [PMV03] J.P.W. Pluim, J.B.A. Maintz, and M.A. Viergever. Mutual-information-based registration of medical images : a survey. *Medical Imaging, IEEE Transactions on*, 22(8) :986–1004, 2003.
- [PP05] P. Pognet and G. Poisson. Session robotique médicale la conception de robots pour l’assistance aux gestes médicaux et chirurgicaux. 2005.
- [PP07] H. Park and K.R. Park. Iris recognition based on score level fusion by using svm. *Pattern Recognition Letters*, 28(15) :2019–2028, 2007.
- [Pre70] J.M.S. Prewitt. *Object enhancement and extraction*, volume 75. Academic Press, New York, 1970.
- [PVR98] D.T. Pham, J. Verron, and M.C. Roubaud. A singular evolutive extended kalman filter for data assimilation in oceanography. *Journal of Marine Systems*, 16(3-4) :323 – 340, 1998.
- [PZR04] N.A. Patronik, M.A. Zenati, and C.N. Riviere. Crawling on the heart : a mobile robotic device for minimally invasive cardiac interventions. *Medical Image Computing and Computer-Assisted Intervention*, pages 9–16, 2004.
- [QGC⁺11] M. Quan, W. Guijin, S. Chenbo, L. Xinggang, and R. Zhiwei. A new framework for on-line object tracking based on surf. *Pattern Recognition Letters*, 32(13) :1564 – 1571, 2011.
- [QH09] C.C. Qu and J. Hahn. Process monitoring and parameter estimation via unscented kalman filtering. *Journal of Loss Prevention in the Process Industries*, 22(6) :703 – 709, 2009.
- [QMXW10] H. Qian, Y. Mao, W. Xiang, and Z. Wang. Recognition of human activities using svm multi-class classifier. *Pattern Recognition Letters*, 31(2) :100–111, 2010.

- [RBC⁺02] J. Rosen, J.D. Brown, L. Chang, M. Barreca, M. Sinanan, and B. Hannaford. The bluedragon-a system for measuring the kinematics and dynamics of minimally invasive surgical tools in-vivo. In *IEEE International Conference on Robotics and Automation*, volume 2, pages 1876–1881, 2002.
- [RC04] Andrei Romanenko and José A.A.M. Castro. The unscented filter as an alternative to the ekf for nonlinear state estimation : a simulation case study. *Computers & Chemical Engineering*, 28(3) :347 – 355, 2004.
- [Rei11] T. Reineking. Particle filtering in the dempster-shafer theory. *International Journal of Approximate Reasoning*, 52(8) :1124 – 1135, 2011.
- [RGH11] A. Roy, N. Gale, and L. Hong. Automated traffic surveillance using fusion of doppler radar and video information. *Mathematical and Computer Modelling*, 54(1-2) :531 – 543, 2011.
- [RGT⁺11] A. Riwan, B. Giudicelli, F. Taha, J.Y. Lazennec, A. Sabhani, P. Kilian, Z. Jabbour, J. VanRhijn, F. Louveau, G. Morel, et al. Projet surgicobot : robot d’assistance la chirurgie du rachis. *IRBM*, 2011.
- [Rob63] L.G. Roberts. Machine perception of three-dimensional solids. Technical report, DTIC Document, 1963.
- [Rob77] G.S. Robinson. Edge detection by compass gradient masks. *Computer Graphics and Image Processing*, 6(5) :492–501, 1977.
- [Ros03] P.L. Rosin. Measuring shape : ellipticity, rectangularity, and triangularity. *Machine Vision and Applications*, 14(3) :172–184, 2003.
- [Rou02] J. Roubert. *Automatic guidance of a laparoscope using computer vision*. PhD thesis, Lund Institute of Technology, 2002.
- [RPR05] E. Ramasso, D. Pellerin, and M. Rombaut. Reconnaissance de sauts d’athlète dans les vidéos : fusion du mouvement de caméra et de la trajectoire de points caractéristiques. 2005.
- [RSMR10] M.R.N. Rao, G.R. Sridhar, K. Madhu, and A.A. Rao. A clinical decision support system using multi-layer perceptron neural network to predict quality of life in diabetes. *Diabetes and Metabolic Syndrome : Clinical Research and Reviews*, 4(1) :57 – 59, 2010.
- [RSSS11] K. Rathnayaka, T. Sahama, M. A. Schuetz, and B. Schmutz. Effects of ct image segmentation methods on the accuracy of long bone 3d reconstructions. *Medical Engineering & Physics*, 33(2) :226 – 233, 2011.
- [RVS11] K.R. Radhika, M.K. Venkatesha, and G.N. Sekhar. An approach for on-line signature authentication using zernike moments. *Pattern Recognition Letters*, 32(5) :749 – 760, 2011.
- [RW10] R.M. Rangayyan and Y. Wu. Screening of knee-joint vibroarthrographic signals using probability density functions estimated with parzen windows. *Biomedical Signal Processing and Control*, 5(1) :53 – 58, 2010.
- [SAP10] B. Soheilian, A. Arlicot, and N. Paparoditis. Extraction de panneaux de signalisation routière dans les images couleurs. *Actes du 17ème congrès francophone en Reconnaissance de Formes et Intelligence Artificielle*, 10 :743–750, 2010.

- [Sau06] M. Sauvée. *Contribution à l'aide aux gestes pour la chirurgie cardiaque à coeur battant. Guidage échographique par asservissement prédictif non linéaire*. PhD thesis, Thèse de l'Université de Montpellier 2, 2006.
- [SB12] E. Simon and L. Bertino. Gaussian anamorphosis extension of the denkf for combined state parameter estimation : Application to a 1d ocean ecosystem model. *Journal of Marine Systems*, 89(1) :1 – 18, 2012.
- [SBL01] E. Shahbazian, D.E. Blodgett, and P. Labbé. The extended ooda model for data fusion systems. In *Proceedings of the International Conference on Information Fusion*, pp. 19-25, 2001.
- [SBW99] A.N. Steinberg, C.L. Bowman, and F.E. White. Revisions to the jdl data fusion model. Technical report, DTIC Document, 1999.
- [SBZ⁺03] M. Shoham, M. Burman, E. Zehavi, L. Joskowicz, E. Batkilin, and Y. Kunicher. Bone-mounted miniature robot for surgical procedures : Concept and clinical applications. *IEEE International Conference on Robotics and Automation*, 19(5) :893–901, 2003.
- [Sca98] B. Scassellati. Eye finding via face detection for a foveated active vision system. In *Proceedings of the National Conference on Artificial Intelligence*, pages 969–976, 1998.
- [SDT09] Y. Shi, I. Dinov, and A. Toga. Cortical shape analysis in the laplace-beltrami feature space. In *Medical Image Computing and Computer-Assisted Intervention*, volume 5762, pages 208–215. Springer Berlin / Heidelberg, 2009.
- [SE11] J.M. Sáez and F. Escolano. 6dof entropy minimization slam for stereo-based wearable devices. *Computer Vision and Image Understanding*, 115(2) :270–285, 2011.
- [SI73] Y. Shirai and H. Inoue. Guiding a robot by visual feedback in assembling tasks. *Pattern Recognition*, 5(2) :99–106, 1973.
- [Sim57] H.A. Simon. *Models of Man : Social and Rational. Mathematical Essays on Rational Human Behavior in a Social Setting*. 1957.
- [SK78] P.V. Sankar and E.V. Krishnamurthy. On the compactness of subsets of digital pictures. *Computer Graphics and Image Processing*, 8(1) :136–143, 1978.
- [Sma05] J. Smarandache, F. Dezert. Proportional conflict redistribution rules for information fusion. *JAIIF Journal*, March 2005.
- [SOA⁺10] A. Schlaefel, C. Otte, R. Ansari, G. Hüttmann, L. Richter, R. Bruder, M. Heinig, M. Sommerauer, and G. Kovacs. Towards high resolution image guided navigation for prostate brachytherapy. In *International Congress and Exhibition, Computer Assisted Radiology and Surgery*, page S24, 2010.
- [Sob78] I. Sobel. Neighborhood coding of binary images for fast contour following and general binary array processing. *Computer Graphics and Image Processing*, 8(1) :127 – 135, 1978.
- [SP09] F. Saint-Pierre. Chirurgie prothétique du genou assistée par ordinateur : apport de la navigation. Technical report, Haute Autorité de Santé, 2009.
- [Spa07] A. Spadaro. Lighting device provided with automatic aiming system, 2007.

- [SPDC06] M. Sauvée, P. Poignet, E. Dombre, and E. Courtial. Image based visual servoing through nonlinear model predictive control. In *IEEE Conference on Decision and Control*, pages 1776–1781, 2006.
- [SRLA11] M Sadeghi, M. Razmara, T.L Lee, and M.S. Atkins. A novel method for detection of pigment network in dermoscopic images using graphs. *Computerized Medical Imaging and Graphics*, 35(2) :137 – 143, 2011.
- [SRMD11] P. Shah, B. Reddy, S. Merchant, and U. Desai. Context enhancement to reveal a camouflaged target and to assist target localization by fusion of multispectral surveillance videos. *Signal, Image and Video Processing*, pages 1–16, 2011.
- [SSS+08] S. Speidel, G. Sudra, J. Senemaud, M. Drentschew, B.P. Müller-Stich, C. Gutt, and R. Dillmann. Recognition of risk situations based on endoscopic instrument tracking and knowledge based situation modeling. In *Proceedings of SPIE*, volume 6918, page 69180X, 2008.
- [SSY+07] J.K. Seon, E.K. Song, T.R. Yoon, S.J. Park, B.H. Bae, and S.G. Cho. Comparison of functional results with navigation-assisted minimally invasive and conventional techniques in bilateral total knee arthroplasty. *Computer Aided Surgery*, 12(3) :189–193, 2007.
- [ST01] O. Schneider and J. Troccaz. A six-degree-of-freedom passive arm with dynamic constraints (padyc) for cardiac surgery application : Preliminary experiments. *Computer aided surgery*, 6(6) :340–351, 2001.
- [Str06] E. Straszecka. Combining uncertainty and imprecision in models of medical diagnosis. *Information Sciences*, 176(20) :3026 – 3059, 2006.
- [Stu97] C. Studholme. *Measures of 3D medical image alignment*. University of London, 1997.
- [SVF11] L. Snidaro, I. Visentini, and G. Foresti. Data fusion in modern surveillance. In *Innovations in Defence Support Systems - 3*, volume 336, pages 1–21. Springer Berlin / Heidelberg, 2011.
- [SW80] A.C. Sanderson and L.E. Weiss. Image-based visual servo control using relational graph error signals. *Proceedings of the IEEE*, 68 :1074–1077, 1980.
- [SW86] M. Subbarao and A.M. Waxman. Closed form solutions to image flow equations for planar surfaces in motion. *Computer Vision, Graphics, and Image Processing*, 36(2-3) :208–228, 1986.
- [SW10] H. Schneider and J. Wahrburg. Simulation model for the dynamics analysis of a surgical assistance robot. *Intech, Vukovar, Croatia*, 2010.
- [TC03] O. Tahri and F. Chaumette. Application of moment invariants to visual servoing. In *IEEE International Conference on Robotics and Automation*, volume 3, pages 4276–4281, 2003.
- [TC05] O. Tahri and F. Chaumette. Point-based and region-based image moments for visual servoing of planar objects. *Robotics, IEEE Transactions on*, 21(6) :1116–1127, 2005.
- [TC11] L. Tzu-Chao. Decision-based fuzzy image restoration for noise reduction based on evidence theory. *Expert Systems with Applications*, 38(7) :8303 – 8310, 2011.

- [TCM⁺11] K. Toyoda, J. Chung, M. Murata, S. Ieiri, and M. Tomikawa. Implementation of a mr compatible laparoscope holding robot for laparoscopic hepatectomy using intraoperative mri. In *International Congress and Exhibition, Computer Assisted Radiology and Surgery*, pages S226–S295, 2011.
- [Tea80] M.R. Teague. Image analysis via the general theory of moments. *JOSA*, 70(8) :920–930, 1980.
- [TGC⁺08] S. Tamaz, R. Gourdeau, A. Chanu, J.B. Mathieu, and S. Martel. Real-time mri-based control of a ferromagnetic core for endovascular navigation. *Biomedical Engineering*, 55(7) :1854–1863, 2008.
- [Tho90] S.C.A. Thomopoulos. Sensor integration and data fusion. *Journal of Robotic Systems*, 7(3) :337–372, 1990.
- [TM01] J. Triesch and C. Malsburg. Democratic integration : Self-organized integration of adaptive cues. *Neural computation*, 13(9) :2049–2074, 2001.
- [TSS⁺08] L. Tesar, A. Shimizu, D. Smutek, H. Kobatake, and S. Nawano. Medical image analysis of 3d ct images based on extension of haralick texture features. *Computerized Medical Imaging and Graphics*, 32(6) :513 – 520, 2008.
- [TWT⁺96] R.C. Thompson, B.J. Wilson, M.L. Tobin, A.S. Hill, and S.J. Hawkins. Biologically generated habitat provision and diversity of rocky shore organisms at a hierarchy of spatial scales. *Journal of Experimental Marine Biology and Ecology*, 202(1) :73 – 84, 1996.
- [Vac11] J. Amat J. Casals A. Vaca, R. Aranda. Safe human-robot co-manipulation in assisted robotic surgery. *15th Annual Conference of the International Society for Computer Aided Surgery*, 6 :281, 2011.
- [Val01] L. Valet. *Un système flou de fusion coopérative : application au traitement d’images naturelles*. PhD thesis, 2001.
- [Vap95] V. Vapnik. The nature of statistical learning theory. *Springer*, 1995.
- [VBA11] T.D. Vu, J. Burlet, and O. Aycard. Grid-based localization and local mapping with moving object detection and tracking. *Information Fusion*, 12(1) :58–69, 2011.
- [VdM73] C. Von der Malsburg. Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik*, 14(2) :85–100, 1973.
- [Vin01] M. Vincze. Robust tracking of ellipses at frame rate. *Pattern Recognition*, 34(2) :487–498, 2001.
- [Vir06] J. Virtanen. Enhancing the compatibility of surgical robots with magnetic resonance imaging. Technical report, Faculty of technology, department of mechanical engineering, university of oulu, 2006.
- [VLC06] S. Voros, J.A. Long, and P. Cinquin. Automatic localization of laparoscopic instruments for the visual servoing of an endoscopic camera holder. In *Medical Image Computing and Computer-Assisted Intervention*, volume 4190, pages 535–542. Springer Berlin / Heidelberg, 2006.
- [VMGBS11] M.C. Valiente, R. Machin, E. Garcia-Barriocanal, and M.A. Sicilia. An ontology-based integrated approach to situation awareness for high-level information fusion in c4isr.

- In *Advanced Information Systems Engineering Workshops*, volume 83, pages 513–527. Springer Berlin Heidelberg, 2011.
- [VMO05] M.A. Vitrani, G. Morel, and T. Ortmaier. Automatic guidance of a surgical instrument with ultrasound based visual servoing. In *IEEE International Conference on Robotics and Automation*, pages 508–513, 2005.
- [VNM53] J. Von Neumann and O. Morgenstern. *Theories of games and economic behavior*. Princeton University Press, 1953.
- [Vor06] S. Voros. *Vers la commande en boucle fermée d’un robot porte-optique : apport de l’analyse contextuelle d’images endoscopiques*. PhD thesis, Université Paris VI, 2006.
- [VR04] B.R. Venkatesh and K.R. Ramakrishnan. Recognition of human actions using motion history information extracted from the compressed video. *Image and vision computing*, 22(8) :597–607, 2004.
- [Wal99] L. Wald. Some terms of reference in data fusion. *Geoscience and Remote Sensing, IEEE Transactions on*, 37(3) :1190–1193, 1999.
- [WC04] J Wu and A. Chung. Multimodal brain image registration based on wavelet transform using sad and mi. In *Medical Imaging and Augmented Reality*, volume 3150, pages 270–277. Springer Berlin / Heidelberg, 2004.
- [WFG⁺03] R. Westphal, J. Faulstich, T. Gösling, S. Winkelbach, T. Hüfner, C. Krettek, and F. Wahl. Fracture reduction using a telemanipulator with haptical feedback. In *17th International Congress and Exhibition, Computer Assisted Radiology and Surgery*, pages 25–28, 2003.
- [WGD09] J. Wang, J. Guo, and L. Ding. An adaptive kalman filtering based state of charge combined estimator for electric vehicle battery pack. *Energy Conversion and Management*, 50(12) :3182 – 3186, 2009.
- [WJLD05] A. Wolf, B. Jaramaz, B. Lisien, and A.M. DiGioia. Mbars : mini bone-attached robotic system for joint arthroplasty. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 1(2) :101–121, 2005.
- [WMT09] B. Wowra, A. Muacevic, and J.C. Tonn. Quality of radiosurgery for single brain metastases with respect to treatment technology : a matched-pair analysis. *Journal of neuro-oncology*, 94(1) :69–77, 2009.
- [WSLV08] M. Wiggins, A. Saad, B. Litt, and G. Vachtsevanos. Evolving a bayesian classifier for ecg-based age classification in medical applications. *Applied Soft Computing*, 8(1) :599–608, 2008.
- [WUW98] Y.F. Wang, D.R. Uecker, and Y. Wang. A new framework for vision-enabled and robotically assisted minimally invasive surgery. *Computerized Medical Imaging and Graphics*, 22(6) :429–437, 1998.
- [WW10] D. Wied and R. Weißbach. Consistency of the kernel density estimator-a survey. *Statistical Papers, to appear DOI*, 10, 2010.
- [WWL10] J. Walters-Williams and Y. Li. Comparative study of distance functions for nearest neighbors. In Khaled Elleithy, editor, *Advanced Techniques in Computing Sciences and Software Engineering*, pages 79–84. Springer Netherlands, 2010.

- [XJ07] N. Xiaolin and F. Jiancheng. An autonomous celestial navigation method for leo satellite based on unscented kalman filter and information fusion. *Aerospace Science and Technology*, 11(2-3) :222 – 228, 2007.
- [XLL⁺11] J. Xi, P. Lamata, J. Lee, P. Moireau, D. Chapelle, and N. Smith. Myocardial transversely isotropic material parameter estimation from in-silico measurements based on a reduced-order unscented kalman filter. *Journal of the Mechanical Behavior of Biomedical Materials*, 4(7) :1090 – 1102, 2011.
- [XSK⁺10] W. Xi, J. Seidel, J.W. Kakareka, T.J. Pohida, D.E. Milenic, J. Proffitt, S. Majewski, A.G. Weisenberger, M.V. Green, and P.L. Choyke. Monica : a compact, portable dual gamma camera system for mouse whole-body imaging. *Nuclear Medicine and Biology*, 37(3) :245 – 253, 2010.
- [XZ10] W. Xin and T. Zhenmin. Modified particle filter-based infrared pedestrian tracking. *Infrared Physics & Technology*, 53(4) :280 – 287, 2010.
- [Yag87] R.R. Yager. On the dempster-shafer framework and new combination rules. *Information sciences*, 41(2) :93–137, 1987.
- [Yam11] T. Yamamoto. Non-regionwise weight cancellation for monte carlo higher order criticality calculations using kernel density estimator. *Annals of Nuclear Energy*, 38(11) :2515 – 2520, 2011.
- [YD10] P.L. Yen and B.L. Davies. Active constraint control for image-guided robotic surgery. *Proceedings of the Institution of Mechanical Engineers, Part H : Journal of Engineering in Medicine*, 224(5) :623, 2010.
- [YHH⁺11] J. Yang, H.Z. Huang, L.P. He, S.P. Zhu, and D. Wen. Risk evaluation in failure mode and effects analysis of aircraft turbine rotor blades using dempster-shafer evidence theory under uncertainty. *Engineering Failure Analysis*, 2011.
- [YK06] B.S. Yang and K.J. Kim. Application of dempster-shafer theory in fault diagnosis of induction motors using vibration and current signals. *Mechanical Systems and Signal Processing*, 20(2) :403–420, 2006.
- [YLA02] P.L.H. Yu, K.F. Lam, and M. Alvo. Nonparametric rank tests for independence in opinion surveys. *Australian Journal of Statistics*, 31 :279–290, 2002.
- [YLZ10] J. Yang, C. Liu, and L. Zhang. Color space normalization : Enhancing the discriminating power of color spaces for face recognition. *Pattern Recognition*, 43(4) :1454–1466, 2010.
- [YNGR07] R.B. Yadav, N.K. Nishchal, A.K. Gupta, and V.K. Rastogi. Retrieval and classification of shape-based objects using fourier, generic fourier, and wavelet-fourier descriptors technique : A comparative study. *Optics and Lasers in engineering*, 45(6) :695–708, 2007.
- [YSO⁺10] U. Yetkin, A. Sahin, B. Ozcem, M. Bademci, I. Yurekli, and A. Gurbuz. Our surgical approach to the high risk left main coronary artery stenosis in a case who received irradiation due to malignancies of two separate systems. *The Internet Journal of Oncology*, 7(1), 2010.

- [YTS11] G.K. Yadav, K. Tarun, and N. Smita. Real time vehicle detection for rear and forward collision warning systems. In *Advances in Computing and Communications*, volume 193, pages 368–377. Springer Berlin Heidelberg, 2011.
- [YWR11] D. Yong, J. Wen, and S. Rehan. Modeling contaminant intrusion in water distribution networks : A new similarity-based dst method. *Expert Systems with Applications*, 38(1) :571 – 578, 2011.
- [YXY10] M. Yu, G. Xiaodong, and W. Yuanyuan. Histogram similarity measure using variable bin size distance. *Computer Vision and Image Understanding*, 114(8) :981 – 989, 2010.
- [YYY93] K. Yamazawa, Y. Yagi, and M. Yachida. Omnidirectional imaging with hyperboloidal projection. In *IEEE International Conference on Intelligent Robots and Systems*, volume 2, pages 1029–1034. IEEE, 1993.
- [Zad65] L.A. Zadeh. Fuzzy sets. *Information and control*, 8(3) :338–353, 1965.
- [Zad99] L. A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 100, Supplement 1(0) :9 – 34, 1999.
- [ZBA09] X. Zabulis, H. Baltzakis, and A. Argyros. Vision-based hand gesture recognition for human-computer interaction. *The Universal Access Handbook, Human Factors and Ergonomics. Lawrence Erlbaum Associates, Inc.(LEA)*, 2009.
- [ZGY⁺06] H.X. Zhou, Y.H. Guo, X.F. Yu, S.Y. Bao, J.L. Liu, Y. Zhang, Y.G. Ren, and Q. Zheng. Clinical characteristics of remote zeus robot-assisted laparoscopic cholecystectomy : a report of 40 cases. *World Journal of Gastroenterology*, 12(16) :2606–2609, 2006.
- [ZJGHHFL11] H. Zhou, J. J. Gomez-Hernandez, H.J. Hendricks-Franssen, and L. Li. An approach to handling non-gaussianity of parameters and state variables in ensemble kalman filtering. *Advances in Water Resources*, 34(7) :844 – 864, 2011.
- [ZL02] D. Zhang and G. Lu. Shape-based image retrieval using generic fourier descriptor. *Signal Processing : Image Communication*, 17(10) :825–848, 2002.
- [ZMA⁺11] E. Zervas, A. Mpimpoudis, C. Anagnostopoulos, O. Sekkas, and S. Hadjiefthymiades. Multisensor data fusion for fire detection. *Information Fusion*, 12(3) :150 – 159, 2011.
- [ZRL⁺11] N. Zhang, S. Ruan, S. Lebonvallet, Q. Liao, and Y. Zhu. Kernel feature selection to fuse multi-spectral mri images for brain tumor segmentation. *Computer Vision and Image Understanding*, 115(2) :256 – 269, 2011.
- [ZS03] J. Zhong and S. Sclaroff. Segmenting foreground objects from a dynamic textured background via a robust kalman filter. 2003.
- [ZvdH06] Z. Zivkovic and F. van der Heijden. Efficient adaptive density estimation per image pixel for the task of background subtraction. *Pattern recognition letters*, 27(7) :773–780, 2006.

Aurélien GAUVIN

Asservissement visuel d'un éclairage opératoire

Les travaux présentés dans ce manuscrit traitent de l'asservissement visuel d'un éclairage opératoire. Il s'agit d'une thèse CIFRE soutenue par l'entreprise MAQUET SAS et en collaboration avec le Laboratoire PRISME de l'Université d'Orléans.

Les éclairages opératoires offrent à l'équipe chirurgicale des conditions d'éclairage suffisantes pour leur permettre d'accomplir leurs gestes. Leur positionnement est difficile en cours d'intervention et engendre fréquemment des heurts entre les membres de l'équipe. Les solutions déjà développées pour résoudre ce problème ne donnent pas satisfaction en raison de l'interaction forte entre le système et l'équipe chirurgicale. Nous proposons dans cette étude un éclairage opératoire asservi visuellement ne nécessitant pas d'information explicite, et opérationnel quel que soit le type de chirurgie. Il s'agit d'un système "intelligent" autrement dit capable de désigner de lui-même la zone à éclairer et "autonome", c'est-à-dire apte à se déplacer seul une fois les coordonnées de la cible connues. Ces deux points constituent la problématique de cette étude.

Pour rendre "intelligent" l'éclairage opératoire, nous proposons un processus de désignation basé sur la reconnaissance d'objets spécifiques : le sang, la peau, les champs stériles et les instruments. Pour ce faire nous utilisons une fusion des attributs forme, couleur et mouvement basée sur le cadre crédibiliste. Nous résolvons les problèmes d'inhomogénéité de l'image dus à la puissance de l'éclairage par l'ajout d'une étape de fusion intermédiaire. Une fois l'ensemble des objets reconnus, nous procédons à la désignation de la zone à éclairer à l'aide de la théorie de la décision.

L'autonomie du système est quant à elle assurée par une boucle d'asservissement visuel 2D, qui permet de faire converger l'éclairage vers la zone précédemment désignée. Nous avons réalisé au cours de cette étude un prototype qui a permis de valider l'approche dans des conditions réelles.

Mots clés : asservissement visuel, fusion de données, chirurgie.

Visual servoing of a surgical light

The work presented in this manuscript is related to the visual servoing of a surgical light. This is a collaborative study between MAQUET SAS and PRISM laboratory (University of Orléans).

Surgical light provides the surgical team enough lighting to perform their activities. The positioning of this equipment during the operation is arduous and leads frequently to disagreement between members of the team. Solutions already developed to solve this problem do not offer satisfaction in reason of the high interaction between the system and the team. This work aims to propose a surgical light visually served which do not require explicit information and which is operational whatever the kind of surgery. It consists in an "intelligent" system able to designate by itself the region where the surgeon is working on, and which is also "autonomous", that is to say the system can move to this target. These two points correspond to the problematic of this study.

To make the operating light "intelligent" we propose an architecture based on the recognition of specific objects : blood, skin, steriled field, instruments. To achieve this we fuse shape, color and movement attributes. We solve the inhomogeneity problem of images due to high illumination thanks to an intermediate fusion step. Once all the objects are recognized we carry out the designation thanks to the decision theory.

The "autonomous" part of the system consists in a 2D visual servoing loop that makes possible the convergence of the surgical light to the region of interest. A prototype has been realized during this work, which enabled us to validate the proposed approach in a real environment.

Keywords : visual servoing, data fusion, surgery.



Laboratoire PRISME
8 rue Léonard de Vinci
45072 Orléans cedex 2

MAQUET
GETINGE GROUP