



HAL
open science

Contribution à l'étude du recalage de données 3D /couleur

Lounis Douadi

► **To cite this version:**

Lounis Douadi. Contribution à l'étude du recalage de données 3D /couleur. Automatique / Robotique. Université Montpellier II - Sciences et Techniques du Languedoc, 2006. Français. NNT: . tel-00141930

HAL Id: tel-00141930

<https://theses.hal.science/tel-00141930>

Submitted on 16 Apr 2007

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.

ACADEMIE DE MONTPELLIER
UNIVERSITE MONTPELLIER II

-Sciences et Techniques du Languedoc -

T H E S E

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE MONTPELLIER II

Discipline : **Génie Informatique, Automatique et Traitement du Signal**

Formation Doctorale : **Systèmes Automatiques et Microélectroniques**

Ecole Doctorale : **Informatique, Structures, Systèmes.**

présentée et soutenue publiquement

par

Lounis DOUADI

le 04 Octobre 2006.

**Contribution à l'étude du recalage
de données 3D /couleur**

JURY

M. André CROSNIER

M. El Mustapha MOUADDIB

M. Simon LACROIX

Mme Marie-José ALDON

Président

Rapporteur

Rapporteur

Directeur de Thèse

A ma famille...A mes parents.

REMERCIEMENTS

Les travaux présentés dans ce document ont été effectués au sein du département Robotique du LIRMM. Je tiens à remercier Messieurs Michel Habib et Michel Robert ancien et actuel directeurs du LIRMM ainsi que Messieurs Etienne Dombre et René Zapata ancien et actuel responsables du département Robotique pour m'avoir accueilli au sein de leur établissement pendant cette thèse.

Je remercie Madame Marie-José Aldon, mon directeur, pour avoir dirigé cette thèse. Je la remercie pour la qualité de son encadrement, ses qualités professionnelles et surtout pour ses qualités humaines. Son aide précieuse a permis de mener à bien ces travaux. Je lui suis, donc, naturellement, très reconnaissant.

Je veux exprimer ma gratitude à Messieurs Olivier Strauss, Sébastien Druon et André Crosnier pour leur aide précieuse, ainsi qu'à toutes les personnes qui ont contribué de près ou de loin à ce travail : Abdellah El-jalaoui, Philippe Amat, Frédéric Comby, William Puech, Jean Triboulet, Etienne Dombre...

Enfin, je tiens à remercier mes collègues et amis : Jean-Mathias Spiewak (Batias), Olivier Parodi (Padawan), Philippe Amat, Yousra Benzaïda, Mohammed Samer, Benoît Telle, Geovany Borges, Meziane Bennour, mes deux amis "Abdellah" El-jalaoui, Nizar Benregba, Riad Kaced, Hassan El-makssoud, Mohammed El-ghazouani, et tous les autres collègues et amis...(plus de places pour les citer tous).

TABLE DES MATIERES

1	INTRODUCTION	15
1.1	INTRODUCTION	17
1.2	PROBLEMATIQUE	17
1.3	APPLICATIONS DU RECALAGE	18
1.4	CONTEXTE DE L'ETUDE	19
1.5	PRESENTATION DU MANUSCRIT	20
2	RECALAGE DE DONNEES 3D	23
2.1	INTRODUCTION	25
2.2	PROBLEMATIQUE DU RECALAGE	25
2.2.1	<i>Nécessité et objectifs du recalage</i>	25
2.2.2	<i>Formalisation du problème de recalage géométrique</i>	26
2.2.3	<i>Représentation de la transformation rigide</i>	28
2.2.4	<i>Recalage local et recalage global</i>	29
2.3	METHODES ITERATIVES	30
2.3.1	<i>Sélection des points</i>	31
2.3.2	<i>Technique d'appariement</i>	32
2.3.3	<i>Pondération des paires de points appariés</i>	34
2.3.4	<i>Rejet des mauvais appariements</i>	34
2.3.5	<i>Critère à minimiser et algorithme de minimisation</i>	35
2.4	MISE EN CORRESPONDANCE D'ELEMENTS GEOMETRIQUES CARACTERISTIQUES	36
2.5	METHODES PAR OPTIMISATION	37
2.6	CONCLUSION	38
3	RECALAGE DE DEUX NUAGES DE POINTS 3D/COULEUR PAR ICP	39
3.1	INTRODUCTION	41
3.2	RECALAGE DE DONNEES 3D AVEC L'ALGORITHME ICP	41
3.2.1	<i>Principe de l'algorithme</i>	41
3.2.2	<i>Calcul de la transformation de repère</i>	43
3.2.3	<i>Elimination des points aberrants par seuillage adaptatif</i>	44
3.2.4	<i>Mise en œuvre d'un k-D tree pour la recherche du plus proche voisin</i>	45
3.3	PRISE EN COMPTE DE LA COULEUR DANS LE RECALAGE	46
3.3.1	<i>Appariement avec un critère mixte</i>	46
3.3.2	<i>Appariement avec seuillage de la distance couleur</i>	46
3.3.3	<i>Choix de l'espace couleur</i>	47
3.4	ETUDE COMPARATIVE POUR UN DEPLACEMENT SIMULE	48
3.4.1	<i>Données expérimentales</i>	48
3.4.2	<i>Influence du seuillage adaptatif pour le recalage 3D</i>	49
3.4.3	<i>Apport de la couleur dans les performances du recalage</i>	50
3.4.3.1	<i>Appariement avec un critère mixte</i>	50
3.4.3.2	<i>Appariement avec seuillage de la distance couleur</i>	51
3.4.4	<i>Influence du recouvrement entre les deux images</i>	52
3.5	ETUDE COMPARATIVE POUR UN DEPLACEMENT REEL	53
3.5.1	<i>Convergence des algorithmes</i>	53
3.5.2	<i>Influence de l'espace couleur</i>	56
3.6	CONCLUSION	58
4	APPARIEMENT DE POINTS COULEUR CARACTERISTIQUES	59
4.1	INTRODUCTION	61
4.2	ETAT DE L'ART	62
4.2.1	<i>Détection de points d'intérêt</i>	62

4.2.1.1	Méthodes basées sur les contours.....	63
4.2.1.2	Méthodes basées sur le signal.....	63
4.2.1.2.1	Détecteur de Kitchen Rosenfeld.....	66
4.2.1.2.2	Détecteur Harris Stephens.....	67
4.2.1.2.3	Détecteur Harris Précis.....	67
4.2.1.2.4	Détecteur Harris Précis Couleur.....	67
4.2.1.2.5	Comparaison des différents détecteurs.....	68
4.2.2	<i>Appariement d'images 2D</i>	69
4.2.2.1	Méthodes basées sur le signal.....	69
4.2.2.1.1	Corrélation du signal.....	70
4.2.2.1.2	Corrélation de phase.....	70
4.2.2.1.3	Mesure de l'information mutuelle.....	71
4.2.2.1.4	Distance de Hausdorff.....	72
4.2.2.2	Méthodes basées sur les histogrammes.....	73
4.2.2.3	Méthodes basées sur les invariants différentiels.....	74
4.2.3	<i>Elimination des faux appariements</i>	74
4.2.4	<i>Conclusion</i>	76
4.3	EXTRACTION DE POINTS D'INTERET.....	77
4.3.1	<i>Exemple 1: peinture à l'huile</i>	79
4.3.2	<i>Exemple 2: Figurine 3D</i>	80
4.3.3	<i>Exemple 3: boîte cylindrique</i>	81
4.4	CARACTERISATION ET APPARIEMENT DES POINTS D'INTERET.....	84
4.4.1	<i>Vecteur d'invariants</i>	84
4.4.2	<i>Normalisation des couleurs</i>	85
4.4.3	<i>Appariement des points d'intérêt</i>	87
4.5	VALIDATION EXPERIMENTALE DE LA METHODE D'APPARIEMENT.....	88
4.5.1	<i>Influence de la normalisation des couleurs</i>	89
4.5.1.1	Appariement sans normalisation.....	90
4.5.1.2	Appariement avec normalisation.....	91
4.5.2	<i>Influence de la taille de la fenêtre normalisation des couleurs</i>	91
4.5.2.1	Exemple 1: Figurine 3D.....	92
4.5.2.2	Exemple 2 : Peinture à l'huile.....	94
4.5.2.3	Exemple 3: Boite cylindrique.....	97
4.6	CONCLUSION.....	98
5	RECALAGE AUTOMATIQUE 3D/COULEUR.....	101
5.1	INTRODUCTION.....	103
5.2	ETAT DE L'ART.....	104
5.3	ALGORITHME RANSAC.....	105
5.3.1	<i>Principe</i>	106
5.3.2	<i>Application au recalage de nuages de points 3D/couleur</i>	108
5.3.3	<i>Mise en œuvre de l'algorithme RANSAC</i>	109
5.4	VALIDATION EXPERIMENTALE AVEC DES DONNEES SIMULEES.....	112
5.4.1	<i>Cas de deux images sans bruit additionnel</i>	112
5.4.2	<i>Cas de deux images avec bruit additionnel</i>	115
5.5	VALIDATION DU RECALAGE COMPLET AVEC DES DONNEES REELLES.....	117
5.6	CONCLUSION.....	122
6	CONCLUSION GENERALE.....	125
6.1	BILAN DES TRAVAUX REALISES.....	127
6.1.1	<i>Recalage itératif utilisant l'information couleur</i>	127
6.1.2	<i>Automatisation du processus de recalage</i>	127
6.2	PERSPECTIVES.....	128
6.2.1	<i>Amélioration du recalage de deux nuages 3D/Couleur</i>	129
6.2.2	<i>Recalage de N vues 3D/Couleur</i>	129
6.2.3	<i>Intégration des données pour la construction d'un modèle surfacique 3D/C</i>	130

LISTE DES FIGURES

FIGURE 2.1 - DEUX NUAGES 3D/COULEUR, (A) AVANT ET (B) APRES LE RECALAGE.....	28
FIGURE 3.1 - EXEMPLE DE K-D TREE (2-D TREE)	45
FIGURE 3.2 - REPRESENTATION DE L'ESPACE COULEUR HSV.	47
FIGURE 3.3 - LES DEUX IMAGES A RECALER IMAGE1 ET IMAGE2	49
FIGURE 3.4 - EVOLUTION DE L'ERREUR RESIDUELLE DE DISTANCE POUR ICP ET ICP_GSA	50
FIGURE 3.5 - EVOLUTION DE L'ERREUR RESIDUELLE DE DISTANCE POUR ICP_GSA ET ICP_MSA	51
FIGURE 3.6 - EVOLUTION DE L'ERREUR RESIDUELLE DE DISTANCE POUR ICP_GSA, ICP_MSA ICP_GSC... 51	
FIGURE 3.7 - EVOLUTION DE L'ERREUR RESIDUELLE DE DISTANCE EN FONCTION DU RECOUVREMENT DES DEUX IMAGES.....	53
FIGURE 3.8 - LES DEUX IMAGES REELLES RECALEES PAR LA TRANSFORMATION INITIALE.....	54
FIGURE 3.9 - ERREURS RESIDUELLES POUR ICP_GSA, ICP_MSA ET ICP_GSC.....	54
FIGURE 3.10 - LES DEUX IMAGES REELLES RECALEES PAR ICP_GSC	55
FIGURE 3.11 - ERREURS RESIDUELLES POUR ICP_GSA, ICP_MSA ET ICP_GSC	55
FIGURE 3.12 - LES DEUX PARTIES DES IMAGES RECALEES PAR LA TRANSFORMATION INITIALE TI.....	56
FIGURE 3.13 - EFFET DE L'ESPACE COULEUR.....	58
FIGURE 4.1 - POINTS D'INTERET DETECTES PAR LE DETECTEUR DE HARRIS PRECIS COULEUR SUR L'IMAGE CALVET1.	78
FIGURE 4.2 - POINTS D'INTERET DETECTES PAR LE DETECTEUR DE HARRIS PRECIS COULEUR SUR L'IMAGE CALVET2.	79
FIGURE 4.3 - POINTS D'INTERET DETECTES PAR LE DETECTEUR DE HARRIS PRECIS COULEUR SUR L'IMAGE OURSON1.	80
FIGURE 4.4 - POINTS D'INTERET DETECTES PAR LE DETECTEUR DE HARRIS PRECIS COULEUR SUR L'IMAGE OURSON2.	81
FIGURE 4.5 - IMAGE IVOIRE1	82
FIGURE 4.6 - IMAGE IVOIRE2	82
FIGURE 4.7 - POINTS D'INTERET DETECTES PAR LE DETECTEUR DE HARRIS PRECIS COULEUR SUR UNE PARTIE DE L'IMAGE IVOIRE1.	83
FIGURE 4.8 - POINTS D'INTERET DETECTES PAR LE DETECTEUR DE HARRIS PRECIS COULEUR SUR UNE PARTIE DE L'IMAGE IVOIRE2.	83
FIGURE 4.9 - EXEMPLE DE NORMALISATION APPLIQUEE A L'IMAGE OURSON1 AVEC UNE FENETRE CIRCULAIRE DE TAILLE RAYON 11 PIXELS.	86
FIGURE 4.10 - EXEMPLE DE NORMALISATION APPLIQUEE A L'IMAGE IVOIRE1 AVEC UNE FENETRE CIRCULAIRE DE RAYON DE TAILLE 9 PIXELS.	87

FIGURE 4.11 - MISE EN CORRESPONDANCE DE DEUX IMAGES IVOIRE AVEC SANS UN CHANGEMENT LINEAIRE DE LA LUMINOSITE (: LES BONS APPARIEMENTS (TOUS) SONT REPRESENTES PAR DES PAVES BLANCS), LES MAUVAIS PAR DES CROIX.	89
FIGURE 4.12 - MISE EN CORRESPONDANCE DES LA MEMES IMAGES IVOIRE AVEC UN CHANGEMENT LINEAIRE DE LA LUMINOSITE (: LES BONS APPARIEMENTS SONT REPRESENTES PAR DES PAVES BLANCS, LES MAUVAIS PAR DES CROIX VERTES).	90
FIGURE 4.13 - MISE EN CORRESPONDANCE DE DEUX IMAGES IVOIRE AVEC UN CHANGEMENT LINEAIRE DE LA LUMINOSITE ET NORMALISATION DES COULEURS (: LES BONS APPARIEMENTS (TOUS) SONT REPRESENTES PAR DES PAVES BLANCS), LES MAUVAIS PAR DES CROIX VERTES.	91
FIGURE 4.14 - MISE EN CORRESPONDANCE DES DEUX IMAGES OURS1 ET OURS2 SANS NORMALISATION DE LA COULEUR: (LES BONS APPARIEMENTS SONT REPRESENTES PAR DES PAVES BLANCS, LES MAUVAIS PAR DES CROIX VERTES. LE TAUX DE BONS APPARIEMENTS EST DE 43%).....	92
FIGURE 4.15 - MISE EN CORRESPONDANCE DES DEUX IMAGES OURS1 ET OURS2 AVEC NORMALISATION DE LA COULEUR (: LES BONS APPARIEMENTS SONT REPRESENTES PAR DES CROIX BLANCHES, LES MAUVAIS PAR DES CROIX NOIRES. LE TAUX DE BONS APPARIEMENTS EST DE 56%).	93
FIGURE 4.16 - MISE EN CORRESPONDANCE DES DEUX IMAGES CALVET1 ET CALVET2 SANS NORMALISATION DE LA COULEUR (: LES BONS APPARIEMENTS CORRECTS SONT REPRESENTES PAR DES PAVES BLANCS, LES MAUVAIS PAR DES CROIX VERTES. LE TAUX DE BONS APPARIEMENTS EST DE 43%).	95
FIGURE 4.17 - MISE EN CORRESPONDANCE DES DEUX IMAGES CALVET1 ET CALVET2 AVEC NORMALISATION DE LA COULEUR: (LES BONS APPARIEMENTS SONT REPRESENTES PAR DES PAVES BLANCS, LES MAUVAIS PAR DES CROIX VERTES. LE TAUX DE BONS APPARIEMENTS EST DE 65%).	96
FIGURE 4.18 - MISE EN CORRESPONDANCE DES DEUX IMAGES CALVET1 ET CALVET2 IVOIRE AVEC NORMALISATION DE LA COULEUR DIAMETRE 11 (: LES BONS APPARIEMENTS SONT REPRESENTES PAR DES CROIX BLANCHES, LES MAUVAIS PAR DES CROIX NOIRES. LE TAUX DE BONS APPARIEMENTS EST DE 12%).	97
FIGURE 4.19 - MISE EN CORRESPONDANCE DES DEUX IMAGES CALVET1 ET CALVET2 IVOIRE AVEC NORMALISATION DE LA COULEUR DIAMETRE 25 (: LES BONS APPARIEMENTS SONT REPRESENTES PAR DES CROIX BLANCHES, LES MAUVAIS PAR DES CROIX NOIRES. LE TAUX DE BONS APPARIEMENTS EST DE 18%).	98
FIGURE 5.1 - PRINCIPE DU RECALAGE AUTOMATIQUE.....	110
FIGURE 5.2 - ORGANIGRAMME DE L'ALGORITHME RANSAC.	111
FIGURE 5.3 - LES DEUX VUES "IVOIRE1" (FIGURE A) ET "IVOIRE2" (FIGURE B) AVEC DEPLACEMENT SIMULE. CHAQUE NUAGE CONTIENT PLUS DE 27000 POINTS 3D/COULEUR.	112
FIGURE 5.5 - RECALAGE 3D AUTOMATIQUE DES DEUX VUES IVOIRE PAR MISE EN CORRESPONDANCE DE POINTS D'INTERET COULEUR.	114
FIGURE 5.6 - DEUX VUES 3DC DE "WALLIS" AVEC DEPLACEMENT SIMULE.	115
FIGURE 5.7 - RESULTAT DE L'APPARIEMENT DES IMAGES COULEUR CORRESPONDANT AUX NUAGES DE LA FIGURE 5.6.	116
FIGURE 5.8 - RESULTAT DU RECALAGE DES IMAGES DE LA FIGURE 5.7 AVEC LA TRANSFORMATION INITIALE ESTIMEE PAR RANSAC.	117

FIGURE 5.9 - LES DEUX NUAGES OURS1 (11650 POINTS) ET OURS2 (10789 POINTS) A RECALER.	118
FIGURE 5.10 - RESULTAT DE LA MISE EN CORRESPONDANCE 2D DES DEUX IMAGES OURS1 ET OURS2 (LES BONS APPARIEMENTS SONT MARQUE AVEC DES CROIX (+) BLANCHES, LES FAUX APPARIEMENTS SONT MARQUES AVEC DES CROIX (X) NOIRES).	119
FIGURE 5.11 - RECALAGE GROSSIER DES NUAGES 3D/COULEUR OURS1 ET OURS2 PAR LA TRANSFORMATION ESTIMEE AVEC RANSAC.	120
FIGURE 5.12 - EVOLUTION DE LA MOYENNE DES DISTANCES ENTRE POINTS APPARIES EN FONCTION DES ITERATIONS DE ICPGCTYIQ POUR LE RECALAGE DES NUAGES OURS1 ET OURS2.	121
FIGURE 5.13 - RECALAGE FIN DES NUAGES 3D/COULEUR OURS1 ET OURS2 PAR LA TRANSFORMATION ESTIMEE PAR ICPGCTYIQ.	122

LISTE DES TABLEAUX

TABLEAU 3.1 - <i>ERREURS DE RECALAGE DES ALGORITHMES EVALUES</i>	52
TABLEAU 4.1 - <i>TAUX DES BONS APPARIEMENTS EN FONCTION DU DIAMETRE DE LA FENETRE DE NORMALISATION DE LA COULEUR : CAS DE LA PAIRE OURS1/OURS2</i>	93
TABLEAU 4.2 - <i>TAUX DES BONS APPARIEMENTS EN FONCTION DU DIAMETRE DE LA FENETRE DE NORMALISATION DE LA COULEUR : CAS DE LA PAIRE CALVET1/CALVET2</i>	94
TABLEAU 5.1 - <i>NOMBRE N DE TIRAGE</i>	109
TABLEAU 5.2 - <i>ERREURS ENTRE LA TRANSFORMATION ESTIMEE ET LA TRANSFORMATION EXACTE POUR LE RECALAGE DE LA FIGURE (5.8)</i>	115
TABLEAU 5.3 <i>PARAMETRES DE TRANSLATION ET DE ROTATION DE LA TRANSFORMATION ESTIMEE ENTRE LES VUES OURS1 ET OURS2 PAR RANSAC</i>	120
TABLEAU 5.4 - <i>PARAMETRES DE TRANSLATION ET DE ROTATION DE LA TRANSFORMATION ESTIMEE ENTRE LES VUES OURS1 ET OURS2 PAR ICPGCTYIQ</i>	121

CHAPITRE *1*

1 INTRODUCTION

Sommaire

1.1	INTRODUCTION	17
1.2	PROBLEMATIQUE	17
1.3	APPLICATIONS DU RECALAGE	18
1.4	CONTEXTE DE L'ETUDE	19
1.5	PRESENTATION DU MANUSCRIT	20

1.1 Introduction

Les travaux présentés dans ce mémoire ont pour objectif de proposer des solutions pour la construction de modèles complexes d'objets permettant de représenter et de visualiser non seulement leur forme, mais aussi leur aspect photométrique (couleur, état de surface, texture). Ils concernent plus précisément le problème du recalage de données visuelles 3D/couleur.

1.2 Problématique

La construction de modèles de scènes ou d'objets réels, en trois dimensions, avec ou sans texture est un problème qui concerne de nombreux domaines d'application. Par exemple, ces modèles peuvent être utilisés pour représenter l'environnement dans lequel navigue un robot, pour construire des scènes de réalité virtuelle à partir de l'observation d'objets réels, pour représenter des monuments historiques qui ont été numérisés en vue de leur restauration ou encore pour alimenter une base de données d'objets de musée consultable par un large public.

La construction d'un modèle 3D comporte trois étapes : l'acquisition d'un ensemble d'images 3D sous différents points de vue, leur recalage dans un repère unique, la simplification des données obtenues pour éliminer les redondances résultant du recouvrement des images et le calcul d'une représentation surfacique de l'objet.

Deux grandes catégories de méthodes sont utilisées pour acquérir les données. La première qui est basée sur la vision active en 3D (mesure du temps de vol, vision en lumière structurée, triangulation laser) fournit des images denses de la scène observée (images de distance et parfois images vidéo). La deuxième catégorie qui fait appel à une ou plusieurs caméras vidéo (stéréovision, vision en mouvement) permet d'acquérir la texture de la scène ainsi qu'une structure 3D minimale (par exemple, un ensemble de points d'intérêt). Mais en règle générale, avec ce type de capteur, on ne dispose pas d'une image 3D dense décrivant précisément la géométrie de la scène.

Quel que soit le capteur utilisé, pour numériser complètement un objet ou une scène 3D il faut acquérir, sous des points de vue différents, un ensemble d'images qui se recouvrent partiellement. Le nombre de vues nécessaires dépend de la complexité et de la taille de l'objet, mais également de la résolution du capteur. On obtient donc un ensemble d'images qu'il faut recaler. Ainsi, un scanner 3D fournit des nuages de points non structurés qu'il faut ramener dans un repère unique. Lorsque ces images sont recalées deux par deux on parle de recalage local, lorsqu'elles sont recalées toutes ensemble, il s'agit d'un recalage global.

Le sujet traité dans cette thèse est plus précisément le recalage de deux nuages de points 3D/couleur denses et non structurés. Notre objectif est de proposer des algorithmes :

- permettant de prendre en compte des masses de données importantes (pouvant aller jusqu'à plusieurs millions de points par image),
- capable d'apporter une solution avec des données dans lesquelles l'information géométrique n'est pas suffisante. C'est le cas par exemple pour des objets au relief peu marqué (peinture à l'huile, ..), avec des objets dont la forme présente des symétries qui peuvent conduire à des solutions multiples,
- qui ne nécessitent pas d'intervention de l'opérateur pour gérer le processus de recalage (par exemple la pose de marqueurs artificiels sur l'objet, ou le pointage manuel de repères dans les images numérisées).

Comme nous le verrons dans le paragraphe 1.4, ce travail de recherche a pour application la création de modèles 3D/couleur à haute résolution d'objets de musée (peintures à l'huile, sculptures, figurines archéologiques, ...). Notre objectif est de développer une approche intégrée pour la construction de modèles 3D texturés qui tire parti du caractère complémentaire des données vidéo et 3D fournies par les scanners à lumière structurée. Ce type de capteur présente l'avantage de fournir simultanément ces deux types de données dans le même repère (repère attaché à la caméra).

1.3 Applications du recalage

Les domaines d'application du recalage 3-D sont nombreux. En effet, cette technique de la vision est aussi bien utilisée en robotique (robots mobiles, asservissement visuel), qu'en imagerie médicale 3-D, en imagerie par satellite, en archéologie, ou encore en urbanisme.

La vision artificielle en trois dimensions occupe une place importante en robotique et notamment dans des applications telles que la localisation de véhicules autonomes et l'asservissement visuel. La localisation de robots mobiles par la vision a, dans la plupart des cas, recours à des méthodes de recalage (mise en correspondance de caractéristiques géométriques ou photométriques). L'une des méthodes les plus utilisées pour la localisation de robots mobiles, en milieu intérieur, est nommée SLAM (*Synchronous Localization And Mapping*). Il s'agit de se localiser en mettant en correspondance une carte globale de l'environnement avec la carte locale acquise par le robot à chaque instant. La carte globale est mise à jour à chaque instant. Ces cartes peuvent être constituées de caractéristiques géométriques 2D ou 3D (segments de droites, points, courbes, éléments de surface, etc.).

En imagerie médicale, des techniques telles que la tomographie à rayon X et l'imagerie par résonance magnétique permettent une visualisation en trois dimensions des différentes structures anatomiques. De nombreux travaux ont été menés dans ce domaine. L'intérêt est de faciliter le diagnostic en permettant de visualiser des pathologies telles que les tumeurs et de voir leur évolution dans le temps. Le recalage de données 3D peut aussi avoir un intérêt en imagerie aérienne pour la construction de cartes géographiques 3D pour des applications de géologie ou dans le domaine de la défense. En archéologie, le recalage concerne la modélisation d'objets ou de sites historiques. Des techniques nouvelles d'acquisition utilisant un capteur de profondeur (scanners 3D, télémètre laser) et un capteur photométrique (caméra CCD ou appareil photo numérique) permettent de fournir des données 3D/couleur. On obtient ainsi des

relevés topographiques des sites archéologiques relativement précis. Grâce aux logiciels de visualisation, il est possible de naviguer virtuellement et librement autour et à l'intérieur de ces sites historiques qui peuvent être mis à la disposition du grand public.

1.4 Contexte de l'étude

Cette étude s'insère dans le cadre du projet national ART3D (ACI « Masse de données ») dont l'objectif général est « la numérisation, l'archivage et la restitution des œuvres d'art en volume du patrimoine national ou mondial ». Dans le cadre de ce projet, le LIRMM a en charge deux aspects :

- l'étude d'un capteur multimodal haute résolution pour la numérisation 3D/couleur,
- l'intégration des données photométriques dans le modèle surfacique 3D.

Le premier point inclut la définition du cahier des charges du système de numérisation, le choix d'une technologie répondant à ce cahier des charges, la définition du mode d'utilisation du capteur multimodal (éclairage de l'objet, technique de balayage permettant la numérisation complète de l'ensemble du volume,...), l'analyse des besoins logiciels puis le choix et/ou le développement des logiciels permettant l'acquisition synchronisée des données, leur prétraitement et leur recalage, la réalisation d'un démonstrateur de capteur 3D/couleur à partir de produits commercialisés.

Une étude bibliographique relativement exhaustive a permis de recenser les différentes solutions matérielles envisageables pour la numérisation, et de sélectionner un capteur de distance à projection de franges qui répond au cahier des charges de l'application [BRE1]. La validation expérimentale de la solution choisie a été effectuée en réalisant une campagne de numérisation sur des objets réels permettant de couvrir la plus grande diversité des cas : objets manufacturés, œuvres d'art de dimensions, de formes et d'aspect variables (vases, sculptures, peintures de chevalet, ...).

La deuxième étape concerne la reconstruction automatique de surfaces multimodales. Il s'agit d'introduire des informations couleur denses dans la maquette numérique 3D de façon à pouvoir visualiser non seulement la géométrie de l'objet, mais aussi son aspect photométrique. Pour cela, il faut développer une méthode de recalage d'images 3D /couleur représentant des vues partielles de l'objet à numériser qui ont été obtenues avec différentes positions du capteur. Les problèmes à considérer pour la construction de ces modèles multimodaux à haute résolution sont liés à la grande quantité de données manipulées, aux zones d'ombre qui empêchent une numérisation complète de la surface, aux erreurs qui entachent les données (variation de l'éclairage de l'objet), à la prise en compte d'informations redondantes (recouvrement des différentes vues). Cette deuxième phase du projet doit aboutir à la réalisation d'un logiciel assurant le recalage d'un ensemble de vues 3D/couleur, et la construction d'un modèle surfacique permettant de visualiser rapidement et d'analyser avec précision la forme et la texture d'un objet. C'est dans cette deuxième phase du projet que s'insère notre travail sur le recalage d'images 3D/couleur.

1.5 Présentation du manuscrit

Ce manuscrit est organisé en six chapitres. Le premier chapitre nous a permis d'introduire le sujet de cette thèse ainsi que la problématique et de situer le contexte de notre travail.

Dans le chapitre 2 nous présentons un état de l'art des différentes méthodes de recalage de données 3D. Ces méthodes sont classées en trois catégories : méthodes itératives, méthodes basées sur la mise en correspondance de caractéristiques et méthodes par optimisation. Nous verrons aussi que quelques solutions exploitent la couleur ou le niveau de gris des points 3D pour améliorer leur recalage. Il apparaît que les approches itératives de type ICP sont parmi les plus utilisées pour recalibrer des nuages de points 3D. C'est ce type d'approche que nous avons choisi d'étudier. Cependant, il nécessite une initialisation préalable de la transformation entre les deux images, et ce problème n'est généralement pas résolu de manière automatique.

Nous décrivons dans le chapitre 3 les méthodes de recalage que nous avons développées. Il s'agit de méthodes itératives inspirées de l'algorithme ICP (*Iterative Closest Point*) que nous appliquons au cas du recalage de deux nuages de points 3D/couleur. Les solutions que nous avons implémentées exploitent l'information couleur dans la phase de mise en correspondance afin d'améliorer la robustesse des appariements et la convergence de l'algorithme itératif de recalage. Ces algorithmes sont évalués tout d'abord avec des images 2D/couleur réelles auxquelles nous avons fait subir un déplacement relatif simulé. La deuxième étape de la validation est réalisée avec des images réelles obtenues avec un déplacement inconnu du capteur de numérisation.

L'algorithme ICP nécessite une initialisation préalable de la transformation entre les deux images. Ceci implique une intervention de l'utilisateur. Nous proposons à travers le quatrième et le cinquième chapitre une solution pour réaliser cette initialisation de manière entièrement automatique.

La première étape de notre approche automatique est présentée dans le quatrième chapitre. Celle-ci consiste à exploiter l'information couleur attachée aux points 3-D pour extraire un ensemble d'appariements de points caractéristiques dans la paire d'images 2-D couleur correspondant aux nuages de points à recalibrer. Nous donnons, d'abord, un état de l'art sur les principales méthodes de mise en correspondance d'images 2-D couleur. Nous détaillons, ensuite, l'approche que nous avons choisie basée sur l'appariement de points d'intérêt couleur. Celle-ci nécessite une première étape d'extraction de points caractéristiques couleur suivie d'une étape de mise en correspondance basée sur le calcul de vecteurs d'invariants différentiels couleur. Il résulte de cette mise en correspondance, une liste de paires de points 3D/couleur, qui, compte tenu du bruit sur l'information couleur, comporte de faux appariements qu'il est nécessaire d'éliminer pour rendre le processus d'initialisation plus robuste.

Les paires de points couleur identifiées précédemment ont donc permis d'établir une liste d'appariements 3D plausibles. Un algorithme basé sur le tirage aléatoire de ces paires de points (RANSAC) est mis en œuvre pour éliminer les faux appariements 3D et rechercher, ainsi, le sous-ensemble de paires qui permettent d'atteindre le meilleur consensus pour initialiser la transformation 3D entre les deux images. On obtient ainsi une estimation grossière de la transformation entre les deux nuages à recaler. Cette transformation est ensuite affinée par l'une des variantes de ICP que nous avons présentées dans le troisième chapitre. Cette procédure d'initialisation automatique est expérimentée sur données simulées avec déplacement connu, puis sur données réelles.

Enfin, la conclusion générale et les perspectives de ce travail sont données dans le sixième chapitre du manuscrit.

2 RECALAGE DE DONNEES 3D

Sommaire

2.1	INTRODUCTION	25
2.2	PROBLEMATIQUE DU RECALAGE	25
2.2.1	<i>Nécessité et objectifs du recalage</i>	25
2.2.2	<i>Formalisation du problème de recalage géométrique</i>	26
2.2.3	<i>Représentation de la transformation rigide</i>	28
2.2.4	<i>Recalage local et recalage global</i>	29
2.3	METHODES ITERATIVES.....	30
2.3.1	<i>Sélection des points</i>	31
2.3.2	<i>Technique d'appariement</i>	32
2.3.3	<i>Pondération des paires de points appariés</i>	34
2.3.4	<i>Rejet des mauvais appariements</i>	34
2.3.5	<i>Critère à minimiser et algorithme de minimisation</i>	35
2.4	MISE EN CORRESPONDANCE D'ELEMENTS GEOMETRIQUES CARACTERISTIQUES.....	36
2.5	METHODES PAR OPTIMISATION	37
2.6	CONCLUSION.....	38

2.1 Introduction

Dans les applications de modélisation de l'environnement, on dispose généralement de plusieurs images 3D de la même scène prises de points de vue différents. Le recalage de ces vues peut être défini comme étant le processus d'estimation des transformations rigides (rotations et translations) qui permettent de ramener ces différentes vues dans un référentiel commun. Le recalage de deux vues est souvent considéré comme un problème d'optimisation, dont la fonction de coût est basée sur la mesure d'une distance entre les parties communes de ces deux vues. Les différentes méthodes proposées pour résoudre ce problème diffèrent essentiellement selon la métrique utilisée pour formuler cette distance et selon la technique de minimisation mise en œuvre.

Dans ce chapitre nous dressons un état de l'art sur les méthodes de recalage de données 3D. Ces méthodes sont classées en trois catégories : méthodes itératives, méthodes basées sur la mise en correspondance de caractéristiques et méthodes par optimisation. Cette classification, inspirée par celle proposée par Chen et al [CHE99], n'est pas unique et ne répertorie pas la totalité des méthodes de recalage existantes.

2.2 Problématique du recalage

2.2.1 Nécessité et objectifs du recalage

La reconstruction de modèles 3-D complets (création de modèles d'objets, reconstruction de l'environnement en robotique mobile, imagerie médicale 3-D, etc.) nécessite l'acquisition de plusieurs images de l'objet sous des points de vue différents. Cette contrainte résulte tout d'abord des limitations du champ du capteur utilisé. Ainsi, plus la résolution du capteur est importante, plus son champ est petit. Dans ce cas, il faut acquérir un nombre important d'images pour numériser l'ensemble de l'objet. Par ailleurs, la présence de zones d'ombre liées à la complexité de la forme 3D numérisée se traduit également par un accroissement du nombre de vues. Il en résulte des images 3-D qui doivent être transformées dans le même système de coordonnées. Cette opération est appelée "recalage". Le référentiel commun peut être choisi arbitrairement comme celui de l'une des images. Pour que le recalage soit possible, les points de vue choisis pour la numérisation doivent être tels que deux images voisines se recouvrent partiellement.

Le recalage vise à fournir un ensemble de données exploitables pour la reconstruction du modèle numérique d'un objet ou d'un environnement de telle sorte que l'on puisse, d'une part le visualiser sur un écran, et d'autre part traiter ces données (simplification, échantillonnage systématique ou sélectif,...) pour obtenir une représentation adaptée à l'application envisagée. Mais ce n'est pas le seul domaine

d'application du recalage. Il peut également servir pour la reconnaissance de formes 3D ou pour résoudre des problèmes de localisation en robotique mobile.

Lorsque les transformations de repères entre les différentes images à recaler (qui correspondent aux différents points de vue utilisés pendant l'acquisition des données) sont connues de façon relativement précise, le recalage de ces vues est une tâche en partie résolue par un certain nombre de logiciels. C'est le cas lorsque les différentes positions du capteur ou de l'objet sont mesurables pendant la numérisation (utilisation d'un plateau tournant sur lequel est positionné l'objet, ou bien d'un bras instrumenté pour déplacer le capteur). Il est aussi possible de fixer des marqueurs sur l'objet à numériser ou des balises dans l'environnement à reconstruire (par exemple des boules calibrées dont le centre servira de point de référence). Ces approches ont l'avantage d'être précises, mais elles sont contraignantes du point de vue instrumentation, et elles ne peuvent pas être utilisées dans tous les cas.

Lorsque l'on ne dispose pas de marqueur artificiel ou de système instrumenté pour mesurer le déplacement du capteur, l'initialisation du recalage peut être effectuée soit manuellement par l'opérateur, soit automatiquement par traitement logiciel des images. Dans le premier cas, l'opérateur repère sur une paire d'images de la scène des éléments caractéristiques communs : coins, arêtes, ... En les mettant en correspondance, il est possible d'identifier la transformation rigide permettant de passer d'un repère à l'autre. Dans le deuxième cas, on dispose uniquement des deux images brutes fournies par le capteur de numérisation. Une image est constituée généralement par un nuage de points 3D structuré ou non structuré (suivant la technologie du capteur employé). Avec certains types de capteurs, on a également une information photométrique (niveau de gris ou couleur) pour chaque point. C'est le cas des capteurs à lumière structurée qui associent une caméra à un système de projection de lumière pour estimer la profondeur des points mesurés dans l'image [SAL98, BRE1, BRE2, OSU].

Ce chapitre est consacré à un état de l'art des méthodes développées pour recaler automatiquement deux nuages de points 3D. Nous verrons également comment les informations photométriques sont utilisées lorsqu'elles sont présentes.

2.2.2 Formalisation du problème de recalage géométrique

De manière générale, les méthodes de recalage conçues pour recaler des éléments de formes 3D s'appliquent à divers types de données géométriques. On peut citer de manière non exhaustive [SEE00] :

- des ensembles de points,
- des lignes polygonales,
- des éléments de courbes,
- des surfaces facettées (ensembles de triangles),
- des surfaces paramétriques, ...

Rappelons que dans le cadre de cette thèse nous nous intéressons au recalage des données brutes extraites d'une séquence d'images fournie par un capteur de numérisation. Il s'agit donc d'ensembles de points 3D qui se recouvrent partiellement.

Il faut noter qu'il s'agit d'un problème différent de celui qui consiste à recalcr de nouvelles données dans un modèle existant (par exemple : intégrer de nouveaux points dans une surface connue). Dans ce cas, tous les points du nuage à recalcr ont un correspondant dans le modèle [BES92].

Considérons le cas où les ensembles de données à recalcr sont deux nuages de points P_1 et P_2 qui sont les images 3D d'un même objet ou d'une même scène, acquises par un capteur sous deux points de vue différents et de telle manière que ces images se recouvrent partiellement.

Pour chaque paire de points correspondants p_1 et p_2 , l'objectif du recalage est d'estimer la matrice de rotation R de dimension (3x3) et le vecteur de translation t de dimension (3x1) tels que :

$$p_1 = R p_2 + t \quad (2.1)$$

où p_1 et p_2 sont implicitement les vecteurs de dimension (3x1) qui représentent les points dans les repères associés respectivement aux deux nuages P_1 et P_2 .

Pour simplifier l'écriture, nous écrirons cette relation en utilisant les coordonnées homogènes :

$$p_1 = T p_2 \quad (2.2)$$

où T est la matrice de transformation homogène définie par :

$$T = \begin{bmatrix} & & t_1 \\ & R & t_2 \\ & & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Nous devons donc déterminer les 6 paramètres de la matrice T . Théoriquement, trois paires de points suffisent pour cette estimation. Cependant, compte tenu des bruits qui entachent les données, il est nécessaire d'utiliser un nombre de paires $n > 3$ et de chercher la meilleure estimée au sens des moindres carrés d'un système surdéterminé de n équations. Le critère à minimiser s'écrit :

$$\mathcal{E} = \frac{1}{n} \sum_{i=1}^n \|p_1(i) - T p_2(i)\|^2 \quad (2.3)$$

On voit donc que le recalage 3D de deux nuages de points soulève deux problèmes inhérents à toutes les méthodes que nous allons présenter dans cet état de l'art :

1. la mise en correspondance de points de l'ensemble P_1 et de l'ensemble P_2 ,
2. l'estimation de la transformation rigide T qui minimise l'erreur résiduelle entre les points appariés.

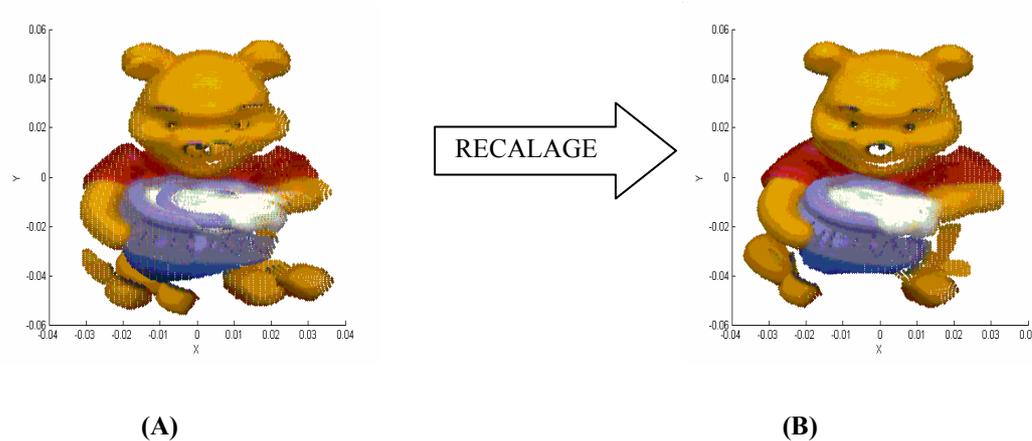


Figure2.1 - Deux nuages 3D/couleur, (a) avant et (b) après le recalage.

2.2.3 Représentation de la transformation rigide

Il existe différentes façons de représenter les changements d'orientation d'un corps dans l'espace. Parmi les plus utilisées en vision 3D et en particulier dans les problèmes de recalage [SEE00], nous pouvons citer:

- les angles de roulis tangage lacet (RTL).
- les angles d'Euler.
- les quaternions et les quaternions duaux.

Nous avons opté pour les représentations utilisant les angles RTL et celle basée sur les quaternions. Et ce dans un souci de vérification de l'exactitude de nos résultats. La représentation RTL a l'intérêt d'être plus simple à manipuler comparé aux quaternions (en commande de robots cette représentation nous permet d'accéder directement aux rotations des articulation, la rotation autour de chaque axe signifie l'envoi direct d'une commande au moteur spécifique) mais a l'inconvénient de présenter plusieurs solutions et des positions singulières lorsque l'on souhaite résoudre le problème inverse (retrouver les valeurs des trois rotations à partir de la matrice R). Les angles de Roulis – Tangage – Lacet expriment le changement d'orientation par trois rotations successives d'un repère autour des trois axes principaux (z, y, x). La matrice de rotation R est donnée par l'expression suivante :

$$R = \begin{bmatrix} C \phi C \theta & C \phi S \theta S \psi - S \phi C \psi & C \phi S \theta C \psi + S \phi S \psi \\ S \phi C \theta & S \phi S \theta S \psi + C \phi C \psi & S \phi S \theta C \psi - C \phi S \psi \\ -S \theta & C \theta S \psi & C \theta C \psi \end{bmatrix} \quad (2.4)$$

Où ψ , θ et ϕ sont, respectivement les angles de rotations autour des axes x, y et z.

Les quaternions (paramètres d'Euler, ou paramètres d'Olindes-Rodrigues) nous permettent d'éviter ces singularités. Les rotations sont exprimées par quatre paramètres qui correspondent à une rotation unique α ($-\pi \leq \alpha \leq +\pi$) autour d'un axe de vecteur unitaire \vec{u} . Ces paramètres sont définis par :

$$\begin{cases} q_1 = C(\alpha/2) \\ q_2 = u_x S(\alpha/2) \\ q_3 = u_y S(\alpha/2) \\ q_4 = u_z S(\alpha/2) \end{cases} \quad (2.5)$$

Les quaternions satisfont l'équation suivante :

$$q_1^2 + q_2^2 + q_3^2 + q_4^2 = 1 \quad (2.6)$$

La matrice de rotation est donnée par :

$$R = \begin{bmatrix} 2(q_1^2 + q_2^2) - 1 & 2(q_2 q_3 - q_1 q_4) & 2(q_2 q_4 + q_1 q_3) \\ 2(q_2 q_3 + q_1 q_4) & 2(q_1^2 + q_3^2) - 1 & 2(q_3 q_4 - q_1 q_2) \\ 2(q_2 q_4 - q_1 q_3) & 2(q_3 q_4 + q_1 q_2) & 2(q_1^2 + q_4^2) - 1 \end{bmatrix} \quad (2.7)$$

2.2.4 Recalage local et recalage global

La construction du modèle 3-D d'un objet nécessite le recalage de plusieurs vues. Pour recaler l'ensemble de ces vues dans un référentiel unique, on peut utiliser soit une approche locale soit une approche globale.

Le recalage local [CHE91, BES92, ZHA94] consiste à recaler les différentes vues deux par deux. Ainsi, pour un modèle composé de N vues, il faut recaler successivement N-1 paires d'images. Cette approche a l'inconvénient de ne pas prendre en compte toutes les interactions entre les divers recouvrements mutuels des différentes surfaces ; en effet chaque vue chevauche généralement plusieurs autres vues. Un autre inconvénient de cette approche est que les erreurs de recalage par paires se propagent de surface en surface [BEN02].

Le recalage global contrairement au recalage local, prend en considération toutes les vues disponibles à la fois [BEN02]. Il prend en compte tous les

recouvrements possibles entre les différentes vues. Il permet une meilleure distribution des erreurs résiduelles entre les différents recouvrements. Les méthodes de recalage global décrites dans la littérature [BER96, MAS95, PUL99] sont pour la plupart des extensions des méthodes de recalage simples présentées dans la suite de ce chapitre.

2.3 Méthodes itératives

Ces méthodes sont les plus utilisées pour résoudre le problème du recalage de données 3-D. Nous pouvons citer l'algorithme ICP (Iterative Closest Point) proposé par Besl et McKay [BES92] qui reste l'une des méthodes les plus utilisées pour recalculer des données 3D sur un modèle 3D. Cet algorithme fonctionne avec des ensembles de points 3D, mais aussi avec des données 3D plus complexes : segments de droites, courbes, surfaces ... Basé sur la recherche du point le plus proche, il calcule, de façon itérative, la matrice de transformation à six degrés de liberté permettant d'effectuer le recalage. Le principe de l'algorithme est d'itérer les deux étapes du recalage : la mise en correspondance des données et l'estimation de la transformation de repères entre les images à recalculer. Au bout de chaque itération l'algorithme fournit une liste de points appariés et une estimation de la transformation de repère entre les images. Cette transformation est utilisée à l'itération suivante pour la mise à jour de la liste des points appariés. Ces derniers serviront, à leur tour, pour calculer une nouvelle estimation de la transformation. Ces étapes sont répétées jusqu'à convergence de l'algorithme. Celui-ci converge lorsque l'erreur résiduelle de distance entre les points appariés est inférieure à un certain seuil. Les auteurs montrent qu'en utilisant un critère de distance de type moindre carré, l'algorithme converge de manière monotone vers le minimum local le plus proche. Pour atteindre le minimum global, il faut initialiser correctement la matrice de transformation. La solution proposée ici consiste à tester un ensemble de transformations initiales dont le nombre dépend de la complexité du modèle 3D et du taux de recouvrement des deux images à recalculer.

Les méthodes de recalage itératives procèdent généralement toutes, à l'instar de l'algorithme ICP, en deux étapes principales : la mise en correspondance et l'estimation de la transformation de repère entre les ensembles de données 3D à recalculer. Cependant, il existe plusieurs variantes de l'algorithme ICP. En effet, l'algorithme original proposé par Besl & McKay présentant des limitations, de nombreuses améliorations y ont été apportées par la suite. Rusinkiewicz et Levoy [RUS01] proposent une étude comparative de plusieurs variantes de ICP, en les appliquant à trois exemples d'images de synthèse qui sont bien représentatives des différents types de formes 3D rencontrées. À partir du bilan de cette étude, ils proposent une version optimisée de ICP dans le but d'améliorer la vitesse de convergence.

L'étude concerne les améliorations apportées par chaque auteur à chaque étape de l'algorithme original. Ces étapes sont :

1. la sélection des points à appairer (échantillonnage des données, utilisation de points de contrôle),
2. la technique d'appariement,

3. la pondération des paires de points obtenues pour estimer la transformation de repère,
4. le rejet des mauvais appariements (points aberrants),
5. le critère à minimiser et l'estimateur utilisé pour le calcul de la transformation de repère recherchée.

De très nombreux travaux ont été menés pour l'amélioration de l'algorithme de base au niveau de chaque point cité ci dessus. Un des objectifs les plus importants à atteindre est l'amélioration de la robustesse de l'algorithme vis à vis des mauvais appariements (erreurs de mesure générées par une imperfection du capteur utilisé) en utilisant des méthodes d'appariement robustes aux points aberrants, notamment par le biais d'un seuillage de distance dynamique [ZHA94] ou statique. D'autres travaux ont été effectués dans le but d'améliorer la vitesse de convergence de l'algorithme [RUS01]. En effet, le recalage 3D concerne souvent des masses de données 3D (ou dans notre cas 3D/couleur) qui occupent un espace mémoire important et qui ralentissent considérablement l'algorithme ICP classique pendant la phase de recherche du plus proche voisin. D'où le recours à des méthodes d'optimisation de cette recherche telles que l'utilisation des arbres k-D (k-D tree) [ZHA94] ou des Octree-spline [LAV95].

2.3.1 Sélection des points

Cette sélection concerne les points de chaque image à recalcr. La façon dont les points en entrée de l'algorithme sont sélectionnés peut avoir un impact sur la vitesse de convergence de celui ci. Est-il préférable de traiter la totalité des données disponibles, faut-il les échantillonner ou faut-il sélectionner seulement des données pertinentes ?

Lorsque les données à recalcr sont des nuages de points, il existe plusieurs approches pour sélectionner ces données. Masuda et al [MAS95] utilisent l'algorithme ICP avec une technique de tirage aléatoire des points pour améliorer la robustesse de l'algorithme aux faux appariements. Dans ce cas, l'échantillonnage ne concerne qu'un seul ensemble de points. Les auteurs effectuent plusieurs tirages aléatoires et calculent, pour chaque échantillon, la transformation de repère correspondant au recalage de l'échantillon avec le deuxième ensemble de données. Le nombre et la taille des échantillons sont déterminés selon une approche probabiliste, de manière à assurer l'obtention d'au moins un sous-ensemble de paires de points ne contenant pas de mauvais appariement. Il en résulte un nombre de transformations estimées égal au nombre d'échantillons prélevés. On ne garde que la transformation qui génère l'erreur quadratique de distance la moins importante entre les deux images à recalcr.

Seung-Hwan Kim et al [KIM04] proposent une version de ICP utilisant un sous échantillonnage des points suivant la direction radiale en partant du centre de gravité du nuage de points à échantillonner. Les auteurs montrent que leur algorithme est aussi

efficace que l'algorithme classique (sans échantillonnage), et a l'avantage d'être considérablement plus rapide.

Gelfand et al [GEL03] proposent une méthode originale pour la sélection des points à utiliser dans ICP lors du recalage de deux ensembles de points 3D. Cette méthode se base sur la notion de "stabilité géométrique" du recalage. Cette stabilité concerne le glissement des surfaces à recaler l'une par rapport à l'autre. Les auteurs sont capables d'estimer les transformations de repère susceptibles de causer un glissement important entre les deux surfaces et donc une instabilité, et choisissent les points des ces surfaces pour lesquels ce glissement est minimum.

Rusinkiewitz et Levoy [RUS01] montrent, à travers plusieurs approches étudiées, qu'il est plus intéressant au sens de la robustesse de sélectionner les points selon l'orientation de leurs normales, que d'effectuer un tirage aléatoire sur l'ensemble des données. Ils proposent de trier d'abord ces points et de les regrouper dans des ensembles différents, en fonction de l'orientation de leurs normales. Ensuite, un échantillonnage uniforme est effectué sur chaque ensemble.

2.3.2 Technique d'appariement

La convergence de ICP dépend beaucoup de la qualité des appariements utilisés. En effet, la présence de faux appariements, dans le meilleur des cas, ralentit la convergence de l'algorithme et peut, au pire des cas, causer sa divergence. Une méthode de mise en correspondance robuste est donc indispensable. La distance géométrique euclidienne "classique" n'est, peut être pas suffisante pour établir des appariements relativement corrects et nécessaires à la convergence de ICP. Ainsi, d'autres critères qu'une simple distance euclidienne (ou en plus de celle ci) peuvent être utilisés. Les données photométriques constituent une information supplémentaire intéressante pour améliorer la qualité des appariements. Il est possible de les intégrer en utilisant une distance mixte comportant un terme de distance couleur et un terme de distance euclidienne, au lieu de la distance classique. Il est aussi possible d'utiliser la couleur en introduisant la notion des plus proches points compatibles. Lorsque l'on ne dispose pas de la couleur, il est possible d'utiliser des caractéristiques géométriques telles que les normales des surfaces, les courbes, les moments, etc.

Chen et Medioni [CHE91] utilisent, pour la mise en correspondance d'une paire de nuage de points 3D, une distance entre un point d'une surface et un plan tangent dans l'autre. Ce plan est perpendiculaire à la normale associée au point du second ensemble. Il s'agit d'une distance point-plan, alors que la distance euclidienne utilisée habituellement entre un point de la première surface et un point de la deuxième est une distance point-point.

Pulli [PUL99] utilise pour le recalage de deux surfaces la distance point-plan. L'auteur conclut que l'utilisation de la distance point-plan améliore le recalage et la vitesse de convergence de ICP. Dans le cas d'une distance point-point, le processus du recalage (par moindres carrés) peut être assimilé au fait d'attacher chaque point de la première surface à son correspondant dans la deuxième surface par un ressort idéal. Ces ressorts permettent de recaler, au mieux, les deux ensembles de points. Quant au cas de la distance point-plan, il peut être défini par le fait d'attacher une extrémité d'un ressort à chaque point du premier nuage, et laisser glisser la deuxième extrémité le long du plan tangent dans la seconde surface [PUL99].

A.N Stein [STE02] utilise pour la modélisation 3D d'objets une variante de ICP basée, elle aussi, sur le calcul des normales des points pour l'appariement. L'auteur calcule les produits scalaires entre les normales des points dans les surfaces à recalcr. Ce produit est maximal lorsque les normales des deux points à apparier ont la même orientation.

G.C Sharp *et al* [SHA02] proposent L'ICPIF (*ICP using Invariant Features*) une nouvelle variante de ICP utilisant pour la phase de mise en correspondance des caractéristiques géométriques de la surface de l'objet calculées au voisinage des points à apparier. Celles ci doivent rester invariantes au mouvement du capteur. Les auteurs utilisent comme caractéristiques géométriques, des courbures, des moments invariants ou des harmoniques sphériques. Cet algorithme apporte une amélioration dans la phase d'appariement de ICP. Au lieu d'utiliser une simple distance euclidienne, G.C Sharp *et al* introduisent un nouveau terme de distance entre caractéristiques géométriques, pour effectuer l'appariement. Celle ci est pondérée afin de contrôler la contribution de ces invariants lors de l'appariement. Cette pondération traduit la confiance qu'on accorde aux caractéristiques géométriques vis à vis des données 3D. Ce poids peut être constant, ou adaptatif (variant à chaque itération de l'algorithme, en fonction de l'erreur moyenne quadratique entre les points appariés). En évaluant leur algorithme ICPIF sur données synthétiques et réelles, les auteurs montrent que celui ci converge mieux et nécessite moins d'itérations que ICP. D'autre part, cet algorithme ne nécessite aucune connaissance de la transformation de repère initiale. Pour certains cas particuliers, il est même possible d'utiliser uniquement une distance basée sur les invariants géométriques pour la mise en correspondance. Ceci dit, dans la plus part des cas, il apparaît qu'il est préférable d'utiliser les deux termes de distance (distance euclidienne et distance basée sur les caractéristiques géométriques).

Lorsque l'information couleur est disponible, elle peut être utilisée pour améliorer l'appariement dans ICP. C'est le cas en particulier lorsque les données géométriques ne sont pas assez significatives (cas de surfaces avec un relief très plat). Les données photométriques peuvent constituer une information supplémentaire intéressante pour améliorer la qualité des appariements. Il est possible de les intégrer, pendant la phase d'appariement, en utilisant une distance mixte comportant un terme de distance couleur et un terme de distance géométrique, au lieu de la distance géométrique classique. L'algorithme CICP (*color ICP*) proposé par Johnson *et al* [JOH96] concerne le recalage de données 3D/couleur pour la construction de modèles d'objets. Il utilise une distance mixte dans la phase d'appariement de ICP. Comme pour ICPIF [SHA02], un coefficient *couleur* est utilisé pour moduler la contribution des données photométriques dans la phase d'appariement. Les auteurs montrent l'intérêt de ces données en appliquant CICP à des données synthétiques. Leurs résultats montrent une amélioration du recalage comparé à ICP classique.

Il est aussi possible d'utiliser la couleur en introduisant la notion des plus proches *points compatibles* [JOS02]. Dans ce cas, les paires de points sont classées une première fois selon leur distance couleur qui ne doit pas dépasser un certain seuil, ensuite parmi ces points qu'on appelle *compatibles* on calcule les points les plus proches en utilisant une distance géométrique.

2.3.3 Pondération des paires de points appariés

La pondération des couples de points appariés a pour objectif de renforcer, dans la phase d'estimation de la transformation géométrique, l'apport des appariements supposés être corrects et d'atténuer l'effet des faux appariements. On peut considérer deux types de pondération possibles [JOS02].

Une pondération binaire où le poids affecté vaut un lorsque l'appariement est considéré correct sinon ce poids prend la valeur nulle. Ceci revient à rejeter les appariements considérés comme faux. Le second type de pondération revient à ne pas considérer les appariements comme étant soit exclusivement "correct" ou exclusivement "faux", mais à considérer aussi les couples de points dont la qualité de l'appariement se situe entre ces deux catégories.

Luck et al [LUC00] utilisent une version de l'algorithme ICP avec une pondération des paires de points appariés inversement proportionnelle à la médiane du carré de la distance entre ces points. Cette pondération agit en amplifiant l'erreur de distance des *mauvais* appariements. Les *bons* appariements ne sont pas affectés par cette pondération, leur erreur n'est pas amplifiée.

Rusinkiewitz & Levoy [RUS01] concluent, sur la base de leur étude comparant différentes fonctions de pondération non binaire (poids constants, pondération selon la distance entre les points à apparier, pondération basée sur les normales des points), que celle ci n'a pas d'impact sur la vitesse de convergence de ICP.

Ainsi, comme nous allons le voir dans le paragraphe suivant, la plupart des variantes de ICP utilisent une pondération binaire, c'est à dire un simple rejet des mauvais appariements, plus efficace qu'une pondération générale.

2.3.4 Rejet des mauvais appariements

Il est très important de rejeter les appariements supposés faux. Ceux ci induisent en erreur les estimateurs par moindres carrés. La difficulté réside dans la définition même d'un "mauvais appariement". Sur quel critère se baser pour décider qu'un appariement est faux ? La façon la plus basique d'éliminer ces points, est d'utiliser un seuil fixe de distance géométrique ou mixte (comportant un terme photométrique). Ce seuil serait difficile à fixer dans le cas de ICP, étant donné que d'une itération à l'autre la distance moyenne entre les points appariés est sensée décroître.

Il est possible d'utiliser un seuil adaptatif statistique comme dans [ZHA94]. Le seuil proposé par Zhang s'adapte en fonction des caractéristiques statistiques (moyenne, écart-type et médiane) des distances entre les points appariés. Dans [ZHA94] l'auteur utilise une variante de ICP avec seuillage adaptatif statistique pour l'élimination des points aberrants. Ce seuil s'avère très efficace et son utilisation

améliore les performances de ICP vis à vis de la vitesse de convergence. Ce seuil sera utilisé dans les méthodes que nous proposons.

Pulli [PUL99] applique plusieurs contraintes pour éliminer les mauvais appariements. La première impose que les points mis en correspondance soient compatibles. Cette compatibilité est liée aux normales de ces points (l'écart d'orientation des normales de deux points appariés ne doit pas dépasser 45 degrés). En plus de cette contrainte, les points se trouvant sur les bords des deux surfaces à recalcr ne sont pas pris compte (ils ne sont pas appariés, ces points ne sont pas fiables car ils sont souvent entachés d'erreurs de mesure). Les deux dernières contraintes consistent à appliquer deux seuils différents. Le premier (dynamique) a pour principe d'éliminer un certain pourcentage de points appariés ayant passé les deux premiers tests. Le second seuil est ensuite appliqué pour éliminer les paires de points dont les distances sont relativement grandes. Ce seuil est de l'ordre de la résolution du capteur 3D.

Masuda et al. [MAS95] classent les points à apparier en quatre catégories: occultés, non appariés, aberrants et corrects. Les points occultés sont des points qui ne sont pas visibles dans une des images, car ils sont cachés par d'autres parties de l'objet (ils sont dans une zone d'ombre du capteur). Les points d'une image non appariés sont des points qui ne peuvent avoir de correspondants dans l'autre. Les points aberrants concernent les paires de points appariés dont la distance dépasse un certain seuil. Les points corrects sont donc les points qui n'appartiennent à aucune des trois catégories citées ci dessus. Le rejet des points aberrants permet de se libérer de la condition qui exige, dans l'algorithme de Besl et McKay [BES92] que le second ensemble de données (points de la deuxième image) soit un sous-ensemble du premier. Le recalage est amélioré du fait que les mauvais appariements ne contribuent plus dans l'estimation du déplacement.

Rusinkewitz et Al [RUS01] comparent plusieurs approches pour l'élimination des faux appariements, et concluent que rejeter ces appariements améliore la qualité du recalage et la stabilité de ICP, par contre ceci n'a pas d'impact sur la vitesse de l'algorithme.

2.3.5 Critère à minimiser et algorithme de minimisation

Le calcul de la meilleure transformation revient à minimiser un critère de distance. Le critère le plus simple est la somme des erreurs de distance quadratiques entre les paires de points appariés (équation 2.3). Cette erreur est une fonction de la matrice de rotation R et du vecteur de translation t (qui constituent la matrice de transformation). Dans ce cas il existe plusieurs approches pour l'estimation de la transformation rigide telles que la décomposition en valeurs singulières (SVD) [ARU87], l'utilisation des quaternions unitaires [BES92], ou des quaternions duaux [WAL91].

Luck et al [LUC00] proposent une méthode hybride de recalage basée sur ICP et le recuit simulé. Ce dernier est utilisé, ici, pour éviter le problème des minimums locaux dont souffre ICP. Le recuit simulé fourni à ICP une "bonne" transformation initiale, lorsque celui ci est pris dans un minimum local. Chen et al [CHE92] utilisent un critère de distance différent de celui utilisé par Besl et al [BES92]. Celui ci se base

sur la somme des carrés des distances entre les points d'une image et le plan (la surface) contenant les points de l'autre [CHE92]. Ce critère s'avère plus intéressant que la distance euclidienne classique entre deux points, mais il impose un calcul des normales à l'élément de surface qui entoure chaque point considéré.

2.4 Mise en correspondance d'éléments géométriques caractéristiques

Cette catégorie de méthodes de recalage est basée sur l'extraction de points d'intérêt ou d'éléments géométriques remarquables dans la surface ou le modèle à recalier. Pour ces points ou ces éléments 3D on calcule ensuite un ensemble de caractéristiques géométriques qui permettront de les identifier et de les comparer. Nous avons vu que de telles caractéristiques 3D pouvaient être utilisées dans certaines variantes d'ICP et qu'elles étaient alors calculées pour tous les points examinés pendant la phase de mise en correspondance [SHA02], [GOD01]. Ici, ces caractéristiques sont calculées pour des éléments particuliers de l'image 3D (points d'intérêt). La mise en correspondance est donc basée sur une mesure de ressemblance entre ces caractéristiques. Pour que la mise en correspondance soit robuste, celles-ci doivent rester invariantes au déplacement du capteur pendant l'acquisition des différentes vues à recalier. On parle alors d'"invariants" géométriques ([SEE02], [SIS00]).

Stein et Medioni [STE92] utilisent des cartes locales nommées *splash*. Elles représentent la distribution des normales des surfaces (à appairer) le long d'un cercle géodésique.

Chua et Jarvis [CHU96] proposent une méthode de mise en correspondance utilisant des invariants géométriques tels que les repères de Darboux et les courbures principales. Ils supposent que l'angle entre deux points de vues, correspondant aux deux positions du capteur, ne dépasse pas un certain seuil (angle relativement petit). Les auteurs sélectionnent un ensemble de points (*seed*), non redondants et supposés appartenir à la zone de recouvrement des deux vues. Pour cela ils se basent sur une méthode appelée analyse de l'hémisphère de vue (*view-emysphere analysis*). Le premier point *seed* d'une surface étant sélectionné, les auteurs calculent son correspondant dans la seconde surface en se basant sur leurs caractéristiques géométriques. Un ensemble de transformations approximatives sont calculées en utilisant les repères de Darboux estimés en ces deux points. Ces transformations sont utilisées, ainsi que la technique de l'analyse de l'hémisphère de vue pour trouver deux autres points *seed* dans la zone de recouvrement. Les repères de Darboux et les courbures principales de ces trois points d'un premier ensemble sont des invariants locaux, et sont utilisés pour trouver des points candidats à l'appariement dans le second ensemble. Pour chaque candidat, une transformation rigide est calculée. Toutes les transformations possibles sont finalement évaluées par un processus de transformation/alignement. La convergence de cet algorithme dépend beaucoup de la précision avec laquelle sont calculés les repères de Darboux.

Thirion [THI96] et al proposent un nouveau type de caractéristiques géométriques nommées *the extremal points of 3-D surface*, basées sur des invariants géométriques. Leurs positions relatives sont invariantes à la transformation rigide. Les auteurs utilisent donc, des invariants géométriques tels que les courbures principales, les distances et les orientations des paires ou triplets des *extremal points*, pour calculer un ensemble de transformations rigides possibles. La meilleure transformation est, alors, déterminée par une méthode itérative.

Bendels [BEN04] et al utilisent, pour un recalage multi-vues, la mise en correspondance de nouvelles caractéristiques géométriques appelées "caractéristiques de surface" (feature surface, ou éléments caractéristiques de surface) contenant l'information d'échelle (scale data). Les auteurs proposent une technique de recalage multi-vues entièrement automatique. Aucune connaissance a priori sur la transformation de repère n'est nécessaire. Cette initialisation est calculée dans une première étape de recalage grossier de *points caractéristiques*. Une étape utilisant un recalage fin est réalisée sur des caractéristiques de surfaces basée sur une variante de ICP.

Les méthodes basées sur la mise en correspondance d'éléments géométriques remarquables ont l'avantage de ne nécessiter aucune connaissance a priori d'une approximation de la transformation à estimer. Mais elles présentent deux inconvénients majeurs:

- elles ne permettent pas de recaler des images 3D qui ne contiennent pas d'éléments remarquables,
- le calcul des vecteurs d'invariants 3D pénalise souvent le temps d'exécution.

2.5 Méthodes par optimisation

Les méthodes par optimisation fournissent un recalage (ou une estimation de la transformation de repère) "optimal" en minimisant une fonction d'erreur par des techniques de minimisation telles que les moindres carrés. Nous pouvons citer, comme méthode classique d'optimisation, l'algorithme de Levenberg-Marquardt (LM), utilisée par Fitzgibbon dans [FIT01].

Blais et Levine [BLA95] utilisent une méthode par optimisation pour le recalage de deux nuages de points 3D. La fonction d'erreur utilisée est la distance euclidienne entre les points dits *de contrôle* d'une image avec ceux de la deuxième. Les points de contrôle sont des points sélectionnés de façon aléatoire dans chaque nuage de points à recaler. L'algorithme d'optimisation utilisé pour trouver le minimum global de la fonction d'erreur est le recuit simulé VFSR (*Very Fast Simulated Reannealing*), algorithme connu pour sa robustesse aux minimums locaux. Parmi ces méthodes d'optimisation nous pouvons, également, citer les méthodes du simplexe.

Les méthodes d'optimisation non linéaire classiques nécessitent, en général, la connaissance d'une transformation initiale proche de la transformation réelle. Elles ne convergent pas de manière certaine vers un minimum global [MOR96].

Il existe des méthodes d'optimisation non linéaires plus appropriées au problème de mise en correspondance telles que la décomposition en valeurs singulières, la méthode des quaternions, la méthode des quaternions doubles ou encore les réseaux de neurones artificiels (réseaux de Kohonen). Ces derniers présentent l'avantage de garantir une solution, et d'être très adaptatifs; leur inconvénient réside dans la longue phase d'apprentissage.

La méthode RANSAC (random sample consensus) développée par Fischler et Bolles [FIS81] a été utilisée pour recalibrer un modèle sur des données expérimentales bruitées. Elle peut être utilisée pour mettre en correspondance deux nuages de points 3-D, afin d'estimer ensuite la transformation rigide entre ces deux ensembles. Sa particularité est qu'elle ne traite pas toutes les données disponibles en même temps mais plutôt des échantillons sélectionnés de manière aléatoire sur l'ensemble des données.

Chu-Song Chen et al [CHU99] proposent une extension de cette méthode pour effectuer une mise en correspondance efficace et rapide. Cette méthode est appelée RANSAC-based DARCES. Son avantage par rapport aux autres méthodes d'optimisation est qu'elle ne nécessite ni extraction de primitives ni pré-modélisation. Un autre avantage très important est que cette méthode n'utilise aucune estimation initiale de la transformation.

2.6 Conclusion

L'état de l'art nous a permis d'avoir un aperçu sur différentes méthodes de mise en correspondance de données 3D (ou 3D/couleur). Ces méthodes ont été classées en trois classes: méthodes itératives, méthodes basées sur des caractéristiques et méthodes par optimisation. Cependant, cette classification n'est pas unique. En effet, les méthodes de mise en correspondance peuvent être classées suivant le type de données traitées, la taille de ces données, le critère de minimisation utilisé, etc. Nous avons vu que la méthode la plus utilisée dans le domaine du recalage de points 3D est l'algorithme ICP. Plusieurs versions de cet algorithme ont été proposées apportant à ce dernier des améliorations à tous les niveaux.

Ceci nous a encouragé à utiliser une approche similaire pour recalibrer deux ensembles de points 3-D/couleurs. Dans le chapitre qui suit, nous évaluerons les performances de trois variantes de ICP. L'une utilisant un seuil dynamique dans la phase d'appariement, la deuxième utilise en plus de ce seuil l'information photométrique, et la dernière ICP_GSC, effectue une mise en correspondance basée sur une distance euclidienne suivie d'une contrainte sur les distances couleur.

3 RECALAGE DE DEUX NUAGES DE POINTS 3D/COULEUR PAR ICP

Sommaire

3.1	INTRODUCTION	41
3.2	RECALAGE DE DONNEES 3D AVEC L'ALGORITHME ICP	41
3.2.1	<i>Principe de l'algorithme</i>	41
3.2.2	<i>Calcul de la transformation de repère</i>	43
3.2.3	<i>Elimination des points aberrants par seuillage adaptatif</i>	44
3.2.4	<i>Mise en œuvre d'un k-D tree pour la recherche du plus proche voisin</i>	45
3.3	PRISE EN COMPTE DE LA COULEUR DANS LE RECALAGE	46
3.3.1	<i>Appariement avec un critère mixte</i>	46
3.3.2	<i>Appariement avec seuillage de la distance couleur</i>	46
3.3.3	<i>Choix de l'espace couleur</i>	47
3.4	ETUDE COMPARATIVE POUR UN DEPLACEMENT SIMULE	48
3.4.1	<i>Données expérimentales</i>	48
3.4.2	<i>Influence du seuillage adaptatif pour le recalage 3D</i>	49
3.4.3	<i>Apport de la couleur dans les performances du recalage</i>	50
3.4.4	<i>Influence du recouvrement entre les deux images</i>	52
3.5	ETUDE COMPARATIVE POUR UN DEPLACEMENT REEL.....	53
3.5.1	<i>Convergence des algorithmes</i>	53
3.5.2	<i>Influence de l'espace couleur</i>	56
3.6	CONCLUSION.....	58

3.1 Introduction

Dans ce chapitre, nous nous intéressons au recalage de deux ensembles de données 3D/couleur. Nous avons vu, dans l'état de l'art précédent (chapitre 2), qu'il existe de nombreuses approches pour le recalage de données 3D. Parmi ces méthodes, l'algorithme ICP semble correspondre le mieux à notre problématique. Dans ce qui suit nous proposerons de nouvelles variantes de l'algorithme ICP et nous évaluerons leurs performances comparées à celles de l'algorithme classique proposé par Besl et McKay [BES92]. Le but du travail présenté dans ce chapitre est d'exploiter l'information couleur pour améliorer les performances de ICP. Nous adapterons cet algorithme, initialement développé pour le recalage de données 3D au recalage de données 3D/couleur.

Dans la première partie de ce chapitre, nous présentons les différentes variantes de ICP évaluées. Nous donnons le principe de l'algorithme ICP classique. Celui-ci est amélioré en appliquant un seuil adaptatif statistique sur les distances géométriques dans la phase de mise en correspondance de ICP, ceci correspond à la variante ICP_GAT. La phase de recherche du plus proche voisin est accélérée grâce à l'utilisation d'un KDTree décrit plus loin. Nous décrivons ensuite les variantes de ICP qui nous permettent d'exploiter l'information couleur: ICP_MSA et ICP_GSC.

Dans la deuxième partie de ce chapitre, ces variantes sont appliquées au recalage de deux nuages de points 3D/couleur obtenus en simulant un déplacement sur une image réelle fournie par le capteur MINOLTA de l'OSU (Ohio State university) ainsi qu'à deux images réelles 3D/couleur obtenues avec un déplacement réel du capteur MINOLTA. Le but est d'évaluer les performances de ces variantes dans le cadre d'une étude comparative.

3.2 Recalage de données 3D avec l'algorithme ICP

3.2.1 Principe de l'algorithme

Dans l'état de l'art du chapitre précédent, nous avons vu un aperçu sur l'algorithme ICP et quelques variantes de celui ci à travers les améliorations apportées au niveau de chaque étape. Dans ce qui suit nous allons décrire l'algorithme original proposé par Besl et McKay [BES92]. Les variantes de ICP utilisées et évaluées, dans notre étude, seront décrites plus loin. Celles ci concernent l'intégration des données couleur dans le processus de mise en correspondance des deux ensembles de données. De façon générale l'algorithme ICP peut être formulé comme suit:

Nous disposons de deux ensembles de points 3D : $D1$ et $D2$ et d'une connaissance approximative \tilde{T} de la transformation de repères entre ces deux ensembles. Initialiser :

$$D_2' = \tilde{T} \times D_2, \text{ et convergence} = 0.$$

K_{max} étant le nombre d'itérations maximal à effectuer si l'algorithme n'atteint pas la convergence avant,

convergence est une variable binaire, elle vaut 1 lorsque la condition de convergence de l'algorithme est satisfaite, et zéro sinon.

WHILE ($k < K_{max}$ or *convergence* = 0)

1. Calculer les points les plus proches (phase d'appariement) basé sur la distance euclidienne d entre les points $D1$ et $D2'$, il en résulte la liste de points appariés $[CP1, CP2]$; Pour chaque point du second ensemble chercher le point le plus proche dans le premier ensemble en se basant sur le calcul de la distance minimale entre chaque point $D2$ et l'ensemble des points $D1$.
2. Utiliser la liste des points appariés $[CP1, CP2]$ en entrée de l'algorithme d'estimation de la position (algorithme de recalage) en minimisant l'erreur quadratique:

$$\varepsilon = \frac{1}{N} \sum_{i=1}^N \|T * cp_2(i) - cp_1(i)\|^2 \quad (3.1)$$

Il en résulte l'estimation \hat{T} ,

3. mise à jour de : $T_k = \hat{T}$, $D2' = T_k * D2$,
4. $k = k+1$,
si condition de convergence satisfaite \Rightarrow *convergence* = 1,

FIN

L'algorithme converge lorsque l'erreur résiduelle quadratique ε (équation 3.1) de distance entre les points appariés atteint un certain seuil s , généralement pris au voisinage de la résolution du capteur de profondeur.

$$\text{convergence} = 1 \Leftrightarrow \varepsilon \leq s.$$

Cet algorithme, basique, est sensible aux mauvais appariements. Il est nécessaire d'utiliser un seuillage qui puisse permettre de pallier ce problème. Nous avons opté

pour un seuil adaptatif statistique. Celui ci, nous semble plus approprié qu'un seuil de distance fixe, du fait de la nature itérative de l'algorithme.

3.2.2 Calcul de la transformation de repère

Pour l'estimation de la transformation rigide \hat{T} de la deuxième étape de ICP il est possible d'utiliser la méthode de Horn [HOR88] basée sur les quaternions ou alors la méthode de Arun [ARU87] basée sur une décomposition en valeurs singulières SVD décrite dans ce paragraphe. Le but est de retrouver la matrice de rotations R et le vecteur de translations t qui lient l'ensemble des points $y_1 y_2 \dots y_N$ et leurs correspondants $x_1 x_2 \dots x_N$ par la relation suivante :

$$y_n = R.x_n + t + \eta_n \quad (3.2)$$

η_n étant un bruit.

R et t sont calculés en minimisant le critère d'erreur suivant :

$$\sum_{n=1}^N \|y_n - (R.x_n + t)\|^2 \quad (3.3)$$

Les centre de gravité des deux ensembles de points mis en correspondance sont donnés par :

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_n \quad (3.4)$$

$$\bar{y} = \frac{1}{N} \sum_{i=1}^N y_n \quad (3.5)$$

On définit les points centrés de chaque ensembles comme suit :

$$x'_n = x_n - \bar{x} \quad (3.6)$$

$$y'_n = y_n - \bar{y} \quad (3.7)$$

et la matrice M d'inter-covariance des ensembles x et y :

$$M = \sum_{i=1}^N x'_n y_n{}^t \quad (3.8)$$

On démontre [HOR88] que si \hat{R} et \hat{t} minimisent l'équation (3.3) alors :

$$\hat{R} = \text{ArgMax}_R \text{Trace}(R^t M) \quad (3.9)$$

$$\hat{t} = \bar{x} - \hat{R}\bar{y} \quad (3.10)$$

(S, V, D) est la décomposition en valeurs singulières de la matrice M tel que $S'MD = V$, alors la solution de l'équation (équation minimisation) est donnée par :

$$\hat{R} = DS' \quad (3.11)$$

3.2.3 Elimination des points aberrants par seuillage adaptatif

ICP_GSA signifie variante de ICP utilisant une distance euclidienne géométrique (G) pendant la phase d'appariement (recherche du point le plus proche) suivi d'un seuillage statistique adaptatif (SA) [ZHA94] pour le rejet des points aberrants. Il est noté D_{max} et est défini comme suit :

si $\mu < D$

$$D_{max} = \mu + 3\sigma \Leftrightarrow \text{Le recalage est très bon} \quad (3.12)$$

sinon $\mu < 3D$

$$D_{max} = \mu + 2\sigma \Leftrightarrow \text{Le recalage est bon} \quad (3.13)$$

sinon $\mu < 6D$

$$D_{max} = \mu + \sigma \Leftrightarrow \text{Le recalage est assez bon} \quad (3.14)$$

sinon

$$D_{max} = \xi \Leftrightarrow \text{Le recalage est mauvais} \quad (3.15)$$

avec :

$$\mu = \frac{1}{N} \sum_{i=1}^N d_i, \quad (3.16)$$

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (d_i - \mu)^2}, \quad (3.17)$$

N est le nombre de paires de points appariés,

μ est la moyenne des distances d_i entre points appariés,

σ est l'écart type de ces distances,

d_i est la distance entre deux points appariés (cp_i^1, cp_i^2)

D est une constante prise de l'ordre de la résolution géométrique du capteur. Dans notre cas nous fixerons $D = r / 5$, r étant la résolution du capteur 3-D.

Toutes les paires mises en correspondance dont la distance euclidienne dépasse le seuil D_{max} sont rejetées. Seules les autres sont considérées pour l'étape d'estimation de la transformation.

3.2.4 Mise en œuvre d'un k-D tree pour la recherche du plus proche voisin

La nature itérative de l'algorithme ICP constitue un véritable handicap du point de vue temps de calcul. La phase de recherche du plus proche voisin est le processus le plus coûteux en temps. Ce temps de calcul augmente rapidement avec la taille des données traitées, ceci est le cas du recalage de masses de données 3D/couleur (exemple : une image 3D/couleur fournie par le capteur MINOLTA de l'OSU peut contenir plus de 15000 points). En effet, la recherche "naïve" du plus proche voisin consiste à balayer, pour chaque point d'un nuage, les points du second nuage. Ce processus a une complexité de l'ordre $O(n_1.n_2)$, n_1 et n_2 étant, respectivement, le nombre de points du premier et du second nuage (exemple : $225e^6$ balayages nécessaires pour deux nuages de 15000 points). Pour des applications telles que la localisation en robotique mobile, l'algorithme ICP n'est généralement pas confronté à de grandes masses de données (recalage de nuages de points épars : tels que les amers artificiels). Il est cependant, nécessaire d'accélérer cet algorithme pour satisfaire les contraintes temps réel du fonctionnement du robot.

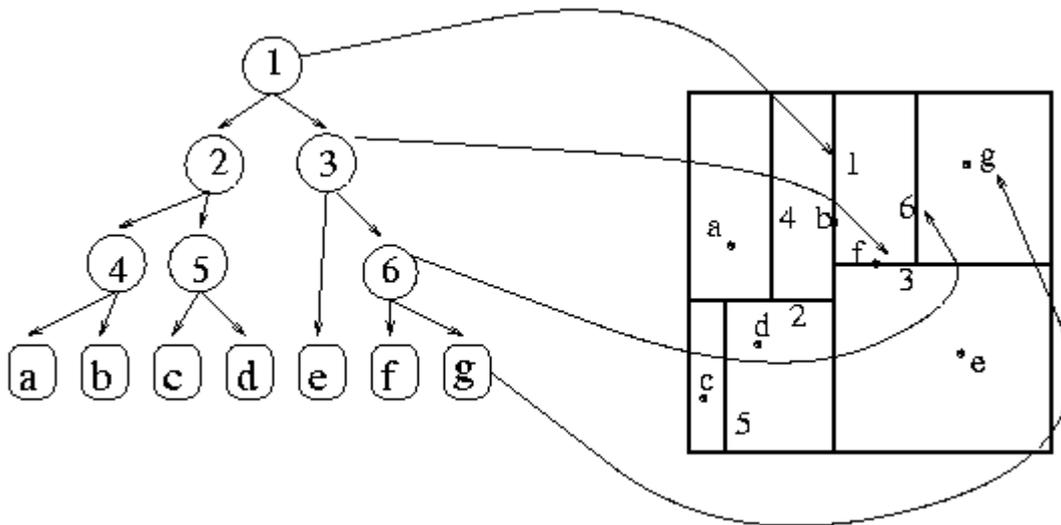


Figure 3.1 - Exemple de k-D tree (2-D tree)

La solution la plus utilisée dans la littérature pour accélérer la phase d'appariement de ICP est l'utilisation de l'arbre k-D tree. Le k-D tree est un arbre

binaire de recherche, il a été introduit pour la première fois par Bentley et al [BEN79]. La complexité moyenne de la recherche du plus proche voisin par cet arbre est en $O(n)$. Chaque nœud de cet arbre représente une partition de l'espace k -dimensionnel. La racine de l'arbre représente l'espace entier. Les feuilles de l'arbre (les extrémités) correspondent à des sous espaces contenant des sous ensembles de points mutuellement exclusifs du nuage traité. La construction d'un k -D tree est assez simple sa complexité est en $O(n \log n)$.

Pour un nuage de n points de dimension k , les points sont triés en fonction de chaque dimension. La racine de l'arbre (le point de départ) est la médiane des points suivant le premier axe (ou dimension). Les fils de la racine (le nœud père) seront les médianes des deux sous ensembles de points, cette médiane est calculée, cette fois sur le deuxième axe (suivant la deuxième dimension). Cette opération est répétée pour tous les nœuds non terminaux. La figure 3.1 illustre le processus de construction d'un k -D tree pour le cas 2-D.

3.3 Prise en compte de la couleur dans le recalage

3.3.1 Appariement avec un critère mixte

ICP_MSA est une variante de ICP qui intègre l'information couleur en utilisant une distance mixte (M) pour la recherche du point le plus proche. Cette étape est suivie d'un seuillage adaptatif (SA) pour rejeter les faux appariements. La distance mixte utilisée comporte un terme géométrique et un terme photométrique :

$$d_m = \sqrt{((1-\alpha)A) + (\alpha B)} \quad (3.18)$$

Où :

$$A = (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 \quad (3.19)$$

et

$$B = ((c_{11} - c_{21})^2 + (c_{21} - c_{22})^2 + (c_{31} - c_{32})^2) \quad (3.20)$$

Où (x_i, y_i, z_i) sont les coordonnées du point de l'image i et (c_{1i}, c_{2i}, c_{3i}) sont les composantes couleur de ce point. Le coefficient α nous permet de moduler l'apport et l'influence de l'information couleur (ou géométrique) dans la phase d'appariement. Ce coefficient, qui prend des valeurs comprises entre zéro et un, indique la confiance donnée à l'information couleur par rapport à la géométrie de l'objet à numériser. Lorsque l'objet présente une forme peu significative (pas de reliefs, exp. cas d'un tableau) on donnera plus d'importance à la texture (couleur), α sera proche de un. Ce coefficient indique, aussi, la qualité des données vis à vis des bruits de mesure. Les données les moins entachées de bruit auront un poids plus élevé.

3.3.2 Appariement avec seuillage de la distance couleur

Nous avons développé ICP_GSC qui est une autre variante de ICP. Sa particularité est qu'elle utilise la couleur indépendamment de la géométrie.

L'algorithme utilise, dans ce cas, un critère de distance géométrique (G) suivi d'une contrainte sur la couleur (un seuillage couleur : SC). Le seuil couleur est appliqué aux distances euclidiennes photométriques (RVB, ou dans un autre espace). Celui-ci est fixé, dans notre cas, à la médiane de ces distances.

3.3.3 Choix de l'espace couleur

Une couleur est généralement représentée par trois composantes. Ces composantes définissent un espace couleur. Un espace couleur est une représentation géométrique de la couleur dans l'espace (généralement en trois ou quatre dimensions) pour un modèle de couleur donné. Un modèle de couleur est un modèle mathématique pour décrire, classer et comparer les couleurs. Il existe de très nombreux espaces couleur, le plus populaire est l'espace RVB utilisé notamment par les caméras couleur ou par les systèmes d'affichage vidéo (écrans à tube cathodique, LCD ou encore à plasma). Pour l'œil humain, les couleurs sont une combinaison pseudo linéaire de petites, moyennes et grandes longueurs d'ondes ce qui correspond approximativement aux trois couleurs primaires utilisées dans l'espace RVB [BOV00]. Un autre espace couleur très utilisé en traitement de l'image et pour l'affichage est l'espace YIQ (luminance, in-phase chromatic, quadratic chromatic). Cet espace a pour particularité de pouvoir séparer la luminance de la composante de chromaticité. Nous pouvons citer, également, l'espace HSV (Hue, Saturation, Value). Cet espace correspond, en quelques sortes, à la définition, au quotidien, que nous avons de la couleur. En effet, la première composante H traduit la teinte qui correspond à la nuance de couleur (où la couleur se situe dans le spectre des couleurs rouge, jaune et pourpre) elle varie de 0 à 360. La saturation S décrit la pureté de la teinte de la couleur par rapport au blanc, elle varie de 0 à 100 (%). Une couleur de teinte rouge entièrement saturée correspond à une couleur ne contenant pas de blanc. Un rouge moins saturé tend au rose. La dernière composante V correspond à la brillance de la couleur, elle varie de 0 à 100 (%).

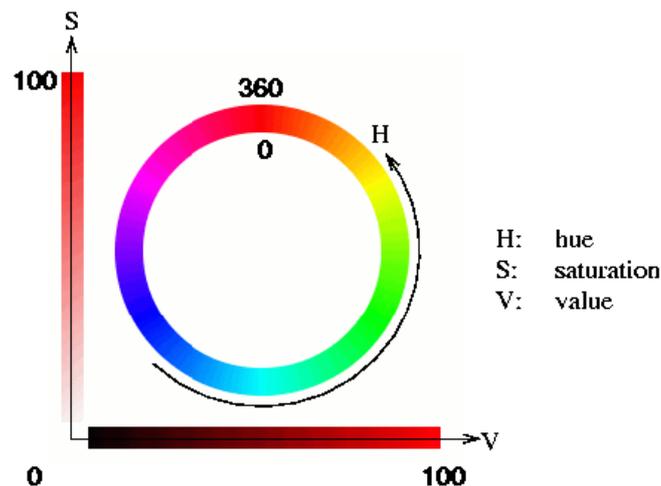


Figure 3.2 - Représentation de l'espace couleur HSV.

Tous les espaces couleur sont liés, entre eux, par des relations linéaires ou non-linéaires. Nous pouvons ainsi passer d'une représentation à une autre grâce aux transformations liant ces espaces. Le passage de l'espace RVB vers l'espace YIQ (et inversement de YIQ vers RVB) se fait par une transformation linéaire (équation 3.21) contrairement à la transformation entre les espaces RVB et HSV.

$$\begin{pmatrix} Y \\ I \\ Q \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & 0.311 \end{pmatrix} \begin{pmatrix} R \\ V \\ B \end{pmatrix} \quad (3.21)$$

Le choix de l'espace couleur utilisé dans les variantes de ICP couleur dépend de la qualité de l'information couleur dont nous disposons. Lors de l'acquisition des deux vues à recalcr, la principale difficulté concernant l'acquisition des données photométrique concerne (en plus des bruits de mesure liés au capteur) les changements d'illumination dus au changement du point de vue ainsi qu'aux variations de lumière liées à la source lumineuse. Dans le cas idéal où la luminance reste constante un espace couleur tel que RVB peut être utilisé. Par contre, pour le cas général, il est préférable de travailler dans un espace couleur qui puisse permettre de séparer la luminosité de la chromaticité pour pouvoir ainsi annuler l'effet du changement de luminosité en ne considérant que les deux autres composantes. Nous verrons plus loin (paragraphe 3.4.6) comment exploiter cet avantage en utilisant l'espace YIQ pour une variante de ICP couleur au lieu de travailler dans l'espace RVB.

3.4 Etude comparative pour un déplacement simulé

Dans cette section nous proposons une étude comparative (§3.4.2, §3.4.3 et §3.4.4) des trois variantes de ICP détaillées plus haut (ICP_GSA, ICP_MSA et ICP_GSC) les performances de ces variantes seront également comparées à l'algorithme classique. Le paragraphe suivant présentera les données expérimentales utilisées pour cette évaluation.

3.4.1 Données expérimentales

Le choix des données 3-D/couleur à utiliser pour notre étude est important. Nous avons, dans le cadre d'autres travaux, utilisé des données synthétiques générées par ordinateur. Nous étions alors confronté au problèmes liés à la forme à générer. Celle ci ne devait pas présenter de symétries, cas particulier, qui pénalise l'algorithme ICP. Ce problème peut, bien sur, être résolu avec des formes 3-D aléatoires, qui peuvent, par exemple, être générées à l'aide de fractales 3-D (génération de terrains 3-D, fractales de Mandelbrot). Les données de l'OSU [OSU] ont été choisies pour se rapprocher du cas réel.

Pour notre étude nous utilisons une partie (2000 points) d'une image 3-D/couleur notée Image1. Celle ci subit un déplacement simulé T_s (défini plus loin) qui

nous permet d'obtenir la deuxième image à recalier: Image2 (figure 3.3). Dans ce cas, aucun bruit de mesure n'est ajouté à la deuxième image.

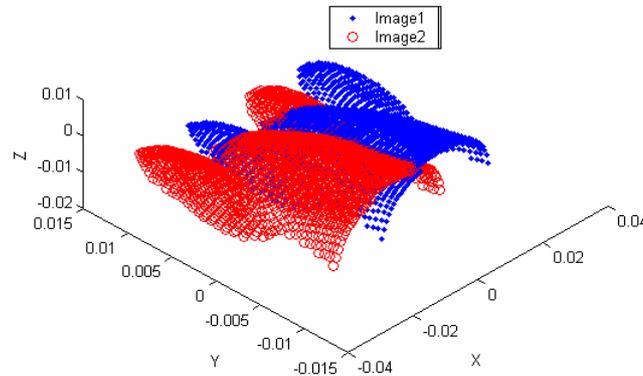


Figure 3.3 - Les deux images à recalier Image1 et Image2

Le déplacement simulé T_s est défini par une translation suivant l'axe X : $t_x = 30$ mm. Ceci pour une étendue des données en X de 45 mm. Le recouvrement entre les deux images, dans ce cas, est de l'ordre de 40 % (recouvrement relativement petit). La transformation fournie aux algorithmes à évaluer, T_i , est obtenue en biaisant T_s avec un biais sur les translations égal à 5 mm (soit 1/3 du déplacement, biais important), et un biais sur les rotations égal à 5° .

3.4.2 Influence du seuillage adaptatif pour le recalage 3D

La figure.3.4 montre les erreurs résiduelles de distance, entre les points appariés, pour ICP et pour ICP_GSA, en fonction des itérations de l'algorithme. Dans le cas de ICP (classique), l'erreur décroît de manière monotone vers un minimum local et l'atteint au bout de 16 itérations. Par contre, nous pouvons remarquer une allure différente concernant le cas de ICP_GSA. En effet, celle ci augmente pendant les 4 premières itérations pour se stabiliser à une erreur résiduelle relativement importante pendant plus de 25 itérations, ensuite celle ci chute brusquement vers une valeur inférieure à celle de ICP au bout de plus de 40 itérations. Cette non monotonie constatée au démarrage de l'algorithme ICP_GSA est due au fait qu'à la première itération nous forçons le seuil D_{max} à une valeur exagérée égale à $50D$. Ceci pour éviter de rejeter la totalité des points appariés à la première itération (les deux images étant, alors, relativement éloignées).

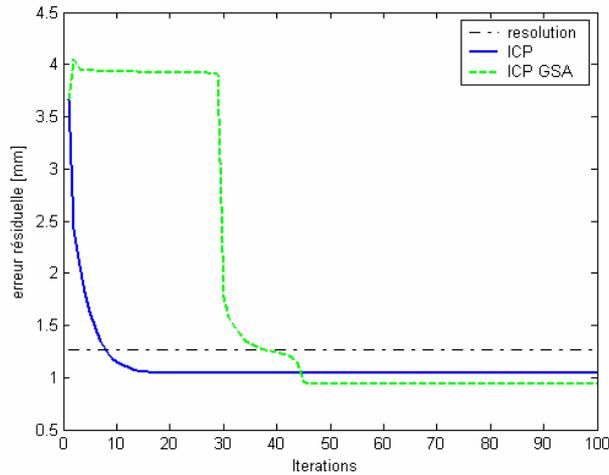


Figure 3.4 - Evolution de l'erreur résiduelle de distance pour ICP et ICP_GSA

L'algorithme ICP_GSA nécessite la définition du paramètre D . Ce dernier reste un handicap lorsqu'on veut utiliser le seuil adaptatif. Il est, en effet, difficile à ajuster. Prendre une valeur trop grande induit, dans la plupart des cas un comportement imprévisible (non monotone). Si D est, par contre pris trop petit, l'algorithme a tendance à ne pas converger (le seuil adaptatif ne fonctionne plus correctement, D_{max} prendra en permanence la valeur de la médiane : seuil trop sévère).

Il est, aussi, important de noter que l'erreur résiduelle dans les deux cas reste inférieure à la résolution du capteur. Ceci est du à l'absence des bruits de mesure additif sur la deuxième vue.

Remarque : Le lecteur peut remarquer la différence dans la valeur de la résolution entre la figure 3.4 et la figure 3.5 (elle vaut le double sur la figure 3.4). Ceci est du au fait que pour certains essais les données ont été échantillonnées (prise en compte d'un point sur deux au lieu de tous les points de l'image d'origine), et ce, dans un souci de rapidité d'exécution. Ceci n'affecte, en aucun cas, les conclusions tirées des études comparatives.

3.4.3 Apport de la couleur dans les performances du recalage

3.4.3.1 Appariement avec un critère mixte

Les courbes de la figure 3.5 décrivent les erreurs résiduelles générées dans le cas de ICP_GSA et ICP_MSA. De la comparaison entre les deux courbes, nous pouvons conclure que ICP_MSA est plus performant que ICP_GSA. L'erreur générée par ICP_MSA (utilisant une distance mixte et un seuil adaptatif) est nettement inférieure à celle générée par ICP_GSA (utilisant une distance géométrique et un seuil adaptatif). L'information couleur améliore, dans ce cas, les performances de l'algorithme.

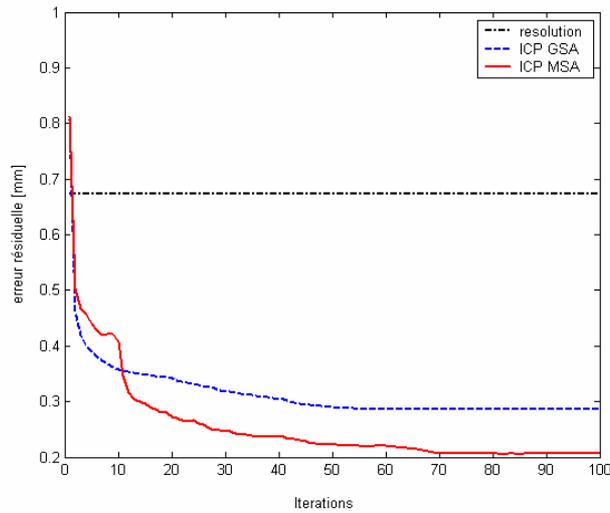


Figure 3.5 - Evolution de l'erreur résiduelle de distance pour ICP_GSA et ICP_MSA

ICP_MSA nécessite, en plus de l'ajustage du paramètre D , la définition du coefficient α . Dans le cas de la figure (3.5) celui ci est pris égal à 0.5 car vu le type d'objet numérisé, nous accordons le même poids aux données couleur et aux données géométriques. Nous verrons, plus loin, que dans le cas général, il sera délicat de définir ce coefficient de façon pertinente.

3.4.3.2 Appariement avec seuillage de la distance couleur

La variante ICP_GSC, décrite dans le paragraphe 3.3.2, nous permet de pallier la difficulté liée à la définition du paramètre D pour ICP_GSA et ICP_MSA, et du coefficient α pour ICP_MSA.

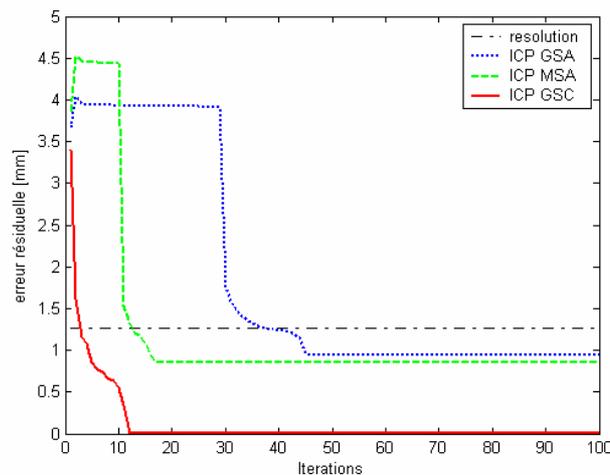


Figure 3.6 - Evolution de l'erreur résiduelle de distance pour ICP_GSA, ICP_MSA ICP_GSC

Nous pouvons constater sur la figure 3.6 l'efficacité de la variante ICP_GSC. L'erreur atteint une valeur proche de zéro en moins de 15 itérations. Ceci confirme, donc que la couleur n'est pas exploitée correctement pour ICP_MSA (coefficient α mal réglé). On peut aussi considérer que les mauvais appariements ne sont pas rejetés de façon efficace par le seuil adaptatif dans ICP_GSA et ICP_MSA (influence du paramètre D).

Algorithmes	ICP_GSA	ICP_MSA	ICP_GSC
Erreurs			
tx [mm]	1.8e-005	1.2e-005	~0
ty [mm]	1.7e-005	4.1e-005	~0
tz [mm]	7.1e-005	6.6e-005	~0
rho [d°]	0.024	0.2	~0
theta [d°]	0.07	0.04	~0
phi [d°]	0.3	0.23	~0
Itération de convergence	45	18	12

Tableau 3.1 - Erreurs de recalage des algorithmes évalués

Le tableau ci dessus (tableau.1) contient les erreurs de position et d'orientation calculées entre la transformation correcte T_s (connue en vérité terrain) et les transformations fournies par chaque algorithme à l'itération de stabilisation (l'itération pour laquelle ces erreurs et l'erreur résiduelle de distance se stabilisent). On peut remarquer des erreurs proches de zéro dans le cas de ICP_GSC. Les deux autres variantes (GSA et MSA) génèrent des erreurs du même ordre de grandeur, même si celles générées par ICP_MSA restent, globalement, légèrement inférieures à celles générées par ICP_GSA. Ceci dit, lorsqu'on compare le nombre d'itérations nécessaire pour atteindre ces minimums locaux, on constate qu'il a fallu 45 itérations pour ICP_GSA contre seulement 18 pour ICP_MSA. Par ailleurs, ICP_GSC ne nécessite que 12 itérations pour atteindre des erreurs nettement moins importantes. Cet algorithme reste, de loin, le plus performant des trois, en ce qui concerne : les erreurs de recalage, l'erreur résiduelle et le nombre d'itérations nécessaires pour la convergence (la vitesse de convergence).

3.4.4 Influence du recouvrement entre les deux images

Il s'agit d'évaluer les algorithmes proposés vis à vis de l'importance du recouvrement entre les deux images à recaler. Dans ce qui suit nous faisons varier le

déplacement simulé Ts de manière à faire varier le recouvrement entre les deux images.

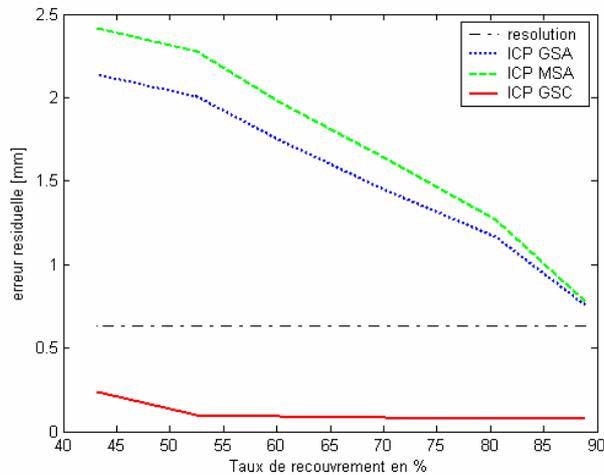


Figure 3.7 - Evolution de l'erreur résiduelle de distance en fonction du recouvrement des deux images

ICP_GSC reste, jusqu'à un recouvrement de 55 %, insensible aux variations de celui-ci. L'erreur augmente légèrement pour un recouvrement de 45 %. Quant aux deux autres variantes (ICP_MGA et ICP_MSA), les erreurs résiduelles varient plus. ICP_MSA est plus sensible à la qualité du recouvrement. Ceci dit, lorsque celui-ci est très important (de l'ordre de 90 %) les performances des deux variantes sont, globalement, les mêmes.

3.5 Etude comparative pour un déplacement réel

3.5.1 Convergence des algorithmes

Pour l'étude des performances des variantes de ICP face aux bruits de mesure nous utilisons dans ce qui suit des données réelles. Nous n'allons plus simuler de déplacement pour générer une nouvelle image considérée comme la seconde vue à recalculer. Nous allons recalculer deux images réelles acquises par le capteur MINOLTA de l'OSU [OSU]. Ces deux images sont acquises pour deux positions différentes de l'objet en rotation sur un plateau tournant (rotation de 20° autour de l'axe y). Il est, donc, évident que dans ce cas les algorithmes évalués seront confrontés aux bruits de mesures (géométrique et photométrique). Sur la figure 3.8 sont affichées l'image1, et l'image2 transformée dans le repère de la première, avec la transformation approximative fournie par l'OSU [OSU]. Nous pouvons constater que cette approximation ne recalcule pas correctement les deux images. C'est la transformation utilisée pour initialiser les algorithmes évalués.

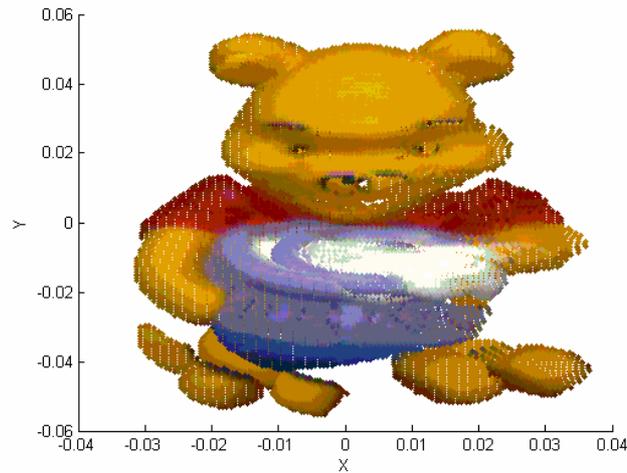


Figure 3.8 - Les deux images réelles recalées par la transformation initiale

La figure 3.9 montre les erreurs résiduelles générées par les variantes évaluées dans les paragraphes précédents. Ces courbes ont été obtenues en recalant les 2000 premiers points de chaque image (balayage dans le plan XY suivant l'axe X) et qui correspondent, approximativement, suivant l'axe Y, au premier cinquième de chaque image (le haut de la tête de l'oursin). Nous pouvons constater, dans ce cas que ICP_GSA donne les meilleures performances comparé aux autres variantes ICP_MSA ($\alpha = 0.5$ dans l'équation.7) et ICP_GSC. Cette dernière génère, cependant, une erreur résiduelle du même ordre de grandeur que celle générée par ICP_GSA.

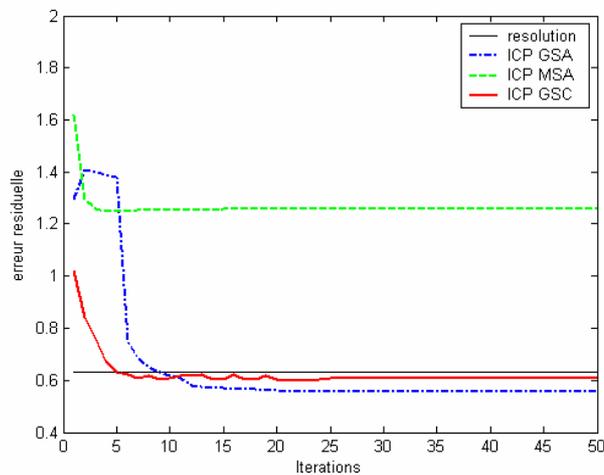


Figure 3.9 - Erreurs résiduelles pour ICP_GSA, ICP_MSA et ICP_GSC

Pour ces deux variantes l'erreur reste inférieure à la résolution du capteur. ICP_MSA est de loin le moins performant des trois algorithmes. Ces résultats sont justifiés par le fait que pour les portions des deux images recalées, l'information couleur n'est pas très significative. En effet cette partie de l'image n'est pas très riche en couleurs. Ce qui n'est pas le cas de la forme géométrique. Le coefficient α dans ICP_MSA n'a, donc pas été correctement choisi, il est pris égal à 0.5.

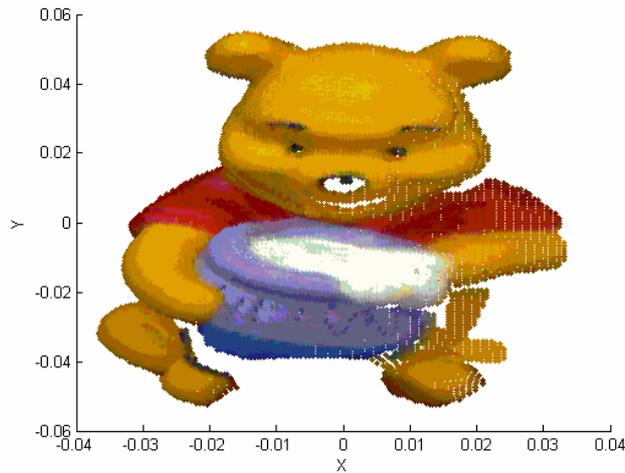


Figure 3.10 - Les deux images réelles recalées par ICP_GSC

Sur la figure ci-dessus (figure.3.10) nous avons affiché le résultat du recalage avec ICP_GSC. On peut vérifier la qualité du recalage en comparant ces images avec celle de la figure 3.8.

Un deuxième recalage a été effectué sur deux autres parties des images réelles de la figure 3.8. Cette fois, nous recalons les parties des deux images où les couleurs sont plus significatives que le cas précédent (près du centre de chaque image). La figure 3.11 illustre les erreurs résiduelles des trois variantes utilisées. Dans ce cas, les courbes des erreurs des deux variantes ICP_GSA et ICP_MSA ont une forme similaire. L'erreur résiduelle pour ces deux variantes ne décroît pas de façon monotone mais augmente et se stabilise à une valeur relativement importante (supérieure à $3.r$, r étant la résolution). Ceci est dû à un mauvais choix du paramètre D dans le calcul du seuil adaptatif, celui-ci pourrait être trop petit. ICP_GSC est le plus performant des trois algorithmes. L'erreur résiduelle, pour cette variante, décroît de façon monotone vers un minimum local de l'ordre du double de la résolution.

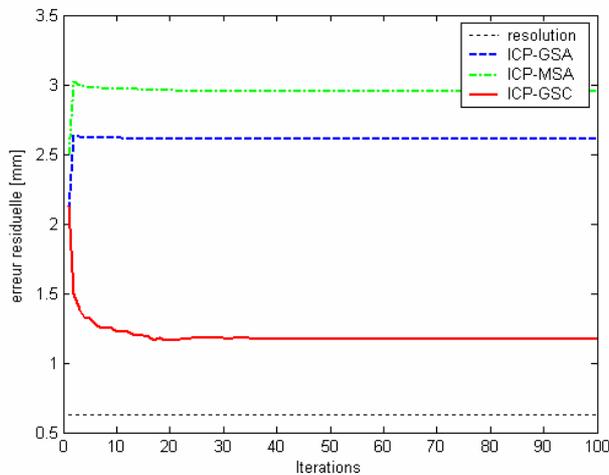


Figure 3.11 - Erreurs résiduelles pour ICP_GSA, ICP_MSA et ICP_GSC

3.5.2 Influence de l'espace couleur

Dans ce qui suit nous nous intéressons à l'influence de l'espace couleur utilisé pour le calcul des distances (distance mixte ou distance couleur) dans la phase de mise en correspondance des variantes couleur de ICP. Notre but est de proposer une variante ICP couleur robuste aux changements d'illumination. Nous utilisons dans ce qui suit deux parties des deux images réelles de l'OSU où l'information couleur est significative. La figure 3.12 montre ces deux vues recalées par la transformation approximative utilisée dans les paragraphes précédents pour initialiser les variantes de ICP.

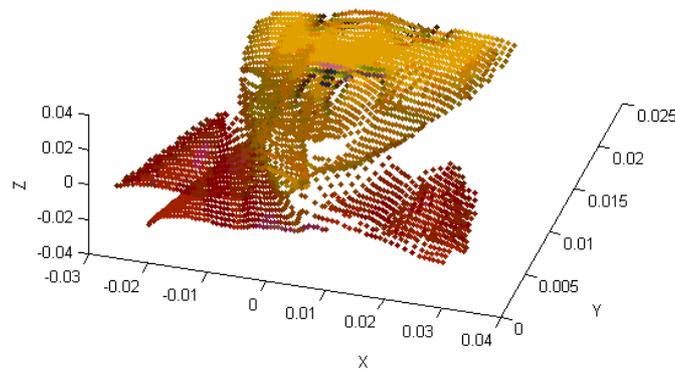


Figure 3.12 - Les deux parties des images recalées par la transformation initiale T_i

L'erreur résiduelle de la figure 3.11 est, dans le meilleur des cas, sensiblement égale au double de la résolution du capteur. Ceci est dû aux bruits de mesures géométrique et photométrique. L'erreur liée à la variation des couleurs est due à la variation de l'éclairage d'une image à l'autre. Ceci se répercute sur l'intensité de la

lumière émise par l'objet éclairé (à numériser), sans, pour autant, affecter la couleur intrinsèque de celui-ci [JOH96]. Nous pouvons réduire l'effet de la variation de l'intensité de la lumière en utilisant un espace couleur qui puisse nous permettre d'intervenir sur cette grandeur. Au lieu de travailler dans l'espace RGB, nous utiliserons, dans ce qui suit, l'espace YIQ (Y: représente la luminance, I et Q sont les composantes chromatiques). On peut, ainsi, séparer l'intensité Y et la couleur intrinsèque I et Q. L'expression de la distance mixte d_m dans ICP_MSA (RGB et YIQ) devient alors :

$$d_m = \sqrt{((1-\alpha)A) + (\alpha C)} \quad (3.22)$$

La distance euclidienne couleur utilisée dans la phase de rejet des faux appariements dans ICP_GSC est définie comme suit :

$$d_c = \sqrt{C} \quad (3.23)$$

Où A est donné par l'équation (3.19) et :

$$C = (a(c_{11} - c_{21})^2 + b(c_{21} - c_{22})^2 + c(c_{31} - c_{32})^2) \quad (3.24)$$

a , b et c sont les poids de chaque composante couleur. Leurs valeurs sont comprises entre 0 et 1 (inclus). Ils traduisent la contribution de chaque composante couleur dans le calcul des distances dm et dc .

$$\text{Pour l'espace couleur RGB on prend } \begin{cases} a = 1, \\ b = 1, \\ \text{et} \\ c = 1 \end{cases}$$

$$\text{Pour l'espace couleur YIQ on prend } \begin{cases} a = 0, \\ b = 1, \\ \text{et} \\ c = 1 \end{cases}$$

Les courbes de la figure 3.13 représentent les erreurs résiduelles des cinq variantes de ICP utilisées pour le recalage des images de la figure.10. On s'intéresse, ici, à l'effet de l'espace couleur (RGB ou YIQ) sur les performances des deux variantes de ICP/couleur (ICP_MSA et ICP_GSC) utilisées dans les paragraphes précédents. ICP_MSA_RGB et ICP_GSC_RGB correspondent aux variantes testées précédemment (travaillant dans l'espace couleur RGB), ICP_MSA_YIQ et ICP_GSC_YIQ correspondent aux variantes de ICP/couleur utilisant l'espace couleur YIQ. Nous pouvons constater que les trois variantes utilisant un seuillage adaptatif génèrent une courbe d'erreur se stabilisant à une valeur nettement supérieure à la résolution du capteur. Ceci est dû au mauvais choix de D . Par contre l'erreur générée dans le cas de ICP_GSC_RGB et ICP_GSC_YIQ décroît de manière monotone vers un

minimum local. Ceci dit, nous pouvons remarquer que cette erreur est moins importante dans le cas de ICP_GSC_YIQ comparée à celle générée par ICP_GSC_RGB. L'intensité de la lumière a, donc, un réel effet sur les performances de notre variante de ICP. Le changement de l'espace couleur nous a permis d'atténuer cet effet et d'améliorer les performances de ICP_GSC.

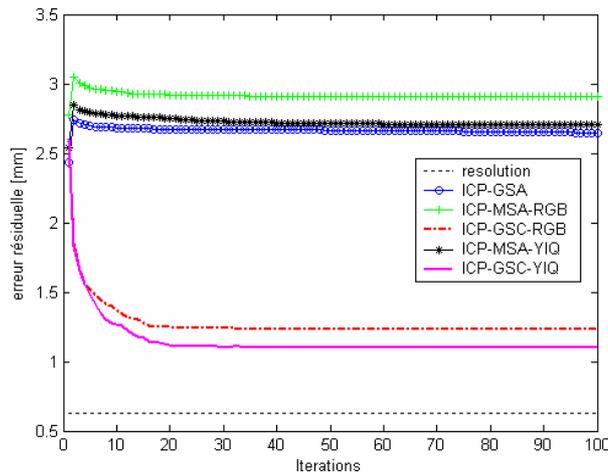


Figure 3.13 - Effet de l'espace couleur

3.6 Conclusion

Nous avons présenté des variantes de ICP pour le recalage de deux ensembles de données 3D/couleur. Ces variantes intègrent la couleur avec des approches différentes. Le but du travail effectué étant de montrer l'intérêt de l'information photométrique, nous avons comparé ces variantes avec une autre version de ICP n'utilisant pas la couleur : ICP_GSA. Cette dernière est une version améliorée de l'algorithme classique utilisant un seuillage adaptatif pour rejeter les mauvais appariements. Nous avons montré qu'elle améliore les performances de ICP.

L'information couleur a été intégrée par deux approches différentes. La première utilise une distance mixte pour le calcul des mises en correspondances : elle correspond à la variante ICP_MSA (distance mixte et seuillage adaptatif). La seconde intègre la couleur séparément de l'information géométrique : ICP_GSC utilise un critère de distance géométrique pour l'appariement suivi d'une contrainte sur les distances couleur des points appariés pour rejeter les faux appariements. Cette dernière reste la plus performante des trois en l'absence de bruits de mesure. Nous avons remarqué que le seuillage statistique adaptatif restait très sensible au choix du paramètre D (c'est le point faible de cette méthode). De même, la nécessité de régler le coefficient α en fonction de la nature des données à recalculer rend délicate l'utilisation d'une distance mixte pour l'intégration de la couleur.

4 APPARIEMENT DE POINTS COULEUR CARACTERISTIQUES

Sommaire

4.1	INTRODUCTION	61
4.2	ETAT DE L'ART.....	62
4.2.1	<i>Détection de points d'intérêt</i>	62
4.2.2	<i>Appariement d'images 2D</i>	69
4.2.3	<i>Elimination des faux appariements</i>	74
4.2.4	<i>Conclusion</i>	76
4.3	EXTRACTION DE POINTS D'INTERET	77
4.3.1	<i>Exemple 1: peinture à l'huile</i>	79
4.3.2	<i>Exemple 2: Figurine 3D</i>	80
4.3.3	<i>Exemple 3: boîte cylindrique</i>	81
4.4	CARACTERISATION ET APPARIEMENT DES POINTS D'INTERET.....	84
4.4.1	<i>Vecteur d'invariants</i>	84
4.4.2	<i>Normalisation des couleurs</i>	85
4.4.3	<i>Appariement des points d'intérêt</i>	87
4.5	VALIDATION EXPERIMENTALE DE LA METHODE D'APPARIEMENT	88
4.5.1	<i>Influence de la normalisation des couleurs</i>	89
4.5.2	<i>Influence de la taille de la fenêtre normalisation des couleurs</i>	91
4.6	CONCLUSION.....	98

4.1 Introduction

Dans ce chapitre nous nous intéressons à la mise en correspondance d'images 2D couleur. Il existe plusieurs approches pour la mise en correspondance d'images 2D. Ces approches peuvent être classées en deux principales catégories :

- approches basées sur les données photométriques,
- approches basées sur les données géométriques,

Le premier type d'approche utilise l'information de texture pour la mise en correspondance d'images. Parmi ces méthodes nous pouvons citer les méthodes par corrélation du signal, les méthodes utilisant des descripteurs statistiques (histogrammes) et les méthodes utilisant les invariants différentiels. La section 4.2.2.2 donne un aperçu sur ces différentes méthodes.

La deuxième catégorie d'approches utilise les données géométriques dans chaque image à mettre en correspondance. Dans ce cas, la ressemblance est calculée sur la structure de contour globale de la scène avec des techniques telles que la relaxation. Nous ne nous étendrons pas davantage sur ce type d'approche. En effet, notre choix s'est porté sur les méthodes utilisant l'information photométrique des images. Ce choix est justifié par la difficulté d'extraire des contours de façon correcte dans les images couleur bruitées délivrées par le capteur de numérisation 3D/couleur.

Notre objectif est donc d'exploiter l'information couleur attachée aux points 3D pour extraire un ensemble de points caractéristiques dans la paire d'images 2D couleur correspondant aux deux nuages de points 3D/Couleur à recalcr. Ces points caractéristiques sont ensuite appariés à l'aide de méthodes de mise en correspondance 2D (telles que l'utilisation de vecteur d'invariants différentiels § 4.4.1). A partir de ces appariements 2D, nous pouvons déduire les appariements 3D correspondant pour initialiser les méthodes de recalage 3D/Couleur présentées dans le chapitre 5.

Dans la deuxième section de ce chapitre nous présentons un état de l'art des détecteurs de points d'intérêt (détecteurs de points caractéristiques) les plus utilisés pour la mise en correspondance d'images 2D :

- le détecteur de Kitchen Rosenfeld,
- le détecteur de Harris Stephens,
- et le détecteur de Harris précis.

Nous passerons, également, en revue les principales méthodes de mise en correspondance d'images 2D basées sur :

- la corrélation de fenêtres de voisinage,
- la distance entre descripteurs statistiques (par exemple : distance d'histogrammes),
- la distance entre vecteurs d'invariants différentiels.

Nous donnons, en conclusion de l'état de l'art, notre choix de détecteur de points d'intérêts ainsi que le choix de la méthode de mise en correspondance que nous allons utiliser dans les sections suivantes et dans le chapitre 5.

La particularité du travail que nous présentons dans ce chapitre est l'utilisation d'images couleur 2D correspondant à des points 3DC. En effet, dans notre cas, les images couleurs 2D à mettre en correspondance proviennent de capteurs 3D/Couleur utilisés pour l'acquisition des données 3DC à recaler, et ce, dans le but d'une reconstruction 3DC.

Dans la littérature les méthodes de détection de points d'intérêts et les méthodes de mise en correspondance 2D sont appliquées à des images 2D obtenues avec un déplacement, relativement simple, entre les images à recaler ou alors pour des applications telles que le calibrage de systèmes de stéréovision.

Dans notre cas, les images 2D à recaler proviennent d'images 3DC obtenues avec un déplacement 3D quelconque du capteur entre les deux prises de vue. Ceci a pour conséquence une déformation possible des motifs couleur des images. Par ailleurs, lors de l'acquisition des différentes vues 3DC de l'objet à reconstruire les déplacements entre les vues sont généralement plus importants qu'en stéréovision avec plus de degrés de liberté. Ceci a pour conséquence que les images subissent des variations relativement importantes et non uniformes de l'intensité, ce qui constitue une vraie difficulté dans la phase de mise en correspondance.

4.2 Etat de l'art

4.2.1 Détection de points d'intérêt

La notion de point d'intérêt fut introduite pour la première fois dès 1977 par H. Moravec [MOR77]. Un point d'intérêt est un point qui doit avoir des caractéristiques plus significatives que les autres. Ces points correspondent à des changements bidimensionnels du signal. Nous pouvons en citer, les coins en 'L' et les points de jonctions en 'T'. De très nombreux travaux ont été menés pour l'extraction de ces points dans l'image

Les méthodes d'extraction de points d'intérêt en 2D peuvent être classées en deux principales catégories :

- Les méthodes basées sur les contours. Ces méthodes se basent sur l'extraction au préalable de primitives géométriques telles que les contours, les segments de droite et les courbes. Dans une seconde étape des coins sont extraits correspondant aux courbures maximales de ces contours. les méthodes basées sur le signal. Celles-ci, connues aussi sous le nom de *méthodes iconiques*, ont pour principe d'extraire directement les points d'intérêt à partir de l'information photométrique.

4.2.1.1 Méthodes basées sur les contours

Ce type de méthode existe depuis longtemps. Elles sont basées sur la recherche des points de courbure maximale le long des chaînes de contour. Ces méthodes nécessitent, donc, un prétraitement des images qui correspond à l'étape d'extraction de contours (étape de segmentation a priori).

Berzins [BER84], Asada et Brady [ASA86] et Medioni et Yasumoto [MED86] ont développé des approches basées sur les contours au milieu des années quatre-vingt. Les contours sont extraits à l'aide d'un opérateur de contour gaussien. Un contour représente le passage à zéro du Laplacien symétrique de cet opérateur. Les auteurs obtiennent ainsi une représentation en courbes de l'image. Ces méthodes extraient les points d'intérêt (les coins) correspondant aux points de courbure maximale en utilisant un seuil approprié.

L'un des inconvénients de ces méthodes est le fait de dépendre fortement de la qualité de l'extraction des contours et de la qualité de ces contours même. En effet, les coins extraits comme points d'intérêt proviennent de courbes connectées dans l'image à traiter. Il est impossible de faire la différence entre un point qui provient d'un 'véritable' coin existant physiquement dans la scène 3D et un coin résultant de l'intersection de deux courbes dans l'image 2D et qui n'a aucun sens physique. Aussi ces méthodes sont très sensibles aux bruits.

4.2.1.2 Méthodes basées sur le signal

Les méthodes basées sur le signal extraient les points d'intérêt directement à partir du signal. Ceci signifie qu'elles ne nécessitent aucun prétraitement de l'image contrairement aux méthodes basées sur les contours. Ceci rend généralement les méthodes basées signal plus performantes que celles basées contour, car le prétraitement des images dans le cas des méthodes basées contour provoque une perte de l'information.

La classification des méthodes d'extraction de points d'intérêt que nous utilisons regroupe les méthodes basées signal et d'autres méthodes, non citées précédemment et qui sont basées sur un modèle théorique du signal (modèle d'un point d'intérêt). En effet, pour nous, ces deux types de méthodes se basent sur le signal.

Ces méthodes considèrent l'image comme une surface d'intensité $I(x,y)$. Ainsi, les dérivées des courbures du signal peuvent être calculées pour détecter les points d'intérêt. Moravec [MOR77] extrait des *coins* dans l'image en se basant sur le calcul de l'autocorrection du signal sur une fenêtre de l'image déplacée le long des deux directions u et v . Un point d'intérêt correspond à un point pour lequel l'intensité de son voisinage varie considérablement dans les deux directions (verticale et horizontale). Autrement dit, un point d'intérêt (x,y) de la surface $I(x,y)$ correspond à un point pour lequel la fonction $E(x,y)$ (équation 4.1) prend une valeur relativement grande.

$$E(x, y) = \sum_{u,v} w(u, v) |I(x + v, y + u) - I(u, v)|^2, \quad (4.1)$$

w correspond à la fenêtre de voisinage à considérer, elle vaut 1 à l'intérieur du voisinage et zéro à l'extérieur.

Beudet [BEA78] fut le premier à introduire un détecteur de points d'intérêt basé sur la dérivée deuxième du signal. Il calcul la mesure DET (equation.4.2) liée indirectement à la courbure Gaussienne de $I(x,y)$ pour déterminer la présence d'un point d'intérêt :

$$DET = I_{xx}I_{yy} - I_{xy}^2 \quad (4.2)$$

I_{xx} et I_{yy} sont respectivement les dérivées secondes de $I(x,y)$ en fonction de x , y ; I_{xy} est la dérivée mixte.

Cette mesure correspond au déterminant du Hessien de $I(x,y)$. Beudet [BEA78] utilise pour le calcul de ce Hessien le développement de Taylor de second ordre de $I(x,y)$. Les points d'intérêts correspondent aux points où cette mesure est maximale. La mesure DET est invariante aux rotations. Ceci dit, cette technique a deux principaux inconvénients. Le premier concerne le fait que cette méthode utilise des dérivées secondes de l'image ce qui la rend vulnérable au bruit. Le second concerne le fait que ce détecteur extrait les points à l'intérieur des coins mais non sur ceux là [DER93]. Dreschler et Nagel [DRE82] reprennent la technique de Beudet [BEA78] pour déterminer les *courbures gaussiennes* au niveau des pixels se situant entre les extrema de cette grandeur dans l'image entière. En effet, les auteurs constatèrent que la *courbure gaussienne* pouvait prendre des valeurs importantes sur des contours marqués. Cette grandeur étant le résultat du produit des deux courbures principales d'une surface [SCH96]. A partir d'une étude sur le comportement de la *courbure gaussienne* autour d'un coin (simulé par son modèle théorique) les auteurs définissent un point d'intérêt comme un point où la *courbure gaussienne* s'annule et change de signe [SCH96].

D'autres approches de détection de points d'intérêt consistent à calculer un indice *d'intérêt* (nommé *cornerness* d'un point à partir du produit de l'amplitude du gradient avec le taux de variation de la direction de cette amplitude pour ce point. Kitchen et Rosenfeld [KIT82] utilisent cette approche pour extraire des coins dans l'image (cf. section 4.2.1.2.1).

Zuniga et Haralick [ZUN83] proposent un détecteur de coins très proche de celui proposé par Kitchen et al [KIT82]. Un coin est défini par un changement significatif au niveau de la courbure d'une courbe plane. Au lieu du polynôme quadratique utilisé par Kitchen et al (section Kitchen), les auteurs utilisent le modèle de facettes proposé par Haralick [ZUN83]. A chaque pixel est associée une surface bi-cubique polynomiale par moindres carrés. Dans Shah et Jain [SHA84] et Nagel [NAG83] toutes les méthodes basées sur le calcul des dérivées sont équivalentes à quelques détails près.

L'un des inconvénients de ces méthodes est qu'elles se basent sur un modèle de coin en 'L'. La détection et la localisation d'autres types de points d'intérêt, ne sont pas fiables. L'amplitude du gradient, information sur laquelle sont basées ces méthodes, n'est pas localisée correctement au voisinage immédiat des coins et le devient encore moins lorsque l'angle du coin en question est fermé. Ceci implique des erreurs de localisation du coin recherché. Aussi, à cause de l'amplitude du gradient, ces méthodes sont très sensibles au bruit. Une étape de lissage préalable est nécessaire pour réduire l'effet des bruits de l'image. Ceci dit, l'effet non désiré de ce lissage est qu'il affecte la bonne localisation des coins. Généralement, les points détectés sont à l'intérieur du coin en question et non sur ce coin. Le problème de localisation de ces coins a été abordé par Deriche et Giraudon [DER90] qui ont mené une étude analytique sur plusieurs détecteurs en utilisant le modèle théorique d'un coin. Les auteurs constatent que la localisation du point d'intérêt à l'intérieur du coin est liée à l'écart type σ du filtre de lissage. Autrement dit, le comportement de ces détecteurs varie suivant l'échelle. Pour trouver la position exacte d'un coin, les auteurs calculent la réponse du détecteur pour deux échelles différentes. La droite reliant les deux points obtenus est considérée comme la bissectrice de l'angle du coin. La véritable position du coin se trouve sur cette droite à l'endroit où le Laplacien s'annule. Les auteurs précisent que pour des coins dont l'angle est inférieur à $\pi/4$ et pour σ grand, la méthode n'est plus sûre.

Harris [HAR88] propose une amélioration de l'approche de Moravec. Le détecteur de Harris se base sur le calcul d'une fonction locale d'auto-corrélation du signal. Cette fonction mesure la variation locale du signal (intensité) sur une fenêtre. Harris utilise la matrice M suivante (équation 4.3) liée à la fonction d'auto-corrélation du signal pour extraire ces points d'intérêt :

$$M = \exp \frac{x^2+y^2}{2\sigma^2} \otimes \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad (4.3)$$

Ce détecteur sera détaillé dans le paragraphe (§ 4.2.1.2.2).

Förstner [FOR86] propose une méthode de détection de points d'intérêt basée sur une analyse locale du gradient en un point de l'image pour en extraire des caractéristiques bas niveau. En travaillant en précision sub-pixel, il utilise un modèle de point d'intérêt construit dans l'espace objet au lieu de l'espace image. Ses points d'intérêt sont détectés par projection du modèle de point d'intérêt dans le domaine image. Un filtre non-linéaire est utilisé pour extraire les points d'intérêt. Ainsi, la localisation des points d'intérêt est insensible à l'angle des coins en 'L'.

Rohr [ROH92¹, ROH92²] utilise plusieurs modèles de points caractéristiques formés à la base de coins. Ceci a pour inconvénient que l'on se place dans un cas très restrictif quant au type de points à détecter. Un autre inconvénient de l'approche de Rohr est que sa méthode est relativement coûteuse en temps de calcul. Blaszkza et Deriche [BLA94] améliorent l'approche de Rohr en utilisant les mêmes modèles de primitives que Rohr sauf qu'au lieu de modéliser le flou de l'image par un filtre

Gaussien, les auteurs utilisent un filtre exponentiel. Les auteurs améliorent la phase d'initialisation du processus de minimisation lors de la recherche des points d'intérêt correspondant aux modèles théoriques.

Heitger et Al [HEI92] introduisent une approche originale basée sur la simulation de cellules de systèmes visuels biologiques. Les auteurs combinent des filtres symétriques pairs et impairs de même direction pour une mesure de l'énergie locale pour la détection de caractéristiques 1D de l'image. Une pseudo dérivée de cette énergie (orientée) est calculée pour chaque direction pour détecter les caractéristiques 2D de l'image. Leur méthode est basée sur des hypothèses neurophysiologiques.

Nous allons détailler dans les sections suivantes les détecteurs de points d'intérêt les plus utilisés en traitement d'images à niveaux de gris, ainsi qu'une extension au cas des images couleur:

- détecteur de Kitchen Rosenfeld,
- détecteur de Harris Stephens,
- détecteur de Harris Précis,
- détecteur de Harris précis couleur [GOU00].

Dans ce qui suit nous n'allons pas évaluer tous ces détecteurs du fait que de nombreuses études comparatives concernant ces détecteurs ont été effectuées, nous pouvons citer les travaux de Schmid [SCH96] et ceux de Gouet [GOU00]. Nous nous contentons de décrire le principe de chaque méthode.

4.2.1.2.1 Détecteur de Kitchen Rosenfeld

Le détecteur de Kitchen et Rosenfeld est basé sur la courbure des contours et sur l'amplitude du gradient pour définir une mesure K (équation 4.4) liée au produit de cette courbure par l'amplitude du gradient. Les auteurs utilisent cette mesure pour la recherche des maxima de courbure des isophotes du signal.

$$K = \frac{I_{xx}I_y^2 + I_{yy}I_x^2 - 2I_{xy}I_xI_y}{I_x^2 + I_y^2} \quad (4.4)$$

I_x, I_y, I_{xx}, I_{xy} et I_{yy} sont les dérivées premières et secondes de l'image par rapport à x et à y

Ce détecteur procède de la façon suivante :

- détection de contour,
- calcul de la mesure K,
- calcul de l'orientation des isophotes donnée par :

$$\theta = \arctan\left(\frac{-I_x}{I_y}\right) \quad (4.5)$$

- recherche des maxima locaux de la mesure K dans la direction des isophotes,
- rehaussement des coins extraits en multipliant l'image des maxima avec l'image des contours.
- seuillage pour extraire les points d'intérêt.

4.2.1.2.2 Détecteur Harris Stephens

Ce détecteur [HAR88], connu aussi sous le nom de détecteur de Plessey est l'un des détecteurs les plus utilisés. Il est basé sur le calcul d'une matrice (équation 4.3) liée à la fonction d'auto-correlation du signal calculée sur une fenêtre. Celle-ci est calculée à partir des dérivées de premier ordre de l'image.

4.2.1.2.3 Détecteur Harris Précis

Pour une meilleure stabilité du détecteur de Harris Stephens, une version dite *précise* [HAR88] utilise pour le calcul des dérivées de l'image une fonction de lissage gaussien. La matrice M de Harris s'écrit alors :

$$M = G(\tilde{\sigma}) \otimes \begin{bmatrix} I_x^2(\sigma) & I_x(\sigma)I_y(\sigma) \\ I_x(\sigma)I_y(\sigma) & I_y^2(\sigma) \end{bmatrix} \quad (4.6)$$

$G(\tilde{\sigma})$ est un filtre de lissage gaussien de taille $\tilde{\sigma}$ (écart type de la gaussienne 2D). σ est l'écart type de l'opérateur de dérivée gaussienne utilisé pour calculer les dérivées de l'image.

Un point d'intérêt est un point pour lequel la courbure principale de la fonction d'auto-correlation du signal est maximale. Les valeurs propres de la matrice M représentent ces courbures principales. Harris propose une mesure R (équation 4.7) basée sur le calcul du déterminant et de la trace de M lui évitant de calculer les valeurs propres de celle-ci.

$$R = \det(M) - K.trace^2(M) \quad (4.7)$$

K est une constante prise généralement égale à 0.04. Cette constante représente le taux de combinaison de l'information de contour donnée par la trace de M et de l'information d'angularité donnée par le déterminant de M . Un point d'intérêt correspond à un point pour lequel la mesure R est relativement grande (en plus d'être positive).

4.2.1.2.4 Détecteur Harris Précis Couleur

Les détecteurs cités précédemment travaillent tous en niveau de gris. Aucun n'exploite l'information photométrique qui nous intéresse particulièrement. Dans cette section nous nous intéressons au détecteur de Harris Précis Couleur développé

par Gouet et al [GOU00, MON98] qui travaille dans l'espace couleur RVB. La matrice M devient :

$$M = G(\tilde{\sigma}) \otimes \begin{bmatrix} R_x^2(\sigma) + G_x^2(\sigma) + B_x^2(\sigma) & R_x R_y(\sigma) + G_x G_y(\sigma) + B_x B_y(\sigma) \\ R_x R_y(\sigma) + G_x G_y(\sigma) + B_x B_y(\sigma) & R_y^2(\sigma) + G_y^2(\sigma) + B_y^2(\sigma) \end{bmatrix} \quad (4.8)$$

L'expression de R reste inchangée comparée à Harris précis. Et la recherche des points d'intérêt suit le même principe.

Gouet a adapté le détecteur de Harris précis et le détecteur de coins de Kitchen et Rosenfeld, qui travaillent normalement en niveaux de gris, pour qu'ils puissent tenir compte de l'information couleur. L'auteur a fait dans [GOU00] une étude comparative de ces deux nouvelles variantes couleur et du détecteur Harris Précis en niveau de gris. Cette étude est basée sur les critères de répétabilité et de localisation des points d'intérêt. La conclusion de Gouet est que les deux nouvelles variantes donnent de meilleures performances pour des transformations dans l'image telle que la rotation, le changement de point de vue, le changement d'échelle ainsi que la perturbation par un bruit. Par contre pour un changement affine de luminosité, le détecteur en niveau de gris obtient de meilleurs résultats que les deux autres. L'auteur justifie ces derniers résultats (concernant le changement affine de luminosité) "par le fait que le cadre de l'évaluation est simplifié en passant aux niveaux de gris".

4.2.1.2.5 Comparaison des différents détecteurs

Une étude comparative des détecteurs suivants, en niveau de gris, a été menée par Schmid [SCH96] :

- Harris Précis [BAU96],
- Heitger [HEI92],
- Forstner [FOR94],
- Horaud [HOR90],
- Cottier [COT94],

Dans cette étude, l'auteur compare les performances de ces différents détecteurs du point de vue de leur répétabilité (la répétabilité vérifie si la détection d'un point se répète correctement d'une image à l'autre, c'est une mesure visuelle). Elle démontre que les méthodes basées sur les contours sont les moins performantes. Schmid conclut que le détecteur de Harris Précis donne les meilleurs résultats, comparé aux quatre autres détecteurs. Ceci concerne les transformations de l'image liées à la rotation, au changement d'échelle, au changement de points de vue, au changement de luminosité et au bruit lié au capteur.

Gouet [GOU00] a comparé également les performances des détecteurs basés signal (Harris/Harris Précis, Harris Précis Couleur et Kitchen Rosenfeld) et des détecteurs basés contours (utilisation du Gradient mono-spectral, gradient de Di-

Zenzo). Sa conclusion rejoint celles citées plus haut concernant le fait que les détecteurs basés signal donnent les résultats de détection les plus précis et les plus stables comparé aux détecteurs basés contours.

4.2.2 Appariement d'images 2D

Nous allons à présent nous intéresser aux méthodes d'appariement d'images 2D. Nous donnons un bref état de l'art des méthodes d'appariement les plus utilisées en traitement d'images dans la section suivante.

Dans cet état de l'art nous présenterons brièvement les principales approches de mise en correspondance d'images 2D dont la classification peut être faite selon différents critères. L'un des premiers états de l'art sur le sujet est celui de Gottesfeld Brown [GOT92] où les méthodes sont classées en quatre catégories principales:

- méthodes basées sur la corrélation,
- méthodes basées sur la transformée de Fourier,
- méthodes basées sur la correspondance de points,
- méthodes basées sur un modèle élastique (Elastic Model-Based Matching).

Gouet [GOU00] propose une autre classification des méthodes de mise en correspondance d'image qui nous semble assez complète, et que nous avons adoptée dans l'état de l'art. Nous présenterons donc les techniques qui sont basées sur une comparaison des images utilisant le signal, les histogrammes ou les invariants différentiels. Ce qu'il faut retenir concernant la classification utilisée dans [GOU00] c'est que les méthodes de mise en correspondance d'images 2D sont regroupées en deux principales catégories : les méthodes d'appariement épars et les méthodes d'appariement dense. L'appariement dense signifie que l'appariement concerne l'ensemble des pixels des images traitées contrairement à un appariement épars qui est appliqué à certains points des images traitées (par exemple, les points d'intérêt).

4.2.2.1 Méthodes basées sur le signal

Ces méthodes utilisent directement le signal de l'image. Elles utilisent soit des fenêtres de taille définie, soit l'image entière. La simplicité de ces méthodes fait leur principal défaut. En effet, l'utilisation d'une fenêtre rectangulaire limite la robustesse de ces méthodes à la mise en correspondance entre images ayant subi des simples translations (pas de rotations). Une solution efficace pour pallier ce problème est d'utiliser des fenêtres circulaires. Ceci dit, des transformations d'image plus complexes telles que celles liées aux changements d'échelle, aux changements de point de vue ou aux occultations posent des problèmes à ce type de méthode. Par ailleurs, le contenu informatif des fenêtres n'est pas, tout le temps, très riche. Ainsi, les mauvaises correspondances peuvent être dues au fait qu'une fenêtre appliquée sur une zone non significative (zone lisse) de la première image soit mise en correspondance avec une fenêtre de la deuxième image avec un contenu non significatif, mais qui ne correspondent pas au même point physique. La mesure de similarité entre deux images ou entre deux sous-images peut se faire de plusieurs manières. Nous en présentons quatre dans la suite de ce paragraphe.

4.2.2.1.1 Corrélation du signal

Cette méthode utilise une mesure de la corrélation sur une fenêtre autour du point à mettre en correspondance. C'est la plus ancienne méthode de mise en correspondance d'images 2D. La mesure de la corrélation CC (*cross-correlation*) classique est donnée par l'équation suivante :

$$CC(u, v) = \frac{\sum_x \sum_y w(x, y) I(x - u, y - v)}{(\sum_x \sum_y I^2(x - u, y - v))^{1/2}} \quad (4.9)$$

w est la fenêtre utilisée pour le calcul de la corrélation au point (x, y) considéré, I est l'intensité de l'image à traiter. L'équation précédente correspond à l'expression basique de la corrélation. Des variantes plus robustes ont été proposées. Nous citons les plus fréquemment rencontrées dans la littérature :

- SSD : Sum of Squared Differences
- NSSD : Normalized SSD,
- NCC : Normalized Cross Correlation,
- ZNSSD : Zero mean NSSD
- ZNCC : Zero mean NCC.

Le lecteur peut se référer à [GOU00] pour plus de détails sur ces différentes formules de corrélation.

Le principe de base de la mise en correspondance par corrélation consiste à rechercher pour un point d'une image son correspondant dans l'autre image pour lequel la mesure de corrélation est maximale. Cette mesure est calculée sur des images en précision pixel. Ceci dit, il est possible de passer en précision sub-pixel par interpolation des valeurs de la mesure de la corrélation. La corrélation donne de très bons résultats pour des images différant seulement par une simple translation. Elle peut, cependant, tolérer une légère rotation. Un grand nombre de travaux ont été menés dans le but d'améliorer la robustesse de la méthode de corrélation à des transformations image (géométriques ou affines) plus complexes qu'une simple translation. Jung et Lacroix [JUN01] ont développé un algorithme robuste pour l'appariement de points d'intérêt en niveaux de gris basé sur la corrélation du signal.

4.2.2.1.2 Corrélation de phase

Un inconvénient majeur de la méthode par corrélation est son coût en temps de calcul. Les approches basées sur la transformée de Fourier sont plus rapides que la corrélation du signal brut. Elles exploitent la représentation de l'image dans le domaine spectral. Ces méthodes, nommées aussi *méthodes par corrélation de phase* utilisent le théorème du décalage de Fourier. Nous donnons dans ce qui suit le principe de la méthode. $f_1(x, y)$ et $f_2(x, y)$ sont deux signaux bidimensionnels. En vision, ces signaux représentent les deux images à mettre en correspondance. Ces signaux sont liés par (ou différent de) la translation (X, Y) . Soient $F_1(w_x, w_y)$ et

$F_2(w_x, w_y)$ leurs transformées de Fourier respectives. En considérant le théorème de décalage de Fourier, le spectre de puissance croisé des deux signaux s'écrit :

$$\frac{F_1(w_x, w_y)\overline{F_2(w_x, w_y)}}{|F_1(w_x, w_y)\overline{F_2(w_x, w_y)}|} = \exp(j(Xw_x + Yw_y)) \quad (4.10)$$

où \overline{F} représente le conjugué de la transformée F .

L'inverse du spectre de puissance croisé de l'équation précédente est donnée par :

$$F^{-1}\left(\frac{F_1(w_x, w_y)\overline{F_2(w_x, w_y)}}{|F_1(w_x, w_y)\overline{F_2(w_x, w_y)}|}\right) = F^{-1}(\exp(j(Xw_x + Yw_y))) \quad (4.11)$$

$F^{-1}(\exp(j(Xw_x + Yw_y))) = \delta(X, Y)$ est une impulsion de Dirac centrée en (X, Y) .

Ainsi, le calcul de l'inverse du spectre de puissance croisé nous permet, à la fois, de vérifier si deux signaux, liés par une translation, sont identiques et d'identifier les paramètres de cette translation.

Les méthodes basées sur la transformée de Fourier sont plus robustes aux bruits. Elles sont aussi moins coûteuses en temps de calcul que les méthodes par corrélation du signal. Néanmoins, comme pour ces méthodes, leur robustesse est démontrée pour des déplacements en translation et/ou pour des rotations légères. De nombreuses améliorations de la technique de base ont vu le jour en vue de rendre robuste la méthode à des transformations plus complexes.

4.2.2.1.3 Mesure de l'information mutuelle

Ces méthodes, plus récentes, sont très utilisées dans le domaine du recalage multimodal pour des applications telles que l'imagerie médicale. La notion d'information mutuelle est une notion de la théorie de l'information. C'est une mesure de la dépendance statistique entre deux ensembles de données. L'expression de l'information mutuelle entre deux variables aléatoires x et y est donnée par l'équation suivante :

$$MI(x, y) = H(y) - H(y|x) = H(x) + H(y) - H(x, y) \quad (4.12)$$

$H(x)$ est l'entropie de la variable aléatoire x . La recherche des correspondances est

basée sur la maximisation de l'expression de MI . Pour plus de détails sur l'utilisation de l'information mutuelle le lecteur peut se référer à [VIO97].

4.2.2.1.4 Distance de Hausdorff

La distance de Hausdorff est une mesure de similarité qui permet de comparer deux ensembles de points. Pour deux ensembles de points A et B la distance de Hausdorff h est donnée par l'expression suivante :

$$H(A, B) = \max(h(A, B), h(B, A)) \quad (4.13)$$

où $h(A, B)$ correspond donc à la distance maximum d'un point de l'ensemble A à son plus proche voisin dans l'ensemble B :

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\|$$

Il faut noter que cette distance $H(A, B)$ devient très importante lorsqu'un seul point de A est significativement éloigné de B . Elle est donc très sensible au bruit, et n'est donc pas très représentative de la distance entre deux images. L'expression suivante permet de calculer une distance partielle sur les K meilleures paires de points appariés entre A et B . Cette distance partielle est la plus utilisée.

$$h_K(A, B) = K_{a \in A}^{me} \min_{b \in B} \|a - b\| \quad (4.14)$$

Cette expression nous conduit à la distance de Hausdorff partielle $H_K(A, B)$ des K points de A qui s'apparient le mieux avec des points de l'ensemble B .

Le but est de trouver un nombre maximum de paires de points appariés pour lesquelles la distance partielle de Hausdorff est inférieure à un certain seuil δ :

$$H_K(A, B) \leq \delta \quad (4.15)$$

Soit $K_\delta(A, B)$ ce nombre maximum pour lequel l'équation 3.15 est vérifiée.

$F_\delta(A, B) = \frac{K_\delta(A, B)}{|A|}$ est la proportion des points de A qui sont appariés aux points de

B et pour lesquels la distance de Hausdorff vérifie l'équation précédente. Cette proportion est appelée *fraction de Hausdorff*. Le principe des mises en correspondance basées sur la distance de Hausdorff est de maximiser cette *fraction* pour qu'elle atteigne un certain seuil T , pour une certaine erreur δ de la distance de Hausdorff. Un très grand nombre de travaux [HUT93] ont été menés sur des applications en recalage d'image, en reconnaissance/localisation d'objets ou encore en indexation d'image. Cette distance est très robuste aux bruits et aux faux appariements dûs notamment aux occultations.

4.2.2.2 Méthodes basées sur les histogrammes

L'histogramme des couleurs dans une image peut être considéré comme une approximation de la distribution des couleurs [SWA91]. On entend par couleur aussi bien le vecteur couleur à 3 composantes associé à chaque point, que toute autre représentation de l'aspect photométrique du point (niveau de gris, teinte). L'histogramme est une représentation globale qui ne contient pas d'information sur la position des points.

La mise en correspondance des images est basée sur la comparaison de leurs histogrammes. Les histogrammes ont l'avantage d'être invariants aux transformations géométriques de l'image. Il existe même certaines variantes de ces histogrammes les rendant invariants aux changements d'illumination [FUN95]. Cependant, l'inconvénient de telles approches réside dans le fait qu'il est difficile de définir un critère de comparaison d'histogrammes. Par ailleurs, du point de vue occupation mémoire, il faut noter que les histogrammes contiennent souvent une masse de données relativement importante.

Swain et Ballard [SWA91] proposent une méthode de comparaison basée sur l'intersection d'histogrammes. Pour le cas de l'indexation d'images, l'équation suivante donne l'expression de l'intersection de l'histogramme de l'image du modèle avec celui d'une image de la base. Ces histogrammes sont notés respectivement M et H .

$$Int(H, M) = \frac{\sum_{i=1}^n \min(h_{ci}, m_{ci})}{\sum_{i=1}^n m_{ci}} \quad (4.13)$$

où n représente le nombre de couleurs dans un espace discrétisé, h_{ci} et m_{ci} sont les valeurs correspondantes des deux histogrammes pour la couleur i .

Une estimation limitée aux couleurs les plus représentatives rend cette approche moins coûteuse en temps de calcul.

Une approche plus classique pour comparer les histogrammes de deux images est basée sur le calcul de distances entre histogrammes. Différentes expressions de la distance ont été utilisées, comme par exemple : la distance L_1 [Huang 99], la distance euclidienne L_2 pondérée ou non [SCH97]. Ainsi, la distance L_1 est donnée par :

$$d_{L_1}(H, M) = \sum_{i=1}^n |h_{ci} - m_{ci}| \quad (4.14)$$

Le lecteur trouvera des études comparatives sur ces différentes approches dans [BRU99] et dans [JEO01, JEO04].

4.2.2.3 Méthodes basées sur les invariants différentiels

Dans ce type d'approche, il s'agit de procéder par une étape de recherche de points d'intérêt suivie d'une étape de caractérisation de ces points. A chaque point est alors associé un descripteur basé sur l'information géométrique et/ou sur l'information photométrique du voisinage du point. Ces descripteurs sont généralement des vecteurs d'attributs indépendants si possible des transformations géométriques appliquées à l'image. La mise en correspondance, dans ce cas, revient donc à comparer les vecteurs d'invariants d'un point d'intérêt d'une image avec ceux des points de la deuxième image. Un critère de comparaison de base serait la distance euclidienne entre ces vecteurs. Mais une telle distance ne peut prendre en considération le fait que les composantes des vecteurs d'invariants n'ont pas le même ordre de grandeur. Dans son état de l'art, Gouet [GOU00] passe en revue les deux principales approches pour comparer deux vecteurs d'invariants différentiels :

- la corrélation entre les vecteurs,
- la distance de Mahalanobis.

La première consiste à calculer un score de corrélation entre les vecteurs d'attributs à comparer. La distance de Mahalanobis est une grandeur statistique. C'est une distance entre deux variables aléatoires qui tient compte de la matrice de covariance de ces variables. Cette distance est donc très robuste aux bruits. Malheureusement, il est très difficile de calculer la matrice de covariance des vecteurs d'invariants à comparer. Le seul moyen d'évaluer cette matrice est de procéder de façon empirique. Ceci nécessite d'avoir plusieurs observations du vecteur d'invariants du même point d'intérêt apparaissant sur plusieurs images correspondant à différentes conditions de prise de vue, de luminosité, et de bruit. Une condition très importante pour que cette estimation empirique puisse donner des résultats réalistes est que le point d'intérêt choisi soit localisé sans erreur sur chaque image. Cette opération est effectuée pour plusieurs points d'intérêt avec différentes séquences d'images. Il en résulte autant de matrices de covariance que de points suivis. La matrice de covariance finale est la moyenne de toutes les matrices estimées.

4.2.3 Elimination des faux appariements

Les techniques de mise en correspondance citées dans ce chapitre ne peuvent être robustes à toutes les transformations (photométriques ou géométriques) que peuvent subir les images. La présence de bruits est aussi l'une des raisons qui peut induire en erreur ces méthodes. Ceci implique la présence de mauvaises correspondances ou de faux appariements. Ces faux appariements induisent en erreur à leur tour l'estimation de la transformation entre les deux images à recalculer. Il est nécessaire d'éliminer ces faux appariements pour ne garder que les appariements corrects. Suivant la méthode de mise en correspondance utilisée, il existe plusieurs approches pour palier ce problème.

Pour les méthodes par corrélation, il est par exemple possible de ne considérer que les paires ayant un score de corrélation important. Les autres sont

rejetées. Pour les méthodes basées sur les invariants différentiels il est, aussi, possible d'appliquer un seuil de distance entre les vecteurs d'invariants. Seules les paires pour lesquels la distance entre vecteurs d'invariants est inférieure à ce seuil sont prises en compte.

Il est aussi possible d'imposer des contraintes sur les appariements telles que des contraintes géométriques ou des contraintes de voisinage et d'unicité. La contrainte d'unicité signifie qu'un point apparié ne peut être impliqué que dans un seul appariement. Autrement dit, cette contrainte élimine les multiples appariements. Pour s'assurer qu'un point n'est considéré qu'une seule fois il suffit d'appliquer un appariement croisé. Celui-ci consiste à effectuer une première mise en correspondance de chaque point de la première image avec tous les points de la deuxième. Cette étape est suivie par une mise en correspondance de chaque point de la deuxième image avec tous les points de la première. Seules les paires qui ont été sélectionnées dans les deux étapes sont retenues.

Parmi les contraintes géométriques rencontrées dans la littérature, nous pouvons citer la contrainte épipolaire en stéréoscopie et la contrainte de voisinage. Cette contrainte relie les deux caméras du système stéréoscopique. On suppose un point de la scène visible par les deux caméras. A partir d'une image fournie par l'une des caméras, la géométrie épipolaire lui fait correspondre une ligne 1D dans l'image de la deuxième caméra. Ceci ramène le problème de recherche du point correspondant d'un problème 2D à un problème 1D. Ainsi la *dimensionnalité* du problème est réduite. Cette contrainte est très intéressante lorsque la géométrie épipolaire est connue. Ceci n'est pas le cas dans notre application.

La contrainte de voisinage se base sur des relations géométriques entre le point à appairier et les points d'intérêt se situant dans son voisinage. Cette relation ou caractéristique peut être, par exemple, l'angle existant entre deux points d'intérêt voisins du point en question (par exemple les deux plus proches voisins), le sommet de cet angle étant le point à appairier. Cet angle est supposé se conserver d'une image à l'autre pour un bon appariement [SCH96]. Une paire de points appariés est conservée si l'angle formé par le point de la première image avec ses deux voisins correspond approximativement à celui formé par son correspondant et ses voisins dans la deuxième image. La contrainte de conservation des angles citée ici et utilisée par Schmid [SCH96] tolère un écart entre les angles à comparer de 20 degrés.

Il est également possible d'envisager d'imposer que pour qu'un point soit apparié il faut qu'un certain nombre de points d'intérêt, du voisinage exclusif de celui-ci, soient appariés correctement. Cette dernière contrainte est assez sévère, car elle ne tolère aucune erreur de détection des points d'intérêt. Il faut donc prendre en considération les erreurs de détection dues essentiellement aux imperfections du détecteur utilisé et au bruit de l'image. Ainsi, au lieu d'imposer que tous les points d'un certain voisinage des points à appairier soient mis en correspondance on peut se contenter seulement d'un appariement partiel (par exemple 50%) des points du voisinage.

Une approche classique d'élimination de faux appariements ou d'appariements ambigus est la *méthode de relaxation*. La relaxation est un algorithme itératif qui minimise une fonction d'énergie globale appelée critère de relaxation. L'algorithme de relaxation a pour but d'éliminer les mauvais appariements de façon itérative en partant d'un ensemble d'appariements initial. Le critère de relaxation est lié à une contrainte de voisinage de la caractéristique mise en correspondance. Cette contrainte est propagée dans l'ensemble des appariements de façon itérative. Gouet [GOU00] utilise la relaxation pour éliminer les appariements ambigus. Elle parle de contrainte semi-locale. Plus de détails sur la méthode de relaxation peuvent être trouvés dans [ZHA94⁺].

4.2.4 Conclusion

L'état de l'art nous a permis, entre autres, de passer en revue certaines méthodes de mise en correspondance d'images 2D. Les méthodes qui nous semblent intéressantes sont les méthodes éparées. Ce choix est basé, d'une part, sur leur simplicité de mise en œuvre et d'autre part sur le fait qu'elle utilise des points d'intérêt au lieu de la totalité des données qui pourraient ralentir la mise en correspondance pour des images de grandes résolutions.

Nous avons vu dans l'état de l'art précédent différents détecteurs de points d'intérêt et nous avons particulièrement détaillé le détecteur de coins de Kitchen et Rosenfeld, le détecteur de Harris Stephens ainsi que le détecteur de Harris Précis Couleur. Dans la deuxième partie de cet état de l'art, nous avons donné un aperçu de certaines méthodes de mise en correspondance d'images 2D.

Plusieurs études comparatives concernant ces détecteurs et les méthodes de mise en correspondance ont été effectuées. Nous nous basons sur l'une de ces études pour le choix du détecteur ainsi que de la méthode de mise en correspondance à utiliser. L'étude comparative menée par Gouet [GOU00] nous semble assez complète. Nous optons pour une méthode de mise en correspondance basée sur les invariants différentiels. L'appariement est effectué sur des points d'intérêt pour lesquels sont calculés les vecteurs d'invariants. L'appariement est basé sur la distance euclidienne normalisée entre ces vecteurs.

Notre choix de détecteur de points d'intérêt s'est porté sur le détecteur de Harris Précis couleur développé par Gouet [GOU00]. L'auteur, démontre, dans son étude comparative que celui-ci est en effet le plus performant des détecteurs cités dans l'état de l'art. Il est relativement stable du fait qu'il n'est basé que sur des dérivées du premier ordre de l'image. Aussi, Gouet [GOU00] conclut que le détecteur de Harris Précis Couleur possède une *excellente* répétitivité comparé aux autres détecteurs évalués.

4.3 Extraction de points d'intérêt

Dans ce qui suit, nous allons appliquer le détecteur de points d'intérêt Harris Précis Couleur [GOU00] sur des images réelles. Les images choisies pour effectuer cette validation expérimentale ainsi que celles des chapitres suivants, représentent trois cas de figure qui soulèvent des problèmes différents :

- Les images du tableau "Calvet" ont été obtenues avec un appareil photo numérique de résolution 3 méga pixels, la taille de ses images est de 640x480 pixels. Elles ont été réalisées avec un éclairage naturel. Nous n'avons pas effectué de numérisation de cet objet car la peinture ne présentait pas de relief 3D significatif.
- La figurine "Ourson" a été numérisée avec un scanner 3D/couleur MINOLTA. Les différents points de vue ont été obtenus en la plaçant sur un plateau tournant dont la rotation est mesurée par un codeur. Ces données dont la partie 3D a déjà été utilisée dans le chapitre 3, sont fournies par l'Ohio State University sur le site [OSU]. La taille de ces images couleurs est de 200x200 pixels.
- Les images 3D/couleur de la boîte cylindrique "Ivoire" ont été réalisées au LIRMM avec le capteur de numérisation TRITOS développé par la société BREUKMANN [BRE2]. Il s'agit d'un exemple où la prise en compte des données 3D seules ne permet pas le recalage exact. Nous allons utiliser l'extraction de points d'intérêt dans l'image couleur RGB fournie par la caméra intégrée au capteur, celle-ci fournit des images couleur de taille 1280x960 pixels (figures 4.5 et 4.6). Nous avons utilisé l'image couleur initiale codée sur 3x8bits, et non les composantes couleur des points 3D/couleur car, ici, l'information couleur codée sur 3x5bits était trop bruitée.



Figure 4.1 - Points d'intérêt détectés par le détecteur de Harris Précis Couleur sur l'image Calvet1.

La détection des points d'intérêt se fait donc par le calcul de la matrice M de l'équation 4.8. Pour chaque pixel de l'image nous calculons le coefficient R donné par l'équation 4.7. Un point d'intérêt correspond à un point pour lequel R est positif et suffisamment grand. Gouet applique un seuil pour R au dessous duquel le point n'est pas considéré comme point d'intérêt.

La détermination de ce seuil n'est pas facile car elle dépend du type d'image. Nous évitons ce choix du seuil de la manière suivante. Les valeurs positives classées par ordre décroissant de R sont stockées dans un tableau noté R_p . Ensuite nous ne considérons que les N_{pi} points d'intérêt correspondant aux N_{pi} premières valeurs R_p .



Figure 4.2 - Points d'intérêt détectés par le détecteur de Harris Précis Couleur sur l'image Calvet2.

4.3.1 Exemple 1: peinture à l'huile

Les figures 4.1 et 4.2 montrent les résultats du détecteur Harris Précis Couleur appliqué aux images Calvet1 et Calvet2. Ces images ont été acquises à l'aide d'un appareil photo numérique dans des conditions d'éclairage naturel (à l'extérieur). Ces deux paires de vues diffèrent d'une rotation d'environ trente degrés autour de l'axe optique. Dans cet exemple nous prenons $\sigma = 1$ pour une fenêtre de dérivation de taille égale à 7. La taille de la fenêtre de lissage est également égale à 7 pour $\tilde{\sigma} = 1$. Pour des soucis de visibilité nous n'affichons pas ici les points d'intérêt qui sont sur le voisinage immédiat d'un autre point d'intérêt. Ceci dit, lorsque nous effectuons la mise en correspondance des deux images nous considérons tous les points d'intérêt. Nous affichons dans la première image 61 points d'intérêt et nous en affichons 87 dans la seconde.

4.3.2 Exemple 2: Figurine 3D

Les figures 4.3 et 4.4 représentent les points d'intérêt détectés par le détecteur de Harris Précis couleur appliqué aux images 3D/couleur obtenues avec un scanner MINOLTA [OSU]. Il s'agit, cette fois, de deux images d'un objet 3D prises pour deux positions différentes de celui-ci. La deuxième image est obtenue après rotation de l'objet d'un angle d'environ 20 degrés autour de l'axe vertical (perpendiculaire à l'axe optique) par rapport à la première image. Il s'agit, donc, d'un changement de perspective. Nous pouvons constater, à l'oeil, une bonne répartition spatiale des points d'intérêt sur les deux images, ainsi qu'une répétitivité relativement satisfaisante de ces points. L'absence de points d'intérêt en dehors de la forme de l'objet (autour de celui ci) est due au fait que l'on ne traite que la zone de l'image correspondant à l'objet à numériser (les pixels correspondant aux points 3D/couleur fournis par le scanner 3D).

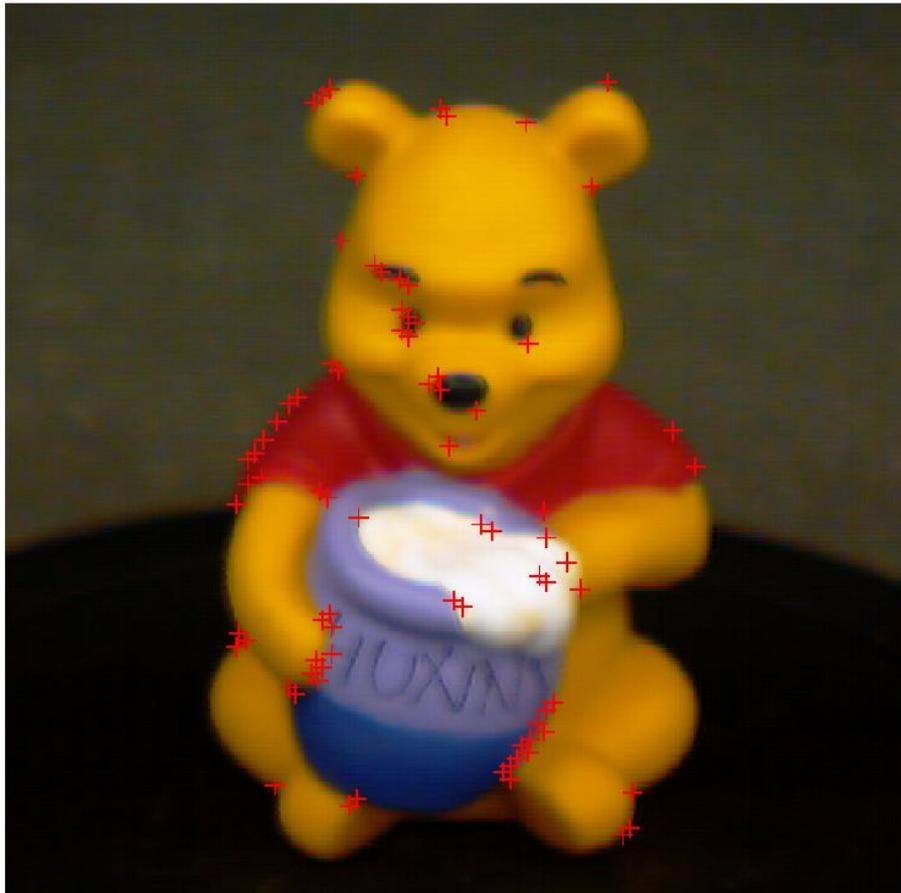


Figure 4.3 - Points d'intérêt détectés par le détecteur de Harris Précis Couleur sur l'image Ourson1.

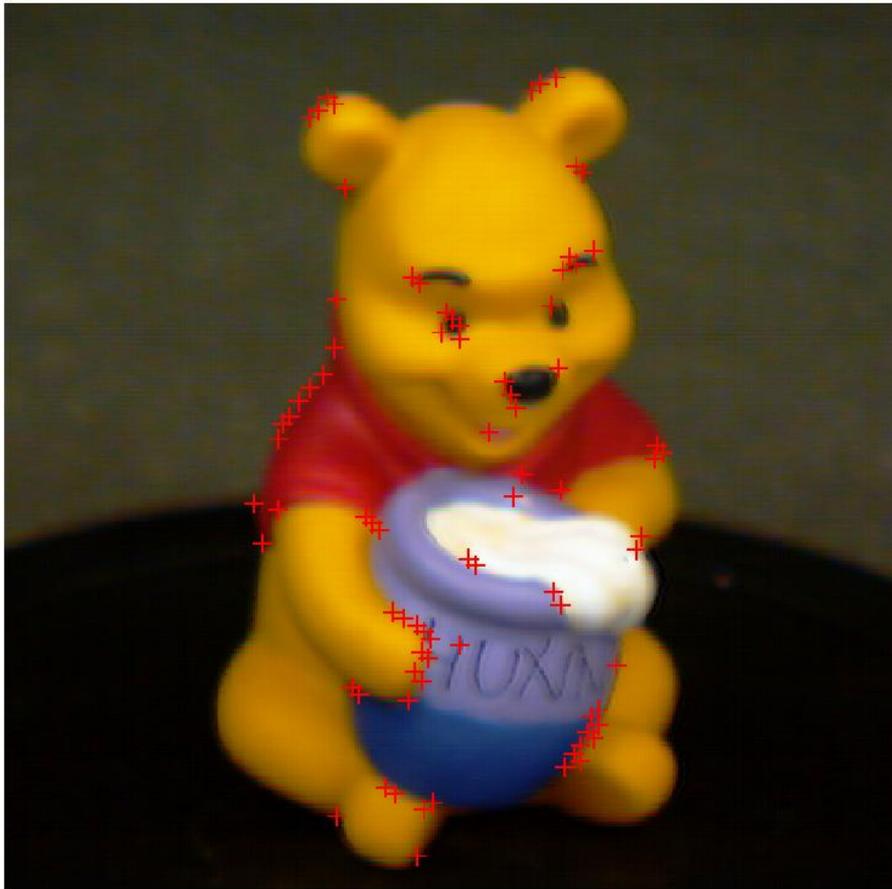


Figure 4.4 - Points d'intérêt détectés par le détecteur de Harris Précis Couleur sur l'image *Ourson2*.

4.3.3 Exemple 3: boîte cylindrique

Les figures 4.5 et 4.6 montrent deux images de l'objet *Ivoire* acquises par le capteur Breukmann. Ce capteur fournit des données 3D/couleur desquelles nous avons tiré les images 2D correspondant à chaque vue de l'objet. Nous ne travaillons que sur une partie de chaque image (figures 4.7 et 4.8) à cause de la quantité de données relativement imposante que contiennent ces images. Le problème de la quantité de données n'intervient pas au niveau du traitement 2D de l'image mais au niveau du recalage 3D/couleur. En effet, une image 3D/couleur du capteur Breukmann peut contenir jusqu'à un million de points 3D/couleur. Nous préférons ainsi tester les images qui correspondent aux données 3D/couleur à recalcer (pour lesquelles chaque pixel a un correspondant en 3D).



Figure 4.5 - *Image Ivoire1*



Figure 4.6 - *Image Ivoire2*

Les figures suivantes (figures 4.7 et 4.8) montrent le résultat de la détection de points d'intérêt par Harris Précis Couleur. Nous prenons $\sigma = 1$ pour une fenêtre de dérivation de taille égale à 7. $\tilde{\sigma} = 1$, la taille de la fenêtre de lissage est également égale à 7.



Figure 4.7 - Points d'intérêt détectés par le détecteur de Harris Précis Couleur sur une partie de l'image Ivoire1.



Figure 4.8 - Points d'intérêt détectés par le détecteur de Harris Précis Couleur sur une partie de l'image Ivoire2.

Nous pouvons constater sur ces images que les points d'intérêt sont relativement moins bien répartis en comparaison aux résultats précédents (cas des images "Calvet" et "Ourson"). Nous pouvons également remarquer une faible répétabilité de ces points. Ceci est dû, essentiellement, au fait que ces images sont considérablement bruitées. Lors de l'acquisition de ces images au sein du LIRMM,

aucune précaution particulière concernant l'éclairage n'a été prise, ce qui peut expliquer le changement considérable de luminosité entre ces deux images Ivoire.

4.4 Caractérisation et appariement des points d'intérêt

En conclusion de l'état de l'art précédent nous avons précisé notre choix concernant l'approche à utiliser pour la mise en correspondance des images 2D pour l'initialisation des algorithmes de recalage 3D/couleur. Nous nous sommes inspiré des travaux de Gouet [GOU00] pour la mise en correspondance de nos images. La méthode de mise en correspondance consiste à apparier les points d'intérêt que nous avons extrait par le détecteur de Harris Précis Couleur. Nous avons vu dans l'état de l'art que l'appariement des points d'intérêt nécessite la caractérisation locale de ces points. Cette caractérisation consiste à calculer pour chaque point d'intérêt un vecteur d'invariants. L'appariement est basé sur la comparaison de ces vecteurs. L'étape qui suit, donc, la détection des points d'intérêt concerne la caractérisation de ces points.

La caractérisation des points d'intérêt est une caractérisation locale (elle concerne le voisinage de ces points). Cette caractérisation doit être invariante aux transformations de l'image telles que les transformations affines, le changement d'illumination et le bruit de l'image.

Dans cette étude nous exploitons l'information couleur lorsque celle-ci est disponible. Ainsi au lieu d'utiliser une caractérisation en niveau de gris telle que le jet local ou les invariants de Hilbert nous préférons une caractérisation couleur [GOU00]. L'un des avantages de cette caractérisation est qu'elle exploite l'information inter canal et que le calcul des vecteurs d'invariants ne nécessite que des dérivées d'ordre un pour un vecteur de dimension 8, alors qu'il est nécessaire d'aller à l'ordre 2 voire à l'ordre 3 pour un vecteur de la même dimension lorsque seuls les niveaux de gris sont utilisés (exemple : le calcul d'invariants de Hilbert en niveaux de gris nécessite des dérivées de l'image de l'ordre 3 pour un vecteur de dimension 9). La dimension du vecteur d'invariant est importante. En effet, plus grand est le vecteur d'attributs plus grand est son pouvoir discriminant.

4.4.1 Vecteur d'invariants

Les vecteurs d'invariants que nous allons utiliser pour la phase d'appariement des points d'intérêt couleur ont été introduits par Gouet [GOU00]. L'expression de ce vecteur est donnée par l'équation 4.15. L'auteur a évalué l'invariance de la caractérisation locale des points d'intérêt par ce vecteur pour des transformations image telles que - la rotation image, - le changements d'illumination (interne¹, externe² et complexe³), ainsi qu'en présence de bruit image. Elle montre que la caractérisation locale utilisant le vecteur de l'équation 4.15 reste invariante aux rotations image.

¹ Changement au niveau des composantes internes de la source lumineuse

² Changement du par exemple à un déplacement de la source lumineuse

³ Changement interne et externe en même temps

Cependant, il est nécessaire de procéder à une normalisation locale de la couleur pour que cette caractérisation reste robuste aux changements de luminosité. Nous avons adopté cette normalisation des couleurs pour améliorer la robustesse des invariants couleur que nous utilisons. La méthode de normalisation est décrite dans le paragraphe suivant.

$$\begin{pmatrix} R \\ \|\nabla R\|^2 \\ V \\ \|\nabla V\|^2 \\ B \\ \|\nabla B\|^2 \\ \nabla R \cdot \nabla V \\ \nabla R \cdot \nabla B \end{pmatrix} \quad (4.15)$$

Où :

∇R , ∇V et ∇B sont, respectivement, les gradients des plans couleur R, V et B,

$\|\nabla R\|^2$, $\|\nabla V\|^2$ et $\|\nabla B\|^2$ sont, respectivement, les carrés des modules de ces gradients,

$\nabla R \cdot \nabla V$ et $\nabla R \cdot \nabla B$ sont les termes inter canaux.

4.4.2 Normalisation des couleurs

La normalisation des couleurs a pour objectif de réduire l'effet des changements de luminosité sur la caractérisation des points d'intérêt. L'intérêt de la normalisation des couleurs est de rendre robuste l'appariement des points d'intérêt. Nous verrons, dans ce qui suit, son impact sur la qualité des appariements.

Une normalisation basique serait de ramener les valeurs des composantes couleur dans l'intervalle $[0, 1]$. Dans [MON00] les couleurs sont normalisées de façon locale. La normalisation est faite dans une fenêtre circulaire dans le but de préserver les propriétés locales des pixels. Les résultats obtenus par l'auteur avec cette normalisation nous semblent intéressants. Nous adoptons cette méthode de normalisation dont l'expression suivante donne la valeur du niveau d'intensité normalisée propre à un plan couleur.

$$\overline{I_n(x, y)} = \frac{I_n(x, y) - med_{12}}{med_{34} - med_{14}} \quad (4.16)$$

med_{12} est la médiane des niveaux de gris du plan I_n (ou des pixels situés à l'intérieur de la fenêtre circulaire).

med_{14} et med_{34} sont respectivement la médiane du premier et du deuxième quartier de l'ensemble des niveaux de gris considérés (ceux du plan I_n si l'on considère toute l'image ou seulement ceux appartenant à la fenêtre circulaire).



Figure 4.9 - Exemple de normalisation appliquée à l'image *Ourson1* avec une fenêtre circulaire de taille rayon 11 pixels.

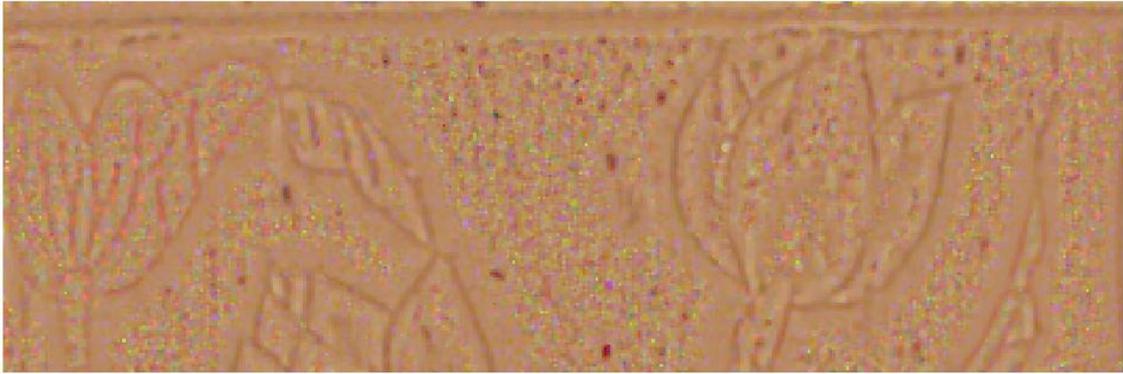


Figure 4.10 - Exemple de normalisation appliquée à l'image *Ivoire1* avec une fenêtre circulaire de rayon de taille 9 pixels.

Les figures ci-dessus (4.9 et 4.10) illustrent deux exemples de normalisation des couleurs appliquée aux images *Ourson2* et *Ivoire1*. Le rayon de la fenêtre circulaire de normalisation pour la première image est de 11, celui pour la deuxième image est de 9. Cette taille doit être adaptée en fonction de la différence de luminosité entre les deux images à mettre en correspondance. Plus cette différence est grande plus petite doit être la taille de la fenêtre de normalisation. Nous verrons dans les paragraphes suivant l'impact de la normalisation sur la qualité des appariements.

4.4.3 Appariement des points d'intérêt

La dernière phase de la mise en correspondance des images couleur est l'appariement des points d'intérêt basé sur une distance euclidienne (équation 4.17) entre les vecteurs d'invariants différentiels normalisés présentés dans le paragraphe 4.4.1. Nous n'utilisons pas la distance de Mahalanobis à cause des difficultés liées au calcul de la matrice de covariance des vecteurs d'invariants. Par ailleurs, nous n'imposons aucune contrainte géométrique (contrainte semi locale de voisinage) pour éliminer les faux appariements, car cette élimination sera faite sur les points 3D avec l'algorithme RANSAC dans la phase d'initialisation des algorithmes de recalage 3D (chapitre 5).

$$d_{inv} = \|V_{inv}^1 - V_{inv}^2\| \quad (4.17)$$

Où V_{inv}^1 et V_{inv}^2 sont les deux vecteurs d'invariants correspondant aux points d'intérêt à apparier.

La normalisation des vecteurs d'attributs est nécessaire du fait que les composantes de ces vecteurs ne varient pas toutes dans la même plage de grandeurs. Sans normalisation sur ces vecteurs la distance euclidienne ne peut pas traduire correctement le degré de ressemblance de deux vecteurs d'attributs pour deux points d'intérêt.

4.5 Validation expérimentale de la méthode d'appariement

Dans cette section, il s'agit de mettre en correspondance deux images 2D en appariant les points d'intérêt fournis par le détecteur de Harris Précis Couleur (paragraphes 4.3 et 4.4). Nous allons considérer pour notre étude les trois exemples d'images utilisées dans la section précédente.

Les principaux paramètres à fixer pour l'ensemble des processus de détection de points d'intérêt et d'appariement sont :

- les tailles des fenêtres de lissage utilisées pour la détection de points d'intérêt (détecteur de Harris),
- le nombre de points d'intérêt à retenir pour la phase d'appariement,
- la taille des fenêtres de lissage utilisées dans le calcul des vecteurs d'invariants des points d'intérêt,
- et la taille de la fenêtre de normalisation de la couleur,

Ainsi, nous devons faire varier ces différents paramètres séparément dans le but d'évaluer leur impact sur la qualité des appariements. Dans ce qui suit nous donnons les résultats concernant l'impact de la taille de la fenêtre de normalisation. Les trois autres paramètres ont été fixés aux valeurs pour lesquelles nous avons obtenu un nombre satisfaisant de points d'intérêt avec une bonne répartition spatiale ainsi qu'une meilleure qualité de l'appariement. Nous donnons ci-dessous les valeurs de ces paramètres qui restent les mêmes pour les expériences qui suivent, dans le cas contraire ces valeurs sont précisées :

- *détecteur de Harris* :

$\sigma = 1$, la taille de la fenêtre de dérivée gaussienne est de 7,

$\tilde{\sigma} = 1$, la taille de la fenêtre de lissage gaussien est de 7,

- *calcul des vecteurs d'invariants* :

$\sigma_{vi} = 2$, la taille de la fenêtre de calcul des dérivées gaussiennes est de 15,

- *nombre de points d'intérêt à conserver pour l'appariement* :

Dans la littérature plusieurs approches ont été utilisées pour choisir les points d'intérêt les plus significatifs pour l'étape de mise en correspondance. En effet, on ne peut garder tous les points pour lesquels la mesure R (équation 4.2) est positive. Nous avons vu qu'il est possible d'envisager de fixer un seuil pour la mesure R en dessus duquel un point d'intérêt est conservé. Ceci dit, pour éviter de définir un tel seuil, nous utilisons la méthode décrite dans le paragraphe 4.3 qui consiste à considérer les N_{PI} (300 dans ce cas) points d'intérêt correspondant aux N_{PI} plus grandes valeurs positives de R .



Figure 4.11 - Mise en correspondance de deux images Ivoire avec sans un changement linéaire de la luminosité (: les bons appariements (tous) sont représentés par des pavés blancs), les mauvais par des croix.

4.5.1 Influence de la normalisation des couleurs

La figure 4.11 montre le résultat de la mise en correspondance de la séquence test que nous utilisons. Ces images sont obtenues en considérant deux parties différentes d'une même image Ivoire. Dans le cas de la figure 4.11 aucun changement d'illumination n'est effectué sur la deuxième image. Ainsi, en absence de bruit l'appariement est parfait. Le taux de bons appariements est de 100%. Ceci dans le but de valider l'approche de mise en correspondance que nous avons adoptée.

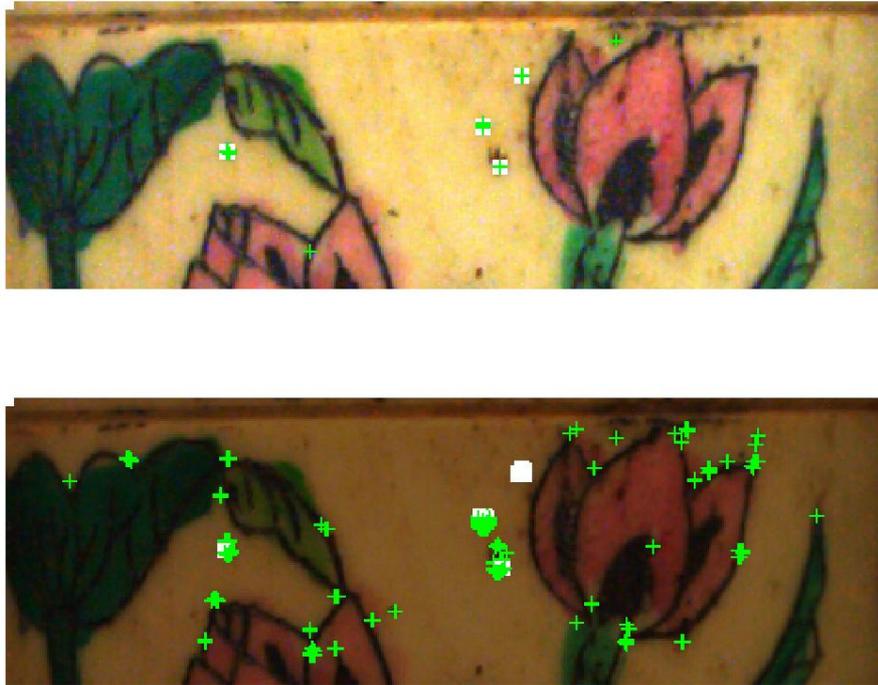


Figure 4.12 - Mise en correspondance des la mêmes images Ivoire avec un changement linéaire de la luminosité (: les bons appariements sont représentés par des pavés blancs, les mauvais par des croix vertes).

4.5.1.1 Appariement sans normalisation

La figure 4.12 montre le résultat de la mise en correspondance des deux images Ivoire précédentes dont la deuxième a subi un changement linéaire de luminosité obtenu par la transformation suivante :

$$\begin{cases} r_c = 0.6.r + 0.3, \\ v_c = 0.5.v + 0.2, \\ b_c = 0.4.b + 0.1, \end{cases} \quad (4.18)$$

r, v, b sont les composantes couleur de l'image d'origine.

r_c, v_c, b_c sont les composantes couleur de l'image modifiée (qui a subit le changement d'illumination).

Par cet essai nous voulons démontrer la dégradation de la qualité des appariements lorsque la luminosité change entre deux prises de vue. Aucune normalisation de la couleur n'a été faite. Nous constatons à l'œil un taux de bons appariements égal à 9%. Ce taux est extrêmement faible.

4.5.1.2 Appariement avec normalisation

La figure 4.13 montre le résultat de la mise en correspondance de deux images Ivoire différant d'une translation et d'un changement d'illumination appliqué à la deuxième image. Dans ce cas la couleur est normalisée par la méthode citée dans le paragraphe 4.4.2, le diamètre de la fenêtre de normalisation est pris égal à 11 pixels. Le taux de bons appariements en absence de bruit est ici de 100 %.

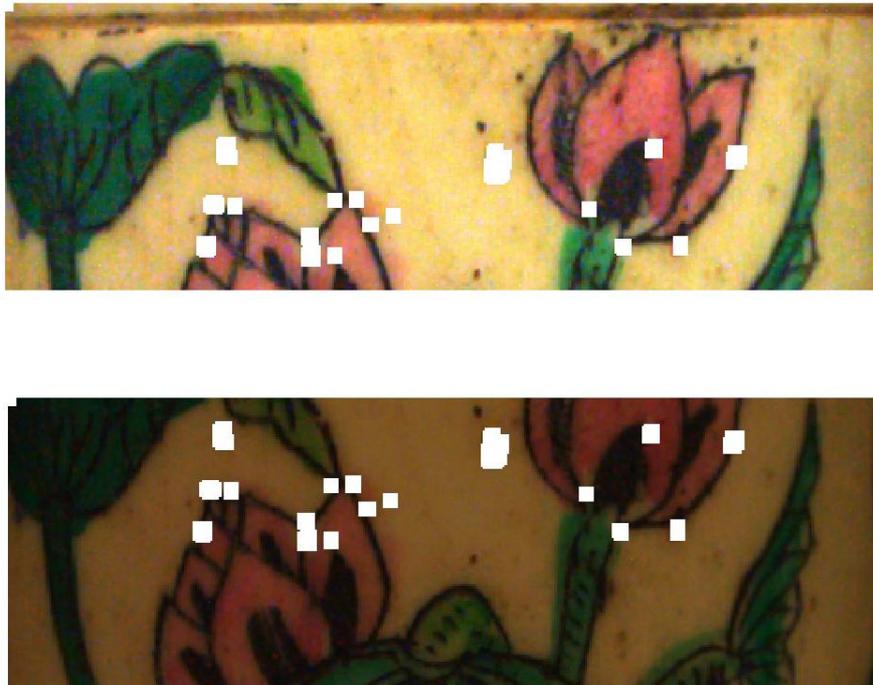


Figure 4.13 - Mise en correspondance de deux images Ivoire avec un changement linéaire de la luminosité et normalisation des couleurs (: les bons appariements (tous) sont représentés par des pavés blancs), les mauvais par des croix vertes.

Ces deux résultats montrent clairement l'intérêt de normaliser la couleur. Ceci dit, la difficulté réside dans le choix du diamètre de la fenêtre de normalisation. Il est donc nécessaire d'évaluer la qualité des appariements d'une paire d'images pour différentes valeurs du diamètre de cette fenêtre.

4.5.2 Influence de la taille de la fenêtre normalisation des couleurs

La figure 4.14 illustre le résultat de la mise en correspondance de la paire d'images Ours1/Ours2 sans normalisation de la couleur (images réelles acquises avec deux positions différentes du capteur MINOLTA [OSU]). Dans ce cas le taux de bons appariements est de l'ordre de 43%. Ce taux est amélioré et atteint plus de 56% en

appliquant une normalisation de la couleur avec une fenêtre dont le diamètre est égal à 33 pixels. Le tableau 4.1 résume les résultats obtenus pour ce taux pour différentes valeurs du diamètre de la fenêtre de normalisation.

4.5.2.1 Exemple 1: Figurine 3D



Figure 4.14 - Mise en correspondance des deux images Ours1 et Ours2 sans normalisation de la couleur: (les bons appariements sont représentés par des pavés blancs, les mauvais par des croix vertes. Le taux de bons appariements est de 43%).



Figure 4.15 - Mise en correspondance des deux images Ours1 et Ours2 avec normalisation de la couleur (: les bons appariements sont représentés par des crois blanches, les mauvais par des crois noires. Le taux de bons appariements est de 56%).

Taille fenêtre normalisation en pixels	11	15	17	19	21	23	25	27	29	31	33
Taux de bons appariements en %	27	29	37	23	25	28	30	42	47	52	56

Tableau 4. 1 - Taux des bons appariements en fonction du diamètre de la fenêtre de normalisation de la couleur : cas de la paire Ours1/Ours2.

La figure 4.15 illustre le résultat de l'appariement pour la séquence Ours1/Ours2 lorsqu'on applique une normalisation avec un diamètre de la fenêtre égal à 33 pixels.

4.5.2.2 Exemple 2 : Peinture à l'huile

Le résultat des appariements pour la séquence Calvet1/Calvet2 est donné figure 4.16 lorsqu'il n'y a pas de normalisation sur les couleurs et figure 4.17 lorsque la couleur est normalisée. La figure 4.17 correspond au résultat obtenu pour une normalisation dont le diamètre de la fenêtre est de 41 pixels. Dans ce cas le taux de bons appariements atteint plus de 65% contre seulement 43% quand la couleur n'est pas normalisée. Le tableau suivant (tableau 4.2) contient les résultats avec différentes valeurs du diamètre de la fenêtre normalisation.

Taille fenêtre normalisation en pixels	9	15	21	23	25	31	33
Taux de bons appariements en %	15	30	41	46	47	59	65

Tableau 4.2 - *Taux des bons appariements en fonction du diamètre de la fenêtre de normalisation de la couleur : cas de la paire Calvet1/Calvet2.*



Figure 4.16 - Mise en correspondance des deux images Calvet1 et Calvet2 sans normalisation de la couleur (: les bons appariements corrects sont représentés par des pavés blancs, les mauvais par des croix vertes. Le taux de bons appariements est de 43%).



Figure 4.17 - Mise en correspondance des deux images Calvet1 et Calvet2 avec normalisation de la couleur: (les bons appariements sont représentés par des pavés blancs, les mauvais par des croix vertes. Le taux de bons appariements est de 65%).

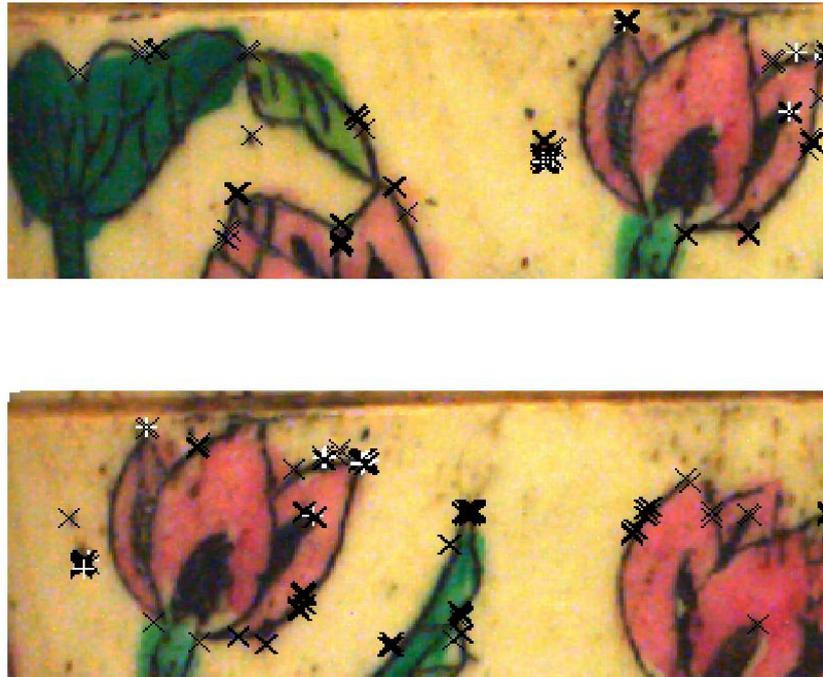


Figure 4.18 - Mise en correspondance des deux images Calvet1 et Calvet2Ivoire avec normalisation de la couleur diamètre 11 (: les bons appariements sont représentés par des croix blanches, les mauvais par des croix noires. Le taux de bons appariements est de 12%).

4.5.2.3 Exemple 3: Boite cylindrique

Les figures 4.18 et 4.19 illustrent les résultats des appariements de la paire d'images Ivoire1/Ivoire2 pour une fenêtre de normalisation de diamètre égal respectivement à 11 et 25. Les taux de bons appariements sont respectivement de l'ordre de 12% et 18%. Le tableau suivant résume les résultats pour des valeurs intermédiaires de ce diamètre.

Taille fenêtre normalisation en pixels	11	15	17	21	23	25	27	29	31	41
Taux de bons appariements en %	12	12	14	10	14	18	13	12	12	13

Tableau 4. - Taux des bons appariements en fonction du diamètre de la fenêtre de normalisation de la couleur (: cas de la paire Calvet1/Calvet2Ivoire).

Le résultat décevant correspondant aux pourcentages faibles de bons appariements pour cette séquence Ivoire est expliqué par la présence de bruit image relativement important, ainsi que de changements internes et externes de l'illumination. En effet, ces images ont été acquises sans précautions particulières quant aux changements de l'illumination. Aussi, la normalisation permet d'être robuste aux changements internes de l'illumination mais n'est pas efficace lors de changements externes et en présence de bruits image.



Figure 4.19 - Mise en correspondance des deux images Calvet1 et Calvet2Ivoire avec normalisation de la couleur diamètre 25 (: les bons appariements sont représentés par des croix blanches, les mauvais par des croix noires. Le taux de bons appariements est de 18%).

4.6 Conclusion

Dans ce chapitre nous avons vu, à travers l'état de l'art, les principales approches de mise en correspondance d'images 2D. Pour les méthodes de mise en correspondance éparse nous avons donné un aperçu des principales méthodes d'extraction de points d'intérêt. Quelques méthodes d'appariement de points d'intérêt ont été présentées. Nous avons opté pour une mise en correspondance éparse. Autrement dit, nous mettons en correspondance les images couleur 2D en se basant sur l'appariement des points d'intérêt extraits de ces images. Pour l'extraction des points d'intérêt nous avons opté pour le détecteur de Harris précis couleur. L'appariement des points d'intérêt est basé sur le calcul de vecteurs d'invariants couleur (paragraphe 4.4.1). Cet appariement est basé sur la distance euclidienne normalisée entre les vecteurs d'invariants associés aux points d'intérêt des deux images à mettre en correspondance. Nous avons d'abord montré l'intérêt d'appliquer une normalisation de la couleur et l'importance du choix du diamètre de la fenêtre de normalisation.

Nous avons obtenu des résultats satisfaisants concernant les images fournies par le capteur MINOLTA de l'OSU [OSU] ainsi que pour la paire d'images fournies par un appareil photo numérique du commerce. Ceci dit, le taux de bons appariements peut être amélioré en utilisant une contrainte de voisinage (géométrique ou photométrique) dans le but d'éliminer les appariements ambigus. Nous n'utilisons pas de contrainte de voisinage pour améliorer l'appariement en 2D. En effet, le but de cette étape est de fournir un pourcentage correct d'appariements 2D en vue d'une initialisation de la transformation 3-D entre deux images 3D/couleur. Dans le chapitre suivant, un algorithme basé sur le tirage aléatoire des paires 3-D associées permettra d'éliminer l'influence des faux appariements dans l'estimation de la transformation 3-D entre les deux images. Les résultats décevants obtenus pour Ivoire1/Ivoire2 sont expliqués par la présence de bruit important et par un changement complexe d'illumination (interne et externe). Ceci est du, aussi, à la déformation des motifs d'une image à l'autre à cause du changement de perspective.

5 RECALAGE AUTOMATIQUE 3D/COULEUR

Sommaire

5.1	INTRODUCTION	103
5.2	ÉTAT DE L'ART.....	104
5.3	ALGORITHME RANSAC	105
5.3.1	<i>Principe.....</i>	<i>106</i>
5.3.2	<i>Application au recalage de nuages de points 3D/couleur</i>	<i>108</i>
5.3.3	<i>Mise en œuvre de l'algorithme RANSAC</i>	<i>109</i>
5.4	VALIDATION EXPERIMENTALE AVEC DES DONNEES SIMULEES	112
5.4.1	<i>Cas de deux images sans bruit additionnel.....</i>	<i>112</i>
5.4.2	<i>Cas de deux images avec bruit additionnel.....</i>	<i>115</i>
5.5	VALIDATION DU RECALAGE COMPLET AVEC DES DONNEES REELLES	117
5.6	CONCLUSION.....	122

5.1 Introduction

Dans ce chapitre nous allons nous intéresser à l'initialisation automatique du recalage lorsque celui-ci est effectué par une méthode itérative telle que l'algorithme ICP. Nous donnons un bref état de l'art des travaux menés pour l'automatisation du processus de recalage 3D dans le premier paragraphe. Dans le deuxième paragraphe nous décrivons la méthode RANSAC que nous utilisons pour calculer une estimation initiale de la transformation 3D. La validation expérimentale de notre approche est présentée dans le troisième paragraphe.

L'approche classique pour le recalage de deux nuages de points 3D nécessite une connaissance approximative de la transformation entre ces deux nuages. Les méthodes de recalage itératives telles que l'algorithme ICP procèdent donc à un affinage de cette transformation initiale. Cette transformation peut être obtenue par différents moyens. Par exemple, lorsqu'il s'agit de la reconstruction d'objets de petites tailles, il est très fréquent d'utiliser un plateau tournant pour l'acquisition des différentes vues à recalcr. Dans ce cas, la transformation initiale correspond approximativement à la mesure de l'angle de rotation du plateau tournant entre les deux prises de vue.

Une autre technique similaire consiste à déplacer le capteur qui effectue les prises de vue 3D avec un bras de mesure. C'est généralement un bras articulé passif dont les articulations sont pourvues de codeurs. Pour calculer le déplacement du capteur à partir du mouvement de l'organe terminal du bras de mesure il faut avoir, au préalable, identifié précisément les paramètres de la transformation entre le repère de cet organe terminal et celui du capteur. Cette solution laisse à l'opérateur un grand choix d'angles de prise de vue. On peut citer, à titre d'exemple, les bras de mesure construits par la société FARO [FARO] qui intègrent ou non le système de mesure 3D (ex : Laser Line Probe de FARO). Ils possèdent jusqu'à sept degrés de liberté. Leur précision peut atteindre quelques microns pour un volume de travail de trois à quatre mètres.

Dans la plupart des cas, les logiciels de recalage commercialisés estiment la transformation initiale à partir de la mise en correspondance de points de contrôle définis par l'utilisateur. Grâce à une interface graphique, l'utilisateur peut sélectionner manuellement un certain nombre de paires de points (trois paires peuvent être suffisantes) sur les deux images à recalcr. Il est ainsi possible d'initialiser l'algorithme de recalage. Un autre moyen d'initialiser cet algorithme est l'utilisation de marqueurs et d'amers artificiels placés sur l'objet. Ceci a l'inconvénient de modifier la texture et parfois la forme de l'objet.

Dans ce chapitre nous présentons notre approche pour l'initialisation automatique du recalage. Le but du travail présenté ici, est de pouvoir procéder à un recalage de deux nuages de points 3D/couleur de façon totalement automatique. Notre approche se base sur la mise en correspondance des images 2D correspondant aux nuages 3D/couleur à recalcr. Cette mise en correspondance dans l'espace couleur a été effectuée avec la méthode présentée dans le chapitre 4. Nous avons accès aux

coordonnées 3D des points résultant de cette mise en correspondance. Nous pouvons donc constituer une liste de paires de points 3D. Cette liste contient un certain nombre de faux appariements qu'il est nécessaire d'éliminer pour estimer un modèle de transformation entre les deux vues qui soit le plus proche possible de la solution exacte. Pour cela, nous avons choisi une méthode de tirage aléatoire, l'algorithme RANSAC qui sera présenté dans la troisième partie de ce chapitre.

5.2 Etat de l'art

Dans cet état de l'art nous donnons un aperçu des solutions logicielles développées pour l'automatisation du processus de recalage 3D ou 3DC. Un grand nombre de travaux ont été menés pour le recalage de nuages de points 3D utilisant des méthodes itératives nécessitant une connaissance a priori de la transformation entre les nuages de points. L'initialisation automatique des algorithmes de recalage constitue actuellement un axe de recherche à part entière. En effet, cette initialisation reste l'élément manquant dans le processus de recalage. C'est la seule étape qui nécessite une intervention de la part de l'utilisateur.

Gelfand et al [GEL05] proposent une méthode de recalage automatique de deux nuages de points 3D. La méthode est basée sur le calcul de descripteurs géométriques robustes au bruit. Ces descripteurs sont calculés sur un certain nombre de points caractéristiques extraits de chaque nuage de points à recaler. Les points caractéristiques sont mis en correspondance par comparaison de leurs descripteurs géométriques. Ces correspondances sont utilisées pour calculer une transformation grossière fournie par un algorithme d'optimisation (branch-and-bound algorithm). Cette transformation grossière sert d'initialisation à l'algorithme ICP qui réalise le recalage fin entre les deux nuages initiaux.

La méthode proposée par Sang-Hoon Kim et al [KIM02] permet d'effectuer le recalage automatique d'un ensemble de nuages de points 3D de manière itérative, sachant que l'on dispose des images d'intensité correspondantes. La méthode s'inspire de l'algorithme proposé par Horn [HOR87]. Il consiste à aligner grossièrement les deux nuages de points 3D en calculant leurs axes d'inertie à partir des matrices de covariance qui décrivent les distributions des points dans l'espace. Compte tenu des problèmes liés aux symétries éventuelles et aux occlusions, cette solution ne fournit qu'une estimation grossière de la transformation 3D. Le recalage obtenu en utilisant cette transformation est affiné en calculant les matrices de projection 3D/2D dans le but de retrouver les positions exactes du capteur. Cette méthode fonctionne pour des paires d'images qui se recouvrent relativement bien, mais elle nécessite la connaissance ou le calcul des matrices de projection de l'espace 3D vers l'espace image (2D).

Ameesh Makadia et al [MAK06] ont développé une nouvelle technique de recalage automatique dont la particularité est, selon les auteurs, d'être efficace même pour des taux de recouvrements relativement faibles (pouvant atteindre 45%), entre les deux nuages de points 3D à recaler. Cette méthode utilise les images gaussiennes étendues (EGIs pour Extended Gaussian Images). Une estimation initiale de la transformation est obtenue en calculant la rotation 3D entre les deux images gaussiennes étendues. Cette technique est basée sur le calcul de la corrélation de ces

deux images gaussiennes dans le domaine de Fourier et fait usage des transformées harmoniques sphériques et rotationnelles. Les auteurs procèdent, enfin, à un affinage de la transformation initiale par ICP. Lorsque les nuages de points ont un faible recouvrement, une solution particulière basée sur une nouvelle représentation de l'EGI est mise en œuvre.

Chu-Song Chen et al [CHU99] proposent une méthode de recalage de deux images 3D qui se recouvrent partiellement. Cette méthode, nommée RANSAC-based-DARCES (Data Aligned Rigidity Constraint), ne nécessite pas d'initialisation de la transformation 3D. Elle utilise la contrainte de rigidité entre un certain nombre de points de contrôle présélectionnés, afin de réduire l'espace de recherche exploré pour réaliser la mise en correspondance des points 3D. Selon les auteurs, le temps de calcul est considérablement réduit, et l'algorithme converge toujours vers la bonne solution en l'absence de bruit sur les données.

Les paragraphes précédents nous ont permis d'avoir un aperçu sur quelques méthodes de recalage 3D automatique. L'approche de Gelfand et al est celle qui se rapproche le plus de l'approche que nous souhaitons adopter. Au lieu d'utiliser des caractéristiques géométriques pour calculer les correspondances permettant de fournir une transformation grossière entre les deux images à recaler, nous utilisons des caractéristiques photométriques pour mettre en correspondance les images 2D correspondantes aux nuages 3D à recaler (chapitre 4). Nous utilisons l'algorithme RANSAC pour éliminer les faux appariements résultant de la mise en correspondance couleur. Ainsi, nous obtenons grâce à l'algorithme RANSAC la transformation grossière permettant d'initialiser l'Algorithme ICP. Dans ce qui suit nous présentons l'algorithme RANSAC et son principe.

5.3 Algorithme RANSAC

Il existe de nombreuses techniques proposées dans la littérature qui permettent d'obtenir des estimateurs robustes. Parmi celles-ci nous pouvons citer la transformée de Hough, les LMS (*Least Median of Squares*), les M-estimateurs et la méthode RANSAC. Le lecteur peut se référer à [MAL05] pour un état de l'art et une description plus complète des estimateurs robustes les plus utilisés en vision robotique.

Le choix de la méthode RANSAC comme moyen d'éliminer les faux appariements (ou une partie) n'est pas fortuit, mais il est basé sur un travail antérieur dans le cadre duquel nous l'avons appliquée pour le recalage de deux nuages de points 3D épars (problème de localisation en robotique mobile) [DOU01]. Nous avons, dans le cadre de ces travaux, comparé en simulation les performances du paradigme RANSAC à celles d'un estimateur robuste (moindres carrés pondérés) pour estimer le déplacement d'un capteur embarqué. L'appariement 3D/3D était basé sur la distance de Mahalanobis. L'étude comparative nous a permis, d'une part, d'évaluer la robustesse de ces deux solutions face aux bruits de mesure (pour plusieurs niveaux de ces bruits), ainsi qu'en présence de points aberrants (à différents taux). En conclusion de cette étude, nous avons constaté que la méthode RANSAC était la plus robuste aux bruits de mesure et aux faux appariements.

L'algorithme RANSAC (RANdom SAMple Consensus) est une méthode d'estimation robuste développée par Fishler et Bolles en 1981 [FIS81]. Cet algorithme est l'une des méthodes les plus robustes dans le domaine du recalage de données visuelles. Sa particularité est d'être robuste à la présence de points aberrants à l'origine des faux appariements dans la phase de mise en correspondance du processus de recalage. RANSAC procède par tirage aléatoire sur des données contenant des points aberrants pour estimer le modèle correspondant à ces données.

Fishler et Bolles l'ont utilisé pour résoudre un problème de localisation classique en cartographie : "Etant donné un ensemble de balises (ou points de contrôle) dont la position est connue dans un repère de référence, déterminer dans ce repère, la position du point où se trouvait le capteur qui a mesuré ces balises". Ce problème les a conduit à estimer les paramètres d'un arc de cercle (le rayon et les coordonnées du centre) décrit par un ensemble de points 2D qui contient des points aberrants. Pour cet exemple, des échantillons de trois points (trois points sont suffisants pour déterminer un cercle) sur l'ensemble des données disponibles sont tirés de façon aléatoire. Pour chaque échantillon les auteurs calculent les paramètres du cercle (ou du modèle) correspondant. Les paramètres du modèle retenus sont ceux pour lesquels ils trouvent le plus grand nombre de points compatibles, c'est à dire de points dont la distance au modèle (distance à l'arc estimé) ne dépasse pas un certain seuil.

L'algorithme de Fishler et Bolles décrit ci-dessous, concerne l'estimation du modèle correspondant à un ensemble de points de mesure S contenant un certain nombre de points aberrants.

5.3.1 Principe

Par opposition aux méthodes classiques d'identification de modèles qui utilisent le maximum de mesures possibles pour éliminer progressivement celles qui sont invalides, l'algorithme de RANSAC s'appuie au départ sur le plus petit ensemble de données nécessaires pour élargir ensuite cet ensemble avec le maximum de données compatibles avec le modèle identifié. On peut résumer l'algorithme mis au point par Fishler et Bolles [FIS81] de la manière suivante :

Algorithme RANSAC :

On considère un modèle dont l'identification des paramètres nécessite un minimum de s points de mesure, et un ensemble de S mesures disponibles.

- (i) *Sélectionner de façon aléatoire un échantillon de s points sur l'ensemble des S mesures et identifier les paramètres du modèle à partir de cet échantillon.*
- (ii) *Déterminer le sous-ensemble de S_k points parmi les S mesures compatibles avec le modèle identifié dans l'étape précédente. Ces points sont déterminés en utilisant un critère de validité d prédéfini. Le sous-ensemble de S_k points est le consensus C correspondant à l'échantillon actuel.*

- (iii) Si la taille S_k (nombre des points compatibles) est plus grand qu'un certain seuil S_{\min} , recalculer le modèle en utilisant tous les S_k points et sauvegarder le modèle et le consensus C correspondant à ce tirage.
- (iv) Dans le cas contraire (S_k inférieur à S_{\min}), sélectionner un autre échantillon s et répéter les étapes (ii) et (iii).

Les paramètres à définir dans l'algorithme RANSAC sont les suivants :

- **La taille s d'un échantillon :**

Ce paramètre dépend du type de modèle recherché. La taille de l'échantillon s correspond au nombre minimum de données nécessaire pour déterminer ce modèle. Par exemple, dans le cas d'un cercle ou d'un arc de cercle, la taille de l'échantillon s est de trois points. Prendre une taille pour s , plus importante que la taille minimum, n'est pas conseillé. En effet, ceci réduit la probabilité de sélectionner un échantillon s ne contenant aucun point aberrant et fournissant une bonne estimée du modèle. Ainsi, la taille minimum nécessaire pour s maximise les chances (ou la probabilité) de procéder à un tirage aléatoire pour lequel s ne contient aucun point aberrant [CANTZ].

- **Le seuil de distance d :**

Ce seuil de distance est un critère de validité qui permet de sélectionner les points compatibles avec le modèle estimé à chaque tirage aléatoire. Pour l'exemple de Fischer et Bolles, les distances de tous les points de l'ensemble S au modèle sont calculées. Seuls les points dont la distance euclidienne est inférieure à d sont retenues. Par exemple d peut être pris égal à la moyenne de ces distances, le seuil est dans ce cas large et imprécis (un point trop éloigné du modèle a une influence considérable sur le reste des points). Un seuil plus robuste aux points aberrants est la médiane de ces distances. Lorsque les caractéristiques des bruits de mesures sont connues, il peut être envisagé d'utiliser une distance de Mahalanobis qui tient compte de ces bruits.

- **Le nombre de tirages aléatoires N :**

Le nombre de tirages aléatoires idéal serait de considérer toutes les combinaisons de points possibles. Or ceci, n'est pas nécessaire et peut être très coûteux (si ce n'est impossible) en temps de calcul. Le nombre de tirages N est pris suffisamment grand tel qu'on ait une probabilité p , souvent égale à 0.99, qu'au moins l'un des tirages ne contienne aucun point aberrant. Ce nombre dépend donc fortement de la proportion de points aberrant contenus dans les données. Plus grand est le taux des points aberrants plus important est le nombre de tirages nécessaire pour trouver un échantillon ne contenant pas de point aberrant. L'expression suivante donne le nombre de tirages N en fonction de cette probabilité p et de la proportion ε de points aberrants contenus dans les données.

$$N = \log(1 - p) / \log(1 - (1 - \varepsilon)^s) \quad (5.1)$$

Cette expression est utilisable lorsque la proportion des points aberrants contenus dans les données est connue. Dans la plupart des cas cette information n'est pas disponible. N peut être calculé de façon adaptative. Le nombre N est initialisé, dans ce cas, à une valeur initiale, il est ensuite mis à jour à chaque tirage aléatoire, la condition d'arrêt de RANSAC est également modifiée, comme suit :

- $N = \infty$, *compteur* = 0.
- *While* $N > \text{compteur}$, *répéter* :
 - choisir un échantillon et compter le nombre de points compatibles (inliers).*
 - $\varepsilon = 1 - (\text{nombre de points compatibles}) / (\text{nombre total de points}).$
 - définir* N *correspondant à* ε .
 - Incrémenter compteur.*
- *Fin.*

- **Le nombre S_{\min} des points compatibles :**

Ce paramètre permet de décider si un tirage est intéressant ou non. A chaque tirage est calculé le nombre de points compatibles avec le modèle. Si ce nombre est supérieur à S_{\min} , les points compatibles du tirage sont conservés. S_{\min} est inversement proportionnel au taux de points aberrants.

5.3.2 Application au recalage de nuages de points 3D/couleur

Nous utilisons la méthode RANSAC pour éliminer les points aberrants (faux appariements) dans la liste de paires de points issue de la mise en correspondance des images 2D telle que nous l'avons décrit dans le chapitre précédent. Le but est d'obtenir une estimation approximative de la transformation 3D correspondant au déplacement entre les deux vues à recalculer. Cette transformation servira à initialiser l'algorithme ICP pour affiner le recalage. Ainsi, nous serons en mesure d'effectuer le recalage des deux nuages de points 3D/couleur de façon entièrement automatique, autrement dit, sans aucune intervention de l'utilisateur. Notre méthode de recalage automatique comporte trois étapes :

- 1- création d'une liste de paires de points 3-D à partir de la mise en correspondance des images couleur correspondant aux deux vues 3D/couleur,
- 2- estimation de la transformation 3D initiale avec la méthode RANSAC qui permet d'éliminer les outliers (faux appariements) dans la liste de paires 3-D.
- 3- recalage fin avec l'algorithme ICP amélioré,

La figure 4.1 présente le diagramme détaillé de cette approche.

5.3.3 Mise en œuvre de l'algorithme RANSAC

L'algorithme RANSAC est appliqué, ici, à l'estimation de la transformation de repère T (modèle à identifier) entre deux nuages de points 3-D/couleur notés Nuage1 et Nuage2. Dans ce cas, il utilise en entrée les coordonnées 3-D de S paires de points mis en correspondance selon la méthode décrite dans le chapitre 4. Il fournit en sortie une liste de n paires de points correctement appariés (inliers), ainsi que la transformation T_{ransac} estimée à partir de cet ensemble de n paires de points.

Le critère de validité des appariements, c'est-à-dire le seuil d qui permet de décider si une paire de points 3-D est compatible avec la transformation T_k , identifiée à partir d'un échantillon s , est généralement appliqué à la distance euclidienne calculée dans l'espace 3-D.

Pour exécuter l'algorithme RANSAC, nous devons fixer au préalable deux paramètres :

- la taille s de l'échantillon, c'est-à-dire, le nombre de paires de points prélevés aléatoirement à chaque tirage, parmi les S paires initiales.
- le nombre de tirages N qui assure une probabilité suffisante (99%) d'avoir au moins un échantillon dont toutes les paires sont valides.

L'estimation des six inconnues (trois rotations et trois translations) de la transformation de repère T requiert, théoriquement, la connaissance des coordonnées 3-D de 3 paires de points non alignés. Cependant, compte tenu du bruit qui entache ces mesures, nous utiliserons des échantillons qui contiennent un nombre s de paires supérieur à 3.

Nous avons vu que la taille s de l'échantillon conditionne le nombre N de tirages - Ainsi, l'équation (5.1) permet de calculer le nombre N qui assure une probabilité $p = 99\%$ d'avoir au moins un échantillon correct - La table (5.1) donne un aperçu des valeurs de N en fonction de s et de la proportion ε de paires de points non valides contenue dans la liste initiale.

Taille d'un échantillon	Proportion ε de paires non valides					
	20 %	30 %	40 %	50 %	60 %	70 %
4	8.74	16.76	33.18	71.36	177.58	566.234
5	11.6	25.03	56.89	145.05	447.42	1892.83
6	15.15	36.8	96.384	292.422	1122.005	6314.803
7	19.57	53.59	162.2	587.16	2808.47	21054.72

Tableau 5. 1 - Nombre N de tirage

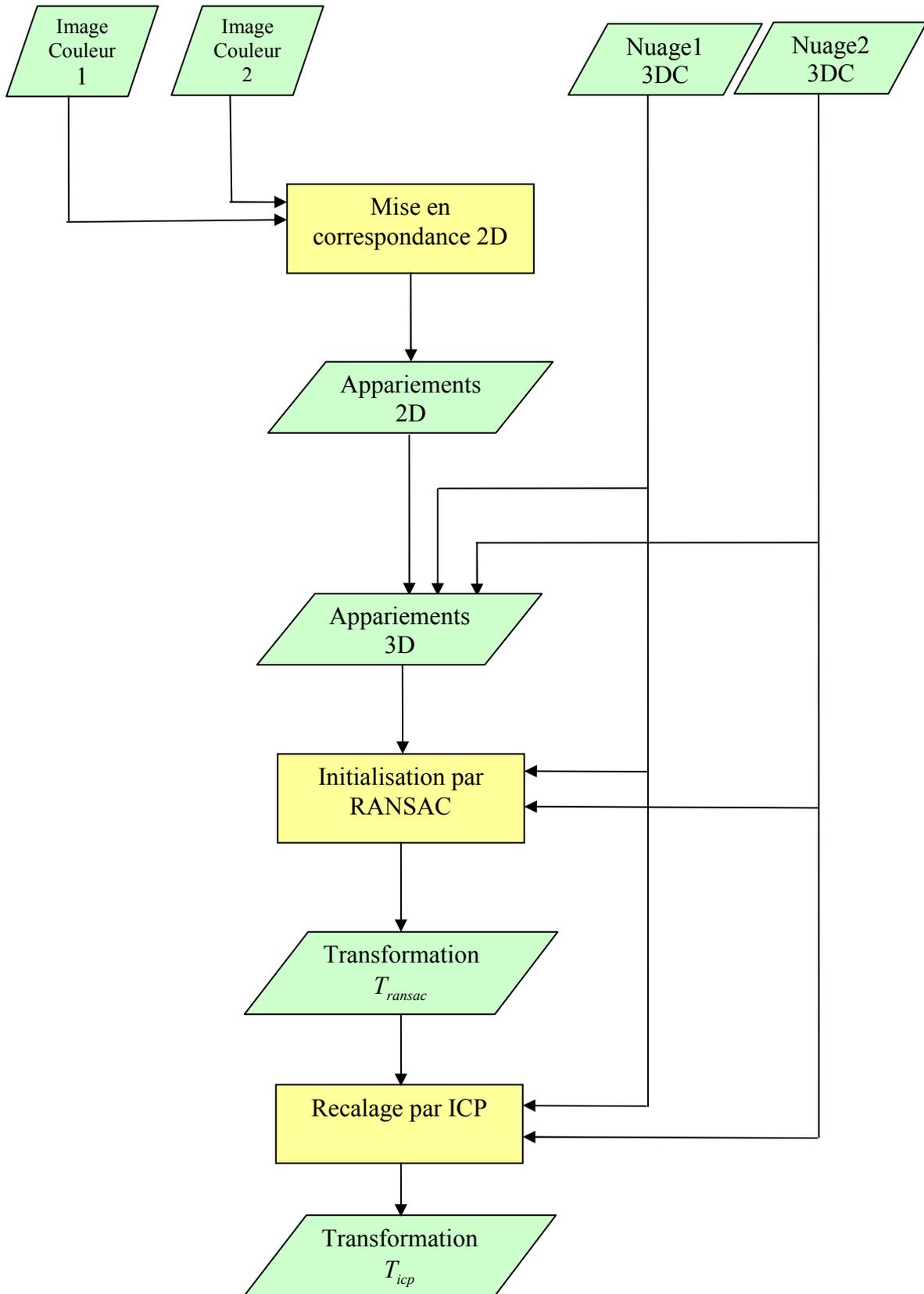


Figure 5.1 - Principe du recalage automatique

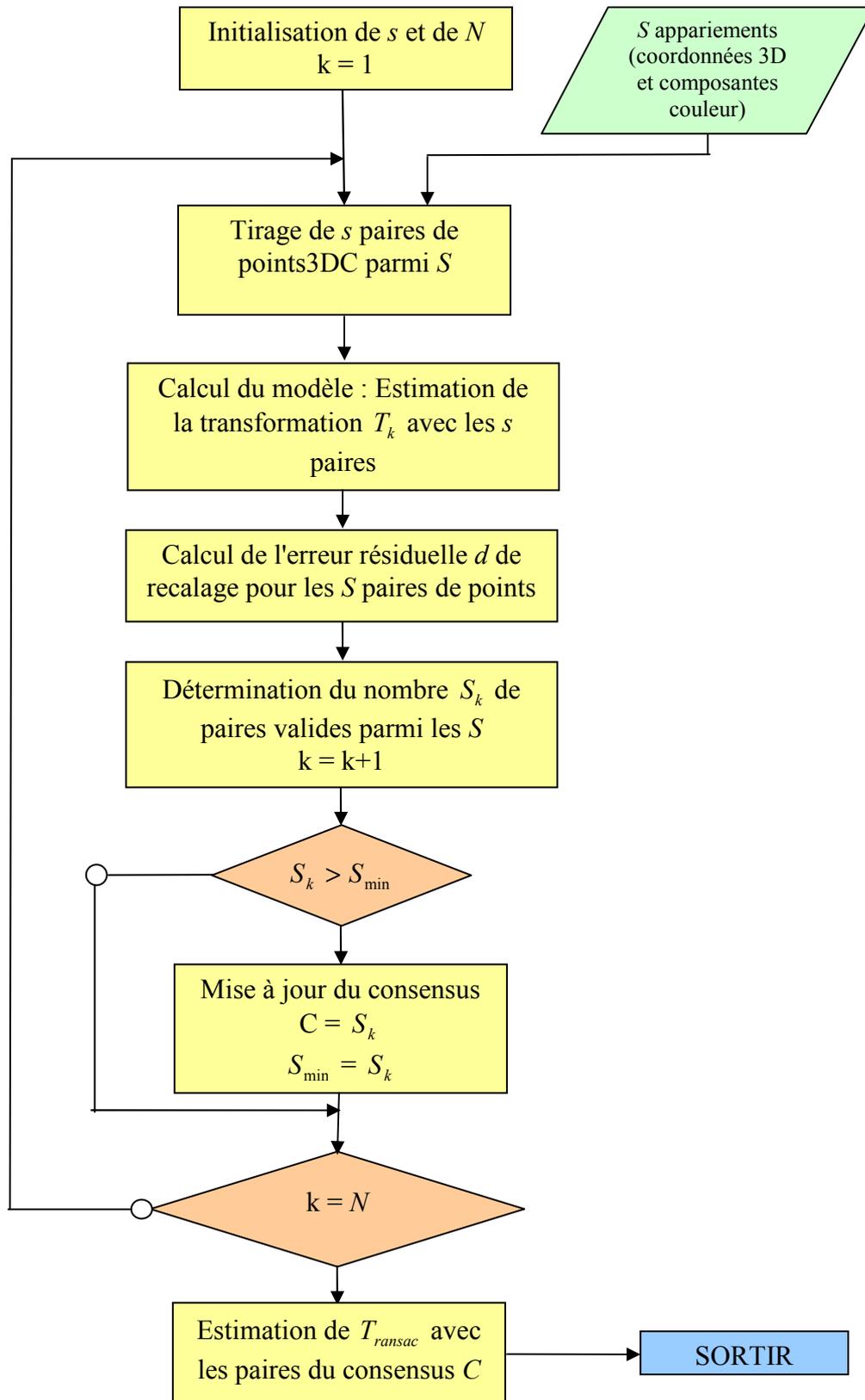


Figure 5.2 - Organigramme de l'algorithme RANSAC.

5.4 Validation expérimentale avec des données simulées

Pour une première validation de notre méthode d'initialisation nous testons la procédure d'initialisation avec des données simulées. Il s'agit de générer deux vues d'un objet à partir d'une même image 3D/couleur réelle. Pour cela, nous sélectionnons deux sous-ensembles de ce nuage de points qui se recouvrent partiellement (figure 5.3). Ces deux sous-ensembles sont obtenus en simulant un déplacement du nuage initial. On connaît, ainsi, la transformation réelle entre les deux vues à recaler, autrement dit, la vérité terrain.

Ainsi, le simulateur que nous avons développé nous permet, à partir d'une image réelle ou simulée, de générer deux vues qui se recouvrent partiellement en fonction du déplacement simulé. Il est, également, possible de simuler du bruit sur les données photométrique ainsi que sur les données géométriques.

5.4.1 Cas de deux images sans bruit additionnel

Dans un premier temps nous appliquons notre méthode d'initialisation à des données sans bruit additionnel. Nous utilisons l'une des images Ivoire pour extraire deux vues ("ivoire1" et "ivoire2") se recouvrant partiellement en simulant un déplacement comme décrit plus haut. Les figures 5.3 (a et b) illustrent ces deux vues simulées.

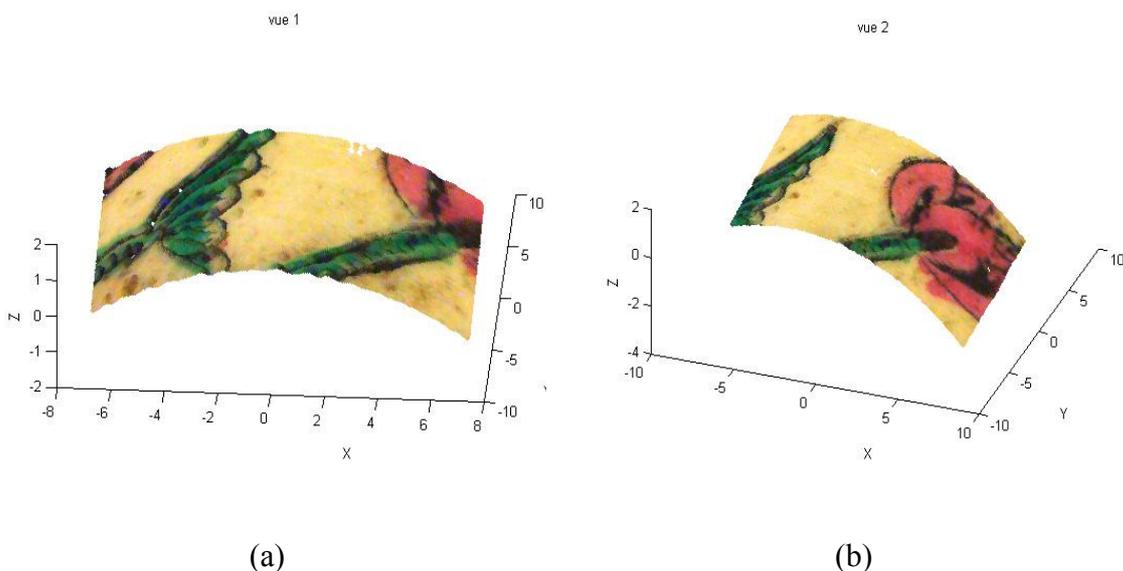


Figure 5.3 - Les deux vues "ivoire1" (figure a) et "ivoire2" (figure b) avec déplacement simulé. Chaque nuage contient plus de 27000 points 3D/couleur.

Dans ce cas le déplacement simulé est la combinaison de deux translations suivant les deux axes X et Y : respectivement t_x qui vaut 3 centimètres $t_y = 3$ centimètres. Le recouvrement entre les deux images est dans ce cas de l'ordre de 45%.

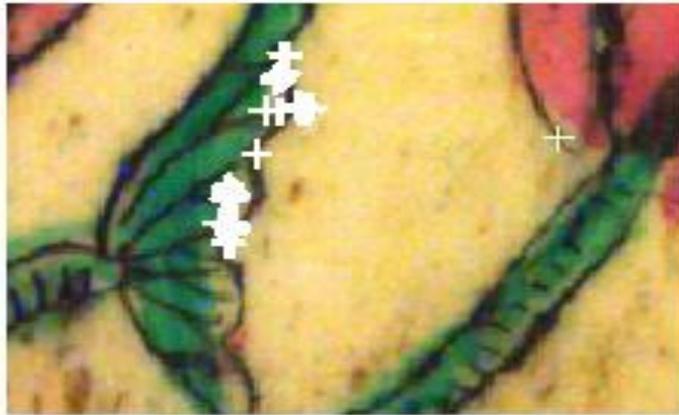


Figure 5.4 - Points 2D appariés dans les des deux images "ivoire1" et "ivoire2", Les appariements corrects (tous, dans ce cas) sont représentés par les croix "+" blanches.

La figure 5.4 illustre le résultat de la première phase du recalage automatique qui est la phase de mise en correspondance 2D des images couleur correspondant aux deux vues simulées ("ivoire1" et "ivoire2"). Dans ce cas la mise en correspondance est parfaite, 100% des appariements sont corrects. Il n'est donc pas nécessaire d'effectuer un tirage aléatoire sur les appariements (ceci ne contenant pas de faux appariements). Lorsque l'on utilise la totalité des appariements 3D pour estimer la transformation entre les deux vues, nous retrouvons la transformation que nous avons simulée. Le résultat de ce recalage est illustré par la figure 5.5. Par contre, nous constatons que lorsqu'on utilise l'algorithme RANSAC pour estimer l'initialisation nous ne retrouvons pas cette

transformation avec précision mais une transformation proche de celle-ci. Ce résultat est très intéressant car il met en évidence un problème très important que peut rencontrer l'algorithme RANSAC. Il s'agit du cas où les bons appariements sont concentrés sur une seule zone de l'image. En effet, pour que RANSAC fonctionne correctement avec des échantillons de 4 ou 5 paires, il est nécessaire que ces paires soient bien réparties dans l'espace. Autrement dit, en dehors du fait que celles-ci ne doivent pas être alignées, elles doivent également être relativement éloignées les unes des autres.

Nous pouvons constater sur la figure 5.4 la concentration des points appariés sur une zone de l'image (le bord droit de la feuille) et une répartition dans l'espace qui favorise les tirages aléatoires de paires de points alignées ou trop rapprochées. Ceci peut être du, soit à un problème de détection de points d'intérêt (parmi les points détectés par Harris Précis Couleur –voir chapitre 4- nous n'avons retenu que les 300 points dont le coefficient R est le plus grand), soit à un seuillage trop sévère sur les distances entre les vecteur d'attributs des points d'intérêt retenus. Dans ce cas exemple nous n'avons pas jugé utile d'effectuer des essais sur les paramètres liés à la mise en correspondance 2D (la taille des filtres de lissage, la taille de la fenêtre de normalisation des couleur, le seuillage des distances d'attributs...etc). Dans les exemples suivants ces paramètres ont été définis de façon empirique.

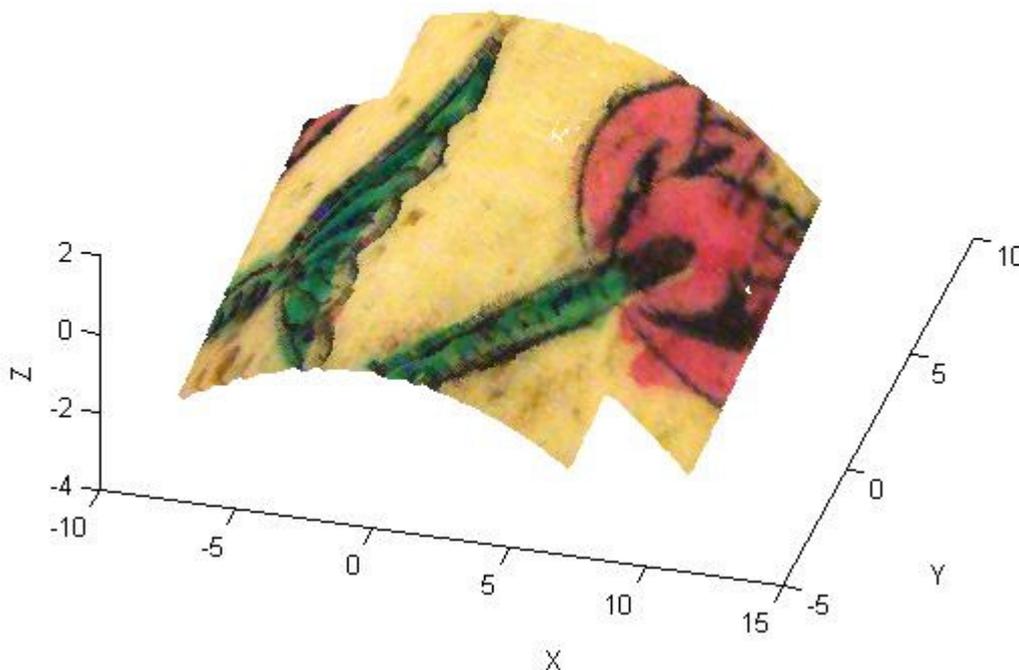


Figure 5.5 - Recalage 3D automatique des deux vues Ivoire par mise en correspondance de points d'intérêt couleur.

5.4.2 Cas de deux images avec bruit additionnel

Nous allons à présent nous intéresser au recalage de deux vues obtenues par déplacement simulé mais nous faisons subir à la deuxième vue un changement d'illumination (cf. chapitre 4, équation 4.18). Les figures 5.6 (a et b) montrent les deux vues à recaler dans ce cas. Il s'agit, de deux vues obtenues en simulant deux acquisition voisines sur une image du tableau Wallis1. Le changement d'illumination a été effectué sur la deuxième image. Le déplacement simulé entre les deux vues est une combinaison de deux translations t_x et t_y égales à 5 centimètres, ce qui donne un taux de recouvrement de l'ordre de 60%.

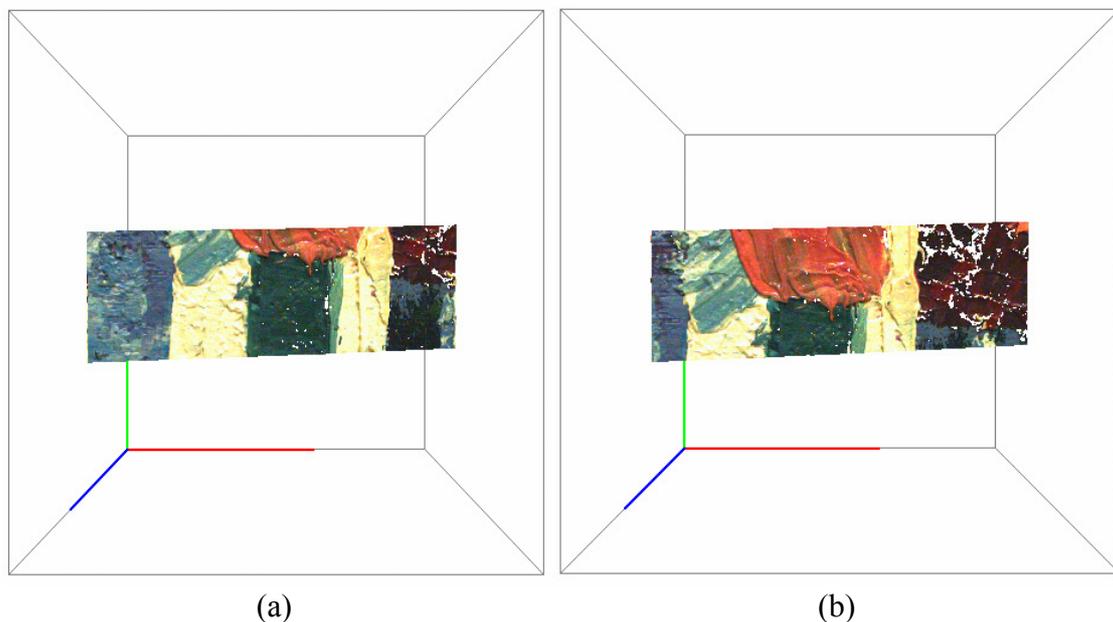


Figure 5.6 - Deux vues 3DC de "Wallis" avec déplacement simulé.

unités	translations en [cm]			rotations en [rad]		
	t_x	t_y	t_z	rho(axe Z)	theta (axe X)	phi (axe Y)
Degrés de liberté						
Valeur de l'erreur	0.6897	0.7402	1.0179	0.0427	0.0125	3.7100e-004

Tableau 5. 2 - Erreurs entre la transformation estimée et la transformation exacte pour le recalage de la figure (5.8).

La figure 5.7 montre le résultat de la mise en correspondance 2D des images 2D couleur correspondant aux deux nuages 3D/couleur à recaler. Le taux de bons appariements est, dans ce cas, de l'ordre de 80%. Nous pouvons, ainsi, déduire le nombre de tirages minimum nécessaire au fonctionnement de RANSAC (équation 5.1). Nous obtenons un nombre de tirages minimum égal à 19,56 que nous arrondissons à 20 tirages. La figure 5.8 montre le résultat du recalage obtenu avec la transformation de repère issue de ces 20 tirages. Le tableau ci-dessus (tableau 5.1) donne les erreurs des rotations et des translations entre la transformation initiale estimée par RANSAC et la transformation exacte (vérité terrain).



Figure 5.7 - Résultat de l'appariement des images couleur correspondant aux nuages de la figure 5.6.

Nous constatons pour cet exemple que la qualité de la transformation estimée par RANSAC est relativement bonne. Le nombre de paires s de l'échantillon aléatoire est pris égal à 4 points. RANSAC donne dans ce cas des résultats intéressants. Ceci est du, d'une part à une 'bonne' répartition dans l'espace des paires de points appariés ainsi qu'à un recouvrement de l'ordre de 60%. A partir de cette transformation nous pouvons affiner le recalage avec l'une des variantes de l'algorithme ICP présentées dans le chapitre (chapitre 3). Dans la section suivante nous appliquons la procédure complète de recalage automatique (l'initialisation et l'affinage) à deux images réellement acquises sous deux points de vue différents.

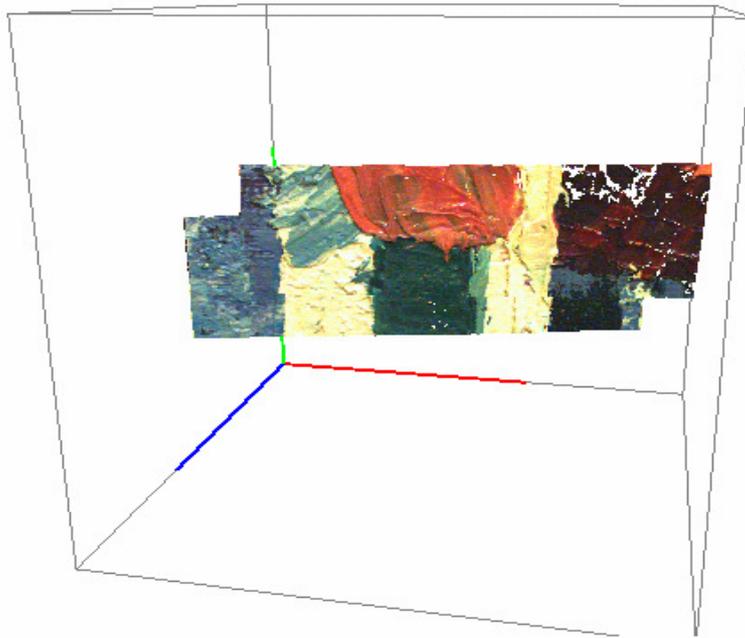


Figure 5.8 - Résultat du recalage des images de la figure 5.7 avec la transformation initiale estimée par RANSAC.

5.5 Validation du recalage complet avec des données réelles

Dans cette section nous validons notre méthode de recalage automatique avec des données réelles fournies par le capteur MINOLTA de l'OSU [OSU]. Il s'agit des deux nuages 3D/couleur Ours1 et Ours2 que nous avons utilisés dans le chapitre (chapitre ICP) pour l'étude comparative des variantes de ICP avec initialisation manuelle. Le but ici est d'effectuer un recalage automatique de ces deux nuages. Les données fournies au système sont les deux vues 3D/couleur ainsi que les images 2D couleur leur correspondant. En sortie nous récupérons la transformation qui recalle le mieux ces deux vues. Cette transformation est donc calculée par ICP initialisé par la transformation approximative fournie par RANSAC, lui-même initialisé par l'appariement 2D effectué par la méthode décrite dans le chapitre 4.

La figure 5.9 (a et b) illustre les deux nuages de points 3D/couleur à recaler. Leur résolution spatiale est de 0.62 mm. Dans la première phase du recalage automatique, c'est à dire, la mise en correspondance 2D, les paramètres de lissage du détecteur Harris Précis Couleur sont $\sigma = 1$ et $\tilde{\sigma} = 1$, la taille des fenêtres des lissages est égale à 7 pixels (7x7 pour le lissage 2D). Le rayon de la fenêtre de normalisation des couleurs est pris égal à 13 pixels. Le lissage appliqué pour le calcul des descripteurs des points d'intérêt est $\sigma = 2$ pour une fenêtre de lissage de taille égale à 15 pixels.

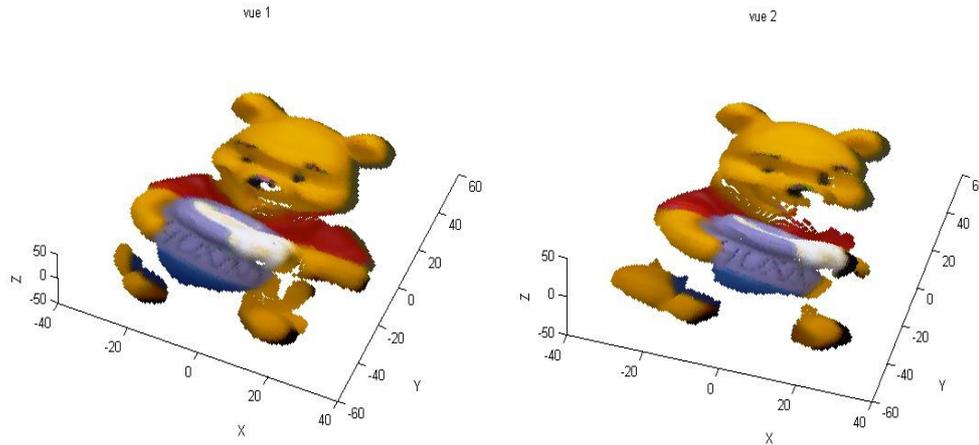


Figure 5.9 - Les deux nuages Ours1 (11650 points) et Ours2 (10789 points) à recalcer.

La figure 5.10 montre le résultat de la mise en correspondance 2D des images Ours1 et Ours2 correspondant aux deux vues 3D/couleur à recalcer. Nous retenons pour la mise en correspondance 300 points d'intérêt dans chaque image (correspondant aux 300 plus grandes valeurs du coefficient de Harris 'R'). Nous obtenons, dans ce cas, un taux de faux appariements de l'ordre de 40%. Ce taux est supérieur à celui retrouvé dans l'exemple précédent (cas des images du tableau : Wallis1 et Wallis2 avec déplacement simulé).

L'équation 5.1 nous donne un nombre de tirages minimum pour RANSAC égal à 57 tirages. A chaque tirage, nous en effectuons 60. Nous prenons des échantillons dont le nombre de paires s est égal à 5. On peut constater une bonne répartition des appariements dans l'espace (voir figure 5.10). Ceci augmente la probabilité de tirer un échantillon de 5 paires correctes relativement "bien" réparties dans l'image (éviter les paires alignées). La figure 5.11 illustre le résultat de l'estimation de la transformation initiale par l'algorithme RANSAC.

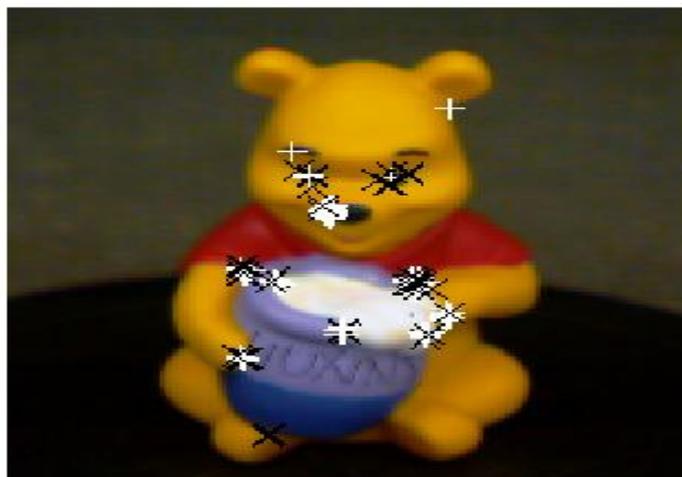


Figure 5.10 - Résultat de la mise en correspondance 2D des deux images Ours1 et Ours2 (les bons appariements sont marqué avec des croix (+) blanches, les faux appariements sont marqués avec des croix (x) noires).

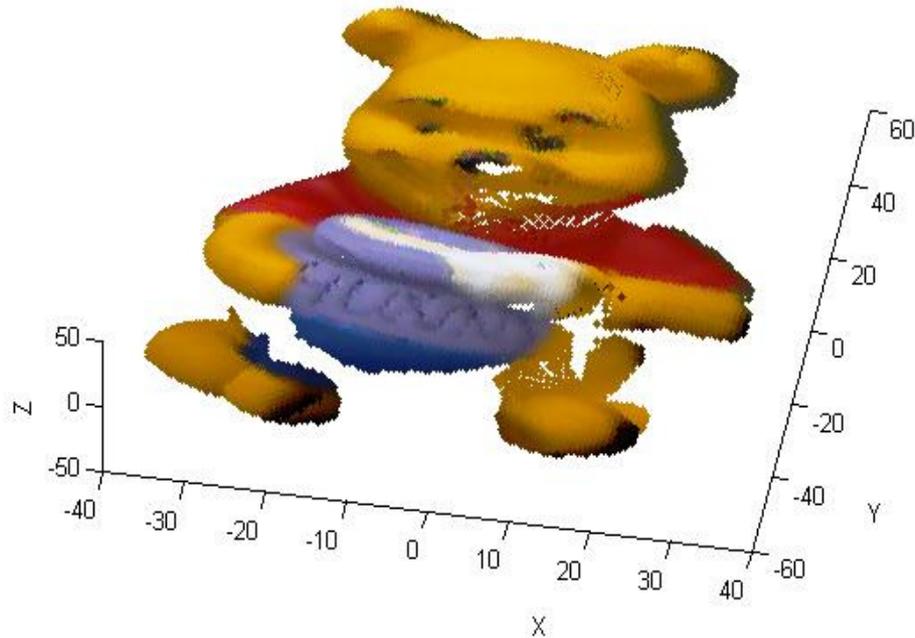


Figure 5.11 - Recalage grossier des nuages 3D/couleur Ours1 et Ours2 par la transformation estimée avec RANSAC.

Au bout des 60 tirages aléatoires, nous obtenons la transformation approximative T_0 dont les valeurs des paramètres (de rotation et de translation) sont données dans le tableau ci-dessous (tableau 5.2). N'ayant aucune connaissance de la transformation exacte entre ces deux vues nous ne pouvons calculer les erreurs de rotation et de translation. Ceci dit, la moyenne des distances entre les paires appariées, égale à 1,07 millimètres, nous donne une idée de la qualité du recalage. Nous pouvons, également, apprécier la qualité du recalage sur l'image de la figure 5.11.

unités	translations en [cm]			rotations en [rad]		
	tx	ty	tz	rho(axe Z)	theta (axe X)	phi (axe Y)
Degrés de liberté						
Valeur du paramètre	7.9381	0.4956	0.8150	0.1285	0.0317	0.2316

Tableau 5.3 Paramètres de translation et de rotation de la transformation estimée entre les vues Ours1 et Ours2 par RANSAC.

La transformation fournie par RANSAC nous servira dans ce qui suit à affiner le recalage avec la variante la plus performante de ICP couleur décrite dans le chapitre (chapitre ICP). Cette variante est basée sur le calcul d'une distance couleur entre les points 3D/couleur exprimée dans l'espace YIQ. La convergence de ICPGCTYIQ est atteinte dans ce cas au bout de 73 itérations. La moyenne des distances entre les points appariés est de l'ordre de 0.76 millimètres pour une résolution des données de l'ordre de 0.62 millimètres. La figure 5.12 donne l'évolution de cette erreur en fonction des itérations de ICPGCTYIQ.

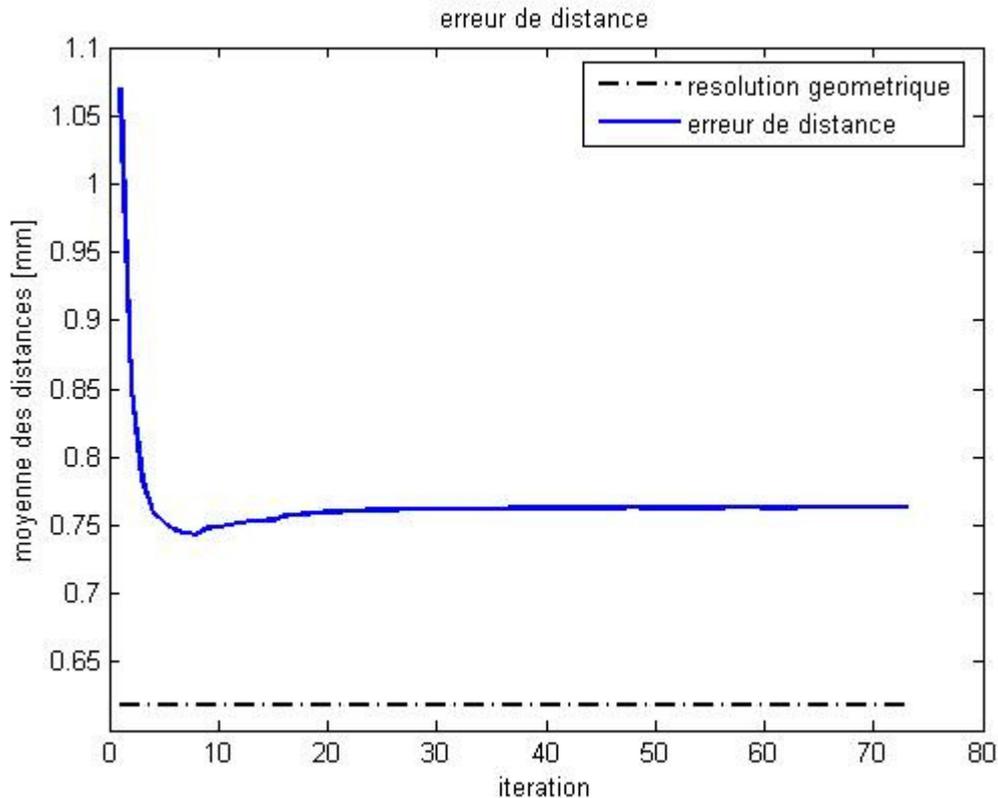


Figure 5.12 - Evolution de la moyenne des distances entre points appariés en fonction des itérations de ICPGCTYIQ pour le recalage des nuages Ours1 et Ours2.

unités	translations en [cm]			rotations en [rad]		
	tx	ty	tz	rho(axe Z)	theta (axe X)	phi (axe Y)
Degrés de liberté						
Valeur du paramètre	6.5050	0.5178	0.4719	0.1277	0.0023	0.2686

Tableau 5.4 - Paramètres de translation et de rotation de la transformation estimée entre les vues Ours1 et Ours2 par ICPGCTYIQ.

Le nombre de points appariés à l'issue du recalage fin est de 5395 paires. La vitesse de convergence est de 487 secondes (avec un processeur à 2 GHz). Le tableau suivant (tableau 5.3) donne les paramètres de la transformation \hat{T} estimée par ICPGCTYIQ. La figure 5.13 illustre le résultat du recalage des deux nuages Ours1 et Ours2 avec la transformation \hat{T} .



Figure 5.13 - Recalage fin des nuages 3D/couleur Ours1 et Ours2 par la transformation estimée par ICPGCTYIQ.

On peut constater la qualité du recalage qui est améliorée avec ICPGCTYIQ par rapport au recalage fourni par RANSAC. En effet, la moyenne des distances entre points appariés est réduite à 0.76 millimètres (contre – millimètres pour RANSAC) qui est proche de la résolution géométrique du capteur 3D de l'ordre de 0.63 millimètres. On peut aussi constater la qualité du recalage amélioré par ICPGCTYIQ sur la figure 5.13 en la comparant au résultat de la figure 5.11.

5.6 Conclusion

Nous avons présenté dans ce chapitre notre approche pour automatiser entièrement le recalage de deux nuages de points 3D/couleur. Le but était d'effectuer un recalage fin sans aucune intervention de l'utilisateur. Notre méthode comporte trois

étapes principales : la mise en correspondance des deux images 2D correspondant aux deux nuages de points 3D/couleur à recalcr, l'estimation de la transformation approximative (recalage grossier) par RANSAC en utilisant les appariements 3D correspondant aux appariements 2D de la première étape et le recalage fin avec la variante ICPGCTYIQ initialisée par la transformation approximative calculée dans la deuxième étape.

Le bon fonctionnement de notre approche automatique dépend fortement de l'étape de mise en correspondance 2D. En fonction de la qualité et de la répartition spatiale des appariements résultants de cette étape, l'algorithme RANSAC peut donner un résultat plus ou moins bon. Dans le cas où, par exemple, les paires de points mis en correspondance dans la première étape sont concentrées et mal réparties dans l'espace ou lorsque ces paires sont alignées, la transformation initiale estimée peut être fautive. Il est donc très important de bien ajuster les paramètres liés à cette première phase (taille des fenêtres de lissage gaussien du détecteur Harris Précis Couleur, paramètres de la caractérisation couleur, taille de la fenêtre de normalisation des couleurs et seuil de distance entre vecteurs d'attributs lors de l'appariement 2D). Par ailleurs, pour des images couleur trop bruitées, le taux de mauvais appariements peut atteindre des niveaux importants (au-delà de 50% des appariements). Alors, la probabilité pour que RANSAC puisse tirer un échantillon aléatoire exempt de mauvais appariements devient trop faible. Il faut alors envisager un nombre de tirages aléatoires trop important pour que RANSAC réussisse à converger vers une solution "acceptable" dans un temps raisonnable.

Le recalage grossier effectué avec RANSAC est affiné en utilisant l'une des variantes de ICP que nous avons développées et présentées dans le chapitre précédent ICPGCTYIQ. Cette dernière est initialisée avec la transformation approximative fournie par RANSAC. Nous avons testé notre méthode avec des données simulées ainsi qu'avec des données réelles. Les données simulées ont permis d'apprécier la qualité du recalage (grossier et fin) en comparant la transformation estimée à la transformation exacte connue. L'utilisation de données réelles permet de tester la robustesse de notre méthode face aux bruits de mesure (géométriques et photométriques). Nous avons obtenu des résultats très intéressants pour le cas des données 3D/couleur fournies par le capteur MINOLTA. Les résultats significatifs obtenus pour les données simulées ainsi que pour les données réelles montrent la faisabilité de notre méthode de recalage automatique de paires de nuages de points 3D/couleur.

CHAPITRE 6

6 CONCLUSION GENERALE

Sommaire

6.1	BILAN DES TRAVAUX REALISES.....	127
6.1.1	<i>Recalage itératif utilisant l'information couleur.....</i>	<i>127</i>
6.1.2	<i>Automatisation du processus de recalage</i>	<i>127</i>
6.2	PERSPECTIVES	128
6.2.1	<i>Amélioration du recalage de deux nuages 3D/Couleur.....</i>	<i>129</i>
6.2.2	<i>Recalage de N vues 3D/Couleur.....</i>	<i>129</i>
6.2.3	<i>Intégration des données pour la construction d'un modèle surfacique 3D/C</i>	<i>130</i>

Le travail réalisé dans cette thèse concerne le recalage de nuages de points denses et non structurés acquis par un scanner 3D à haute résolution, en vue de la construction de modèles 3D texturés d'objets complexes. L'une des principales contributions de cette thèse concerne l'utilisation de l'information couleur pour améliorer les performances du recalage en utilisant l'algorithme ICP, initialement conçu pour le traitement de données géométriques pures. La deuxième contribution majeure de notre travail concerne l'initialisation de l'algorithme ICP. Nous avons développé une méthode qui exploite la couleur et qui est basée sur la mise en correspondance d'images 2-D pour nous fournir cette initialisation. Nous avons ainsi proposé une méthode de recalage de deux nuages de points 3-D/couleur entièrement automatique. Les sections suivantes donnent un résumé du travail présenté dans cette thèse suivi des perspectives.

6.1 Bilan des travaux réalisés

6.1.1 Recalage itératif utilisant l'information couleur

Nous avons développé une nouvelle variante de l'algorithme ICP pour le recalage de deux nuages de points 3-D/couleur. Celle-ci utilise l'information couleur dans la phase d'appariement de l'algorithme. Nous avons montré l'efficacité de cette variante comparée à l'algorithme classique ainsi qu'à d'autres variantes de ICP que nous avons adaptées à notre problème. Notre méthode donne de meilleures performances en terme de vitesse de convergence et donne une erreur résiduelle de recalage inférieure à celle atteinte par les autres méthodes évaluées. Ceci est dû à la capacité de notre algorithme à rejeter de manière efficace les mauvais appariements, et ce grâce à l'utilisation d'une contrainte de distance couleur dans la phase de mise en correspondance. Nous avons ainsi démontré que, lorsque la couleur est disponible, celle-ci constitue une information supplémentaire pouvant améliorer considérablement la qualité du recalage. Grâce à l'information couleur, nous avons pu développer une méthode de recalage possédant un critère discriminant robuste. Par ailleurs, nous avons montré l'importance du choix de l'espace couleur utilisé et son impact sur la qualité des appariements. Pour pallier le problème de l'influence du changement d'illumination sur la robustesse de l'appariement, nous avons utilisé un espace couleur perceptuel qui nous permet de séparer les composantes de chromaticité de la composante d'intensité. Ces composantes sont pondérées de façon à annuler l'effet du changement d'illumination. L'espace choisi, pour cela, est l'espace YIQ. Les résultats que nous avons obtenus avec cette nouvelle méthode confirment l'intérêt d'un choix judicieux de l'espace couleur et montrent son impact sur la qualité du recalage.

6.1.2 Automatisation du processus de recalage

La plupart des logiciels de recalage itératif sont initialisés manuellement par l'opérateur. Les rares méthodes d'initialisation proposées pour résoudre le problème de manière automatique ne nous ayant pas convaincus, nous avons cherché à proposer une solution exploitable chaque fois qu'une information couleur suffisamment riche est présente dans les données (images texturées). En effet, pour éviter que le processus

de recalage converge vers un minimum local, il est nécessaire d'initialiser ICP en lui fournissant une transformation de repère proche de la solution exacte. Pour estimer cette transformation de manière automatique, nous avons proposé une méthode robuste basée sur la recherche automatique de paires de points couleur caractéristiques dans les images 3-D/couleur à recaler. Notre méthode fonctionne en deux étapes principales :

- la première étape a pour but de fournir une liste initiale d'appariements basée sur la mise en correspondance des images 2-D couleur correspondant aux nuages de points 3-D/couleur à recaler. Nous avons d'abord extrait des points d'intérêt couleur avec le détecteur de Harris Précis couleur dans les deux images à appairer. Ensuite, nous avons calculé pour ces points leurs vecteurs d'invariants différentiels couleur. Les points d'intérêt des deux images sont mis en correspondance par comparaison de leurs vecteurs d'invariants. Pour améliorer la robustesse de la mise en correspondance des images 2-D couleur vis-à-vis des changements de luminosité, nous avons normalisé les couleurs des images 2-D.
- la deuxième étape de notre méthode d'initialisation est basée sur l'algorithme RANSAC que nous avons appliqué aux paires 3-D correspondant aux appariements 2-D obtenues à l'issue de l'appariement couleur de la première étape. La méthode RANSAC élimine les mauvais appariements contenus dans cette liste de paires 3D, et fournit une estimation grossière de la transformation entre les deux images. C'est la transformation initiale recherchée.

La transformation initiale, obtenue par notre méthode d'initialisation, est affinée ensuite par ICP. Nous avons, ainsi, été capable de procéder à un recalage entièrement automatique de données réelles sans aucune initialisation manuelle.

6.2 Perspectives

Le travail présenté dans cette thèse apporte des contributions dans le domaine du recalage de données 3-D. Notamment, l'utilisation de l'information photométrique pour le recalage et la proposition d'une solution de recalage automatique exploitant l'information de texture. De nombreuses améliorations peuvent être apportées à ces travaux. La première concerne l'amélioration de la robustesse de l'algorithme ICP développé face à la présence de mauvais appariements. Il est nécessaire, également, améliorer la phase d'initialisation automatique en rendant la mise en correspondance 2-D plus robuste aux bruits de mesure et aux changements d'illumination. Par ailleurs, il est souhaitable de diminuer les temps d'exécution des différentes procédures développées en optimisant les phases de recherche du plus proche voisin.

6.2.1 Amélioration du recalage de deux nuages 3D/Couleur

Nous avons, dans le cadre d'autres travaux, utilisé, pour le recalage de deux nuages de points 3-D, une distance de Mahalanobis pour la phase de mise en correspondance ainsi qu'un estimateur par moindres carrés pondérés [DOU01]. Nous

avons, alors, démontré l'efficacité de la distance de Mahalanobis, qui a pour particularité de prendre en considération les différents bruits de mesure dans la matrice de covariance qui sert au calcul de cette distance. Cette distance améliore considérablement la qualité des appariements et la robustesse de la mise en correspondance en présence de faux appariements. Par ailleurs, l'utilisation d'un estimateur robuste (par moindres carrés pondérés) permet, également, de réduire, de façon itérative, l'effet des faux appariements sur le processus d'estimation de la transformation de repère. Nous n'avons pas utilisé cette approche dans le travail de cette thèse par souci de rapidité d'exécution. En effet, le calcul de la distance de Mahalanobis fait intervenir les matrices de covariances des bruits de mesures.

Nous pouvons, également, envisager d'utiliser l'information couleur pour sélectionner les points avant l'appariement, de façon à diminuer réellement, dès le départ, la masse de données à traiter [DRU06]. Ainsi, en mettant grossièrement, en correspondance les centres de régions couleur de caractéristiques similaires il serait possible d'initialiser approximativement la transformation de repère.

Lorsque la forme de l'objet à numériser présente davantage de zones d'intérêt que la texture (peu de points d'intérêt dans l'image couleur), on peut envisager d'utiliser des points d'intérêt 3-D. Ces points seront appariés en utilisant des vecteurs d'attributs 3D tels que les courbures [SIS00, GOD01], les normales, etc.

Dans ICPGCT, il serait intéressant d'évaluer d'autres espaces couleur et leur impact sur la réduction de l'influence des changements d'illumination.

6.2.2 Recalage de N vues 3D/Couleur

L'objectif du recalage de N vues est la construction d'un modèle complet de l'objet à numériser. Nous pouvons envisager d'utiliser les approches développées dans le travail présenté dans cette thèse pour recalibrer les N vues paire par paire (recalage global), ou alors adapter ces approches (l'utilisation de la couleur et l'automatisation du recalage) à des méthodes de recalage global.

6.2.3 Intégration des données pour la construction d'un modèle surfacique 3D/C

L'objectif à plus long terme de notre travail concerne la reconstruction d'un modèle surfacique précise de l'objet à numériser. Ceci nécessite entre autre de travailler sur la simplification des données dans les zones de recouvrement des nuages de points et sur l'ajout de données 3D/Couleur manquantes dans les zones non accessibles par le capteur et ce après la phase du recalage (local ou global).

BIBLIOGRAPHIE

- [ARU87] K. Arun, T. Huang, S. Blostein, "Least Squares Fitting of two 3D Point Sets", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 9, n° 5, pp. 698-700, 1987.
- [ASA86] Asada H, Brady M, "The Curvature Primal Sketch", *IEEE PAMI*, n° 8, pp. 2-14, 1986.
- [BAU96] Bauckhage C, Schmid C, "Evaluation of Keypoints Detectors", *Rapport technique INRIA*, 1996.
- [BEA78] Beaudet P.R, "Rotationally Invariant Image Operators", *International Joint Conference on Pattern Recognition*, pp. 579-583, 1978.
- [BEN04] Bendels G.H, Degener P, Wahl R, Körtgen M, Klein R, "Image-Based Registration of 3D-Range Data Using Feature Surface Elements", *VAST*, pp. 115-124, 2004.
- [BEN02] Ben-Jemaa R, Schmitt F, "Recalage 3D", *Traité IC2 "Information, Commande, Communication"*, Volume "Images de Profondeur", Ed. HERMES, ISBN : 2-7462-0316-2, 2002, 2002, pp127-180.
- [BEN79] Bentley J.L, Friedman J.H, "Data Structures for Range Searching", *ACM Computing Surveys*, 11(4), 1979.
- [BER96] Bergevin R, Soucy M, Gagnon H, Larendeau D, "Towards a General Multi-View Registration Technique", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 18, n° 5, pp. 540-547, may 1996
- [BER84] Berzins V, "Accuracy of Laplacian Edge Detector", *Computer Vision, Graphics and Image Processing*, n° 27, pp. 195-210, 1984.
- [BES92] Besl P.J, McKay N.D, "A Method for Registration of 3D Shapes", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 14, n° 2, pp.239-256, 1992
- [BLA95] Blais G, Levin M.D, "Registering Multiview Range Data to Create 3D Computer Objects", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 17, n° 8, pp. 820-824, 1995.
- [BLA94] Blaszkas T, Deriche R, "Recovering and Characterizing Image Features using an Efficient Model Based Approach", *Rapport Technique, INRIA*, n° 2422, 1994.

- [BOV00] Bovik A, "Handbooks of : Image and Video Processing", Academic Press, 2000.
- [BRE1] <http://www.breuckmann.com/index.php?id=methodology&L=2>
- [BRE2] <http://www.breuckmann.com/index.php?id=tritos&L=2>
- [BRU99] R. Brunelli, O. Mich, "On the Use of Histograms for Image Retrieval", *Proceedings of ICMCS*, vol. 2, pp. 143-147, 1999.
- [CANTZ] Random Sample Consensus (RANSAC),
H.Cantzler,[http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/CANTZLER2/ransac.pdf]
- [CHE91] Y. Chen, G.Medioni, "Object Modelling by Registration of Multiple Range Views", *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 2724-2729, April 1991
- [CHE92] Y. Chen, G.Medioni, "Object Modelling by Registration of Multiple Range Images", *Image and Vision Computing*, Vol. 10, n° 3, pp. 145-155, 1992.
- [CHU99] Chu-Song Chen, Yi-Ping Hung, Jen-Bo Cheng, "RANSAC-based DARCES: a new approach to fast automatic registration of partially overlapping range images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 21, n° 11, Nov 1999, Page(s):1229-1234.
- [CHU96] Chua C.S, "3-D Free-Form Surface Registration and Object Recognition", *International Journal of Computer Vision*, n° 17, pp. 77-99, 1996.
- [COT94] Cottier J.C, "Extraction et Appariement Robustes des Points d'Interet de deux Images non Etalonnees", stage de maitrise, 1994.
- [DER90] Deriche R, Giraudon G, "Accurate Corner Detection: An Analytical Study" *IEEE International Conference on Computer Vision*, pp. 60-70, Osaka, Japan, December 1990.
- [DER93] Deriche R, Giraudon G, "A Computational Approach to Corner and Vertex detection", *International Journal of Computer Vision*, vol. 10, n° 2, pp.101-124, 1993.
- [DOU01] Douadi L, "Recalage de Nuages de Points 3-D Fournis par un Capteur Embarqué", *Rapport de Stage de DEA, LIRMM, Université de Montpellier2*, 2001.
- [DOU06] Douadi Lounis, Aldon Marie-José, Crosnier André, "Pair-Wise Registration of 3D Color data Sets with ICP", dans *Proceedings IEEE IROS*, Beijing, 2006.

- [DRE82] Dreschler L, Nagel H-H, "Volumetric Model and 3-D Trajectory of a Moving Car Derived from Monocular TV-frame Sequence of a Street scene", *Computer Vision, Graphics and Image Processing*, n° 20, vol. 3, pp. 199-228, 1982.
- [DRU06] Druon S, Aldon M-J, Crosnier A, "Color Constrained ICP for Registration of Large Unstructured 3D Color Data Sets", *IEEE International Conference on Information Acquisition (ICIA'06)*, pp. 20-23, August 2006, Shandong, China, pp. 249-255.
- [FARO] www.faro.com.
- [FAU86] Faugeras O, Hebert M, "The Representation, Recognition and Localisation of 3D Objects", *International Journal of Robotics Research*, Vol. 5, n° 3, pp. 27-51, 1986.
- [FIS81] Fischler M.A, Bolles R.S, "Random Sample Consensus: A Paradigm For Model Fitting With Applications to Image Analysis and Automated Cartography", *Communications of the ACM*, Vol. 24, no. 6, June 1981.
- [FIT01] Fitzgibbon A.W, "Robust Registration of 2D and 3D Point Sets", *Proceedings, British Machine Vision Conference*, pp. 411-420, 2001.
- [FOR86] Förstner W, Gülch E, "A Fast Operator for Detection and Precise Location of Distinct Points, Corners and Centers of Circular Features", *Proceedings of the ISPRS Workshop on Fast Processing of Photogrammetric Data, Interlaken, Switzerland*,
- [FUN95] Funt B, Finlayson G, "Color Constant Color Indexing", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, n° 17, pp. 522-529.
- [GEL03] Gelfand N, Ikemoto L, Rusinkiewicz S, Levoy M. "Geometrically Stable Sampling for the ICP Algorithm". *Fourth International Conference on 3D Digital Imaging and Modeling (3DIM)*, pp 260-267. October 2003.
- [GEL05] Gelfand, N, Mitra, N.J, Guibas, L.J, and Pottmann H, "Robust global registration", dans *SGP'05*, pp. 197-206, 2005.
- [GOD01] Godin G, Laurendeau D, Bergevin R, "A method for the registration of Attributed Range Images", *3DIM*, pp. 179-186, Québec, 2001.
- [GOT92] Gottesfeld Brown L, "A Survey of image Registration Techniques", *ACM, Computing Surveys*, n° 24, pp. 326-376, 1992.
- [GOU00] Gouet V, "Mise en Correspondance d'Images en Couleur", Thèse de Doctorat, *LIRMM, Université de Montpellier2*, 2000.

- [HAR88] Harris C. G, Stephens M, "A combined Corner et Edge Detector", *4th Alvey vision Conference*, pp. 147-151, 1988.
- [HEI92] Heitger F, Rosenthaler L, von der Heydt R, Peterhans E, Kübler O, "Simulation of Neural contour Mechanism: From Simple to End-Stopped Cells", *Vision Research*, n° 32, vol. 5, pp. 963-981, 1992.
- [HOR87] Horn B.K.P, "Closed-Form Solution of Absolute Orientation Using Unit Quaternions," *Journal of the Optical Society of America*, vol. 4, no. 4, pp. 629-642, April 1987.
- [HOR88] B. Horn, H. Hilden, S. Negahdaripour. Closed-form solution of absolute orientation using orthonormal matrices. *J. Opt. Soc. Am. A*, 5:1127--1135, July 1988.
- [HOR90] Horaud R, Skordas T, veillon F, "Finding Geometric and Relational Structures in an Image", dans *Proceedings of the first European Conference on Computer Vision*, pp. 374-384, 1990.
- [HUA99] Huang J, Kumar S.R, Mitra M, Zhu W.J, Zabih L, "Spatial Color Indexing and Applications", *Internationale Journal Of Computer Vision*, n° 35, vol. 3, pp. 245-268.
- [HUT93] Huttenlocher D.P, Klanderman G.A, Rucklidge W.J, "Comparing Images Using the Hausdorff Distance", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, n° 9, vol. 15, pp. 850-863, Septembre 1993.
- [JEO01] Jeong S, Won C.S, Gray R.M, "image Retrieval Using color Histograms Generated by Gauss Mixture vector Quantization", *CVIU*, n° 1, vol. 3, pp. 44-66, April-June 2004.
- [JOH96] Andrew Johnson, Sing Bing Kang, "Registration and Integration of Textured 3D Data", *Cambridge Research Lab, Technical Research Series*, CRL 96/4, October 1996.
- [JOS02] Timothée Jost, "Fast Geometric Matching for Shape Registration", *Thèse, Faculté des Sciences de l'Université de Neuchâtel*, 2002.
- [JUN01] Jung K, Lacroix S, "A Robust Interest Points Matching Algorithm", *8th International Conference on Computer Vision (ICCV'01)*, Vancouver (Canada), 9-12 Juillet 2001, vol. 2, pp.538-543.
- [KIM02] Sang-Hoon Kim, Jung-Kak Seo, Hyun-Ki Hong, Min-Hyung Choi, "Iterative Registration of Multiple 3D Data Sets Using Covariance Matrix" *Proceedings of International Conference on Virtual Systems and MultiMedia*, Sept. 2002.
- [KIM04] Seung-Hwan Kim, Dong-O Kim, Sang Wook Lee, Rae-Hong Park, "Reducing Computation Time for Range Image Registration Using Radial-

- Distance Down-Sampling", *Proceedings. 10th Korea-Japan Joint Workshop Frontiers of Computer Vision FCV 2004*, pp. 275-280, Kyushu University, Fukuoka, Japan, Feb. 2004.
- [KIT82] Kitchen L, Rosenfeld A, "Grey Level Corner Detection", *Pattern Recognition Letters*, pp. 95-102, 1982.
- [LAV95] Lavallée S, Szeliski R, "Recovering the Position and the Orientation of Free-Form Objects from Image Contours Using 3D Distance Maps", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 17, n° 4, pp. 378-390, 1995.
- [LUC00] Jason Luck, Charles Little, William Hoff, "Registration of Range Data Using a Hybrid simulated Annealing and Iterative Closest Point Algorithm", *IEEE International Conference on Robotics and Automation*, San Francisco, pp. , 2000.
- [MAK06] Makadia A, Patterson A, Daniilidis K, "Fully Automatic Registration of 3D Point Clouds", *IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 1297-1304, 2006.
- [MAL05] Malis E, Marchand E, "Méthodes robustes d'estimation pour la vision robotique". *Journées nationales de la recherche en robotique, JNRR'05*, Guidel, France, Octobre 2005.
- [MAS95] Takeshi Masuda, Naokazu Yokoya, "A Robust Method for Registration and Segmentation of Multiple Range Images", *Computer Vision And Image Understanding*, Vol. 61, n° 3, pp. 295-307, May 1995.
- [MED86] Medioni G, Yasumoto Y, "Corner Detection and Curve Representation Using B-splines", *International Conference on Robotics and Automation*, pp. 764-769, 1986.
- [MON98] Montesinos P, Gouet V, Deriche R, "Differential Invariants for Color Images", dans *Proceedings of 14th International Conference on Pattern Recognition*, Brisbane, Australie, 1998.
- [MON00] Montesinos P, Gouet V, Deriche R, Pelé D, "Matching Color Uncalibrated Images Using Differential Invariants", *Image and Vision Computing*, n° 18, vol. 9, pp. 659-672.
- [MOR77] Moravec H.P, "Towards automatic visual obstacle avoidance", *Proceedings of the 5th Int. Joint Conference Artificial Intelligence*, vol. 2, pp. 584, August 1977
- [MOR96] MORON V, "Mise En Correspondance de Données 3-D Avec Un Modèle CAO, Application à L'inspection Automatique", *thèse INSA Lyon*, 1996.

- [NAG83] Nagel H. H, "Constraints for the Estimation of Displacement Vector Fields from Image Sequences", *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 945-951, Karlsruhe, West Germany, August 1983.
- [OSU] OHIO State University, Range Image Data Base,
<http://sampl.ece.ohio-state.edu/data/3DDB/RID/minolta/>.
- [PUL99] Kari Pulli, "Multiview Registration for Large Data Sets". *3DIM*, pp. 160-168, Ottawa, 1999.
- [ROH92¹] Rohr K, "Modelling and identification of Characteristic Intensity Variations", *Image and Vision Computing*, vol. 10, pp. 66-76, 1992
- [ROH92²] Rohr K, "Recognizing Corners by fitting Parametric Models", *International journal of computer vision*, n° 9, vol. 3, pp. 212-230, 1992.
- [RUS01] Szymon Rusinkiewicz, Mark Levoy, "Efficient Variant of the ICP Algorithm". *3DIM*, pp. 145-152, Québec, 2001.
- [SAL98] Salvi J, Batlle J, Mouaddib E-M, "A robust-coded pattern projection for dynamic 3D scene measurement", *Pattern Recognition Letters*, n° 19, pp. 1055-1065, 1998.
- [SCH97] Schiele B, "Reconnaissance d'Objets Utilisant des Histogrammes Multidimensionnels de Champs Réceptifs", *Thèse de Doctorat, INPG, GRAVIR – IMAG*, 1997.
- [SCH96] Schmid C, "Appariement d'Images par Invariants Locaux de Niveaux de Gris", *Thèse de Doctorat, INPG, France*, 1996.
- [SEE02] Seeger S, Laboureaux X, "Feature extraction and registration: An overview". *Principles of 3D Image Analysis and Synthesis*, pp. 153–166, 2002.
- [SHA84] Shah M. A, Jain R, "Detecting Time Varying Corners", *Computer Vision, Graphics and Image Processing*, n° 28, vol. 3, pp. 345-355, 1984.
- [SHA02] Gregory C. Sharp, Sang W. Lee and David K. Wehe, "ICP Registration Using Invariant Features", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, n° 1, pp. 90-102, 2002.
- [SIS00] SISTIAGA M, "Navigation référencée images de terrain pour engins sous-marins", *Thèse de Doctorat, Université de Montpellier2, LIRMM*, 2000.
- [STE02] Andrew Neil Stein, "Modeling Real-World Object from Noisy Data Using ICP", december 3, 2002.

- [STE92] Stein, F, Medioni, G, "Structural Indexing: Efficient 3-D Object Recognition," ", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 14, n° 2, pp. 125 - 145, 1992.
- [SWA91] Swain M, Ballard D, "Color Indexing", *International Journal of Computer Vision*, n° 7, vol. 1, pp. 11-32.
- [THI96] Thirion J.P, "New feature points based on geometric invariants for 3d image registration", *International Journal of Computer Vision*, 18(2):121 {137, May 1996.
- [VIO97] Viola P, Wells W.M, "Alignment by Maximization of Mutual Information", *International Journal of Computer Vision*, n° 24, pp.137-154, 1997.
- [WAL91] M. Walker, L. Shao, R. Volz, "Estimating 3D Location Parameters Using Dual Number Quaternions", *CVGIP: Image Understanding*, Vol.5, No. 3, 1991.
- [ZHA94] Zhengyou Zhang, "Iterative Point Matching for Registration of Free Form Curves", *IJCV*, vol. 13, n° 2, pp 119-152, 1994.
- [ZHA94⁺] Zhang Z, Deriche R, Faugeras O, Luong Q.T, "A Robust Technique for Matching Two Uncalibrated Images Through the Recovery of the Unknown Epipolar Geometry". *Rapport de Recherche INRIA*, n° 2273, May 1994.
- [ZUN83] Zuniga O. A, Haralick R. M, "Corner detection using the facet model". *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 30-37, 1983.

Résumé :

Dans les applications de modélisation de l'environnement, on dispose généralement de plusieurs images 3D de la même scène prises de points de vue différents. Le recalage de ces vues consiste à estimer les transformations rigides qui permettent de les ramener dans un référentiel commun. Cette thèse propose des solutions pour réaliser le recalage automatique de nuages de points denses et non structurés acquis par un scanner 3D/couleur à haute résolution, en vue de la construction de modèles 3D texturés d'objets complexes. La première contribution de ce travail concerne l'utilisation de l'information couleur pour améliorer les performances du recalage en utilisant l'algorithme ICP (Iterative Closest Point) initialement conçu pour le traitement de données géométriques pures. Lorsque le niveau du bruit sur la couleur n'est pas trop élevé, l'utilisation de cette information photométrique permet d'améliorer la convergence de l'algorithme et de réduire l'erreur résiduelle de recalage dans l'espace 3D. Une étude concernant la robustesse de la méthode au bruit sur la couleur montre qu'on peut diminuer l'effet négatif des changements d'illumination en choisissant un espace couleur perceptuel qui permet de séparer les composantes chromatiques de la composante intensité. La deuxième contribution de ce travail concerne l'automatisation du processus de recalage. En effet, pour éviter que l'algorithme itératif de recalage ne converge vers un minimum local, il est nécessaire de lui fournir une transformation initiale proche de la solution exacte. Pour estimer cette transformation de manière automatique on recherche un ensemble de paires de points couleur dans les deux nuages à appairer. Ces paires sont construites à partir de points d'intérêt extraits des images couleur à l'aide du détecteur de Harris Précis Couleur. La transformation 3D initiale est ensuite estimée en appliquant l'algorithme RANSAC aux paires 3D résultant de l'appariement précédent. Cette méthode permet d'éliminer l'influence des mauvais appariements couleur et fournit généralement une initialisation correcte du processus de recalage. Les tests effectués sur des images réelles illustrent les performances et la faisabilité des solutions proposées.

Titre (en anglais) :

Solutions for registering 3D-color point sets

Résumé (en anglais) :

Creating a 3D object model requires generally the acquisition of multiple partially overlapping range images captured with different sensor locations. The 3D registration of these images that are assumed to be pair-wisely overlapping consists in estimating the rigid transformation that places them in a common coordinate system. This thesis presents solutions for automatically registering two unstructured 3D-color point sets provided by a high-resolution scanner. Our first aim was to improve performance of classical ICP (Iterative Closest Point) approaches by taking into account colour information for point matching. When colour noise is not too important, using photometric information allows us to improve the convergence and to decrease the residual error of the registration process. Moreover, it is possible to reduce the drawbacks of intensity variations by choosing a perceptual colour space where intensity and chromatic components are separated. The second aim of this work was to develop a fully automatic registration method, which does not require any pose measuring hardware or manual intervention. In fact, ICP requires an a-priori knowledge of an approximate estimate of the rigid transformation. This has been solved using a set of interest point pairs extracted from the two colour images with the Harris accurate detector. The initial 3D transform is then estimated by applying the RANSAC algorithm to the set of 3D point pairs issued from the previous colour matching. This randomised estimator is robust to outliers and provides a rough estimation that is enough accurate to initialise the iterative registration process. Experimental results show the capabilities of the proposed solutions with colour range images of real-world objects.

Mots-Clés :

Modélisation 3D, images 3D-couleur, recalage automatique, mise en correspondance d'images.

Discipline :

Génie Informatique, Automatique et Traitement du Signal

Laboratoire d'Informatique, de Robotique et de Micro-électronique de Montpellier (LIRMM)
UMR CNRS/Université Montpellier II, n° C55060
161, Rue Ada, 34392-MONTPELLIER – Cedex 5 – France.