



**HAL**  
open science

# Alignement paramétrique d'images : proposition d'un formalisme unifié et prise en compte du bruit pour le suivi d'objets

Jean-Baptiste Authesserre

## ► To cite this version:

Jean-Baptiste Authesserre. Alignement paramétrique d'images : proposition d'un formalisme unifié et prise en compte du bruit pour le suivi d'objets. Traitement du signal et de l'image [eess.SP]. Université Sciences et Technologies - Bordeaux I, 2010. Français. NNT : . tel-00672589

**HAL Id: tel-00672589**

**<https://theses.hal.science/tel-00672589>**

Submitted on 21 Feb 2012

**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.

N° d'ordre : 4136

# THÈSE

PRÉSENTÉE À

L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE DES SCIENCES PHYSIQUES ET DE L'INGÉNIEUR

pour obtenir le titre de

**DOCTEUR**

SPÉCIALITÉ : AUTOMATIQUE ET PRODUCTIQUE, SIGNAL ET IMAGE

Présentée et soutenue par

Jean-Baptiste AUTHESSERRE

## Alignement paramétrique d'images : proposition d'un formalisme unifié et prise en compte du bruit pour le suivi d'objets

Thèse dirigée par Yannick BERTHOUMIEU et Rémi MÉGRET

préparée à l'IMS UMR 5218, Groupe SIGNAL ET IMAGE

Soutenue le 02 Décembre 2010

**Devant la commission d'examen formée de :**

Philippe RÉFRÉGIER	Professeur à l'Université d'Aix-Marseille	<i>Président</i>
Michel BERTHIER	Professeur à l'Université de La Rochelle	<i>Rapporteur</i>
Ezio MALIS	HDR, Président de Robocortex S.A.S.	<i>Rapporteur</i>
Mohamed NAJIM	Professeur à l'ENSEIRB-MATMECA	<i>Examinateur</i>
Yannick BERTHOUMIEU	Professeur à l'ENSEIRB-MATMECA	<i>Examinateur</i>
Rémi MÉGRET	Maître de Conférences à l'ENSEIRB-MATMECA	<i>Examinateur</i>



## Remerciements

L'aboutissement d'une thèse est rarement le fruit du travail d'une personne, mais plutôt le résultat d'une synergie au sein d'une équipe. L'heure est enfin venue de rendre à César ce qui est à César.

Durant ces trois dernières années, j'ai été porté par deux esprits : le professeur Yannick Berthoumieu et le Docteur Rémi Mégret. Ils ont su me hisser vers le monde des Idées alors que, les pieds fermement ancrés au sol, je regardais le ciel... Je les remercie chaleureusement pour l'attention qu'ils ont portée à mes travaux.

C'est parfois dans les situations les plus improbables que la science fait son brin de chemin. C'est peut-être devant la machine à café du sixième étage qu'ont été levés les verrous de ce travail de thèse. Je remercie l'ensemble des doctorants, permanents et personnels administratifs qui ont croisé ma route devant cette merveille de technologie, et avec qui j'ai pu débattre des mystères de la science.

Je souhaite également remercier mes parents. Ils ont su m'inculquer l'importance de l'instruction et m'ont permis de faire ces longues études. Encore aujourd'hui cela est une chance dont je suis conscient.

Le manuscrit de thèse est un plat copieux. J'espère que sa digestion n'a pas été pesante pour le professeur Michel Berthier et le docteur Ezio Malis, qui ont accepté de rapporter mes travaux, ainsi que pour le professeur Philippe Réfrégier, qui a bien voulu présider mon jury de thèse. Les remarques qu'ils ont formulé m'ont permis de prendre du recul sur mon travail. Merci à eux.

Je tiens également à remercier le professeur Mohamed Najim pour m'avoir accueilli dans l'équipe Signal et Images et pour avoir accepté d'être examinateur de mes travaux.

Ma dernière pensée va au Duc qui, bien loin des problèmes de convergence (au sens large), s'abandonne sur les hauteurs blanches et brumeuses de la Russie.





# Résumé

L'alignement paramétrique d'images est une technique fréquemment utilisée dans de nombreuses applications comme la réalité augmentée, la compression vidéo, la super-résolution ou encore le suivi d'objets. Il s'agit de recalibrer plusieurs images entre elles, en compensant les déplacements en fonction d'un modèle paramétrique, de façon à aligner les contenus visuels. Parmi les nombreuses approches de la littérature, nous nous intéressons notamment à la technique de recalage d'imagette (*template matching*) particulièrement adaptée pour le suivi d'objets. Étudié depuis près de trente ans, le recalage d'imagette a donné naissance à de nombreux algorithmes.

Notre première contribution consiste en l'élaboration d'un nouveau formalisme bidirectionnel permettant d'unifier les algorithmes existants et utilisant une optimisation locale. Ainsi nous proposons notamment une classification des approches de l'état de l'art reposant sur les concepts de paramétrisation incrémentale et de classes d'équivalence des paramètres bidirectionnels. Les approches existantes sont réinterprétées sous la forme de quatre familles de méthodes à savoir Directes, Inverses, Symétriques et Bidirectionnelles qui étendent la taxonomie Forward (Directe) et Inverse couramment utilisée dans le domaine. Ce cadre théorique nous amène de plus à porter un éclairage nouveau sur un certain nombre d'approches de la littérature et à disposer d'une base solide facilitant la définition de nouvelles approches. Il en découle un premier résultat : l'équivalence au premier ordre entre les approches d'une même famille. Dans le cadre de la méthode d'optimisation Gauss-Newton, un second résultat est la mise en évidence des relations entre des approches reposant sur l'optimisation de critères a priori hétérogènes, notamment entre différents types d'approches symétriques et l'algorithme ESM. Ce formalisme permet enfin de définir une méthodologie de conception simple pour la création de nouveaux algorithmes de recalage d'images par optimisation locale.

Notre deuxième contribution est de proposer une étude théorique approfondie de l'influence du bruit sur le processus d'alignement. Deux nouveaux critères d'alignement recouvrant les principales méthodes existantes sont définis. L'approche ACL (Asymmetric Composition on Lie Group) est une extension des approches de l'état de l'art qui introduit un coefficient permettant de pondérer l'utilisation des gradients des deux images lors de l'alignement. Plusieurs heuristiques sont proposées pour l'estimation automatique de ce coefficient en fonction de la qualité des images en présence. Une étude de ce nouveau critère apporte des éléments nouveaux pour confirmer l'intérêt des approches symétriques en présence de bruits de RSB (Rapport Signal sur Bruit) identiques pour chaque image et révèle qu'en présence d'images ayant une forte asymétrie en terme de niveau de bruit, la plupart des approches de l'état de l'art produisent des résultats sous-optimaux. L'approche BCL (bidirectional Composition On Lie Group) formalise le problème d'alignement de manière moins contrainte que les approches de l'état de l'art en laissant à l'algorithme le soin de choisir dans quel repère les images sont alignées. Ceci a pour effet de permettre à l'algorithme de pondérer de façon adaptée la contribution des gradients de chaque image. L'étude théorique des approches ACL et BCL permet d'identifier l'intérêt de ces dernières notamment en présence de niveaux de bruits asymétriques affectant les images. Une étude expérimentale sur données synthétiques confirme les résultats théoriques obtenus. Du point de vue des performances géné-

---

rales, elle révèle que les nouvelles méthodes produisent des résultats comparables aux approches de l'état de l'art dans le pire des cas avec une nette amélioration des performances en présence de bruits de niveaux importants et de niveaux différents sur les images. Ce constat est validé lors de la mise en place d'une procédure de suivi d'objets sur des données semi synthétiques simulant l'acquisition de vidéos dans des conditions de faible luminosité (low light).

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Contexte . . . . .	1
1.2	Motivation . . . . .	3
1.2.1	Imagerie à faible rapport signal sur bruit . . . . .	3
1.2.2	Suivi paramétrique . . . . .	3
1.2.3	Problématique . . . . .	4
1.3	Objectifs de la thèse . . . . .	5
<b>2</b>	<b>Alignement paramétrique d'images</b>	<b>7</b>
2.1	Généralités . . . . .	7
2.1.1	Position du problème . . . . .	7
2.1.2	Modèles 2D de transformations globales géométriques . . . . .	8
2.1.3	Critère d'alignement . . . . .	14
2.2	Alignement par minimisation en ligne d'une fonctionnelle d'erreur . . . . .	18
2.2.1	Principes généraux de la minimisation d'une fonctionnelle d'erreur . . . . .	19
2.2.2	L'algorithme de Lucas et Kanade . . . . .	21
2.2.3	Approches à faible temps de calcul . . . . .	22
2.2.4	Modification du critère d'alignement pour être robuste à des perturbations photométriques . . . . .	27
2.3	L'alignement d'images, un problème asymétrique? . . . . .	30
2.3.1	Asymétrie lors de la formalisation du problème d'alignement . . . . .	31
2.3.2	Asymétrie lors de la résolution du problème d'alignement . . . . .	32
2.4	Conclusion . . . . .	34
<b>3</b>	<b>Formalisation bidirectionnelle du problème d'alignement géométrique</b>	<b>35</b>
3.1	Formalisation idéale . . . . .	35
3.1.1	Formalisation usuelle . . . . .	35
3.1.2	Formalisation bidirectionnelle . . . . .	36
3.2	Formalisation pratique . . . . .	36
3.2.1	Définition d'une région d'intérêt comme support d'alignement . . . . .	36
3.2.2	Définition d'une mesure de similarité . . . . .	37
3.3	Contrainte sur l'espace bidirectionnel . . . . .	40
3.3.1	Taxonomie . . . . .	40
3.3.2	Nature de la paramétrisation incrémentale . . . . .	41
3.4	Mise à jour des paramètres . . . . .	44
3.4.1	Règle de mise à jour et applications . . . . .	45
3.4.2	Convergence . . . . .	48
3.5	Etapes de conception d'un algorithme d'alignement d'images . . . . .	48

<b>4</b>	<b>Application du formalisme Bidirectionnel aux approches existantes</b>	<b>51</b>
4.1	Application de l'algorithme Gauss-Newton aux approches SSD . . . . .	52
4.2	Classification des approches de l'état de l'art : méthodes unidirectionnelles . . . . .	53
4.2.1	Les approches Directes (F) (Forward) . . . . .	53
4.2.2	Les approches Inverses (I) . . . . .	54
4.3	Classification des approches de l'état de l'art : méthodes bidirectionnelles . . . . .	57
4.3.1	Approches Symétriques Compositionnelles Exactes (SCE) . . . . .	57
4.3.2	Approches approximativement symétriques . . . . .	61
4.3.3	Approches Bidirectionnelles Pures . . . . .	62
4.4	Équivalence des algorithmes d'une même famille dans le cas d'une optimisation par l'algorithme Gauss-Newton . . . . .	66
4.4.1	Hypothèses . . . . .	66
4.4.2	Fondements de l'équivalence . . . . .	66
4.4.3	Équivalence des approches au sein d'une même famille . . . . .	67
4.4.4	Contraintes sur l'espace bidirectionnel et règle de mise à jour . . . . .	71
4.5	Conclusions . . . . .	72
4.5.1	Bilan . . . . .	72
4.5.2	Conception d'approches pour le suivi d'objet . . . . .	72
<b>5</b>	<b>Composition sur groupe de Lie pour l'alignement d'images de qualités diffé-</b> <b>rentes</b>	<b>75</b>
5.1	Composition asymétrique sur groupe de Lie . . . . .	75
5.1.1	Analyse théorique de l'erreur ACL . . . . .	77
5.1.2	Minimisation de variance pour l'estimation de $\alpha$ . . . . .	79
5.1.3	ACL Géométrique pour l'estimation de $\alpha$ . . . . .	80
5.1.4	ACL Analytique pour l'estimation de $\alpha$ (AACL) . . . . .	81
5.2	Composition Bidirectionnelle sur groupe de Lie . . . . .	83
5.2.1	Propriété d'invariance . . . . .	83
5.2.2	L'algorithme BCL projeté (PBCL) . . . . .	84
5.2.3	Analyse théorique de l'influence du bruit . . . . .	85
5.3	Résumé des nouvelles approches et approximations efficaces . . . . .	86
<b>6</b>	<b>Validation expérimentale</b>	<b>91</b>
6.1	Performances des approches ACL sur données synthétiques . . . . .	91
6.1.1	Présentation du banc de test . . . . .	91
6.1.2	Comportement des algorithmes ACL à $\alpha$ fixe en présence de bruit . . . . .	92
6.1.3	Influence de l'asymétrie du bruit sur les approches ACL . . . . .	94
6.2	Performances de l'approche BCL sur données synthétiques . . . . .	99
6.2.1	Équivalence des approches BCL et PBCL . . . . .	99
6.2.2	Impact de la projection pour les approches BCL et PBCL . . . . .	99
6.2.3	Influence de l'asymétrie des niveaux de bruit : BCL vs ACL . . . . .	100
6.3	Application au suivi dans des conditions de faible luminosité . . . . .	103

6.3.1	Illustration de suivi sur des données réelles . . . . .	103
6.3.2	Analyse moyennée sur des séquences réelles . . . . .	105
6.3.3	Analyse statistique sur des séquences semi-synthétiques . . . . .	111
<b>7</b>	<b>Conclusion</b> . . . . .	<b>115</b>
7.1	Bilan . . . . .	115
7.2	Intérêt . . . . .	115
7.3	Perspective . . . . .	116
<b>A</b>	<b>Définition d'une erreur pour l'approche IA</b> . . . . .	<b>117</b>
<b>B</b>	<b>Procédure de traduction des gradients de Brooks</b> . . . . .	<b>121</b>
B.1	Principe général . . . . .	121
B.2	Équivalence à l'ordre 1 de la règle de mise à jour de l'approche GESM avec la règle de mise à jour de l'approche FA+IAD . . . . .	122
<b>C</b>	<b>Équivalence des approches au sein d'une même famille</b> . . . . .	<b>125</b>
C.1	Approches Directes (F) . . . . .	125
C.2	Approches Inverses (I) . . . . .	126
C.3	Approches Bidirectionnelles Pures . . . . .	126
C.4	Approches Symétriques compositionnelles exactes . . . . .	127
C.5	Équivalence à l'ordre 1 de l'approche SCO avec les approches symétriques exactes	129
C.5.1	Relation entre les incréments . . . . .	129
C.5.2	Équivalence entre les approches . . . . .	129
<b>D</b>	<b>Banc de test de Baker et Matthews pour les transformations affines</b> . . . . .	<b>133</b>
<b>E</b>	<b>Combinaison de fonctionnelles</b> . . . . .	<b>135</b>
E.1	Interprétation géométrique . . . . .	135
E.2	Comparaison des approches symétriques compositionnelles et des approches sy- métriques par combinaison pour le suivi d'objets . . . . .	137
<b>F</b>	<b>Approximation à l'ordre 2 de l'erreur ACL</b> . . . . .	<b>141</b>
	<b>Bibliographie</b> . . . . .	<b>145</b>



# Table des figures

2.1	Processus d'alignement d'images. . . . .	8
3.1	Principe général du formalisme bidirectionnel. . . . .	38
3.2	Taxonomie des approches Directe, Inverse et bidirectionnelle. . . . .	41
3.3	Interprétation géométrique des différentes paramétrisations incrémentales. . . . .	43
3.4	Exemple d'alignement pour une approche bidirectionnelle pure avec mise à jour identité et mise à jour directe. . . . .	47
3.5	Illustration de la fonction erreur $E(\boldsymbol{\mu}_I, \boldsymbol{\mu}_T)$ et trajectoires des approches FCL, ICL, SCL et BCL durant l'alignement de deux images. . . . .	49
4.1	Interprétation géométrique de la fonctionnelle <i>Symmetric Composition on Lie Group</i> SCL. . . . .	58
4.2	Fréquence moyenne de convergence et vitesse moyenne de convergence pour les approches des catégories Directe, Inverse, Symétrique et Bidirectionnelle. . . . .	70
4.3	Paramétrisation incrémentale et règle de mise à jour. . . . .	71
5.1	Interprétation géométrique de la formulation asymétrique. . . . .	76
5.2	Interprétation géométrique des différentes approches dans le cas d'une optimisation Gauss-Newton pour un mouvement caractérisé par un paramètre. . . . .	81
5.3	Interprétation générique de la valeur $\alpha^{\text{GACL}}$ . . . . .	82
6.1	Images utilisées pour les expériences sur données synthétiques. . . . .	93
6.2	Fréquence moyenne de convergence des différentes approches ACL relativement au bruit dans l'image pour un Sigma Point valant 12 pixels. . . . .	94
6.3	Fréquence moyenne de convergence en fonction de l'asymétrie du bruit pour les différentes approches. . . . .	96
6.4	Fréquence de convergence moyenne et vitesse de convergence pour les approches ESM, ICL, FCL, $F\text{-AACL}^{\text{ESM}}$ et $F\text{-GACL}$ en présence de bruit sur l'image. . . . .	98
6.5	Fréquence de convergence et vitesse de convergence en échelle logarithmique pour les approches FCL, ICL, SCL, BCL et PBCL utilisant une optimisation Gauss-Newton et une règle de mise à jour Directe. . . . .	101
6.6	Fréquence de convergence des algorithmes ESM, ICL et PBCL comparés aux approches synthétiques $\text{PBCL}_-$ , $\text{PBCL}_\varepsilon$ et $\text{ESM}_{nf}$ pour un Sigma Point égal à 12 et un bruit sur l'image décroissant. . . . .	101
6.7	Fréquence moyenne de convergence en fonction de l'asymétrie du bruit pour les approches ESM, BCL, $\text{AACL}^{\text{FCL}}$ , $\text{AACL}^{\text{ICL}}$ , $\text{AACL}^{\text{ESM}}$ et GACL. . . . .	102
6.8	Exemple d'acquisitions réelles dans des conditions de faible luminosité. . . . .	103
6.9	Illustration de suivi dans des conditions réelles pour les approches FCL, ICL, ESM et BCL sur une vidéo acquise dans des conditions de faible luminosité. . . . .	104



6.10	Vitesse de convergence en termes d'Erreur RMS Point et d'Erreur RMS moyenne en fonction du nombre d'itérations sur une séquence d'images réelles capturée dans des conditions de basse luminosité. Seule la première image est utilisée pour construire le template. . . . .	106
6.11	Fréquence de convergence en fonction du nombre d'itérations sur une séquence d'images réelles capturée dans des conditions de basse luminosité. Seule la première image est utilisée pour construire le template. . . . .	107
6.12	Exemples de templates moyens créés par l'algorithme BCL. . . . .	108
6.13	Vitesse de convergence en termes d'Erreur RMS Point moyenne et d'Erreur RMS moyenne en fonction du nombre d'itérations sur une séquence d'images réelles capturée dans des conditions de basse luminosité. . . . .	108
6.14	Fréquence de convergence en fonction du nombre d'itérations sur une séquence d'images réelles capturée dans des conditions de basse luminosité. Le template est créé en moyennant 3 ou 5 images. . . . .	110
6.15	Fréquence de convergence en fonction du temps de calcul sur une séquence d'images réelles capturée dans des conditions de basse luminosité. . . . .	110
6.16	Images utilisées afin de simuler les conditions de basse luminosité. . . . .	112
6.17	Fréquence de convergence en fonction du seuil $\tau$ sur l'Erreur RMS Point. . . . .	113
D.1	Principe du banc de test de Baker et Matthews. . . . .	133
E.1	Interprétation géométrique de la fonctionnelle <i>Forward Inverse Symmetric Combination</i> FISC. . . . .	137
E.2	Fréquence moyenne de convergence et vitesse moyenne de convergence pour les approches FISCLT et ESM. . . . .	139

# Liste des tableaux

2.1	Principales évolutions de l’algorithme de Lucas et Kanade pour améliorer les performances en termes de temps de calcul. . . . .	26
2.2	Principales fonctionnelles utilisées pour adapter l’algorithme de Lucas et Kanade à des images de différentes natures. . . . .	30
3.1	Propriétés principales des paramétrisations incrémentales. . . . .	43
3.2	Étapes de conception d’un algorithme d’alignement d’images. . . . .	50
4.1	Taxonomie des paramétrisations incrémentales bidirectionnelles et règles de mise à jour directe correspondantes. . . . .	64
4.2	Classification des approches SSD de la littérature. . . . .	65
5.1	Caractéristiques des approches par composition sur groupe de Lie. . . . .	87
5.2	Schéma itératif des approches par composition sur groupe de Lie. . . . .	87
5.3	Temps de calcul (en ms) des différentes approches. . . . .	88
6.1	Fréquence moyenne de convergence en fonction du bruit pour les différentes approches pour des valeurs caractéristiques de $\beta$ . . . . .	97
6.2	Matrice de projection $P^\perp$ , matrice Jacobienne $J$ et erreur initiale $\mathbf{e}$ associées aux algorithmes synthétiques. . . . .	100
6.3	Erreur RMS Point des différentes approches à la 20e itération en fonction du nombre d’images utilisées pour construire le template moyen. . . . .	109
6.4	Fréquence de convergence et précision exprimée par la moyenne et l’écart-type de l’Erreur RMS Point. Le template moyen a été construit avec $j = 5$ images. . . . .	112
6.5	Fréquence de convergence et précision exprimée par la moyenne et l’écart-type de l’Erreur RMS Point. Le template moyen a été construit avec $j = 9$ images. . . . .	112
E.1	Estimation des matrices $\mathbf{A}$ et $\mathbf{B}$ associées aux approches FISC et SCL . . . . .	138



# Glossaire des sigles utilisés

- ACL : Asymmetric Compositional Lie (p. 76)
- AACL : Analytic ACL (p. 81)
- B : Bidirectional (p. 40)
- BA : Bidirectional Additive (p. 63)
- BC : Bidirectional Compositional (p. 63)
- BCL : Bidirectional Compositional Lie (p. 83)
- BCO : Bidirectional Compositional Opposite (p. 63)
- BDGM : Bidirectional Gradient Method (p. 24, 63)
- ESM : Efficient Second-order Minimization (p. 25, 58, 76)
- F : Forward (Directe) (p. 40)
- FA : Forward Additive (p. 21, 53)
- F*-AACL : Fast AAACL (p. 88)
- FA+IAD : Mélange des paramétrisations FA et IAD (p. 60)
- FC : Forward Compositional (p. 54, 76)
- FCL : Forward Compositional Lie (p. 54, 76)
- F*-GACL : Fast-GACL (p. 88)
- FIAC : Forward Inverse Asymmetric Combination (p. 33, 133)
- FIACLT : FIAC on Lie Group for Tracking (p. 135)
- FISC : Forward Inverse Symmetric Combination (p. 32, 133)
- FISCLT : FISC on Lie Group for Tracking (p. 135)
- GACL : Geometric ACL (p. 81)
- GESM : Generalized ESM (p. 26, 59)
- GIC : Generalized Inverse Compositional (p. 26, 56)
- GN : optimisation de type Gauss-Newton
- I : Inverse (p. 40)
- IA : Inverse Additive (p. 24, 56)
- IAD : Inverse Additive Direct (p. 56)
- IAR : Inverse Additive Reverse (p. 55)
- IC : Inverse Compositional (p. 23, 54)
- ICD : Inverse Compositional Direct (p. 55)
- ICL : Inverse Compositional Lie (p. 55, 76)
- ICR : Inverse Compositional Reverse (p. 54)
- MVACL : Minimum-Variance ACL (p. 79)
- PBCL : Projected BCL (p. 84)
- SAD : Symmetric Additive Direct (p. 62)
- SC : Symmetric Compositional (p. 61)
- SCL : Symmetric Composition on Lie Group (p. 58)
- SCO : Symmetric Compositional Opposite (p. 62)
- SCE : Symétrie Compositionnelle Exacte (p. 57)

SGM : Symmetric Gradient Method (p. 23)

SSD : Sum of Squared Difference (Erreur quadratique)

# Introduction

---

## 1.1 Contexte

L'alignement d'images consiste à mettre en correspondance les éléments présents dans différentes images. Les applications sont nombreuses et peuvent être classées dans plusieurs rubriques en fonction de la nature des images [Zitova 2003] :

**Différents points de vue** Mettre en correspondance différents points de vue d'une même scène permet d'obtenir une information plus riche sur la scène observée (que celle qui est portée par chaque image individuellement). Ce principe est utilisé pour répondre à plusieurs problématiques :

- Les appareils photos disposent d'un champ de vision réduit et ne peuvent capturer en une seule image l'ensemble de l'environnement qui les entoure. L'alignement de plusieurs points de vue permet d'obtenir une seule image "grand angle" de l'environnement de la caméra. On parle alors de construction de panorama [Szeliski 2006]. Les travaux sur ces questions ont donné naissance à des applications grand public comme PhotoFit [TekMate 2000], Panorama Builder, REALVIZ ou encore Autostich.
- Il arrive fréquemment que l'information d'importance dans une scène ne soit pas capturée par suffisamment de pixels dans l'image. En effet, les capteurs CCD au coeur des appareils numériques étant limités en termes de résolution, de nombreux détails peuvent être absents (phénomène de repliement dont une caractéristique est l'effet "escalier"). La fusion de plusieurs images de la scène, après alignement, peut permettre de faire apparaître l'information manquante en allant au delà de la résolution du capteur. On parle alors de super-résolution [Farsiu 2005a].
- De nombreuses applications nécessitent aujourd'hui la reconstruction de modèles 3D de l'environnement. La mise en correspondance des images permet l'estimation de la position relative des caméras utilisées pour capturer les différents points de vue [Silveira 2008]. Cette étape est indispensable afin de construire l'objet 3D par triangulation. Ainsi par exemple pour la navigation autonome de robots, l'utilisation des différentes vues occasionnées par les mouvements même du robot permet à ce dernier de reconstruire un modèle partiel de son environnement et de se positionner relativement à ce dernier.

**Différents moments** L'alignement d'images capturées à différents instants permet d'identifier ce qui a (n'a pas) changé dans la scène observée.

- En compression vidéo, l'estimation du mouvement entre images permet de coder efficacement les images d'une séquence sous la forme d'un panorama ou bien, dans des standards comme MPEG-4 [Watkinson 2004], de diminuer l'erreur de prédiction.

- En imagerie médicale, l’alignement de clichés de différentes périodes permet de surveiller les modifications au sein du métabolisme d’un patient, comme l’évolution d’une tumeur.
- Comparer des images satellites prises à différents instants permet de surveiller les changements à la surface du globe terrestre. Ainsi [Shah 2008] étudie l’alignement d’images de l’Arctique pour faciliter le contrôle du cycle des eaux dans cette région. En effet l’évolution des lacs dans cette région du monde est une donnée importante pour analyser les changements climatiques.

**Différents capteurs.** L’alignement multi-capteurs a pour objectif de mettre en correspondance spatiale des images capturant des phénomènes différents :

- En imagerie médicale, les images CTscan (*Computed Tomography*) mettent en évidence les structures macroscopiques à l’intérieur du patient comme les os. Les images PET (*Positron Emission Tomography*) localisent en revanche les activités du métabolisme mais pas bien les structures. Les images étant prises à différents moments et points de vue, il est impossible de simplement superposer les deux images. Leur mise en correspondance par alignement permet d’identifier les sites structurels d’activité métabolique [Brown 1992].
- En imagerie satellite, les images SAR (*Synthetic Aperture Radar*) permettent de surveiller ce qui se passe au sol même en présence de nuages ou bien en pleine nuit. En présence de crise (catastrophe naturelle...), l’alignement d’une image SAR avec une image optique peut constituer une importante aide à la décision [Suri 2010].

**Alignement à un modèle.** L’alignement à un modèle consiste à mettre en correspondance une image avec une représentation de la scène observée. Ce modèle peut être synthétisé sur ordinateur ou bien être construit à partir d’autres images de la scène.

- En imagerie médicale, la mise en correspondance de clichés inter-patients permet d’étudier la variabilité au sein de leurs anatomies [Joshi 2004]. La création d’atlas (banque d’images de patients alignées dans le même repère), permet alors au médecin de disposer d’un ensemble de modèles auxquels comparer les images du patient traité.
- L’alignement des images acquises par une caméra avec des marqueurs visuels insérés dans l’environnement permet de localiser la caméra relativement à son entourage. Il est alors possible de surimposer sur l’image des informations virtuelles. Reposant sur ce principe, une visite virtuelle du cabinet de travail de Charles V pour le château de Vincennes a par exemple été proposée au grand public<sup>1</sup>.
- L’alignement à des modèles 3D CAO (Conception Assistée par Ordinateur) pour le suivi d’objets 3D ou bien à des modèles 2D (comme une image) pour le suivi d’objets 2D est la source de nombreuses applications comme l’asservissement visuel de bras motorisés [Marchand 2005].

La liste présentée ici n’est pas exhaustive. Elle met cependant en lumière une grande variété d’applications possibles. Chacune d’elle possède des spécificités propres. Nous précisons dans la suite le champ de nos investigations.

---

1. <http://www.axyz.fr/realite-augmentee.htm>.

## 1.2 Motivation

### 1.2.1 Imagerie à faible rapport signal sur bruit

Pour des applications comme l'imagerie radar, l'amélioration de la vision nocturne dans les trafics routier et aérien, l'imagerie thermique ou encore l'imagerie médicale, l'alignement d'images constitue un problème difficile conduisant à mettre en correspondance des images de faibles RSB (Rapport Signal sur Bruit) en raison de dégradations spécifiques (respectivement bruit de speckle, fluctuations du nombre de photons incidents, bruit électronique, bruit de quantification). En particulier, un système de vision nocturne constitue une alternative intéressante pour appréhender son environnement dans des conditions de faible luminosité. En raison du fait que la lumière réfléchiée par la cible est très faible, l'image capturée est entachée d'un bruit important. Dans ces conditions de faible RSB, des applications de vision comme la super-résolution, la segmentation ou bien la reconnaissance d'objets nécessitent l'alignement d'images de qualité différente, à savoir par exemple l'image courante fortement bruitée avec une image de référence de meilleure qualité, qui peut avoir été obtenue elle-même à partir des images bruitées.

Cet environnement motivera le cadre applicatif de cette thèse visant à étudier le problème du suivi d'objets par alignement d'images fortement dégradées.

### 1.2.2 Suivi paramétrique

Deux groupes d'approches sont couramment envisagées pour effectuer le suivi.

- Le suivi par mise en correspondance de primitives repose sur la détection et la mise en correspondance d'amers (point d'intérêt, segments...) entre un modèle de référence (*Template*) et l'image courante. Ces approches ont l'avantage d'être robustes à des déplacements importants et aux occultations car le suivi repose sur des correspondances locales. Cependant, en présence de bruit de niveau important, la détection de structures caractéristiques devient difficile.
- Le suivi par optimisation d'une fonctionnelle d'erreur définit une mesure de dissimilarité entre le modèle complet et l'image courante. Une stratégie d'optimisation est utilisée pour minimiser cette erreur. La force de ces approches est de pouvoir fonctionner en présence de motifs complexes difficiles à modéliser par des amers [Jurie 2002b, Benhimane 2007b]. En l'absence de précautions, ces approches sont plus sensibles aux perturbations liées au bruit et sujettes à des minima locaux.

Dans cette thèse, nous nous focalisons sur les méthodes du second groupe. Deux grandes familles sont souvent envisagées. La première repose sur un apprentissage hors-ligne [Jurie 2002a, Bayro-Corrochano 2007][Matthews 2004b, Amberg 2009] de la relation entre l'erreur mesurée et les paramètres de mouvement. Ces méthodes associent la capacité d'estimer des mouvements importants à un faible temps de calcul pour le suivi. La phase d'apprentissage étant relativement coûteuse, la disponibilité préalable d'informations sur l'objet à suivre est nécessaire, ce qui n'est pas toujours le cas en pratique (par exemple dans le cas où l'objet à suivre apparaît pour la première fois dans le champ de vision de la caméra).



Une seconde famille d’approches, que nous appelons approches en ligne, repose sur la minimisation de la fonctionnelle d’erreur par annulation des gradients de cette dernière (*gradient-based method*). La fonctionnelle d’erreur étant la plupart du temps non linéaire, ces approches utilisent un schéma itératif dans lequel, à chaque itération, une approximation de la fonctionnelle est calculée et minimisée. Ces approches sont opérationnelles pour une grande variété d’environnements en s’affranchissant de la nécessité d’informations préalables, c’est pourquoi elles ont déjà été étudiées de façon intensive, comme en témoignent les nombreuses contributions qui ont été apportées pour améliorer les performances de suivi des approches.

Les contributions concernent notamment :

- la prise en compte de différents modèles de mouvement : pour le suivi d’objets rigides, des méthodes permettant l’estimation de différents modèles de mouvements 2D [Tomasì 1991][Hager 1998][Buenaposada 2002] ou bien 3D [Drummond 1999b][Bayro-Corrochano 2007] ont été proposées. Des modèles déformables ont également été étudiés [Matthews 2004b].
- la robustesse à des perturbations extérieures : différents travaux se sont intéressés à rendre le suivi robuste aux changements de luminosité [Hager 1998, Baker 2003a, Bartoli 2008][Evangelidis 2008], ou encore aux occultations [Hager 1998, Jurie 2002b, Schreiber 2007].
- la prise en compte du flou dû au mouvement [Jin 2005, Dai 2006].
- la limitation de l’influence du bruit en termes de biais [Pham 2005][Chen 2009].
- la prise en compte du problème des variations de l’apparence de l’objet au cours du temps et mise à jour du template [Matthews 2004a, Schreiber 2007].
- la limitation du coût calculatoire pour des approches temps-réel en précalculant un certain nombre d’informations [Hager 1998, Baker 2001] ou bien en sélectionnant un sous-ensemble de pixels adapté [Averbuch 2003, Brooks 2007, Benhimane 2007a].
- l’amélioration du rayon de convergence et de la vitesse de convergence des algorithmes [Keller 2004, Benhimane 2007b].

Nous précisons à présent la problématique combinant l’utilisation de ces familles d’approches avec des images bruitées.

### 1.2.3 Problématique

Les approches par optimisation d’une fonctionnelle d’erreur reposent le plus souvent sur un schéma itératif reposant sur des approximations locales de la fonctionnelle. Il est donc important de limiter au maximum le coût calculatoire par itération. C’est pourquoi la plupart des approches de la littérature utilisent uniquement les termes d’ordre strictement inférieur à 2 de la fonctionnelle. Cela permet en outre d’obtenir une certaine stabilité par rapport aux approches d’ordre supérieur [Malis 2005].

Une propriété importante des algorithmes concerne le rayon de convergence, c’est à dire la capacité de ces derniers à converger vers la bonne estimation malgré des mouvements importants, c’est pourquoi les algorithmes sont souvent intégrés dans un schéma multirésolution (par exemple [Keller 2004, Malis 2005]). Ceci a également pour effet de limiter l’influence du bruit

### 1.3. Objectifs de la thèse

---

en moyennant fortement ce dernier pour les échelles les moins fines. Cependant, cette stratégie suppose que les structures ne sont pas dégradées par un tel traitement.

Les approches envisagées dans ce mémoire reposent sur l'estimation des gradients de la fonctionnelle et peuvent être assimilées à des extensions de l'algorithme de [Lucas 1981]. Elles se distinguent dans le choix de/des images(s) utilisées pour le calcul des gradients. Bien que de nombreuses approches de la littérature aient été testées dans des situations d'images bruitées, peu d'approches ont été formalisées pour tirer parti au mieux des informations de chaque image dans cette situation. Ainsi dans [Baker 2004], les auteurs préconisent l'utilisation de la méthode calculant les gradients de la fonctionnelle d'erreur à partir de l'image la moins bruitée. Leur analyse repose sur une étude expérimentale de différentes approches de l'état de l'art. [Pham 2005] présente une analyse de l'influence du bruit en termes de biais sur l'algorithme de [Lucas 1981] et montre que cette influence diminue à une vitesse exponentielle au cours des itérations. Cependant, cette analyse est proposée dans un contexte d'estimation de mouvement subpixelique pour laquelle le mouvement à estimer est toujours inférieur à 1 pixel. Cette étude porte ainsi sur la précision et néglige l'influence du bruit en termes de robustesse. Dans [Chen 2009], les auteurs proposent un algorithme adapté aux images bruitées, cependant l'approche proposée nécessite une connaissance a priori de la variance du bruit sur chacune des images impliquées.

De manière générale, les approches proposées (à part celle de [Chen 2009]) supposent implicitement un alignement entre images dégradées de même RSB (Rapport Signal sur Bruit), ou bien entre une image de très bonne qualité avec une image dégradée par du bruit. Ainsi aucune approche n'a été proposée pour tenir compte des niveaux de bruit dans chaque image et ce sans connaissance a priori des niveaux de bruit.

### 1.3 Objectifs de la thèse

Dans cette thèse, nous nous concentrons surtout sur l'estimation de transformations globales et géométriques entre deux images. Notre objectif applicatif est de développer de nouveaux algorithmes adaptatifs pour le suivi d'objets en présence d'images fortement dégradées par du bruit.

Dans le chapitre 2, nous présentons les travaux sur lesquels s'appuie ce mémoire. Après une introduction rapide des notions théoriques fondamentales liées au problème d'alignement, nous abordons les approches de l'état de l'art. Nous nous focalisons en particulier sur les approches reposant sur une optimisation locale. Cette étude nous permet d'identifier deux problématiques. D'une part, les liens existants entre les approches de l'état de l'art ne sont pas toujours clairs. Un effort de formalisation semble donc nécessaire pour mettre en évidence les bénéfices des différentes approches. D'autre part, la plupart des approches semblent faire des hypothèses implicites fortes concernant le contenu des images, notamment en termes de niveau de bruit présent. Ces hypothèses ne sont pas toujours justifiées d'un point de vue applicatif. Il semble donc possible de proposer de nouvelles méthodes tirant parti de cette constatation pour utiliser au mieux les informations portées par les images.

Afin de répondre à la première problématique, nous proposons dans le chapitre 3 un formalisme générique, que nous nommons "bidirectionnel", pour exprimer le problème de l'alignement

itératif. Nous définissons en particulier trois grandes familles qui recouvrent la taxonomie de [Baker 2004] et permettent d’interpréter les différentes approches de la littérature : les approches Directes, Inverses et Bidirectionnelles. Nous proposons enfin une grille de conception générique d’algorithmes d’alignement itératif, qui formalise pour la première fois de façon explicite les liens entre la paramétrisation locale du mouvement à l’aide d’un incrément et la règle de mise à jour à utiliser dans le cadre rigoureux de la fonctionnelle bidirectionnelle.

Le formalisme bidirectionnel est utilisé au chapitre 4 pour revisiter les approches de l’état de l’art. Nous établissons notamment des connections fortes entre différentes approches de la littérature. Dans le contexte particulier d’une optimisation de type Gauss-Newton, nous démontrons l’équivalence des approches à l’ordre 1 au sein d’une même famille. Cet effort de formalisation nous permet enfin de clarifier un certain nombre de principes pour la conception d’algorithmes de suivi d’objet.

En s’appuyant sur le formalisme bidirectionnel, nous étudions au chapitre 5 la seconde problématique identifiée, à savoir l’alignement d’images dégradées par du bruit. Nous proposons deux nouvelles approches que nous nommons ACL (*Asymmetric Composition on Lie Group*) et BCL (*Bidirectional Composition on Lie Group*), s’adaptant automatiquement aux niveaux de bruit dégradant les deux images. Nous établissons des connections fortes entre ces deux nouvelles approches et l’algorithme ESM [Benhimane 2007b], ce qui nous permet d’identifier théoriquement quelles améliorations elles apportent par rapport aux approches de l’état de l’art.

Dans le dernier chapitre de cette thèse, nous évaluons expérimentalement les nouvelles approches proposées en les comparant aux approches représentatives de l’état de l’art. Dans un premier temps, nous modifions le banc de test de [Baker 2004] afin de pouvoir évaluer plus finement le comportement des algorithmes testés en présence de bruit affectant les images. Ceci donne lieu à un nombre important d’expérimentations sur données synthétiques permettant de confirmer les mérites des différentes approches. Dans un second temps, nous proposons un mécanisme de construction d’image de référence permettant de profiter au mieux des nouvelles approches dans un contexte de suivi d’objets sur une vidéo capturée dans des conditions de faible luminosité (fortement dégradée par du bruit). Les différents algorithmes sont alors évalués sur des données réelles.

# Alignement paramétrique d'images

---

## Sommaire

<b>2.1</b>	<b>Généralités</b>	<b>7</b>
2.1.1	Position du problème	7
2.1.2	Modèles 2D de transformations globales géométriques	8
2.1.3	Critère d'alignement	14
<b>2.2</b>	<b>Alignement par minimisation en ligne d'une fonctionnelle d'erreur</b>	<b>18</b>
2.2.1	Principes généraux de la minimisation d'une fonctionnelle d'erreur	19
2.2.2	L'algorithme de Lucas et Kanade	21
2.2.3	Approches à faible temps de calcul	22
2.2.4	Modification du critère d'alignement pour être robuste à des perturbations photométriques	27
<b>2.3</b>	<b>L'alignement d'images, un problème asymétrique ?</b>	<b>30</b>
2.3.1	Asymétrie lors de la formalisation du problème d'alignement	31
2.3.2	Asymétrie lors de la résolution du problème d'alignement	32
<b>2.4</b>	<b>Conclusion</b>	<b>34</b>

---

## 2.1 Généralités

### 2.1.1 Position du problème

Nous considérons une série d'images  $\{\mathbf{I}_i\}_{i=1..n}$  correspondant à des prises de vue d'un même objet d'intérêt (ici objet peut désigner une scène statique, ou bien un objet en mouvement). On appelle alignement d'images le processus permettant de mettre en correspondance les éléments visuels issus des différentes images. Dans le contexte de l'alignement paramétrique d'images, cette mise en correspondance est réalisée par l'estimation de transformations modélisant les variations entre images. On parle d'alignement paramétrique car les transformations mises en jeu dépendent d'un nombre limité de paramètres. Les modèles de transformation les plus souvent utilisés sont :

- Déformation géométrique : la transformation  $\mathbf{W}$ , paramétrisée par un vecteur de paramètres  $\boldsymbol{\mu}$ , modélise un changement de point de vue de la caméra, ou bien le déplacement de l'objet d'intérêt. Cette transformation agit donc sur les coordonnées spatiales des images. Plus de détails sur ces transformations sont données dans la section 2.1.2.
- Transformation photométrique : la transformation  $\mathfrak{P}$ , paramétrisée par un vecteur de paramètres  $\mathbf{p}$ , modélise un changement de luminosité, ou bien un bruit de mesure modifiant les intensités des pixels de l'image.

Dans le contexte de l'alignement paramétrique d'images, nous allons donc chercher à estimer les paramètres des transformations.

Les différents algorithmes proposés dans la littérature utilisent des modèles pour  $\mathbf{W}$  et  $\mathfrak{P}$  spécifiquement adaptés au domaine applicatif ciblé. Ainsi la plupart des approches adaptent par exemple le modèle de mouvement  $\mathbf{W}$  en fonction d'une connaissance *a priori* du mouvement étudié ou bien le modèle de transformation photométrique  $\mathfrak{P}$  en fonction d'un contexte particulier de prise d'images.

Une fois le modèle de transformation fixé, il est nécessaire de définir un critère  $C$  permettant d'estimer les paramètres de la transformation. Comme pour les modèles de transformation, ce dernier est choisi en fonction du contexte applicatif. On pourra par exemple le choisir de manière à être robuste à des transformations que nous ne cherchons pas à estimer (variation d'éclairage par exemple lorsque nous voulons estimer une déformation géométrique seulement), ou bien pour tenir compte d'un bruit perturbant les données dont le modèle est supposé connu. La figure 2.1 résume d'un point de vue système le problème d'alignement d'images. Les entrées sont

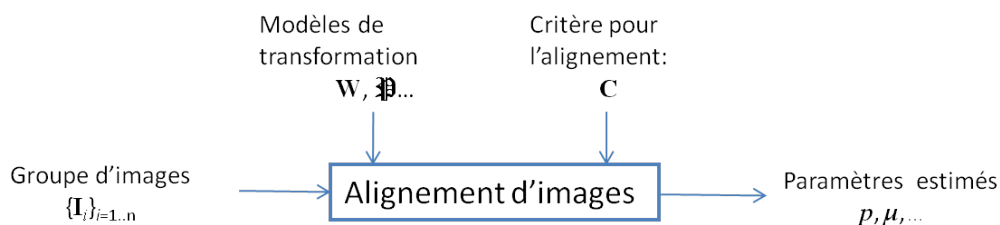


FIGURE 2.1 – Processus d'alignement d'images.

les observations, à savoir les images à aligner ainsi que notre connaissance *a priori* du problème traduite sous la forme de modèles de transformation et d'un critère d'alignement. Les sorties sont les paramètres des transformations. Afin d'estimer  $\mathbf{p}$  et  $\boldsymbol{\mu}$ , il est nécessaire de mettre en place une stratégie pour résoudre le critère mis en jeu. Ce dernier étant généralement complexe, nous ne savons pas le résoudre de manière exacte. De plus, nous sommes souvent amenés, pour des raisons de temps de calcul, à volontairement approximer le problème. La plupart des stratégies de résolution étant développées spécifiquement en fonction du critère d'alignement choisi, nous présentons dans la suite conjointement les critères et les stratégies de résolution.

Ainsi les approches d'alignement d'images peuvent être analysées et classifiées en fonction de la grille de lecture suivante :

- Modèles de transformation mis en jeu.
- Critère d'alignement et méthode de résolution du critère.

Dans cette thèse, nous nous concentrons sur les méthodes d'alignement entre deux images que nous notons dans la suite  $I$  et  $T$ . Nous nous intéressons particulièrement à l'estimation de transformations rigides géométriques entre ces images.

### 2.1.2 Modèles 2D de transformations globales géométriques

Afin de simplifier le discours, nous supposons dans la suite que nous travaillons sur des images à 2 dimensions en niveaux de gris, bien que le formalisme puisse s'étendre sans difficulté

## 2.1. Généralités

---

particulière à des blocs de données volumiques 3D ou à des images à plusieurs canaux (couleur, image multispectrale). Une image  $I$  est définie à travers la fonction continue qui associe une valeur réelle à une position spatiale :

$$\begin{aligned}\mathbb{R}^2 &\longmapsto \mathbb{R} \\ \mathbf{x} &\longmapsto I(\mathbf{x})\end{aligned}\tag{2.1}$$

Considérons  $\mathbf{x}$  une position spatiale dans l'image  $I$  et  $\mathbf{y}$  la position correspondante dans l'image  $T$ . On note  $\mathbf{W}$  l'application qui transforme  $\mathbf{x}$  en  $\mathbf{y}$  :

$$\begin{aligned}\mathbb{R}^2 &\longmapsto \mathbb{R}^2 \\ \mathbf{x} &\longmapsto \mathbf{y} = \mathbf{W}(\boldsymbol{\mu}, \mathbf{x})\end{aligned}\tag{2.2}$$

$\boldsymbol{\mu}$  est un vecteur permettant de paramétrer la transformation.

Nous présentons dans la suite les transformations les plus couramment utilisées dans le contexte d'alignement paramétrique 2D, à savoir celles qui préservent la dimensionalité (un sous espace est transformé en un sous espace de même dimension) et la colinéarité (des points alignés dans la première image, le sont également dans la seconde image).

### 2.1.2.1 Représentation vectorielle des transformations

Nous reprenons ici les paramétrisations des transformations 2D présentées dans [Szeliski 2006]. Le vecteur  $\mathbf{x} = (x, y)^T$  désigne une position spatiale 2D. Ces paramétrisations utilisent un vecteur de paramètres  $\boldsymbol{\mu} \in \mathbb{R}^p$ .

**Translation** Une translation 2D peut être représentée par un vecteur de paramètre  $\boldsymbol{\mu} = (\mu_1, \mu_2)^T \in \mathbb{R}^2$ . La transformation associée est :

$$\mathbf{W}(\boldsymbol{\mu}, \mathbf{x}) = \begin{pmatrix} x + \mu_1 \\ y + \mu_2 \end{pmatrix}\tag{2.3}$$

**Transformation Euclidienne** Une transformation Euclidienne 2D peut être représentée par un vecteur de paramètre  $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3)^T \in \mathbb{R}^3$ .

Le modèle de transformation est défini par :

$$\mathbf{W}(\boldsymbol{\mu}, \mathbf{x}) = \begin{pmatrix} \cos(\mu_1)x - \sin(\mu_1)y + \mu_2 \\ \sin(\mu_1)x + \cos(\mu_1)y + \mu_3 \end{pmatrix}\tag{2.4}$$

Cette transformation est la composition d'une rotation plane et d'une translation. Cette transformation préserve les distances et par conséquent leur rapport, ainsi que les angles et donc le parallélisme.

**Similarité** Cette transformation peut être représentée par un vecteur de 4 paramètres  $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3, \mu_4)^T \in \mathbb{R}^4$ . Son action est définie par la relation :

$$\mathbf{W}(\boldsymbol{\mu}, \mathbf{x}) = \begin{pmatrix} \mu_1 & -\mu_2 & \mu_3 \\ \mu_2 & \mu_1 & \mu_4 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} \mu_1x - \mu_2y + \mu_3 \\ \mu_2x + \mu_1y + \mu_4 \end{pmatrix}\tag{2.5}$$

Cette transformation préserve le rapport des distances, les angles et donc le parallélisme.

**Transformation Affine** Une transformation affine 2D peut être représentée par un vecteur de 6 paramètres  $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3, \mu_4, \mu_5, \mu_6)^T \in \mathbb{R}^6$ . Son application sur une position spatiale conduit à :

$$\mathbf{W}(\boldsymbol{\mu}, \mathbf{x}) = \begin{pmatrix} \mu_1 & \mu_2 & \mu_3 \\ \mu_4 & \mu_5 & \mu_6 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} \mu_1 x + \mu_2 y + \mu_3 \\ \mu_4 x + \mu_5 y + \mu_6 \end{pmatrix} \quad (2.6)$$

Une transformation affine préserve le parallélisme.

**Homographie** Une homographie est une transformation bijective du plan projectif vers lui même qui conserve l'alignement. Elle peut-être représentée par un vecteur de 9 paramètres  $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3, \mu_4, \mu_5, \mu_6, \mu_7, \mu_8, \mu_9)^T \in \mathbb{R}^9$  mais ne possède que 8 degrés de liberté :

$$\mathbf{W}(\boldsymbol{\mu}, \mathbf{x}) \propto H \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} \mu_1 x + \mu_2 y + \mu_3 \\ \mu_4 x + \mu_5 y + \mu_6 \\ \mu_7 x + \mu_8 y + \mu_9 \end{pmatrix} \quad (2.7)$$

avec :

$$H = \begin{pmatrix} \mu_1 & \mu_2 & \mu_3 \\ \mu_4 & \mu_5 & \mu_6 \\ \mu_7 & \mu_8 & \mu_9 \end{pmatrix} \quad (2.8)$$

Les points manipulés ici possèdent une troisième composante appelée composante homogène. Elle traduit le fait que le plan projectif modélise les points situés à l'infini, évitant ainsi le traitement spécifique de certains cas particuliers (intersection de droites parallèles...) [Faugeras 2001, Hartley 2004]. La matrice  $H$  est définie à un coefficient multiplicatif non nul près. Lorsque la composante homogène est non nulle, les coordonnées euclidiennes du point correspondant à  $\mathbf{x} = (x, y)^T$  sont  $\mathbf{y}$  telles que :

$$\mathbf{y} = \begin{pmatrix} \frac{\mu_1 x + \mu_2 y + \mu_3}{\mu_7 x + \mu_8 y + \mu_9} \\ \frac{\mu_4 x + \mu_5 y + \mu_6}{\mu_7 x + \mu_8 y + \mu_9} \end{pmatrix} \quad (2.9)$$

L'homographie ne préserve que le birapport (à savoir le rapport du rapport des distances).

**Gestion de la surparamétrisation** L'homographie est paramétrée par 9 paramètres bien qu'elle ne possède que 8 degrés de liberté. C'est pourquoi une contrainte supplémentaire est imposée sur les paramètres de l'homographie. La contrainte est généralement choisie en fonction de la nature du critère d'alignement (cf. section 2.1.3) mis en jeu afin d'en faciliter la résolution. Ainsi, lorsque l'homographie est estimée à partir d'un ensemble de paires de primitives appariées entre les images, comme des points, des droites, ou des coniques, une approche classique [Hartley 2004] consiste à rajouter la contrainte  $\|\boldsymbol{\mu}\| = 1$  et à utiliser l'algorithme DLT (*Direct Linear Transform*) qui permet d'estimer les paramètres de l'homographie à partir de  $n$  couples de points ( $n > 3$ ) (cf. [Hartley 2004], pp. 88-91) en résolvant un système linéaire. Si cette paramétrisation semble adaptée pour l'utilisation de l'algorithme DLT, elle l'est beaucoup moins lorsque l'on veut améliorer l'estimation obtenue par l'algorithme DLT en minimisant une fonctionnelle

## 2.1. Généralités

---

d'erreur faisant intervenir des distances géométriques et qui s'avère souvent non linéaire. En effet l'utilisation de la contrainte  $\|\boldsymbol{\mu}\| = 1$  avec une fonctionnelle non linéaire impose de mettre en place une procédure d'optimisation sous contrainte [Guangwei 2008, Vercauteren 2009] ce qui complique la résolution du problème et augmente la complexité calculatoire.

Lorsque l'estimation de l'homographie passe par l'optimisation d'une fonctionnelle d'erreur non linéaire, deux paramétrisations ont été proposées. Dans [Baker 2004], les auteurs utilisent la contrainte  $\mu_9 = 1$ . Avec cette contrainte l'homographie est paramétrée par 8 paramètres appartenant à  $\mathbb{R}^8$ . Cela permet la mise en place de procédures simples d'optimisation itérative sur des espaces vectoriel, comme celles présentées à la section 2.2. Cependant cette paramétrisation peut également engendrer des instabilités, notamment si le mouvement à estimer correspond à  $\mu_9$  proche de 0. Cela arrive par exemple lorsque l'homographie entre deux images  $I$  et  $T$  projette l'origine de  $T$  sur la droite à l'infini dans  $I$  (cf. [Hartley 2004]). Cette situation ne correspond pas forcément à un cas généré si la région d'intérêt considérée pour l'alignement ne contient pas l'origine.

Une troisième paramétrisation, présentée dans [Benhimane 2007b], utilise la contrainte  $\det(H) = 1$  avec  $H$  définie par (2.8). Cette dernière approche présente l'avantage de pouvoir représenter l'ensemble des transformations possibles, tout en étant particulièrement adaptée au problème de minimisation d'une fonctionnelle non linéaire. En effet, en utilisant le fait que l'ensemble des homographies forme un groupe de Lie, il est possible de convertir le problème d'optimisation sur un groupe (sous contrainte) en un problème d'optimisation sur un espace vectoriel (non contraint). Le paragraphe 2.1.2.3 détaille la théorie derrière une telle paramétrisation par algèbre de Lie.

### 2.1.2.2 Groupe et action de groupe

Dans la suite, nous faisons l'hypothèse que  $\boldsymbol{\mu}$  appartient à un groupe  $\mathcal{G}$ , c'est à dire à un ensemble muni d'une loi de composition interne notée  $\circ$ , d'un élément neutre noté  $\mathbf{0}_{\mathcal{G}}$  et vérifiant :

$$\forall \boldsymbol{\mu}, \boldsymbol{\nu} \in \mathcal{G} \quad \boldsymbol{\mu} \circ \boldsymbol{\nu} \in \mathcal{G} \quad (2.10)$$

$$\forall \boldsymbol{\mu} \in \mathcal{G} \quad \boldsymbol{\mu} \circ \mathbf{0}_{\mathcal{G}} = \boldsymbol{\mu} \quad (2.11)$$

$$\forall \boldsymbol{\mu} \in \mathcal{G} \quad \exists \boldsymbol{\mu}^{-1} \in \mathcal{G}, \quad \boldsymbol{\mu} \circ \boldsymbol{\mu}^{-1} = \mathbf{0}_{\mathcal{G}} \quad (2.12)$$

La transformation  $\mathbf{W}$  est une action de groupe si elle vérifie :

$$\forall \boldsymbol{\mu}, \boldsymbol{\nu} \in \mathcal{G} \quad \mathbf{W}(\boldsymbol{\mu} \circ \boldsymbol{\nu}, \mathbf{x}) = \mathbf{W}(\boldsymbol{\mu}, \mathbf{W}(\boldsymbol{\nu}, \mathbf{x})) \quad (2.13)$$

$$\mathbf{W}(\mathbf{0}_{\mathcal{G}}, \mathbf{x}) = \mathbf{x} \quad (2.14)$$

$$\mathbf{y} = \mathbf{W}(\boldsymbol{\mu}, \mathbf{x}) \iff \mathbf{x} = \mathbf{W}(\boldsymbol{\mu}^{-1}, \mathbf{y}) \quad (2.15)$$

Les transformations présentées à la section 2.1.2.1 sont des actions de groupe. Les transformations affines 3D vérifient également ce modèle.



### 2.1.2.3 Paramétrisation locale dans une algèbre de Lie.

Dans cette partie, nous présentons une façon particulière permettant de paramétrer certaines transformations proches de l'identité. Cette paramétrisation des transformations infinitésimales joue un rôle important lorsqu'il est question de mettre en place une procédure d'optimisation. En effet comme nous le verrons à la section 2.2, le problème d'alignement se ramène souvent à déterminer un incrément  $\delta\boldsymbol{\mu} \in \mathcal{G}$  proche de l'identité permettant de corriger une estimation courante des paramètres d'alignement. Cela passe par la mise en place d'une procédure d'optimisation qui peut-être complexe lorsque le groupe des transformations considéré n'est pas un espace vectoriel mais une variété quelconque. Dans ce cas une représentation matricielle de la transformation (comme celles présentées section 2.1.2) nécessite d'imposer une contrainte supplémentaire sur les paramètres d'alignement ce qui conduit à des algorithmes d'optimisation sous contraintes. Comme nous le voyons dans cette section, il existe un certain nombre de ces groupes, réunis sous la désignation *groupes de Lie*, dont les propriétés spécifiques permettent de grandement faciliter l'estimation de  $\delta\boldsymbol{\mu}$ .

Un groupe de Lie  $\mathcal{G}$  est une variété différentielle munie de l'opération de composition ( $\circ$ ). Les transformations comme les mouvements affines sur le plan Euclidien  $\mathbf{x} \in \mathbb{R}^2$  et les homographies sur le plan projectif  $\mathbf{x} \in \mathbb{P}^2$  [Bayro-Corrochano 2007] ou encore l'ensemble des rotations 3D sur  $\mathbb{R}^3$  [Y. Ma 2004] sont des exemples de groupes de Lie.

Une algèbre de Lie  $\mathfrak{G}$  peut être associée à un groupe de Lie de dimension finie  $\mathcal{G}$  dont l'espace vectoriel sous-jacent est l'espace tangent à  $\mathcal{G}$  au niveau de l'élément neutre  $\mathbf{0}_{\mathcal{G}}$ . Le principal intérêt d'une paramétrisation dans un groupe de Lie est que localement un élément  $\delta\boldsymbol{\mu} \in \mathcal{G}$  autour de  $\mathbf{0}_{\mathcal{G}}$  peut être paramétrisé par un incrément  $\mathbf{v} \in \mathfrak{G}$  en utilisant l'application exponentielle (*exponential map*) :

$$\delta\boldsymbol{\mu}(\mathbf{v}) = \exp(\mathbf{v}), \quad (2.16)$$

avec les propriétés suivantes, pour tout  $\alpha, \beta \in \mathbb{R}$  :

$$\exp(-\mathbf{v}) = \exp(\mathbf{v})^{-1} \quad (2.17)$$

$$\exp(\alpha\mathbf{v}) \circ \exp(\beta\mathbf{v}) = \exp((\alpha + \beta)\mathbf{v}). \quad (2.18)$$

L'intérêt d'une telle paramétrisation apparait lorsque l'on s'intéresse à estimer l'incrément  $\delta\boldsymbol{\mu}$  proche de  $\mathbf{0}$  permettant d'optimiser une fonctionnelle d'erreur. En reparamétrant ce dernier par  $\mathbf{v}$  dans l'algèbre de Lie associée, le problème d'optimisation sur le groupe des paramètres  $\mathcal{G}$  se ramène à un problème équivalent où l'on cherche à optimiser la fonctionnelle sur  $\mathfrak{G}$  qui est un espace vectoriel. La relation (2.16) nous permet alors au final d'estimer un  $\delta\boldsymbol{\mu}$  appartenant au groupe des transformations recherchées. Ces propriétés de groupe de Lie facilitent l'estimation de transformations courantes comme les transformations affines et les transformations euclidiennes 3D (Rotation + Translation) [Drummond 1999b, Drummond 1999a, Bayro-Corrochano 2007] ou les homographies [Benhimane 2007b, Bayro-Corrochano 2007].

Dans la suite, nous donnons un exemple de reparamétrisation dans l'algèbre de Lie pour le cas des homographies. Le groupe des homographies peut être représenté par l'ensemble des matrices

## 2.1. Généralités

---

$3 \times 3$  dont le déterminant vaut 1. Comme pour tous les groupes matriciels de dimension finie, l'application exponentielle est définie par l'exponentielle de matrice *expm*. Ainsi pour  $\boldsymbol{\mu}$  proche de l'identité, on peut écrire :

$$H(\boldsymbol{\mu}) = \text{expm}\left(\sum_{i=1}^8 v_i \mathbf{A}_i\right) \quad (2.19)$$

avec pour toute matrice  $\mathbf{X}$  :

$$\text{expm}(\mathbf{X}) = \sum_{i=1}^{\infty} \frac{1}{i!} (\mathbf{X}^i) \quad (2.20)$$

Les  $\mathbf{A}_i$  sont des matrices  $3 \times 3$  appelés les générateurs de l'algèbre de Lie. La condition  $\det(H(\boldsymbol{\mu})) = 1$  impose la condition suivante sur les générateurs :

$$\forall i, \quad \text{Tr}(\mathbf{A}_i) = 0 \quad (2.21)$$

où *Tr* désigne la trace de la matrice. Les  $\mathbf{A}_i$  représentent les directions tangentes au groupe de Lie au voisinage de l'identité. On choisit généralement les  $\mathbf{A}_i$  de telle sorte qu'elles aient une interprétation géométrique :

$$\mathbf{A}_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \mathbf{A}_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad (2.22)$$

sont les générateurs associés aux translations (il suffit d'appliquer (2.20) à  $\mathbf{A}_1$  ou bien  $\mathbf{A}_2$  pour obtenir des matrices de translation).

$$\mathbf{A}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \quad \mathbf{A}_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (2.23)$$

sont les générateurs associés respectivement aux changements d'échelle isotropes et à la rotation plane par rapport à l'origine.

$$\mathbf{A}_5 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \mathbf{A}_6 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (2.24)$$

sont les générateurs associés respectivement aux cisaillements (*shear*) à 0 et 45 degrés.

$$\mathbf{A}_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \mathbf{A}_8 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad (2.25)$$

sont les générateurs associés aux aspects purement projectif (induit par la droite à l'infini). Il est intéressant de noter qu'en sélectionnant un sous-ensemble des  $\mathbf{A}_i$  il est possible de définir une reparamétrisation locale de Lie pour n'importe quelle transformation présentée à la section 2.1.2.

### 2.1.3 Critère d'alignement

Concernant la formalisation du problème d'alignement, deux tendances majoritaires se dégagent :

- Approches utilisant des primitives (*Feature-based approaches*) : l'estimation du mouvement peut être formalisée comme la résolution d'un problème géométrique impliquant l'extraction préalable de primitives géométriques comme des points, des droites...
- Approches directes (*Direct alignment*) : l'estimation du mouvement met en jeu la définition d'une fonctionnelle à optimiser tenant compte de toute l'information portée par les images.

#### 2.1.3.1 Alignement géométrique par mise en correspondances de primitives

Les approches utilisant des primitives reposent essentiellement sur 3 étapes :

1. Extraction des primitives : il s'agit de détecter dans l'image des objets géométriques caractéristiques comme des points d'intérêt (coins), des segments (arêtes d'objets) et dans certains cas des primitives plus évoluées comme des cercles, ou bien encore des coniques.
2. Mise en correspondance des primitives : il s'agit d'estimer les correspondances entre primitives de 2 images distinctes (ou plus).
3. Estimation des paramètres d'alignement : utilisation de l'information extraite pour déterminer les paramètres  $\mu$  de la transformation géométrique.

**Extraction de primitives.** Les points d'intérêt sont utilisés le plus couramment car ces primitives sont présentes dans les images naturelles. Le fait que ces primitives soient distribuées sur toute l'image permet de concevoir des systèmes génériques et robustes aux occlusions. Ce type d'approches repose souvent sur le parcours de l'image à l'aide d'une fenêtre de dimensions variables. L'analyse du contenu de la fenêtre doit permettre de déterminer si il y a présence à la position courante d'un point d'intérêt.

La famille de détecteurs la plus utilisée repose sur l'analyse du tenseur de structure associé à la fenêtre courante. Ce dernier permet de calculer simplement les deux directions principales (vecteurs propres du tenseur de structure) et la dispersion associée (valeurs propres du tenseur de structure) selon lesquelles a lieu la variation d'intensité lumineuse. L'analyse des valeurs propres de ce tenseur permet donc efficacement de déterminer la présence d'un coin (2 valeurs propres proches et grandes) ou bien d'un contour (une valeur propre négligeable devant l'autre). Deux approches couramment employées et reposant sur ce constat sont le détecteur de Harris [Harris 1988] ou encore celui de [Shi 1994]. Dans [Shi 1994] les auteurs ont montré que les positions dont le tenseur associé présentait des valeurs propres grandes et de même amplitude sont des régions susceptibles d'être suivies de manière fiable dans le temps ce qui justifie ce type de mesures. Ces approches ont été combinées avec des décompositions multi-échelles afin de les rendre robustes à des changements d'échelle importants apparaissant lorsque la caméra s'est fortement éloignée ou rapprochée de la scène considérée, ou bien lorsque l'on cherche à appareiller des images prises à des résolutions différentes. Une technique souvent employée est de construire un espace échelle en convoluant l'image avec un noyau Gaussien de différentes tailles

## 2.1. Généralités

---

puis d'appliquer le détecteur sur chacun des palliers [Mikolajczyk 2004]. Une sélection automatique de l'échelle caractéristique permet alors d'éliminer la redondance lorsque une structure a été détectée à plusieurs échelles [Lindeberg 1998]. Cette échelle caractéristique permet également d'extraire un descripteur adapté pour décrire la structure détectée.

Une deuxième approche de détection consiste à étudier la réponse à un simple filtrage de l'image mettant en évidence les hautes fréquences et donc les zones présentant un contraste. Cette approche se prête également bien à une implémentation multi-échelles. Ainsi l'opérateur LoG (*Laplacian of Gaussian*) applique un filtre Laplacien à chaque échelle, ce qui a pour effet de mettre en évidence les contours, les extrema et les coins. La variante DoG (*Difference of Gaussian*) [Lindeberg 1998] consiste à remplacer le filtre par une simple différence d'images lissées à deux échelles séparées d'un facteur constant. Cette approche plus rapide est notamment utilisée par le détecteur SIFT (*Scale Invariant Feature Transform*) [Lowe 2004].

Une troisième famille de détecteur repose sur l'analyse des niveaux de gris dans la fenêtre considérée. Ces détecteurs sont pour la plupart des extensions du détecteur SUSAN (Smallest Univalued Segment Assimilating Nucleus) [Smith 1997]. Ce dernier utilise une fenêtre circulaire qui évalue les pixels dans la fenêtre dont l'intensité est proche de celle du pixel central. La zone constituée par ces pixels est appelée *Univalued Segment Assimilating Nucleus* (USAN). Les zones caculées dont l'aire atteint un minimum local correspondent à des zones centrées sur des coins. Ce détecteur a l'avantage d'être plus performant en temps de calcul que le détecteur de Harris.

**Mise en correspondance.** Afin de mettre en correspondance les primitives, à ces dernières est associé un descripteur qui résume les informations présentes au voisinage de la structure détectée. Généralement la dimension du voisinage considéré est choisie en fonction de l'échelle associée au point d'intérêt. Une manière simple de représenter la structure détectée est d'utiliser l'intensité des pixels voisins de la structure comme descripteur. Cependant ceci conduit à des descripteurs de dimensionnalité importante ce qui implique un coût calculatoire élevé lors de l'appariement. C'est pourquoi les travaux se sont principalement dirigés ces dernières années vers des descripteurs reposant sur des distributions d'attributs. Ces derniers utilisent des histogrammes afin de capturer certaines caractéristiques de forme et d'apparence du voisinage de la structure. Le descripteur probablement le plus connu est le descripteur SIFT (*Scale Invariant Feature Transform*) [Lowe 2004] qui repose sur le calcul d'un histogramme 3D de la position et de l'orientation des gradients du voisinage. Une évaluation expérimentale détaillée de plusieurs descripteurs sur des images en niveaux de gris a été proposée dans [Mikolajczyk 2005].

**Éstimation des paramètres d'alignement.** Deux types de fonction coût sont envisagés pour estimer les paramètres de l'alignement géométrique. La première consiste à utiliser les couples de correspondance pour définir une distance algébrique de la forme  $\|\mathbf{A}\boldsymbol{\mu}\|$  que l'on va chercher à minimiser.  $\mathbf{A}$  est une matrice construite à partir des coordonnées des différents couples de correspondance. Cette approche est possible pour les transformations définies dans la partie 2.1.2. Par exemple l'algorithme *DLT* (Direct Linear Transform), utilisé pour l'estimation des paramètres d'une homographie [Hartley 2004], minimise une telle distance sous la contrainte

supplémentaire  $\|\boldsymbol{\mu}\| = 1$  (les homographies sont représentées par des matrices  $3 \times 3$  bien qu'elles ne possèdent que 8 degrés de liberté, cf. partie 2.1.2). Ceci conduit à un système homogène qui est traditionnellement résolu simplement en utilisant des approches comme la SVD (*Singular Value Decomposition*) qui permet de trouver efficacement le noyau de la matrice  $\mathbf{A}$ . Il a été montré dans [Hartley 1997] que ce type d'approche était très sensible au bruit sur les positions spatiales des points d'intérêt. Les auteurs proposent une renormalisation des coordonnées permettant de fortement réduire les erreurs. Dans [Harker 2005], les auteurs montrent que la matrice  $\mathbf{A}$  contient à la fois des éléments déterministes (qui peuvent être réarrangés sous la forme d'une matrice creuse) et des éléments bruités. L'algorithme DLT ne tient pas compte de la nature différente des éléments présents dans  $\mathbf{A}$ . Les auteurs proposent alors un algorithme de résolution conservant le caractère non-itératif de l'approche DLT mais qui s'avère plus robuste aux erreurs et en moyenne cinq fois plus rapide. L'avantage principal des approches algébriques est d'estimer les paramètres par la résolution d'un problème linéaire ce qui les rend efficace en temps de calcul et non sujettes à des minima locaux.

Une deuxième approche, souvent utilisée pour améliorer les résultats obtenus par les approches algébriques, consiste à minimiser une distance géométrique à savoir une distance entre coordonnées de points dans un repère précis. Par exemple, dans le cas de deux images  $I$  et  $T$  pour lesquelles ont été détectés et appareillés les points  $\{\mathbf{x}_i\}_{i=1..N}$  et  $\{\mathbf{y}_i\}_{i=1..N}$ , on peut chercher les paramètres  $\boldsymbol{\mu}$  qui minimisent l'erreur de transfert [Hartley 2004] :

$$\sum_{i=1}^N (\mathbf{y}_i - \mathbf{W}(\boldsymbol{\mu}, \mathbf{x}_i))^2 \quad (2.26)$$

Lorsque le modèle de mouvement  $\mathbf{W}$  n'est pas linéaire (par rapport à  $\boldsymbol{\mu}$ ), l'équation (2.26) requiert l'utilisation d'approches itératives. Ces dernières supposent la connaissance approximative  $\boldsymbol{\mu}_0$  des paramètres d'alignement ce qui les rend dépendantes d'une première étape d'estimation de ces paramètres. L'estimation approximative est généralement réalisée en utilisant une méthode algébrique. La nature de la fonctionnelle (2.26) dépend essentiellement de notre connaissance du problème. Ainsi l'erreur (2.26) suppose un bruit Gaussien seulement sur les positions détectées dans l'image  $I$ . Ce type d'erreur est parfaitement adaptée lorsque nous alignons une image  $I$  à un modèle  $T$  pour lequel est connu précisément la position des structures d'intérêt. Cette fonctionnelle peut être facilement adaptée à du bruit sur les mesures des deux images en utilisant une distance de transfert symétrique, ou bien une erreur de reprojection [Hartley 2004].

Afin de prendre en compte la présence éventuelle, dans les observations, de données aberrantes (*outliers*), à savoir des points appareillés qui ne se correspondent pas en réalité, il est également possible [Malis 2006] soit d'effectuer à priori un classement des couples de points en *inliers* (correspondant au modèle) ou *outliers* par des algorithmes de vote comme la transformée de Hough ou bien l'algorithme RANSAC [Fischler 1981] puis d'estimer les paramètres finaux, soit d'estimer simultanément les paramètres d'alignement et les outliers en remplaçant la fonctionnelle (2.26) par une métrique robuste comme les M-estimateurs.

### 2.1.3.2 Alignement Direct

L'alignement direct repose sur deux choix principaux :

## 2.1. Généralités

---

1. Définition d'une fonctionnelle d'erreur  $E(\boldsymbol{\mu})$  décrivant le processus d'alignement.
2. Choix d'un algorithme d'optimisation.

**Définition d'une fonctionnelle d'erreur.** La famille de fonctionnelles la plus étudiée pour l'estimation des paramètres de mouvement entre deux images par alignement directe est la famille du *template matching*. Cette famille repose sur la définition d'un critère de comparaison directe des intensités lumineuses des images. Le problème d'alignement est alors formulé comme l'optimisation de ce critère. Les méthodes généralement envisagées sont des approches itératives en raison de la nature non linéaire de la fonctionnelle d'erreur.

Parmi les approches de *template matching*, la fonctionnelle la plus répandue est l'erreur quadratique SSD (*Sum of Squared Difference*) utilisée dès les travaux de [Lucas 1981] pour le calcul de flot optique, qui est optimale au sens du maximum de vraisemblance sous l'hypothèse que les mesures sont entachées par un bruit blanc Gaussien centré. Plusieurs variantes de cette fonctionnelle SSD ont été proposées dans la littérature, soit pour augmenter la robustesse lors de l'estimation du mouvement [Keller 2004], soit pour augmenter les performances en temps de calcul en permettant de réaliser un certain nombre de calculs hors-ligne [Hager 1998, Baker 2004], soit enfin en limitant le nombre d'itérations nécessaire pour converger [Benhimane 2004, Keller 2004, Benhimane 2007b].

Afin de rendre le processus d'alignement robuste à des transformations photométriques qui n'ont pas d'intérêt à être estimées, d'autres fonctionnelles de *template matching* ont été proposées. Ainsi dans [Evangelidis 2008][Chen 2009], les auteurs utilisent le coefficient de corrélation normalisé au lieu de la SSD rendant le processus robuste aux variations d'illumination.

Dans le cas de la présence de données aberrantes (occultation, reflets spéculaires,...) une solution consiste à utiliser les M-estimateurs [Black 1993]. Ces derniers permettent le développement de systèmes robustes sans nécessiter la modélisation explicite de toutes les sources de désalignement possibles dans le critère d'alignement ou dans les modèles de transformation des images. Cependant ce type d'approches conduit la plupart du temps à des algorithmes plus lents à converger [Baker 2003b].

Modifier la fonctionnelle d'erreur permet également de gagner en robustesse en présence de déformations géométriques importantes. Ainsi sont présentées des approches similaires aux approches développées pour le *template matching* dans les travaux de [Comaniciu 2003, Hager 2004, Mégret 2006] où le critère à minimiser est une erreur entre distributions de couleurs pondérées. Ces fonctionnelles permettent d'obtenir des zones de convergence plus importantes et d'être plus robuste dans les situations où la déformation géométrique de la scène observée n'est pas parfaitement prise en compte par le modèle de transformation utilisé lors de l'estimation du mouvement.

Enfin pour s'adapter au contexte de l'alignement d'images multimodales, des méthodes reposant sur la maximisation de l'information mutuelle entre les images [Maes 1997, Dowson 2008, Dame 2009] ont été proposées. Ce critère permet de mesurer la dépendance statistique entre les intensités des pixels des images comparées et est supposé être maximal quand les deux images sont alignées.

**Optimisation de la fonctionnelle.** Une fois la fonctionnelle d'erreur déterminée, il est nécessaire de mettre en place une procédure pour optimiser cette dernière. Une première famille d'approches repose sur l'estimation en ligne des paramètres de mouvement. La fonctionnelle à optimiser étant non-linéaire, ces approches reposent sur une approximation de la fonctionnelle d'erreur afin de se ramener à un système linéaire. Des algorithmes génériques d'optimisation itérative comme la méthode de descente du gradient [Amberg 2009], ou encore des variantes de l'algorithme de Newton [Lucas 1981, Shum 2000, Xiao 2008] sont les méthodes les plus utilisées. Ces méthodes estiment les extrema de la fonctionnelle en calculant les points stationnaires de cette dernière, à savoir les points pour lesquels les gradients de la fonctionnelle s'annulent. Nous reviendrons plus en détail sur ces approches à la section 2.2.

Une deuxième famille envisagée sont les approches supervisées qui se fondent sur une phase d'apprentissage. Introduites par [Gleicher 1997] sous le nom d'approches par Décomposition des Différences (*Difference Decomposition*), ces méthodes reposent sur l'utilisation d'une image de référence pour générer, à l'aide de déformations choisies, un ensemble d'erreurs de référence. Ces erreurs sont obtenues en calculant la différence pixel à pixel entre l'image de référence et l'image obtenue en appliquant les déformations sélectionnées à l'image de référence. Dans l'approche initiale, ces erreurs de référence sont utilisées afin de décomposer l'erreur réelle observée comme une combinaison linéaire des erreurs de référence. En faisant l'hypothèse que l'erreur observée varie linéairement avec les paramètres de mouvement, les auteurs proposent une approche où les paramètres de mouvement estimés sont obtenus par une combinaison linéaire de ceux employés pour générer les images d'erreur de référence. Une approche similaire est proposée dans [Jurie 2002a][Bayro-Corrochano 2007] où les erreurs de référence sont employées pour précalculer une matrice qui, appliquée à l'erreur observée, produit directement les paramètres de mouvement correspondant. Ces approches se sont révélées efficaces par rapport aux approches en ligne (utilisant les gradients) pour estimer des mouvements d'amplitude importante. En outre, une fois la phase d'apprentissage réalisée, le coût calculatoire introduit durant l'estimation du mouvement est très faible.

Cependant comme cela a été décrit dans [Benhimane 2006], ces approches ne sont pas adaptées à toutes les applications. En effet, la phase d'apprentissage peut s'avérer très coûteuse en temps de calcul ce qui rend ces approches inutilisables dans un contexte de suivi où l'on ne dispose pas à l'avance d'images de l'objet à suivre (par exemple si l'objet est sélectionné par l'utilisateur au début d'un flux vidéo).

## 2.2 Alignement par minimisation en ligne d'une fonctionnelle d'erreur

Dans ce paragraphe, nous nous intéressons plus particulièrement aux approches d'alignement reposant sur l'optimisation d'une fonctionnelle d'erreur. Cette fonctionnelle d'erreur est construite à partir des images  $I$  et  $T$  à aligner. Considérons la fonctionnelle générique  $E(\boldsymbol{\mu})$  à valeur réelle et dépendant de paramètres représentés par le vecteur  $\boldsymbol{\mu}$ . Les paramètres  $\boldsymbol{\mu}^*$  qui optimisent cette fonctionnelle (selon les cas minimum, ou maximum) vont être au centre de notre étude. Dans un premier temps, nous présentons un certain nombre de concepts généraux utilisés



## 2.2. Alignement par minimisation en ligne d'une fonctionnelle d'erreur

---

par l'ensemble des approches de la littérature. Nous revenons ensuite sur l'algorithme de Lucas et Kanade [Lucas 1981] qui est à la source des approches en ligne, puis nous présentons les extensions les plus marquantes de cet algorithme pour nos travaux. Les approches sont présentées depuis le point de vue original des auteurs. Le formalisme présenté au chapitre 3 nous permettra de proposer une réinterprétation des différentes approches.

### 2.2.1 Principes généraux de la minimisation d'une fonctionnelle d'erreur

#### 2.2.1.1 Optimisation itérative

La plupart des fonctionnelles rencontrées en pratique, que ce soit en *template matching* [Baker 2004], en suivi d'objet par mise en correspondance de distributions de couleurs [Comaniciu 2003, Hager 2004], ou encore en alignement d'images multimodales [Maes 1997, Viola 1997] sont non linéaires. Nous sommes donc amenés à utiliser une procédure d'optimisation itérative reposant sur les étapes suivantes :

1. Définir une relation incrémentale entre l'estimé courant des paramètres  $\boldsymbol{\mu}^n$  et l'estimé suivant  $\boldsymbol{\mu}^{n+1}$  en introduisant un incrément  $\mathbf{v}$  :

$$\boldsymbol{\mu}^{n+1} = \boldsymbol{\mu}(\boldsymbol{\mu}^n, \mathbf{v}). \quad (2.27)$$

L'introduction d'un incrément supposé infinitésimal  $\mathbf{v}$  permet de simplifier la fonctionnelle d'erreur. C'est ce dernier qui est estimé à chaque itération.

2. Approximer la fonctionnelle d'erreur : généralement en utilisant un développement de Taylor de cette dernière.
3. Résoudre le problème approximé par la recherche des points stationnaires (points pour lesquels les gradients de la fonctionnelle sont nuls) de ce dernier, donnant lieu à l'estimation des paramètres  $\hat{\mathbf{v}}$ .
4. Répéter cette procédure jusqu'à convergence après avoir mis à jour les paramètres :  $\boldsymbol{\mu}^{n+1} \leftarrow \boldsymbol{\mu}(\boldsymbol{\mu}^n, \hat{\mathbf{v}})$ .

#### 2.2.1.2 Paramétrisation incrémentale

La paramétrisation incrémentale dépend essentiellement de la nature de l'ensemble  $\mathcal{G}$  dans lequel évoluent les paramètres à estimer.

**Paramétrisation additive vectorielle.** Lorsque l'ensemble des paramètres est isomorphe à un espace vectoriel de dimension  $p$  (translation, changement d'échelle, transformation affine...), où  $p$  désigne le nombre de paramètres à estimer, une relation incrémentale additive peut être utilisée :

$$\boldsymbol{\mu}(\boldsymbol{\mu}^n, \mathbf{v}) = \boldsymbol{\mu}^n + \mathbf{v}. \quad (2.28)$$

**Paramétrisation compositionnelle.** Dans le cas où l'ensemble des paramètres  $\mathcal{G}$  forme un groupe, une relation incrémentale compositionnelle sera préférée :

$$\boldsymbol{\mu}(\boldsymbol{\mu}^n, \mathbf{v}) = \boldsymbol{\mu}^n \circ \mathbf{v}. \quad (2.29)$$



Remarquons que la composition  $\circ$  peut être confondue avec l'addition lorsque le groupe  $\mathcal{G}$  est un espace vectoriel, comme c'est le cas pour l'ensemble des translations dans  $\mathbb{R}^2$ . Ce type de paramétrisation incrémentale est directement lié à la notion de composition de transformations [Vercauteren 2009]. Pour une paramétrisation compositionnelle,  $\mathbf{v}$  appartient au groupe des paramètres  $\mathcal{G}$  garantissant ainsi également l'appartenance de  $\boldsymbol{\mu}(\boldsymbol{\mu}^n, \mathbf{v})$  à ce groupe.

**Paramétrisation par groupe de Lie.** Dans le cas particulier où l'ensemble des paramètres forme un groupe, mais qu'en plus ce dernier est une variété (espace courbe) différentielle (on parle de groupe de Lie), une relation incrémentale compositionnelle sera également préférée :

$$\boldsymbol{\mu}(\boldsymbol{\mu}^n, \mathbf{v}) = \boldsymbol{\mu}^n \circ \exp(\mathbf{v}). \quad (2.30)$$

Dans ce cas, comme nous l'avons vu section 2.1.2.3, l'incrément  $\mathbf{v}$  appartient à l'algèbre de Lie  $\mathfrak{G}$  associée au groupe  $\mathcal{G}$ , ce qui garantit également l'appartenance de  $\boldsymbol{\mu}(\boldsymbol{\mu}^n, \mathbf{v})$  au groupe.

Ainsi l'incrément  $\mathbf{v}$  appartient soit au groupe  $\mathcal{G}$  soit à l'algèbre de Lie  $\mathfrak{G}$  associée si  $\mathcal{G}$  est un groupe de Lie. Dans la suite, l'espace dans lequel évolue l'incrément  $\mathbf{v}$  est noté  $\mathcal{P}$ . Dans la paramétrisation incrémentale, les paramètres courants  $\boldsymbol{\mu}^n$  sont connus et le mouvement à estimer est entièrement caractérisé par  $\mathbf{v}$ . Ainsi la fonctionnelle d'erreur peut être optimisée uniquement en fonction de l'inconnue  $\mathbf{v}$ . Par abus de notation nous écrivons dans la suite  $E^n(\mathbf{v}) = E(\boldsymbol{\mu}(\boldsymbol{\mu}^n, \mathbf{v}))$  ou simplement  $E(\mathbf{v})$  quand il n'y a pas d'ambiguïté. Une fois l'estimation de l'incrément  $\mathbf{v}$  réalisée, les paramètres de mouvement correspondants dans  $\mathcal{G}$  sont obtenus grâce à l'équation (2.27). Dans la partie suivante, nous illustrons le schéma d'optimisation itérative sur une fonctionnelle d'erreur quadratique SSD (*Sum Squared Difference*) en introduisant l'algorithme Gauss-Newton.

### 2.2.1.3 Illustration de l'optimisation itérative : L'Approche SSD Gauss-Newton

Nous nous focalisons dans cette partie sur les approches utilisant une fonctionnelle d'erreur quadratique SSD (*Sum Squared Difference*). Ce type d'approches est répandu et a donné lieu à une large littérature :

$$E(\mathbf{v}) = \mathbf{e}(\mathbf{v})^T \mathbf{e}(\mathbf{v}), \quad (2.31)$$

où  $\mathbf{e}(\mathbf{v})$  est un vecteur d'erreur dépendant de  $\mathbf{v}$  dont chaque entrée  $e_i(\mathbf{v})$  comptabilise l'erreur d'alignement pour les différentes observations mises en correspondance entre les images.  $e_i(\mathbf{v})$  peut être définie de nombreuses manières. Par exemple dans le contexte des approches par *Template Matching*,  $e_i(\mathbf{v})$  est une erreur définie entre les intensités lumineuses des images (cf. sections 2.2.2, 2.2.3 et 2.2.4.1). Dans le contexte d'alignement par mise en correspondance de distributions d'attributs,  $e_i(\mathbf{v})$  représente l'erreur entre deux classes des histogrammes utilisés pour représenter l'information de couleur contenue dans les deux images (cf section 2.2.4.4).

Pour minimiser (2.31), il existe de nombreuses possibilités. Dans le contexte de l'alignement d'images, l'approche Gauss-Newton (GN) permet d'obtenir une vitesse de convergence et une fréquence de convergence moyenne importantes à un coût calculatoire raisonnable comparé à la plupart des approches en ligne. L'approche GN est au centre des travaux d'unification de [Baker 2004] et nous la retrouvons dans de nombreux travaux en *Template Matching* [Lucas 1981,

## 2.2. Alignement par minimisation en ligne d'une fonctionnelle d'erreur

[Hager 1998, Shum 2000, Baker 2001, Buenaposada 2002, Keller 2004, Guangwei 2008] et en suite d'objet par mise en correspondances de distributions d'attributs [Hager 2004, Mégret 2006], c'est pourquoi cette méthode est au centre des travaux présentés dans cette thèse.

Dans la suite nous introduisons les notations suivantes :

$$\mathbf{J}(\mathbf{w}) = \left. \frac{\partial \mathbf{e}(\mathbf{v})}{\partial \mathbf{v}} \right|_{\mathbf{w}}. \quad (2.32)$$

$\mathbf{J}(\mathbf{w})$  est la matrice jacobienne de l'erreur  $\mathbf{e}(\mathbf{v})$  en  $\mathbf{w}$ .

L'algorithme Gauss-Newton repose sur l'approximation de l'erreur par son développement de Taylor à l'ordre 1 :

$$\mathbf{e}(\mathbf{v}) = \mathbf{e}(\mathbf{0}) + \mathbf{J}(\mathbf{0})\mathbf{v} + O(\|\mathbf{v}\|^2). \quad (2.33)$$

En utilisant (2.33) dans (2.31) et en négligeant les termes d'ordres supérieurs à 1, puis en recherchant les points stationnaires de la fonctionnelle obtenue, nous obtenons :

$$\mathbf{v} = -\mathbf{J}(\mathbf{0})^+ \mathbf{e}(\mathbf{0}) \quad (2.34)$$

+ désigne ici l'opérateur pseudo-inverse. Si  $\mathbf{J}(\mathbf{0})$  est de rang plein nous avons :

$$\mathbf{J}(\mathbf{0})^+ = (\mathbf{J}(\mathbf{0})^T \mathbf{J}(\mathbf{0}))^{-1} \mathbf{J}(\mathbf{0})^T = \mathbf{H}(\mathbf{0})^{-1} \mathbf{J}(\mathbf{0})^T. \quad (2.35)$$

$\mathbf{H}(\mathbf{0}) = (\mathbf{J}(\mathbf{0})^T \mathbf{J}(\mathbf{0}))$  est une approximation, par l'approche Gauss-Newton, de la Hessienne de l'erreur  $E(\mathbf{v})$  utilisant uniquement des dérivées d'ordre 1. Dans la suite, nous appelons abusivement ce terme Hessienne bien que ce ne soit qu'une approximation.

Rappelons qu'une fois l'incrément  $\mathbf{v}$  estimé, la mise à jour des paramètres de mouvement est réalisée en utilisant la règle de mise à jour reposant sur la paramétrisation incrémentale (2.27) :

$$\boldsymbol{\mu}^{n+1} = \boldsymbol{\mu}(\boldsymbol{\mu}^n, \mathbf{v}). \quad (2.36)$$

### 2.2.2 L'algorithme de Lucas et Kanade

L'algorithme de Lucas et Kanade [Lucas 1981] est un cas particulier des approches SSD Gauss-Newton (cf. section 2.2.1.3) utilisant une erreur pixel à pixel reposant sur la comparaison directe des intensités lumineuses des images  $I$  et  $T$  :

$$e_i(\mathbf{v}) = I(\mathbf{W}(\boldsymbol{\mu}^n + \mathbf{v}, \mathbf{x}_i)) - T(\mathbf{x}_i). \quad (2.37)$$

$e_i$  est nulle lorsque nous avons trouvé une déformation  $\mathbf{W}(\boldsymbol{\mu}^n + \mathbf{v}, \cdot)$  qui appliquée à  $I$  produit une image résultante semblable à  $T$ . L'erreur (2.37) est utilisée pour construire  $\mathbf{e}(\mathbf{v})$  :

$$\mathbf{e}(\mathbf{v}) = [e_1(\mathbf{v}), \dots, e_N(\mathbf{v})]^T, \quad (2.38)$$

le vecteur d'erreur sur la zone d'intérêt considérée. La zone d'intérêt est ici caractérisée par les positions spatiales  $\{\mathbf{x}_i\}_{i=1..N}$  sélectionnées dans l'image  $T$ . Elle correspond à la portion de  $T$  dont on cherche la zone correspondante dans  $I$ . Deux façons de définir la région d'intérêt sont discutées à la section 3.4.1. Le critère choisi afin de juger de la qualité de l'alignement est la fonctionnelle 2.31 minimisée par l'approche Gauss-Newton.

Dans le cas de l'approche de Lucas et Kanade, la matrice Jacobienne  $\mathbf{J}(\mathbf{0})$  est obtenue en concaténant verticalement les matrices jacobiennes  $\mathbf{J}_{I,i}^A(\mathbf{0})$  associées à l'erreur en chaque pixel (2.37) :

$$\mathbf{J}_{I,i}^A(\mathbf{0}) = \left. \frac{\partial I(\mathbf{W}(\boldsymbol{\mu}^n + \mathbf{v}, \mathbf{x}_i))}{\partial \mathbf{v}} \right|_{\mathbf{0}}. \quad (2.39)$$

La règle de mise à jour des paramètres de mouvement associée est :

$$\boldsymbol{\mu}^{n+1} = \boldsymbol{\mu}^n + \mathbf{v} \quad (2.40)$$

La définition de l'erreur (2.37) conduit à déformer  $I$  vers  $T$ . Les deux images ne jouent donc pas tout à fait le même rôle :  $T$  est utilisée ici comme modèle vers lequel s'aligner d'où l'emploi de la lettre  $T$  (*Template*) pour désigner cette image. Dans un contexte de suivi d'objet par exemple,  $I$  peut désigner l'image courante de la vidéo traitée et  $T$  la première image de la vidéo dans laquelle a été sélectionné l'objet d'intérêt. Cette approche a été nommée **approche FA** dans [Baker 2004] pour *Forwards Additive*, que nous traduisons en français par approche Directe Additive. Le terme Directe est utilisé pour signifier que nous cherchons à aligner l'image courante  $I$  vers l'image de référence  $T$  (dans laquelle est sélectionnée l'objet d'intérêt à travers le choix de la région  $R$ ). Le terme additif correspond à la nature de la paramétrisation incrémentale utilisée.

**L'approche FC (*Forward Compositional*)** (que nous nommons en Français Directe Compositionnelle) est une variante de l'algorithme de Lucas et Kanade attribuée à [Shum 2000]. Elle repose sur une erreur pixel à pixel utilisant une paramétrisation incrémentale compositionnelle :

$$e_i(\mathbf{v}) = I(\mathbf{W}(\boldsymbol{\mu}^n \circ \mathbf{v}, \mathbf{x}_i)) - T(\mathbf{x}_i). \quad (2.41)$$

Cette fois, la matrice jacobienne associée est donc la suivante :

$$\mathbf{J}_{I,i}^C(\mathbf{0}) = \left. \frac{\partial I(\mathbf{W}(\boldsymbol{\mu}^n \circ \mathbf{v}, \mathbf{x}_i))}{\partial \mathbf{v}} \right|_{\mathbf{0}} \quad (2.42)$$

et la règle de mise à jour :

$$\boldsymbol{\mu}^{n+1} = \boldsymbol{\mu}^n \circ \mathbf{v}. \quad (2.43)$$

L'introduction de la composition a deux avantages majeurs : cela permet d'une part l'estimation aisée des paramètres appartenant à des espaces courbes munis d'une structure de groupe (si l'on souhaite utiliser une paramétrisation additive pour ce type d'espace il est nécessaire d'imposer lors de l'optimisation une contrainte supplémentaire pour garantir qu'à l'issue de la règle de mise à jour les paramètres estimés sont toujours dans le groupe), et d'autre part un coût calculatoire moins important puisque une partie de la matrice Jacobienne peut être précalculée [Shum 2000, Baker 2004]. Une adaptation de cette approche sur groupe de Lie est proposée dans [Guangwei 2008].

Différentes évolutions ont été apportées afin d'adapter les approches FA et FC à différents contextes qui font l'objet des paragraphes suivants.

### 2.2.3 Approches à faible temps de calcul

L'algorithme de Lucas et Kanade est gourmand en temps de calcul car à chaque nouvelle image et à chaque itération il est nécessaire de recalculer la matrice jacobienne et sa pseudo-inverse. Afin d'obtenir des algorithmes temps réel, différentes variantes ont été proposées reposant

## 2.2. Alignement par minimisation en ligne d'une fonctionnelle d'erreur

soit sur une définition différente des images fixes et mobiles (section 2.2.3.1) soit sur la définition d'heuristique remplaçant avantageusement l'algorithme Gauss-Newton pour estimer la matrice Jacobienne où Hessienne associée à l'optimisation du critère (section 2.2.3.2). Une autre amélioration a consisté à limiter le nombre de pixels utilisés lors de la minimisation du critère (section 2.2.3.3).

### 2.2.3.1 Modification de la définition des images fixes et mobiles

Dans les approches FA et FC, l'image  $T$  est fixe (aucune déformation ne lui est appliquée) alors que  $I$  est déformée. Selon les contextes applicatifs envisagés, une définition différente des images fixes et mobiles permet de mettre en place des approches plus efficaces.

**L'approche IC (*Inverse Compositional*).** Proposée par [Baker 2004], cette méthode consiste à minimiser la fonctionnelle (2.31) en utilisant :

$$e_i(\mathbf{v}) = I(\mathbf{W}(\boldsymbol{\mu}^n, \mathbf{x}_i)) - T(\mathbf{W}(\mathbf{v}, \mathbf{x}_i)). \quad (2.44)$$

La matrice jacobienne associée vérifie :

$$\mathbf{J}_i(\mathbf{0}) = \mathbf{J}_{T,i}^C(\mathbf{0}) = \left. \frac{\partial T(\mathbf{W}(\mathbf{v}, \mathbf{x}_i))}{\partial \mathbf{v}} \right|_{\mathbf{0}}, \quad (2.45)$$

avec la règle de mise à jour :

$$\boldsymbol{\mu}^{n+1} = \boldsymbol{\mu}^n \circ \mathbf{v}^{-1} \quad (2.46)$$

Cette approche, en appliquant la déformation  $\mathbf{v}$  sur  $T$  au lieu de  $I$ , permet d'entièrement pré-calculer la pseudo-inverse  $(\mathbf{J}(\mathbf{0}))^+$ . En effet à la différence des approches précédentes, la matrice Jacobienne introduite ne dépend plus des paramètres initiaux  $\boldsymbol{\mu}^n$ , mais uniquement des gradients de l'image  $T$  qui peuvent être précalculés dans les situations où  $T$  n'est pas amenée à changer. Cela fait de cette méthode une technique particulièrement efficace, par exemple, pour le suivi d'objets dans une vidéo où l'objet d'intérêt peut être délimité dans l'image de référence  $T$  avant de commencer le suivi. Cette méthode nécessite cependant que la fonction  $\mathbf{W}$  soit une action de groupe ce qui limite son cadre applicatif essentiellement à l'estimation de transformations rigides.

Les approches FA, FC et IC ont été démontrées équivalentes au premier ordre [Baker 2004], justifiant ainsi le choix de la définition de l'erreur alternative (2.44) à l'erreur plus classique (2.37).

**L'approche SGM (*Symmetric Gradient Methods*).** Cette méthode proposée par [Keller 2004] consiste à modifier l'erreur pixel à pixel pour qu'elle fasse intervenir les gradients des deux images. Pour cela  $e_i$  est définie par :

$$e_i(\mathbf{v}) = I(\mathbf{W}(\boldsymbol{\mu}^n \circ \frac{\mathbf{v}}{2}, \mathbf{x}_i)) - T(\mathbf{W}(-\frac{\mathbf{v}}{2}, \mathbf{x}_i)), \quad (2.47)$$

Dans le contexte d'une optimisation Gauss-Newton ils obtiennent la matrice jacobienne associée à un pixel :

$$\mathbf{J}_i(\mathbf{0}) = \frac{1}{2}(\mathbf{J}_{I,i}^C(\mathbf{0}) + \mathbf{J}_{T,i}^C(\mathbf{0})). \quad (2.48)$$

La règle de mise à jour associée est la règle additive (2.43). Comme nous le verrons au chapitre 4.4.4, cette règle de mise à jour est très restrictive et n'est compatible qu'avec un petit nombre de transformations.

**L'approche BDGM (*Bidirectional Gradient Methods*)** Dans [Keller 2004], les auteurs proposent également une deuxième définition de l'erreur pixel à pixel permettant de faire apparaître les gradients des deux images par la prise en compte de la compensation des deux images de façon indépendantes :

$$e_i(\mathbf{v}) = I(\mathbf{W}(\boldsymbol{\mu}^n \circ \mathbf{v}_I, \mathbf{x}_i)) - T(\mathbf{W}(-\mathbf{v}_T, \mathbf{x}_i)). \quad (2.49)$$

Les auteurs appliquent l'approche Gauss-Newton pour l'estimation de la matrice Jacobienne conduisant à la concaténation des matrices Jacobiennes pixel à pixel :

$$\mathbf{J}_i(\mathbf{0}) = [\mathbf{J}_{I,i}^C(\mathbf{0}) \quad \mathbf{J}_{T,i}^C(\mathbf{0})]. \quad (2.50)$$

Tout comme pour l'approche SGM, la loi de mise à jour utilisée est additive et restreint fortement les modèles de mouvement pouvant être estimés (cf. section 4.4.4). Les auteurs montrent théoriquement et expérimentalement que l'utilisation des erreurs (2.47) et (2.49) conduit à une vitesse de convergence plus rapide. Ainsi, si ces critères ne permettent pas de diminuer le temps de calcul par itération, ils permettent de diminuer le nombre d'itérations nécessaires pour converger.

### 2.2.3.2 Introduction d'heuristiques pour l'estimation de la matrice Jacobienne et/ou Hessienne

Dans cette partie, nous introduisons des approches alternatives à l'approche Gauss-Newton permettant, pour une erreur  $e_i$  fixée, de calculer astucieusement les gradients de la fonctionnelle SSD (2.31) associée. Ces approches, tout comme l'approche Gauss-Newton, permettent d'approximer la fonctionnelle SSD par une quadrique en utilisant uniquement des dérivées d'ordre 1 des images mises en jeux. Le terme heuristique est ici employé car ces approches reposent sur des hypothèses supplémentaires par rapport à l'approche Gauss-Newton pour leur justification théorique.

**L'approche IA (*Inverse Additive*).** Les auteurs de [Hager 1998] introduisent une approche différente de l'algorithme de Lucas et Kanade pour approximer l'erreur (2.37). Sous l'hypothèse que les images sont approximativement identiques à l'itération  $n$  :

$$I(\mathbf{W}(\boldsymbol{\mu}^n, \mathbf{x}_i)) \approx T(\mathbf{x}_i), \quad (2.51)$$

les auteurs montrent que l'optimisation se ramène à utiliser dans (2.34) :

$$\mathbf{J}_i(\mathbf{0}) = \nabla T(\mathbf{x}_i) \left( \left. \frac{\partial \mathbf{W}(\boldsymbol{\mu}^n, \mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}_i} \right)^{-1} \left. \frac{\partial \mathbf{W}(\boldsymbol{\mu}, \mathbf{x}_i)}{\partial \boldsymbol{\mu}} \right|_{\boldsymbol{\mu}^n}. \quad (2.52)$$

Cette forme conduit à l'utilisation des gradients calculés sur le template qui peuvent donc être précalculés. De plus, les auteurs montrent que pour un certain nombre de transformations, dont

## 2.2. Alignement par minimisation en ligne d'une fonctionnelle d'erreur

les transformations affines, la matrice jacobienne peut se factoriser comme le produit entre une matrice dépendant uniquement des positions spatiales  $\mathbf{x}_i$  et une matrice dépendant des paramètres courants  $\boldsymbol{\mu}^n$ . Dans ce cas, la matrice dépendant des positions spatiales et des gradients de  $T$  peut être précalculées. Ainsi seule la seconde matrice nécessite d'être calculée à chaque itération ce qui conduit à une complexité calculatoire très faible. La règle de mise à jour utilisée est :

$$\boldsymbol{\mu}^{n+1} = \boldsymbol{\mu}^n + \mathbf{v}. \quad (2.53)$$

Avec cette approche les auteurs considèrent que seule  $I$  est mobile (déformée). Cependant l'optimisation proposée inverse le rôle de  $I$  et  $T$  pour le calcul des gradients. Pour cette raison cette approche est nommée IA (*Inverse Additive*) dans [Baker 2004].

**L'approche ESM (Efficient Second-order Minimization).** A la différence de l'approche précédente qui diminue le coût calculatoire par itération, l'intérêt de l'algorithme ESM proposé par [Benhimane 2004, Benhimane 2007b] repose sur une diminution du nombre d'itérations nécessaires pour converger. Cette approche consiste à minimiser (2.31) en utilisant cette fois une paramétrisation par groupe de Lie dans l'erreur pixel à pixel :

$$e_i(\mathbf{v}) = I(\mathbf{W}(\boldsymbol{\mu}^n \circ \exp(\mathbf{v}), \mathbf{x}_i)) - T(\mathbf{x}_i) \quad (2.54)$$

En faisant l'hypothèse que pour les paramètres de transformation vrais  $\mathbf{v}^*$  nous avons :

$$\forall \mathbf{x}_i \quad I(\mathbf{W}(\boldsymbol{\mu}^n \circ \exp(\mathbf{v}^*), \mathbf{x}_i)) = T(\mathbf{x}_i), \quad (2.55)$$

et en utilisant les propriétés de l'application exponentielle (2.17)(2.18), les auteurs montrent que l'utilisation de la matrice Jacobienne :

$$\mathbf{J}(\mathbf{0}) = \frac{1}{2}(\mathbf{J}_I^{CL}(\mathbf{0}) + \mathbf{J}_T^{CL}(\mathbf{0})) \quad (2.56)$$

avec :

$$\mathbf{J}_{I,i}^{CL}(\mathbf{0}) = \left. \frac{\partial I(\mathbf{W}(\boldsymbol{\mu}_I^n \circ \exp(\mathbf{v}), \mathbf{x}_i))}{\partial \mathbf{v}} \right|_{\mathbf{0}}, \text{ et :} \quad (2.57)$$

$$\mathbf{J}_{T,i}^{CL}(\mathbf{0}) = \left. \frac{\partial T(\mathbf{W}(\exp(\mathbf{v}), \mathbf{x}_i))}{\partial \mathbf{v}} \right|_{\mathbf{0}}, \quad (2.58)$$

conduit à une approximation de  $\mathbf{e}(\mathbf{v})$  à l'ordre 2. La démonstration de cette propriété est présentée dans [Benhimane 2007b, Vercauteren 2007]. Une généralisation de cette démonstration sera proposée à la section 5.1.1.1.

Nous pouvons remarquer ici que la matrice jacobienne dépend des gradients de  $I$ , on perd donc la possibilité du précalcul introduite par l'approche Inverse Compositionnelle. L'intérêt de l'ESM est d'approximer la fonctionnelle d'erreur de manière plus fine que l'approche Gauss-Newton, ce qui conduit à une vitesse de convergence plus rapide (nombre d'itérations nécessaires avant convergence). [Malis 2009] revient sur les propriétés de convergence de l'ESM et montre que cette approche a une vitesse de convergence cubique, tandis que l'approche de Gauss-Newton a au mieux une vitesse de convergence quadratique. De plus [Mei 2008] montre expérimentalement que malgré un coût calculatoire par itération de l'ESM plus élevé que celui de l'approche

Hypothèse	Amélioration	Référence
Considérer $T$ comme image mobile (équation (2.44))	temps de calcul par itération	[Baker 2001, Baker 2004]
Considérer $I$ et $T$ comme images mobiles (équations (2.47)(2.49))	nombre d'itérations avant convergence	[Keller 2004]
Similarité des deux images à l'itération $n$ (équation (2.51))	temps de calcul par itération	[Hager 1998]
Similarité des deux images à convergence (équation (2.55))	nombre d'itérations avant convergence	[Benhimane 2004, Benhimane 2007b]
Un sous-ensemble de pixels est suffisamment représentatif	temps de calcul par itération	[Averbuch 2003, Baker 2003a] [Benhimane 2007a, Brooks 2007]

TABLE 2.1 – Principales évolutions de l'algorithme de Lucas et Kanade [Lucas 1981] pour améliorer les performances en termes de temps de calcul.

Inverse Compositionnelle [Baker 2004], le coût calculatoire global avant convergence (nombre d'itération  $\times$  coût par itération) est en faveur de l'ESM. Enfin, cette approche conduit aussi à une plus grande zone de convergence permettant d'estimer des mouvements plus importants.

### Les approches GIC (*Generalized Inverse Compositional*) et GESM (*Generalized ESM*).

Dans [Brooks 2009], les auteurs remarquent que l'avantage des approches IC et ESM repose sur l'utilisation de la matrice Jacobienne du template  $T$  permettant soit un coût calculatoire par itération plus faible, soit d'améliorer les propriétés de convergence en estimant de manière adaptée les gradients de la fonctionnelle d'erreur. Cependant ces algorithmes nécessitent une architecture spécifique d'optimisation : ils ont été développés tous les deux avec une règle de mise à jour compositionnelle. De plus l'ESM repose sur l'utilisation d'une paramétrisation de l'incrément dans une algèbre de Lie. Afin d'étendre ces approches à des techniques d'optimisation utilisant une règle de mise à jour additive et une paramétrisation du mouvement sur un groupe quelconque, les auteurs proposent une procédure permettant de traduire la matrice Jacobienne associée à une image en fonction de paramètres de mouvement affectant l'autre image. Cela leur permet de définir une généralisation de l'approche IC [Baker 2004] et de l'ESM [Benhimane 2007b] qu'ils nomment *Generalized Inverse Compositional* et *Generalized ESM* utilisant toutes deux une règle de mise à jour additive. Le formalisme développé au chapitre 3 nous permettra de revenir plus en détail sur ces approches à la section 4.2.2.4, 4.3.1.2 et en annexe B.

#### 2.2.3.3 Sélection de pixels

Une approche complémentaire permettant de limiter le temps de calcul est d'utiliser uniquement un sous-ensemble de l'information visuelle disponible en sélectionnant astucieusement les positions spatiales de l'image de référence  $T$  que nous souhaitons mettre en correspondance. Une approche simple à mettre en place est d'utiliser uniquement les pixels de fort gradient dans l'image de référence  $T$  [Averbuch 2003, Baker 2003a]. Des méthodes déterminant, par apprentissage, un ensemble de pixels adapté peuvent également être envisagées. Ainsi [Benhimane 2007a],



## 2.2. Alignement par minimisation en ligne d'une fonctionnelle d'erreur

constatant que les algorithmes d'optimisation des sections 2.2.2 et 2.2.3 reposent sur une approximation linéaire (par exemple approche directes et inverses) ou bien quadratique (approche ESM) de l'erreur  $\mathbf{e}(\mathbf{v})$ , ils proposent par apprentissage de déterminer les pixels du template  $T$  pour lesquels l'approximation de l'erreur n'affecte pas trop le processus d'alignement en terme de précision. Dans [Brooks 2007], les auteurs abordent le problème de la sélection des pixels comme un problème de compromis entre la précision souhaitée et le temps de calcul autorisé.

Dans le tableau 2.1, nous présentons la synthèse des améliorations apportées à l'algorithme de Lucas et Kanade en termes d'efficacité calculatoire.

### 2.2.4 Modification du critère d'alignement pour être robuste à des perturbations photométriques

Bien que les plus répandues, les approches utilisant une erreur SSD ne sont pas toujours adaptées aux images observées. D'autres critères ont été proposés pour être invariant à un certain nombre de problèmes pratiques. L'optimisation de ces critères pouvant reposer sur des approches similaires à celles proposées dans le contexte des fonctionnelles SSD, nous détaillerons principalement le critère utilisé qui constitue la modification principale expliquant l'intérêt de ces approches.

#### 2.2.4.1 Variations de luminosité

Dans [Evangelidis 2008] les auteurs cherchent à maximiser le coefficient de corrélation normalisé noté ECC (*Enhanced Correlation Coefficient*) afin d'être robuste aux variations de luminosité. Ce critère est en effet invariant à l'application d'un gain et d'un offset sur les valeurs d'intensité des images considérées. Ils définissent notamment une variante de l'approche FA noté FA-ECC dont l'objectif est de maximiser le critère :

$$C(\mathbf{v}) = \frac{\sum_{i=1}^N (I(\mathbf{W}(\boldsymbol{\mu}^n + \mathbf{v}, \mathbf{x}_i)) - \bar{I})(T(\mathbf{x}_i) - \bar{T})}{\sqrt{\sum_{i=1}^N (I(\mathbf{W}(\boldsymbol{\mu}^n + \mathbf{v}, \mathbf{x}_i)) - \bar{I})^2} \sqrt{\sum_{i=1}^N (T(\mathbf{x}_i) - \bar{T})^2}} \quad (2.59)$$

en fonction de  $\mathbf{v}$ .  $\bar{T}$  et  $\bar{I}$  désignent les valeurs moyennes de l'intensité lumineuse dans la région d'intérêt (sur laquelle se fait l'alignement) dans le template et dans l'image. Les auteurs proposent leur propre algorithme d'optimisation qui repose sur un développement de Taylor à l'ordre 1 de l'image déformée  $\bar{I}(\boldsymbol{\mu} + \mathbf{v})$ . En utilisant les mêmes principes, les auteurs combinent également le critère ECC et l'approche inverse compositionnelle donnant naissance à l'approche IC-ECC.

Une approche similaire est envisagée dans [Chen 2009] qui inclut dans le critère une connaissance des statistiques du bruit affectant les images. Les auteurs s'intéressent à l'estimation de translations et considèrent les images perturbées par deux bruits blancs Gaussien de moyenne nulle et de variance respective  $\sigma_I^2$  et  $\sigma_T^2$ . Sous ces hypothèses, le coefficient de corrélation peut s'écrire :

$$C^2(\mathbf{v}) = \frac{\sum_{i=1}^N (I(\mathbf{x}_i + \mathbf{v}) - \bar{I})(T(\mathbf{x}_i) - \bar{T})^2}{(\sum_{i=1}^N ((I(\mathbf{x}_i + \mathbf{v}) - \bar{I})^2 - N\sigma_I^2)) \sum_{i=1}^N ((T(\mathbf{x}_i) - \bar{T})^2 - N\sigma_T^2)} \quad (2.60)$$

où  $N$  est le nombre de pixels considérés dans la zone d'intérêt. Afin d'assurer, lors de la résolution, une utilisation symétrique des informations statistiques du bruit affectant les images, les auteurs



modifient le critère (2.60) afin d'appliquer symétriquement sur chaque image la déformation spatiale :

$$C^2(\mathbf{v}) = \frac{\sum_i^N (I(\mathbf{x}_i + \frac{\mathbf{v}}{2}) - \bar{I})(T(\mathbf{x}_i - \frac{\mathbf{v}}{2}) - \bar{T})^2}{\sum_{i=1}^N ((I(\mathbf{x}_i + \frac{\mathbf{v}}{2}) - \bar{I})^2 - N\sigma_I^2) \sum_{i=1}^N ((T(\mathbf{x}_i - \frac{\mathbf{v}}{2}) - \bar{T})^2 - N\sigma_T^2)} \quad (2.61)$$

Les auteurs approximent alors le coefficient par un développement de Taylor à l'ordre 1 et étudient les points stationnaires associés à la fonctionnelle approximée.

### 2.2.4.2 Multi-modalité et variations de luminosité non-linéaires

D'autres approches ont été développées dans le contexte d'images multi-modales s'appuyant sur l'optimisation de l'information mutuelle entre les images [Viola 1997, Maes 1997]. Ces approches reposent sur l'estimation de la densité de probabilité conjointe de l'attribut visuel. Cette probabilité est représentée par un histogramme bidimensionnel dont chaque bin est identifié par un couple  $(u, v)$ . Les entrées de l'histogramme vérifient :

$$p_{u,v}(I, T, \boldsymbol{\mu}) = \frac{1}{N} \sum_{i=1..N} \delta_u(I(\mathbf{W}(\boldsymbol{\mu}, \mathbf{x}_i))) \delta_v(T(\mathbf{x}_i)), \quad (2.62)$$

où  $\delta_w$  est une fonction de quantification valant 1 si l'intensité lumineuse passée en argument appartient au bin  $w$ . Ces approches conservent à travers cette probabilité conjointe une information sur la configuration spatiale des images. En notant :

$$p_u(I, \boldsymbol{\mu}) = \frac{1}{N} \sum_{i=1..N} \delta_u(I(\mathbf{W}(\boldsymbol{\mu}, \mathbf{x}_i))), \text{ et} \quad (2.63)$$

$$p_v(T, \mathbf{0}) = \frac{1}{N} \sum_{i=1..N} \delta_v(T(\mathbf{x}_i)), \quad (2.64)$$

les densités marginales associées à chaque image, la fonctionnelle d'erreur est alors définie par :

$$E(\boldsymbol{\mu}) = - \sum_{u,v} p_{u,v}(I, T, \boldsymbol{\mu}) \log \left( \frac{p_{u,v}(I, T, \boldsymbol{\mu})}{p_u(I, \boldsymbol{\mu}) p_v(T, \mathbf{0})} \right). \quad (2.65)$$

Ces approches ont été récemment adaptées pour des problématiques de vision comme le suivi d'objet [Dowson 2008] (adaptation de l'approche IC) ou d'asservissement visuel [Dame 2009] du fait de leur grande robustesse à une variété de transformations non-linéaires affectant les intensités des images, comme en présence de fortes variations de luminosité ou bien d'occultation.

### 2.2.4.3 Occultations

Les M-estimateurs permettent une estimation robuste aux occultations. Ces fonctionnelles ont été initialement introduites par [Black 1993] pour l'estimation du flot optique. L'idée repose sur l'utilisation d'une fonction statistique  $\rho$  permettant d'écarter les mesures aberrantes non prises en compte par le modèle. Le critère à optimiser devient :

$$E(\mathbf{v}) = \sum_i \rho(\mathbf{e}_i(\mathbf{v})) \quad (2.66)$$

## 2.2. Alignement par minimisation en ligne d'une fonctionnelle d'erreur

---

où  $e_i(\mathbf{v})$  caractérise l'erreur d'alignement pixel à pixel et peut être définie de plusieurs façons comme par (2.37) dans [Hager 1998, Dufaux 2000] ou encore (2.44) dans [Baker 2003a].

L'équation (2.66) est une généralisation de l'erreur SSD, puisque en choisissant  $\rho$  telle que :  $\rho(t) = t^2$  on retrouve la fonctionnelle SSD (2.31). En pratique, une fonction  $\rho$  qui croît moins vite que la fonction  $t \mapsto t^2$  est utilisée afin de pondérer l'influence des pixels pour lesquels l'erreur  $e_i$  est importante et qui correspondent potentiellement à des mesures aberrantes. Plusieurs exemples de fonctions sont étudiées dans [Black 1996].

Dans le cas où la fonction  $\rho$  est symétrique à savoir  $\rho(t) = -\rho(-t)$ , l'approximation du critère par un développement de Taylor conduit à l'algorithme *IRLS* (*Iterative Reweighted Least Square*) très proche de l'algorithme Gauss-Newton SSD. En effet la seule différence est que la contribution de chaque pixel  $i$  aux matrices Jacobienne et Hessienne est pondérée par  $\rho(e_i(\mathbf{0}))$ .

### 2.2.4.4 Déformations géométriques parasites

Les méthodes proposées jusqu'à présent reposent sur la comparaison directe des intensités lumineuses entre image. Lorsque les images ne respectent pas exactement le modèle de transformation géométrique choisi (par exemple dans le contexte de suivi de personnes pour de la vidéo surveillance), ces méthodes conduisent à de mauvaises estimation car le modèle géométrique utilisé ne leur permet pas de mettre en correspondance les observations issues de chaque image. Une approche populaire développée ces dernières années, notamment pour le suivi d'objets, repose sur la mise en correspondance d'attributs visuels [Comaniciu 2003]. Cette approche combine la robustesse apportée par les mesures de similarité entre distributions d'attributs (principalement des histogrammes de couleur) avec des algorithmes d'optimisation itératif conduisant à des systèmes performants. A la différence des critères évoqués aux sections précédentes, la mise en correspondance d'histogrammes conduit à intégrer spatialement l'information permettant un meilleur comportement de l'approche dans le cas où la scène observée subit des transformations géométriques non modélisées. Ces approches sont donc adaptées à l'estimation du mouvement 3D ou de suivi d'objets déformables [Guskov 2006].

L'utilisation d'histogrammes de couleur ne permet pas de conserver des informations sur la configuration spatiale de l'objet, car les pixels sont regroupés par couleur sans tenir compte de leur position, ce qui rend difficile le suivi de déplacements complexes. [Comaniciu 2003] ajoute une information de nature spatiale en utilisant un noyau donnant un poids plus élevé aux pixels proches du centre de l'objet. Cependant les noyaux symétriques utilisés sont peu sensibles aux rotations. Des approches utilisant plusieurs noyaux ont été développées pour résoudre ce problème [Hager 2004, Guskov 2006, Mégret 2006]. Ces approches introduisent une information contextuelle en positionnant plusieurs noyaux sur l'objet d'intérêt.

Ainsi l'objet est représenté par un ensemble de  $\mathcal{K}$  descripteurs correspondant à différentes régions spatiales. L'information portée par chaque région est intégrée spatialement à l'aide d'une fonction de pondération appelée noyau et notée  $K_k(\mathbf{x})$  qui définit également le support spatial de la région considérée :

$$q_{k,u}(I, \boldsymbol{\mu}) = C_k^I \sum_{i=1}^N K_k(\mathbf{x}_i) \delta_u(I(\mathbf{W}(\boldsymbol{\mu}, \mathbf{x}_i))) \quad (2.67)$$

Contexte	Fonctionnelle	Référence
Variation d'éclairage	Coefficient de corrélation normalisé	[Evangelidis 2008, Chen 2009]
Images multimodales	Information mutuelle	[Viola 1997, Maes 1997] [Dowson 2008, Dame 2009]
Occultation	M-estimateur	[Hager 1998, Dufaux 2000, Baker 2003a]
Déformations géométriques parasites	Métriques sur des distributions d'attributs	[Hager 2004, Mégret 2006] [Comaniciu 2003, Guskov 2006]

TABLE 2.2 – Principales fonctionnelles utilisées pour adapter l'algorithme de Lucas et Kanade à des images de différentes natures.

avec  $\delta_u$  une fonction de quantification valant 1 si l'intensité lumineuse passée en argument appartient à la classe  $u$  de l'histogramme.  $C_k^I$  est une constante de normalisation choisie de telle sorte que  $\sum_{k,u} q_{k,u} = 1$ . En particulier le template  $T$  est caractérisé par :

$$p_{k,u} = q_{k,u}(T, \mathbf{0}) = C_k^T \sum_{i=1}^N K_k(\mathbf{x}_i) \delta_u(T(\mathbf{x}_i)) \quad (2.68)$$

Deux fonctionnelle principales ont été proposées pour mesurer la similarité des histogrammes associés aux images  $I$  et  $T$ . La première repose sur le coefficient de Bhattacharyya [Comaniciu 2003, Guskov 2006] et peut s'écrire :

$$F_B(\boldsymbol{\mu}) = \sum_{k=1}^{\mathcal{K}} \sum_{u=1}^m \sqrt{p_{k,u} q_{k,u}(I, \boldsymbol{\mu})}. \quad (2.69)$$

La seconde repose sur la distance de Matusita [Hager 2004, Mégret 2006] :

$$E_M(\boldsymbol{\mu}) = \mathbf{e}(\boldsymbol{\mu})^T \mathbf{e}(\boldsymbol{\mu}), \quad (2.70)$$

avec  $\mathbf{e}(\boldsymbol{\mu})$  le vecteur obtenu par la concaténation des erreurs  $e_{k,u}$  définies par :

$$e_{k,u} = \sqrt{q_{k,u}(I, \boldsymbol{\mu})} - \sqrt{p_{k,u}}, \quad (2.71)$$

Ces deux fonctionnelles sont optimisées par des méthodes similaires à celles présentées dans les paragraphes précédents. On note notamment l'existence d'une adaptation de l'approche de Lucas et Kanade (approche FA) dans les travaux de [Hager 2004] et de l'approche de Baker et Matthews [Baker 2004] (approche IC) dans les travaux de [Mégret 2006].

Le tableau 2.2 propose une synthèse des principales fonctionnelles d'erreur alternatives proposées pour adapter l'algorithme de Lucas et Kanade à différent contexte.

### 2.3 L'alignement d'images, un problème asymétrique ?

Une question essentielle lorsque nous cherchons à aligner deux images et de définir la meilleure façon d'utiliser l'information visuelle présente dans chaque image pour arriver à nos fins. Dans

## 2.3. L’alignement d’images, un problème asymétrique ?

de nombreuses applications comme la super-résolution, le suivi d’objet 3D par alignement à un modèle (joue le rôle de template), la création d’images mosaïques, nous sommes souvent amenés à aligner deux images de qualités différentes. Pour ce type d’application, il semble naturel de vouloir privilégier l’information provenant de l’image de meilleure qualité et l’alignement d’images apparaît alors comme un problème asymétrique.

Cet objectif peut se traduire au moins à deux niveaux différents au sein d’un algorithme : dès la définition de la fonctionnelle à optimiser, et dans la façon de l’optimiser.

### 2.3.1 Asymétrie lors de la formalisation du problème d’alignement

Nous avons vu précédemment que les auteurs avaient proposé des algorithmes reposant sur différentes définitions des images fixes et mobiles. Dans le contexte de l’alignement SSD, nous avons distingué notamment les fonctionnelles Directes [Lucas 1981, Shum 2000, Benhimane 2007b], Inverses [Baker 2004, Bayro-Corrochano 2007], Symétriques et Bidirectionnelles [Keller 2004]. Comme nous l’illustrons maintenant avec les approches directe et inverse compositionnelle, ces fonctionnelles ne sont pas rigoureusement équivalentes car elles proposent d’aligner les images dans des repères différents.

**Problématique du changement de repère.** Supposons de manière idéale que  $I$  et  $T$  soient deux images à support infini et continues et que la fonction de déformation  $\mathbf{W}$  soit de classe  $C^1$ . L’approche directe compositionnelle FC peut être exprimée dans le domaine continu par :

$$E^{\text{FC}}(\mathbf{v}) = \iint_{\mathbf{x}} (I(\mathbf{W}(\boldsymbol{\mu}^n \circ \mathbf{v}, \mathbf{x})) - T(\mathbf{x}))^2 d\mathbf{x}, \quad (2.72)$$

et l’approche inverse compositionnelle IC par :

$$E^{\text{IC}}(\mathbf{v}) = \iint_{\mathbf{z}} (I(\mathbf{W}(\boldsymbol{\mu}^n, \mathbf{z})) - T(\mathbf{W}(\mathbf{v}, \mathbf{z})^{-1}))^2 d\mathbf{z}. \quad (2.73)$$

Les deux fonctionnelles précédentes ne sont pas rigoureusement équivalentes. En effet, en appliquant le changement de variable  $\mathbf{z} = \mathbf{W}(\mathbf{v}, \mathbf{x})$  dans (2.72), nous obtenons l’expression équivalente suivante :

$$E^{\text{FC}}(\mathbf{v}) = \iint_{\mathbf{z}} (I(\mathbf{W}(\boldsymbol{\mu}^n, \mathbf{z})) - T(\mathbf{W}(\mathbf{v}, \mathbf{z})^{-1}))^2 j(\mathbf{v}, \mathbf{z}) d\mathbf{z} \quad (2.74)$$

où :

$$j(\mathbf{v}, \mathbf{z}) = \det \left( \frac{\partial \mathbf{W}^{-1}(\mathbf{v}, \mathbf{z})}{\partial \mathbf{z}} \Big|_{\mathbf{z}=\mathbf{W}(\mathbf{v}, \mathbf{x})} \right), \quad (2.75)$$

caractérise le changement de volume lors du passage des coordonnées  $\mathbf{x}$  dans le template aux coordonnées  $\mathbf{z}$  dans l’image déformée par  $\boldsymbol{\mu}_T^n$ . D’un point de vue pratique, ce terme de changement de volume est souvent négligé dans les applications de suivi d’objets et la fonctionnelle (2.73) est utilisée ce qui permet de diminuer la complexité calculatoire. [Baker 2004, Dedeoglu 2007] ont discuté la validité de cette approximation. En dehors des cas où le mouvement implique de gros changement d’échelles d’une image à l’autre, ce terme a peu d’influence. C’est le cas lorsque l’incrément  $\mathbf{v}$  est proche de l’élément neutre, c’est pourquoi il est négligé en pratique dans le cadre d’une optimisation itérative.

Le terme de changement de volume introduit précédemment permet de normaliser l'erreur inverse pour qu'elle soit identique à l'erreur directe. Cela revient donc à considérer que le repère le plus adapté pour l'alignement est le repère du template (celui vers lequel s'effectue l'alignement). En l'absence d'informations supplémentaires sur les deux images à aligner, il n'y a aucune raison de privilégier un repère de coordonnées plutôt qu'un autre. C'est pourquoi de nombreux auteurs, notamment en imagerie médicale [Christensen 2001, Cachier 2000] ont formalisé le problème d'alignement comme la minimisation de la somme d'une fonctionnelle directe et inverse :

$$E^{\text{FISC}}(\boldsymbol{\mu}) = \iint_{\mathbf{x}} (I(\mathbf{W}(\boldsymbol{\mu}, \mathbf{x})) - T(\mathbf{x}))^2 d\mathbf{x} + \iint_{\mathbf{z}} (I(\mathbf{z}) - T(\mathbf{W}(\boldsymbol{\mu}, \mathbf{z})^{-1}))^2 d\mathbf{z} \quad (2.76)$$

Remarquons ici que la partie inverse de cette fonctionnelle est différente de la fonctionnelle inverse (2.73) puisque reposant sur le changement de variable  $\mathbf{z} = \mathbf{W}(\boldsymbol{\mu}, \mathbf{x})$ . On nomme dans la suite *Forward Inverse Symmetric Combination* FISC la fonctionnelle (2.76). Cette dernière considère les deux images à aligner comme interchangeables (l'inversion du rôle de  $I$  et  $T$  dans le critère (2.76) conduit à l'estimation d'un  $\boldsymbol{\mu}$  exactement inverse de celui intervenant dans (2.76)).

Cependant, cette erreur est rarement minimisée sous cette forme dans la communauté médicale car elle implique le calcul de l'inverse de la transformation ce qui peut être difficile à obtenir pour des modèles déformables. C'est pourquoi les auteurs [Cachier 2000, Tagare 2009] introduisent un terme de changement de volume similaire à celui introduit à l'équation (2.74) afin de transformer la partie inverse de la fonctionnelle en une fonctionnelle directe équivalente ce qui conduit à l'erreur totale suivante :

$$E^{\text{FISC}}(\boldsymbol{\mu}) = \iint_{\mathbf{x}} (I(\mathbf{W}(\boldsymbol{\mu}, \mathbf{x})) - T(\mathbf{x}))^2 (1 + j(\boldsymbol{\mu}^{-1}, \mathbf{x})) d\mathbf{x} \quad (2.77)$$

Pour des transformations non rigides impliquant des changements d'échelle, le terme de changement de volume peut avoir une influence sur le résultat final pour des déformations importantes en modifiant l'importance relative des différentes zones du support spatial. C'est pourquoi ce terme est rarement négligé dans cette communauté. Cela a donné lieu à une étude détaillée de la signification et de l'influence de ce terme [Tagare 2009].

**Problématique du changement de résolution.** Les équations (2.72) et (2.74) illustrent comment le théorème de changement de variable permet de rendre les fonctionnelles Directe Compositionnelle et Inverse Compositionnelle équivalentes, en présence d'une déformation purement géométrique sur des images définies sur un espace continu et de résolution infinie. Dans [Dedeoglu 2007], les auteurs s'intéressent au choix de l'image à déformer lorsque l'on est en présence d'images reliées par un changement d'échelle. Les auteurs montrent qu'il faut déformer l'image de plus haute résolution vers l'image de plus basse résolution, en prenant en compte un filtrage anti-repliement modélisant l'intégration spatiale sur les pixels du capteur, pour éviter d'obtenir une estimation biaisée de la transformation. Ils proposent un algorithme *Resolution Aware Formulation* fonctionnant selon ce principe pour le suivi de visages.

### 2.3.2 Asymétrie lors de la résolution du problème d'alignement

Lors de la minimization du critère choisi, nous avons identifié deux causes potentielles pratiques d'asymétrie.

### 2.3. L’alignement d’images, un problème asymétrique ?

**Choix de la région d’intérêt.** En pratique, une image est bornée et seule une portion de l’image, appelée région d’intérêt et notée  $R$ , est utilisée pour réaliser l’alignement. Dans l’expression (2.72), la somme est donc limitée à l’ensemble des  $\mathbf{x} \in R$ . Pour que (2.74) soit équivalent à (2.72), la somme doit être effectuée théoriquement sur l’ensemble  $R' = \{\mathbf{z}, \mathbf{z} = \mathbf{W}(\mathbf{v}, \mathbf{x}) \text{ pour } \mathbf{x} \in R\}$  qui dépend de l’inconnue  $\mathbf{v}$ . Une approximation à l’ordre 0 est donc introduite  $R' \approx \{\mathbf{z}, \mathbf{z} = \mathbf{W}(\mathbf{0}, \mathbf{x}) \text{ pour } \mathbf{x} \in R\}$ . Remarquons cependant qu’en pratique,  $\mathbf{v}$  étant supposé infinitésimal, cette approximation n’a que peu d’influence sur les résultats.

**Choix des gradients utilisés.** Une seconde raison d’asymétrie produisant des performances vraiment différentes en pratique est le choix des gradients utilisés. En effet, le calcul des gradients s’opère par dérivation de l’image. Ces derniers sont donc potentiellement sensibles à la présence de bruit, notamment haute fréquence. Si nous revenons aux approches présentées précédemment section 2.2, nous constatons que soit :

- la matrice jacobienne  $\mathbf{J}$  de l’erreur dépend uniquement des gradients de l’image  $I$  (les approches appelées Directes (*Forwards*) par [Baker 2004]),
- la matrice jacobienne  $\mathbf{J}$  dépend uniquement des gradients du Template (les approches Inverses [Baker 2004]),
- la matrice jacobienne  $\mathbf{J}$  dépend de manière égale des gradients des deux images (SGM [Keller 2004], ESM [Benhimane 2007b]).

Dans les deux premiers cas, le problème semble résolu en faisant l’hypothèse implicite qu’une des deux images possède des informations de meilleure qualité. Dans une situation de suivi d’objet par exemple, ou un template de bonne qualité a été obtenu soit par la connaissance d’un modèle idéal, soit de manière ad hoc par des approches de fusion de plusieurs images d’exemple, le parti pris pour une approche inverse peut être justifié. Cependant, la construction d’un template off-line n’est pas toujours possible.

[Sabuncu 2009] propose de minimiser plutôt le critère obtenu par la moyenne des fonctionnelles suivante :

$$E^{\text{FIAC}}(\boldsymbol{\mu}) = \frac{1}{2} \iint_{\mathbf{x}} (I(\mathbf{W}(\boldsymbol{\mu}, \mathbf{x})) - T(\mathbf{x}))^2 d\mathbf{x} + \frac{1}{2} \iint_{\mathbf{z}} (I(\mathbf{z}) - T(\mathbf{W}(\boldsymbol{\mu}, \mathbf{z})^{-1}))^2 j(\boldsymbol{\mu}, \mathbf{z}) d\mathbf{z} \quad (2.78)$$

Le critère ainsi défini que l’on nomme fonctionnelle *Forward Inverse Asymmetric Combination* est rigoureusement équivalent au critère direct compositionnel classique (2.72) si l’on se place dans un cadre itératif compositionnel en posant  $\boldsymbol{\mu} = \boldsymbol{\mu}^n \circ \mathbf{v}$ , tout en conduisant en pratique à une résolution utilisant les gradients des deux images. Ici, bien que les gradients des deux images soient utilisés, cette approche ne considère pas les deux images comme interchangeables. En effet, en raison du terme de changement de volume, le poids associé aux gradients de  $I$  et  $T$  n’est pas le même. L’utilisation des gradients des deux images pour l’optimisation d’un critère direct est aussi effectuée dans l’approche ESM [Benhimane 2007b]. Les liens entre ces approches sont discutés en annexe E car ils nécessitent d’abord l’introduction de plusieurs concepts des chapitres 3 et 4.

Les approches SGM et ESM utilisent les gradients des deux images de manière parfaitement symétrique ce qui est justifié si les deux images sont de même qualité. Cependant il arrive fréquemment en pratique, par exemple dans un contexte de suivi d'objet ou bien dans un contexte de construction de panoramas, que les images soient de qualité différente. Doit on alors utiliser uniquement les gradients de l'image de meilleure qualité? Doit-on favoriser une approche mélangeant les gradients des deux images de manière homogène?

### 2.4 Conclusion

Nous avons précédemment dressé un état de l'art relatif au problème d'alignement d'images. Nous nous sommes notamment intéressés aux approches d'alignement directe en nous focalisant sur les approches de recalage de modèle utilisant une fonctionnelle SSD.

L'analyse des travaux antérieurs nous a permis d'identifier un certain nombre de problématiques ouvertes. D'une part, concernant la formulation du problème d'alignement, la taxonomie proposée par [Baker 2004] ne permet pas d'intégrer des approches plus récentes comme celles présentées dans les travaux de [Keller 2004][Benhimane 2007b]. Un travail d'unification est donc à envisager. D'autre part, nous avons mis en évidence que, à la fois lors de la sélection d'un critère d'alignement, mais également dans le choix de la procédure d'optimisation de ce dernier, les approches proposées font un a priori concernant la qualité relative des informations portées par les images.

Dans cette thèse, nous nous attacherons donc à apporter des réponses à ces deux problématiques en proposant :

- un nouveau cadre théorique, nommé formalisme *bidirectionnel*, permettant d'unifier les approches d'alignement paramétrique (section 3).
- une étude détaillée de la réinterprétation des méthodes existantes à l'aune de ce formalisme 4.
- deux nouvelles approches s'adaptant à la qualité des images en présence d'images affectées par du bruit (section 5).
- une étude expérimentale exhaustive comparant les nouvelles approches aux méthodes de l'état de l'art (section 6).

# Formalisation bidirectionnelle du problème d'alignement géométrique

---

## Sommaire

---

<b>3.1</b>	<b>Formalisation idéale</b>	<b>35</b>
3.1.1	Formalisation usuelle	35
3.1.2	Formalisation bidirectionnelle	36
<b>3.2</b>	<b>Formalisation pratique</b>	<b>36</b>
3.2.1	Définition d'une région d'intérêt comme support d'alignement	36
3.2.2	Définition d'une mesure de similarité	37
<b>3.3</b>	<b>Contrainte sur l'espace bidirectionnel</b>	<b>40</b>
3.3.1	Taxonomie	40
3.3.2	Nature de la paramétrisation incrémentale	41
<b>3.4</b>	<b>Mise à jour des paramètres</b>	<b>44</b>
3.4.1	Règle de mise à jour et applications	45
3.4.2	Convergence	48
<b>3.5</b>	<b>Étapes de conception d'un algorithme d'alignement d'images</b>	<b>48</b>

---

Dans cette partie, nous développons un nouveau formalisme pour unifier les différentes techniques d'alignement reposant sur une optimisation locale. Des concepts génériques sont introduits pour élargir la classification des approches proposée dans [Baker 2004]. Nous identifions enfin les différents choix à réaliser pour la conception d'algorithmes d'alignement itératifs.

Dans tout ce chapitre, nous faisons l'hypothèse que l'espace des paramètres  $\mathcal{G}$  est un groupe qui agit sur les positions spatiales à travers l'action de groupe  $\mathbf{W}$  (cf. section 2.1.2.2).

## 3.1 Formalisation idéale

Dans cette partie, nous supposons que  $I$  (resp.  $T$ ) est une fonction à support spatial infini qui associe à une position spatiale  $\mathbf{x}$  l'intensité lumineuse correspondante dans l'image  $I$  (resp.  $T$ ).

### 3.1.1 Formalisation usuelle

La formalisation usuelle du problème de l'alignement géométrique repose sur l'hypothèse de l'invariance des niveaux de gris lors du déplacement à savoir qu'il est possible de trouver  $\mu^*$  tel



que :

$$\forall \mathbf{x} \in \mathbb{R}^2, \quad I(\mathbf{W}(\boldsymbol{\mu}^*, \mathbf{x})) = T(\mathbf{x}) \quad (3.1)$$

Avec cette formulation, l'image  $T$  n'est jamais déformée et l'image  $I$  est compensée par  $\boldsymbol{\mu}^*$  pour coïncider avec  $T$ . Le repère de coordonnées dans lequel les images sont alignées, que nous appelons dans la suite repère de compensation, est donc dans ce cas celui de l'image  $T$ .

### 3.1.2 Formalisation bidirectionnelle

Nous proposons maintenant une formalisation idéale plus générique impliquant la déformation des deux images. En supposant l'invariance des niveaux de gris, nous pouvons également écrire en posant  $\mathbf{z} = \mathbf{W}(\boldsymbol{\mu}_T^{-1}, \mathbf{x})$  :

$$\forall \mathbf{z} \in \mathbb{R}^2, \quad I(\mathbf{W}(\boldsymbol{\mu}_I, \mathbf{z})) = T(\mathbf{W}(\boldsymbol{\mu}_T, \mathbf{z})) \quad (3.2)$$

avec :

$$\boldsymbol{\mu}_I = \boldsymbol{\mu}^* \circ \boldsymbol{\mu}_T \quad (3.3)$$

Le problème d'alignement consiste désormais à trouver les paramètres bi-directionnels  $(\boldsymbol{\mu}_I, \boldsymbol{\mu}_T)$  tels que (3.2) soit vraie. L'équation (3.2) repose sur l'idée qu'appliquer une déformation à  $I$  pour l'aligner sur  $T$  revient au même qu'appliquer la déformation inverse à  $T$  pour l'aligner sur  $I$ . De manière encore plus générale, les paramètres  $\boldsymbol{\mu}_I$  et  $\boldsymbol{\mu}_T$  traduisent le fait qu'il existe une infinité de repères de coordonnées au delà des repères associés aux images  $I$  et  $T$  dans lesquels il est possible d'aligner les deux images.

## 3.2 Formalisation pratique

En pratique, l'alignement d'image est soumis à deux contraintes :

- Les images sont à support borné.
- Les images observées sont soumises à des perturbations photométriques. Ainsi même à l'issu d'une estimation parfaite de la transformation géométrique, (3.1) ou (3.2) ne sont pas exactes.

### 3.2.1 Définition d'une région d'intérêt comme support d'alignement

Les images réelles étant bornées spatialement, il n'est possible de les compenser qu'à partir des informations présentes dans leurs zones de recouvrement (à savoir les pixels présents dans les deux images et se correspondant). Une zone d'intérêt sur laquelle s'effectue l'alignement est donc classiquement définie conduisant à l'équation :

$$\forall \mathbf{x} \in R, \quad I(\mathbf{W}(\boldsymbol{\mu}^*, \mathbf{x})) = T(\mathbf{W}(\boldsymbol{\mu}_T^n, \mathbf{x})). \quad (3.4)$$

$R$  représente le support spatial de la région d'intérêt dans le repère de compensation. Les paramètres  $\boldsymbol{\mu}_T^n$  introduits ici permettent de distinguer le repère de compensation du repère du template. Il est souvent possible en pratique de faire coïncider ces deux repères, c'est pourquoi dans de nombreux articles [Hager 1998, Baker 2004, Benhimane 2007b, Keller 2004] ce paramètre

## 3.2. Formalisation pratique

---

n'est pas introduit. Nous verrons cependant que ce paramètre joue un rôle important pour formaliser le processus d'alignement d'images de manière générique. Dans les approches standards,  $\mu_T^n$  reste fixé tout au long du processus d'alignement. La portion d'image sur laquelle on souhaite s'aligner est donc fixée et  $I$  est potentiellement déformée par application de  $\mu$  pour ressembler à la portion extraite. Avec la formulation usuelle (3.4), les informations extraites de  $T$  et utilisées pour l'alignement sont figées.

En appliquant le changement de variable suivant :

$$\mathbf{z} = \mathbf{W}(\mu_T^{-1} \circ \mu_T^n, \mathbf{x}), \quad (3.5)$$

avec un  $\mu_T$  quelconque on peut retomber sur une formalisation bidirectionnelle :

$$\forall \mathbf{z} \in R, \quad I(\mathbf{W}(\mu_I, \mathbf{z})) = T(\mathbf{W}(\mu_T, \mathbf{z})) \quad (3.6)$$

avec :

$$\mu_I = \mu^* \circ (\mu_T^n)^{-1} \circ \mu_T \quad (3.7)$$

La zone d'intérêt dans  $T$  est désormais paramétrée par  $\mu_T$ . Ces derniers ne sont pas figés a priori. Nous revenons en détail sur cette propriété à la section 3.4.

**Définition d'une région d'intérêt en pratique.** Il existe plusieurs façons de délimiter la région d'intérêt  $R$  qui dépendent des applications visées. Une méthode usuellement employée est de délimiter la zone d'intérêt  $R$  par une boîte englobante dans le repère de compensation. La boîte englobante permet à l'aide de  $\mu_I$  et  $\mu_T$  de mettre en correspondance les informations d'intérêt dans  $I$  et  $T$ . Par exemple les auteurs de [Baker 2004] utilisent une boîte englobante rectangulaire définie dans un repère de compensation coïncidant avec le repère de l'image  $T$  (ils utilisent implicitement  $\mu_T^n = \mathbf{0}$  dans l'équation (3.4)). Cette approche permet de calculer une fois pour toute l'échantillonnage spatial de la boîte englobante. L'échantillonnage est souvent choisi pour correspondre à la grille de pixels entiers du template  $T$  afin de ne pas introduire d'aliasing lors de l'extraction de l'objet d'intérêt à partir de cette image. L'utilisation d'une boîte englobante permet de sélectionner un objet particulier et est donc particulièrement adaptée pour le suivi d'objets.

La figure 3.1 illustre les différents repères discutés jusqu'à présent. Pour une image  $\mathbf{L}$  quelconque,  $\text{warp}(\mathbf{L}; \mu_L)$  désigne l'image obtenue en compensant l'image  $\mathbf{L}$  vers le repère de compensation.  $\text{warp}(\mathbf{L}; \mu_L; R)$  désigne alors la portion d'intérêt dans l'image compensée à savoir l'image composée des intensités  $L(\mathbf{W}(\mu_L, \mathbf{z}))$  calculées pour l'ensemble de positions spatiales  $\mathbf{z}$  dans la région  $R$ . Dans cette illustration, la région d'intérêt est définie à l'aide d'une boîte englobante rectangulaire dans le repère de compensation.

### 3.2.2 Définition d'une mesure de similarité

En pratique en raison de dégradations multiples (bruit capteur, occultations,...) l'équation (3.4) n'est pas rigoureusement vraie et le problème d'alignement prend classiquement la forme de l'optimisation d'une fonctionnelle d'erreur. En général, la nature de la fonctionnelle va dépendre des hypothèses faites sur les dégradations des images. Plusieurs fonctionnelles ont été présentées

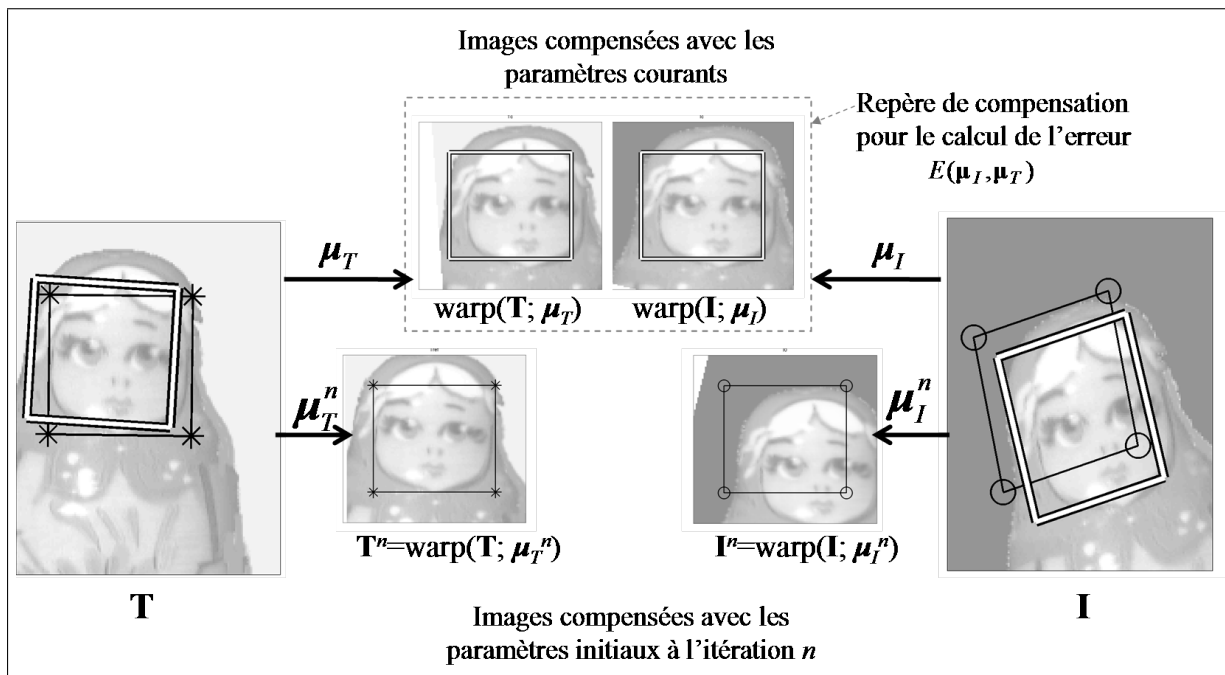


FIGURE 3.1 – Principe général du formalisme bidirectionnel, donnant une vision des différents repères de coordonnées considérés. La région d'intérêt  $R$  est affichée sur  $\text{warp}(T, \mu_T^n)$  et  $\text{warp}(I, \mu_I^n)$ , et sur  $\text{warp}(T, \mu_T)$  et  $\text{warp}(I, \mu_I)$ . Ces régions sont déformées sur  $I$  et  $T$ . Pour l'approche Directe,  $\mu_T$  et  $\mu_T^n$  sont égaux. Pour l'approche Inverse,  $\mu_I$  et  $\mu_I^n$  sont égaux. Dans le cas général présenté ici, le repère de compensation commun correspond à la compensation des deux images  $I$  et  $T$  respectivement avec les paramètres  $\mu_I$  et  $\mu_T$ .

### 3.2. Formalisation pratique

à la section 2.2.4 comme l'erreur SSD, le coefficient de corrélation normalisé, les M-estimateurs, l'information mutuelle ou encore des fonctionnelles reposant sur la comparaison d'histogrammes.

En combinant la définition du problème d'alignement sous une forme bidirectionnelle (3.6) et les fonctionnelles d'erreur usuelles, il est possible de construire de nouvelles fonctionnelles à optimiser. Afin de prendre en compte les différentes fonctionnelles évoquées précédemment, nous définissons la fonctionnelle d'erreur générique suivante :

$$E(\boldsymbol{\mu}_I, \boldsymbol{\mu}_T) = f(\text{att}(\text{warp}(I, \boldsymbol{\mu}_I, \mathcal{S})), \text{att}(\text{warp}(T, \boldsymbol{\mu}_T, \mathcal{S}))) \quad (3.8)$$

$\text{warp}(I, \boldsymbol{\mu}_I, \mathcal{S})$  et  $\text{warp}(T, \boldsymbol{\mu}_T, \mathcal{S})$  désignent les portions d'images extraites par application des paramètres  $\boldsymbol{\mu}_I$  et  $\boldsymbol{\mu}_T$  à l'échantillonnage  $\mathcal{S} = \{\mathbf{x}_i\}_{i=1..N}$  effectué dans le repère de compensation (à l'intérieur de la boîte englobante).  $N$  correspond au nombre d'échantillons. Dans le cas idéal (les variations entre les images sont uniquement dues à la déformation spatiale), ces portions sont identiques à convergence. Le sigle "att" représente la fonction qui calcule un certain nombre d'attributs visuels sur les images compensées. Ce sont ces attributs qui sont ensuite comparés. La fonction  $f$  à valeur réelle permet de mesurer la similarité entre les attributs des images  $I$  et  $T$ . Par exemple dans le cas de la fonctionnelle SSD, att vaut l'identité et la fonction  $f$  vérifie :

$$E(\boldsymbol{\mu}_I, \boldsymbol{\mu}_T) = f(\text{att}(\text{warp}(I, \boldsymbol{\mu}_I, \mathcal{S})), \text{att}(\text{warp}(T, \boldsymbol{\mu}_T, \mathcal{S}))) = \mathbf{e}(\boldsymbol{\mu}_I, \boldsymbol{\mu}_T)^T \mathbf{e}(\boldsymbol{\mu}_I, \boldsymbol{\mu}_T), \quad (3.9)$$

avec  $\mathbf{e}(\boldsymbol{\mu}_I, \boldsymbol{\mu}_T)$  définie par la concaténation des erreurs pixels à pixels  $e_i$  :

$$e_i(\boldsymbol{\mu}_I, \boldsymbol{\mu}_T) = I(\mathbf{W}(\boldsymbol{\mu}_I, \mathbf{x}_i)) - T(\mathbf{W}(\boldsymbol{\mu}_T, \mathbf{x}_i)) \quad (3.10)$$

L'objectif consiste désormais à trouver les paramètres bidirectionnels  $(\boldsymbol{\mu}_I, \boldsymbol{\mu}_T)$  qui permettent d'optimiser la fonctionnelle (3.8). Pour cela, nous adaptions le schéma itératif présenté à la section 2.2.1.1 en le reformulant de la façon suivante :

1. A l'itération  $n$ , le problème est reparamétrisé en fonction des estimés courants  $\boldsymbol{\mu}_I^n$  et  $\boldsymbol{\mu}_T^n \in \mathcal{G}$  et de l'incrément  $\mathbf{v} \in \mathcal{P}$  :

$$E^n(\mathbf{v}) = f(\text{att}(I(\mathbf{W}(\boldsymbol{\mu}_I(\boldsymbol{\mu}_I^n, \mathbf{v})), \mathcal{S})), \text{att}(T(\mathbf{W}(\boldsymbol{\mu}_T(\boldsymbol{\mu}_T^n, \mathbf{v})), \mathcal{S}))), \quad (3.11)$$

avec  $\boldsymbol{\mu}_I(\boldsymbol{\mu}_I^n, \mathbf{0}) = \boldsymbol{\mu}_I^n$  et  $\boldsymbol{\mu}_T(\boldsymbol{\mu}_T^n, \mathbf{0}) = \boldsymbol{\mu}_T^n$ . L'introduction d'un incrément supposé infinitésimal  $\mathbf{v}$  permet de simplifier la fonctionnelle d'erreur. C'est ce dernier que l'on estime à chaque itération.

2. La fonctionnelle d'erreur  $E^n$  est optimisée localement. Ceci se fait typiquement par la définition d'une fonctionnelle approximée  $\hat{E}^n$  dont on recherche un point stationnaire  $\hat{\mathbf{v}} = \arg \min(\hat{E}^n)$ .
3. Les paramètres sont mis à jour :

$$(\boldsymbol{\mu}_I^{n+1}, \boldsymbol{\mu}_T^{n+1}) \leftarrow \text{maj}(\boldsymbol{\mu}_I(\boldsymbol{\mu}_I^n, \hat{\mathbf{v}}), \boldsymbol{\mu}_T(\boldsymbol{\mu}_T^n, \hat{\mathbf{v}})) \quad (3.12)$$

La nature des fonctions de paramétrisation incrémentale  $\boldsymbol{\mu}_I(\boldsymbol{\mu}_I^n, \cdot)$  et  $\boldsymbol{\mu}_T(\boldsymbol{\mu}_T^n, \cdot)$  conditionne fortement les propriétés de convergence et de coût calculatoire des algorithmes d'alignement mis en place. Deux choix sont déterminants :

- Quelles restrictions souhaitons-nous imposer quant à l’exploration de l’espace bidirectionnel. Cela consiste essentiellement à choisir les dimensions de l’espace bidirectionnel selon lesquelles on s’autorise à se déplacer. Cette piste sera étudiée à la section 3.3.1.
- Quel type de paramétrisation est utilisé (additive, compositionnelle...) pour lier l’incrément  $\mathbf{v}$  et les estimés initiaux  $\mu_I^n$  et  $\mu_T^n$  aux paramètres courants  $\mu_I$  et  $\mu_T$ . Cette piste sera évoquée à la section 3.3.2.

De la même manière, la mise à jour des paramètres par la fonction maj dépend de l’application ciblée, mais devra rester cohérente avec les deux choix précédents. Ceci sera discutée section 3.4.

### 3.3 Contrainte sur l’espace bidirectionnel

#### 3.3.1 Taxonomie

##### 3.3.1.1 Type de compensation

Il existe trois façons génériques de contraindre l’espace bidirectionnel pour lesquelles des algorithmes ont déjà été développés. Nous proposons donc ici une nouvelle taxonomie concernant les algorithmes d’alignement d’images. Cette taxonomie étend la classification de [Baker 2004] et se présente sous la forme de trois classes :

**Compensation Directe (F) (*Forwards*) :** L’image courante est déformée par un incrément  $\mathbf{v} \in \mathcal{P}$  pour s’aligner avec le template qui reste fixe.

$$\mu_I = \mu_I(\mu_I^n, \mathbf{v}) \quad \mu_T = \mu_T^0. \quad (3.13)$$

**Compensation Inverse (I) (*Inverse*).** Le template courant est déformé par un incrément  $\mathbf{v} \in \mathcal{P}$  pour s’aligner avec l’image  $I^n = \text{warp}(I, \mu_I^n)$  qui reste fixe.

$$\mu_I = \mu_I^n \quad \mu_T = \mu_T(\mu_T^n, \mathbf{v}). \quad (3.14)$$

Les classes directe et inverse correspondent aux classes du même nom introduite par [Baker 2004].

**Compensation Bidirectionnelle (B).** L’image et le template sont alignés l’un vers l’autre (les deux images sont déformées). L’approche bidirectionnelle peut être encore décomposée en fonction des contraintes que l’on impose aux corrections  $\mathbf{v} \in \mathcal{P}$ . On peut mettre en avant notamment les sous-familles suivantes :

- La sous-classe **Bidirectionnelle Pure** considère que deux corrections complètement indépendantes sont apportées à l’image et au template. Dans ce cas, l’incrément appartient à un espace de dimension double par rapport aux approches directes et inverses. Il peut être séparé en deux parties distinctes  $\mathbf{v} = [\mathbf{v}_I, \mathbf{v}_T] \in \mathcal{P}^2$  :

$$\mu_I = \mu_I(\mu_I^n, \mathbf{v}_I) \quad \mu_T = \mu_T(\mu_T^n, \mathbf{v}_T). \quad (3.15)$$

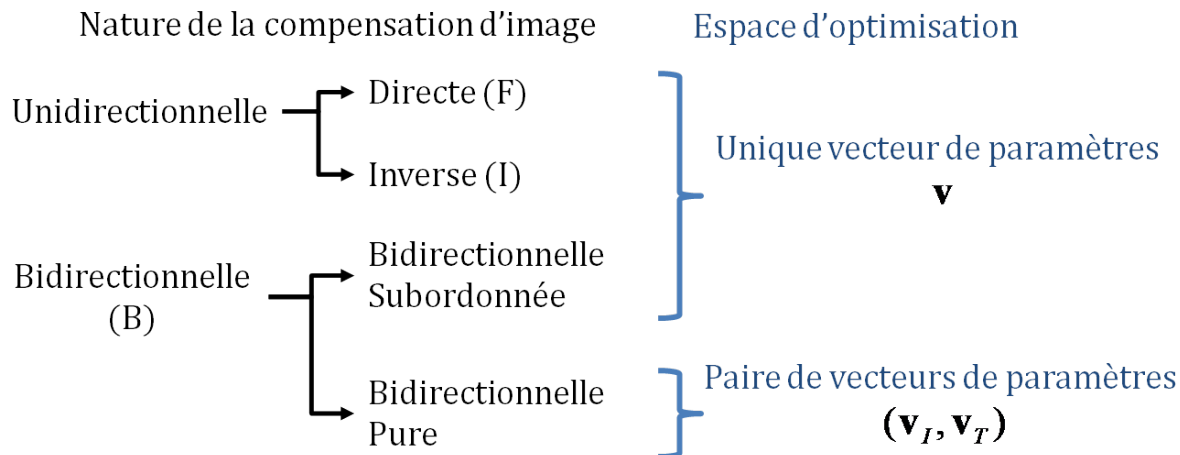


FIGURE 3.2 – Taxonomie des approches Directe, Inverse et bidirectionnelle mettant en évidence deux façons de classifier les méthodes.

- La sous-classe **Bidirectionnelle Subordonnée** suppose que les deux images sont déformées l'une vers l'autre de façon subordonnée à un unique incrément  $\mathbf{v} \in \mathcal{P}$  :

$$\boldsymbol{\mu}_I = \boldsymbol{\mu}_I(\boldsymbol{\mu}_I^n, \mathbf{v}) \quad \boldsymbol{\mu}_T = \boldsymbol{\mu}_T(\boldsymbol{\mu}_T^n, \mathbf{v}). \quad (3.16)$$

C'est une approche bidirectionnelle puisque les deux images sont déformées. Le terme subordonné met en évidence que  $\boldsymbol{\mu}_I$  et  $\boldsymbol{\mu}_T$  sont dépendants du même incrément  $\mathbf{v} \in \mathcal{P}$ . Dans cette catégorie seront notamment rangées les approches symétriques abordées à la section 4.3.1 ainsi que les approches asymétriques détaillées au chapitre 5.

La classification proposée ici tient compte des images affectées par l'incrément de compensation de mouvement  $\mathbf{v}$ . Ainsi, les approches directe et inverse considèrent une compensation unidirectionnelle (une image est fixe, l'autre est déformée). Les approches bidirectionnelles impliquent quant à elles une compensation bidirectionnelle (l'image et le template sont déformés).

#### 3.3.1.2 Espace d'optimisation locale

Une seconde classification possible repose sur la nature de l'incrément  $\mathbf{v}$  utilisé pour l'optimisation locale. L'incrément est en effet homogène à un simple vecteur de  $\mathcal{P}$  dans les approches Directe, Inverse et Bidirectionnelle Subordonnée alors qu'il est homogène à une paire de vecteur  $[\mathbf{v}_I, \mathbf{v}_T] \in \mathcal{P}^2$  pour l'approche Bidirectionnelle Pure.

Ces deux classifications complémentaires sont résumées figure 3.3.1.2. L'utilisation d'une paire de paramètres et donc potentiellement de deux fois plus de degrés de liberté que nécessaire pour aligner les images fait de l'approche Bidirectionnelle Pure un cas à part des différentes approches.

#### 3.3.2 Nature de la paramétrisation incrémentale

Comme introduit à la section 2.2.1.2, plusieurs solutions peuvent être envisagées pour combiner un incrément avec une estimation courante  $\boldsymbol{\mu}^n$  des paramètres. Rappelons que nous considé-

rons que les paramètres  $\mu$  appartiennent à un groupe  $\mathcal{G}$ . Trois approches apparaissent couramment dans la littérature.

**L'approche additive** utilise un loi du type :

$$\mu^A(\mu^n, \mathbf{v}) = \mu^n + \mathbf{v}, \quad \mathbf{v} \in \mathbb{R}^p. \quad (3.17)$$

**L'approche compositionnelle** s'écrit :

$$\mu^C(\mu^n, \mathbf{v}) = \mu^n \circ \mathbf{v}, \quad \mathbf{v} \in \mathcal{G} \quad (3.18)$$

Cette approche a été introduite par [Shum 2000].

**L'approche compositionnelle sur groupe de Lie** utilise :

$$\mu^{CL}(\mu^n, \mathbf{v}) = \mu^n \circ \exp(\mathbf{v}), \quad \mathbf{v} \in \mathfrak{G}, \quad (3.19)$$

où  $\mathfrak{G}$  est l'algèbre de Lie associée au groupe  $\mathcal{G}$ .

Remarquons tout d'abord que pour un espace de paramètres formant un groupe additif (translation par exemple), ces trois lois sont identiques. De plus, la paramétrisation (3.17) est utilisée dans de nombreux algorithmes [Hager 1998, Hager 2004, Buenaposada 2002] même pour des groupes non-additifs, en considérant l'ensemble des paramètres  $\mathcal{G}$  comme un sous-espace d'un espace vectoriel. Comme cela a été évoqué dans [Vercauteren 2009], il paraît plus naturel d'utiliser une paramétrisation incrémentale compositionnelle pour des transformations géométriques en raison de la signification géométrique que l'on peut donner à la composition de transformations. Les approches sur groupe de Lie se distinguent des approches plus conventionnelles en introduisant un incrément appartenant à un espace différent du groupe des paramètres.

Dans la suite nous appelons orbite d'une image  $I$ , sous l'opération du groupe de transformations géométriques  $\mathcal{G}$ , le nouvel ensemble obtenu en appliquant les éléments de  $\mathcal{G}$  sur  $I$  en utilisant l'action de groupe  $\mathbf{W}$ . Ce concept a été notamment introduit sous le nom de "video orbit" [Mann 1997]. L'orbite de  $I$  correspond donc à l'ensemble des versions déformées de  $I$ . On note cette dernière orbite( $I$ ). A l'intérieure de l'orbite on peut distinguer trois images particulières :

- $I = \text{warp}(I, \mathbf{0})$  correspond à l'image initiale.
- $I^n = \text{warp}(I, \mu_I^n)$  correspond à l'image  $I$  compensée par les paramètres initiaux  $\mu_I^n$ . Pour un pixel  $\mathbf{z} \in R$  nous avons  $I^n(\mathbf{z}) = I(\mathbf{W}(\mu_I^n, \mathbf{z}))$
- $\hat{I} = \text{warp}(I, \mu_I(\mu_I^n, \hat{\mathbf{v}}))$  correspond à l'image  $I$  compensée par les paramètres  $\mu_I(\mu_I^n, \hat{\mathbf{v}})$ . Pour un pixel  $\mathbf{z} \in R$  nous avons  $\hat{I}(\mathbf{z}) = I(\mathbf{W}(\mu_I(\mu_I^n, \hat{\mathbf{v}}), \mathbf{z}))$

La figure 3.3 (en haut) illustre le déplacement opéré sur l'orbite de  $I$  lorsque l'on utilise une paramétrisation additive. Sur ce schéma une ligne horizontale noire représente l'orbite de l'image. Un point représente une version de l'image sur l'orbite. Une flèche sur laquelle figure les paramètres  $\mu$  signifie un déplacement sur l'orbite transformant une version de l'image en une autre. La version de l'image à l'extrémité de la flèche est obtenue en compensant la version précédente par  $\mu$ . Nous constatons qu'il n'existe pas de lien géométrique simple entre  $I^n$  et  $\hat{I}$  pour la paramérisation additive. En effet en notant  $\mathbf{z}$  la position d'un pixel dans la région

### 3.3. Contrainte sur l'espace bidirectionnel

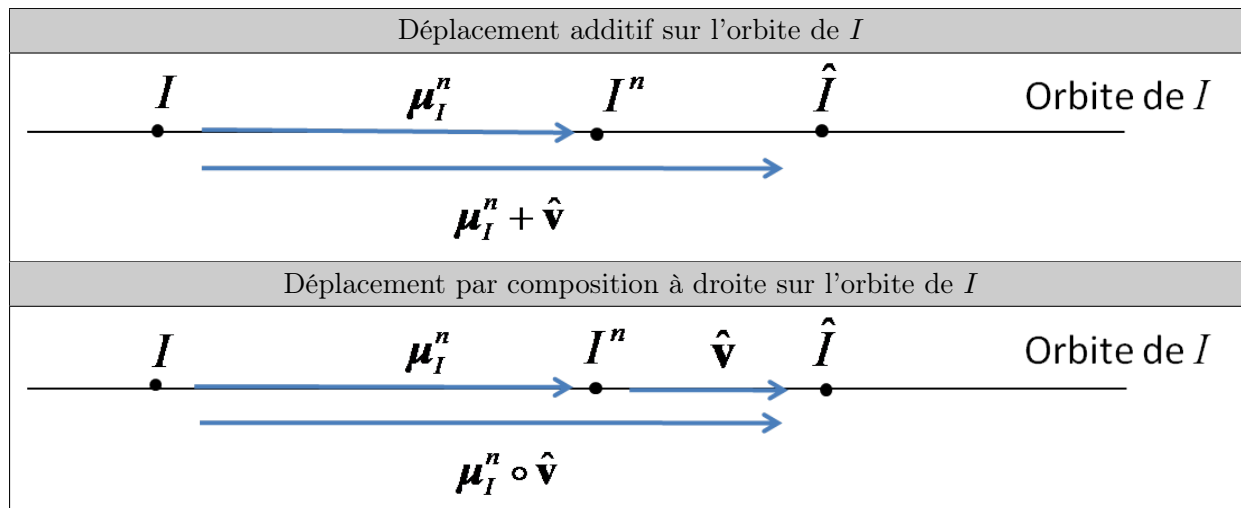


FIGURE 3.3 – Interprétation géométrique des différentes paramétrisations incrémentales. Seule la composition permet d'exprimer simplement un lien entre l'image obtenue par compensation de  $I$  avec les paramètres finaux  $\mu_I(\mu_I^n, \hat{\mathbf{v}})$  à savoir  $\hat{I} = \text{warp}(I, \mu_I(\mu_I^n, \hat{\mathbf{v}}))$  et l'image obtenue par compensation de  $I$  avec les paramètres initiaux  $I^n = \text{warp}(I, \mu_I^n)$ .

	paramétrisation incrémentale	Espace d'optimisation $\mathcal{P}$
Additive	$\mu_L = \mu_L^n + \mathbf{v}$	$\mathbf{v} \in \mathcal{G} = \mathbb{R}^p$
Compositionnelle	$\mu_L = \mu_L^n \circ \mathbf{v}$	$\mathbf{v} \in \mathcal{G}$
Compositionnelle Lie	$\mu_L = \mu_L^n \circ \exp(\mathbf{v})$	$\mathbf{v} \in \mathcal{G}$

TABLE 3.1 – Propriétés principales des paramétrisations incrémentales.

d'intérêt dans le repère de  $\hat{I}$  et  $\mathbf{y}$  la position correspondante dans l'image  $I^n$ , nous avons  $\mathbf{y} = \mathbf{W}((\mu_I^n)^{-1} \circ (\mu_I^n + \hat{\mathbf{v}}), \mathbf{z})$  et donc :

$$\begin{aligned} \hat{I}(\mathbf{z}) &= I^n(\mathbf{W}((\mu_I^n)^{-1} \circ (\mu_I^n + \hat{\mathbf{v}}), \mathbf{z})) \\ &= I(\mu_I^n + \hat{\mathbf{v}}, \mathbf{z}). \end{aligned} \tag{3.20}$$

Le lien entre  $\hat{I}$  et  $I^n$  fait donc intervenir  $\mu_I^n$ .

La figure 3.3 (milieu) illustre le déplacement opéré sur l'orbite de  $I$  lorsque l'on utilise une paramétrisation compositionnelle. Le lien entre  $\hat{I}$  et  $I^n$  peut être exprimé uniquement à l'aide de l'incrément  $\hat{\mathbf{v}}$  sans utilisation de  $\mu_I^n$  :  $\hat{I}(\mathbf{z}) = I^n(\mathbf{W}(\hat{\mathbf{v}}, \mathbf{z})) = I(\mathbf{W}(\mu_I^n \circ \hat{\mathbf{v}}, \mathbf{z}))$ .

Le tableau 3.1 résume les propriétés des différentes paramétrisations incrémentales évoquées. Nous développerons plus en détail à la section 4.4 les liens entre ces paramétrisations dans le cadre des fonctionnelles d'erreur SSD.



### 3.4 Mise à jour des paramètres

Supposons que les deux images  $I$  et  $T$  soient identiques à la compensation géométrique près en utilisant les paramètres d'alignement  $(\boldsymbol{\mu}_I, \boldsymbol{\mu}_T)$  :

$$\text{warp}(I, \boldsymbol{\mu}_I) = \text{warp}(T, \boldsymbol{\mu}_T) \quad (3.21)$$

Dans ce cas, nous avons également  $\forall \Delta\boldsymbol{\mu} \in \mathcal{P}$  :

$$\text{warp}(I, \boldsymbol{\mu}_I \circ \Delta\boldsymbol{\mu}) = \text{warp}(T, \boldsymbol{\mu}_T \circ \Delta\boldsymbol{\mu}) \quad (3.22)$$

Ainsi, une infinité de paires de paramètres, que nous notons  $\text{equiv}(\boldsymbol{\mu}_I, \boldsymbol{\mu}_T)$ , obtenue par composition à droite avec le même jeu de paramètres  $\Delta\boldsymbol{\mu}$  devrait également être considérée comme des paramètres vrais :

$$\text{equiv}(\boldsymbol{\mu}_I, \boldsymbol{\mu}_T) = \{(\boldsymbol{\mu}_I \circ \Delta\boldsymbol{\mu}, \boldsymbol{\mu}_T \circ \Delta\boldsymbol{\mu}) \mid \Delta\boldsymbol{\mu} \in \mathcal{P}\}. \quad (3.23)$$

L'équation de mise à jour (3.12) doit donc sélectionner  $\boldsymbol{\mu}_I^{n+1}$  et  $\boldsymbol{\mu}_T^{n+1}$  parmi l'ensemble défini par (3.23). Plusieurs règles de mise à jour ont été utilisées dans la littérature. Nous définissons dans la suite deux règles de mise à jour génériques recouvrant ces dernières.

La règle de mise à jour usuelle associée à l'optimisation itérative d'une erreur  $\mathbf{e}(\boldsymbol{\mu}_I, \boldsymbol{\mu}_T)$  sur l'espace des paramètres bidirectionnel  $\mathcal{P}^2$  est la règle de mise à jour identité.

**Mise à jour Identité.** Cette règle de mise à jour, notée  $\text{maj}^{Id}(\hat{\boldsymbol{\mu}}_I, \hat{\boldsymbol{\mu}}_T)$  conserve les paramètres qui sont passés en argument (qui ont été estimés par une approche itérative quelconque) de la fonction de mise à jour :

$$(\boldsymbol{\mu}_I^{n+1}, \boldsymbol{\mu}_T^{n+1}) = \text{maj}^{Id}(\hat{\boldsymbol{\mu}}_I, \hat{\boldsymbol{\mu}}_T) = (\hat{\boldsymbol{\mu}}_I, \hat{\boldsymbol{\mu}}_T). \quad (3.24)$$

Les paramètres initiaux  $(\boldsymbol{\mu}_I^n, \boldsymbol{\mu}_T^n)$  évoluent à chaque itération conduisant à la déformation des deux images.

**Mise à jour Directe.** Lorsque le temps de calcul entre en jeu, la mise à jour à la fois de  $\boldsymbol{\mu}_I$  et  $\boldsymbol{\mu}_T$  induit de déformer chaque image et de calculer pour chacune les informations utilisées pour réaliser l'alignement à chaque nouvelle itération, ce qui conduit à des algorithmes de plus forte complexité calculatoire. Afin de prendre en compte ce problème, nous introduisons une mise à jour que nous appellons directe :

$$(\boldsymbol{\mu}_I^{n+1}, \boldsymbol{\mu}_T^{n+1}) = \text{maj}^D(\hat{\boldsymbol{\mu}}_I, \hat{\boldsymbol{\mu}}_T) = (\hat{\boldsymbol{\mu}}_I \circ \hat{\boldsymbol{\mu}}_T^{-1} \circ \boldsymbol{\mu}_I^0, \boldsymbol{\mu}_T^0). \quad (3.25)$$

Ainsi les paramètres  $\boldsymbol{\mu}_T$  sont réinitialisés avec leur valeur initiale  $\boldsymbol{\mu}_T^0$  à chaque itération. Cette approche permet de précalculer un certain nombre d'éléments sur l'image  $T$  pour l'ensemble des itérations, ce qui améliore grandement les temps de calcul. L'approche Inverse Compositionnelle proposée par [Baker 2004] bénéficie de cet avantage en offrant la possibilité de précalculer complètement l'inverse de la matrice Jacobienne de l'erreur. De plus l'algorithme d'alignement est alors contraint d'utiliser dans l'image  $T$  uniquement les informations portées par les alentours de la région localisée par  $\boldsymbol{\mu}_T^0$ .

Le comportement au cours du processus itératif introduit par les règles de mise à jour Directe et Identité dépend de la famille d'algorithmes utilisée :

### 3.4. Mise à jour des paramètres

---

- Pour les approches directes (F), les règles de mise à jour identité et directe sont identiques (car  $\mu_T = \mu_T^n = \mu_T^0$ ).
- Les approches de la famille inverses (I) utilisent toujours dans les approches existantes une règle de mise à jour directe. En effet, l'utilisation d'une règle de mise à jour identité avec une approche ou seule l'image  $T$  est considérée comme mobile (ce qui est la particularité des approches inverses) correspond en fait à une approche directe dans laquelle le rôle de  $I$  et  $T$  est inversé.
- Pour les approches bidirectionnelles, différents comportements sont introduits par les différentes règles de mise à jour. La figure 3.4 illustre les comportements des deux lois de mise à jour dans le cas où l'objet d'intérêt est délimité à l'aide d'une boîte englobante rectangulaire. On constate que la règle de mise à jour identité conduit à s'éloigner de la région initialement sélectionnée dans  $T$  pour l'alignement. L'algorithme est en effet libre de sélectionner la région qui lui paraît la plus opportune pour minimiser la fonctionnelle d'erreur. En revanche, avec la mise à jour directe, la portion de  $T$  sur laquelle s'effectue l'alignement est figée.

#### 3.4.1 Règle de mise à jour et applications

Le choix d'une règle de mise à jour dépend essentiellement de l'application ciblée. Dans le cas de l'étude de mouvements rigides, on peut distinguer deux situations :

- la scène complète observée est statique et les deux images considérées sont reliées dans leur ensemble par une transformation rigide.
- Des objets sont en mouvement devant une scène statique. Ce sont les paramètres de mouvement de ces objets qui sont recherchés.

**Scène statique** Pour ce type d'application, les deux règles de mises à jour semblent possibles. La particularité de la règle de mise à jour identité est de laisser l'algorithme libre de faire évoluer la région d'intérêt, c'est à dire de sélectionner le contenu sur lequel s'aligner. Dans le cas idéal (images continues identiques), il y a a priori une infinité de régions se correspondant à l'intérieur de la zone de recouvrement, le problème semble donc mal posé. En pratique, les images étant imparfaites (bruit...), l'algorithme va choisir la région qui minimise l'erreur.

L'avantage de la mise à jour directe est un gain en temps de calcul en limitant le nombre de déformations des images. Cependant, il est nécessaire pour un bon alignement que  $\mu_T^0$  désigne une région dans le template  $T$  qui soit entièrement incluse dans la zone de recouvrement réelle des deux images. Dans le cas contraire, il est impossible de mettre en correspondance tous les pixels de la région d'intérêt sélectionnée dans le template avec ceux de l'image, ce qui peut nuire à l'alignement.

**Suivi d'objets** Dans le contexte de suivi d'objets, les deux images ne sont pas conceptuellement interchangeables puisque l'une des deux images est utilisée pour délimiter l'objet d'intérêt que l'on cherche à suivre dans la vidéo. Re-initialiser  $\mu_T^n$  à  $\mu_T^0$  permet de garder la région d'intérêt  $R$  proche de la position de l'objet d'intérêt dans le Template, et donc d'éviter de prendre en compte

des informations non pertinentes dans  $I$  quand l'optimisation s'approche de l'optimum. Ce type d'application penche donc clairement en faveur de l'utilisation d'une règle de mise à jour directe.

Cela est illustré Figure 3.5 (à gauche). Dans cet exemple les deux images utilisées sont celles de la figure 3.1. La zone d'intérêt est déterminée par un rectangle de dimensions fixes dans le repère de compensation. L'objet d'intérêt, à savoir le visage de la poupée, est sélectionné initialement par l'intermédiaire des paramètres  $\mu_T^0$  dans l'image  $T$ . Supposons que nous disposions des paramètres vrais  $\mu_I^*$  permettant d'aligner  $I$  sur la portion extraite dans  $T$  par  $\mu_T^0$ . Si les deux images étaient identiques à la compensation géométrique près, l'ensemble  $\text{equiv}(\mu_I^*, \mu_T^0)$  correspondrait à autant de paires de paramètres valides pour aligner les images  $I$  et  $T$ . Comme les deux images ont un arrière plan différent (gris dans  $I$  et blanc dans  $T$ ), la vallée formée sur la figure 3.5 autour de la variété des paramètres vrais  $\text{equiv}(\mu_I^*, \mu_T^0)$  possède une faible erreur  $E(\mu_I, \mu_T)$  dans l'espace bidirectionnel  $\mathcal{P}^2$  seulement pour  $\mu_T$  suffisamment proche de  $\mu_T^0$ . Supposons maintenant qu'une dégradation photométrique déplace le minimum de la fonctionnelle d'erreur vers l'arrière plan (bruit, reflets spéculaires...). Avec l'utilisation d'une règle de mise à jour identité, la zone d'intérêt extraite du template se déplacerait vers l'arrière plan pour minimiser au mieux l'erreur  $E(\mu_I, \mu_T)$ , conduisant à l'estimation de paramètres ne correspondant pas nécessairement au paramètres de mouvement du visage de la poupée.

Itération	Mise à jour Identité			Mise à jour Directe		
	$I^n$	$T^n$	$I^n - T^n$	$I^n$	$T^n$	$I^n - T^n$
1						
10						
50						
100						

FIGURE 3.4 – Exemple d’alignement pour une approche bidirectionnelle pure avec mise à jour identité (à gauche) et directe (à droite). La zone d’intérêt est définie à l’aide d’une boîte englobante rectangulaire de dimensions fixes dans le repère de compensation. Alors qu’avec la règle de mise à jour directe la zone d’intérêt extraite depuis  $T$  reste inchangée, avec la règle de mise à jour identité cette dernière évolue au cours des itérations.

### 3.4.2 Convergence

Concernant les propriétés de convergence, la règle de mise à jour Identité correspond en fait à la règle de mise à jour traditionnelle des algorithmes itératifs. Les propriétés de convergence de l’algorithme d’optimisation appliqué à chaque itération sont ainsi conservées.

La règle de mise à jour directe en revanche correspond à la règle de mise à jour traditionnelle uniquement pour les approches de la famille Directe (F). Pour les approches des familles Inverses ou Bidirectionnelles la valeur de la fonctionnelle d’erreur peut changer durant l’étape de mise à jour, comme illustré Figure 3.5 par un saut le long de l’orbite des paramètres estimés à la fin de chaque itération. Sur cette figure, les approches dont les trajectoires sont tracées appartiennent aux classes respectives Directe (F), Inverse (I), Symétrique (sous-classe de la famille Bidirectionnelle Subordonnée dont le détail est abordé à la section 4.3.1) et Bidirectionnelle (B). Alors que l’approche Directe a une trajectoire linéaire, les autres approches, en raison de la règle de mise à jour directe utilisée, effectuent un saut de l’erreur  $E(\boldsymbol{\mu}_I(\boldsymbol{\mu}_I^n, \hat{\mathbf{v}}), \boldsymbol{\mu}_T(\boldsymbol{\mu}_T^n, \hat{\mathbf{v}}))$  à l’erreur  $E(\boldsymbol{\mu}_I^{n+1}, \boldsymbol{\mu}_T^{n+1})$  lors de la mise à jour des paramètres.

Pour assurer en pratique que cela ne nuit pas à la minimisation de l’erreur, un critère d’arrêt basé sur la décroissance lors du passage de  $E(\boldsymbol{\mu}_I^n, \boldsymbol{\mu}_T^n)$  à  $E(\boldsymbol{\mu}_I^{n+1}, \boldsymbol{\mu}_T^{n+1})$  est utilisé.

Rien ne garantit que le critère d’arrêt ne stoppe trop rapidement le processus d’optimisation en raison d’une différence entre l’erreur  $E(\boldsymbol{\mu}_I(\boldsymbol{\mu}_I^n), \hat{\mathbf{v}}), \boldsymbol{\mu}_T(\boldsymbol{\mu}_T^n, \hat{\mathbf{v}}))$  obtenue après estimation de l’incrément  $\mathbf{v}$  à l’étape  $n$  et sa version mise à jour  $E(\boldsymbol{\mu}_I^{n+1}, \boldsymbol{\mu}_T^{n+1})$ . De manière théorique, ces deux valeurs peuvent être rendues identiques, si nous négligeons la dérive de la région d’intérêt, en utilisant une fonctionnelle d’erreur, tenant compte des changements de volume (cf. section 2.3.1) comme cela a été fait pour les fonctionnelles SSD dans [Cachier 2003] [Tagare 2009]. Cette approche génère des termes additionnels qui compensent le biais associé au changement d’échelle éventuel du passage d’un repère de coordonnées à l’autre. Cette invariance théorique introduit un coût calculatoire supplémentaire, et ne sera pas étudiée dans cette thèse. Le lecteur intéressé est invité à lire [Tagare 2009] pour une discussion détaillée des problèmes associés à la prise en compte des changements de volume.

Pour les applications de suivi d’objet qui sont développées dans cette thèse, le problème du biais ne semble pas être un problème majeur. En effet, l’application d’une règle de mise à jour Directe combinée avec l’erreur (3.10) a été mise en place avec succès avec l’approche Inverse Compositionnelle [Baker 2004] pour laquelle Baker et Matthews ont proposé une preuve d’équivalence à l’ordre 1 avec les approches Directes Compositionnelle et Additive.

## 3.5 Etapes de conception d’un algorithme d’alignement d’images

Nous revenons dans cette partie sur les différents concepts indépendants identifiés jusqu’à présent et ayant une importance pour la conception d’un algorithme itératif d’alignement d’images.

**Choix 1 : Fixer le modèle de transformation** c’est définir la nature du mouvement à estimer. Ce choix est essentiellement dépendant de l’application visée. Les mouvements affines et homographiques étant à la fois très couramment utilisés, et pouvant être mis en oeuvre de

### 3.5. Etapes de conception d'un algorithme d'alignement d'images

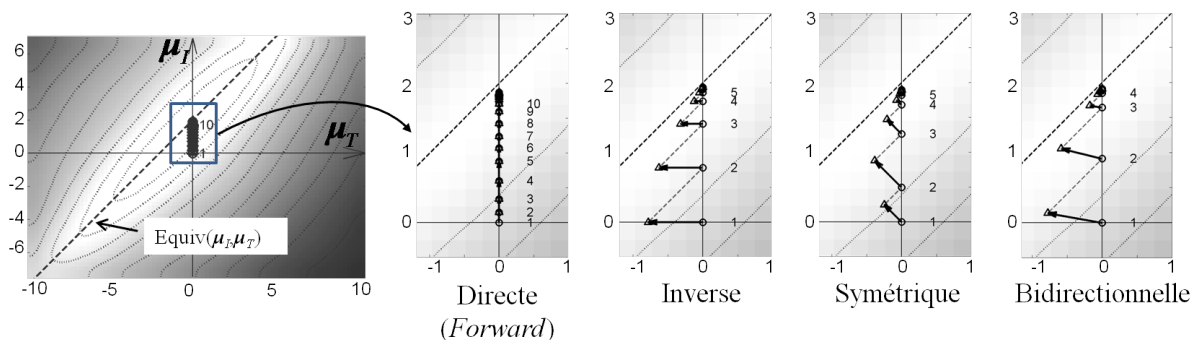


FIGURE 3.5 – Illustration de la fonction erreur  $E(\mu_I, \mu_T)$  et trajectoires des approches FCL, ICL, SCL et BCL durant l'alignement des deux images de la figure 3.1, pour une déformation réelle de 2 pixels en translation horizontale  $\mu_{I,1}^* = 2$ . Seuls les coefficients de translation horizontale  $(\mu_{I,1}, \mu_{T,1})$  sont affichés, ce qui correspond à une coupe de l'espace bidirectionnel  $\mathcal{P}^2$ . Pour l'erreur  $E$ , des niveaux de gris plus sombres indiquent une erreur plus importante. La translation a été estimée à l'aide de l'algorithme d'optimisation Gauss-Newton employant une règle de mise à jour directe. Chaque itération  $n$  est tracée à l'aide d'une flèche qui relie  $(\mu_I^n, \mu_T^n)$  (marqueurs numérotés  $\odot$ ) à  $(\mu_I(\mu_I^n, \hat{\mathbf{v}}), \mu_T(\mu_T^n, \hat{\mathbf{v}}))$  ( $\triangle$ ), et d'un segment en pointillés vers l'itération suivante  $(\mu_I^{n+1}, \mu_T^{n+1})$  utilisant une règle de mise à jour directe. Les approches sont initialisées avec  $(\mu_{I,1}^0, \mu_{T,1}^0) = (0, 0)$ . La droite en pointillés  $(\mu_{I,1}, \mu_{T,1}) = (\mu_1, 2 + \mu_1)$  représente la variété des estimés corrects.

nombreuses façons différentes, ils seront choisis pour illustrer les différents résultats de la partie expérimentale 6.

**Choix 2 : Choisir la paramétrisation** c'est définir l'espace  $\mathcal{G}$  dans lequel résident les paramètres du mouvement  $\mu_I$  et  $\mu_T$ . Une fois la paramétrisation choisie, il est possible de définir clairement l'expression de  $\mathbf{W}(\mu, \mathbf{x})$ . Différents modèles pour les transformations rigides sont présentés à la section 2.1.2. Remarquons que pour pouvoir appliquer le formalisme bidirectionnel présenté dans ce chapitre avec des modèles déformables, notamment si l'on veut utiliser une règle de mise à jour directe avec une approche inverse ou bidirectionnelle, les propriétés de composition et d'inversion compatibles avec celles d'une action de groupe (2.13)(2.14)(2.15) sont nécessaires afin de pouvoir utiliser (3.25). Une heuristique pour estimer ces opérations de façon efficace dans le cas de mouvement déformable est présentée dans [Gay-Bellile 2007].

**Choix 3 : Définir la fonctionnelle d'erreur**  $E(\mu_I, \mu_T)$  dépend grandement de l'application et des performances visées : une simple norme L2 dans [Baker 2004], une métrique robuste dans [Baker 2003a]. Il est à noter que la métrique d'erreur n'est pas nécessairement une erreur pixel à pixel, mais peut prendre une forme plus complexe. Par exemple, dans [Mégret 2006], l'erreur est fondée sur la dissimilarité entre distributions de couleurs locales après compensation de mouvement.

Il apparaît ainsi qu'une grande liberté est laissée au concepteur dans le choix de la métrique

Choix 1	Fixer le modèle de transformation
Choix 2	Choisir la paramétrisation
Choix 3	Définir la fonction d’erreur $E(\boldsymbol{\mu}_I, \boldsymbol{\mu}_T)$
Choix 4	Fixer des contraintes sur l’espace bidirectionnel
Choix 5	Choisir une règle de mise à jour
Choix 6	Décider de la méthode d’optimisation

TABLE 3.2 – Étapes de conception d’un algorithme d’alignement d’images.

à cette étape pour la conception d’algorithmes plus robustes ou plus performants.

**Choix 4 : Fixer les contraintes sur l’espace bidirectionnel,** c’est choisir l’expression des fonctions  $\boldsymbol{\mu}_I = \boldsymbol{\mu}_I(\boldsymbol{\mu}_I^n, \mathbf{v})$  et  $\boldsymbol{\mu}_T = \boldsymbol{\mu}_T(\boldsymbol{\mu}_T^n, \mathbf{v})$ . C’est donc essentiellement :

- Définir des contraintes sur l’exploration de l’espace bidirectionnel en choisissant une classe parmi (F), (I), et (B). Nous mettons en évidence dans cette thèse à la section 4.4.3 un certain nombre de caractéristiques de ces familles d’approches dans le contexte d’une optimisation Gauss-Newton d’une fonctionnelle SSD.
- Définir la nature de la paramétrisation incrémentale parmi les différentes façons développées à la section 3.1. Nous verrons à la section 4.4.3 plus en détail les liens entre ces différentes paramétrisations incrémentales dans le contexte d’une optimisation Gauss-Newton d’une fonctionnelle SSD.

A l’issue de cette étape, la fonctionnelle d’erreur dépend uniquement du paramètre vectoriel  $\mathbf{v}$ . C’est ce dernier qu’il faut estimer.

**Choix 5 : Définir une règle de mise à jour,** La règle de mise à jour dépend également de l’application. Nous avons vu à la section 3.4 qu’une règle de mise à jour directe (3.25) permettait de diminuer la complexité calculatoire tout en conservant la région d’intérêt proche de la région initialement sélectionnée à travers  $\boldsymbol{\mu}_T^0$ . Dans le cas de l’alignement d’images d’une scène fixe comme pour la création de panorama, une règle de mise à jour identité (3.24) paraît également envisageable avec alors l’intérêt d’autoriser l’algorithme à sélectionner lui même (en partant de l’objet d’intérêt initialement déterminée par  $\boldsymbol{\mu}_T^0$ ) la portion d’image sur laquelle s’aligner. Nous nous intéressons dans cette thèse plutôt aux applications de suivi visuel et donc aux approches utilisant une règle de mise à jour directe.

**Choix 6 : La méthode d’optimisation** est la méthode utilisée pour minimiser  $E(\mathbf{v})$ , la fonctionnelle d’erreur caractérisant le processus d’alignement. Ceci peut être réalisé en utilisant des approches basées gradients, comme les méthodes Gauss-Newton (GN) et Newton (N) [Baker 2004] ou encore des méthodes d’ordre 3 [Keller 2008]. Dans la section suivante, nous donnons un exemple d’optimisation utilisant l’algorithme GN. Soulignons ici le fait que la méthode d’optimisation à utiliser est indépendante a priori des contraintes fixées sur l’espace bidirectionnel comme la paramétrisation incrémentale.

# Application du formalisme Bidirectionnel aux approches existantes

---

## Sommaire

---

<b>4.1</b>	<b>Application de l'algorithme Gauss-Newton aux approches SSD . . . . .</b>	<b>52</b>
<b>4.2</b>	<b>Classification des approches de l'état de l'art : méthodes unidirectionnelles . . . . .</b>	<b>53</b>
4.2.1	Les approches Directes (F) (Forward) . . . . .	53
4.2.2	Les approches Inverses (I) . . . . .	54
<b>4.3</b>	<b>Classification des approches de l'état de l'art : méthodes bidirectionnelles</b>	<b>57</b>
4.3.1	Approches Symétriques Compositionnelles Exactes (SCE) . . . . .	57
4.3.2	Approches approximativement symétriques . . . . .	61
4.3.3	Approches Bidirectionnelles Pures . . . . .	62
<b>4.4</b>	<b>Équivalence des algorithmes d'une même famille dans le cas d'une optimisation par l'algorithme Gauss-Newton . . . . .</b>	<b>66</b>
4.4.1	Hypothèses . . . . .	66
4.4.2	Fondements de l'équivalence . . . . .	66
4.4.3	Équivalence des approches au sein d'une même famille . . . . .	67
4.4.4	Contraintes sur l'espace bidirectionnel et règle de mise à jour . . . . .	71
<b>4.5</b>	<b>Conclusions . . . . .</b>	<b>72</b>
4.5.1	Bilan . . . . .	72
4.5.2	Conception d'approches pour le suivi d'objet . . . . .	72

---

Après avoir défini le problème de recalage du point de vue usuel au chapitre 2, nous avons introduit dans le chapitre précédent la notion de compensation bidirectionnelle pour formaliser le processus d'alignement itératif. Dans ce chapitre, nous utilisons ces concepts pour porter un éclairage nouveau sur les approches de l'état de l'art.

Dans un premier temps, nous montrons qu'il est possible d'intégrer les approches présentées à la section 2.2 aux différentes familles (F), (I) et (B) évoquées à la section 3.3.1. La classification repose sur le type de compensation utilisée sur les images (section 3.3.1.1) et la nature de la paramétrisation incrémentale (section 3.3.2).

Dans un second temps, nous étendons la démonstration d'équivalence des approches FA, FC, IC et IA de [Baker 2004] pour les fonctionnelles SSD avec optimisation Gauss-Newton en démontrant notamment l'équivalence à l'ordre 1 des approches au sein d'une même famille.



A l'issue de ce travail de réinterprétation de l'existant, nous déterminons un certain nombre de lignes directrices pour la conception d'algorithmes itératifs dans le contexte particulier du suivi d'objet.

Les différentes approches de la section 2.2 utilisent une règle de mise à jour directe. C'est pourquoi dans la suite nous nous intéressons uniquement à cette règle de mise à jour. Par conséquent, dans tout ce chapitre  $\boldsymbol{\mu}_T^n$  reste fixe au cours des itérations et est égal à  $\boldsymbol{\mu}_T^0$ .

## 4.1 Application de l'algorithme Gauss-Newton aux approches SSD

A l'issue de l'étape de paramétrisation incrémentale, la fonctionnelle ne dépend que de l'incrément  $\mathbf{v}$ . Dans le cas particulier d'une fonctionnelle SSD, cette dernière peut s'écrire :

$$E(\mathbf{v}) = \mathbf{e}(\mathbf{v})^T \mathbf{e}(\mathbf{v}), \quad (4.1)$$

avec  $\mathbf{e}(\mathbf{v})$  défini par la concaténation des erreurs pixels à pixels  $e_i$  :

$$e_i(\mathbf{v}) = I(\mathbf{W}(\boldsymbol{\mu}_I(\boldsymbol{\mu}_I^n, \mathbf{v}), \mathbf{x}_i)) - T(\mathbf{W}(\boldsymbol{\mu}_T(\boldsymbol{\mu}_T^n, \mathbf{v}), \mathbf{x}_i)) \quad (4.2)$$

L'application de l'algorithme Gauss-Newton (présenté à la section 2.2) à cette fonctionnelle conduit à la solution :

$$\mathbf{v} = -\mathbf{J}_e(\mathbf{0})^+ \mathbf{e}(\mathbf{0}) \quad (4.3)$$

avec :

$$\mathbf{J}_e(\mathbf{0}) = \left. \frac{\partial \mathbf{e}(\mathbf{v})}{\partial \mathbf{v}} \right|_{\mathbf{0}} \quad (4.4)$$

Selon la nature des contraintes imposées sur l'espace bidirectionnel,  $\mathbf{J}_e(\mathbf{0})$  va s'exprimer comme une combinaison des matrices jacobiniennes associées à chaque image (cf. sections 4.2 et 4.3). Ces matrices Jacobiennes sont elles-mêmes dépendantes de la paramétrisation incrémentale choisie. Si nous considérons les paramétrisations discutées à la section 3.3.2 et résumées dans le tableau 3.1, nous pouvons définir trois matrices Jacobiennes élémentaires associées à une image quelconque  $L$  (pouvant être  $I$  ou  $T$ ) et aux paramètres initiaux  $\boldsymbol{\mu}_L^n$ .

**Paramétrisation additive (équation (3.17)).**

$$\mathbf{J}_{L,i}^A(\mathbf{0}) = \left. \frac{\partial L(\boldsymbol{\mu}_L^n + \mathbf{v}, \mathbf{x}_i)}{\partial \mathbf{v}} \right|_{\mathbf{0}} \quad (4.5)$$

**Paramétrisation compositionnelle (équation (3.18)).**

$$\mathbf{J}_{L,i}^C(\mathbf{0}) = \left. \frac{\partial L(\mathbf{W}(\boldsymbol{\mu}_L^n \circ \mathbf{v}, \mathbf{x}_i))}{\partial \mathbf{v}} \right|_{\mathbf{0}} \quad (4.6)$$

**Paramétrisation compositionnelle sur groupe de Lie (équation (3.19)).** La version sur algèbre de Lie constitue en fait un cas particulier de la paramétrisation compositionnelle à droite.

$$\mathbf{J}_{L,i}^{CL}(\mathbf{0}) = \left. \frac{\partial L(\mathbf{W}(\boldsymbol{\mu}_L^n \circ \exp(\mathbf{v}), \mathbf{x}_i))}{\partial \mathbf{v}} \right|_{\mathbf{0}} \quad (4.7)$$

La majeure partie des approches de la littérature utilisent une matrice Jacobienne construite à partir de ces matrices élémentaires. Nous discuterons au cas par cas celles qui utilisent une paramétrisation particulière. L'incrément  $\mathbf{v}$  estimé grâce à l'équation (4.3) est injecté dans l'équation (3.25) de mise à jour directe des paramètres d'alignement.

## 4.2 Classification des approches de l'état de l'art : méthodes unidirectionnelles

Dans cette section nous élargissons la taxonomie de [Baker 2004] en classifiant les différentes approches utilisant une compensation directe ou inverse.

### 4.2.1 Les approches Directes (F) (Forward)

#### 4.2.1.1 Les approches directes additives ou *Forwards Additive* (FA)

L'approche FA présentée section 2.2.2 utilise :

$$\boldsymbol{\mu}_I(\mathbf{v}) = \boldsymbol{\mu}_I^n + \mathbf{v} \quad \text{et} \quad \boldsymbol{\mu}_T(\mathbf{v}) = \boldsymbol{\mu}_T^n. \quad (4.8)$$

La matrice Jacobienne associée  $\mathbf{J}_e^{\text{FA}}$  est :

$$\mathbf{J}_e^{\text{FA}}(\mathbf{0}) = \mathbf{J}_I^A(\mathbf{0}). \quad (4.9)$$

L'utilisation de  $\boldsymbol{\mu}_T^n$  généralise ici la formalisation proposée dans [Baker 2004] ou ce dernier vaut  $\mathbf{0}$ . Les deux formalisations sont équivalentes en remplaçant  $T$  par  $T^n = \text{warp}(T, \boldsymbol{\mu}_T^n)$ . La règle de mise à jour directe associée est :

$$\boldsymbol{\mu}_I^{n+1} = \boldsymbol{\mu}_I^n + \mathbf{v} \quad (4.10)$$

La paramétrisation incrémentale FA a été adaptée à d'autres fonctionnelles d'erreur que la SSD comme le coefficient de corrélation normalisé [Evangelidis 2008], les fonctionnelles à noyaux reposant sur des distributions d'attribut [Hager 2004, Guskov 2006] ou encore l'information mutuelle [Dowson 2008].

#### 4.2.1.2 Les approches directes compositionnelles ou *Forwards Compositional* (FC, FCL)

L'approche FC [Shum 2000] présentée section 2.2.2 s'inscrit naturellement dans le formalisme bidirectionnel en posant :

$$\boldsymbol{\mu}_I(\mathbf{v}) = \boldsymbol{\mu}_I^n \circ \mathbf{v} \quad \text{et} \quad \boldsymbol{\mu}_T(\mathbf{v}) = \boldsymbol{\mu}_T^n \quad (4.11)$$

La matrice Jacobienne associée vérifie :

$$\mathbf{J}_e^{\text{FC}}(\mathbf{0}) = \mathbf{J}_I^C(\mathbf{0}) \quad (4.12)$$

Cette approche utilise une loi de composition à droite et la règle de mise à jour :

$$\boldsymbol{\mu}_I^{n+1} = \boldsymbol{\mu}_I^n \circ \mathbf{v} \quad (4.13)$$

Une approche similaire à l'approche FC est proposée dans [Keller 2004] sous le nom de *Basic Gradient Method*. Les auteurs utilisent la même paramétrisation incrémentale et la même matrice Jacobienne que l'approche FC. En revanche ils utilisent une règle de mise à jour additive  $\boldsymbol{\mu}_I^{n+1} = \boldsymbol{\mu}_I^n + \mathbf{v}$  ( $\boldsymbol{\mu}_T^n$  restant fixe). Cette approche ne respecte pas la règle de mise à jour directe (3.25), ce qui peut poser des problèmes de convergence discutés à la section 4.4.

Une variante de l'approche FC, nommée par les auteurs *LIEMANIFOLD*, est proposée par [Guangwei 2008]. Les auteurs utilisent une paramétrisation incrémentale sur groupe de Lie :

$$\boldsymbol{\mu}_I(\mathbf{v}) = \boldsymbol{\mu}_I^n \circ \exp(\mathbf{v}) \quad \text{et} \quad \boldsymbol{\mu}_T(\mathbf{v}) = \boldsymbol{\mu}_T^n. \quad (4.14)$$

Nous renommons cette approche dans la suite FCL (*Forward Composition on Lie Groups*) en accord avec notre taxonomie. La matrice Jacobienne associée avec une optimisation Gauss-Newton  $\mathbf{J}_e^{\text{FCL}}$  vérifie :

$$\mathbf{J}_e^{\text{FCL}}(\mathbf{0}) = \mathbf{J}_I^{\text{CL}}(\mathbf{0}) \quad (4.15)$$

La loi de mise à jour associée est :

$$\boldsymbol{\mu}_I^{n+1} = \boldsymbol{\mu}_I^n \circ \exp(\mathbf{v}). \quad (4.16)$$

$\boldsymbol{\mu}_T^n$  reste fixe à la valeur  $\boldsymbol{\mu}_T^0$ . L'intérêt de cette approche repose sur l'utilisation d'un incrément dans l'algèbre de Lie. Comme nous l'avons discuté à la section 2.1.2, cela permet d'utiliser une paramétrisation de l'homographie plus adaptée, dans certaines situations, que la paramétrisation matricielle classique et cela sans ajout d'une contrainte supplémentaire lors de l'optimisation.

Nous avons vu à la section 2.2.3.2 que l'algorithme ESM a été initialement présenté dans [Benhimane 2007b] comme une alternative à l'approche Gauss-Newton pour minimiser la fonctionnelle SSD utilisant la paramétrisation incrémentale FCL. Nous choisissons cependant de classer cette approche comme une approche symétrique, ce qui sera détaillé et justifié à la section 4.3.1.

Les approches directes ont été combinées avec des méthodes d'optimisation de type Newton dans [Baker 2004], et des approches d'ordre 3 dans [Keller 2008].

## 4.2.2 Les approches Inverses (I)

### 4.2.2.1 Les approches inverses compositionnelles (ICR, ICD et ICL)

L'approche Inverse Compositionnelle IC [Baker 2001, Baker 2004] utilise les contraintes :

$$\boldsymbol{\mu}_I(\mathbf{v}) = \boldsymbol{\mu}_I^n \quad \text{et} \quad \boldsymbol{\mu}_T(\mathbf{v}) = \boldsymbol{\mu}_T^n \circ \mathbf{v}. \quad (4.17)$$

ce qui conduit à :

$$\mathbf{J}_e^{\text{ICR}}(\mathbf{0}) = -\mathbf{J}_T^C(\mathbf{0}). \quad (4.18)$$

L'approche initiale utilise une règle de mise à jour directe :

$$\boldsymbol{\mu}_I^{n+1} = \boldsymbol{\mu}_I^n \circ \mathbf{v}^{-1}, \quad (4.19)$$

$\boldsymbol{\mu}_T^n$  restant inchangé. Cette mise à jour faisant intervenir l'inverse de l'incrément, nous renommons cette approche ICR (*Inverse Compositional Reverse*). Cette approche a également été appliquée [Matthews 2004b] avec succès pour l'estimation de transformations complexes comme les modèles d'apparence actifs. Une adaptation de cette méthode pour fonctionner avec le coefficient de corrélation normalisé a été réalisée dans [Evangelidis 2008]. Les auteurs de [Mégret 2006] utilisent également la paramétrisation incrémentale IC avec une fonctionnelle de similarité entre distributions d'attributs. Cette approche a enfin été adaptée pour maximiser l'information mutuelle [Dowson 2008].

En remplaçant  $\mathbf{v}$  par  $\mathbf{v}^{-1}$  dans  $\boldsymbol{\mu}_T$  :

$$\boldsymbol{\mu}_I(\mathbf{v}) = \boldsymbol{\mu}_I^n \quad \text{et} \quad \boldsymbol{\mu}_T(\mathbf{v}) = \boldsymbol{\mu}_T^n \circ \mathbf{v}^{-1}. \quad (4.20)$$

nous pouvons définir une nouvelle approche nommée ICD (*Inverse Compositional Direct*). L'utilisation de l'algorithme d'optimisation Gauss-Newton conduit à une approche de même complexité que l'ICR avec  $\mathbf{J}_e^{\text{ICD}}(\mathbf{0}) = -\mathbf{J}_e^{\text{ICR}}(\mathbf{0})$ . L'utilisation d'une règle de mise à jour directe conduit alors à ne plus inverser  $\mathbf{v}$ , ce qui peut être un avantage par exemple dans des applications d'imagerie médicale où certains modèles de mouvement ne permettent pas un calcul aisé de l'inverse d'une transformation.

Enfin, en utilisant les mêmes principes, nous pouvons introduire la paramétrisation Inverse Compositionnelle sur Groupe de Lie ICL qui utilise une paramétrisation sur groupe de Lie et dont les caractéristiques sont présentées dans le tableau 4.1. Nous n'avons pas trouvé dans la littérature de références explicites à cette paramétrisation incrémentale.

### 4.2.2.2 Approche IAR (Inverse Additive Reverse)

Il existe plusieurs façon d'introduire un incrément additif correctif sur  $T$ . L'approche IAR (*Inverse Additive Reverse*) est la plus intuitive, car elle est la version duale de l'approche FA, en inversant les rôles de  $I$  et  $T$  :

$$\boldsymbol{\mu}_I(\mathbf{v}) = \boldsymbol{\mu}_I^n \quad \text{et} \quad \boldsymbol{\mu}_T(\mathbf{v}) = \boldsymbol{\mu}_T^n + \mathbf{v}. \quad (4.21)$$

La règle de mise à jour directe associée est cependant plus complexe car il est nécessaire de traduire l'incrément additif sur  $\boldsymbol{\mu}_T^n$  vers le référentiel de  $I$  :

$$\boldsymbol{\mu}_I^{n+1} = \boldsymbol{\mu}_I^n \circ (\boldsymbol{\mu}_T^n + \mathbf{v})^{-1} \circ \boldsymbol{\mu}_T^n, \quad (4.22)$$

$\boldsymbol{\mu}_T^n$  restant fixe. Le Jacobien associé à cette approche est :

$$\mathbf{J}_e^{\text{IAR}}(\mathbf{0}) = -\mathbf{J}_T^A(\mathbf{0}). \quad (4.23)$$

Nous n'avons pas trouvé de référence explicite à cette paramétrisation incrémentale. Nous l'abordons cependant ici afin de rendre claire les implications, sur la règle de mise à jour, de l'utilisation de la matrice Jacobienne  $\mathbf{J}_T^A(\mathbf{0})$  associée à une paramétrisation incrémentale additive.

### 4.2.2.3 Approche IAD (Inverse Additive Directe)

L'algorithme introduit par [Hager 1998] est nommé *Inverse Additive* dans [Baker 2004]. Dans la suite, nous nous référons à cet algorithme par l'acronyme IA. Comme présenté à la section 2.2.3.2, la justification initiale de cet algorithme repose sur la minimisation d'une fonctionnelle utilisant une paramétrisation incrémentale FA, mais où le rôle de  $I$  et  $T$  est inversé lors du calcul de la matrice Jacobienne, en faisant l'hypothèse (2.51) que les images sont quasiment identiques une fois compensées par les paramètres d'alignement estimés à l'itération  $n$  :

$$I(\mathbf{W}(\boldsymbol{\mu}^n, \mathbf{x}_i)) \approx T(\mathbf{x}_i), \quad (4.24)$$

hypothèse très forte sur le contenu des deux images.

Nous proposons dans la suite une nouvelle paramétrisation incrémentale pour laquelle l'application de l'algorithme Gauss-Newton conduit naturellement à l'approche IA, sans nécessiter aucune hypothèse supplémentaire :

$$\boldsymbol{\mu}_I(\mathbf{v}) = \boldsymbol{\mu}_I^n, \quad (4.25)$$

$$\boldsymbol{\mu}_T(\mathbf{v}) = \boldsymbol{\mu}_T^n \circ (\boldsymbol{\mu}_I^n + \mathbf{v})^{-1} \circ \boldsymbol{\mu}_I^n. \quad (4.26)$$

Nous nommons cette approche IAD, car la mise à jour directe implique l'addition de l'incrément :

$$\boldsymbol{\mu}_I^{n+1} = \boldsymbol{\mu}_I^n + \mathbf{v}. \quad (4.27)$$

La matrice Jacobienne associée est obtenue par concaténation verticale des contributions de chaque pixel :

$$\mathbf{J}_{e,i}^{\text{IAD}}(\mathbf{0}) = \left. \frac{\partial T(\mathbf{W}(\boldsymbol{\mu}_T^n \circ (\boldsymbol{\mu}_I^n + \mathbf{v})^{-1} \circ \boldsymbol{\mu}_I^n, \mathbf{x}_i))}{\partial \mathbf{v}} \right|_{\mathbf{0}}. \quad (4.28)$$

Il est possible de montrer que cette matrice Jacobienne vérifie (cf. annexe A) :

$$\mathbf{J}_{e,i}^{\text{IAD}}(\mathbf{0}) = \nabla T(\mathbf{W}(\boldsymbol{\mu}_T^n, \mathbf{x}_i)) \left( \left. \frac{\partial \mathbf{W}(\boldsymbol{\mu}_I^n, \mathbf{z})}{\partial \mathbf{z}} \right|_{\mathbf{x}_i} \right)^{-1} \left. \frac{\partial \mathbf{W}(\boldsymbol{\mu}, \mathbf{x}_i)}{\partial \boldsymbol{\mu}} \right|_{\boldsymbol{\mu}_I^n}. \quad (4.29)$$

avec  $\nabla T(\mathbf{W}(\boldsymbol{\mu}_T^n, \mathbf{x}_i))$  les gradients du template :

$$\nabla T(\mathbf{W}(\boldsymbol{\mu}_T^n, \mathbf{x}_i)) = \left. \frac{\partial T(\mathbf{W}(\boldsymbol{\mu}_T^n, \mathbf{z}))}{\partial \mathbf{z}} \right|_{\mathbf{x}_i}. \quad (4.30)$$

L'expression obtenue généralise l'expression (2.52) de l'approche IA de [Hager 1998], en introduisant les paramètres initiaux sur le template  $\boldsymbol{\mu}_T^n$ . Les deux expressions sont rigoureusement exactes si  $\boldsymbol{\mu}_I^n = \mathbf{0}$ . Nous remarquons également que les deux approches utilisent la même règle de mise à jour. Cette analyse révèle ainsi que l'algorithme de [Hager 1998] est un cas particulier de minimisation d'une erreur bidirectionnelle utilisant la paramétrisation incrémentale IAD. Cette formulation est valide même quand (4.24) n'est pas vérifiée.

### 4.2.2.4 Approche GIC

Nous montrons dans cette partie que l'approche GIC présentée initialement par les auteurs de [Brooks 2009] comme une extension de l'approche IC correspond en réalité à une expression différente de l'approche IAD présentée à la section précédente.

### 4.3. Classification des approches de l'état de l'art : méthodes bidirectionnelles

**Point de vue initial.** En se reposant sur ce que nous appelons *Procédure de traduction de Brooks* que nous présentons à l'annexe B, les auteurs de [Brooks 2009] proposent l'approche *Generalized Inverse Compositional* GIC utilisant une loi de mise à jour additive :

$$\boldsymbol{\mu}_I^{n+1} = \boldsymbol{\mu}_I^n + \mathbf{v} \quad (4.31)$$

A partir de l'équation (B.4), ils paramètrent  $\boldsymbol{\mu}_T$  sous la forme :

$$\boldsymbol{\mu}_T(\mathbf{v}) = (\boldsymbol{\mu}_I^n + \mathbf{v})^{-1} \circ \boldsymbol{\mu}_I^n \quad (4.32)$$

ce qui conduit à la matrice Jacobienne associée au template :

$$\mathbf{J}_e^{\text{GIC}}(\mathbf{0}) = \mathbf{J}_T^C(\mathbf{0})\mathcal{M}_{I \rightarrow T}, \quad (4.33)$$

avec :

$$\mathcal{M}_{I \rightarrow T} = \left. \frac{\partial (\boldsymbol{\mu}_I^n + \mathbf{v})^{-1} \circ \boldsymbol{\mu}_I^n}{\partial \mathbf{v}} \right|_{\mathbf{0}}. \quad (4.34)$$

L'expression (4.33) permet de traduire la matrice Jacobienne inverse compositionnelle  $\mathbf{J}_e^{\text{IC}}(\mathbf{0}) = \mathbf{J}_T^C(\mathbf{0})$  afin qu'elle soit compatible avec des paramètres de mouvement additif  $\mathbf{v}$  affectant l'image  $I$ .

**Point de vue du formalisme.** La matrice Jacobienne (4.33) correspond à la matrice Jacobienne associée à la fonctionnelle SSD  $E(\boldsymbol{\mu}_I^n, (\boldsymbol{\mu}_I^n + \mathbf{v})^{-1} \circ \boldsymbol{\mu}_I^n)$ . Nous avons vu à la section 4.2.2.3, que cette erreur correspondait à l'erreur minimisée par l'approche IA de [Hager 1998]. Ainsi l'approche GIC et l'approche IA, bien que présentée de manière très différente, correspondent en réalité au calcul des mêmes quantités. La différence, en pratique, se situe uniquement dans la façon de calculer la matrice Jacobienne. Théoriquement les matrices Jacobiennes utilisées sont identiques :

$$\underbrace{\mathbf{J}_e^{\text{GIC}}(\mathbf{0})}_{(4.33)} = \underbrace{\mathbf{J}_e^{\text{IA}}(\mathbf{0})}_{(2.52)}. \quad (4.35)$$

La généralisation de ces approches pour  $\boldsymbol{\mu}_T^n \neq \mathbf{0}$  conduit à la matrice Jacobienne de la paramétrisation incrémentale IAD (4.29).

## 4.3 Classification des approches de l'état de l'art : méthodes bidirectionnelles

### 4.3.1 Approches Symétriques Compositionnelles Exactes (SCE)

Nous appelons approche symétrique compositionnelle exacte toute approche Bidirectionnelle Subordonnée vérifiant :

$$(\boldsymbol{\mu}_I^n)^{-1} \circ \boldsymbol{\mu}_I(\boldsymbol{\mu}_I^n, \mathbf{v}) = \boldsymbol{\mu}_T(\boldsymbol{\mu}_T^n, \mathbf{v})^{-1} \circ \boldsymbol{\mu}_T^n \quad (4.36)$$

Cette paramétrisation incrémentale est appelée symétrique car le repère de compensation choisi se situe exactement "au milieu" des repères des images compensées  $I^n$  et  $T^n$  au sens compositionnel du terme. Ce phénomène est illustré à la figure 4.1 pour une approche symétrique compositionnelle particulière, l'approche SCL qui est présentée dans la partie suivante.

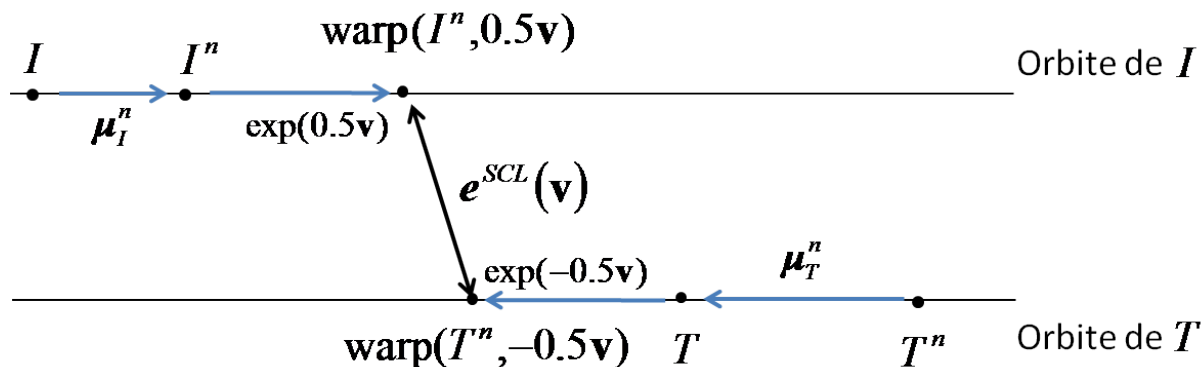


FIGURE 4.1 – Interprétation géométrique de la fonctionnelle *Symmetric Composition on Lie Group* SCL, le repère de compensation est exactement au milieu des deux repères associés aux images compensées avec les estimations courantes des paramètres.

#### 4.3.1.1 L'approche SCL-ESM

Nous avons vu à la section 2.2.3.2 que l'algorithme ESM se différencie de la plupart des approches de l'état de l'art, d'une part par l'utilisation de paramètres appartenant nécessairement à un groupe de Lie, et d'autre part en faisant l'hypothèse forte que les deux images  $I$  et  $T$  sont identiques à la compensation de mouvement près (equation (2.55))

$$\forall \mathbf{x}_i \quad I(\mathbf{W}(\boldsymbol{\mu}^n \circ \exp(\mathbf{v}^*), \mathbf{x}_i)) = T(\mathbf{x}_i), \quad (4.37)$$

c'est à dire que  $I$  et  $T$  partagent la même orbite. Initialement, les auteurs de [Benhimane 2007b] présentent cette approche pour minimiser la fonctionnelle SSD utilisant une paramétrisation incrémentale FCL.

Nous proposons maintenant une nouvelle interprétation de l'algorithme ESM. Nous nous intéressons pour cela à la paramétrisation incrémentale symétrique sur groupe de Lie SCL (*Symmetric Compositional on Lie Group*) utilisant :

$$\boldsymbol{\mu}_I(\mathbf{v}) = \boldsymbol{\mu}_I^n \circ \exp\left(\frac{1}{2}\mathbf{v}\right), \quad (4.38)$$

$$\boldsymbol{\mu}_T(\mathbf{v}) = \boldsymbol{\mu}_T^n \circ \exp\left(-\frac{1}{2}\mathbf{v}\right). \quad (4.39)$$

L'utilisation des équations (4.38) et (4.39) dans les expressions (4.4) conduisent au Jacobien  $\mathbf{J}_e^{\text{SCL}}$  :

$$\mathbf{J}_e^{\text{SCL}}(\mathbf{0}) = \frac{1}{2}(\mathbf{J}_I^{\text{CL}}(\mathbf{0}) + \mathbf{J}_T^{\text{CL}}(\mathbf{0})). \quad (4.40)$$

En raison des propriétés de l'application exponentielle (2.18), la règle de mise à jour directe associée est :

$$\boldsymbol{\mu}_I^{n+1} = \boldsymbol{\mu}_I^n \circ \exp(\mathbf{v}) \quad (4.41)$$

En comparant l'expression de  $\mathbf{J}_e^{\text{SCL}}(\mathbf{0})$  (4.40) avec l'expression (2.56), il apparaît que l'approche SCL utilise la même matrice Jacobienne que l'ESM quand  $\boldsymbol{\mu}_T^n$  vaut  $\mathbf{0}$ . Les deux approches utilisent également la même loi de mise à jour.

Cette relation, entre les deux approches, met en lumière de nouvelles propriétés de l'algorithme ESM. En effet, lorsque l'hypothèse (2.55) n'est pas vérifiée, l'analyse précédente révèle que les équations de l'ESM conduisent alors à minimiser l'erreur SCL en utilisant l'approche Gauss-Newton avec une règle de mise à jour directe. L'algorithme ESM peut donc être vu sous deux angles différents :

- soit comme une procédure d'optimisation de l'erreur FCL lorsque les deux images partagent la même orbite,
- soit comme l'application de la procédure d'optimisation Gauss-Newton à l'erreur SCL dans le cas plus général où les deux images ne partagent pas la même orbite (ce qui est le cas en pratique, le modèle de déformation géométrique ne modélisant pas à lui seul les variations d'intensité lumineuse de l'image).

Nous approfondirons ces liens au sein du chapitre 5, où nous verrons comment concilier les deux points de vue dans le cadre de l'analyse des approches *Asymmetric Compositional on Lie Group* (ACL) qui généralisent les approches FCL et SCL.

#### 4.3.1.2 L'approche GESM

L'ESM nécessite une règle de mise à jour compositionnelle et une paramétrisation sur groupe de Lie ce qui peut être une limitation dans certaines applications, notamment si l'on souhaite implanter cet algorithme dans un programme d'optimisation générique utilisant une règle de mise à jour standard additive [Brooks 2009].

**Point de vue des auteurs.** Comme pour l'approche GIC, en se reposant sur ce que nous appelons *Procédure de traduction de Brooks* que nous présentons à l'annexe B, les auteurs utilisent l'équation (B.4) pour rendre interprétable la matrice Jacobienne sur  $T$  relativement à la transformation directe  $\boldsymbol{\mu}_I(\boldsymbol{\mu}_I^n, \mathbf{v})$ . Par analogie avec l'équation de la matrice Jacobienne de l'ESM (4.40), les auteurs proposent alors de moyenniser les matrices Jacobiennes associées à  $I$  et  $T$  pour une même transformation directe  $\boldsymbol{\mu}_I(\boldsymbol{\mu}_I^n, \mathbf{v})$ . Ils définissent ainsi une extension de l'ESM [Benhimane 2007b], l'algorithme GESM (*Generalized ESM*) qui utilise la règle de mise à jour (nous reprenons ici l'expression utilisant  $\mathbf{v}^{-1}$  telle que présentée par les auteurs, bien que le même raisonnement puisse être fait avec  $\mathbf{v}$  à la place) :

$$\boldsymbol{\mu}_I^{n+1} = \boldsymbol{\mu}_I^n + \mathbf{v}^{-1} \quad (4.42)$$

et la matrice Jacobienne moyenne :

$$\mathbf{J}_e^{\text{GESM}}(\mathbf{0}) = \frac{1}{2}(\mathbf{J}_I(\mathbf{0}) + \mathbf{J}_T(\mathbf{0})) \quad (4.43)$$

avec, en interprétant l'explication des auteurs avec nos notations :

$$\mathbf{J}_{I,i}(\mathbf{0}) = \left. \frac{\partial I(\mathbf{W}(\boldsymbol{\mu}_I^n + \mathbf{v}^{-1}, \mathbf{x}_i))}{\partial \mathbf{v}} \right|_{\mathbf{0}} \quad (4.44)$$

$$\mathbf{J}_{T,i}(\mathbf{0}) = \left. \frac{\partial T(\mathbf{W}((\boldsymbol{\mu}_I^n + \mathbf{v}^{-1})^{-1} \circ \boldsymbol{\mu}_I^n, \mathbf{x}_i))}{\partial \mathbf{v}} \right|_{\mathbf{0}} = \mathbf{J}_T^C(\mathbf{0}) \mathcal{M}_{I \rightarrow T} \quad (4.45)$$



avec :

$$\mathcal{M}_{I \rightarrow T} = \left. \frac{\partial(\boldsymbol{\mu}_I^n + \mathbf{v}^{-1})^{-1} \circ \boldsymbol{\mu}_I^n}{\partial \mathbf{v}} \right|_{\mathbf{0}} \quad (4.46)$$

Les auteurs soulignent qu'il n'est pas garanti que cette approche conserve les bonnes propriétés de convergence de l'ESM [Brooks 2009] d'un point de vue théorique. Les expérimentations effectuées par les auteurs montrent que cette approche permet d'augmenter les performances en termes de fréquence de convergence (mesurée sous la forme d'un taux d'échec *Failure Rate*) par rapport aux approches FA et GIC, notamment lorsque le mouvement à estimer est important. Cependant ils n'observent pas de gain en termes de vitesse de convergence (mesuré sous la forme du temps de calcul avant convergence).

**Formalisme bidirectionnel : un autre point de vue** Considérons la paramétrisation incrémentale suivante :

$$\boldsymbol{\mu}_I(\mathbf{v}) = \boldsymbol{\mu}_I^n + \left( \frac{1}{2} \mathbf{v}^{-1} \right), \quad (4.47)$$

$$\boldsymbol{\mu}_T(\mathbf{v}) = \boldsymbol{\mu}_T^n \circ \left( \boldsymbol{\mu}_I^n + \left( \frac{1}{2} \mathbf{v}^{-1} \right) \right)^{-1} \circ \boldsymbol{\mu}_I^n. \quad (4.48)$$

Cette paramétrisation correspond à une symétrie compositionnelle exacte, telle que définie à l'équation (4.36). Afin de bien mettre en évidence les paramétrisations incrémentales de l'approche, nous renommons cette dernière FA+IAD car elle mélange des paramétrisations incrémentales homogènes à celles des approches FA et IAD, faisant ainsi apparaître une matrice Jacobienne combinant les matrices Jacobiennes de ces deux approches. La matrice Jacobienne associée correspond à :

$$\mathbf{J}_e^{\text{FA+IAD}}(\mathbf{0}) = -\frac{1}{2}(\mathbf{J}_e^{\text{FA}}(\mathbf{0}) + \mathbf{J}_e^{\text{IAD}}(\mathbf{0})), \quad (4.49)$$

qui est identique à (4.43) si nous imposons  $\boldsymbol{\mu}_T^n = \mathbf{0}$ . La règle de mise à jour directe correspondante devrait être :

$$\boldsymbol{\mu}_I^{n+1} = \left( \boldsymbol{\mu}_I^n + \left( \frac{1}{2} \mathbf{v}^{-1} \right) \right) \circ (\boldsymbol{\mu}_I^n)^{-1} \circ \left( \boldsymbol{\mu}_I^n + \left( \frac{1}{2} \mathbf{v}^{-1} \right) \right). \quad (4.50)$$

Les auteurs utilisent pourtant (4.42). Nous montrons à l'annexe B.2 que la règle de mise à jour employée par les auteurs correspond à une approximation à l'ordre 1 de la règle de mise à jour exacte (4.50). Ainsi nous avons :

$$\boldsymbol{\mu}_I^{n+1} = \boldsymbol{\mu}_I^n + \mathbf{v}^{-1} + O(\|\mathbf{v}\|^2) \quad (4.51)$$

Du point de vue de notre formalisme, l'approche GESM peut être interprétée comme l'optimisation de l'erreur utilisant les paramétrisations incrémentales (4.47) et (4.48), mais utilisant une règle de mise à jour approximative afin de pouvoir l'intégrer au sein de solveurs qui imposent d'utiliser une règle de mise à jour additive.

Comme nous l'avons illustré à travers les approches GIC et GESM, la procédure de traduction de Brooks permet d'introduire les gradients de  $T$  dans la minimisation de la fonctionnelle  $E(\boldsymbol{\mu}_I, \mathbf{0})$ . En cela ces approches peuvent être considérées comme des heuristiques alternatives à l'approche Gauss-Newton (dans le cas d'une fonctionnelle SSD) pour minimiser l'erreur directe. C'est pourquoi nous avons classé ces approches avec les algorithmes ESM et IA dans l'état de l'art (section 2.2.3.2).

### 4.3. Classification des approches de l'état de l'art : méthodes bidirectionnelles

L'utilisation du formalisme bidirectionnel nous permet de porter un éclairage nouveaux sur ces approches :

- en montrant à la section 4.2.2.4 que l'approche GIC est en fait une expression différente de l'approche IA.
- en identifiant dans cette section une paramétrisation incrémentale associée à l'approche GESM. Nous montrons également que cette méthode procède à une approximation à l'ordre 1 sur la règle de mise à jour.

De manière plus générale, en envisageant toujours le mouvement relativement au repère d'une des deux images, les auteurs de [Brooks 2009] imposent l'utilisation, lors de l'optimisation, d'un vecteur  $\mathbf{v}$  de longueur  $p$  ou  $p$  correspond au nombre de degrés de liberté de la transformation, c'est à dire à la dimension de  $\mathcal{P}$ . Ils s'interdisent ainsi l'utilisation d'un vecteur de longueur plus importante comme ceux de longueur  $2 \times p$  utilisés par les approches de la famille bidirectionnelle pure. Cela constitue de notre point de vue une limite par rapport à notre formalisme. Nous verrons notamment dans le chapitre 5.2 que l'utilisation d'un vecteur incrément bidirectionnel permet d'améliorer les performances en présence d'images bruitées.

De plus, en faisant clairement apparaître la possibilité de définir des fonctionnelles symétriques ou bidirectionnelles pour lesquelles le repère de compensation est différent du repère associé à chaque image, notre formalisme permet la définition d'algorithmes *Inverse Consistent* qui sont utilisés en imagerie médicale [Joshi 2004].

#### 4.3.1.3 Approche SC

L'approche GESM utilise une règle de mise à jour additive mais suppose une formalisation compositionnelle pour pouvoir mélanger les gradients de  $I$  et  $T$ . Lorsque aucune contrainte sur la forme de la règle de mise à jour n'est imposée, une solution plus simple pour exprimer une approche symétrique compositionnelle ne nécessitant pas de paramétrisations incrémentales sur l'algèbre de Lie est d'utiliser l'approche SC (*Symmetric Compositional*) dont la paramétrisation incrémentale est :

$$\boldsymbol{\mu}_I(\mathbf{v}) = \boldsymbol{\mu}_I^n \circ \left(\frac{1}{2}\mathbf{v}\right) \quad \text{et} \quad \boldsymbol{\mu}_T(\mathbf{v}) = \boldsymbol{\mu}_T^n \circ \left(\frac{1}{2}\mathbf{v}\right)^{-1}. \quad (4.52)$$

La matrice jacobienne associée vérifie :

$$\mathbf{J}_e^{\text{SC}}(\mathbf{0}) = \frac{1}{2}(\mathbf{J}_I^C(\mathbf{0}) + \mathbf{J}_T^C(\mathbf{0})), \quad (4.53)$$

La loi de mise à jour directe associée est la suivante :

$$\boldsymbol{\mu}_I^{n+1} = \boldsymbol{\mu}_I^n \circ \left(\frac{1}{2}\mathbf{v}\right) \circ \left(\frac{1}{2}\mathbf{v}\right) \quad (4.54)$$

#### 4.3.2 Approches approximativement symétriques

Les approches présentées ici sont nommées -approximativement symétriques- car elles font intervenir de manière équitable les gradients de chaque image. Cependant elles ne vérifient pas la condition de symétrie exacte (4.36).

Une approche d'apparence très similaire aux approches symétriques compositionnelles a été proposée dans [Keller 2004]. Initialement nommée SGM (*Symmetric Gradient Method*), nous renommons cette approche SCO (*Symmetric Compositional Opposite*), le O correspondant au terme  $-\frac{1}{2}\mathbf{v}$  qui apparaît dans  $\boldsymbol{\mu}_T$ . La formulation présentée ici est une simple extension de celle présentée dans [Keller 2004] qui considérait  $\boldsymbol{\mu}_T^n = \mathbf{0}$  :

$$\boldsymbol{\mu}_I(\mathbf{v}) = \boldsymbol{\mu}_I^n \circ \frac{1}{2}\mathbf{v} \quad \text{et} \quad \boldsymbol{\mu}_T(\mathbf{v}) = \boldsymbol{\mu}_T^n \circ \left(-\frac{1}{2}\mathbf{v}\right). \quad (4.55)$$

La matrice Jacobienne associée vérifie :

$$\mathbf{J}_e^{\text{SCO}}(\mathbf{0}) = \frac{1}{2}(\mathbf{J}_I^C(\mathbf{0}) + \mathbf{J}_T^C(\mathbf{0})). \quad (4.56)$$

En se reposant sur l'équation (3.25), la règle de mise à jour directe associée devrait être :

$$\boldsymbol{\mu}_I^{n+1} = \boldsymbol{\mu}_I^n \circ \left(\frac{1}{2}\mathbf{v}\right) \circ \left(-\frac{1}{2}\mathbf{v}\right)^{-1}. \quad (4.57)$$

Les auteurs de [Keller 2004] utilisent quant à eux une règle de mise à jour additive  $\boldsymbol{\mu}_I^{n+1} = \boldsymbol{\mu}_I^n + \mathbf{v}$  qui n'est pas cohérente avec les paramétrisations incrémentales utilisées. Comme nous le voyons à la section 4.4, ceci a une influence importante sur les propriétés de convergence de l'algorithme.

Dans [Chen 2009], les auteurs ont introduit l'utilisation de deux paramétrisation incrémentales additives dans le cas de l'estimation d'une translation 2D. Dans notre formalisme, nous pouvons inclure leur approche sous la forme de la paramétrisation incrémentale :

$$\boldsymbol{\mu}_I(\mathbf{v}) = \boldsymbol{\mu}_I^n + \frac{1}{2}\mathbf{v} \quad \text{et} \quad \boldsymbol{\mu}_T(\mathbf{v}) = \boldsymbol{\mu}_T^n - \frac{1}{2}\mathbf{v}. \quad (4.58)$$

Nous nommons l'utilisation de ces paramétrisations l'approche SAD (*Symmetric Additive Direct*). A la différence des approches précédentes, les auteurs utilisent une erreur de type coefficient de corrélation normalisée. Cette approche peut être vue comme une extension symétrique des approches additives FA et IAR précédemment proposées. Les auteurs utilisent la règle de mise à jour  $\boldsymbol{\mu}_I^{n+1} = \boldsymbol{\mu}_I^n + \mathbf{v}$  qui semble différente de celle issue de la règle de mise à jour directe :

$$\boldsymbol{\mu}_I^{n+1} = (\boldsymbol{\mu}_I^n + \frac{1}{2}\mathbf{v}) \circ (\boldsymbol{\mu}_T^n - \frac{1}{2}\mathbf{v})^{-1} \circ \boldsymbol{\mu}_T^n \quad (4.59)$$

Cependant dans le cas de l'estimation d'une translation, la composition de paramètre correspond à l'addition et l'inversion de paramètre vérifie  $\boldsymbol{\mu}^{-1} = -\boldsymbol{\mu}$ . Par conséquent, la règle de mise à jour utilisée par les auteurs est valide. Ainsi, l'approche de [Chen 2009] peut être étendue de deux façons pour un mouvement plus générique, soit en utilisant la paramétrisation SC, soit en utilisant la paramétrisation SAD (en veillant à utiliser la règle de mise à jour 4.59).

### 4.3.3 Approches Bidirectionnelles Pures

Une approche purement bidirectionnelle compositionnelle a été proposée dans [Keller 2004] et présentée à la section 2.2.3.1. L'approche initiale utilise implicitement  $\boldsymbol{\mu}_T^n = \mathbf{0}$ , nous la présentons ici avec une initialisation  $\boldsymbol{\mu}_T^n$  quelconque. Elle est intégrée au formalisme via les paramétrisations :

$$\boldsymbol{\mu}_I(\mathbf{v}) = \boldsymbol{\mu}_I^n \circ \mathbf{v}_I \quad \text{et} \quad \boldsymbol{\mu}_T(\mathbf{v}) = \boldsymbol{\mu}_T^n \circ (-\mathbf{v}_T). \quad (4.60)$$

### 4.3. Classification des approches de l'état de l'art : méthodes bidirectionnelles

Pour rester cohérent avec le reste de la taxonomie, nous renommons cette approche BCO (*Bidirectional Compositional Opposite*). La matrice Jacobienne associée est obtenue par concaténation horizontale des matrices Jacobiennes de chaque image :

$$\mathbf{J}_e^{\text{BCO}}(\mathbf{0}) = [\mathbf{J}_I^C(\mathbf{0}) \quad \mathbf{J}_T^C(\mathbf{0})]. \quad (4.61)$$

Dans l'article original [Keller 2004], les auteurs utilisent une règle de mise à jour additive  $\boldsymbol{\mu}_I^{n+1} = \boldsymbol{\mu}_I^n + \mathbf{v}_I + \mathbf{v}_T$  à la place de la règle de mise à jour directe issue du formalisme bidirectionnel :

$$\boldsymbol{\mu}_I^{n+1} = (\boldsymbol{\mu}_I^n \circ \mathbf{v}_I) \circ (-\mathbf{v}_T)^{-1} \quad (4.62)$$

De la même façon que pour l'approche SCO, nous revenons sur cette particularité à la section 4.4.

Afin de fournir un tour d'horizon complet, nous pouvons également introduire une approche bidirectionnelle compositionnelle que l'on nomme BC (Bidirectional Compositional) qui étend les approches compositionnelles unidirectionnelles classiques :

$$\boldsymbol{\mu}_I(\mathbf{v}) = \boldsymbol{\mu}_I^n \circ \mathbf{v}_I \quad \text{et} \quad \boldsymbol{\mu}_T(\mathbf{v}) = \boldsymbol{\mu}_T^n \circ \mathbf{v}_T^{-1}, \quad (4.63)$$

avec la règle de mise à jour :

$$\boldsymbol{\mu}_I^{n+1} = \boldsymbol{\mu}_I^n \circ \mathbf{v}_I \circ \mathbf{v}_T. \quad (4.64)$$

La matrice Jacobienne associée est obtenue par :

$$\mathbf{J}_e^{\text{BC}}(\mathbf{0}) = [\mathbf{J}_I^C(\mathbf{0}) \quad \mathbf{J}_T^C(\mathbf{0})], \quad (4.65)$$

ou encore une l'approche BA utilisant :

$$\boldsymbol{\mu}_I(\mathbf{v}) = \boldsymbol{\mu}_I^n + \mathbf{v}_I \quad \text{et} \quad \boldsymbol{\mu}_T(\mathbf{v}) = \boldsymbol{\mu}_T^n + \mathbf{v}_T, \quad (4.66)$$

avec la règle de mise à jour :

$$\boldsymbol{\mu}_I^{n+1} = (\boldsymbol{\mu}_I^n + \mathbf{v}_I) \circ (\boldsymbol{\mu}_T^n + \mathbf{v}_T)^{-1} \circ \boldsymbol{\mu}_T^{n-1}. \quad (4.67)$$

La matrice Jacobienne associée est :

$$\mathbf{J}_e^{\text{BA}}(\mathbf{0}) = [\mathbf{J}_I^A(\mathbf{0}) \quad -\mathbf{J}_T^A(\mathbf{0})]. \quad (4.68)$$

#### 4.3.3.1 Bilan

Le tableau 4.1 reprend les principales caractéristiques des paramétrisations incrémentales introduites dans cette partie. Le tableau 4.2 établit le lien entre ces paramétrisations incrémentales et les méthodes de l'état de l'art. L'introduction de la famille Bidirectionnelle (Subordonnée et Pure) permet d'étendre la classification de [Baker 2004] en intégrant des approches récentes (les approches symétriques et l'algorithme BCO). Nous avons également introduit des approches qui sont apparues naturellement au cours de la discussion.

C.	App.	$\mu_I$	$\mu_T$	$\mu_I^{n+1} = \mu_I \circ \mu_T^{-1} \circ \mu_T^n$	$\mathbf{J}_e(\mathbf{0})$	Sec
F	FC	$\mu_I^n \circ \mathbf{v}$	$\mu_T^n$	$\mu_I^n \circ \mathbf{v}$	$\mathbf{J}_T^C(\mathbf{0})$	4.2.1.2
	FCL	$\mu_I^n \circ \exp(\mathbf{v})$	$\mu_T^n$	$\mu_I^n \circ \exp(\mathbf{v})$	$\mathbf{J}_T^{CL}(\mathbf{0})$	4.2.1.2
	FA	$\mu_I^n + \mathbf{v}$	$\mu_T^n$	$\mu_I^n + \mathbf{v}$	$\mathbf{J}_T^A(\mathbf{0})$	4.2.1.1
I	ICR	$\mu_I^n$	$\mu_T^n \circ \mathbf{v}$	$\mu_I^n \circ \mathbf{v}^{-1}$	$-\mathbf{J}_T^C(\mathbf{0})$	4.2.2.1
	ICL*	$\mu_I^n$	$\mu_T^n \circ \exp(\mathbf{v})$	$\mu_I^n \circ \exp(-\mathbf{v})$	$-\mathbf{J}_T^{CL}(\mathbf{0})$	4.2.2.1
	ICD <sup>+</sup>	$\mu_I^n$	$\mu_T^n \circ \mathbf{v}^{-1}$	$\mu_I^n \circ \mathbf{v}$	$\mathbf{J}_T^C(\mathbf{0})$	4.2.2.1
	IAR <sup>+</sup>	$\mu_I^n$	$\mu_T^n + \mathbf{v}$	$\mu_I^n \circ (\mu_T^n + \mathbf{v})^{-1} \circ \mu_T^n$	$-\mathbf{J}_T^A(\mathbf{0})$	4.2.2.2
	IAD*	$\mu_I^n$	$\mu_T^n \circ (\mu_I^n + \mathbf{v})^{-1} \circ \mu_I^n$	$\mu_I^n + \mathbf{v}$	éq. (4.29)	4.2.2.3
BS	SC*	$\mu_I^n \circ (\frac{1}{2}\mathbf{v})$	$\mu_T^n \circ (\frac{1}{2}\mathbf{v})^{-1}$	$\mu_I^n \circ (\frac{1}{2}\mathbf{v}) \circ (\frac{1}{2}\mathbf{v})$	$\frac{1}{2}(\mathbf{J}_T^C(\mathbf{0}) + \mathbf{J}_T^C(\mathbf{0}))$	4.3.1.3
	SCL/ESM*	$\mu_I^n \circ \exp(\frac{1}{2}\mathbf{v})$	$\mu_T^n \circ \exp(-\frac{1}{2}\mathbf{v})$	$\mu_I^n \circ \exp(\mathbf{v})$	$\frac{1}{2}(\mathbf{J}_T^{CL}(\mathbf{0}) + \mathbf{J}_T^{CL}(\mathbf{0}))$	4.3.1.3
	FA+IAD*	$\mu_I^n + \frac{1}{2}\mathbf{v}$	$\mu_T^n \circ (\mu_I^n + \frac{1}{2}\mathbf{v})^{-1} \circ \mu_I^n$	$(\mu_I^n + \frac{1}{2}\mathbf{v}) \circ (\mu_I^n)^{-1} \circ (\mu_I^n + \frac{1}{2}\mathbf{v}) \approx \mu_I^n + \mathbf{v}$	$\frac{1}{2}(\mathbf{J}_T^{\text{FA}}(\mathbf{0}) + \mathbf{J}_T^{\text{IAD}}(\mathbf{0}))$	4.3.1.3
	SCO*	$\mu_I^n \circ (\frac{1}{2}\mathbf{v})$	$\mu_T^n \circ (-\frac{1}{2}\mathbf{v})$	$(\mu_I^n \circ (\frac{1}{2}\mathbf{v}) \circ (-\frac{1}{2}\mathbf{v})^{-1})$	$\frac{1}{2}(\mathbf{J}_T^C(\mathbf{0}) + \mathbf{J}_T^C(\mathbf{0}))$	4.3.2
	SAD	$\mu_I^n + \frac{1}{2}\mathbf{v}$	$\mu_T^n - \frac{1}{2}\mathbf{v}$	$(\mu_I^n + \frac{1}{2}\mathbf{v}) \circ (\mu_T^n - \frac{1}{2}\mathbf{v})^{-1} \circ \mu_T^n$	$\frac{1}{2}(\mathbf{J}_T^A(\mathbf{0}) + \mathbf{J}_T^A(\mathbf{0}))$	4.3.2
	ACL**	$\mu_I^n \circ \exp((1 - \alpha)\mathbf{v})$	$\mu_T^n \circ \exp(-\alpha\mathbf{v})$	$\mu_I^n \circ \exp(\mathbf{v})$	$(1 - \alpha)\mathbf{J}_T^{CL}(\mathbf{0}) + \alpha\mathbf{J}_T^{CL}(\mathbf{0})$	5.1
B	BC <sup>+</sup>	$\mu_I^n \circ \mathbf{v}_I$	$\mu_T^n \circ \mathbf{v}_T^{-1}$	$\mu_I^n \circ \mathbf{v}_I \circ \mathbf{v}_T$	$[\mathbf{J}_T^C(\mathbf{0}) \quad \mathbf{J}_T^C(\mathbf{0})]$	4.3.3
	BCO*	$\mu_I^n \circ \mathbf{v}_I$	$\mu_T^n \circ (-\mathbf{v}_T)$	$\mu_I^n \circ \mathbf{v}_I \circ (-\mathbf{v}_T)^{-1}$	$[\mathbf{J}_T^C(\mathbf{0}) \quad \mathbf{J}_T^C(\mathbf{0})]$	4.3.3
	BA <sup>+</sup>	$\mu_I^n + \mathbf{v}_I$	$\mu_T^n + \mathbf{v}_T$	$(\mu_I^n + \mathbf{v}_I) \circ (\mu_T^n + \mathbf{v}_T)^{-1} \circ \mu_T^n$	$[\mathbf{J}_T^A(\mathbf{0}) \quad -\mathbf{J}_T^A(\mathbf{0})]$	4.3.3
	BCL**	$\mu_I^n \circ \exp(\mathbf{v}_I)$	$\mu_T^n \circ \exp(-\mathbf{v}_T)$	$\mu_I^n \circ \mathbf{v}_I \circ \mathbf{v}_T$	$[\mathbf{J}_T^{CL}(\mathbf{0}) \quad \mathbf{J}_T^{CL}(\mathbf{0})]$	5.2

TABLE 4.1 – Taxonomie des paramétrisations incrémentales bidirectionnelles et règles de mise à jour directe correspondantes. La matrice Jacobienne  $\mathbf{J}_e$  correspondant à une optimisation Gauss-Newton de l'erreur SSD utilisant ces paramétrisations incrémentales est également fournie. Les différentes catégories ont été présentées à la section 3.3.1.1. Les conventions utilisées sont les suivantes : Première lettre : F=Forwards, I=Inverse, S=Symmetric, A=Asymmetric, B=Bi-directional; Second lettre : C=Compositional, A=Additive; Troisième lettre : R=Reverse, D=Direct, L=Lie, O=Opposite. Les approches suivies d'un \* correspondent à une nouvelle interprétation d'approches existantes. Les approches suivies d'un + correspondent à des formalisations simples apparues naturellement au cours de l'analyse. Les approches suivies de \*\* sont des nouvelles approches détaillées dans le chapitre 5.

C.	Méthode	P.I.	Optim.	Sect.
F	Forwards Additive [Lucas 1981]	FA	GN	4.2.1.1
	Forwards Compositional [Shum 2000]	FC	GN	4.2.1.2
	Basic Third-order Gradient Method [Keller 2008]	FC	O3	4.2.1.2
	LJEMANIFOLD [Guangwei 2008]	FCL	GN	4.2.1.2
I	Inverse Compositional [Baker 2004] [Matthews 2004b]	ICR	GN, N...	4.2.2.1
	Inverse Additive [Hager 1998], Generalized Inverse Compositional [Brooks 2009]	IAD	GN	4.2.2.3*
S	Efficient Second-order Minimization [Benlimane 2004][Vercauteren 2007]	SCL	GN	4.3.1.3*
	Symmetric Gradient Method [Keller 2004]	SCO	GN	4.3.2*
	Symmetric third-order Gradient Method [Keller 2008]	SCO	O3	4.3.2*
	Generalized ESM [Brooks 2009]	FA+IAD	GN	4.3.1.3*
B	Bi-directional Gradient Method [Keller 2004]	BCO	GN	4.3.3*

TABLE 4.2 – Classification des approches SSD de la littérature. Catégorie : cf. sections 4.2 et 4.3. Paramétrisation Incrementale P.I. : cf. Table 4.1. Optimization : GN=Gauss Newton, N=Newton, O3=Third Order. Les méthodes suivies d'un \* , sont des approches pour lesquelles le formalisme proposé donne un nouveau point de vue. Ceci est discuté dans les paragraphes correspondant à chaque approche.

## 4.4 Équivalence des algorithmes d'une même famille dans le cas d'une optimisation par l'algorithme Gauss-Newton

Dans cette partie, nous nous intéressons aux relations existant entre les différentes paramétrisations incrémentales lorsqu'on emploie l'algorithme d'optimisation Gauss-Newton avec une règle de mise à jour directe. Nous montrons notamment que le formalisme bidirectionnel permet d'étendre élégamment les résultats d'équivalence entre les approches proposés dans [Baker 2004].

### 4.4.1 Hypothèses

Rappelons que lorsque la matrice Jacobienne est de rang plein, l'algorithme Gauss-Newton conduit à l'expression :

$$\mathbf{v} = -\mathbf{J}_e(\mathbf{0})^+ \mathbf{e}(\mathbf{0}). \quad (4.69)$$

Nous faisons dans la suite les hypothèses suivantes :

- le groupe  $\mathcal{G}$  considéré est le même pour l'ensemble des approches considérées au sein d'une même classe (F), (I) ou (B).
- pour une paramétrisation incrémentale  $M$  quelconque, la matrice  $\mathbf{J}_e^M(\mathbf{0})$  associée est de rang plein (pas de problème d'ouverture).
- pour les approches unidirectionnelles, l'incrément  $\mathbf{v}$  introduit lors de la paramétrisation incrémentale appartient à un espace de même dimensionnalité  $p$  que le groupe  $\mathcal{G}$  et pour les approches bidirectionnelles pures,  $\mathbf{v}$  appartient à un espace de dimensionnalité  $2 \times p$  (c'est le cas pour l'ensemble des approches du tableau 4.2).

### 4.4.2 Fondements de l'équivalence

**Matrices Jacobiennes et paramétrisation incrémentale.** Pour une image quelconque  $L$  et une paramétrisation incrémentale  $\mu_L^M(\mu_L^n, \mathbf{v})$  associée à une approche quelconque  $M$  (FC, SCL par exemple...), nous avons :

$$\mathbf{J}_{L,i}^M(\mathbf{0}) = \left. \frac{\partial L(\mathbf{W}(\mu_L^M(\mu_L^n, \mathbf{v}), \mathbf{x}_i))}{\partial \mathbf{v}} \right|_{\mathbf{0}} = \left. \frac{\partial L(\mathbf{W}(\mu, \mathbf{x}_i))}{\partial \mu} \right|_{\mu_L^n} \mathcal{M}_L^{A \rightarrow M} = \mathbf{J}_{L,i}^A(\mathbf{0}) \mathcal{M}_L^{A \rightarrow M} \quad (4.70)$$

avec :

$$\mathcal{M}_L^{A \rightarrow M} = \left. \frac{\partial \mu_L^M(\mu_L^n, \mathbf{v})}{\partial \mathbf{v}} \right|_{\mathbf{0}} \quad (4.71)$$

$\mathcal{M}_L^{A \rightarrow M}$  est une matrice de dimension  $p \times p$  avec  $p$  la longueur de l'incrément  $\mathbf{v}$ , c'est à dire le nombre de dimensions de l'espace  $\mathcal{P}$ . L'expression (4.70) montre qu'il est possible d'exprimer la matrice Jacobienne associée à une paramétrisation incrémentale quelconque en fonction de la matrice Jacobienne associée à la paramétrisation incrémentale additive.

**Approximation des paramétrisations incrémentales** A l'ordre 1 autour de  $\mathbf{v} = \mathbf{0}$  nous avons :

$$\mu_I^M(\mu_I^n, \mathbf{v}) \approx \mu_I^n + \mathcal{M}_I^{A \rightarrow M} \mathbf{v} + O(\|\mathbf{v}\|^2) \quad (4.72)$$

$$\mu_T^M(\mu_T^n, \mathbf{v}) \approx \mu_T^n + \mathcal{M}_T^{A \rightarrow M} \mathbf{v} + O(\|\mathbf{v}\|^2) \quad (4.73)$$

## 4.4. Équivalence des algorithmes d'une même famille dans le cas d'une optimisation par l'algorithme Gauss-Newton

### 4.4.3 Équivalence des approches au sein d'une même famille

Dans cette partie, nous présentons les équations clés de la preuve d'équivalence à l'ordre 1 des approches au sein d'une même famille (Directe, Inverse, Bidirectionnelle Pure et Symétrique Compositionnelle Exacte). Le détail de la démonstration peut être consulté en annexe C.

**Approches Directes (cf. annexe C.1).** On désigne dans la suite par  $\mathbf{v}^F$  l'incrément estimé par une méthode directe F quelconque. Nous avons, en utilisant (3.25), la règle de mise à jour suivante pour les approches directes :

$$\text{maj}^D(\boldsymbol{\mu}_I, \boldsymbol{\mu}_T) = (\boldsymbol{\mu}_I^F(\boldsymbol{\mu}_I^n, \mathbf{v}^F), \boldsymbol{\mu}_T^0). \quad (4.74)$$

Par ailleurs, nous montrons en annexe que :

$$\mathbf{v}^F = (\mathcal{M}_I^{A \rightarrow F})^{-1} \mathbf{v}^{\text{FA}}, \quad (4.75)$$

où  $\mathbf{v}^{\text{FA}}$  est l'incrément estimé par la méthode directe additive.

L'équation (4.75) montre que les approches directes utilisant des paramétrisations incrémentales différentes conduisent avec l'approche Gauss-Newton à l'estimation d'incrément  $\mathbf{v}^F$  différents. Cette différence est due à la matrice  $\mathcal{M}_I^{A \rightarrow F}$  qui varie pour chaque approche. Cependant, en utilisant l'approximation à l'ordre 1 de la règle de mise à jour directe (4.72), nous avons bien une équivalence à l'ordre 1 sur les paramètres  $\boldsymbol{\mu}_I$  estimés :

$$\boldsymbol{\mu}_I^{F, n+1} = \boldsymbol{\mu}_I^n + \mathbf{v}^{\text{FA}} + O(\|\mathbf{v}^F\|^2). \quad (4.76)$$

Cette analyse étend la démonstration d'équivalence des approches FA et FC proposée dans [Baker 2004] à l'ensemble des approches directes (F).

**Approches Inverses (cf. annexe C.2).** On désigne dans la suite par  $\mathbf{v}^I$  l'incrément estimé par une méthode Inverse I quelconque. En utilisant (3.25), la règle de mise à jour pour les approches inverses est la suivante :

$$\text{maj}^D(\boldsymbol{\mu}_I, \boldsymbol{\mu}_T) = (\boldsymbol{\mu}_I^n \circ \boldsymbol{\mu}_T^I(\boldsymbol{\mu}_T^0, \mathbf{v}^I)^{-1} \circ \boldsymbol{\mu}_T^0, \boldsymbol{\mu}_T^0). \quad (4.77)$$

Nous montrons en annexe que l'incrément d'une méthode inverse quelconque peut s'exprimer sous la forme :

$$\mathbf{v}^I = (\mathcal{M}_T^{A \rightarrow I})^{-1} \mathbf{v}^{\text{IAR}}, \quad (4.78)$$

où  $\mathbf{v}^{\text{IAR}}$  est l'incrément estimé par la méthode *Inverse Additive Reverse*. Cette équation montre comme précédemment que les approches inverses utilisant des paramétrisations incrémentales différentes conduisent avec l'approche Gauss-Newton à l'estimation d'incrément différents. Cette différence est due à la matrice  $\mathcal{M}_T^{A \rightarrow I}$  spécifique à chaque approche I. Cependant, en utilisant l'approximation à l'ordre 1 (4.73), nous avons :

$$\boldsymbol{\mu}_I^{I, n+1} = \boldsymbol{\mu}_I^n \circ (\boldsymbol{\mu}_T^0 + \mathbf{v}^{\text{IAR}})^{-1} \circ \boldsymbol{\mu}_T^0 + O(\|\mathbf{v}^I\|^2). \quad (4.79)$$

Cette analyse révèle que l'ensemble des approches inverses sont équivalentes à l'ordre 1 étendant la démonstration d'équivalence des approches IA et IC proposée dans [Baker 2004] à l'ensemble des approches inverses (I).



**Approches Bidirectionnelles Pures (cf. annexe C.3).** Nous désignons dans la suite par  $\mathbf{v}^B = [\mathbf{v}_I^B \quad \mathbf{v}_T^B]^T$  l'incrément estimé par une méthode Bidirectionnelle B quelconque. Cette fois, la règle de mise à jour est :

$$\text{maj}^D(\boldsymbol{\mu}_I, \boldsymbol{\mu}_T) = (\boldsymbol{\mu}_I^B(\boldsymbol{\mu}_I^n, \mathbf{v}_I^B) \circ \boldsymbol{\mu}_T^B(\boldsymbol{\mu}_T^0, \mathbf{v}_T^B)^{-1} \circ \boldsymbol{\mu}_T^0, \boldsymbol{\mu}_T^0). \quad (4.80)$$

De plus l'incrément  $\mathbf{v}^B$  d'une méthode bidirectionnelle quelconque peut s'exprimer en fonction de l'incrément estimé par la méthode Bidirectionnelle additive  $\mathbf{v}^{BA}$ . Il vient ainsi :

$$\mathbf{v}_I^B = (\mathcal{M}_I^{A \rightarrow B})^{-1} \mathbf{v}_I^{BA}, \text{ et} \quad (4.81)$$

$$\mathbf{v}_T^B = (\mathcal{M}_T^{A \rightarrow B})^{-1} \mathbf{v}_T^{BA}. \quad (4.82)$$

En utilisant  $\mathbf{v}_I^B$  et  $\mathbf{v}_T^B$  dans (4.72) et (4.73), nous montrons que les paramètres estimés quelle que soit la méthode Bidirectionnelle B choisie vérifient :

$$\begin{aligned} \boldsymbol{\mu}_I^{B,n+1} &= \boldsymbol{\mu}_I^B(\boldsymbol{\mu}_I^n, \mathbf{v}_I^B) \circ \boldsymbol{\mu}_T^B(\boldsymbol{\mu}_T^0, \mathbf{v}_T^B)^{-1} \circ \boldsymbol{\mu}_T^0 \\ &= (\boldsymbol{\mu}_I^n + \mathbf{v}_I^{BA}) \circ (\boldsymbol{\mu}_T^0 + \mathbf{v}_T^{BA})^{-1} \circ \boldsymbol{\mu}_T^0 + O(\|\mathbf{v}_I^B\|^2 + \|\mathbf{v}_T^B\|^2). \end{aligned} \quad (4.83)$$

L'équation (4.83) révèle que l'ensemble des approches purement bidirectionnelles sont équivalentes à l'ordre 1.

**Approches Symétriques Compositionnelles Exactes (cf. annexe C.4).** Dans la suite, nous notons  $\mathbf{v}^{SCE}$  l'incrément estimé par une méthode symétrique compositionnelle exacte quelconque et  $\mathbf{J}_e^{SCE}(\mathbf{0})$  la matrice Jacobienne associée. La règle de mise à jour d'une approche symétrique est :

$$\text{maj}^D(\boldsymbol{\mu}_I, \boldsymbol{\mu}_T) = (\boldsymbol{\mu}_I^{SCE}(\boldsymbol{\mu}_I^n, \mathbf{v}^{SCE}) \circ \boldsymbol{\mu}_T^{SCE}(\boldsymbol{\mu}_T^0, \mathbf{v}^{SCE})^{-1} \circ \boldsymbol{\mu}_T^0, \boldsymbol{\mu}_T^0). \quad (4.84)$$

Nous montrons en annexe d'une part que :

$$\mathbf{v}^{SCE} = -(\mathcal{M}_I^{A \rightarrow SCE})^{-1} (\mathbf{J}^{SCE,1}(\mathbf{0}))^+ \mathbf{e}(\mathbf{0}), \quad (4.85)$$

avec  $\mathcal{M}_I^{A \rightarrow SCE}$  vérifiant (4.71) pour  $L = I$  et  $\mathbf{J}^{SCE,1}(\mathbf{0})$  satisfaisant :

$$\mathbf{J}^{SCE,1}(\mathbf{0}) = \frac{1}{2}(\mathbf{J}_I^A(\mathbf{0}) - \mathbf{J}_T^{\boldsymbol{\mu}_I^n, \boldsymbol{\mu}_T^0}(\mathbf{0})), \quad (4.86)$$

où  $\mathbf{J}_I^A(\mathbf{0})$  vérifie l'équation (4.5) et :

$$\mathbf{J}_T^{\boldsymbol{\mu}_I^n, \boldsymbol{\mu}_T^0}(\mathbf{0}) = \left. \frac{\partial T(\mathbf{W}(\boldsymbol{\mu}_T^0 \circ \boldsymbol{\mu}^{-1}, \mathbf{W}(\boldsymbol{\mu}_I^n, \mathbf{x}_i)))}{\partial \boldsymbol{\mu}} \right|_{\boldsymbol{\mu}_I^n}. \quad (4.87)$$

et d'autre part que l'incrément peut aussi s'exprimer sous la forme :

$$\mathbf{v}^{SCE} = -(\mathcal{M}_T^{A \rightarrow SCE})^{-1} (\mathbf{J}^{SCE,2}(\mathbf{0}))^+ \mathbf{e}(\mathbf{0}), \quad (4.88)$$

où  $\mathcal{M}_T^{A \rightarrow SCE}$  vérifie (4.71) pour  $L = I$  et  $\mathbf{J}^{SCE,2}(\mathbf{0})$  satisfait :

$$\mathbf{J}^{SCE,2}(\mathbf{0}) = \frac{1}{2}(\mathbf{J}_I^{\boldsymbol{\mu}_I^n, \boldsymbol{\mu}_T^0}(\mathbf{0}) - \mathbf{J}_T^A(\mathbf{0})), \quad (4.89)$$

#### 4.4. Équivalence des algorithmes d'une même famille dans le cas d'une optimisation par l'algorithme Gauss-Newton

avec  $\mathbf{J}_T^A(\mathbf{0})$  vérifiant l'équation (4.5) et  $\mathbf{J}_I^{\mu_I^n, \mu_T^0}(\mathbf{0})$  :

$$\mathbf{J}_I^{\mu_I^n, \mu_T^0}(\mathbf{0}) = \left. \frac{\partial I(\mathbf{W}(\mu_I^n \circ \mu^{-1}, \mathbf{W}(\mu_T^0, \mathbf{x}_i)))}{\partial \mu} \right|_{\mu_T^n}. \quad (4.90)$$

En insérant les deux expressions de  $\mathbf{v}^{SCE}$  dans la règle de mise à jour (4.84), il vient :

$$\mu_I^{SCE, n+1} = (\mu_I^n - (\mathbf{J}^{SCE, 1}(\mathbf{0}))^+ \mathbf{e}(\mathbf{0})) \circ (\mu_T^n - (\mathbf{J}^{SCE, 2}(\mathbf{0}))^+ \mathbf{e}(\mathbf{0}))^{-1} \circ \mu_T^0 + O(\|\mathbf{v}^{SCE}\|^2). \quad (4.91)$$

L'équation (4.91) révèle que l'ensemble des approches symétriques compositionnelles exactes sont équivalentes à l'ordre 1.

**Approches Bidirectionnelles Subordonnées** Les approches bidirectionnelles subordonnées ne peuvent pas toutes être démontrées équivalentes. Nous pouvons cependant effectuer ce travail pour le cas particulier de l'approche SCO, démontrée équivalente aux approches symétriques exactes à l'ordre 1 en annexe C.5.

**Validation expérimentale.** Afin de valider ces résultats théoriques d'équivalence, nous avons intégré les différentes approches du tableau 4.1 au sein du banc de test de [Baker 2004] dont le principe est rappelé dans l'annexe D. Le modèle de transformation utilisé est le modèle affine (2.6) pour lequel l'ensemble des paramétrisations incrémentales discutées sont possibles. Pour chaque algorithme, quinze itérations sont effectuées avant d'évaluer si oui ou non l'algorithme a convergé. La figure 4.2 présente les courbes de fréquence moyenne de convergence et de vitesse moyenne de convergence obtenues sur les images illustrées à la figure 6.1 pour les familles d'approches Directe, Inverse, Symétrique et Bidirectionnelle Pure. Aucun bruit n'a été ajouté aux images. Pour la fréquence de convergence, 500 réalisations différentes de l'initialisation de la déformation ( $\mu_I^0$ ) ont été testées pour un Sigma Point et pour une image puis les résultats ont été moyennés. Les courbes de vitesse de convergence sont obtenues en moyennant les résultats de 100 réalisations différentes de l'initialisation de la déformation par image.

D'après la figure 4.2, nous constatons que les courbes se superposent parfaitement au sein d'une même famille d'approches, validant ainsi la démonstration d'équivalence théorique. Dans [Baker 2004], les auteurs montrent l'équivalence à l'ordre 1 des approches FA et FC avec les approches IA et IC. Avec les résultats d'équivalence au sein d'une même classe que nous avons proposé, nous pouvons en déduire que l'ensemble des approches directes et inverses sont équivalentes à l'ordre 1, ce qui va dans le sens des résultats observés qui sont similaires pour ces deux familles. Les approches symétriques et bidirectionnelles produisent en revanche de meilleurs résultats, à la fois en fréquence de convergence et de vitesse de convergence, ce qui va dans le sens des observations faites pour les algorithmes particuliers SGM (renommé dans le formalisme SCO), BDGM (renommé dans le formalisme BCO) dans [Keller 2004], et pour l'algorithme ESM [Benhimane 2007b] (que nous interprétons comme une approche bidirectionnelle subordonnée symétrique SCL). Ces résultats semblent prometteurs pour la définition de nouveaux algorithmes dans cette famille.

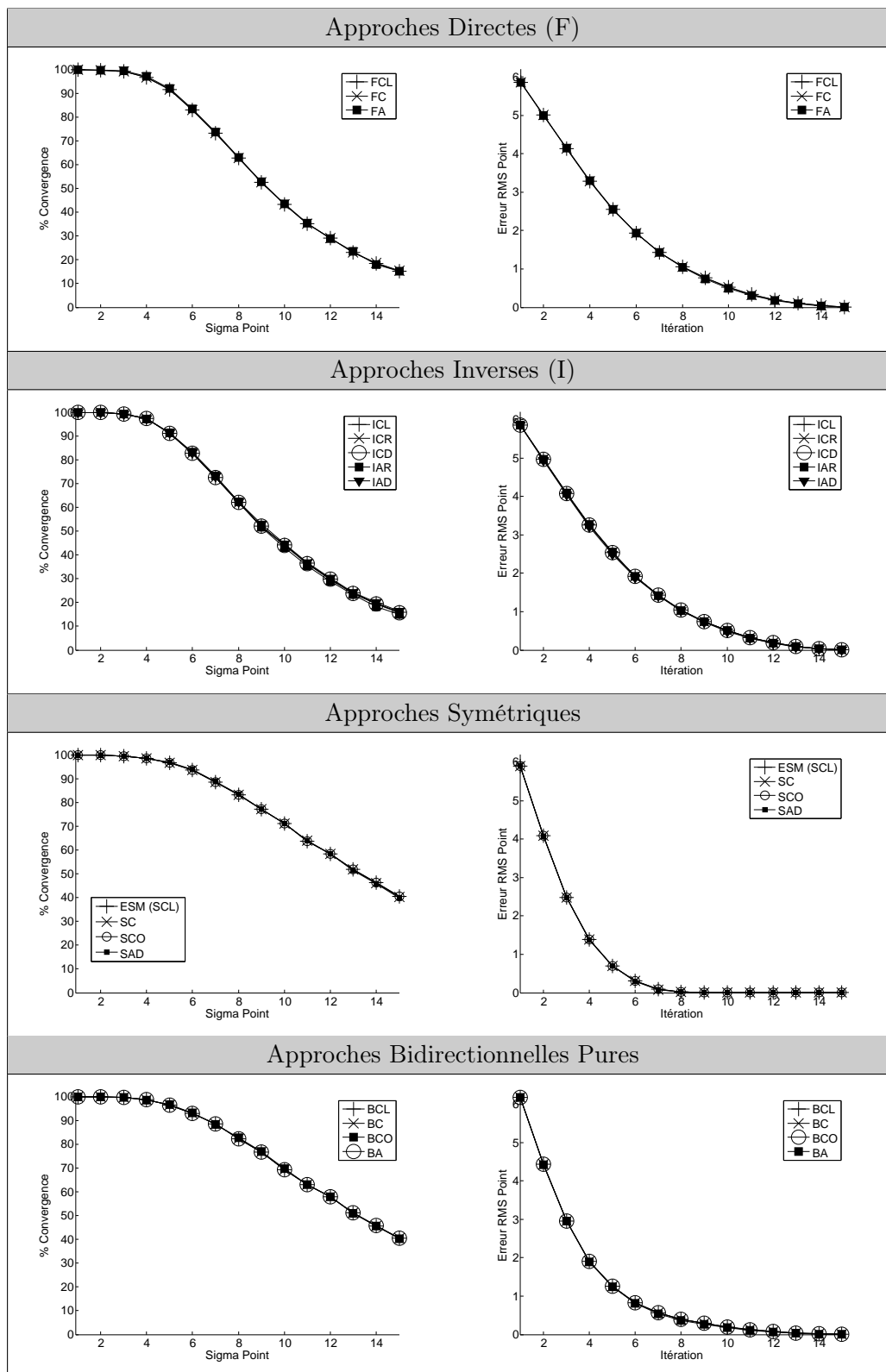


FIGURE 4.2 – Fréquence moyenne de convergence et vitesse moyenne de convergence pour les approches des catégories Directe, Inverse, Symétrique et Bidirectionnelle. Pour la vitesse de convergence Sigma Point = 6. Une règle de mise à jour directe est utilisée.

#### 4.4. Équivalence des algorithmes d'une même famille dans le cas d'une optimisation par l'algorithme Gauss-Newton

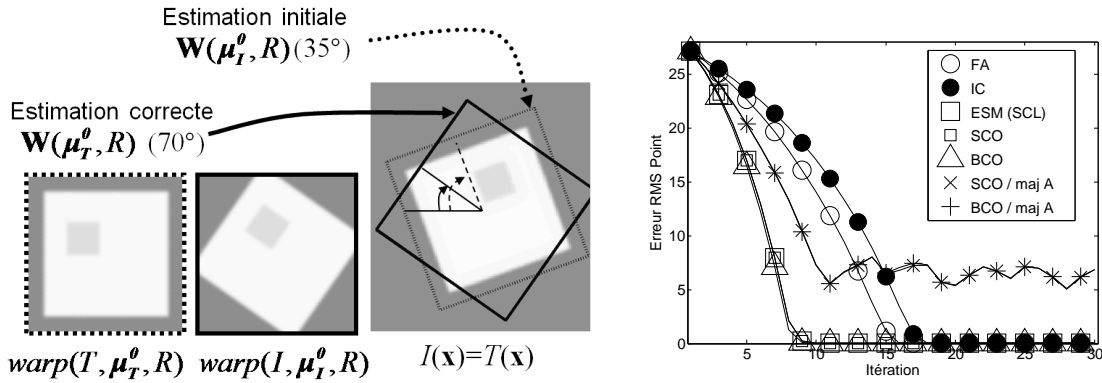


FIGURE 4.3 – Paramétrisation incrémentale et règle de mise à jour. L'équivalence entre deux approches quelconques M1 et M2 au sein d'une même famille ne signifie pas l'égalité de leurs incréments correctifs  $\mathbf{v}_{M1}$  et  $\mathbf{v}_{M2}$ . Ce contreexemple repose sur l'estimation d'un mouvement affine en utilisant le banc de test proposé dans [Baker 2004]. L'erreur est exprimé en pixels. Une rotation  $\mu_I^*$  de  $70^\circ$  autour du centre de l'objet doit être estimée, à partir d'une rotation initiale  $\mu_I^n$  de  $35^\circ$ . Toutes les approches utilisent l'algorithme Gauss-Newton. Les approches hybrides utilisant des paramétrisations incrémentales SCO ou BCO combinées avec une règle de mise à jour additive introduites par [Keller 2004] convergent plus rapidement que des approches des familles (F) et (I) quand  $\mu_I^n$  est suffisamment proche de l'optimum, mais ne peuvent pas converger vers l'estimé correct en raison de l'inadéquation entre la règle de mise à jour utilisée et les paramétrisations incrémentales choisies. L'utilisation de la règle de mise à jour directe issue naturellement du formalisme corrige ce problème, permettant l'obtention de meilleurs résultats tout en conservant les mêmes calculs au niveau de l'estimation de l'incrément  $\mathbf{v}$ .

#### 4.4.4 Contraintes sur l'espace bidirectionnel et règle de mise à jour

Considérons deux approches M1 et M2 quelconques mais appartenant à la même famille parmi les classes (F), (I) ou (B). L'analyse issue de la section 4.4.3 révèle que les estimés finaux à savoir les paramètres  $\mu_I^{n+1}$  produit par ces deux approches sont approximativement identiques. Cependant les incréments associés ayant permis cette estimation  $\mathbf{v}_1$  et  $\mathbf{v}_2$  ne sont pas égaux dans le cas général, en raison de la présence des matrices  $\mathcal{M}_I^{A \rightarrow M}$  et  $(\mathcal{M}_I^{A \rightarrow M})$ . Par conséquent *la règle de mise à jour doit toujours être en accord avec les paramétrisations incrémentales choisies.*

Lorsque cela n'est pas le cas, un estimé correct  $\mathbf{v}$  peut conduire à un estimé incorrect  $\mu_I^{n+1}$ . Cela introduit des problèmes de convergence, surtout lorsque des rotations entrent en jeu. Ce phénomène peut être observé dans certaines expériences de [Keller 2004] et [Keller 2008]. Nous illustrons ce phénomène sur la figure 4.3, où l'erreur correspondant à un mouvement affine oscille sans converger autour de l'estimé correct quand  $\mu_I^n$  correspond à une forte rotation. Le formalisme bidirectionnel propose une méthodologie systématique afin d'éviter de tels problèmes.

## 4.5 Conclusions

### 4.5.1 Bilan

Dans ce chapitre, nous avons utilisé le formalisme bidirectionnel pour unifier une grande variété de méthodes itératives existantes au sein d'un même cadre théorique. Cet effort de formalisation a conduit à un nombre important de résultats :

- Au sein d'une même famille d'algorithmes (parmi les classes directe (F), Inverse (I), Symétrique Compositionnelle Exacte(SCE) ou bidirectionnel pure (B)), les approches ont été démontrées équivalentes à l'ordre 1 (section 4.4.3), ce qui justifie une certaine liberté pour choisir la paramétrisation incrémentale.
- L'approche IA de [Hager 1998] et l'approche ESM proposée par [Benhimane 2007b] peuvent être réinterprétées à l'aide d'une paramétrisation incrémentale qui s'affranchit des hypothèses d'égalité des images à une compensation près.
- Il est possible d'associer aux approches GIC et GESM [Brooks 2009] une paramétrisation incrémentale explicite. Nous montrons de cette manière que les approches GIC [Brooks 2009] et IA calculent théoriquement le même résultat et que l'approche GESM utilise une règle de mise à jour directe approximée à l'ordre 1 sur l'incrément.
- La définition d'une approche à l'aide de paramétrisations incrémentales permet d'assurer l'utilisation d'une règle de mise à jour adaptée. Nous montrons notamment que les approches SCO et BCO [Keller 2004] utilisent une règle de mise à jour inconsistante relativement aux paramétrisations incrémentales ce qui nuit à leurs propriétés de convergence.
- D'un point de vue expérimental, les approches (SCE) et (B) (cf section 4.4.3) semblent apporter un gain en termes de fréquence et de vitesse de convergence.

Il existe un certain nombre d'approches, notamment les méthodes combinant plusieurs fonctionnelles unidirectionnelles qui n'ont pas été intégrées au formalisme. Afin de ne pas alourdir le discours, nous détaillons le lien entre ces approches et le formalisme en annexe E.

Pour conclure ce chapitre, nous effectuons une synthèse des résultats précédents dans le contexte particulier du suivi d'objet.

### 4.5.2 Conception d'approches pour le suivi d'objet

Le problème de suivi d'objet paramétrique repose sur la délimitation de l'objet d'intérêt dans une image de référence (le template  $T$ ). Ce problème peut apparaître a priori comme un problème très asymétrique. L'objet d'intérêt étant fixé dans le template, il paraît naturel de chercher à aligner les images courantes (symbolisées par  $I$ ) issues de la vidéo sur ce template, ce qui joue en faveur de l'utilisation de méthodes de la famille directe (F) (cf. section 4.2.1). C'est d'ailleurs des fonctionnelles d'erreur reposant sur une compensation directe qui sont présentées dans les travaux de [Hager 1998][Benhimane 2007b] et qui demeurent des références pour ce type d'applications.

Concernant l'efficacité calculatoire, les approches inverses semblent plus adaptées puisqu'elles permettent un précalcul (comme la pseudo-inverse de la matrice Jacobienne associée à l'erreur) d'informations sur le template  $T$ . Cependant, en raison des bonnes performances en termes de

## 4.5. Conclusions

---

vitesses de convergence et de fréquence de convergence des approches (B) ou (SCE), les approches utilisant les gradients des deux images semblent être une alternative intéressante aux approches (I). Jusqu'à présent (cf. section 2.3.2), la plupart des approches considèrent intrinsèquement une symétrie dans le problème d'alignement (approches utilisant une paramétrisation incrémentale symétrique) ou une asymétrie totale (approches unidirectionnelles). En présence d'images de qualités différentes la question de l'utilisation au mieux des gradients de chaque image demeure ouverte.

Les approches bidirectionnelles font apparaître naturellement les gradients des deux images lors de l'utilisation de l'algorithme Gauss-Newton. De plus, à travers les sous-classes Bidirectionnelle Subordonnée (BS) et Bidirectionnelle Pure (B), il semble possible de définir des approches permettant un compromis dans l'utilisation des gradients de chaque image. Il paraît donc intéressant de concevoir des approches bidirectionnelles pour le suivi d'objet. Avec ces approches, une règle de mise à Directe doit être utilisée pour contraindre l'algorithme à utiliser, dans le template, les informations visuelles centrées sur la région d'intérêt. Cette règle de mise à jour a été utilisée avec succès dans le cas particulier de l'approche inverse compositionnelle dans les travaux de [Baker 2004]. Les relations fortes entre l'algorithme IA [Hager 1998] et l'optimisation Gauss-Newton de l'approche IAD (cf. section 4.2.2.3) ainsi qu'entre l'algorithme ESM [Benhimane 2007b] et l'optimisation Gauss-Newton de l'approche SCL (cf. section 4.3.1.3) tendent à confirmer la possibilité de mettre en oeuvre le formalisme bidirectionnel avec mise à jour directe pour le suivi d'objet (les approches ESM et IA ayant été conçues pour le suivi d'objet).

L'utilisation d'une paramétrisation incrémentale compositionnelle associe une interprétation claire en termes de composition de transformations (section 3.3.2) et conduisent à des approches moins coûteuses en termes de complexité calculatoire [Shum 2000, Baker 2004]. Etant donné l'équivalence à l'ordre 1 des différentes approches de chaque famille, les approches par composition paraissent donc plus adaptées, notamment dans le contexte du suivi d'objets rigide où le temps de calcul est un facteur limitant et où les transformations mises en jeu sont paramétrables sur des groupes non nécessairement additifs (cf. section 2.1.2). Parmi les approches compositionnelles, celles utilisant une paramétrisation sur groupe de Lie permet la mise en place d'algorithmes d'optimisation plus simples notamment dans le cas d'homographie (cf. section 2.1.2.3), c'est pourquoi nous favorisons cette dernière dans la suite.

Dans la partie suivante, nous proposons deux nouvelles méthodes qui tirent partie des résultats mis en évidence par l'analyse précédente afin d'améliorer les algorithmes de suivi en présence d'images bruitées.



# Composition sur groupe de Lie pour l'alignement d'images de qualités différentes

---

## Sommaire

<b>5.1</b>	<b>Composition asymétrique sur groupe de Lie</b>	<b>75</b>
5.1.1	Analyse théorique de l'erreur ACL	77
5.1.2	Minimisation de variance pour l'estimation de $\alpha$	79
5.1.3	ACL Géométrique pour l'estimation de $\alpha$	80
5.1.4	ACL Analytique pour l'estimation de $\alpha$ (AACL)	81
<b>5.2</b>	<b>Composition Bidirectionnelle sur groupe de Lie</b>	<b>83</b>
5.2.1	Propriété d'invariance	83
5.2.2	L'algorithme BCL projeté (PBCL)	84
5.2.3	Analyse théorique de l'influence du bruit	85
<b>5.3</b>	<b>Résumé des nouvelles approches et approximations efficaces</b>	<b>86</b>

---

Dans ce chapitre, nous proposons deux nouvelles paramétrisations incrémentales permettant une amélioration de performance des approches de recalage en présence d'images dégradées par du bruit. Les deux nouvelles approches utilisent une paramétrisation sur groupe de Lie afin de permettre une procédure d'optimisation aisée et de faciliter les dérivations théoriques. Afin d'alléger les notations, nous noterons  $\mathbf{J}_I$  et  $\mathbf{J}_T$  les matrices jacobiniennes notées précédemment  $\mathbf{J}_I^{CL}$  et  $\mathbf{J}_T^{CL}$  et obtenues en concaténant verticalement les contributions de chaque pixel :

$$\mathbf{J}_{I,i}(\mathbf{w}) = \left. \frac{\partial I(\mathbf{W}(\boldsymbol{\mu}_I^n \circ \exp(\mathbf{v}), \mathbf{x}_i))}{\partial \mathbf{v}} \right|_{\mathbf{v}=\mathbf{w}}, \quad (5.1)$$

$$\mathbf{J}_{T,i}(\mathbf{w}) = \left. \frac{\partial T(\mathbf{W}(\boldsymbol{\mu}_T^n \circ \exp(\mathbf{v}), \mathbf{x}_i))}{\partial \mathbf{v}} \right|_{\mathbf{v}=\mathbf{w}}. \quad (5.2)$$

## 5.1 Composition asymétrique sur groupe de Lie

Dans cette partie, nous proposons une nouvelle approche utilisant la reparamétrisation incrémentale suivante :

$$\boldsymbol{\mu}_I = \boldsymbol{\mu}_I^n \circ \exp((1 - \alpha)\mathbf{v}), \quad (5.3)$$

$$\boldsymbol{\mu}_T = \boldsymbol{\mu}_T^n \circ \exp(-\alpha\mathbf{v}). \quad (5.4)$$



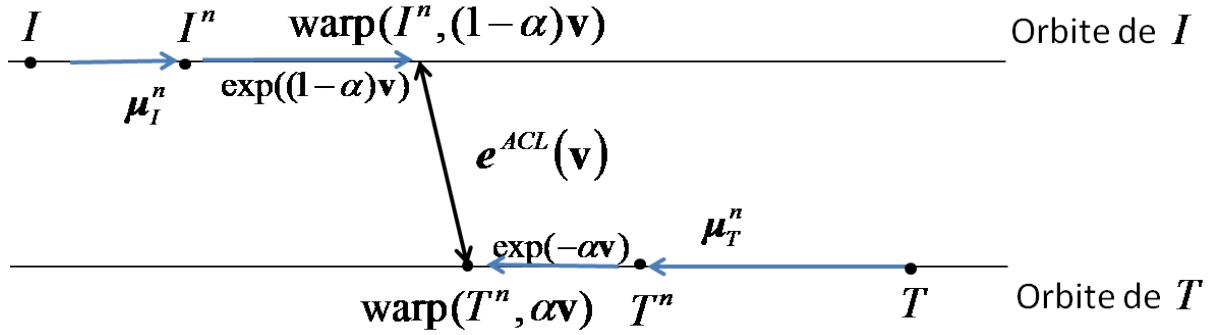


FIGURE 5.1 – Interprétation géométrique de la formulation asymétrique. Le positionnement du repère de compensation relativement aux repères des images  $I^n$  et  $T^n$  est contraint par un paramètre  $\alpha$  entre 0 et 1.

Dans le cadre d’une minimisation SSD, nous obtenons le critère asymétrique :

$$E_\alpha(\mathbf{v}) = \sum_{i=1}^N |e_{\alpha,i}(\mathbf{v})|^2, \quad (5.5)$$

où  $e_{\alpha,i}$  représente l’erreur pour le pixel  $\mathbf{x}_i$  appartenant à la région d’intérêt  $R = (\mathbf{x}_1, \dots, \mathbf{x}_N)$  :

$$e_{\alpha,i}(\mathbf{v}) = I(\mathbf{W}(\mu_I^n \circ \exp((1-\alpha)\mathbf{v}), \mathbf{x}_i)) - T(\mathbf{W}(\mu_T^n \circ \exp(-\alpha\mathbf{v}), \mathbf{x}_i)). \quad (5.6)$$

La particularité de l’équation (5.6) est d’introduire un coefficient de pondération  $\alpha \in [0, 1]$  correspondant à une contrainte asymétrique imposée sur le processus d’alignement.

Le paramètre  $\alpha$  nous permet de formuler un certain nombre d’approches existantes de manière plus générique : les cas particuliers  $\alpha = 0$ ,  $\alpha = 1$  et  $\alpha = 0.5$  correspondent respectivement aux approches FCL, ICL et SCL discutées aux sections 4.2 et 4.3. Rappelons que l’approche SCL est une formulation alternative de l’approche ESM comme cela a été discuté à la section 4.3.1.1. Ainsi la fonctionnelle asymétrique permet d’étendre les approches de l’état de l’art à n’importe quel  $\alpha \in [0, 1]$ . Nous appelons dans la suite  $\alpha$ -ACL (*Asymmetric Compositional Lie*) l’approche asymétrique pour laquelle  $\alpha$  est fixé a priori d’une manière arbitraire.

La figure 5.1 illustre le principe de l’approche ACL. L’approche ACL fait partie de la famille bidirectionnelle subordonnée car les deux images  $I$  et  $T$  (cf. équation (5.3) et (5.4)) sont déformées par l’utilisation du même incrément  $\mathbf{v} \in \mathcal{P}$ .

Aligner deux images revient donc à optimiser localement l’erreur définie par les équations (5.5) et (5.6) à l’étape  $n$  en fonction de  $\mathbf{v}$ . Une fois l’estimé  $\hat{\mathbf{v}}_\alpha$  calculé, une règle de mise à jour directe fournit un nouvel estimé des paramètres de mouvement :

$$\mu_I^{n+1} = \mu_I^n \circ \exp(\hat{\mathbf{v}}_\alpha). \quad (5.7)$$

## 5.1. Composition asymétrique sur groupe de Lie

Dans la suite, nous notons  $\mathbf{e}_\alpha = (e_{\alpha,i})_{i=1..N}$  le vecteur obtenu en concaténant les erreurs  $e_{\alpha,i}$ . L'utilisation de l'algorithme de Gauss-Newton sur (5.5) conduit à :

$$\hat{\mathbf{v}}_\alpha = -(\mathbf{J}_\alpha(\mathbf{0}))^+ \mathbf{e}_\alpha(\mathbf{0}) = -(\mathbf{J}_\alpha(\mathbf{0})^t \mathbf{J}_\alpha(\mathbf{0}))^{-1} \mathbf{J}_\alpha(\mathbf{0})^t \mathbf{e}_\alpha(\mathbf{0}). \quad (5.8)$$

La matrice Jacobienne vérifie :

$$\mathbf{J}_\alpha(\mathbf{w}) = \left. \frac{\partial \mathbf{e}_\alpha(\mathbf{v})}{\partial \mathbf{v}} \right|_{\mathbf{v}=\mathbf{w}}. \quad (5.9)$$

Elle peut être obtenue par la concaténation des gradients  $\mathbf{J}_{\alpha,i}(\mathbf{w})$  des erreurs pixel à pixel  $e_{\alpha,i}$  :  $\mathbf{J}_\alpha(\mathbf{w}) = [\mathbf{J}_{\alpha,1}(\mathbf{w})^t, \dots, \mathbf{J}_{\alpha,N}(\mathbf{w})^t]^t$  avec

$$\mathbf{J}_{\alpha,i}(\mathbf{w}) = (1 - \alpha) \mathbf{J}_{I,i}((1 - \alpha)\mathbf{w}) + \alpha \mathbf{J}_{T,i}(-\alpha\mathbf{w}), \quad (5.10)$$

où  $\mathbf{J}_{I,i}$  and  $\mathbf{J}_{T,i}$  sont les matrices Jacobiennes définies par les équations (5.1) et (5.2). Dans la suite, nous étudions les propriétés théoriques découlant de la paramétrisation incrémentale ACL.

### 5.1.1 Analyse théorique de l'erreur ACL

La plupart du temps, le processus d'alignement est formalisé comme l'optimisation d'une erreur unidirectionnelle directe ( $\alpha = 0$  dans (5.6)) comme dans [Shum 2000][Benhimane 2007b] ou une erreur inverse ( $\alpha = 1$  dans (5.6)) comme dans [Baker 2004]. Nous pensons qu'en présence de bruit fort dans les images, l'utilisation de l'une et/ou de l'autre des images est à préférer en fonction des niveaux respectifs de bruit. Nous montrons à présent que la formulation générique de la fonctionnelle ACL nous permet d'analyser les effets du bruit de façon unifiée. Considérant successivement les cas non bruité et bruité, un développement de Taylor d'ordre 2 est appliqué sur l'erreur  $\mathbf{e}_\alpha(\mathbf{v})$ . Cela nous permet de mettre en évidence à la section 5.1.1.2 le compromis à faire dans la minimisation à la fois des termes d'ordre 2 et de la variance des termes générés par le bruit.

Nous considérons deux images dégradées par des bruits additifs indépendants caractérisés par des variances respectives  $\sigma_I^2$  et  $\sigma_T^2$  :

$$I = I_{nf} + \varepsilon_I \quad T = T_{nf} + \varepsilon_T \quad (5.11)$$

où l'indice  $nf$  fait référence aux termes idéaux non bruités (*noise-free*). Cela conduit à l'expression des matrices Jacobiennes :

$$\mathbf{J}_I(\mathbf{w}) = \mathbf{J}_I^{nf}(\mathbf{w}) + \mathbf{J}_{\varepsilon_I}(\mathbf{w}) \quad \text{et} \quad \mathbf{J}_T(\mathbf{w}) = \mathbf{J}_T^{nf}(\mathbf{w}) + \mathbf{J}_{\varepsilon_T}(\mathbf{w}), \quad (5.12)$$

avec  $\mathbf{J}_{\varepsilon_I}(\mathbf{w})$  et  $\mathbf{J}_{\varepsilon_T}(\mathbf{w})$  obtenues en concaténant :

$$\mathbf{J}_{\varepsilon_I,i}(\mathbf{w}) = \left. \frac{\partial \varepsilon_I(\mathbf{W}(\boldsymbol{\mu}_I^n \circ \exp(\mathbf{v}), \mathbf{x}_i))}{\partial \mathbf{v}} \right|_{\mathbf{v}=\mathbf{w}}, \quad \text{et} \quad (5.13)$$

$$\mathbf{J}_{\varepsilon_T,i}(\mathbf{w}) = \left. \frac{\partial \varepsilon_T(\mathbf{W}(\boldsymbol{\mu}_T^n \circ \exp(\mathbf{v}), \mathbf{x}_i))}{\partial \mathbf{v}} \right|_{\mathbf{v}=\mathbf{w}}. \quad (5.14)$$

### 5.1.1.1 Cas non bruité

Dans cette partie, nous considérons des images non bruitées identiques à une déformation géométrique près caractérisée par le vecteur de paramètres vrais  $\mathbf{v}^*$  :

$$\forall \mathbf{x} \quad I_{nf}(\mathbf{W}(\boldsymbol{\mu}_I^n \circ \exp((1 - \alpha)\mathbf{v}^*)), \mathbf{x}) = T_{nf}(\mathbf{W}(\boldsymbol{\mu}_T^n \circ \exp(-\alpha\mathbf{v}^*)), \mathbf{x}), \quad (5.15)$$

cela induit :

$$\mathbf{e}_\alpha^{nf}(\mathbf{v}^*) = \mathbf{0}. \quad (5.16)$$

Sous cette hypothèse [Benhimane 2007b] a proposé initialement la procédure ESM consistant à utiliser le Jacobien  $\mathbf{J}_{\text{ESM}}$  à la place de  $\mathbf{J}_I(\mathbf{0})$  dans l'équation (5.8) comme une approximation d'ordre 2 de l'erreur directe compositionnelle ( $\alpha = 0$  dans (5.6)) avec :

$$\mathbf{J}_{\text{ESM}} = \frac{1}{2}(\mathbf{J}_I(\mathbf{0}) + \mathbf{J}_T(\mathbf{0})). \quad (5.17)$$

La preuve avancée reposait sur (5.15).

Reposant sur la même hypothèse, le théorème 1 généralise ce résultat en montrant que l'optimisation ESM induit en réalité une approximation à l'ordre 2 de n'importe quelle erreur ACL.

**Theorème 1** (Développement à l'ordre 2 dans des conditions non bruitées aux alentours de  $\mathbf{v}^*$ ).  
Pour tout  $\alpha$  :

$$\mathbf{e}_\alpha^{nf}(\mathbf{v}^*) = \mathbf{e}_\alpha^{nf}(\mathbf{0}) + \mathbf{J}_{\text{ESM}}^{nf} \mathbf{v}^* + O(\|\mathbf{v}^*\|^3) \quad (5.18)$$

avec

$$\mathbf{J}_{\text{ESM}}^{nf} = \frac{1}{2}(\mathbf{J}_I^{nf}(\mathbf{0}) + \mathbf{J}_T^{nf}(\mathbf{0})). \quad (5.19)$$

*Démonstration.* Nous utilisons ici les lemmes présentés à la section F. D'après l'équation (F.2), nous avons à l'ordre 2 :

$$\mathbf{e}_\alpha^{nf}(\mathbf{v}) = \mathbf{e}_\alpha^{nf}(\mathbf{0}) + \frac{1}{2}(\mathbf{J}_\alpha^{nf}(\mathbf{0}) + \mathbf{J}_\alpha^{nf}(\mathbf{v}))\mathbf{v} + O(\|\mathbf{v}\|^3). \quad (5.20)$$

De plus, pour les paramètres vrais  $\mathbf{v}^*$ , l'utilisation de l'équation (F.13) implique :

$$(\mathbf{J}_\alpha^{nf}(\mathbf{0}) + \mathbf{J}_\alpha^{nf}(\mathbf{v}^*))\mathbf{v}^* = (\mathbf{J}_I^{nf}(\mathbf{0}) + \mathbf{J}_T^{nf}(\mathbf{0}))\mathbf{v}^* + O(\|\mathbf{v}^*\|^3) \quad (5.21)$$

L'utilisation de (5.21) dans (5.20) amène vers (5.18).  $\square$

Remarquons que  $\mathbf{e}_\alpha(\mathbf{0})$  est indépendante de  $\alpha$ .

### 5.1.1.2 En présence de bruit

L'équation (5.21) n'est pas valable si l'on considère le Jacobien bruité  $\mathbf{J}_\alpha$  au lieu de  $\mathbf{J}_\alpha^{nf}$ . L'expression utilisée dans (5.21) constitue seulement la partie non bruitée du jacobien. Or durant la procédure d'optimisation, nous avons accès uniquement aux Jacobiens des images bruitées  $\mathbf{J}_I(\mathbf{0})$  et  $\mathbf{J}_T(\mathbf{0})$ . Nous cherchons donc la valeur de  $\alpha$  qui produit un Jacobien  $\mathbf{J}_\alpha(\mathbf{0})$  le plus proche possible de  $\mathbf{J}_{\text{ESM}}^{nf}$ .

## 5.1. Composition asymétrique sur groupe de Lie

**Theorème 2** (Différence entre  $\mathbf{J}_\alpha(\mathbf{0})$  et la matrice Jacobienne idéale  $\mathbf{J}_{\text{ESM}}^{nf}$ ).

$$\mathbf{J}_{\text{ESM}}^{nf} = \mathbf{J}_\alpha(\mathbf{0}) - (\mathbf{A}_\alpha + \mathbf{B}_\alpha), \quad (5.22)$$

avec :

$$\begin{aligned} \mathbf{A}_\alpha &= \left(\frac{1}{2} - \alpha\right)\mathbf{J}_{\ominus nf}(\mathbf{0}), \\ \mathbf{B}_\alpha &= (1 - \alpha)\mathbf{J}_{\varepsilon_I}(\mathbf{0}) + \alpha\mathbf{J}_{\varepsilon_T}(\mathbf{0}), \end{aligned} \quad (5.23)$$

où :

$$\mathbf{J}_{\ominus nf}(\mathbf{0}) = \mathbf{J}_I^{nf}(\mathbf{0}) - \mathbf{J}_T^{nf}(\mathbf{0}). \quad (5.24)$$

et  $\mathbf{J}_{\varepsilon_T}(\mathbf{0})$  et  $\mathbf{J}_{\varepsilon_I}(\mathbf{0})$  vérifient respectivement (5.14) et (5.13) avec  $\mathbf{w} = \mathbf{0}$ .

*Démonstration.*

$$\begin{aligned} \mathbf{J}_{\text{ESM}}^{nf} &= \mathbf{J}_\alpha^{nf}(\mathbf{0}) - \left(\frac{1}{2} - \alpha\right)\mathbf{J}_{\ominus nf}(\mathbf{0}) \\ &= \mathbf{J}_\alpha(\mathbf{0}) - \left(\frac{1}{2} - \alpha\right)\mathbf{J}_{\ominus nf}(\mathbf{0}) - ((1 - \alpha)\mathbf{J}_{\varepsilon_I}(\mathbf{0}) + \alpha\mathbf{J}_{\varepsilon_T}(\mathbf{0})). \end{aligned} \quad (5.25)$$

□

Le théorème 2 met en évidence la présence de deux sortes de termes perturbateurs  $\mathbf{A}_\alpha \mathbf{v}^*$  et  $\mathbf{B}_\alpha \mathbf{v}^*$  quand  $\mathbf{J}_\alpha(\mathbf{0})$  est utilisé à la place du Jacobien ESM non bruité.

**Termes de structure.** Le terme  $\mathbf{A}_\alpha \mathbf{v}^*$  correspond aux termes d'ordre 2 de l'erreur non bruitée  $\mathbf{e}_\alpha^{nf}(\mathbf{v}^*)$  à l'optimum :

$$\mathbf{J}_{\ominus nf}(\mathbf{0})\mathbf{v}^* \stackrel{\text{(F.8)}}{=} (\mathbf{J}_I^{nf}(\mathbf{0}) - \mathbf{J}_I^{nf}(\mathbf{v}^*))\mathbf{v}^* \stackrel{\text{(F.4)}}{=} O(\|\mathbf{v}^*\|^2). \quad (5.26)$$

Ces termes d'ordre 2 sont justement ceux qui sont éliminés par l'algorithme ESM dans le cas non bruité ce qui permet une amélioration de la convergence.

**Termes de bruit.** Le terme  $\mathbf{B}_\alpha$  correspond à la contribution d'ordre 1 du bruit. En supposant que les bruits sur  $I$  and  $T$  sont indépendants et ont des propriétés de gradient similaires, la variance totale de  $\mathbf{B}_\alpha \mathbf{v}^*$  est proportionnelle à :

$$\text{var}(\mathbf{B}_\alpha \mathbf{v}^*) \propto (1 - \alpha)^2 \sigma_I^2 + \alpha^2 \sigma_T^2 \quad (5.27)$$

### 5.1.2 Minimisation de variance pour l'estimation de $\alpha$

La minimisation de  $\text{var}(\mathbf{B}_\alpha \mathbf{v}^*)$  (5.27) en fonction de  $\alpha$  définit un nouvel algorithme ACL à variance de bruit minimale que l'on note MVACL :

$$\alpha^{\text{MVACL}} = \frac{\sigma_I^2}{\sigma_I^2 + \sigma_T^2}. \quad (5.28)$$

Cette valeur de  $\alpha$  a les propriétés suivantes : pour une quantité de bruit répartie de manière symétrique on obtient le comportement de l'ESM ( $\sigma_I = \sigma_T$ ). Lorsque le bruit affecte uniquement une image on obtient le comportement unidirectionnel du FCL (resp. de l'ICL) quand  $\sigma_I \ll \sigma_T$  (resp.  $\sigma_I \gg \sigma_T$ ). Il découle de l'analyse précédente que l'ESM et le MVACL fournissent des résultats identiques pour des niveaux symétriques de bruit, mais que dans les autres cas, ils sont fondés sur des bases opposées en faisant l'hypothèse qu'un des termes  $\mathbf{A}_\alpha$  ou  $\mathbf{B}_\alpha$  est négligeable. L'approche MVACL requiert de plus la connaissance des variances du bruit.

Plusieurs cas typiques peuvent être considérés :

- Le RSB (Rapport Signal sur Bruit) est élevé : le terme  $\mathbf{B}_\alpha$  est quasi nul rendant le terme  $\mathbf{A}_\alpha$  prépondérant : le  $\alpha$  optimal est proche de 0.5, comme celui utilisé par l'ESM.
- Le RSB est faible, mais le bruit est distribué symétriquement ( $\sigma_I \approx \sigma_T$ ) sur  $I$  et  $T$ , alors la variance de  $\mathbf{B}_\alpha \mathbf{v}^*$  est également minimale pour  $\alpha = 0.5$ . Ainsi la valeur optimale est encore  $\alpha = 0.5$ . Nous remarquons d'ailleurs que la variance du terme de bruit est alors deux fois plus petite pour  $\alpha = 0.5$  que la variance de ce même terme pour  $\alpha = 0$  (ou bien 1) grâce au moyennage des matrices Jacobiennes issues des deux images. Ceci est un deuxième élément expliquant le meilleur comportement de l'ESM par rapport aux approches directes et inverses dans le cas de bruit symétrique.
- Le RSB est faible et très asymétriquement distribué  $\sigma_I \gg \sigma_T$  (resp.  $\sigma_I \ll \sigma_T$ ) et l'initialisation est proche de l'optimum ( $\|\mathbf{v}^*\|^2 \ll \|\mathbf{v}^*\|$ ) : pour un niveau de bruit suffisant,  $\mathbf{B}_\alpha$  devient prépondérant devant  $\mathbf{A}_\alpha$ , le  $\alpha$  optimal est celui qui minimise la variance de  $\mathbf{B}_\alpha$  i.e.  $\alpha = 1$  ce qui correspond à l'ICL (resp.  $\alpha = 0$  ce qui correspond au FCL).

Ainsi l'ACL semble capable de gérer l'ensemble de ces situations typiques à condition de prendre en compte à la fois les termes  $\mathbf{A}_\alpha$  et  $\mathbf{B}_\alpha$ . Nous proposons à présent plusieurs approches pour estimer automatiquement le paramètre  $\alpha$ .

### 5.1.3 ACL Géométrique pour l'estimation de $\alpha$

#### 5.1.3.1 Point de vue géométrique sur l'algorithme Gauss-Newton pour l'alignement d'images

Dans la suite, nous notons  $F_\alpha$  le sous-espace affine passant par  $\mathbf{e}_\alpha(\mathbf{0})$  et parallèle à l'espace généré par  $\mathbf{J}_\alpha(\mathbf{0})$ . Les éléments de  $F_\alpha$  sont obtenus par la linéarisation de l'erreur :

$$\tilde{\mathbf{e}}_\alpha(\mathbf{v}) = \mathbf{e}_\alpha(\mathbf{0}) + \mathbf{J}_\alpha(\mathbf{0})\mathbf{v}. \quad (5.29)$$

Nous notons également  $\hat{\mathbf{v}}_\alpha$  l'incrément dans l'algèbre de Lie estimé en utilisant  $\mathbf{J}_\alpha(\mathbf{0})$  dans l'équation (5.8).

L'interprétation géométrique de l'algorithme Gauss-Newton est que l'incrément  $\hat{\mathbf{v}}_\alpha$  est choisi de telle sorte que  $\hat{\mathbf{e}}_\alpha = \tilde{\mathbf{e}}_\alpha(\hat{\mathbf{v}}_\alpha)$  corresponde à la projection orthogonale de l'origine  $\mathbf{O}$  de l'espace d'erreur sur le sous-espace affine  $F_\alpha$ . Ainsi, les différentes approches ACL se distinguent par le choix du sous espace  $F_\alpha$  sur lequel  $\mathbf{O}$  est projeté. Ceci est illustré fig. 5.2 pour le cas d'un modèle de mouvement à un degré de liberté.

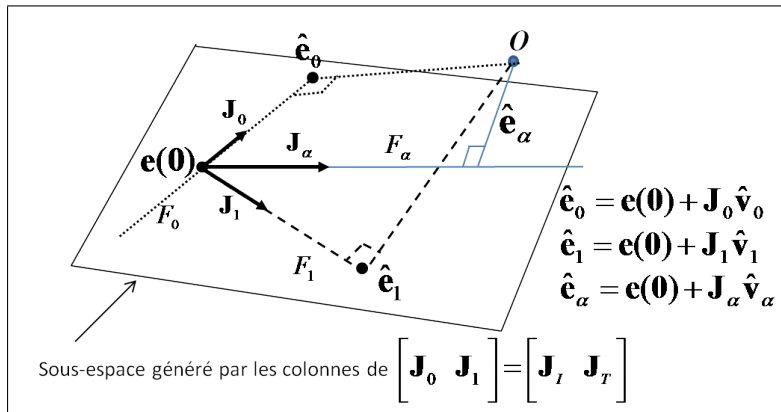


FIGURE 5.2 – Interprétation géométrique des différentes approches dans le cas d'une optimisation Gauss-Newton pour un mouvement caractérisé par un paramètre :  $\mathbf{O}$  est projeté orthogonalement sur un sous espace affine particulier  $F_\alpha$  (doite affine pour un modèle à  $p = 1$  paramètre). La projection de  $\mathbf{O}$  sur  $F_\alpha$  est notée  $\hat{\mathbf{e}}_\alpha$ .

### 5.1.3.2 Algorithme GACL (Geometric ACL)

La valeur de  $\alpha$  de l'approche ACL peut être considérée comme une valeur de confiance sur les directions (dans l'espace erreur), des sous espaces  $F_0$  et  $F_1$ . Cette valeur devrait favoriser le sous-espace le plus proche de l'origine  $\mathbf{O}$  de l'espace erreur. Nous proposons de calculer  $\alpha$  en utilisant les distances entre l'origine  $\mathbf{O}$  et chacun des sous espaces  $F_0$  et  $F_1$ . Considérons le plan passant par les points  $\hat{\mathbf{e}}_0$ ,  $\hat{\mathbf{e}}_1$  et  $\mathbf{O}$  (cf. fig. 5.3). On note  $\mathbf{P}$  la projection orthogonale de  $\mathbf{O}$  sur la droite passant par  $\hat{\mathbf{e}}_1$  et  $\hat{\mathbf{e}}_0$ ,  $l$  la distance entre  $\hat{\mathbf{e}}_1$  et  $\hat{\mathbf{e}}_0$ ,  $d_1$  (respectivement  $d_0$ ) la distance entre  $\mathbf{O}$  et  $\hat{\mathbf{e}}_1$  (respectivement  $\hat{\mathbf{e}}_0$ ). En utilisant le théorème de Pythagore dans les triangles  $(\mathbf{O}, \mathbf{P}, \hat{\mathbf{e}}_0)$  et  $(\mathbf{O}, \mathbf{P}, \hat{\mathbf{e}}_1)$ , nous obtenons :

$$\alpha^{\text{GACL}} = \frac{d_0^2 - d_1^2}{2l^2} + \frac{1}{2} = \frac{\langle \hat{\mathbf{e}}_0 | (\hat{\mathbf{e}}_0 - \hat{\mathbf{e}}_1) \rangle}{\|(\hat{\mathbf{e}}_0 - \hat{\mathbf{e}}_1)\|^2}, \quad (5.30)$$

avec  $\langle . | . \rangle$  (respectivement  $\|.\|$ ) représentant le produit scalaire (respectivement la norme Euclidienne) dans  $\mathbb{R}^N$ .

La valeur de  $\alpha^{\text{GACL}}$  favorise le sous-espace le plus proche de  $\mathbf{O}$ . Si les deux sous-espaces sont à égale distance de  $\mathbf{O}$ ,  $\alpha^{\text{GACL}}$  vaut 0.5, et le comportement de l'ESM est utilisé. Dans sa forme la plus simple cette méthode nécessite le calcul de  $\hat{\mathbf{e}}_0$  et  $\hat{\mathbf{e}}_1$  afin d'estimer  $\alpha^{\text{GACL}}$  ce qui est relativement coûteux.

### 5.1.4 ACL Analytique pour l'estimation de $\alpha$ (AACL)

L'algorithme AACL considère la linéarisation du résidu pixel à pixel (5.6) autour de  $\mathbf{v} = \mathbf{0}$  (cf. eq. (5.29)). Une approximation du résidu de l'erreur complète est obtenue en utilisant eq. (5.29) dans eq. (5.5) :

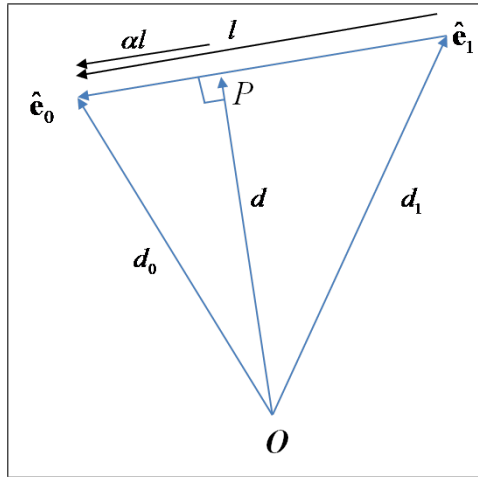


FIGURE 5.3 – Interprétation générique de la valeur  $\alpha^{\text{GACL}}$  : Les distances entre le point  $\mathbf{O}$  et les sous-espaces  $F_0$  et  $F_1$  sont estimées. Le paramètre  $\alpha^{\text{GACL}}$  est choisi pour favoriser le sous-espace le plus proche de  $\mathbf{O}$ .

$$E_{\alpha}^{\text{AACL}}(\mathbf{v}) = \|\mathbf{e}_{\alpha}(\mathbf{0}) + \mathbf{J}_I(\mathbf{0})\mathbf{v}\|^2 - 2\alpha(\mathbf{e}_{\alpha}(\mathbf{0}) + \mathbf{J}_I(\mathbf{0})\mathbf{v})^T \mathbf{J}_{\ominus}(\mathbf{0})\mathbf{v} + \alpha^2 \|\mathbf{J}_{\ominus}(\mathbf{0})\mathbf{v}\|^2, \quad (5.31)$$

avec  $\mathbf{J}_{\ominus} = \mathbf{J}_I(\mathbf{0}) - \mathbf{J}_T(\mathbf{0})$ .

La minimisation de ce résidu en fonction de  $\alpha$  pour un estimé fixe  $\hat{\mathbf{v}}$  permet de calculer le coefficient de pondération utilisé par l'algorithme AACL :

$$\alpha^{\text{AACL}}(\hat{\mathbf{v}}) = \arg \min_{\alpha} (E_{\alpha}^{\text{AACL}}(\hat{\mathbf{v}})) = \frac{(\mathbf{e}_{\alpha}(\mathbf{0}) + \mathbf{J}_I(\mathbf{0})\hat{\mathbf{v}})^T \mathbf{J}_{\ominus}(\mathbf{0})\hat{\mathbf{v}}}{\|\mathbf{J}_{\ominus}(\mathbf{0})\hat{\mathbf{v}}\|^2}, \quad (5.32)$$

que l'on peut réécrire sous une forme similaire à (5.30) :

$$\alpha^{\text{AACL}}(\hat{\mathbf{v}}) = \frac{\langle \tilde{\mathbf{e}}_0(\hat{\mathbf{v}}) | (\tilde{\mathbf{e}}_0(\hat{\mathbf{v}}) - \tilde{\mathbf{e}}_1(\hat{\mathbf{v}})) \rangle}{\|\tilde{\mathbf{e}}_0(\hat{\mathbf{v}}) - \tilde{\mathbf{e}}_1(\hat{\mathbf{v}})\|^2}, \quad (5.33)$$

avec l'erreur  $\tilde{\mathbf{e}}_{\alpha}(\hat{\mathbf{v}})$  défini par l'équation (5.29).

Avec l'approche GACL,  $\alpha$  est obtenu en projetant  $\mathbf{O}$  orthogonalement sur la droite affine  $(\hat{\mathbf{e}}_0, \hat{\mathbf{e}}_1)$ . Avec l'approche AACL,  $\mathbf{O}$  est projeté orthogonalement sur la droite  $(\tilde{\mathbf{e}}_0(\hat{\mathbf{v}}), \tilde{\mathbf{e}}_1(\hat{\mathbf{v}}))$ . Pour l'approche AACL, une valeur fixe pour  $\hat{\mathbf{v}}$  doit être fixée avant d'utiliser l'équation (5.33). Dans la suite, nous notons  $\text{AACL}^M$  l'approche analytique AACL qui utilise la méthode  $M$  pour calculer  $\hat{\mathbf{v}}$ , où  $M$  peut être la méthode FCL ( $\alpha = 0$ ), ICL ( $\alpha = 1$ ) ou bien l'ESM ( $\alpha = 0.5$ ). L'évaluation des approches proposées ( $\alpha$ -ACL, MVACL, AACL et GACL) et leur comparaison

avec les approches de l'état de l'art est présentée au chapitre 6.

## 5.2 Composition Bidirectionnelle sur groupe de Lie

L'approche bidirectionnelle, contrairement à l'ensemble des autres approches, suppose une reparamétrisation bidirectionnelle  $(\mathbf{v}_I, \mathbf{v}_T)$  :

$$\boldsymbol{\mu}_I = \boldsymbol{\mu}_I^n \circ \exp(\mathbf{v}_I), \quad (5.34)$$

$$\boldsymbol{\mu}_T = \boldsymbol{\mu}_T^n \circ \exp(-\mathbf{v}_T). \quad (5.35)$$

Nous nous concentrons désormais sur l'approche BCL (*Bidirectional Compositional Lie*), qui utilise l'erreur  $\mathbf{e}^{\text{BCL}}(\mathbf{v}_I, \mathbf{v}_T)$  avec

$$\begin{aligned} e_i^{\text{BCL}}(\mathbf{v}_I, \mathbf{v}_T) &= I(\mathbf{W}(\boldsymbol{\mu}_I^n \circ \exp(\mathbf{v}_I), \mathbf{x}_i)) \\ &\quad - T(\mathbf{W}(\boldsymbol{\mu}_T^n \circ \exp(-\mathbf{v}_T), \mathbf{x}_i)). \end{aligned} \quad (5.36)$$

Avec une optimisation Gauss-Newton, l'estimation utilise la matrice Jacobienne obtenue en concaténant les contributions de  $I$  et  $T$  :

$$J^{\text{BCL}} = \begin{bmatrix} \mathbf{J}_I(\mathbf{0}) & \mathbf{J}_T(\mathbf{0}) \end{bmatrix}, \quad (5.37)$$

et fournit l'estimation suivante :

$$\begin{bmatrix} \mathbf{v}_I \\ \mathbf{v}_T \end{bmatrix} = \begin{bmatrix} \mathbf{J}_I(\mathbf{0}) & \mathbf{J}_T(\mathbf{0}) \end{bmatrix}^+ \mathbf{e}^{\text{BCL}}(0, 0). \quad (5.38)$$

La loi de mise à jour Directe associée est alors :

$$\boldsymbol{\mu}_I^{n+1} = \boldsymbol{\mu}_I^n \circ \exp(\mathbf{v}_I) \circ \exp(\mathbf{v}_T). \quad (5.39)$$

### 5.2.1 Propriété d'invariance

En reprenant l'interprétation géométrique de la figure 5.2, l'approche BCL constitue une extension des approches ACL pour laquelle, à chaque itération, l'incrément local de dimension  $2p$   $(\mathbf{v}_I, \mathbf{v}_T)$  n'est plus contraint à explorer le sous-espace vectoriel de dimension  $p$  vérifiant  $\alpha \mathbf{v}_I = (1 - \alpha) \mathbf{v}_T$ . Ainsi l'approche FCL explore en effet le sous-espace  $(\mathbf{v}_I, 0)$ , l'approche ICL le sous-espace  $(0, \mathbf{v}_T)$  et l'approche SCL/ESM le sous-espace  $(\frac{1}{2}\mathbf{v}, \frac{1}{2}\mathbf{v})$  (comme cela est illustré sur la figure 3.5). L'approche BCL n'est pas restreinte à un sous-espace spécifique, et semble par conséquent en mesure d'explorer les sous-espaces considérés à la fois par les approches FCL, ICL et SCL. Puisque l'approche ESM/SCL a été démontrée ([Benhimane 2007b] et à la section 5.1.1.1) posséder les meilleures propriétés de convergence en l'absence de bruit, nous allons maintenant tenter d'interpréter la signification des dimensions supplémentaires relativement à cette approche.

Considérons le changement de variable suivant :

$$\mathbf{v}_\oplus = \mathbf{v}_I + \mathbf{v}_T \quad \text{et} \quad \mathbf{v}_\ominus = \mathbf{v}_I - \mathbf{v}_T \quad (5.40)$$

avec  $\mathbf{v}_\oplus$  et  $\mathbf{v}_\ominus$  représentant respectivement les portions symétriques et anti-symétriques.



Le changement de variable (5.40) peut être interprété en notant qu’au premier ordre sur  $\mathbf{v}_I$  et  $\mathbf{v}_T$  nous avons :

$$\begin{aligned}
 \mathbf{e}^{\text{BCL}}(\mathbf{v}_I, \mathbf{v}_T) &\approx \mathbf{e}^{\text{BCL}}(0, 0) + \mathbf{J}_I(0)\mathbf{v}_I + \mathbf{J}_T(0)\mathbf{v}_T \\
 &\approx \frac{1}{2}\mathbf{e}^{\text{BCL}}(0, 0) + \frac{1}{2}\mathbf{J}_I(0)\mathbf{v}_\oplus + \frac{1}{2}\mathbf{J}_T(0)\mathbf{v}_\oplus \\
 &\quad + \frac{1}{2}\mathbf{e}^{\text{BCL}}(0, 0) + \frac{1}{2}\mathbf{J}_I(0)\mathbf{v}_\ominus - \frac{1}{2}\mathbf{J}_T(0)\mathbf{v}_\ominus \\
 &\approx \frac{1}{2}(\mathbf{e}^{\text{BCL}}(\mathbf{v}_\oplus, \mathbf{v}_\oplus) + \mathbf{e}^{\text{BCL}}(\mathbf{v}_\ominus, -\mathbf{v}_\ominus)).
 \end{aligned} \tag{5.41}$$

Le premier terme correspond à l’erreur standard SCL, avec  $\mathbf{v}_I = \mathbf{v}_T = \frac{1}{2}\mathbf{v}_\oplus$ , c’est à dire que les 2 images sont compensées symétriquement l’une vers l’autre. Le second terme est spécifique à l’approche BCL. Il correspond à l’erreur obtenue en compensant les 2 images sur la classe d’équivalence des paramètres initiaux  $\text{equiv}(\boldsymbol{\mu}_I^n, \boldsymbol{\mu}_T^n)$  et définie section 3.4, où la transformation relative  $\boldsymbol{\mu}_I \circ \boldsymbol{\mu}_T^{-1}$  demeure constante.

La linéarisation de l’erreur relativement à  $(\mathbf{v}_\oplus, \mathbf{v}_\ominus)$  est obtenue à partir de (5.41) :

$$\mathbf{e}^{\text{BCL}}(\mathbf{v}_I, \mathbf{v}_T) \approx \mathbf{e}^{\text{BCL}}(0, 0) + \mathbf{J}_\oplus\mathbf{v}_\oplus + \mathbf{J}_\ominus\mathbf{v}_\ominus, \tag{5.42}$$

avec

$$\begin{aligned}
 \mathbf{J}_\oplus &= \frac{1}{2}(\mathbf{J}_I(\mathbf{0}) + \mathbf{J}_T(\mathbf{0})) \text{ et} \\
 \mathbf{J}_\ominus &= \frac{1}{2}(\mathbf{J}_I(\mathbf{0}) - \mathbf{J}_T(\mathbf{0})).
 \end{aligned} \tag{5.43}$$

La minimisation de  $E(\mathbf{v}_I, \mathbf{v}_T)$  relativement aux nouvelles variables  $(\mathbf{v}_\oplus, \mathbf{v}_\ominus)$  correspond alors à :

$$\begin{bmatrix} \mathbf{v}_\oplus \\ \mathbf{v}_\ominus \end{bmatrix} = - \begin{bmatrix} \mathbf{J}_\oplus & \mathbf{J}_\ominus \end{bmatrix}^+ \mathbf{e}^{\text{BCL}}(0, 0), \tag{5.44}$$

Etant donné que  $\mathbf{v}_\oplus$  et  $\mathbf{v}_\ominus$  peuvent être considérés comme des incréments infinitésimaux [Bayro-Corrochano 2007], la règle de mise à jour (5.39) peut s’approximer :

$$\boldsymbol{\mu}_I^{n+1} \approx \boldsymbol{\mu}_I^n \circ \exp(\mathbf{v}_\oplus). \tag{5.45}$$

Il s’en suit que  $\mathbf{v}_\ominus$  est un paramètre de transformation qui est modélisé durant l’estimation mais dont la valeur ne modifie pas l’estimation finale. Ce principe est similaire à celui utilisé dans les approches modélisant les variations de luminosité [Hager 1998], afin de rendre l’estimation de mouvement insensible à ce type de transformation. L’invariance que le BCL impose à chaque itération est au contraire une invariance locale autour de  $\mathbf{v} = 0$  relativement au sous espace  $\mathbf{J}_\ominus$ , ce qui correspond à un changement des coordonnées du repère de référence. Cette propriété est au coeur de l’algorithme PBCL que nous définissons au paragraphe suivant.

### 5.2.2 L’algorithme BCL projeté (PBCL)

L’algorithme BCL projeté (PBCL) tire partie de la loi de mise à jour (5.45) afin d’éviter de résoudre l’équation (5.38) avec la totalité de la matrice Jacobienne  $\mathbf{J}^{\text{BCL}}$ .

## 5.2. Composition Bidirectionnelle sur groupe de Lie

---

En raison des propriétés de la pseudo-inverse d'une matrice quand  $\begin{bmatrix} \mathbf{J}_\oplus & \mathbf{J}_\ominus \end{bmatrix}$  est de rang plein :

$$\mathbf{v}_\oplus = (\mathbf{P}_\ominus^\perp \mathbf{J}_\oplus)^\dagger \mathbf{e}^{\text{BCL}}(0, 0), \quad (5.46)$$

avec  $\mathbf{P}_\ominus^\perp$  représentant la projection (au sein de l'espace vectoriel de l'erreur à  $N$  dimensions) sur le sous-espace orthogonal à l'espace image de  $\mathbf{J}_\ominus$  (espace généré par les colonnes de  $\mathbf{J}_\ominus$ ) :

$$\mathbf{P}_\ominus^\perp = Id_N - \mathbf{J}_\ominus (\mathbf{J}_\ominus^t \mathbf{J}_\ominus)^{-1} \mathbf{J}_\ominus^t. \quad (5.47)$$

La différence entre les approches PBCL et SCL réside dans l'utilisation de  $\mathbf{P}_\ominus^\perp \mathbf{J}_\oplus$  au lieu de  $\mathbf{J}_\oplus$ , c'est à dire dans la projection dans l'espace d'erreur des directions correspondantes à une compensation bidirectionnelle qui transforme le repère de référence sans modifier la transformation relative entre  $I$  et  $T$ .

De plus, en notant que nous avons d'une part  $\mathbf{J}_\alpha(\mathbf{0}) = \mathbf{J}_\oplus + (\frac{1}{2} - \alpha)\mathbf{J}_\ominus$ , et d'autre part  $\mathbf{P}_\ominus^\perp \mathbf{J}_\ominus = 0$ , nous obtenons pour tout  $\alpha$  :

$$\mathbf{P}_\ominus^\perp \mathbf{J}_\alpha = \mathbf{P}_\ominus^\perp \mathbf{J}_\oplus = \mathbf{P}_\ominus^\perp \mathbf{J}_I = \mathbf{P}_\ominus^\perp \mathbf{J}_T. \quad (5.48)$$

En effet, les différences entre les approches ACL reposent sur une pondération différente de la contribution de  $\mathbf{J}_\ominus$ , qui est éliminée par la projection  $\mathbf{P}_\ominus^\perp$ . Cette propriété nous permet de définir la matrice Jacobienne de l'approche PBCL relativement à la matrice Jacobienne précalculée  $\mathbf{J}^{\text{ICL}} = \mathbf{J}_T$ , sans calculer  $\mathbf{J}_\oplus$  de manière explicite :

$$\mathbf{J}^{\text{PBCL}} = \mathbf{P}_\ominus^\perp \mathbf{J}_T. \quad (5.49)$$

Dans le cas non bruité, en supposant que (3.1) est vraie (invariance des niveaux de gris) et qu'une paramétrisation sur groupe de Lie est utilisée, la projection de l'équation fondamentale de l'ESM (5.18) conduit grâce à (5.48) à :

$$\mathbf{P}_\ominus^\perp \mathbf{e}(\mathbf{v}^*) = \mathbf{P}_\ominus^\perp \mathbf{e}(0) + \mathbf{J}^{\text{PBCL}} \mathbf{v}^* + O(\|\mathbf{v}^*\|^3). \quad (5.50)$$

Afin de simplifier les notations,  $\mathbf{e}_\alpha(\mathbf{0})$  est notée simplement  $\mathbf{e}(\mathbf{0})$  puisque cette erreur est indépendante de  $\alpha$ . La différence avec l'approche ESM d'origine est que la nouvelle équation est exprimée dans le sous-espace orthogonal à l'espace image de  $\mathbf{J}_\ominus$ , ce qui induit une invariance locale de la fonction objective relativement à un changement du repère de coordonnées de référence.

### 5.2.3 Analyse théorique de l'influence du bruit

Considérons deux images bruitées par deux bruits indépendants :

$$I = I_{nf} + \varepsilon_I \quad T = T_{nf} + \varepsilon_T \quad (5.51)$$

Ceci impacte sur les termes d'ordre 0 et d'ordre 1 :

$$\mathbf{e}(0) = \mathbf{e}_{nf}(0) + \varepsilon_I - \varepsilon_T \quad (5.52)$$

$$\mathbf{J}_\oplus = \mathbf{J}_{\oplus nf} + \frac{1}{2}(\mathbf{J}_{\varepsilon_I}(\mathbf{0}) + \mathbf{J}_{\varepsilon_T}(\mathbf{0})) \quad (5.53)$$

$$\mathbf{J}_\ominus = \mathbf{J}_{\ominus nf} + \frac{1}{2}(\mathbf{J}_{\varepsilon_I}(\mathbf{0}) - \mathbf{J}_{\varepsilon_T}(\mathbf{0})), \quad (5.54)$$

où  $\mathbf{J}_{\varepsilon_I}$  (resp.  $\mathbf{J}_{\varepsilon_T}$ ) représente la matrice Jacobienne de l'image de bruit  $\varepsilon_I$  (resp.  $\varepsilon_T$ ) à savoir la concaténation en  $\mathbf{w} = \mathbf{0}$  des matrices Jacobiennes de bruit par pixel (5.13) (resp. (5.14)).

Dans le cas bruité, l'erreur (5.18) devient une approximation du premier ordre seulement, avec un Jacobien bruité :

$$\mathbf{e}(\mathbf{v}^*) = \mathbf{e}(0) + (\mathbf{J}_{\oplus nf} + \frac{1}{2}(\mathbf{J}_{\varepsilon_I}(\mathbf{0}) + \mathbf{J}_{\varepsilon_T}(\mathbf{0})))\mathbf{v}^* + O(\|\mathbf{v}^*\|^2). \quad (5.55)$$

On note  $\sigma_I^2$  et  $\sigma_T^2$  les variances du bruit dans  $\mathbf{J}_{\varepsilon_I}(\mathbf{0})\mathbf{v}^*$  et  $\mathbf{J}_{\varepsilon_T}(\mathbf{0})\mathbf{v}^*$ . La variance du bruit dans le Jacobien de l'ESM  $\mathbf{J}_{\oplus}\mathbf{v}^*$  vaut alors  $\frac{1}{4}(\sigma_I^2 + \sigma_T^2)$ , ce qui est la moitié de la variance du bruit des matrices Jacobiennes  $\mathbf{J}^{\text{FCL}}\mathbf{v}^*$  ou  $\mathbf{J}^{\text{ICL}}\mathbf{v}^*$  pour des niveaux de bruit symétriques.

En projetant (5.55), l'erreur projetée PBCL (5.50) devient également uniquement une approximation du premier ordre dans le cas bruité :

$$\mathbf{P}_{\ominus}^{\perp}\mathbf{e}(\mathbf{v}^*) = \mathbf{P}_{\ominus}^{\perp}\mathbf{e}(0) + \mathbf{P}_{\ominus}^{\perp}\mathbf{J}_{\oplus}\mathbf{v}^* + O(\|\mathbf{v}^*\|^2). \quad (5.56)$$

Nous montrons maintenant que  $\mathbf{P}_{\ominus}^{\perp}$  élimine par projection une partie du bruit touchant  $\mathbf{J}_{\oplus}$ . Nous avons précédemment vu que (cf. equation (5.26)) :

$$\mathbf{J}_{\ominus nf}\mathbf{v}^* = (\mathbf{J}_I^{nf}(0) - \mathbf{J}_T^{nf}(0))\mathbf{v}^* \quad (5.57)$$

$$= (\mathbf{J}_I^{nf}(0) - \mathbf{J}_I^{nf}(\mathbf{v}^*))\mathbf{v}^* = \mathcal{O}(\|\mathbf{v}^*\|^2). \quad (5.58)$$

De plus, par définition  $\mathbf{P}_{\ominus}^{\perp}\mathbf{J}_{\ominus} = \mathbf{0}$ , ce qui induit :

$$\mathbf{P}_{\ominus}^{\perp}(\mathbf{J}_{\varepsilon_I}(\mathbf{0}) - \mathbf{J}_{\varepsilon_T}(\mathbf{0}))\mathbf{v}^* = 2\mathbf{P}_{\ominus}^{\perp}(\mathbf{J}_{\ominus} - \mathbf{J}_{\ominus nf})\mathbf{v}^* \quad (5.59)$$

$$= 2\mathbf{P}_{\ominus}^{\perp}\mathbf{J}_{\ominus nf}\mathbf{v}^* = \mathcal{O}(\|\mathbf{v}^*\|^2). \quad (5.60)$$

Supposons maintenant par exemple que le template est moins bruité que l'image. A l'aide de (5.60), nous pouvons modifier l'expression  $\mathbf{P}_{\ominus}^{\perp}\mathbf{J}_{\oplus}$  comme suit :

$$\mathbf{P}_{\ominus}^{\perp}\mathbf{J}_{\oplus}\mathbf{v}^* = \mathbf{P}_{\ominus}^{\perp}(\mathbf{J}_{\oplus nf} + \frac{1}{2}(\mathbf{J}_{\varepsilon_I}(\mathbf{0}) - \mathbf{J}_{\varepsilon_T}(\mathbf{0}) + 2\mathbf{J}_{\varepsilon_T}(\mathbf{0})))\mathbf{v}^* \quad (5.61)$$

$$= \mathbf{P}_{\ominus}^{\perp}(\mathbf{J}_{\oplus nf} + \mathbf{J}_{\varepsilon_T}(\mathbf{0}))\mathbf{v}^* + \mathcal{O}(\|\mathbf{v}^*\|^2). \quad (5.62)$$

Dans le cas de niveaux de bruit asymétriques, l'équation (5.62) met en évidence que l'influence du bruit sur la matrice Jacobienne  $\mathbf{J}_{\oplus}$  peut être réduite à l'intérieur du sous-espace engendré par  $\mathbf{P}_{\ominus}^{\perp}$ . En effet, de la perturbation complète  $\frac{1}{2}(\mathbf{J}_{\varepsilon_I}(\mathbf{0}) + \mathbf{J}_{\varepsilon_T}(\mathbf{0}))$  sur la matrice Jacobienne  $\mathbf{J}_{\oplus}$ , seulement la partie la moins bruitée  $\mathbf{J}_{\varepsilon_T}(\mathbf{0})$  est conservée. En particulier, les dégradations de la matrice Jacobienne projetée ont une influence à l'ordre 2 seulement, lorsque le template n'est pas bruité.

Ces résultats théoriques seront validés expérimentalement dans les sections 6.2.2 et 6.2.3.

### 5.3 Résumé des nouvelles approches et approximations efficaces

Nous avons introduit dans ce chapitre deux nouvelles paramétrisations incrémentales sur groupe de Lie, la paramétrisation ACL et la paramétrisation BCL. Ces approches se distinguent

### 5.3. Résumé des nouvelles approches et approximations efficaces

		Espace bidirectionnel contraint						Espace bidirectionnel complet	
		Erreur ACL						Erreur BCL	
		$\alpha$ fixe			$\alpha$ adaptatif			–	
	FCL	ICL	ESM	$\alpha$ -ACL	MVACL	GACL	AAACL <sup>M</sup>	BCL	
$\alpha$	0	1	0.5	Arbitraire	$\alpha^{\text{MVACL}}$ eq. (5.28)	$\alpha^{\text{GACL}}$ eq. (5.30)	$\alpha^{\text{AAACL}}(\hat{\mathbf{v}})$ eq. (5.33)	–	
$\mathbf{J}(\mathbf{0})$	$(1 - \alpha)\mathbf{J}_I(\mathbf{0}) + \alpha\mathbf{J}_T(\mathbf{0})$						$[\mathbf{J}_I(\mathbf{0}) \quad \mathbf{J}_T(\mathbf{0})]$		

TABLE 5.1 – Caractéristiques des approches sur groupe de Lie étudiées. Pour l’approche AAACL,  $M$  correspond à l’algorithme (FCL, ICL or ESM) utilisé pour initialiser  $\hat{\mathbf{v}}$  dans l’équation (5.33).

1. Précalcul :

- (a) Compenser  $T$  pour calculer  $T^0 = T(\mathbf{W}(\boldsymbol{\mu}_T^0, \cdot))$ ,
- (b) Calculer les gradients de  $T^0$  :  $\nabla T^0$ ,
- (c) Calculer la matrice Jacobienne du template  $\mathbf{J}_T(\mathbf{0})$  (eq. (5.2)),
- (d) Estimer  $\alpha$  pour les méthodes  $F$ -AAACL,  $F$ -GACL et MVACL (eq. (5.33), (5.30) and (5.28)).

2. Répéter jusqu’à convergence :

- (a) Compenser  $I$  par  $\mathbf{W}(\boldsymbol{\mu}_I^n, \mathbf{x})$  pour calculer  $I^n = I(\mathbf{W}(\boldsymbol{\mu}_I^n, \cdot))$ .
- (b) Calculer l’erreur courante  $\mathbf{e}_\alpha(\mathbf{0})$  (eq. (5.6)) et la matrice Jacobienne de l’image  $\mathbf{J}_I(\mathbf{0})$  (eq. (5.1)),
- (c) Uniquement pour les approches AAACL et GACL, calculer  $\alpha$  (eq. (5.33) ou (5.30)),
- (d) Calculer la matrice Jacobienne de l’erreur  $\mathbf{J}(\mathbf{0})$  (eq. (5.10) pour les approches asymétriques, eq. (5.37) pour l’approche BCL et eq. (5.49) pour l’algorithme PBCL),
- (e) Résoudre pour  $\mathbf{v}$  (eq. (5.8)),
- (f) Mettre à jour les paramètres  $\boldsymbol{\mu}^{n+1}$  (eq. (5.7)).

TABLE 5.2 – Schéma itératif des approches par composition sur groupe de Lie.

au niveau des contraintes imposées dans l’exploration de l’espace bidirectionnelle : les approches ACL sont des approches bidirectionnelles subordonnées pour lesquelles à chaque itération l’incrément local de dimension  $2p$  ( $\mathbf{v}_I, \mathbf{v}_T$ ) est contraint à explorer le sous-espace vectoriel de dimension  $p$  vérifiant  $\alpha\mathbf{v}_I = (1 - \alpha)\mathbf{v}_T$ . L’approche BCL est non contrainte et peut donc explorer librement tout l’espace bidirectionnel.

Les approches de l’état de l’art FCL, ICL et ESM peuvent toutes être considérées comme des cas particuliers de la formulation ACL. Elles utilisent des valeurs différentes et fixes pour  $\alpha$  lors du calcul de la matrice Jacobienne (5.10). Au sein de la formulation ACL, nous avons proposé également différentes heuristiques pour une estimation automatique de  $\alpha$  : les approches AAACL, GACL et MVACL. Ces approches estiment  $\alpha$  à chaque itération.

Le tableau 5.1 rappelle les différentes méthodes discutées dans ce chapitre. Les principales étapes de l’algorithme itératif sont fournies dans le tableau 5.2.

Pour l’étape 2c), si le  $\alpha$  calculé est  $> 1$  (resp.  $< 0$ ), nous initialisons ce dernier arbitrairement à

1 (resp. 0) ce qui entraîne le comportement de l’approche ICL (resp. FCL).

Comme cela a été évoqué dans [Baker 2004], les étapes 2c et 2d peuvent être évitées pour l’approche ICL ( $\alpha = 1$ ). De plus pour cette approche, la pseudo-inverse du Jacobien  $\mathbf{J}_{\alpha=1}(\mathbf{0}) = \mathbf{J}_T(\mathbf{0})$  peut être précalculée. L’algorithme ICL est donc l’algorithme le plus efficace des approches ACL en terme de temps de calcul par itération.

L’approche BCL est légèrement plus coûteuse en terme de temps de calcul par rapport aux approches classiques comme l’ESM en raison de l’estimation de deux fois plus de paramètres conduisant à une matrice Jacobienne de largeur  $2 \times p$ . Cela augmente, en particulier, le temps de calcul de la pseudo-inverse de cette matrice (étape 2e)) ainsi que le coût de la mise à jour des paramètres (étape 2f)). L’approche PBCL requiert le calcul de la projection ( $\mathbf{P}_{\ominus}^{\perp} \mathbf{J}_{\oplus}$ ). En pratique, cela peut être implémenté efficacement en pratiquant une décomposition QR de  $\mathbf{J}_{\ominus}$ . Cette étape demeure cependant coûteuse.

Les approches AACL et GACL introduisent une charge supplémentaire pour le calcul de  $\alpha$  (étape 2c) rendant ces approches moins efficaces en termes de temps de calcul si elles sont lancées sur un processeur séquentiel. Pour améliorer les performances de ces approches, nous proposons d’estimer  $\alpha$  seulement une fois lors de la première itération. Cela conduit à une approximation des approches AACL et GACL dénotées  $F$ -GACL (Fast-GACL) et  $F$ -AACL (Fast-AACL). Ces nouvelles approches introduisent uniquement une charge supplémentaire pour calculer  $\alpha$  durant la première itération par rapport aux approches  $\alpha$ -ACL qui utilisent un  $\alpha$  fixé (à part l’approche ICL).

Ce coût supplémentaire est appelé E-CC (*Extra Computational Cost*) dans la suite. De plus nous notons P-CC (*Pre-Computational Cost*) le coût introduit à chaque fois que le template  $T$  change (Etape 1) dans le tableau 5.2) et I-CC (*Iterative Computational Cost*) le coût d’une itération d’une approche  $\alpha$ -ACL (étapes 2a), 2b), 2d, 2e) et 2f)) dans le tableau 5.2). Le tableau 5.3 rassemblent les coûts calculatoires des différentes approches obtenus avec une implémentation Matlab sur un Intel(R) Core(TM)2 duo CPU 3.0 GHz avec 4 GB de RAM. Le nombre de pixels  $N$  du template est 10000. Le nombre de paramètres à estimer pour l’homographie est  $p = 8$ .

Méthodes	P-CC	E-CC	I-CC
FCL	4.22	0	7.00
ICL	7.60	0	4.53
ESM	6.73	0	7.16
AACL <sup>ESM</sup>	6.38	0	9.08
$F$ – AACL <sup>ESM</sup>	6.40	1.33	7.15
GACL	7.57	0	8.99
$F$ – GACL	7.61	1.25	7.15
BCL	6.33	0	8.13
PBCL	6.33	0	10.25

TABLE 5.3 – Temps de calcul (en ms) des différentes approches. Trois étapes peuvent être distinguées : le précalcul sur le template  $T$  (P-CC), la charge supplémentaire introduite par les nouvelles approches efficaces  $F$ -GACL et  $F$ -AACL à chaque fois qu’une nouvelle image  $I$  doit être utilisée (E-CC), et le coût par itération de chaque approche (I-CC).

Le E-CC des approches efficaces représente environ un cinquième du temps nécessaire pour une

### 5.3. Résumé des nouvelles approches et approximations efficaces

---

itération de l'ESM. Ainsi le coût supplémentaire introduit pour estimer  $\alpha$  avec les approximations efficaces des approches GACL et AACL est très faible lorsque de nombreuses itérations sont nécessaires pour atteindre la convergence.

L'ensemble des algorithmes discutés seront évalués expérimentalement dans le chapitre suivant.



# Validation expérimentale

---

## Sommaire

---

<b>6.1 Performances des approches ACL sur données synthétiques . . . . .</b>	<b>91</b>
6.1.1 Présentation du banc de test . . . . .	91
6.1.2 Comportement des algorithmes ACL à $\alpha$ fixe en présence de bruit . . . . .	92
6.1.3 Influence de l'asymétrie du bruit sur les approches ACL . . . . .	94
<b>6.2 Performances de l'approche BCL sur données synthétiques . . . . .</b>	<b>99</b>
6.2.1 Équivalence des approches BCL et PBCL . . . . .	99
6.2.2 Impact de la projection pour les approches BCL et PBCL . . . . .	99
6.2.3 Influence de l'asymétrie des niveaux de bruit : BCL vs ACL . . . . .	100
<b>6.3 Application au suivi dans des conditions de faible luminosité . . . . .</b>	<b>103</b>
6.3.1 Illustration de suivi sur des données réelles . . . . .	103
6.3.2 Analyse moyennée sur des séquences réelles . . . . .	105
6.3.3 Analyse statistique sur des séquences semi-synthétiques . . . . .	111

---

Dans ce chapitre, nous évaluons les nouveaux algorithmes introduits dans ce mémoire. Dans un premier temps, nous validons expérimentalement l'intérêt des approches ACL et BCL en présence d'images dégradées par du bruit. Afin de comparer ces approches à celles de l'état de l'art pour une grande variété de conditions de bruit, nous modifions le banc de test de Baker et Matthews [Baker 2004] (section 6.1.1) pour être en mesure de contrôler les niveaux de bruit sur chaque image en termes de RSB (Rapport Signal sur Bruit). Les mérites respectifs de chaque approche sont alors mis en évidence sur les données synthétiques générées (sections 6.1 et 6.2).

Dans un second temps, nous confrontons les approches sur des images réelles acquises dans des conditions de faible luminosité et pour lesquelles le bruit d'acquisition est important. Nous proposons une procédure de construction d'un template moyen permettant de profiter des avantages des approches ACL et BCL dans le contexte du suivi d'objet.

## 6.1 Performances des approches ACL sur données synthétiques

Dans cette partie nous comparons les nouveaux algorithmes proposés et ceux de l'état de l'art sur données synthétiques.

### 6.1.1 Présentation du banc de test

Le banc de test utilisé est inspiré de celui proposé dans [Baker 2004] dont le principe a été rappelé en annexe D. Des transformations aléatoires sont générées en ajoutant un bruit Gaussien



spatial d'écart-type  $\sigma_p$  (appelé Sigma Point) à quatre positions canoniques dans le repère d'une image de référence  $I_{ref}$  : ces 4 paires de points (canoniques et points de test) définissent un vecteur de paramètres homographiques caractérisant la déformation. Ces paramètres sont utilisés pour déformer  $I_{ref}$  en l'image template  $T_{ref}$ .

$I_{ref}$  et  $T_{ref}$  sont ensuite dégradées par un bruit additif Gaussien :  $I = I_{ref} + \varepsilon_I$  et  $T = T_{ref} + \varepsilon_T$ . Le niveau des bruits est caractérisé par leurs écart-type  $\sigma_I$  et  $\sigma_T$  ainsi que par les RSB correspondants  $RSB_I$  et  $RSB_T$  relativement aux images idéales :

$$RSB = 10 \log_{10} \left( \frac{\mathbb{E}(I_{ref}^2)}{Var(\varepsilon)} \right), \quad (6.1)$$

avec  $\mathbb{E}(I_{ref}^2)$  désignant la valeur moyenne du carré des intensités sur l'ensemble de l'image  $I_{ref}$ .

Les algorithmes d'alignement d'images sont alors exécutés afin d'aligner l'image  $I$  sur l'image  $T$ . Le critère de convergence est l'erreur RMS (*Root Mean Square Error*) à savoir la distance quadratique moyenne entre les positions des points de test et les positions finales des points canoniques déplacées à l'aide de la transformation estimée (notée sur les courbes Erreur RMS Point). Deux critères de performance sont envisagés :

- Fréquence moyenne de convergence : pourcentage des tests pour lesquels un algorithme converge vers un estimé correct (estimé pour lequel l'Erreur RMS Point est inférieure à 1 pixel),
- Vitesse moyenne de convergence : pour les tests pour lesquels l'ensemble des méthodes ont convergé, l'Erreur RMS Point moyen est tracé en fonction du nombre d'itérations.

Afin de simuler l'asymétrie du bruit, le bruit est paramétré par une variance totale  $\sigma^2$ , et par un coefficient d'asymétrie  $\beta \in [0,1]$ . La variance totale  $\sigma^2$  est calculée à partir d'un  $RSB$  global afin de pouvoir moyenniser les résultats obtenus sur toutes les images (figure 6.1) dans des conditions similaires. Le bruit est distribué sur  $I$  et  $T$  comme suit :

$$\sigma_I^2 = (1 - \beta)\sigma^2, \quad \sigma_T^2 = \beta\sigma^2, \quad \text{avec } \beta \in [0, 1]. \quad (6.2)$$

avec  $\sigma_I$  (resp.  $\sigma_T$ ) utilisé comme écart-type du bruit blanc Gaussien dégradant l'image  $I$  (resp. le template  $T$ ). Etant donné que les bruits perturbant les deux images sont indépendants,  $\sigma^2 = \sigma_I^2 + \sigma_T^2$  représente la variance de bruit total injecté dans l'image différence ( $I - T$ ). Le RSB total correspondant est calculé selon l'équation (6.1) en fonction de la variance de l'image de référence.

Dans la suite, nous présentons des résultats moyens obtenus en appliquant le banc de test sur les cinq images [A-E] de la figure 6.1. Le modèle de mouvement est une homographie paramétrée comme dans [Benhimane 2007b] : une matrice d'homographie  $3 \times 3$   $\mathbf{H}$  qui appartient au groupe de Lie  $\mathcal{G}$  des matrices vérifiant  $\det(\mathbf{H}) = 1$ . Son algèbre de Lie associée  $\mathfrak{G}$  est l'espace vectoriel des matrices de trace nulle. Pour l'ensemble des expérimentations présentées, 500 tests par Sigma Point et par image sont effectués pour estimer la fréquence de convergence. Pour la vitesse moyenne de convergence, 100 tests sont effectués par image.

### 6.1.2 Comportement des algorithmes ACL à $\alpha$ fixe en présence de bruit

Nous évaluons maintenant les performances de l'approche ACL introduite à la section 5.1 pour plusieurs valeurs fixes de  $\alpha$ . La figure 6.2 présente les comportements des approches ACL,

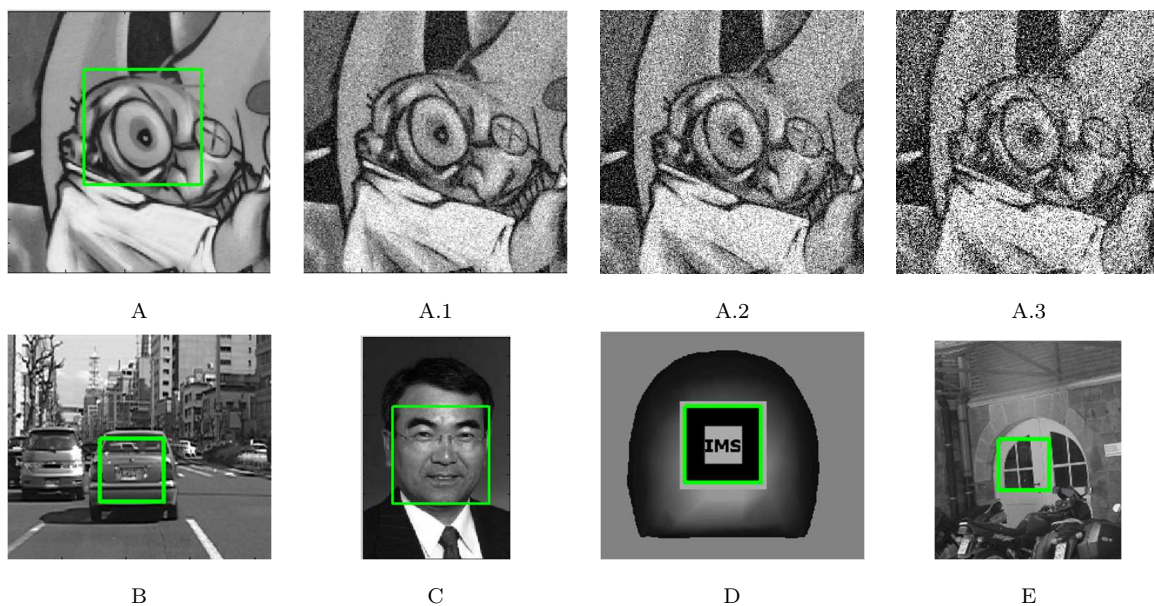


FIGURE 6.1 – Images utilisées pour les expériences sur données synthétiques. Première ligne : L'image non bruitée (à gauche) est utilisée pour extraire l'image template. L'image  $I$  est obtenue en ajoutant un bruit Gaussien à l'image non bruitée. Des exemples d'images dégradées (à droite) sont présentées pour différents RSB : 15 dB, 10 dB et 5 dB. Seconde ligne : autres images utilisées dans les expérimentations. Les images A et E proviennent de la base de données *INRIA Learning and Recognition in Vision (LEAR)* (<http://lear.inrialpes.fr/data>). Les images B et C proviennent du banc de test [Baker 2004]. L'image D a été obtenue à l'aide d'un rendu 3D openGL.

en termes de fréquence de convergence en fonction du bruit sur l'image  $I$ , et ce pour un Sigma Point de 12 pixels. Les meilleurs performances sont obtenues par un compromis entre l'utilisation des gradients des deux images  $T$  et  $I$ , et la pondération des effets du bruit. Ainsi, quand les images sont non bruitées (RSB élevé), les approches FCL ( $\alpha = 0$ ) et ICL ( $\alpha = 1$ ) produisent des résultats identiques et l'algorithme ESM ( $\alpha = 0.5$ ) produit de meilleurs résultats que ces deux méthodes. Cependant, quand le niveau de bruit augmente sur  $I$ , les performances de l'ESM diminuent. Quand le  $RSB$  de  $I$  devient trop faible, alors l'approche ICL produit les meilleurs résultats. Entre ces deux cas limites la meilleure robustesse est obtenue en utilisant une valeur intermédiaire de  $\alpha$ .

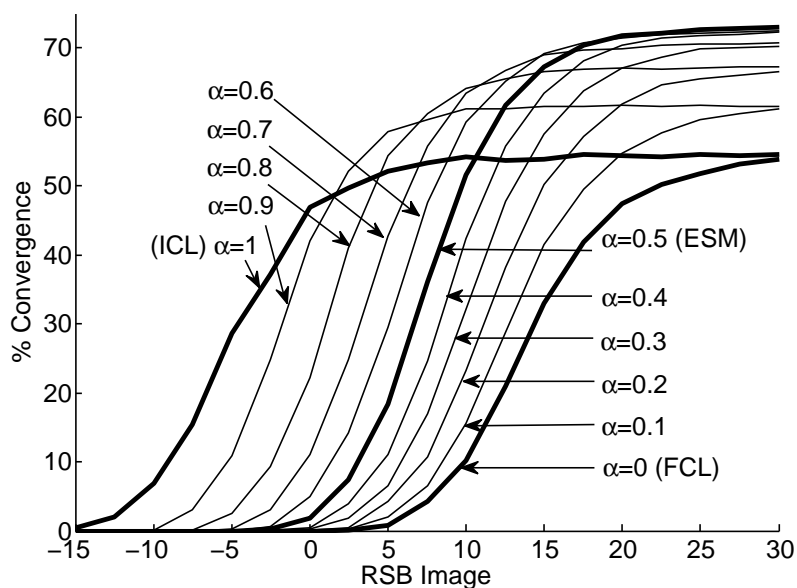


FIGURE 6.2 – Fréquence moyenne de convergence des différentes approches ACL relativement au bruit dans l'image pour un Sigma Point valant 12 pixels. Un RSB de 100 dB a été utilisé pour le template. L'image est dégradée par un bruit dont le RSB varie de -15 dB à 30 dB.

Ces résultats illustrent le fait que les différentes méthodes de l'état de l'art à savoir les approches FCL, ICL et ESM sont fortement affectées par le bruit sur les images. De manière générale si l'ESM a été démontré plus performant que les approches unidirectionnelles en absence de bruit, il existe des situations (notamment le cas où une des images est fortement bruitée) où ce dernier peut fonctionner moins bien qu'une approche unidirectionnelle. De plus, ces premiers résultats expérimentaux mettent en évidence l'existence de conditions pour lesquelles les méthodes de l'état de l'art ne produisent pas les meilleures performances, ce qui justifie l'introduction du paramètre  $\alpha$  et des approches ACL.

### 6.1.3 Influence de l'asymétrie du bruit sur les approches ACL

Quatre configurations sont considérées pour évaluer les performances des approches dans 4 situations typiques :

## 6.1. Performances des approches ACL sur données synthétiques

---

- Exp. 1 : initialisation proche (Sigma Point égal à 6 pixels) et RSB moyen (RSB total=10 dB)
- Exp. 2 : initialisation proche (Sigma Point égal à 6 pixels) et RSB faible (RSB total=5 dB)
- Exp. 3 : initialisation lointaine (Sigma Point égal à 12 pixels) et RSB élevé (RSB total=15 dB)
- Exp. 4 : initialisation lointaine (Sigma Point égal à 12 pixels) et RSB moyen (RSB total=10 dB)

Les résultats de ces expériences sont présentés sur la figure 6.3 et discutés dans la suite.

### 6.1.3.1 Approches conventionnelles

La première colonne de la figure 6.3 montre les résultats pour les algorithmes FCL, ICL, ESM et MVACL. Comme mentionné dans la section théorique (section 5.1.1.1), l'ESM fournit les meilleurs résultats lorsque  $\beta$  est proche de 0.5 (cf. exp. 1-4), ou bien quand le RSB total est élevé (Exp.3). Remarquons également que pour un RSB total fixé, l'ESM a des performances approximativement constantes, quelle que soit la valeur de  $\beta$ . En effet, la variance du bruit dans  $\mathbf{J}^{\text{ESM}}\mathbf{v}^* = \frac{1}{2}(\mathbf{J}_I(\mathbf{0}) + \mathbf{J}_T(\mathbf{0}))\mathbf{v}^*$  vaut  $\frac{1}{4}$  de la somme des variances de bruit de  $\mathbf{J}_I(\mathbf{0})\mathbf{v}^*$  et  $\mathbf{J}_T(\mathbf{0})\mathbf{v}^*$  (cf. sections 5.1.2 et 5.2.3). Cette variance demeure constante dans ces expériences.

En utilisant uniquement les gradients de  $I$  (resp.  $T$ ), le FCL (resp. ICL) produit les meilleurs résultats seulement pour  $\beta$  proche de 1 (resp. 0) ou bien quand  $T$  (resp.  $I$ ) est suffisamment dégradée (Exp. 2, faible RSB).

### 6.1.3.2 L'approche MVACL

Comme l'illustre la première colonne de la figure 6.3, les performances de l'approche MVACL (introduite à la section 5.1.2) passent progressivement des valeurs de celles de l'ICL à celles de l'ESM pour  $\beta \in [0, 0.5]$  puis de celles de l'ESM à celles du FCL pour  $\beta \in [0.5, 1]$ . L'approche MVACL produit de meilleures performances que les approches traditionnelles pour de faibles RSB (Exp.2) car elle pondère de manière appropriée le terme  $\mathbf{B}_\alpha$  de l'équation (5.22). En revanche, cette approche fonctionne moins bien que l'ESM dans des conditions de RSB élevé (Exp.3). En effet, en ne tenant compte que de la variance du bruit pour calculer  $\alpha$ , le MVACL pondère de manière optimale les termes de bruit dans la matrice Jacobienne sans tenir compte du terme  $\mathbf{A}_\alpha$  de l'équation (5.22) qui dans cette situation n'est pas négligeable.

### 6.1.3.3 Influence du choix de $\hat{\mathbf{v}}$ pour les approches AACL

Dans la suite nous évaluons les nouvelles approches proposées. La seconde colonne de la figure 6.3 présente le comportement des approches AACL<sup>FCL</sup>, AACL<sup>ICL</sup>, AACL<sup>ESM</sup> et GACL introduites aux sections 5.1.3 et 5.1.4 en termes de fréquence de convergence relativement à l'asymétrie du bruit. Les performances de l'ESM sont tracées en tant que référence. Nous observons que :

- Quel que soit le Sigma Point et le RSB total et quel que soit  $\beta$ , l'AACL<sup>M</sup> utilisant la méthode  $M$  pour initialiser  $\hat{\mathbf{v}}$  dans (5.33) produit toujours des résultats au moins aussi bons que la méthode  $M$  elle-même (Exp.1-4).

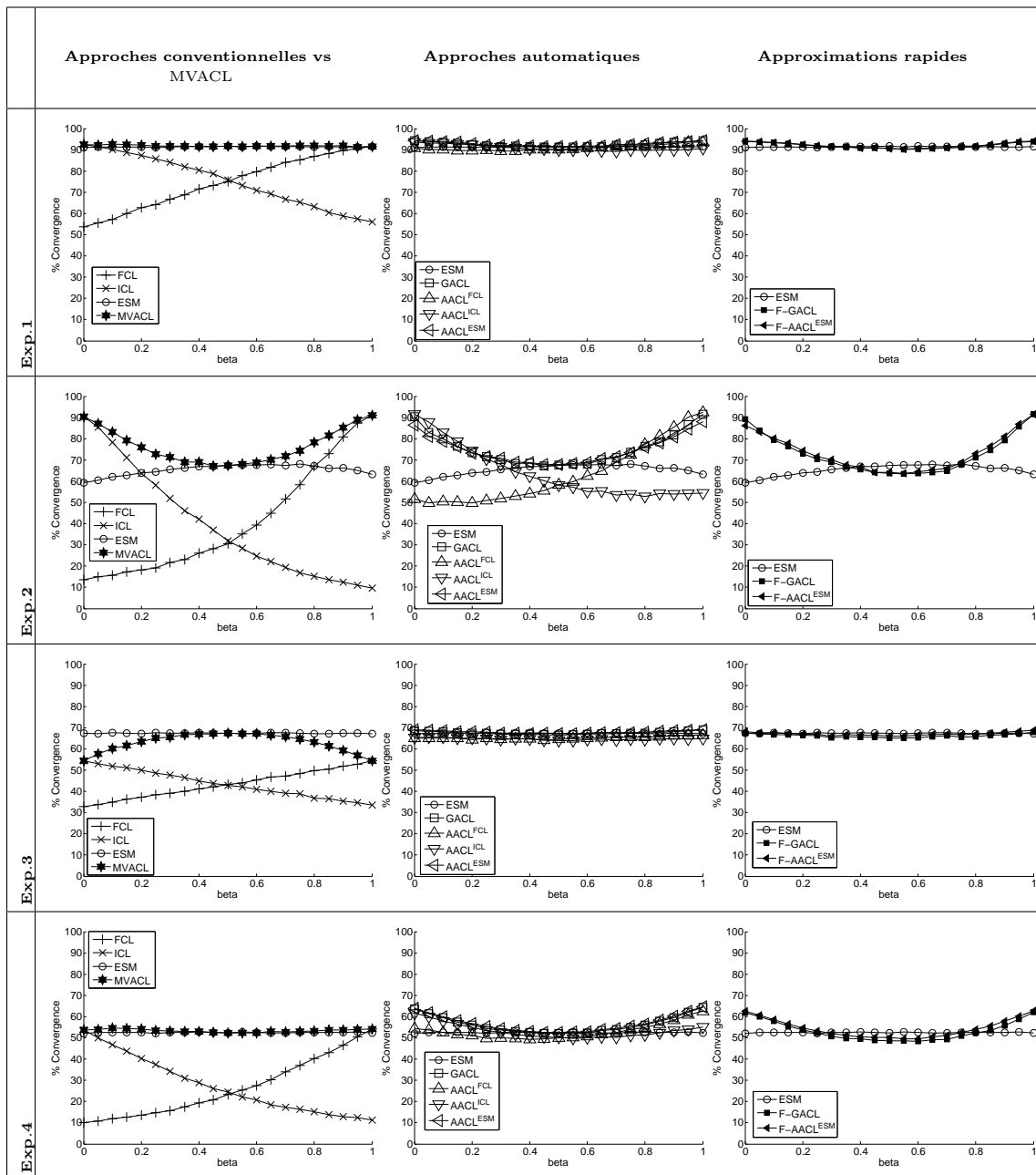


FIGURE 6.3 – Fréquence moyenne de convergence en fonction de l’asymétrie du bruit pour les différentes approches. Exp. 1 : Sigma Point = 6 pixels, RSB=10 dB. Exp. 2 : Sigma Point = 6 pixels, RSB=5 dB. Exp. 3 : Sigma Point = 12 pixels, RSB=15 dB. Exp. 4 : Sigma Point = 12 pixels, RSB=10 dB. Les valeurs numériques correspondant à des valeurs caractéristiques de  $\beta$  sont présentées dans le tableau 6.1.

## 6.1. Performances des approches ACL sur données synthétiques

Conditions	Exp.1 SP = 6 pixels RSB=10 dB			Exp.2 SP = 6 pixels RSB=5 dB			Exp.3 SP = 12 pixels RSB=15 dB			Exp.4 SP = 12 pixels RSB=10 dB			
	$\beta$	0	0.2	0.5	0	0.2	0.5	0	0.2	0.5	0	0.2	0.5
FCL		53.6	62.7	75.0	13.6	18.2	30.6	32.8	37.3	43.2	10.2	13.6	23.3
ICL		<u>92.4</u>	87.4	75.7	<u>90.4</u>	63.6	32.0	54.4	49.9	42.8	<u>53.6</u>	40.2	24.4
ESM		91.0	<u>91.3</u>	<b>91.7</b>	59.4	<u>63.9</u>	<u>67.3</u>	<u>67.4</u>	67.2	<b>67.4</b>	52.2	<u>52.4</u>	<u>52.3</u>
MVACL		92.4	92.2	<u>91.7</u>	90.4	<b>76.0</b>	67.3	54.4	63.4	<u>67.4</u>	53.6	54.2	52.3
GACL		<b>94.4</b>	92.5	91.4	90.5	73.3	67.6	68.5	67.4	66.4	63.6	56.0	52.2
AACL <sup>FCL</sup>		90.7	89.7	90.2	51.6	49.9	58.2	65.1	64.4	64.5	54.4	51.2	50.1
AACL <sup>ICL</sup>		93.7	92.4	90.0	<b>91.6</b>	74.3	58.1	65.3	65.1	64.0	61.5	54.8	50.3
AACL <sup>ESM</sup>		<b>94.4</b>	<b>92.9</b>	91.6	86.4	73.3	<b>67.8</b>	<b>68.9</b>	<b>68.0</b>	67.1	<b>63.9</b>	<b>56.6</b>	<b>52.6</b>
<i>F</i> -GACL		<u>94.1</u>	92.4	<u>90.6</u>	<u>89.1</u>	73.0	63.6	67.4	66.4	65.1	61.8	53.7	48.7
<i>F</i> -AACL <sup>ESM</sup>		94.0	<u>92.5</u>	<u>90.6</u>	86.3	<u>74.4</u>	<u>64.2</u>	<u>68.0</u>	<u>67.0</u>	<u>66.1</u>	<u>62.6</u>	<u>54.6</u>	<u>50.1</u>

TABLE 6.1 – Fréquence moyenne de convergence en fonction du bruit pour les différentes approches pour des valeurs caractéristiques de  $\beta$ . Ici l’abréviation SP est utilisée pour Sigma Point. La meilleure de toute les approches est en gras, la meilleure approche par catégorie est soulignée.

- Les performances de l’AACL<sup>M</sup> sont corrélées aux performances de la méthode *M* elle-même. Ainsi pour les situations où le FCL (resp. ICL) fonctionne bien, comme dans l’expérience 2 pour  $\beta$  proche de 1 (resp. proche de 0), l’AACL<sup>FCL</sup> fournit de meilleurs résultats que les méthodes conventionnelles et les autres approches AAACL en améliorant les performances du FCL (resp. ICL). Dans tous les autres cas, l’AACL<sup>ESM</sup> fournit les meilleurs résultats.

### 6.1.3.4 AAACL vs GACL

Quelle que soit la valeur de  $\beta$  et quel que soit le RSB global, le GACL et le AAACL<sup>ESM</sup> produisent des résultats semblables. De plus ils produisent toujours les meilleurs résultats (Exp. 1, 3 and 4) ou des résultats proches des meilleurs performances (Exp. 2). Le AAACL<sup>FCL</sup> et AAACL<sup>ICL</sup> sont plus robustes seulement dans les situations très asymétriques (Exp.2) où l’image bruitée n’est pas celle utilisée pour calculer les gradients permettant l’estimation de  $\hat{\mathbf{v}}$ . Pour toutes ces raisons, nous focaliserons notre étude dans la suite seulement sur les approches GACL et AAACL<sup>ESM</sup>.

### 6.1.3.5 Performance des approximations plus rapides des approches ACL automatiques

La figure 6.3, troisième colonne, présente les différents comportements de l’ESM, du *F*-AAACL<sup>ESM</sup> et du *F*-GACL (introduits à la section 5.3) en termes de fréquence de convergence en fonction de l’asymétrie du bruit. Nous observons que les approximations rapides produisent des comportements similaires aux comportements des approches AAACL exactes. Cependant les approximations rapides fournissent des résultats légèrement inférieurs à ceux des approches exactes, particulièrement dans les situations où il y a une répartition du bruit quasi-symétrique ( $\beta$  entre 0.35 et 0.65). Elles sont dans ce cas moins performantes que l’ESM (voir Exp. 2 et Exp. 4). Les approximations rapides semblent donc surtout appropriées pour aligner une image forte-

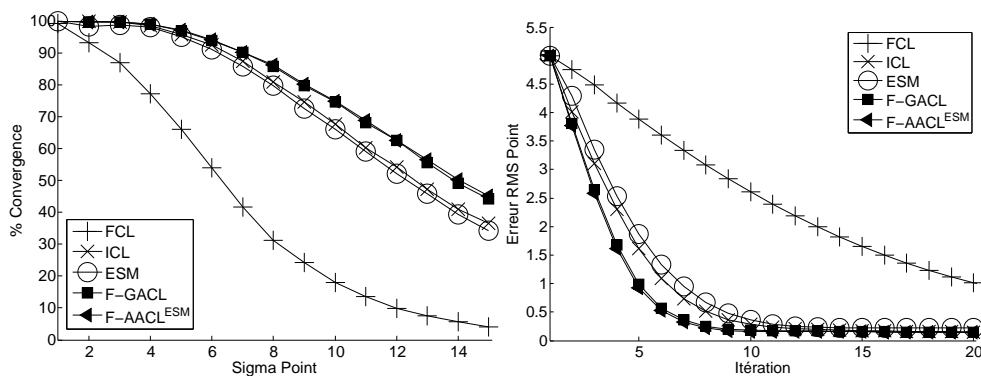


FIGURE 6.4 – Fréquence de convergence moyenne et vitesse de convergence pour les approches ESM, ICL, FCL,  $F\text{-AACL}^{\text{ESM}}$  et  $F\text{-GACL}$  en présence de bruit sur l'image. Un RSB de 10 dB a été utilisé sur l'image et de 100 dB sur le template.

ment bruitée à une autre image de bonne qualité (ce qui en pratique a lieu par exemple lors de la construction de panoramas [Pham 2005] où une image bruitée est alignée sur la mosaïque courante).

Afin d'illustrer ce type de situation, la figure 6.4 présente la vitesse de convergence moyenne (Sigma Point égal à 6 pixels) et fréquence moyenne de convergence en fonction du Sigma Point pour une image bruitée ( $RSB_I = 10dB$ ) et un template de haute qualité ( $RSB_T = 100dB$ ). Pour une meilleure lisibilité, seule les approximations rapides et les approches de l'état de l'art sont présentes.

Les approximations rapides produisent de meilleurs résultats que les approches de l'état de l'art à la fois en termes de fréquence de convergence et de vitesse de convergence. L'amélioration en termes de vitesse de convergence contrebalance le coût E-CC introduit pour le calcul de  $\alpha$  durant la première itération (cf. section 5.3). Ainsi les nouvelles approches sont au moins aussi rapides que les meilleurs algorithmes de la littérature, tout en gérant mieux les cas d'images bruitées. Pour une complexité calculatoire similaire et en fonction de la répartition du bruit sur les images, nous proposons les recommandations suivantes pour le choix d'une approche ACL :

- Dans les situations non bruitées ou lorsque les images sont perturbées par des bruits de niveaux symétriques ( $RSB_I \approx RSB_T$ ), l'ESM produira les meilleurs résultats parmi les approches étudiées ici.
- En présence de bruits fortement asymétriques  $RSB_I \gg RSB_T$  (resp.  $RSB_I \ll RSB_T$ ), le FCL (resp. ICL) fournira les meilleurs résultats.
- Dans toutes les autres situations les approches  $\text{AACL}^{\text{ESM}}$  et GACL devraient être préférées si les exigences en robustesse sont élevées. Faisant un compromis entre robustesse et rapidité de calcul, les approches  $F\text{-AACL}^{\text{ESM}}$  ou  $F\text{-GACL}$  seront choisies en présence de fortes dégradations de niveaux asymétriques.



## 6.2 Performances de l'approche BCL sur données synthétiques

Dans cette partie, nous évaluons l'approche BCL (introduit à la section 5.2) en utilisant le même banc de test que précédemment et introduit à la section 6.1.1.

### 6.2.1 Équivalence des approches BCL et PBCL

La figure 6.5 présente une comparaison des nouvelles approches BCL et PBCL (introduit à la section 5.2.2) avec les approches de l'état de l'art pour trois situations typiques : aucun bruit (à gauche), l'image est bruitée et le template ne l'est pas (au milieu), du bruit dégrade les deux images (à droite).

Les approches BCL et PBCL produisent des performances identiques sur tous ces tests. Ceci confirme de manière expérimentale l'équation (5.45), et justifie le fait que dans la suite ils seront assimilés lors de la discussion.

Les performances en termes de fréquence de convergence et de vitesse de convergence (Fig. 6.5) conduisent aux mêmes conclusions : le BCL produit les meilleures performances dans les situations où du bruit est présent à différents niveaux sur l'image et le template (colonne du milieu et de droite). Cela illustre la capacité de cette approche à éliminer par projection le bruit de la matrice Jacobienne en présence d'asymétrie sur les niveaux de bruit, comme cela a été prédit par l'équation (5.62). Dans la situation non bruitée, ce dernier s'avère légèrement moins performant que l'ESM (colonne de gauche).

### 6.2.2 Impact de la projection pour les approches BCL et PBCL

Afin d'évaluer plus précisément l'impact de la projection  $\mathbf{P}_{\ominus}^{\perp}$  sur le gain en performances, nous définissons trois algorithmes synthétiques de référence supplémentaires, les approches  $\text{PBCL}_{-}$ ,  $\text{PBCL}_{\varepsilon}$  et  $\text{ESM}_{nf}$  qui utilisent toutes les trois à un certain niveau des données non-bruitées (normalement non accessibles) pour l'estimation de la transformation. Tous les algorithmes estiment  $\mathbf{v}$  en utilisant l'équation suivante :

$$\mathbf{v} = (P^{\perp}J)^{+} \mathbf{e}. \quad (6.3)$$

Une synthèse des valeurs spécifiques utilisées pour les matrices  $P^{\perp}$ ,  $J$  et  $\mathbf{e}$  par les approches synthétiques ainsi que par les approches standards est présentée dans le tableau 6.2. La figure 6.6 montre leurs performances en fonction du bruit.

L'algorithme synthétique  $\text{PBCL}_{\varepsilon}$  suppose connue la matrice Jacobienne  $J_{\varepsilon}$  du bruit  $\varepsilon$ , comme défini dans l'équation (5.13). Cette approche fournit les meilleures performances parmi les approches utilisant une erreur bruitée en éliminant par projection le bruit sur  $\mathbf{J}_{\oplus}$ . Elle produit des performances similaires à celle de  $\text{ESM}_{nf}$  sur un large intervalle de niveau de bruit, ce qui confirme que l'estimation est très sensible à la présence de bruit dans la matrice Jacobienne, et justifie donc l'approche par projection.

L'algorithme synthétique  $\text{PBCL}_{-}$  utilise la matrice de projection  $\mathbf{P}_{\ominus}^{\perp}$  non bruitée. Il produit des performances similaires à l'ESM, ce qui montre que la projection par  $\mathbf{P}_{\ominus}^{\perp}$  n'a que peu d'influences négatives sur les performances. Cette approche est seulement légèrement moins robuste que l'ESM dans des conditions de fort RSB. La matrice Jacobienne  $\mathbf{J}_{\ominus}$  n'a par conséquent



Algo	$P^\perp$	$J$	$\mathbf{e}$
ICL	$Id_N$	$J_T^{CL}$	$\mathbf{e}(0)$
ESM	$Id_N$	$\mathbf{J}_\oplus$	$\mathbf{e}(0)$
PBCL	$P_{\mathbf{J}_\ominus}^\perp$	$\mathbf{J}_\oplus$	$\mathbf{e}(0)$
BCL	$Id_N$	$[J_I^{CL} J_T^{CL}]$	$\mathbf{e}(0)$
ESM <sub><i>n<sub>f</sub></i></sub>	$Id_N$	$\mathbf{J}_{\oplus n_f}$	$\mathbf{e}_{n_f}(0)$
PBCL <sub>-</sub>	$P_{\mathbf{J}_{\oplus n_f}}^\perp$	$\mathbf{J}_\oplus$	$\mathbf{e}(0)$
PBCL <sub><math>\varepsilon</math></sub>	$P_{\mathbf{J}_\varepsilon}^\perp$	$\mathbf{J}_\oplus$	$\mathbf{e}(0)$

TABLE 6.2 – Matrice de projection  $P^\perp$ , matrice Jacobienne  $J$  et erreur initiale  $\mathbf{e}$  associées aux algorithmes définis par (6.3).  $Id_N$  désigne la matrice identité  $N \times N$ .

que peu de corrélations avec les gradients spatiaux des images, ce qui entraîne l'élimination de très peu d'informations utiles pour l'alignement.

L'algorithme PBCL produit les meilleurs résultats des approches non-synthétiques sur presque l'ensemble des niveaux de bruit étudié. Tout comme le PBCL<sub>-</sub>, il devient légèrement moins robuste que l'ESM lorsque très peu de bruit dégrade l'image. Tous ces résultats mettent en perspective l'équation (5.62) et confirment que le PBCL utilise une projection adaptée, qui perturbe légèrement les performances en présence de bruit de faible amplitude, mais qui améliore grandement les performances dès qu'une asymétrie importante apparaît sur les niveaux de bruit.

### 6.2.3 Influence de l'asymétrie des niveaux de bruit : BCL vs ACL

Nous avons vu à la section 6.1.3.3 que les approches AACL et GACL produisaient toujours des résultats au moins aussi bons que les approches de l'état de l'art et de meilleurs résultats en présence de bruits de niveaux asymétriques. L'algorithme BCL s'est révélé dans les deux sections précédentes comme une approche particulièrement efficace pour éliminer l'influence du bruit. Dans cette partie, nous comparons l'approche BCL avec les approches GACL et AACL en reprenant les expérimentations Exp. 1-4 présentées section 6.1.1.

La figure 6.7 montre la fréquence moyenne de convergence des différentes approches, en fonction de  $\beta$ .

L'approche BCL a un comportement qui varie entre le comportement de l'approche AACL<sup>ICL</sup> (resp. AACL<sup>ICL</sup>) pour  $\beta = 0$  (resp.  $\beta = 1$ ) vers le comportement de l'approche AACL<sup>ESM</sup> pour  $\beta$  proche de 0.5. Ainsi cette approche produit des résultats légèrement moins bons que l'approche AACL<sup>ESM</sup> ou encore l'ESM lorsque  $\beta$  est proche de 0.5 (Exp. 1-4) ou bien lorsque le RSB total est élevé (Exp. 1,3-4). Cela confirme les résultats observés dans les sections 6.2.1 et 6.2.2. En revanche en présence d'un RSB faible (Exp. 2) et d'un niveau d'asymétrie important, le BCL produit les meilleures performances. Dans tous les cas cette approche produit des résultats proches des résultats optimaux et se révèle donc l'approche la plus en mesure de s'adapter à la plus grande variété de niveaux de bruit.

### 6.3. Application au suivi dans des conditions de faible luminosité

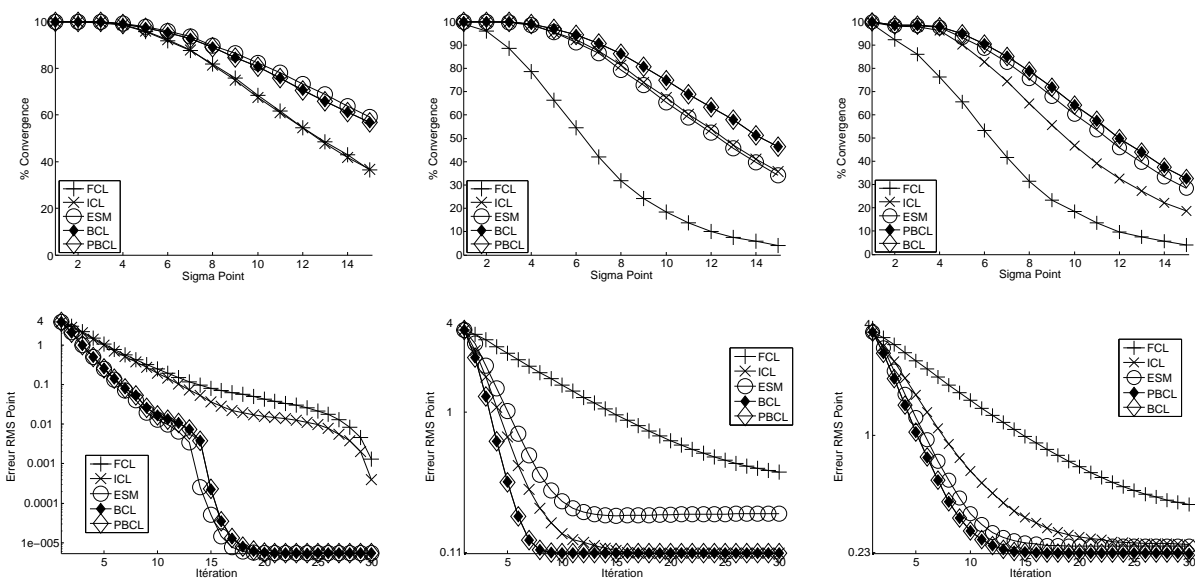


FIGURE 6.5 – Fréquence de convergence (première ligne) et vitesse de convergence en échelle logarithmique pour un Sigma Point fixé à 4 pixels (seconde ligne) pour les approches FCL, ICL, SCL, BCL et PBCL utilisant une optimisation Gauss-Newton et une règle de mise à jour Directe. A gauche :  $I$  et  $T$  sont non bruitées. Au centre : un bruit additif Gaussien de RSB=10dB a été ajouté à  $I$ ,  $T$  n'est pas bruitée. A droite : les deux images ont été dégradées, RSB=10dB pour  $I$ , RSB=15dB pour  $T$ .

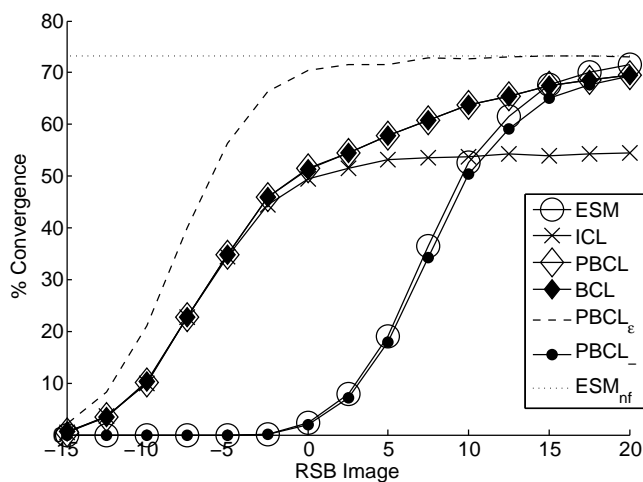


FIGURE 6.6 – Fréquence de convergence des algorithmes ESM, ICL et PBCL comparés aux approches synthétiques  $PBCL_{\epsilon}$ ,  $PBCL_{-}$  et  $ESM_{nf}$  pour un Sigma Point égal à 12 pixels et un bruit sur l'image décroissant (-15dB à 20dB). Aucun bruit n'a été ajouté au template.

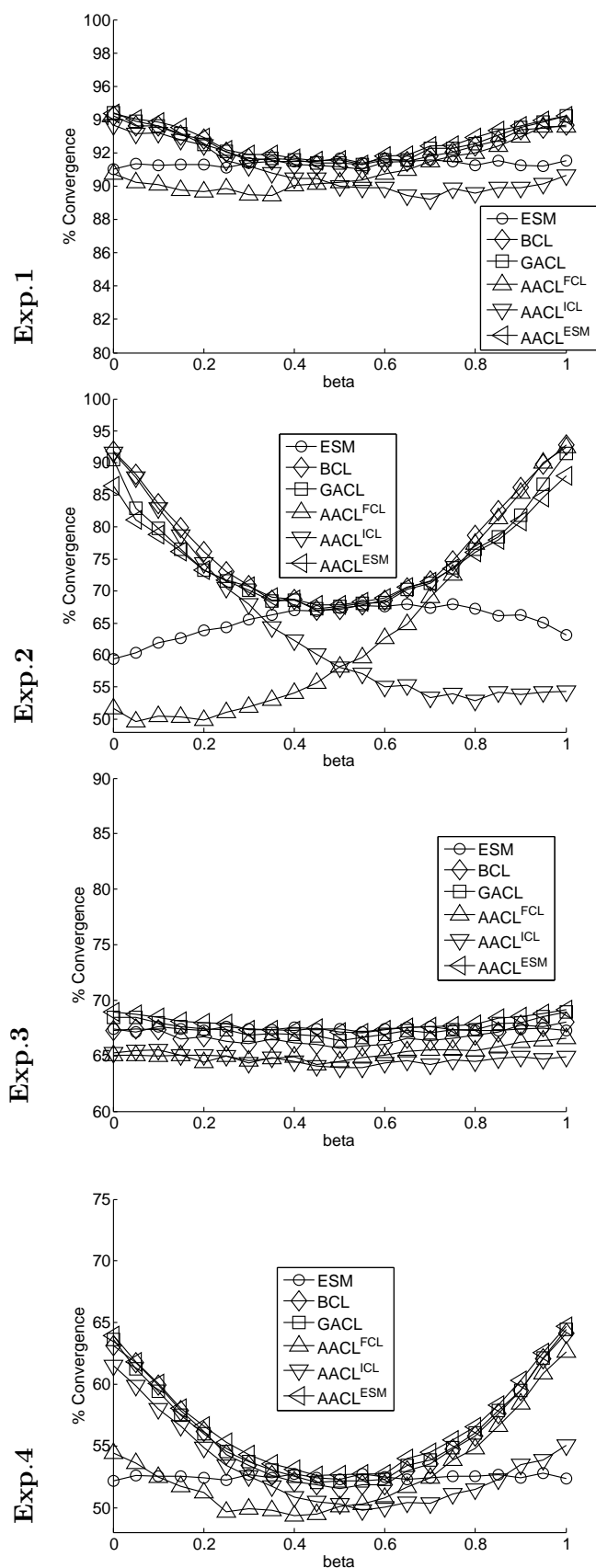


FIGURE 6.7 – Fréquence moyenne de convergence en fonction de l’asymétrie du bruit pour les approches ESM, BCL, AACL<sup>FCL</sup>, AACL<sup>ICL</sup>, AACL<sup>ESM</sup> et GACL. Exp. 1 : Sigma Point = 6 pixels, RSB=10 dB. Exp. 2 : Sigma Point = 6 pixels, RSB=5 dB. Exp. 3 : Sigma Point = 12 pixels, RSB=15 dB. Exp. 4 : Sigma Point = 12 pixels, RSB=10 dB.

## 6.3 Application au suivi dans des conditions de faible luminosité

Dans des conditions de faible luminosité, un système optique produit des images de mauvaise qualité qui peuvent être modélisées suivant les travaux de [Alter 2006] : le nombre de photons observés par un pixel suit une loi de Poisson dont le paramètre caractéristique est proportionnel à l'intensité moyenne reçue. Dans ces conditions, suivre un objet à l'aide d'approches en ligne devient une tâche difficile en raison du bruit de Poisson qui peut dramatiquement affecter les gradients des images. Afin d'améliorer les performances de suivi sur ce type de données, nous proposons ici de diminuer le bruit sur le template en moyennant plusieurs images de la vidéo pour lesquelles le mouvement de l'objet a été estimé. Ce procédé conduit à un problème d'alignement d'images asymétrique où l'image courante  $I$  est mise en correspondance avec un template  $T$  de meilleure qualité. Nous évaluons maintenant l'intérêt des nouveaux algorithmes proposés dans ce contexte.

### 6.3.1 Illustration de suivi sur des données réelles

Nous disposons d'une vidéo capturée dans des conditions réelles de basse luminosité<sup>1</sup>. La figure 6.8 illustre la nature des images observées. Ces images sont marquées par un bruit important (grain) et un faible contraste caractéristique des images prises dans des conditions de faible luminosité.

Afin d'illustrer le principe évoqué en introduction, la figure 6.9 présente les résultats des approches FCL, ICL, ESM et BCL dans des conditions réelles de suivi (30 itérations par image). Un template moyen est construit en moyennant les 5 premières images de la séquence après avoir compensé ces dernières dans le repère associé à la première image. L'approche FCL ne semble pas en mesure de suivre l'objet. L'approche ESM perd le suivi à partir de l'image 20. L'approche ICL accumule une erreur de suivi qui la conduit à décrocher à l'image 30. L'approche BCL suit l'objet jusqu'à la fin de la séquence.

Bien qu'un tel résultat ne puisse être une preuve de la supériorité de l'approche BCL, il illustre cependant, dans un cadre applicatif réaliste possible, le gain prévu lors de l'étude théorique et validé par les expérimentations précédentes. Afin d'évaluer plus en détail les différentes approches sur ce type de données, nous proposons différentes études qui sont l'objet des parties suivantes.

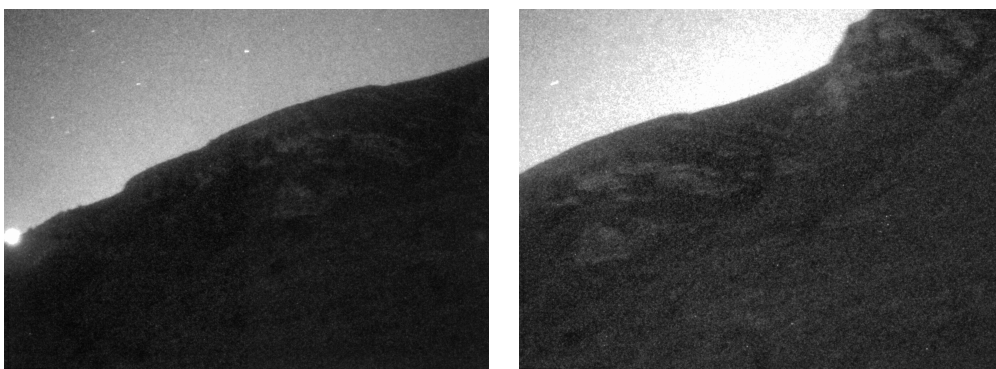


FIGURE 6.8 – Exemple d'acquisitions réelles dans des conditions de faible luminosité.

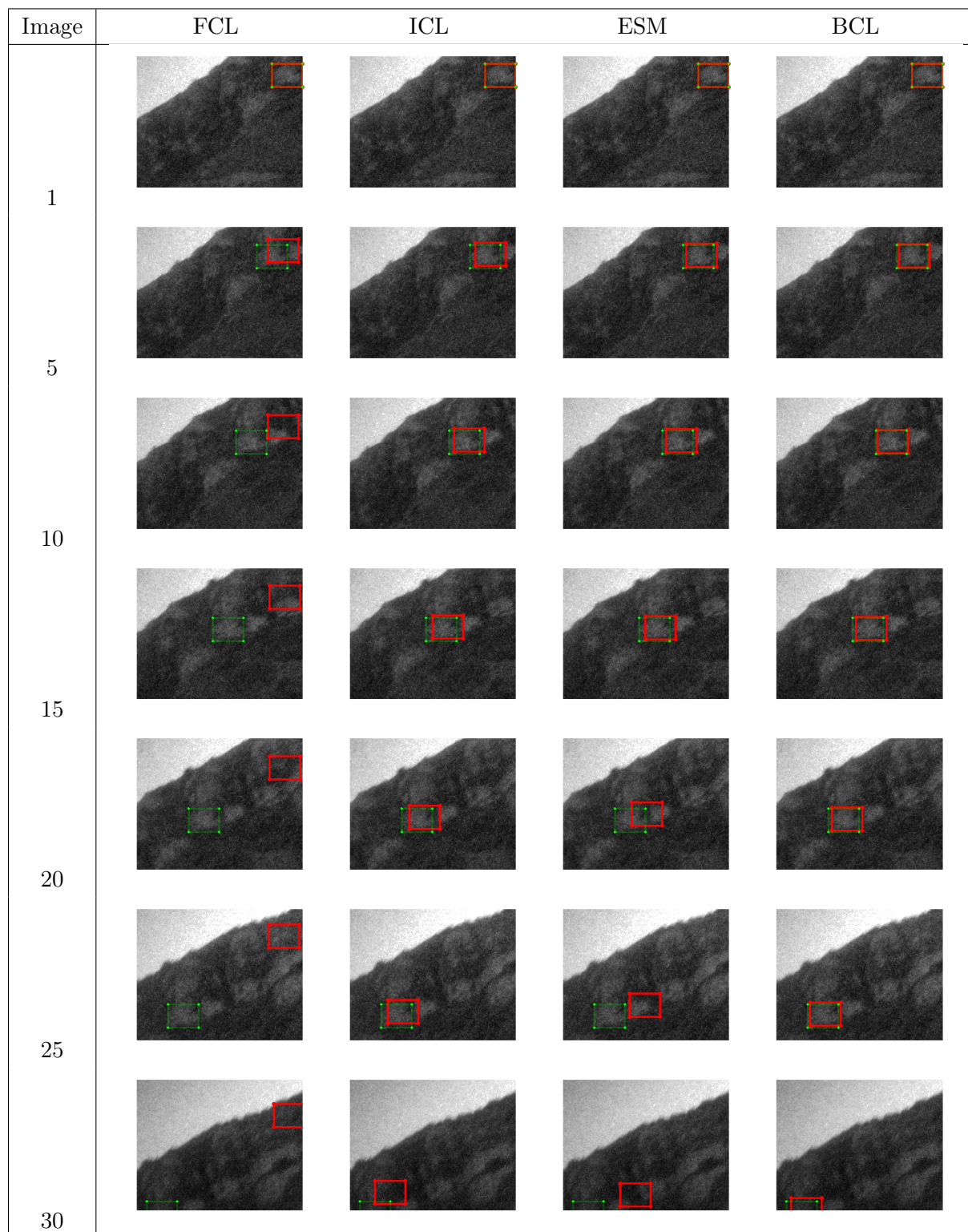


FIGURE 6.9 – Illustration de suivi dans des conditions réelles pour les approches FCL, ICL, ESM et BCL sur une vidéo acquise dans des conditions de faible luminosité. Les images affichées correspondent seulement à des portions des images de la vidéo afin de rendre plus visible les performances du suivi. Les cinq premières images de la séquence sont utilisées pour construire un template moyen. Le rectangle vert correspond à la position de l'objet prédite par la vérité de terrain. Le rectangle rouge correspond au suivi effectué par les différents algorithmes.

### 6.3.2 Analyse moyennée sur des séquences réelles

Nous ne disposons malheureusement pas de la vérité de terrain associée à la vidéo, à savoir le mouvement de la caméra entre chaque image. Nous décrivons dans la section 6.3.2.1 la construction d'une vérité de terrain approximative. Elle est utilisée aux sections 6.3.2.2 et 6.3.2.3 pour évaluer les algorithmes.

#### 6.3.2.1 Constitution d'un corpus d'images test et d'une vérité de terrain

Nous avons sélectionné dans la vidéo réelle deux séquences de 40 images et de 32 images respectivement pour lesquelles étaient visibles des objets à suivre. La figure 6.8 illustre les images utilisées.

Afin de multiplier les tests d'alignement, 3 zones distinctes portant des informations de structure ont été sélectionnées dans la seconde séquence. Nous disposons donc initialement de 136 observations de bruit réel associées à des conditions de basse luminosité (40 pour l'objet de la première séquence et  $3 \times 32$  pour les 3 objets de la seconde). Pour chacune des séquences et pour chaque objet, une vérité de terrain approximative est construite en utilisant la méthode supervisée suivante :

- L'objet d'intérêt est sélectionné dans l'image 1 de la séquence (joue le rôle du template  $T$ ) à l'aide d'une boîte englobante rectangulaire positionnée en réglant les paramètres  $\mu_T^0(1)$ .
- Les paramètres de suivi sont initialisés à  $\bar{\mu}(1) = \mu_T^0(1)$ .
- L'alignement est réalisé entre l'image  $t + 1$  et l'image 1 en initialisant le suivi avec les paramètres  $(\bar{\mu}(t), \mu_T^0)$  et en utilisant l'algorithme ESM avec 200 itérations (ce qui donne naissance aux paramètres vrais  $\bar{\mu}(t + 1)$ ). L'algorithme ESM est utilisé ici car il est le plus adapté en présence de bruit symétrique affectant les images. Le modèle de mouvement estimé est une translation.
- A l'issue de l'alignement par l'ESM, nous vérifions manuellement image par image que ce dernier a été en mesure de suivre l'objet.

La vérité de terrain  $\bar{\mu}$  ainsi constituée est approximative. Nous ne pouvons donc l'utiliser pour faire des mesures de précision. En revanche, nous pouvons l'utiliser pour comparer les différentes propriétés de convergence des algorithmes en termes de vitesse de convergence et de fréquence de convergence.

#### 6.3.2.2 Propriétés de convergence sans construction d'un template "moyen"

Dans cette partie, nous présentons les résultats obtenus en lançant les différents algorithmes sur les deux séquences sélectionnées. Afin d'obtenir des résultats statistiquement significatifs, nous moyennons les performances obtenues pour plusieurs initialisations du mouvement par image et par objet d'intérêt. Nous nous inspirons pour cela du protocole de tests sur données synthétiques présenté à la section 6.1.1. Nous calculons les sommets de la boîte englobante associé à l'objet d'intérêt dans l'image courante  $t$  à l'aide des paramètres  $\bar{\mu}(t)$ . A la position spatiale  $\mathbf{x}$  de chaque sommet est alors ajouté le même offset  $\delta\mathbf{x}$  obtenu aléatoirement selon une

---

1. Nous remercions Matthieu Grossetete de Thales Avionics pour nous avoir fourni ces données.

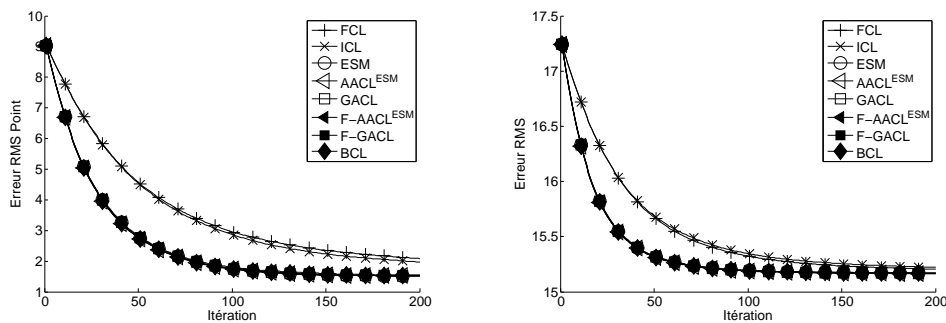


FIGURE 6.10 – Vitesse de convergence en termes d’Erreur RMS Point et d’Erreur RMS moyenne en fonction du nombre d’itérations sur une séquence d’images réelles capturée dans des conditions de basse luminosité. Seule la première image est utilisée pour construire le template.

loi de probabilité normale de moyenne nulle et d’écart-type  $\sigma_p = 10$  (Sigma Point). La nouvelle position de la boîte englobante est utilisée pour calculer les paramètres initiaux  $\mu_Y^0(t)$  utilisés pour initialiser l’alignement. Ce processus est réalisé 20 fois par image donnant naissance à  $20 \times 132 = 2640$  tests d’alignement (132 car la première image de chaque séquence est uniquement utilisée pour extraire le template  $T$ ).

Pour chaque image de la séquence, pour chaque objet et pour chaque initialisation du mouvement, les algorithmes sont évalués en utilisant un template constitué uniquement de la première image de la séquence, ce qui signifie que nous nous plaçons dans des conditions approximativement symétriques de bruit.

La figure 6.10 présente les performances des différentes approches en terme d’Erreur RMS Point et d’Erreur RMS moyenne. Nous calculons l’Erreur RMS Point en calculant la distance euclidienne moyenne entre les quatre coins de la région d’intérêt prédits par l’algorithme testé et ceux issus de la vérité de terrain. L’Erreur RMS à l’itération  $n$  est définie par :

$$E_{RMS}^n = \sqrt{\frac{1}{N} \sum_{i=1}^N (I(\mathbf{W}(\hat{\mu}_T^n, \mathbf{x}_i)) - T(\mathbf{W}(\mu_T^0, \mathbf{x}_i)))^2} \quad (6.4)$$

où  $\hat{\mu}_T^n$  est l’estimé des paramètres de mouvement à l’itération  $n$ . L’Erreur RMS moyenne est obtenue en moyennant l’ensemble des Erreurs RMS obtenues pour un même indice d’itération.

Nous constatons que l’ESM et les nouvelles approches ACL et BCL produisent de meilleurs résultats que les approches unidirectionnelles FCL et ICL ce qui confirme sur données réelles l’intérêt de l’ESM et des nouvelles approches dans des conditions symétriques de bruit.

Afin de compléter cette étude, nous introduisons ici une nouvelle méthodologie permettant d’analyser le compromis robustesse/nombre d’itérations. La figure 6.11 présente l’évolution de la fréquence de convergence en fonction du nombre d’itérations pour les différentes approches. Le seuil de convergence est fixé à 2 et 4 en Erreur RMS Point.

Cette figure confirme les meilleures performances produites par les approches bidirectionnelles (ESM, GACL, AACL<sup>ESM</sup> et BCL ainsi que les approches rapides  $F$ -GACL et  $F$ -AACL<sup>ESM</sup>) par rapport aux approches unidirectionnelles.



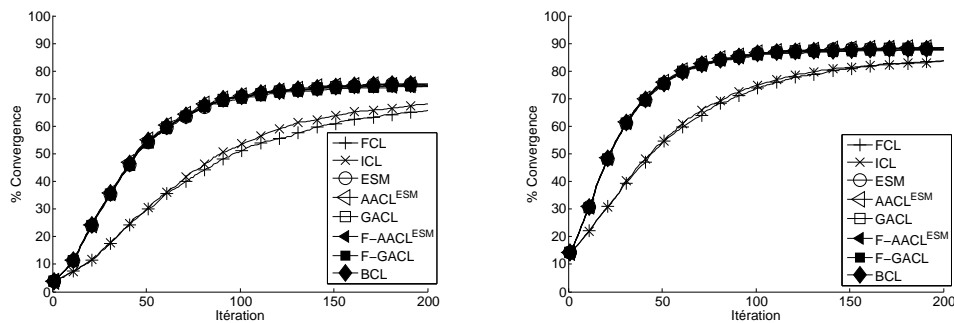


FIGURE 6.11 – Fréquence de convergence en fonction du nombre d’itérations sur une séquence d’images réelles capturée dans des conditions de basse luminosité. Seule la première image est utilisée pour construire le template. Le seuil de convergence est fixé à 2 en Erreur Point RMS à gauche et à 4 à droite.

### 6.3.2.3 Propriétés de convergence avec construction d’un template moyen

Nous utilisons la méthodologie présentée dans la partie précédente afin d’évaluer les algorithmes en présence d’asymétrie du bruit. Dans une première phase, chaque algorithme construit son template en moyennant les  $j$  premières images de la séquence après compensation de ces dernières par le mouvement estimé. A titre d’illustration, la figure 6.12 présente des exemples de template pour  $j = 1$  (aucun moyennage),  $j = 3$  et  $j = 5$  pour l’approche BCL. Alors que les contours de l’objet sont difficilement discernables pour  $j = 1$ , ces derniers apparaissent de plus en plus nettement lorsque  $j$  augmente.

Une fois la première phase réalisée, les algorithmes sont lancés sur le reste des images de la séquence en suivant la même technique d’initialisation multiple utilisée à la section précédente. La figure 6.13 présente les performances des différentes approches en termes d’Erreur RMS Point et d’Erreur RMS moyenne. Ces performances sont moyennées sur l’ensemble des données disponibles ce qui correspond à  $20 \times (136 - 4 \times j)$  tests (20 initialisations du mouvements par image, 3 objets sur 32 images + 1 objet sur 40 images- $4 \times j$ ).

En terme de vitesse de convergence, l’utilisation d’un template moyen semble améliorer les performances de l’ensemble des approches utilisant les gradients de  $T$  pour l’estimation du mouvement. Comme l’illustre le tableau 6.3, pour un nombre faible d’images utilisées, plus le nombre d’images utilisées pour construire le template moyen est important, meilleures sont les performances. De plus, les nouvelles approches AACL, GACL et BCL fournissent non seulement de meilleures performances que dans des conditions symétriques de bruit, mais de plus, elles surpassent ou égalent la meilleure approche de l’état de l’art avec l’utilisation d’un template moyen.

Nous remarquons cependant que l’erreur RMS point diminue jusqu’à atteindre un palier plus élevé avec l’utilisation d’un template moyen ( $>2$  pixels) qu’en utilisant uniquement la première image comme template (1.5 pixels). Cela laisse penser que la création d’un template par intégration de plusieurs images introduit un biais dans l’estimation. La vérité de terrain utilisée étant approximative, il est impossible de dire quelle est la meilleure stratégie de construction du template en termes de précision entre l’utilisation uniquement de la première image et le moyennage



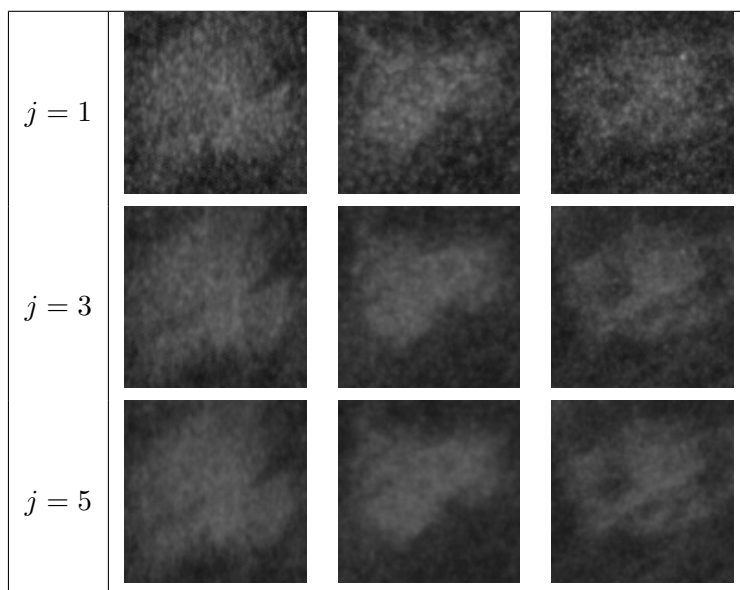


FIGURE 6.12 – Exemples de templates moyens créés par l’algorithme BCL.

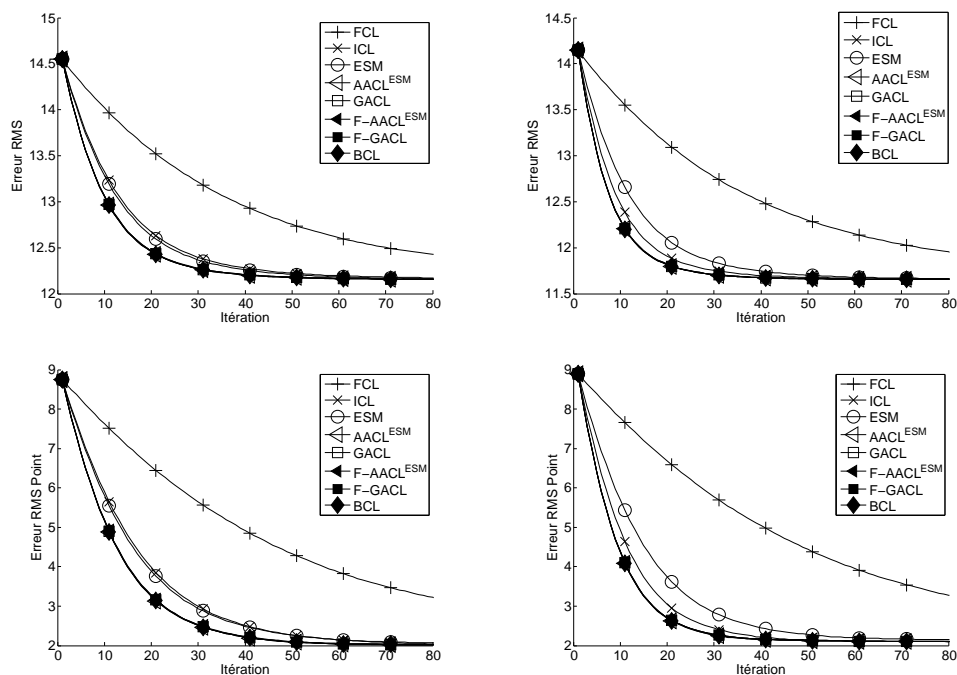


FIGURE 6.13 – Vitesse de convergence en termes d’Erreur RMS Point moyenne et d’Erreur RMS moyenne en fonction du nombre d’itérations sur une séquence d’images réelles capturée dans des conditions de basse luminosité. A gauche, les 3 premières images ont été compensées par chaque approche pour construire un template moyen. A droite, les 5 premières images ont été utilisées pour construire le template.

### 6.3. Application au suivi dans des conditions de faible luminosité

	FCL	ICL	ESM	AACL <sup>ESM</sup>	GACL	<i>F</i> -AACL <sup>ESM</sup>	<i>F</i> -GACL	BCL
$j = 1$	6.82	6.79	5.21	5.18	5.21	5.18	5.22	5.18
$j = 3$	6.54	3.98	3.89	3.25	3.27	3.27	3.28	3.25
$j = 5$	6.69	3.05	3.74	2.70	2.71	2.71	2.71	2.69

TABLE 6.3 – Erreur RMS Point des différentes approches à la 20e itération en fonction du nombre d’images utilisées pour construire le template moyen.

des premières images.

Afin de compléter cette étude, la figure 6.14 présente l’évolution de la fréquence de convergence en fonction du nombre d’itérations pour les différentes approches.

Cette figure confirme les meilleures performances produites par les approches bidirectionnelles (GACL, AACL<sup>ESM</sup> et BCL et les versions rapides) par rapport aux approches de l’état de l’art.

#### 6.3.2.4 Fréquence de convergence et temps de calcul

Les différentes approches étudiées ont un coût calculatoire par itération sensiblement différent (cf. section 5.3). Pour une mise en oeuvre pratique, il est fondamental de s’intéresser aux performances des approches pour un coût calculatoire fixé. La figure 6.15 présente l’évolution de la fréquence de convergence en fonction du temps de calcul des différentes approches, lorsque le template est fortement dégradé ( $j = 1$ ) et lorsque le template est de bien meilleure qualité que l’image courante ( $j = 5$ ). Le seuil de convergence est fixé à 4 pixels en Erreur RMS Point.

Nous constatons que les nouvelles approches proposées produisent des résultats moins bons que l’approche ESM dans des conditions symétriques de bruit (fig. 6.15 à gauche) à l’exception des approximations rapides *F*-GACL et *F*-AACL<sup>ESM</sup> qui produisent des résultats similaires. En présence d’un template de meilleure qualité (fig. 6.15 à droite), l’approche ICL semble produire les meilleurs résultats, suivie des approximations rapides *F*-GACL et *F*-AACL<sup>ESM</sup>.

Ces résultats mettent en perspective les remarques des paragraphes précédents, en montrant que l’utilisation d’un nombre d’itérations plus important pour une approche à faible coût calculatoire peut compenser en partie l’utilisation d’une matrice Jacobienne bruitée. Dans un contexte où les niveaux de bruit dans les images n’est pas maîtrisé précisément, il est cependant difficile de choisir parmi les approches de l’état de l’art. Les approximations rapides *F*-AACL<sup>ESM</sup> et *F*-GACL ont quant à elles l’avantage de produire toujours des résultats satisfaisants pour un coût calculatoire fixé.

Le temps de calcul des approches est très dépendant des choix d’implémentation et de l’architecture matérielle sur laquelle les algorithmes sont mis en oeuvre. Sur une architecture parallèle, il est possible d’envisager le développement efficace des approches AACL et GACL en parallélisant le calcul de  $\alpha$  et la résolution Gauss-Newton du système (5.8). Ceci permettrait la mise en place d’algorithmes possédant un coût calculatoire semblable à l’ESM tout en bénéficiant pleinement du calcul automatique de  $\alpha$  (mis à jour plusieurs fois au cours d’un cycle itératif).

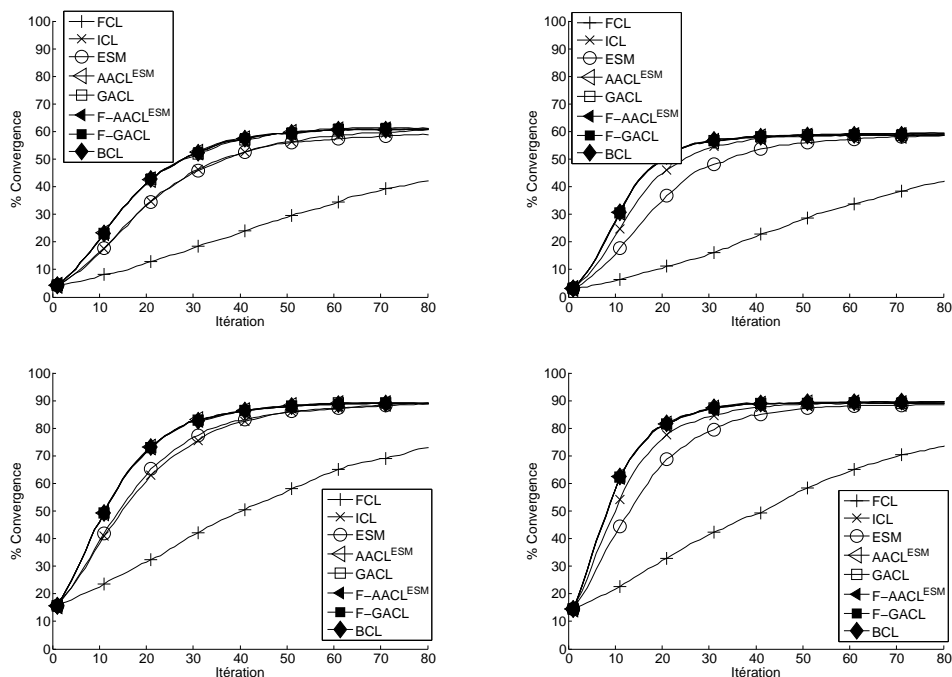


FIGURE 6.14 – Fréquence de convergence en fonction du nombre d’itérations sur une séquence d’images réelles capturée dans des conditions de basse luminosité. A gauche, les 3 premières images sont utilisées pour construire le template. A droite, les 5 premières sont utilisées. Sur la première ligne, le seuil de convergence est fixé à 2 pixels ce qui correspond à peu près au pallier atteint par les approches en moyenne. Sur la seconde ligne, le seuil est fixé à 4 pixels.

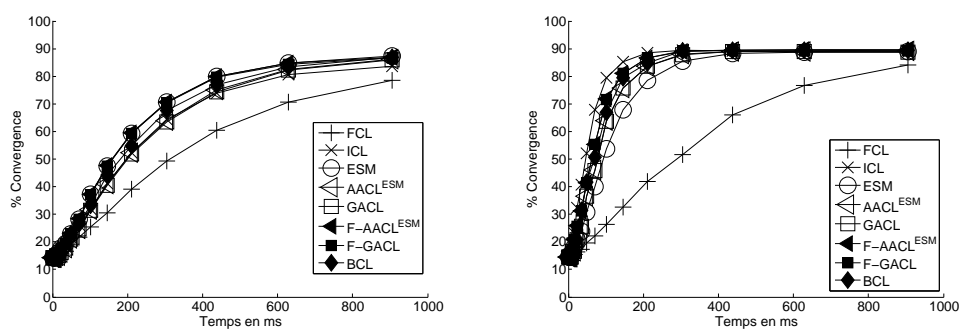


FIGURE 6.15 – Fréquence de convergence en fonction du temps de calcul sur une séquence d’images réelles capturée dans des conditions de basse luminosité. A gauche, seule la première image est utilisée comme template. A droite, les 5 premières sont utilisées. Le seuil de convergence est fixé à 4 pixels sur l’erreur RMS Point.

#### 6.3.3 Analyse statistique sur des séquences semi-synthétiques

Il n'existe pas à notre connaissance de vidéos publiques prises dans des conditions de basse luminosité et pour lesquelles une vérité de terrain est disponible. Afin de poursuivre notre validation sur un plus grand nombre de réalisations de bruit, nous avons simulé des acquisitions basse luminosité en dégradant des vidéos libres d'accès (figure 6.16) prises dans des conditions standards [Lieberknecht 2009]. La procédure de dégradation consiste à assigner à chaque pixel  $I(\mathbf{x})$  une valeur aléatoire tirée d'une distribution de Poisson de moyenne  $(a \cdot I(\mathbf{x}) + b)$ . Les facteurs  $a$  et  $b$  sont choisis de telle sorte que la nouvelle intensité de l'image au pixel  $\mathbf{x}$  soit comprise dans l'intervalle  $[1, 10]$ .

La vérité de terrain est construite à partir de la séquence vidéo initiale qui est cette fois de bonne qualité. La région d'intérêt est définie sur la première image de la vidéo par une boîte englobante positionnée par les paramètres  $\mu_T^0(1)$ , et son contenu utilisé comme première image template  $T$ . L'algorithme ESM est utilisé afin de suivre la région d'intérêt, en utilisant le même modèle de transformation homographique que dans la section 6.1.1. Cette estimation du mouvement est utilisée comme vérité de terrain (paramètres vrais  $\bar{\mu}(t)$ ). Nous avons vérifié image par image que cette estimation était correcte.

Les tests sont effectués sur les séquences dégradées comme suit. L'analyse de performances est effectuée sur les algorithmes FCL, ICL et ESM pour les approches standards ainsi que sur les approches GACL, AACL<sup>ESM</sup>,  $F$ -GACL,  $F$ -AACL<sup>ESM</sup> et BCL pour les nouvelles approches. Les algorithmes sont limités à 40 iterations. Le template est calculé à partir des  $j$  premières images de la vidéo en moyennant les différentes images compensées. L'estimation de mouvement est réalisée de l'image  $t = j + 1$  à 100 en utilisant le template moyen et l'image courante dégradée : pour chaque image de la séquence, les algorithmes d'alignement sont réinitialisés avec la vérité de terrain de l'image précédente  $\mu_T^n(t + 1) = \bar{\mu}(t)$  et  $\mu_T^0(t + 1) = \mu_T^0(1)$ . Le mouvement à estimer correspond donc au mouvement réel de l'objet suivi. Nous calculons l'Erreur RMS Point en comparant les quatre coins de la région d'intérêt prédits par l'algorithme évalué et ceux issus de la vérité de terrain. Afin d'obtenir des résultats significatifs, nous avons moyenné les résultats obtenus sur 6 séquences vidéo issues de l'ensemble "Fast Far" de [Lieberknecht 2009], et généré 20 séquences différentes dégradées à partir de chaque séquence haute qualité initiale. Nous avons testé les configurations  $j = 5$  et  $j = 9$  conduisant respectivement à un total de 11400 et 10920 estimations du mouvement pour chaque approche évaluée.

La figure 6.17 présente la fonction de répartition de l'Erreur RMS Point obtenue par les différentes approches pour  $j = 5$  et  $j = 9$ , qui caractérise la fréquence de convergence obtenue pour différentes valeurs de seuil de convergence fixé par une Erreur RMS Point  $\tau$ . Les statistiques de ces distributions pour  $\tau = 3 \text{ pixels}$  sont résumées dans les tableaux 6.4 et 6.5.

L'approche BCL fournit les meilleures performances en conjuguant la plus petite Erreur RMS Point moyenne, le plus petit écart-type, et la meilleure fréquence de convergence quel que soit le seuil  $\tau$  employé. En termes de fréquence de convergence à  $\tau = 3 \text{ pixels}$ , cette approche est suivie par les algorithmes AACL<sup>ESM</sup> et GACL ainsi que par les approximations rapides de ces approches. Ensuite vient l'ESM puis l'ICL. En terme de précision moyenne, l'algorithme ICL est meilleur que l'ESM pour  $j = 9$  et semble en mesure de concurrencer les approches

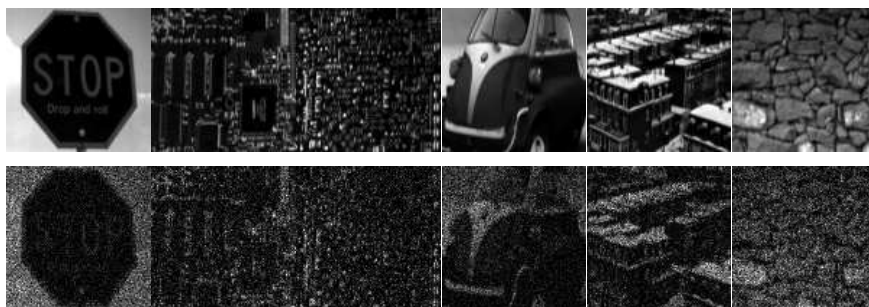


FIGURE 6.16 – Images utilisées afin de simuler les conditions de basse luminosité dans les expériences du paragraphe 6.3. Première ligne : images extraites des séquences vidéos de haute qualité issues de l'ensemble "Fast Far" de [Lieberknecht 2009] et disponible sur <http://www.metaio.com/research>. Seconde ligne : acquisitions simulées dans des conditions de basse luminosité.

<i>Algo</i>	FCL	ICL	ESM	AACL <sup>ESM</sup>	GACL	F-AACL <sup>ESM</sup>	F-GACL	BCL	<i>Init</i>
<i>freq. conv.</i>	93.2%	96.3%	98.4%	98.6%	98.5%	98.5%	98.3%	<b>98.8%</b>	
<i>mean</i>	1.0604	0.6000	0.5314	0.4642	0.4847	0.4691	0.4855	<b>0.3981</b>	1.6842
<i>std</i>	0.5893	0.3470	0.3323	0.3152	0.3145	0.3230	0.3422	<b>0.2833</b>	0.7356
$\frac{std}{\sqrt{nb_{test}}}$	0.0057	0.0033	0.0031	0.0030	0.0030	0.0030	0.0032	0.0027	

TABLE 6.4 – Fréquence de convergence et précision exprimée par la moyenne et l'écart-type de l'erreur RMS Point correspondant aux expériences du paragraphe 6.3. Seulement les tests qui ont convergé pour un seuil de  $\tau = 3$  pixels sont pris en compte. Les statistiques de l'Erreur RMS Point initiale sont également fournies comme références. Le template moyen a été construit avec  $j = 5$  images.

<i>Algo</i>	FCL	ICL	ESM	AACL <sup>ESM</sup>	GACL	F-AACL <sup>ESM</sup>	F-GACL	BCL	<i>Init</i>
<i>freq. conv.</i>	92.7%	96.9%	98.3%	98.8%	98.7%	98.6%	98.4%	<b>98.9%</b>	
<i>mean</i>	1.1530	0.5280	0.6061	0.4875	0.5254	0.4821	0.5094	<b>0.3730</b>	1.7485
<i>std</i>	0.5844	0.2829	0.3435	0.3238	0.3196	0.3122	0.3371	<b>0.2473</b>	0.7108
$\frac{std}{\sqrt{nb_{test}}}$	0.0058	0.0027	0.0033	0.0031	0.0031	0.0030	0.0033	0.0024	

TABLE 6.5 – Fréquence de convergence et précision exprimée par la moyenne et l'écart-type de l'erreur RMS Point correspondant aux expériences du paragraphe 6.3. Seulement les tests qui ont convergé pour un seuil de  $\tau = 3$  pixels sont pris en compte. Les statistiques de l'erreur RMS Point initiale sont également fournies comme références. Le template moyen a été construit avec  $j = 9$  images.

### 6.3. Application au suivi dans des conditions de faible luminosité

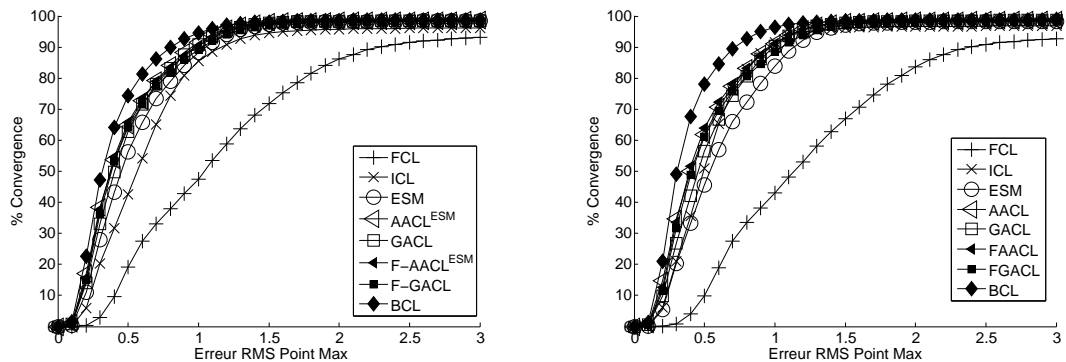


FIGURE 6.17 – Fréquence de convergence des expérimentation du paragraphe 6.3 en fonction du seuil  $\tau$  sur l’Erreur RMS Point. A gauche,  $j = 5$  images sont utilisées pour construire le template moyen. A droite,  $j = 9$  images sont utilisées.

asymétriques GACL et  $AACL^{ESM}$ . Pour  $j = 5$ , c’est l’ESM qui surpasse l’ICL et s’approche des performances des approches AAACL et GACL. En estimant les gradients uniquement à partir de l’image dégradée  $I$ , l’approche FCL produit les plus mauvais résultats dans toutes les situations présentées. Ces résultats prometteurs des approches BCL, GACL et  $AACL^{ESM}$  illustrent les perspectives d’application des approches proposées afin d’améliorer l’alignement d’images à partir de données réelles de faibles RSB.



# Conclusion

---

## 7.1 Bilan

Dans cette thèse, nous avons étudié les algorithmes d’alignement reposant sur une optimisation itérative pour le suivi d’objets. Nos contributions sont de plusieurs ordres :

- Nous avons proposé un cadre théorique bidirectionnel unifiant les approches existantes. Dans notre formalisme, une nouvelle classe d’approches, que nous nommons famille Bidirectionnelle, vient compléter les familles unidirectionnelles Directe (*Forward*) et Inverse (*Inverse*) introduites par [Baker 2004]. Cette nouvelle famille, pour laquelle peu d’approches ont été proposées, offre une plus grande souplesse dans l’utilisation des gradients lors de l’optimisation locale. Cet effort de formalisation nous a permis de tisser des liens forts entre différentes approches reposant sur l’optimisation Gauss-Newton d’une erreur quadratique.
- Nous avons présenté deux nouvelles approches pour le suivi d’objets fondées sur les paramétrisations incrémentales ACL et BCL. Nous avons développé une analyse théorique permettant de justifier le meilleur comportement de ces dernières en présence d’images affectées par des bruits de niveaux différents. Ces approches ont permis de dériver un certain nombre d’algorithmes. L’approche ACL généralise les approches FCL, ICL et SCL/ESM existantes. Elle a donné lieu à plusieurs mises en oeuvre algorithmiques utilisant une optimisation de Gauss-Newton : la méthode  $\alpha$ -ACL pour  $\alpha$  fixe, la méthode MVACL prenant en compte les niveaux de bruit supposés connus, les méthodes GACL et AACL pour une estimation adaptative du facteur  $\alpha$ , et des variantes plus rapides *F*-GACL et *F*-AACL. L’approche BCL est une formulation originale, développée sous la forme de l’algorithme BCL et de sa variante PBCL.
- Des expériences sur données synthétiques ont été proposées pour évaluer les approches de l’état de l’art et les nouvelles approches. Des méthodologies ont été en particulier proposées pour contrôler les niveaux de bruit au sein des évaluations. Ces expériences ont confirmé l’intérêt des nouvelles approches en présence de niveaux de bruit asymétriques. Nous avons enfin illustré l’intérêt des nouvelles approches pour le suivi d’objet dans des conditions de faible luminosité en testant ces dernières sur des acquisitions réelles.

## 7.2 Intérêt

Dans beaucoup d’applications de suivi, l’objet d’intérêt est sélectionné dans la première image de la vidéo. Alors que dans la situation initiale, l’alignement consiste à mettre en correspondance des images de même niveau de bruit, la création d’un template moyen par intégration temporelle permet l’émergence d’un modèle de haute qualité. Dans des situations d’images fortement



bruitées, les approches inverses comme l'ICL deviennent alors très compétitives pour le suivi en associant une estimation fiable des gradients et un coût calculatoire réduit. L'obtention du modèle de haute qualité nécessite la mise en correspondance d'images affectées par des niveaux de bruit asymétriques. Nous pensons que les approches proposées dans ce mémoire sont alors les plus adaptées. Nous pouvons donc imaginer des algorithmes d'alignement hybridant différentes approches et reposant sur deux phases :

- construction du template moyen : utilisation de l'approche BCL ou des approches rapides  $F$ -AACL<sup>ESM</sup> et  $F$ -GACL sur les premières images de la vidéo.
- utilisation de l'approche ICL pour le suivi.

L'apparence de l'objet pouvant varier au fur et à mesure du suivi, une mise à jour du template régulière peut être nécessaire. Le couplage de la procédure décrite précédemment avec les approches proposées dans [Matthews 2004a, Schreiber 2007] peut donc être envisagé. Il est également possible d'adapter la fonctionnelle d'erreur afin de tenir compte des variations de luminosité [Evangelidis 2008, Bartoli 2008], ou bien des occultations [Hager 1998, Jurie 2002b].

### 7.3 Perspective

Nous avons illustré l'intérêt des approches ACL et BCL sur des images acquises dans des conditions de faible luminosité. Nous pensons qu'il existe d'autres situations où les niveaux de bruit mis en jeu sont suffisants pour que les approches discutées apportent un gain significatif. Nous pensons notamment à la mise en correspondance d'images SAR lorsque aucune information auxiliaire (orbite du satellite, DEM (*Digital Elevation Model*)) n'est disponible [Suri 2010].

Le formalisme proposé a permis de mettre en place des approches améliorant les performances d'alignement dans des situations d'images dégradées par du bruit. Dans le contexte de mise en correspondance d'images de résolutions différentes, la qualité du contenu de chaque image peut également influencer sur les performances [Dedeoglu 2007]. Dans le cas de la super-résolution d'images très bruitées, il apparaît nécessaire de modéliser l'effet combiné des asymétries liées à la résolution et de celles liées au bruit. Un cadre formel bidirectionnel semble naturellement s'appliquer pour la prise en compte de ces deux aspects.

Nous avons consacré cette thèse à l'alignement entre deux images. La généralisation des concepts introduits à l'alignement de plusieurs images [Pires 2005] semble une piste prometteuse. La formalisation de la règle de mise à jour identité présentée dans ce mémoire permet la définition d'algorithmes respectant la consistance inverse (*Inverse Consistent*) [Christensen 1999, Cachier 2000] particulièrement adaptés dans les applications d'alignement d'images médicales [Farsiu 2005b] ou la construction de panorama. L'hybridation de telles approches avec l'imposition d'une cohérence géométrique dans l'alignement doit pouvoir permettre d'améliorer la qualité de l'alignement notamment en présence d'images de qualités différentes.

# Définition d'une erreur pour l'approche IA

Dans cette partie, nous montrons que l'application de l'algorithme Gauss-Newton à l'erreur définie par la paramétrisation incrémentale IAD :

$$e_i^{\text{IAD}}(\mathbf{v}) = I(\mathbf{W}(\boldsymbol{\mu}_T^n, \mathbf{x}_i)) - T(\mathbf{W}(\boldsymbol{\mu}_T^n \circ (\boldsymbol{\mu}_T^n + \mathbf{v})^{-1} \circ \boldsymbol{\mu}_T^n, \mathbf{x}_i)), \quad (\text{A.1})$$

conduit naturellement à l'approche *Inverse Additive Direct* nommée *Inverse Additive* IA par [Baker 2004] et initialement proposée dans [Hager 1998]. La démonstration repose sur les deux lemmes suivant :

**Lemme 1** (Effet de l'inversion de paramètre sur les gradients de la transformation.).

$$\left. \frac{\partial \mathbf{W}(\boldsymbol{\mu}, \mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}_i} = \left( \left. \frac{\partial \mathbf{W}(\boldsymbol{\mu}^{-1}, \mathbf{z})}{\partial \mathbf{z}} \right|_{\mathbf{z}=\mathbf{W}(\boldsymbol{\mu}, \mathbf{x}_i)} \right)^{-1}. \quad (\text{A.2})$$

*Démonstration.*

$$\begin{aligned} \left. \frac{\partial \mathbf{W}(\boldsymbol{\mu}^{-1} \circ \boldsymbol{\mu}, \mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}_i} &= \text{Id} \\ &= \left. \frac{\partial \mathbf{W}(\boldsymbol{\mu}^{-1}, \mathbf{z})}{\partial \mathbf{z}} \right|_{\mathbf{z}=\mathbf{W}(\boldsymbol{\mu}, \mathbf{x}_i)} \left. \frac{\partial \mathbf{W}(\boldsymbol{\mu}, \mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}_i} \end{aligned} \quad (\text{A.3})$$

□

**Lemme 2** (Effet de l'inversion de paramètre sur les gradients de la transformation vis à vis de ces paramètres.).

$$\left. \frac{\partial \mathbf{W}(\mathbf{v}^{-1}, \mathbf{x}_i)}{\partial \mathbf{v}} \right|_{\mathbf{0}} = - \left. \frac{\partial \mathbf{W}(\mathbf{v}, \mathbf{x}_i)}{\partial \mathbf{v}} \right|_{\mathbf{0}}. \quad (\text{A.4})$$

*Démonstration.* Nous avons :

$$\mathbf{W}(\mathbf{v}^{-1} \circ \mathbf{v}, \mathbf{x}_i) = \mathbf{W}(\mathbf{v}^{-1}, \mathbf{W}(\mathbf{v}, \mathbf{x}_i)) = \mathbf{g} \circ \mathbf{f}(\mathbf{v}) \quad (\text{A.5})$$

avec :

$$\begin{aligned} \mathbb{R}^p &\longmapsto \mathcal{P} \times \mathbb{R}^2 \\ \mathbf{f} : \mathbf{v} &\longmapsto \begin{pmatrix} \mathbf{v} \\ \mathbf{z} = \mathbf{W}(\mathbf{v}, \mathbf{x}_i) \end{pmatrix} \end{aligned} \quad (\text{A.6})$$

$$\begin{aligned} \mathcal{P} \times \mathbb{R}^2 &\longmapsto \mathbb{R}^2 \\ \mathbf{g} : \begin{pmatrix} \mathbf{v} \\ \mathbf{z} \end{pmatrix} &\longmapsto \mathbf{W}(\mathbf{v}^{-1}, \mathbf{z}) \end{aligned} \quad (\text{A.7})$$

Nous avons par conséquent :

$$\begin{aligned} \left. \frac{\partial \mathbf{W}(\mathbf{v}^{-1} \circ \mathbf{v}, \mathbf{x}_i)}{\partial \mathbf{v}} \right|_{\mathbf{0}} &= \mathbf{0} \\ &= \left. \frac{\partial \mathbf{g} \circ \mathbf{f}(\mathbf{v})}{\partial \mathbf{v}} \right|_{\mathbf{0}} \\ &= \left. \frac{\partial \mathbf{g}(\mathbf{f})}{\partial \mathbf{f}} \right|_{\mathbf{f}=(\mathbf{0}, \mathbf{x}_i)} \left. \frac{\partial \mathbf{f}(\mathbf{v})}{\partial \mathbf{v}} \right|_{\mathbf{0}} \\ &= \begin{bmatrix} \left. \frac{\partial \mathbf{W}(\mathbf{v}^{-1}, \mathbf{x}_i)}{\partial \mathbf{v}} \right|_{\mathbf{0}} & \left. \frac{\partial \mathbf{W}(\mathbf{0}, \mathbf{z})}{\partial \mathbf{z}} \right|_{\mathbf{x}_i} \end{bmatrix} \begin{bmatrix} \left. \frac{\partial \mathbf{v}}{\partial \mathbf{v}} \right|_{\mathbf{0}} \\ \left. \frac{\partial \mathbf{W}(\mathbf{v}, \mathbf{x}_i)}{\partial \mathbf{v}} \right|_{\mathbf{0}} \end{bmatrix} \\ &= \left. \frac{\partial \mathbf{W}(\mathbf{v}^{-1}, \mathbf{x}_i)}{\partial \mathbf{v}} \right|_{\mathbf{0}} + \left. \frac{\partial \mathbf{W}(\mathbf{v}, \mathbf{x}_i)}{\partial \mathbf{v}} \right|_{\mathbf{0}}. \end{aligned} \quad (\text{A.8})$$

□

**Theorème 3** (Expression factorisée de  $\mathbf{J}^{\text{IAD}}(\mathbf{0})$ ).

$$\mathbf{J}_e^{\text{IAD}}(\mathbf{0}) = \nabla T(\mathbf{W}(\boldsymbol{\mu}_T^n, \mathbf{x}_i)) \left( \left. \frac{\partial \mathbf{W}(\boldsymbol{\mu}_I^n, \mathbf{z})}{\partial \mathbf{z}} \right|_{\mathbf{x}_i} \right)^{-1} \left. \frac{\partial \mathbf{W}(\boldsymbol{\mu}, \mathbf{x}_i)}{\partial \boldsymbol{\mu}} \right|_{\boldsymbol{\mu}_I^n}. \quad (\text{A.9})$$

avec :

$$\nabla T(\mathbf{W}(\boldsymbol{\mu}_T^n, \mathbf{x}_i)) = \left. \frac{\partial T(\mathbf{W}(\boldsymbol{\mu}_T^n, \mathbf{z}))}{\partial \mathbf{z}} \right|_{\mathbf{x}_i}. \quad (\text{A.10})$$

*Démonstration.* Considérons la matrice Jacobienne associée à l'erreur IAD (A.1) (cf. section 4.2.2.3) :

$$\begin{aligned} \mathbf{J}_e^{\text{IAD}}(\mathbf{0}) &= - \left. \frac{\partial T(\mathbf{W}(\boldsymbol{\mu}_T^n \circ (\boldsymbol{\mu}_I^n + \mathbf{v})^{-1} \circ \boldsymbol{\mu}_I^n, \mathbf{x}_i))}{\partial \mathbf{v}} \right|_{\mathbf{0}} \\ &= - \nabla T(\mathbf{W}(\boldsymbol{\mu}_T^n, \mathbf{x}_i)) \left. \frac{\partial \mathbf{W}((\boldsymbol{\mu}_I^n + \mathbf{v})^{-1} \circ \boldsymbol{\mu}_I^n, \mathbf{x}_i)}{\partial \mathbf{v}} \right|_{\mathbf{0}} \\ &= - \nabla T(\mathbf{W}(\boldsymbol{\mu}_T^n, \mathbf{x}_i)) \left. \frac{\partial \mathbf{W}(\boldsymbol{\mu}, \mathbf{x}_i)}{\partial \boldsymbol{\mu}} \right|_{\mathbf{0}} \left. \frac{\partial (\boldsymbol{\mu}_I^n + \mathbf{v})^{-1} \circ \boldsymbol{\mu}_I^n}{\partial \mathbf{v}} \right|_{\mathbf{0}} \end{aligned} \quad (\text{A.11})$$

En appliquant le lemme A.4, il vient :

$$\begin{aligned} \mathbf{J}_e^{\text{IAD}}(\mathbf{0}) &= \nabla T(\mathbf{W}(\boldsymbol{\mu}_T^n, \mathbf{x}_i)) \left. \frac{\partial \mathbf{W}(\boldsymbol{\mu}^{-1}, \mathbf{x}_i)}{\partial \boldsymbol{\mu}} \right|_{\mathbf{0}} \left. \frac{\partial (\boldsymbol{\mu}_I^n + \mathbf{v})^{-1} \circ \boldsymbol{\mu}_I^n}{\partial \mathbf{v}} \right|_{\mathbf{0}} \\ &= \nabla T(\mathbf{W}(\boldsymbol{\mu}_T^n, \mathbf{x}_i)) \left. \frac{\partial \mathbf{W}((\boldsymbol{\mu}_I^n)^{-1} \circ (\boldsymbol{\mu}_I^n + \mathbf{v}), \mathbf{x}_i)}{\partial \mathbf{v}} \right|_{\mathbf{0}} \\ &= \nabla T(\mathbf{W}(\boldsymbol{\mu}_T^n, \mathbf{x}_i)) \left. \frac{\partial \mathbf{W}((\boldsymbol{\mu}_I^n)^{-1}, \mathbf{W}(\boldsymbol{\mu}_I^n + \mathbf{v}, \mathbf{x}_i))}{\partial \mathbf{v}} \right|_{\mathbf{0}} \\ &= \nabla T(\mathbf{W}(\boldsymbol{\mu}_T^n, \mathbf{x}_i)) \left. \frac{\partial \mathbf{W}((\boldsymbol{\mu}_I^n)^{-1}, \mathbf{z})}{\partial \mathbf{z}} \right|_{\mathbf{W}(\boldsymbol{\mu}_I^n, \mathbf{x}_i)} \left. \frac{\partial \mathbf{W}(\boldsymbol{\mu}_I^n + \mathbf{v}, \mathbf{x}_i)}{\partial \mathbf{v}} \right|_{\mathbf{0}}. \end{aligned} \quad (\text{A.12})$$

En appliquant le lemme A.2, il vient :

□

---

L'expression obtenue généralise l'expression (2.52) de l'approche IA de [Hager 1998] en introduisant les paramètres initiaux sur le template  $\boldsymbol{\mu}_T^n$ . Les deux expressions sont rigoureusement exactes si  $\boldsymbol{\mu}_T^n = \mathbf{0}$ . De plus, les deux approches utilisent la même règle de mise à jour. Ainsi l'algorithme de [Hager 1998] consiste de manière générale en la minimisation de l'erreur IAD notamment lorsque l'égalité des images à une compensation près (équation (2.51)) n'est pas vraie.



# Procédure de traduction des gradients de Brooks

---

La procédure développée dans [Brooks 2009] est fortement liée au formalisme que nous proposons et repose sur les mêmes hypothèses. En effet, les auteurs considèrent que :

- appliquer une déformation sur l'image  $I$  est équivalent à appliquer la déformation inverse sur  $T$ .
- $\mathbf{W}$  est une action de groupe.

Afin de pouvoir en faire une description rigoureuse et précise, nous pensons que la formalisation proposée au chapitre 3 est nécessaire, ce qui explique que cette partie ait été déplacée en annexe.

## B.1 Principe général

les auteurs s'intéressent à la solution de :

$$\mathbf{v}^* = \arg \min_{\mathbf{v}} (E(\boldsymbol{\mu}_I(\boldsymbol{\mu}_I^n, \mathbf{v}), \mathbf{0})) \quad (\text{B.1})$$

utilisant la règle de mise à jour :

$$\boldsymbol{\mu}_I^{n+1} = \boldsymbol{\mu}_I(\boldsymbol{\mu}_I^n, \mathbf{v}) \quad (\text{B.2})$$

Ils souhaitent profiter des avantages des approches IC et ESM en utilisant les gradients de  $T$  pour minimiser la fonctionnelle (B.1) tout en s'affranchissant des contraintes spécifiques de chaque approche à savoir :

- les approches IC et ESM reposent sur une paramétrisation incrémentale compositionnelle.
- l'approche IC nécessite une règle de mise à jour transformant des paramètres appliqués à  $T$  en des paramètres appliqués à  $I$ . Cette règle de mise à jour n'est pas usuelle.
- l'algorithme ESM nécessite une paramétrisation sur groupe de Lie.

Il s'agit donc de trouver une solution pour que les gradients de  $T$  soit interprétable relativement à la règle de mise à jour (B.2). Pour cela, les auteurs expriment le fait qu'il est équivalent de considérer que  $I$  est mobile par rapport à  $T$  ou bien que  $T$  est mobile par rapport à  $I$ , c'est à dire que la paramétrisation incrémentale  $\boldsymbol{\mu}_I = \boldsymbol{\mu}_I(\boldsymbol{\mu}_I^n, \mathbf{v})$  (qui exprime la déformation de l'image  $I$  vers le repère du template  $T$ ) peut s'exprimer relativement à un mouvement inverse  $\boldsymbol{\mu}_T$  du template  $T$  vers l'image  $I$  :

$$\boldsymbol{\mu}_I(\boldsymbol{\mu}_I^n, \mathbf{v}) = \boldsymbol{\mu}_I^n \circ \boldsymbol{\mu}_T^{-1} \quad (\text{B.3})$$

Ceci permet de paramétrer le mouvement relatif du template par rapport à l'image en fonction de l'incrément  $\mathbf{v}$  apparaissant dans la règle de mise à jour B.2 :

$$\boldsymbol{\mu}_T(\mathbf{v}) = \boldsymbol{\mu}_I(\boldsymbol{\mu}_I^n, \mathbf{v})^{-1} \circ \boldsymbol{\mu}_I^n \quad (\text{B.4})$$

Nous pouvons interpréter l'équation (B.4) comme la reparamétrisation de l'erreur inverse  $E(\boldsymbol{\mu}_I^n, \boldsymbol{\mu}_T)$  en  $E(\boldsymbol{\mu}_I^n, \boldsymbol{\mu}_I^{-1} \circ \boldsymbol{\mu}_I^n)$ . Les auteurs considèrent alors que cette dernière est équivalente à l'erreur directe  $E(\boldsymbol{\mu}_I, \mathbf{0})$ .

## B.2 Équivalence à l'ordre 1 de la règle de mise à jour de l'approche GESM avec la règle de mise à jour de l'approche FA+IAD

Nous montrons dans cette section que la règle de mise à jour employée par les auteurs pour l'approche GESM :

$$\boldsymbol{\mu}_I^{n+1} = \boldsymbol{\mu}_I^n + \mathbf{v}^{-1}, \quad (\text{B.5})$$

correspond à une approximation à l'ordre 1 de la règle de mise à jour exacte (4.50) issue de notre formalisme :

$$\boldsymbol{\mu}_I^{n+1} = \left( \boldsymbol{\mu}_I^n + \left( \frac{1}{2} \mathbf{v}^{-1} \right) \right) \circ (\boldsymbol{\mu}_I^n)^{-1} \circ \left( \boldsymbol{\mu}_I^n + \left( \frac{1}{2} \mathbf{v}^{-1} \right) \right). \quad (\text{B.6})$$

Nous définissons le vecteur  $\mathbf{w}$  proche de  $\mathbf{0}$  par :

$$\mathbf{w} = (\boldsymbol{\mu}_I^n)^{-1} \circ \left( \boldsymbol{\mu}_I^n + \frac{1}{2} \mathbf{v}^{-1} \right) \quad (\text{B.7})$$

Nous avons en insérant  $\mathbf{w}$  dans (B.6) :

$$\boldsymbol{\mu}_I^{n+1} = \boldsymbol{\mu}_I^n \circ \mathbf{w} \circ \mathbf{w}. \quad (\text{B.8})$$

De plus, nous avons d'une part

$$\boldsymbol{\mu}_I^n \circ \mathbf{w} = \left( \boldsymbol{\mu}_I^n + \frac{1}{2} \mathbf{v}^{-1} \right) \quad (\text{B.9})$$

et d'autre part pour  $\mathbf{u} \in \mathcal{G}$  proche de  $\mathbf{0}$  :

$$\boldsymbol{\mu}_I^n \circ \mathbf{u} = \boldsymbol{\mu}_I^n + \mathbf{A}_{\boldsymbol{\mu}_I^n} \mathbf{u} + O(\|\mathbf{u}\|^2) \quad (\text{B.10})$$

avec  $\mathbf{A}_{\boldsymbol{\mu}_I^n}$  vérifiant :

$$\mathbf{A}_{\boldsymbol{\mu}_I^n} = \left. \frac{\partial \boldsymbol{\mu}_I^n \circ \mathbf{u}}{\partial \mathbf{u}} \right|_{\mathbf{0}}. \quad (\text{B.11})$$

Par identification, nous obtenons :

$$\mathbf{A}_{\boldsymbol{\mu}_I^n} \mathbf{u} + O(\|\mathbf{u}\|^2) = \frac{1}{2} \mathbf{v}^{-1} \quad (\text{B.12})$$

De plus à l'ordre 1 autour de  $\mathbf{0}$  :

$$\mathbf{w} \circ \mathbf{w} = 2\mathbf{w} + O(\|\mathbf{w}\|^2), \quad (\text{B.13})$$

## B.2. Équivalence à l'ordre 1 de la règle de mise à jour de l'approche GESM avec la règle de mise à jour de l'approche FA+IAD

---

Par conséquent en posant  $\mathbf{u} = \mathbf{w} \circ \mathbf{w}$  dans (B.8) :

$$\begin{aligned}
 \boldsymbol{\mu}_I^{n+1} &= \boldsymbol{\mu}_I^n \circ \mathbf{w} \circ \mathbf{w} \\
 &= \boldsymbol{\mu}_I^n + \mathbf{A}_{\boldsymbol{\mu}_I^n}(\mathbf{w} \circ \mathbf{w}) + O(\|(\mathbf{w} \circ \mathbf{w})\|^2) \\
 &\stackrel{\text{(B.13)}}{=} \boldsymbol{\mu}_I^n + 2\mathbf{A}_{\boldsymbol{\mu}_I^n}\mathbf{w} + O(\|\mathbf{w}\|^2) \\
 &\stackrel{\text{(B.12)}}{\approx} \boldsymbol{\mu}_I^n + \mathbf{v}^{-1}
 \end{aligned} \tag{B.14}$$

Du point de vue de notre formalisme, l'approche GESM peut être interprétée comme l'optimisation de l'erreur utilisant les paramétrisations incrémentales (4.47) et (4.47), mais utilisant une règle de mise à jour approximative.





# Équivalence des approches au sein d'une même famille

---

Dans cette partie, nous présentons le détail de la preuve d'équivalence à l'ordre 1 des approches au sein d'une même famille (Directe, Inverse, Bidirectionnelle Pure et Symétrique Compositionnelle Exacte).

## C.1 Approches Directes (F)

On désigne dans la suite par  $\mathbf{v}^F$  l'incrément estimé par une méthode directe F quelconque. Nous avons, en utilisant (3.25), la règle de mise à jour suivante pour les approches directes :

$$\text{maj}^D(\boldsymbol{\mu}_I, \boldsymbol{\mu}_T) = (\boldsymbol{\mu}_I^F(\boldsymbol{\mu}_I^n, \mathbf{v}^F), \boldsymbol{\mu}_T^0). \quad (\text{C.1})$$

Or d'après l'équation (4.70), la matrice Jacobienne  $\mathbf{J}_e^F(\mathbf{0})$  d'une approche directe F quelconque est liée à la matrice Jacobienne de la méthode FA par :

$$\mathbf{J}_{e,i}^F(\mathbf{0}) = \left. \frac{\partial I(\mathbf{W}(\boldsymbol{\mu}_I^F(\boldsymbol{\mu}_I^n, \mathbf{v}), \mathbf{x}_i))}{\partial \mathbf{v}} \right|_{\mathbf{0}} = \mathbf{J}_{I,i}^A(\mathbf{0}) \mathcal{M}_I^{A \rightarrow F} = \mathbf{J}_{I,i}^{\text{FA}}(\mathbf{0}) \mathcal{M}_I^{A \rightarrow F}, \quad (\text{C.2})$$

avec  $\mathcal{M}_I^{A \rightarrow F}$  vérifiant (4.71) pour  $L = I$ . Ceci conduit à l'incrément suivant estimé par une méthode directe quelconque :

$$\mathbf{v}^F = -(\mathbf{J}_e^F(\mathbf{0}))^+ \mathbf{e}(\mathbf{0}) = -(\mathcal{M}_I^{A \rightarrow F})^{-1} (\mathbf{J}_I^A(\mathbf{0}))^+ \mathbf{e}(\mathbf{0}), \quad (\text{C.3})$$

d'où :

$$\mathbf{v}^F = (\mathcal{M}_I^{A \rightarrow F})^{-1} \mathbf{v}^{\text{FA}}, \quad (\text{C.4})$$

où  $\mathbf{v}^{\text{FA}}$  est l'incrément estimé par la méthode directe additive. L'hypothèse que les matrices Jacobiennes  $\mathbf{J}_e^F$  sont de rang plein assure l'inversibilité de la matrice  $\mathcal{M}_I^{A \rightarrow F}$  de dimension  $p \times p$ .

L'équation (C.3) montre que les approches directes utilisant des paramétrisations incrémentales différentes conduisent avec l'approche Gauss-Newton à l'estimation d'incrément  $\mathbf{v}^F$  différents. Cette différence est due à la matrice  $\mathcal{M}_I^{A \rightarrow F}$  qui varie pour chaque approche. Cependant, en utilisant l'approximation à l'ordre 1 (4.72) dans (C.1), nous avons bien une équivalence à l'ordre 1 sur les paramètres  $\boldsymbol{\mu}_I$  estimés :

$$\boldsymbol{\mu}_I^{\text{F},n+1} = \boldsymbol{\mu}_I^n + \mathcal{M}_I^{A \rightarrow F} \mathbf{v}^F + O(\|\mathbf{v}^F\|^2) = \boldsymbol{\mu}_I^n + \mathbf{v}^{\text{FA}} + O(\|\mathbf{v}^F\|^2). \quad (\text{C.5})$$

Cette analyse étend la démonstration d'équivalence des approches FA et FC proposée dans [Baker 2004] à l'ensemble des approches directes (F).

## C.2 Approches Inverses (I)

On désigne dans la suite par  $\mathbf{v}^I$  l'incrément estimé par une méthode Inverse I quelconque. En utilisant (3.25), la règle de mise à jour pour les approches inverses est la suivante :

$$\text{maj}^D(\boldsymbol{\mu}_I, \boldsymbol{\mu}_T) = (\boldsymbol{\mu}_I^n \circ \boldsymbol{\mu}_T^I(\boldsymbol{\mu}_T^0, \mathbf{v}^I)^{-1} \circ \boldsymbol{\mu}_T^0, \boldsymbol{\mu}_T^0). \quad (\text{C.6})$$

Or d'après l'équation (4.70), la matrice Jacobienne  $\mathbf{J}_e^I(\mathbf{0})$  d'une approche inverse quelconque est liée à la matrice Jacobienne de la méthode IAR par :

$$\mathbf{J}_{e,i}^I(\mathbf{0}) = - \left. \frac{\partial T(\mathbf{W}(\boldsymbol{\mu}_T^I(\boldsymbol{\mu}_T^0, \mathbf{v}), \mathbf{x}_i))}{\partial \mathbf{v}} \right|_{\mathbf{0}} = -\mathbf{J}_{T,i}^A(\mathbf{0})\mathcal{M}_T^{A \rightarrow I} = \mathbf{J}_{T,i}^{\text{IAR}}(\mathbf{0})\mathcal{M}_T^{A \rightarrow I}. \quad (\text{C.7})$$

Ceci conduit à l'incrément suivant estimé par une méthode inverse quelconque :

$$\mathbf{v}^I = -(\mathbf{J}_e^I(\mathbf{0}))^+ \mathbf{e}(\mathbf{0}) = (\mathcal{M}_T^{A \rightarrow I})^{-1} (\mathbf{J}_T^A(\mathbf{0}))^+ \mathbf{e}(\mathbf{0}), \quad (\text{C.8})$$

d'où :

$$\mathbf{v}^I = (\mathcal{M}_T^{A \rightarrow I})^{-1} \mathbf{v}^{\text{IAR}}, \quad (\text{C.9})$$

où  $\mathbf{v}^{\text{IAR}}$  est l'incrément estimé par la méthode *Inverse Additive Reverse*. Cette équation montre comme précédemment que les approches inverses utilisant des paramétrisations incrémentales différentes conduisent avec l'approche Gauss-Newton à l'estimation d'incrément différents. Cette différence est dû à la matrice  $\mathcal{M}_T^{A \rightarrow I}$  spécifique à chaque approche I. Cependant, en utilisant l'approximation à l'ordre 1 (4.73), nous avons :

$$\boldsymbol{\mu}_T(\boldsymbol{\mu}_T^0, \mathbf{v}^I) = \boldsymbol{\mu}_T^0 + \mathcal{M}_T^{A \rightarrow I} \mathbf{v}^I + O(\|\mathbf{v}^I\|^2) = \boldsymbol{\mu}_T^0 + \mathbf{v}^{\text{IAR}} + O(\|\mathbf{v}^I\|^2). \quad (\text{C.10})$$

En insérant cette dernière relation dans la règle de mise à jour, nous vérifions bien que les paramètres estimés sont identiques à l'ordre 1 et ce quelque soit la méthode inverse utilisée :

$$\boldsymbol{\mu}_I^{\text{I},n+1} = \boldsymbol{\mu}_I^n \circ (\boldsymbol{\mu}_T^0 + \mathbf{v}^{\text{IAR}})^{-1} \circ \boldsymbol{\mu}_T^0 + O(\|\mathbf{v}^{\text{IAR}}\|^2). \quad (\text{C.11})$$

L'équation (C.11) révèle que l'ensemble des approches inverses sont équivalentes à l'ordre 1 étendant la démonstration d'équivalence des approches IA et IC proposée dans [Baker 2004] à l'ensemble des approches inverses (I).

## C.3 Approches Bidirectionnelles Pures

Nous désignons dans la suite par  $\mathbf{v}^B = [\mathbf{v}_I^B \quad \mathbf{v}_T^B]^T$  l'incrément estimé par une méthode Bidirectionnelle B quelconque. Cette fois, la règle de mise à jour est :

$$\text{maj}^D(\boldsymbol{\mu}_I, \boldsymbol{\mu}_T) = (\boldsymbol{\mu}_I^B(\boldsymbol{\mu}_I^n, \mathbf{v}_I^B) \circ \boldsymbol{\mu}_T^B(\boldsymbol{\mu}_T^0, \mathbf{v}_T^B)^{-1} \circ \boldsymbol{\mu}_T^0, \boldsymbol{\mu}_T^0). \quad (\text{C.12})$$

## C.4. Approches Symétriques compositionnelles exactes

La matrice Jacobienne de l'erreur purement bidirectionnelle associée à cette méthode quelconque est :

$$\begin{aligned}
\mathbf{J}_{\mathbf{e},i}^{\mathbf{B}}(\mathbf{0}) &= \left[ \frac{\partial I(\mathbf{W}(\boldsymbol{\mu}_I(\boldsymbol{\mu}_I^n, \mathbf{v}), \mathbf{x}_i))}{\partial \mathbf{v}} \Big|_{\mathbf{0}} \quad - \quad \frac{\partial I(\mathbf{W}(\boldsymbol{\mu}_T(\boldsymbol{\mu}_T^0, \mathbf{v}), \mathbf{x}_i))}{\partial \mathbf{v}} \Big|_{\mathbf{0}} \right] \\
&= [\mathbf{J}_{I,i}^{\mathbf{A}}(\mathbf{0})\mathcal{M}_I^{\mathbf{A}\rightarrow\mathbf{B}} \quad - \quad \mathbf{J}_{T,i}^{\mathbf{A}}(\mathbf{0})\mathcal{M}_T^{\mathbf{A}\rightarrow\mathbf{B}}] \\
&= [\mathbf{J}_{I,i}^{\mathbf{A}}(\mathbf{0}) \quad - \quad \mathbf{J}_{T,i}^{\mathbf{A}}(\mathbf{0})] \begin{pmatrix} \mathcal{M}_I^{\mathbf{A}\rightarrow\mathbf{B}} & \mathbf{0} \\ \mathbf{0} & \mathcal{M}_T^{\mathbf{A}\rightarrow\mathbf{B}} \end{pmatrix} \\
&= \mathbf{J}_{\mathbf{e},i}^{\mathbf{BA}}(\mathbf{0}) \begin{pmatrix} \mathcal{M}_I^{\mathbf{A}\rightarrow\mathbf{B}} & \mathbf{0} \\ \mathbf{0} & \mathcal{M}_T^{\mathbf{A}\rightarrow\mathbf{B}} \end{pmatrix}
\end{aligned} \tag{C.13}$$

où  $\mathbf{J}_{\mathbf{e},i}^{\mathbf{BA}}(\mathbf{0})$  est la matrice Jacobienne associée à l'approche BA. Nous en déduisons, en utilisant (C.13) dans (4.69), l'expression de l'incrément d'une méthode bidirectionnelle quelconque en fonction de l'incrément estimé par la méthode Bidirectionnelle additive  $\mathbf{v}^{\mathbf{BA}}$  :

$$\mathbf{v}^{\mathbf{B}} = - \begin{pmatrix} (\mathcal{M}_I^{\mathbf{A}\rightarrow\mathbf{B}})^{-1} & \mathbf{0} \\ \mathbf{0} & (\mathcal{M}_T^{\mathbf{A}\rightarrow\mathbf{B}})^{-1} \end{pmatrix} \mathbf{J}_{\mathbf{e},i}^{\mathbf{BA}}(\mathbf{0})^+ \mathbf{e}(\mathbf{0}) = \begin{pmatrix} (\mathcal{M}_I^{\mathbf{A}\rightarrow\mathbf{B}})^{-1} & \mathbf{0} \\ \mathbf{0} & (\mathcal{M}_T^{\mathbf{A}\rightarrow\mathbf{B}})^{-1} \end{pmatrix} \mathbf{v}^{\mathbf{BA}}. \tag{C.14}$$

Il vient ainsi :

$$\mathbf{v}_I^{\mathbf{B}} = (\mathcal{M}_I^{\mathbf{A}\rightarrow\mathbf{B}})^{-1} \mathbf{v}_I^{\mathbf{BA}}, \text{ et} \tag{C.15}$$

$$\mathbf{v}_T^{\mathbf{B}} = (\mathcal{M}_T^{\mathbf{A}\rightarrow\mathbf{B}})^{-1} \mathbf{v}_T^{\mathbf{BA}}. \tag{C.16}$$

En utilisant  $\mathbf{v}_I^{\mathbf{B}}$  et  $\mathbf{v}_T^{\mathbf{B}}$  dans (4.72) et (4.73), il vient :

$$\boldsymbol{\mu}_I^{\mathbf{B}}(\boldsymbol{\mu}_I^n, \mathbf{v}_I^{\mathbf{B}}) = \boldsymbol{\mu}_I^n + \mathbf{v}_I^{\mathbf{BA}} + O(\|\mathbf{v}_I^{\mathbf{B}}\|^2) = \boldsymbol{\mu}_I^n + \mathbf{v}_I^{\mathbf{BA}} + O(\|\mathbf{v}_I^{\mathbf{B}}\|^2 + \|\mathbf{v}_T^{\mathbf{B}}\|^2) \tag{C.17}$$

$$\boldsymbol{\mu}_T^{\mathbf{B}}(\boldsymbol{\mu}_T^n, \mathbf{v}_T^{\mathbf{B}}) = \boldsymbol{\mu}_T^0 + \mathbf{v}_T^{\mathbf{BA}} + O(\|\mathbf{v}_T^{\mathbf{B}}\|^2) = \boldsymbol{\mu}_T^0 + \mathbf{v}_T^{\mathbf{BA}} + O(\|\mathbf{v}_I^{\mathbf{B}}\|^2 + \|\mathbf{v}_T^{\mathbf{B}}\|^2) \tag{C.18}$$

Les paramètres estimés quelle que soit la méthode Bidirectionnelle choisie vérifient donc :

$$\begin{aligned}
\boldsymbol{\mu}_I^{\mathbf{B},n+1} &= \boldsymbol{\mu}_I^{\mathbf{B}}(\boldsymbol{\mu}_I^n, \mathbf{v}_I^{\mathbf{B}}) \circ \boldsymbol{\mu}_T^{\mathbf{B}}(\boldsymbol{\mu}_T^0, \mathbf{v}_T^{\mathbf{B}})^{-1} \circ \boldsymbol{\mu}_T^0 \\
&= (\boldsymbol{\mu}_I^n + \mathbf{v}_I^{\mathbf{BA}}) \circ (\boldsymbol{\mu}_T^0 + \mathbf{v}_T^{\mathbf{BA}})^{-1} \circ \boldsymbol{\mu}_T^0 + O(\|\mathbf{v}_I^{\mathbf{B}}\|^2 + \|\mathbf{v}_T^{\mathbf{B}}\|^2).
\end{aligned} \tag{C.19}$$

L'équation (C.19) révèle que l'ensemble des approches purement bidirectionnelles sont équivalentes à l'ordre 1.

## C.4 Approches Symétriques compositionnelles exactes

Dans la suite, nous notons  $\mathbf{v}^{\mathbf{SCE}}$  l'incrément estimé par une méthode symétrique compositionnelle exacte quelconque et  $\mathbf{J}_{\mathbf{e}}^{\mathbf{SCE}}(\mathbf{0})$  la matrice Jacobienne associée. La règle de mise à jour d'une approche symétrique est :

$$\text{maj}^D(\boldsymbol{\mu}_I, \boldsymbol{\mu}_T) = (\boldsymbol{\mu}_I^{\mathbf{SCE}}(\boldsymbol{\mu}_I^n, \mathbf{v}^{\mathbf{SCE}}) \circ \boldsymbol{\mu}_T^{\mathbf{SCE}}(\boldsymbol{\mu}_T^0, \mathbf{v}^{\mathbf{SCE}})^{-1} \circ \boldsymbol{\mu}_T^0, \boldsymbol{\mu}_T^0). \tag{C.20}$$

Rappelons que les approches symétriques compositionnelles exactes vérifient (4.36) :

$$(\boldsymbol{\mu}_I^n)^{-1} \circ \boldsymbol{\mu}_I^{\mathbf{SCE}}(\boldsymbol{\mu}_I^n, \mathbf{v}) = \boldsymbol{\mu}_T^{\mathbf{SCE}}(\boldsymbol{\mu}_T^0, \mathbf{v})^{-1} \circ \boldsymbol{\mu}_T^0 \tag{C.21}$$

Par conséquent nous avons :

$$\boldsymbol{\mu}_I^{\text{SCE}}(\boldsymbol{\mu}_I^n, \mathbf{v}) = (\boldsymbol{\mu}_I^n) \circ \boldsymbol{\mu}_T^{\text{SCE}}(\boldsymbol{\mu}_T^0, \mathbf{v})^{-1} \circ \boldsymbol{\mu}_T^0 \quad (\text{C.22})$$

et :

$$\boldsymbol{\mu}_T^{\text{SCE}}(\boldsymbol{\mu}_T^n, \mathbf{v}) = \boldsymbol{\mu}_T^0 \circ \boldsymbol{\mu}_I^{\text{SCE}}(\boldsymbol{\mu}_I^n, \mathbf{v})^{-1} \circ (\boldsymbol{\mu}_I^n) \quad (\text{C.23})$$

A partir de (C.23), il est possible d'exprimer la matrice Jacobienne  $\mathbf{J}_e^{\text{SCE}}(\mathbf{0})$  par :

$$\mathbf{J}_e^{\text{SCE},1}(\mathbf{0}) = \frac{1}{2}(\mathbf{J}_I^A(\mathbf{0}) - \mathbf{J}_T^{\boldsymbol{\mu}_I^n, \boldsymbol{\mu}_T^0}(\mathbf{0}))\mathcal{M}_I^{A \rightarrow \text{SCE}} \quad (\text{C.24})$$

avec  $\mathbf{J}_I^A(\mathbf{0})$  vérifiant l'équation (4.5),  $\mathcal{M}_I^{A \rightarrow \text{SCE}}$  vérifiant (4.71) pour  $L = I$  et  $\mathbf{J}_T^{\boldsymbol{\mu}_I^n, \boldsymbol{\mu}_T^0}(\mathbf{0})$  satisfaisant :

$$\mathbf{J}_T^{\boldsymbol{\mu}_I^n, \boldsymbol{\mu}_T^0}(\mathbf{0}) = \left. \frac{\partial T(\mathbf{W}(\boldsymbol{\mu}_T^0 \circ \boldsymbol{\mu}^{-1}, \mathbf{W}(\boldsymbol{\mu}_I^n, \mathbf{x}_i)))}{\partial \boldsymbol{\mu}} \right|_{\boldsymbol{\mu}_I^n}. \quad (\text{C.25})$$

A partir de (C.22), il est possible d'exprimer la matrice Jacobienne  $\mathbf{J}_e^{\text{SCE}}(\mathbf{0})$  par :

$$\mathbf{J}_e^{\text{SCE},2}(\mathbf{0}) = \frac{1}{2}(\mathbf{J}_I^{\boldsymbol{\mu}_I^n, \boldsymbol{\mu}_T^0}(\mathbf{0}) - \mathbf{J}_T^A(\mathbf{0}))\mathcal{M}_T^{A \rightarrow \text{SCE}} \quad (\text{C.26})$$

avec  $\mathbf{J}_T^A(\mathbf{0})$  vérifiant l'équation (4.5),  $\mathcal{M}_T^{A \rightarrow \text{SCE}}$  vérifiant (4.71) pour  $L = I$  et  $\mathbf{J}_I^{\boldsymbol{\mu}_I^n, \boldsymbol{\mu}_T^0}(\mathbf{0})$  satisfaisant :

$$\mathbf{J}_I^{\boldsymbol{\mu}_I^n, \boldsymbol{\mu}_T^0}(\mathbf{0}) = \left. \frac{\partial I(\mathbf{W}(\boldsymbol{\mu}_I^n \circ \boldsymbol{\mu}^{-1}, \mathbf{W}(\boldsymbol{\mu}_T^0, \mathbf{x}_i)))}{\partial \boldsymbol{\mu}} \right|_{\boldsymbol{\mu}_T^n}. \quad (\text{C.27})$$

Ainsi d'une part :

$$\mathbf{v}^{\text{SCE}} = -(\mathcal{M}_I^{A \rightarrow \text{SCE}})^{-1} (\mathbf{J}^{\text{SCE},1}(\mathbf{0}))^+ \mathbf{e}(\mathbf{0}), \quad (\text{C.28})$$

avec :

$$\mathbf{J}^{\text{SCE},1}(\mathbf{0}) = \frac{1}{2}(\mathbf{J}_I^A(\mathbf{0}) - \mathbf{J}_T^{\boldsymbol{\mu}_I^n, \boldsymbol{\mu}_T^0}(\mathbf{0})), \quad (\text{C.29})$$

et d'autre part :

$$\mathbf{v}^{\text{SCE}} = -(\mathcal{M}_T^{A \rightarrow \text{SCE}})^{-1} (\mathbf{J}^{\text{SCE},2}(\mathbf{0}))^+ \mathbf{e}(\mathbf{0}), \quad (\text{C.30})$$

avec :

$$\mathbf{J}^{\text{SCE},2}(\mathbf{0}) = \frac{1}{2}(\mathbf{J}_I^{\boldsymbol{\mu}_I^n, \boldsymbol{\mu}_T^0}(\mathbf{0}) - \mathbf{J}_T^A(\mathbf{0})) \quad (\text{C.31})$$

En utilisant (C.28) dans (4.72), il vient :

$$\boldsymbol{\mu}_I^{\text{SCE}}(\boldsymbol{\mu}_I^n, \mathbf{v}^{\text{SCE}}) = \boldsymbol{\mu}_I^n - (\mathbf{J}^{\text{SCE},1}(\mathbf{0}))^+ \mathbf{e}(\mathbf{0}) + O(\|\mathbf{v}^{\text{SCE}}\|^2) \quad (\text{C.32})$$

En utilisant (C.30) dans (4.73), il vient :

$$\boldsymbol{\mu}_T^{\text{SCE}}(\boldsymbol{\mu}_T^n, \mathbf{v}^{\text{SCE}}) = \boldsymbol{\mu}_T^n - (\mathbf{J}^{\text{SCE},2}(\mathbf{0}))^+ \mathbf{e}(\mathbf{0}) + O(\|\mathbf{v}^{\text{SCE}}\|^2) \quad (\text{C.33})$$

En insérant ces deux expressions dans la règle de mise à jour (C.20), il vient :

$$\boldsymbol{\mu}_I^{\text{SCE},n+1} = (\boldsymbol{\mu}_I^n - (\mathbf{J}^{\text{SCE},1}(\mathbf{0}))^+ \mathbf{e}(\mathbf{0})) \circ (\boldsymbol{\mu}_T^n - (\mathbf{J}^{\text{SCE},2}(\mathbf{0}))^+ \mathbf{e}(\mathbf{0}))^{-1} \circ \boldsymbol{\mu}_T^0 + O(\|\mathbf{v}^{\text{SCE}}\|^2). \quad (\text{C.34})$$

L'équation (C.34) révèle que l'ensemble des approches symétriques compositionnelles exactes sont équivalentes à l'ordre 1.

## C.5 Équivalence à l'ordre 1 de l'approche SCO avec les approches symétriques exactes

Dans cette partie, nous établissons une équivalence à l'ordre 1 entre l'approche SCO définie à la section 4.3.2 et l'approche SC définie à la section 4.3.1.3. Afin de différencier les paramètres issus des deux méthodes, nous notons  $\mathbf{v}^M$  l'incrément estimé par la méthode  $M$  et  $\boldsymbol{\mu}_I^{M,n+1}$  les paramètres correspondant obtenus par mise à jour directe :

$$\boldsymbol{\mu}_I^{M,n+1} = \boldsymbol{\mu}_I^M(\boldsymbol{\mu}_I^n, \mathbf{v}^M) \circ \boldsymbol{\mu}_T^M(\boldsymbol{\mu}_T^n, \mathbf{v}^M)^{-1} \circ \boldsymbol{\mu}_T^n \quad (\text{C.35})$$

### C.5.1 Relation entre les incréments

D'après le tableau 4.1, nous avons d'une part :

$$\mathbf{J}_e^{\text{SCO}}(\mathbf{0}) = \mathbf{J}_e^{\text{SC}}(\mathbf{0}), \quad (\text{C.36})$$

ce qui conduit à :

$$\mathbf{v}^{\text{SCO}} = \mathbf{v}^{\text{SC}}. \quad (\text{C.37})$$

Les approches SC et SCO conduise à l'estimation du même incrément.

### C.5.2 Équivalence entre les approches

**Lemme 3.**

$$\left. \frac{\partial \mathbf{W}(\frac{1}{2}\mathbf{v} \circ (\frac{1}{2}\mathbf{v}), \mathbf{x}_i)}{\partial \mathbf{v}} \right|_{\mathbf{0}} = \left. \frac{\partial \mathbf{W}(\frac{1}{2}\mathbf{v} \circ (-\frac{1}{2}\mathbf{v})^{-1}, \mathbf{x}_i)}{\partial \mathbf{v}} \right|_{\mathbf{0}} = \left. \frac{\partial \mathbf{W}(\mathbf{v}, \mathbf{x}_i)}{\partial \mathbf{v}} \right|_{\mathbf{0}}. \quad (\text{C.38})$$

*Démonstration.* Nous avons :

$$\mathbf{W}\left(\frac{1}{2}\mathbf{v} \circ \frac{1}{2}\mathbf{v}, \mathbf{x}_i\right) = \mathbf{g} \circ \mathbf{f}(\mathbf{v}), \quad (\text{C.39})$$

avec :

$$\mathbb{R}^p \longmapsto \mathcal{P} \times \mathbb{R}^2$$

$$\mathbf{f} : \mathbf{v} \longmapsto \begin{pmatrix} \frac{1}{2}\mathbf{v} \\ \mathbf{z} = \mathbf{W}(\frac{1}{2}\mathbf{v}, \mathbf{x}_i) \end{pmatrix}, \quad (\text{C.40})$$

$$\mathcal{P} \times \mathbb{R}^2 \longmapsto \mathbb{R}^2$$

$$\mathbf{g} : \begin{pmatrix} \mathbf{v} \\ \mathbf{z} \end{pmatrix} \longmapsto \mathbf{W}(\mathbf{v}, \mathbf{z}). \quad (\text{C.41})$$

Par conséquent :

$$\begin{aligned}
 \left. \frac{\partial \mathbf{W}(\frac{1}{2}\mathbf{v} \circ \frac{1}{2}\mathbf{v}, \mathbf{x}_i)}{\partial \mathbf{v}} \right|_{\mathbf{0}} &= \left. \frac{\partial \mathbf{g} \circ \mathbf{f}(\mathbf{v})}{\partial \mathbf{v}} \right|_{\mathbf{0}} \\
 &= \left. \frac{\partial \mathbf{g}(\mathbf{f})}{\partial \mathbf{f}} \right|_{\mathbf{f}=(\mathbf{0}, \mathbf{x}_i)} \left. \frac{\partial \mathbf{f}(\mathbf{v})}{\partial \mathbf{v}} \right|_{\mathbf{0}} \\
 &= \left[ \left. \frac{\partial \mathbf{W}(\mathbf{v}, \mathbf{x}_i)}{\partial \mathbf{v}} \right|_{\mathbf{0}} \quad \left. \frac{\partial \mathbf{W}(\mathbf{0}, \mathbf{z})}{\partial \mathbf{z}} \right|_{\mathbf{x}_i} \right] \left[ \begin{array}{c} \left. \frac{\partial \frac{1}{2}\mathbf{v}}{\partial \mathbf{v}} \right|_{\mathbf{0}} \\ \left. \frac{\partial \mathbf{W}(\frac{1}{2}\mathbf{v}, \mathbf{x}_i)}{\partial \mathbf{v}} \right|_{\mathbf{0}} \end{array} \right] \\
 &= \left. \frac{\partial \mathbf{W}(\mathbf{v}, \mathbf{x}_i)}{\partial \mathbf{v}} \right|_{\mathbf{0}}
 \end{aligned} \tag{C.42}$$

En procédant de même nous montrons que :

$$\left. \frac{\partial \mathbf{W}(\frac{1}{2}\mathbf{v} \circ (-\frac{1}{2}\mathbf{v})^{-1}, \mathbf{x}_i)}{\partial \mathbf{v}} \right|_{\mathbf{0}} = \left. \frac{\partial \mathbf{W}(\mathbf{v}, \mathbf{x}_i)}{\partial \mathbf{v}} \right|_{\mathbf{0}}, \tag{C.43}$$

ce qui achève la démonstration. □

**Équivalence SC et SCO** Nous avons :

$$\mathbf{W} \left( \boldsymbol{\mu}_I^n \circ (\frac{1}{2}\mathbf{v}^{\text{SCO}}) \circ (-\frac{1}{2}\mathbf{v}^{\text{SCO}})^{-1}, \mathbf{x}_i \right) \stackrel{\text{(C.37)}}{=} \mathbf{W} \left( \boldsymbol{\mu}_I^n \circ (\frac{1}{2}\mathbf{v}^{\text{SC}}) \circ (-\frac{1}{2}\mathbf{v}^{\text{SC}})^{-1}, \mathbf{x}_i \right) \tag{C.44}$$

Par ailleurs nous avons :

$$\begin{aligned}
 \mathbf{W} \left( \boldsymbol{\mu}_I^n \circ (\frac{1}{2}\mathbf{v}) \circ (-\frac{1}{2}\mathbf{v})^{-1}, \mathbf{x}_i \right) &= \mathbf{W}(\boldsymbol{\mu}_I^n, \mathbf{x}_i) \\
 &\quad + \left. \frac{\partial \mathbf{W}(\boldsymbol{\mu}_I^n, \mathbf{z})}{\partial \mathbf{z}} \right|_{\mathbf{x}_i} \left. \frac{\partial \mathbf{W}(\frac{1}{2}\mathbf{v} \circ (-\frac{1}{2}\mathbf{v})^{-1}, \mathbf{x}_i)}{\partial \mathbf{v}} \right|_{\mathbf{0}} \mathbf{v} + O(\|\mathbf{v}\|^2) \\
 &\stackrel{\text{(C.38)}}{=} \mathbf{W}(\boldsymbol{\mu}_I^n, \mathbf{x}_i) + \left. \frac{\partial \mathbf{W}(\boldsymbol{\mu}_I^n, \mathbf{z})}{\partial \mathbf{z}} \right|_{\mathbf{x}_i} \left. \frac{\partial \mathbf{W}(\mathbf{v}, \mathbf{x}_i)}{\partial \mathbf{v}} \right|_{\mathbf{0}} \mathbf{v} + O(\|\mathbf{v}\|^2).
 \end{aligned} \tag{C.45}$$

Enfin nous avons aussi :

$$\begin{aligned}
 \mathbf{W} \left( \boldsymbol{\mu}_I^n \circ (\frac{1}{2}\mathbf{v}) \circ (\frac{1}{2}\mathbf{v}), \mathbf{x}_i \right) &= \mathbf{W}(\boldsymbol{\mu}_I^n, \mathbf{x}_i) \\
 &\quad + \left. \frac{\partial \mathbf{W}(\boldsymbol{\mu}_I^n, \mathbf{z})}{\partial \mathbf{z}} \right|_{\mathbf{x}_i} \left. \frac{\partial \mathbf{W}(\frac{1}{2}\mathbf{v} \circ (\frac{1}{2}\mathbf{v}), \mathbf{x}_i)}{\partial \mathbf{v}} \right|_{\mathbf{0}} \mathbf{v} + O(\|\mathbf{v}\|^2) \\
 &\stackrel{\text{(C.38)}}{=} \mathbf{W}(\boldsymbol{\mu}_I^n, \mathbf{x}_i) + \left. \frac{\partial \mathbf{W}(\boldsymbol{\mu}_I^n, \mathbf{z})}{\partial \mathbf{z}} \right|_{\mathbf{x}_i} \left. \frac{\partial \mathbf{W}(\mathbf{v}, \mathbf{x}_i)}{\partial \mathbf{v}} \right|_{\mathbf{0}} \mathbf{v} + O(\|\mathbf{v}\|^2),
 \end{aligned} \tag{C.46}$$

d'où

$$\mathbf{W} \left( \boldsymbol{\mu}_I^n \circ (\frac{1}{2}\mathbf{v}) \circ (-\frac{1}{2}\mathbf{v})^{-1}, \mathbf{x}_i \right) = \mathbf{W} \left( \boldsymbol{\mu}_I^n \circ (\frac{1}{2}\mathbf{v}) \circ (\frac{1}{2}\mathbf{v}), \mathbf{x}_i \right) + O(\|\mathbf{v}\|^2). \tag{C.47}$$

### C.5. Équivalence à l'ordre 1 de l'approche SCO avec les approches symétriques exactes

---

Il vient alors :

$$\mathbf{W} \left( \boldsymbol{\mu}_I^n \circ \left( \frac{1}{2} \mathbf{v}^{\text{SC}} \right) \circ \left( -\frac{1}{2} \mathbf{v}^{\text{SC}} \right)^{-1}, \mathbf{x}_i \right) = \mathbf{W} \left( \boldsymbol{\mu}_I^n \circ \left( \frac{1}{2} \mathbf{v}^{\text{SC}} \right) \circ \left( \frac{1}{2} \mathbf{v}^{\text{SC}} \right), \mathbf{x}_i \right) + O(\|\mathbf{v}^{\text{SC}}\|^2), \quad (\text{C.48})$$

et en utilisant (C.48) dans (C.44) :

$$\mathbf{W} \left( \underbrace{\boldsymbol{\mu}_I^n \circ \left( \frac{1}{2} \mathbf{v}^{\text{SCO}} \right) \circ \left( -\frac{1}{2} \mathbf{v}^{\text{SCO}} \right)^{-1}}_{\boldsymbol{\mu}_I^{\text{SCO},n+1}}, \mathbf{x}_i \right) = \mathbf{W} \left( \underbrace{\boldsymbol{\mu}_I^n \circ \left( \frac{1}{2} \mathbf{v}^{\text{SC}} \right) \circ \left( \frac{1}{2} \mathbf{v}^{\text{SC}} \right)}_{\boldsymbol{\mu}_I^{\text{SC},n+1}}, \mathbf{x}_i \right) + O(\|\mathbf{v}^{\text{SC}}\|^2). \quad (\text{C.49})$$

Les approches SC et SCO sont donc équivalentes à l'ordre 1. Du fait que l'ensemble des approches symétriques compositionnelles exactes sont équivalentes et que l'approche SC fait partie de ces approches, nous en déduisons l'équivalence de l'approche SCO avec cette famille d'approche à l'ordre 1.





# Banc de test de Baker et Matthews pour les transformations affines

Le banc de test de Baker et Matthews [Baker 2004] permet de simuler et tester les algorithmes d’alignement sur des transformations synthétiques appliquées à une image. Le principe pour un modèle affine est le suivant : une image de référence  $\mathbf{I}_{\text{ref}}$  est utilisée à la fois pour simuler les images  $I$  et  $T$ . Une zone d’intérêt  $R$  est localisée dans l’image de référence à l’aide d’une boîte englobante rectangulaire. Trois positions définies comme les sommets d’un triangle isocèle inscrit dans la boîte englobante sont utilisées pour définir trois points canoniques. Des transformations affines aléatoires sont générées en ajoutant un bruit Gaussien spatial d’écart-type  $\sigma_{\text{noise}}$  (appelé Sigma Point) aux positions canoniques donnant naissance aux points de test. Les trois positions canoniques sont utilisées pour définir les paramètres  $\mu_I^0$  initialisant l’alignement dans l’image  $I$ . Les trois positions de test sont utilisées pour définir les paramètres  $\mu_T^0$  initialisant l’alignement dans le template  $T$ .

Les paramètres  $\mu_T^0$  ne changeant pas au cours des itérations (avec une règle de mise à jour directe), il sont utilisés pour extraire l’objet d’intérêt  $\text{warp}(T, \mu_T^0, \mathcal{S})$  avec  $\mathcal{S} = \{\mathbf{x}_i\}_{i=1..N}$  un échantillonnage de  $R$ . La figure D.1 illustre le principe du banc de test.

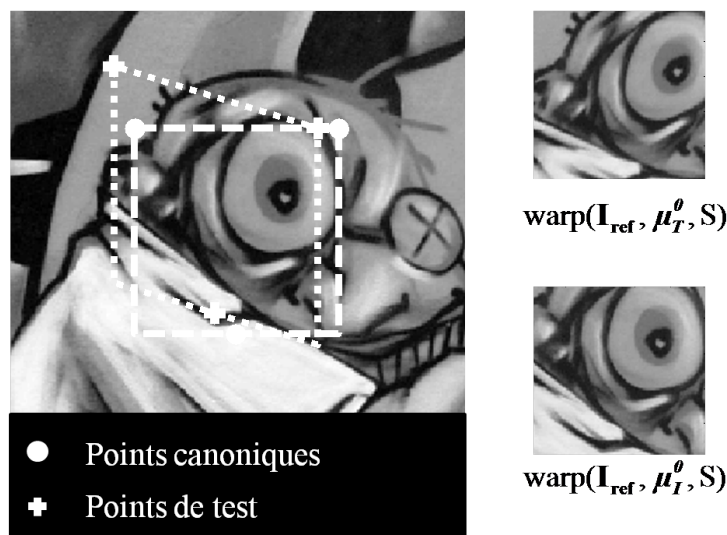


FIGURE D.1 – Principe du banc de test de Baker et Matthews [Baker 2004].

Les algorithmes d’alignement d’images sont alors exécutées afin d’aligner l’image  $I$  sur l’image  $T$ . A chaque nouvelle itération  $n$  d’une approche, les points canoniques correspondant à l’esti-

mation courante  $\mu_j^n$  sont censés se rapprocher des points de test si l'algorithme converge. Le critère de convergence est l'Erreur RMS Point (*Root Mean Squared*) à savoir la distance euclidienne moyenne entre les positions des points de test et les positions finales des points canoniques déplacées à l'aide de la transformation estimée. Deux critères de performance sont envisagés :

- Fréquence moyenne de convergence : pourcentage des tests pour lesquels un algorithme converge vers un estimé correct (estimé pour lequel l'erreur RMS point est inférieure à 1 pixel),
- Vitesse moyenne de convergence : pour les tests pour lesquels l'ensemble des méthodes ont convergé, l'erreur RMS Point moyenne est tracé en fonction du nombre d'itérations.

# Combinaison de fonctionnelles

---

Nous avons vu au chapitre 4 que le formalisme bidirectionnel permettait d'exprimer un large panel de méthodes existantes au sein d'un même cadre théorique. Nous avons également proposé un certain nombre de méthodes qui semblaient émerger naturellement du formalisme. Il existe cependant d'autres voies ou possibilités qui n'ont pas été prises en compte.

Des approches reposant sur la minimisation d'une somme de fonctionnelles unidirectionnelles (Directes ou Inverses) sont courantes dans le domaine du recalage d'images médicales. Dans cette partie, nous développons les liens qui existent entre ces approches et le formalisme bidirectionnel. On s'intéresse notamment à :

- la fonctionnelle que nous appelons *Forward Inverse Symmetric Combination* FISC (2.76) présentée à la section 2.3.1 et visant à rendre les images interchangeables. Nous rappelons ici l'expression de cette dernière :

$$E^{\text{FISC}}(\boldsymbol{\mu}) = \iint_{\mathbf{x}} (I(\mathbf{W}(\boldsymbol{\mu}, \mathbf{x})) - T(\mathbf{x}))^2 d\mathbf{x} + \iint_{\mathbf{z}} (I(\mathbf{z}) - T(\mathbf{W}(\boldsymbol{\mu}^{-1}, \mathbf{z})))^2 d\mathbf{z} \quad (\text{E.1})$$

- la fonctionnelle que nous appelons *Forward Inverse Asymmetric Combination* FIAC (2.78) visant à faire apparaître les gradients des deux images lors de la résolution du problème d'alignement :

$$E^{\text{FIAC}}(\boldsymbol{\mu}) = \frac{1}{2} \iint_{\mathbf{x}} (I(\mathbf{W}(\boldsymbol{\mu}, \mathbf{x})) - T(\mathbf{x}))^2 d\mathbf{x} + \frac{1}{2} \iint_{\mathbf{z}} (I(\mathbf{z}) - T(\mathbf{W}(\boldsymbol{\mu}^{-1}, \mathbf{z})))^2 j(\boldsymbol{\mu}, \mathbf{z}) d\mathbf{z} \quad (\text{E.2})$$

Comme précédemment ces fonctionnelles peuvent être minimisées en utilisant des approches itératives reposant sur l'introduction d'une paramétrisation incrémentale  $\boldsymbol{\mu} = \boldsymbol{\mu}(\boldsymbol{\mu}^n, \mathbf{v})$ . A la différence des approches du formalisme bidirectionnel, ces approches appliquent le même vecteur de paramètres  $\boldsymbol{\mu}$  à  $I$  et  $T$  ce qui a deux conséquences :

- Il n'y a qu'une seule règle de mise à jour possible, la règle de mise à jour Identité traditionnelle  $\boldsymbol{\mu}^{n+1} \leftarrow \boldsymbol{\mu}(\boldsymbol{\mu}^n, \hat{\mathbf{v}})$  avec  $\hat{\mathbf{v}}$  l'estimé issu d'un algorithme d'optimisation quelconque comme par exemple l'algorithme Gauss-Newton.
- Les deux images sont nécessairement déformées à chaque itération.

## E.1 Interprétation géométrique

Par simple changement de variable, la fonctionnelle FIAC introduite dans [Sabuncu 2009] peut être exprimée sous la forme d'une fonctionnelle directe :

$$E^{\text{FIAC}}(\boldsymbol{\mu}) = \iint_{\mathbf{x}} (I(\mathbf{W}(\boldsymbol{\mu}, \mathbf{x})) - T(\mathbf{x}))^2 d\mathbf{x} \quad (\text{E.3})$$

Le repère de compensation associé à cette approche est donc le repère de l'image  $T$  et cette approche peut donc être interprétée comme une approche de la classe Directe (F) avec  $\boldsymbol{\mu}_T^n = \mathbf{0}$ . Cependant, en introduisant la déformation des deux images cette approche implique les gradients des deux images lors de l'optimisation Gauss-Newton utilisée dans [Sabuncu 2009]. Elle associe donc à une interprétation géométrique directe, une estimation des gradients utilisant les deux images. En cela l'optimisation Gauss-Newton de la fonctionnelle FIAC est relativement similaire à la procédure d'optimisation ESM proposée par [Benhimane 2007b] et appliquée à une erreur FCL. La différence principale tient au fait que l'erreur  $E^{\text{FIAC}}$  autorise n'importe quelle paramétrisation incrémentale. De plus, les gradients des images  $I$  et  $T$  sont utilisés de manière asymétrique pour la fonctionnelle FIAC en raison du terme de changement de volume  $j(\boldsymbol{\mu}^{-1}, \mathbf{z})$  qui vient pondérer les gradients de  $T$  et de manière parfaitement symétrique pour l'ESM. Enfin, une grande différence intervient dans le cadre itératif. En effet, comme l'image  $T$  est compensée par un paramètre  $\boldsymbol{\mu}^{-1}$  à chaque itération, la matrice Jacobienne associée à cette image varie à chaque itération et ne peut donc être précalculée comme avec l'ESM.

Nous avons vu à la section 4.3.1 que les équations de l'ESM pouvaient être interprétées également du point de vue de la minimisation Gauss-Newton de l'erreur SCL avec mise à jour Directe. La figure 4.1 illustre la signification de l'approche SCL pour une itération. Pour cette approche, le repère de compensation choisi se situe exactement *au milieu* (au sens compositionnel) des repères des images  $I^n$  et  $T^n$  (pour une image nous avons  ${}^n = \text{warp}(L, \boldsymbol{\mu}_L^n)$ ). Si nous utilisons la fonctionnelle SSD SCL dans un processus itératif utilisant une règle de mise à jour identité (au lieu de la mise à jour directe employée par l'ESM), on obtient également un algorithme où la matrice Jacobienne associée à  $T$  doit être recalculée à chaque itération. La grande différence est alors que le repère de compensation est fixé comme celui de  $T$  au cours des itérations pour l'approche FIAC, alors que ce dernier est au milieu des repères des images  $I^n$  et  $T^n$  pour l'approche SCL avec mise à jour identité.

La figure E.1 illustre le comportement de l'approche FISC. A la différence des approches précédentes, cette approche fait intervenir deux erreurs pixel à pixel intermédiaires qui sont des erreurs directes et inverses. Ainsi, l'erreur globale d'alignement  $E^{\text{FISC}}(\mathbf{v})$  est la moyenne de deux erreurs d'alignement exprimées dans des repères de compensation différents. En effet, l'erreur directe suppose un alignement dans le repère de l'image  $T$  tandis que l'erreur inverse suppose un alignement relativement au repère de l'image compensée  $I$ . En cela, cette approche se révèle très différente des approches du formalisme bidirectionnel qui n'introduit qu'un unique repère de compensation. Concernant l'application de l'algorithme Gauss-Newton, la fonctionnelle FISC introduit une utilisation symétrique des gradients de chaque image ce qui la rapproche sur ce point des approches Bidirectionnelles subordonnées présentées à la section 4.3.2 (selon la paramétrisation incrémentale choisie les approches SC, SCL, SAD...).

**E.2. Comparaison des approches symétriques compositionnelles et des approches symétriques par combinaison pour le suivi d'objets**

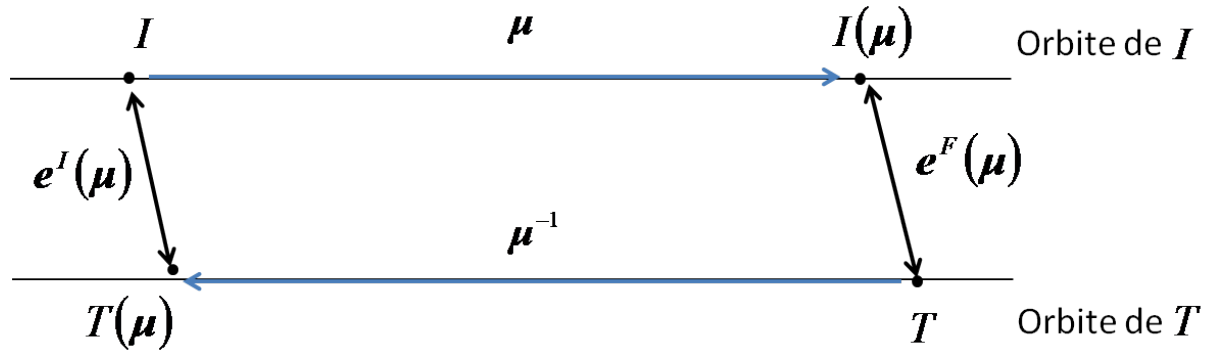


FIGURE E.1 – Interprétation géométrique de la fonctionnelle *Forward Inverse Symmetric Combination* FISC.

**E.2 Comparaison des approches symétriques compositionnelles et des approches symétriques par combinaison pour le suivi d'objets**

Dans le contexte du suivi d'objet, la fonctionnelle FISC tout comme la fonctionnelle FIAC paraissent peu adaptées car elles conduisent à déformer les deux images à chaque itération ce qui implique un temps de calcul plus important et l'assurance de s'éloigner de la région d'intérêt. Ces fonctionnelles peuvent cependant être adaptées au suivi d'objet en modifiant leur définition. Ainsi on peut définir par exemple la fonctionnelle FISCLT (*FISC on Lie Group for Tracking*) utilisant une paramétrisation incrémentale compositionnelle sur groupe de Lie :

$$E^{\text{FISCLT}}(\mathbf{v}) = \iint_{\mathbf{x}} (I(\mathbf{W}(\boldsymbol{\mu}_I^n \circ \exp(\mathbf{v}), \mathbf{x})) - T(\mathbf{x}))^2 d\mathbf{x} + \iint_{\mathbf{z}} (I(\mathbf{z}) - T(\mathbf{W}(\boldsymbol{\mu}_T^0 \circ \exp(-\mathbf{v}), \mathbf{z})))^2 d\mathbf{z} \quad (\text{E.4})$$

ou encore la fonctionnelle FIACLT (*FIAC on Lie Group for Tracking*) :

$$E^{\text{FIACLT}}(\mathbf{v}) = \frac{1}{2} \iint_{\mathbf{x}} (I(\mathbf{W}(\boldsymbol{\mu}_I^n \circ \exp(\mathbf{v}), \mathbf{x})) - T(\mathbf{x}))^2 d\mathbf{x} + \frac{1}{2} \iint_{\mathbf{z}} (I(\mathbf{z}) - T(\mathbf{W}(\boldsymbol{\mu}_T^0 \circ \exp(-\mathbf{v}), \mathbf{z})))^2 j(\mathbf{v}, \mathbf{z}) d\mathbf{z} \quad (\text{E.5})$$

En utilisant avec ces fonctionnelles une règle de mise à jour directe  $\boldsymbol{\mu}_I^{n+1} = \boldsymbol{\mu}_I^n \circ \exp(\mathbf{v})$ , nous garantissons un alignement vers une zone précise du template défini par  $\boldsymbol{\mu}_T^0$  qui ne varie pas au cours des itérations. Nous perdons alors cependant la propriété caractéristique de cette fonctionnelle, à savoir la possibilité d'interchanger les images sans affecter le résultat de l'alignement. Remarquons que pour un mouvement  $\mathbf{v}$  infinitésimal, le terme de changement de volume  $j(\mathbf{v}, \mathbf{z})$  de la fonctionnelle  $E^{\text{FIACLT}}$  devient quasiment constant de valeur 1 et les deux fonctionnelles FISCLT et FIACLT sont alors identiques.

Les deux fonctionnelles FISCLT et SCL sont des fonctionnelles des moindres carrées auxquelles on peut aisément appliquer l'algorithme de Gauss-Newton.

L'application de l'algorithme Gauss-Newton sur les erreurs FISC et SCL conduit alors à

	Fonctionnelle	
	FISCLT	SCL
<b>A</b>	$\frac{1}{2}((\mathbf{J}_I^{CL}(\mathbf{0}))^T \mathbf{J}_I^{CL}(\mathbf{0}) + (\mathbf{J}_T^{CL}(\mathbf{0}))^T \mathbf{J}_T^{CL}(\mathbf{0}))$	$\frac{1}{4}(\mathbf{J}_I^{CL}(\mathbf{0}) + \mathbf{J}_T^{CL}(\mathbf{0}))^T (\mathbf{J}_I^{CL}(\mathbf{0}) + \mathbf{J}_T^{CL}(\mathbf{0}))$
<b>B</b>	$\frac{1}{2}(\mathbf{J}_I^{CL}(\mathbf{0}) + \mathbf{J}_T^{CL}(\mathbf{0}))$	$\frac{1}{2}(\mathbf{J}_I^{CL}(\mathbf{0}) + \mathbf{J}_T^{CL}(\mathbf{0}))$

 TABLE E.1 – Estimation des matrices **A** et **B** associées aux approches FISC et SCL

l'équation :

$$\mathbf{v} = -\mathbf{A}^{-1} \mathbf{B}^T \mathbf{e}(\mathbf{0}) \quad (\text{E.6})$$

où **A** est une matrice  $p \times p$  ( $p$  est le nombre de paramètres de la transformation) homogène à une matrice Hessienne et **B** est une matrice homogène à une matrice Jacobienne de dimension  $N \times p$  où  $N$  est le nombre de pixels utilisés pour l'alignement. Le vecteur  $\mathbf{e}(\mathbf{0})$  correspond à l'erreur et est obtenue par la concaténation des erreurs pixel à pixel :

$$\mathbf{e}_i(\mathbf{0}) = I(\mathbf{W}(\boldsymbol{\mu}_I^n, \mathbf{x}_i)) - T(\mathbf{W}(\boldsymbol{\mu}_T^n, \mathbf{x}_i)). \quad (\text{E.7})$$

Le tableau E.1 fournit les différentes estimations de **A** et **B** associées aux différentes approches.

Nous constatons que l'approche FISCLT et SCL produisent les mêmes matrices **B**. En revanche l'approche FISCLT conduit à l'utilisation d'une matrice **A** obtenue en moyennant les matrices Hessiennes (plus précisément les approximations Gauss-Newton de la Hessienne) produites par les approches FCL et ICL. L'approche SCL rajoute les termes croisés  $(\mathbf{J}_I^{CL}(\mathbf{0}))^T \mathbf{J}_T^{CL}(\mathbf{0})$  et  $(\mathbf{J}_T^{CL}(\mathbf{0}))^T \mathbf{J}_I^{CL}(\mathbf{0})$ .

Nous avons implanté les approches FISCLT et ESM (SCL) au sein du banc d'expérimentation de [Baker 2004] dont le principe est rappelé à la section D. Le modèle de transformation utilisé est le modèle affine (2.6). La figure E.2 présente les courbes de fréquence moyenne de convergence et de vitesse moyenne de convergence obtenues sur les images illustrées à la figure 6.1 pour les familles d'approches Directe, Inverse et purement Bidirectionnelle. Aucun bruit n'a été ajouté aux images. Pour la fréquence de convergence, 500 réalisations différentes de l'initialisation de la déformation ( $\boldsymbol{\mu}_I^0$ ) ont été testées pour un point sigma et pour une image puis l'ensemble des résultats a été moyenné. Les courbes de vitesse de convergence sont obtenues en moyennant les résultats de 100 réalisations différentes de l'initialisation de la déformation par image.

L'algorithme ESM (Gauss-Newton SCL) produit des meilleurs résultats que l'approche Gauss-Newton FISCLT à la fois en termes de fréquence de convergence et de vitesse de convergence. L'ESM est initialement présentée comme une procédure d'optimisation de l'erreur FCL. A la section 5.1.1.1, nous dérivons un théorème important (5.18) impliquant que cette procédure conduit à l'approximation à l'ordre 2 de n'importe quelle erreur asymétrique vérifiant (5.6) en l'absence de bruit. En particulier, nous pouvons appliquer cette procédure à l'erreur ICL. Par conséquent l'application de la procédure ESM à la fonctionnelle FISCLT (E.4) conduit à l'utilisation de la matrice Jacobienne  $\mathbf{J}_{\text{ESM}}$  dans (2.34) pour l'estimation de  $\mathbf{v}$ . Ainsi on peut définir une approche FISCLT-ESM utilisant l'approximation ESM au lieu de l'approximation Gauss-Newton pour minimiser la fonctionnelle FISCLT. Cette approche conduit exactement aux mêmes équations que l'application de l'algorithme Gauss-Newton à l'erreur SCL, c'est à dire aux équations de l'algorithme ESM.

## E.2. Comparaison des approches symétriques compositionnelles et des approches symétriques par combinaison pour le suivi d'objets

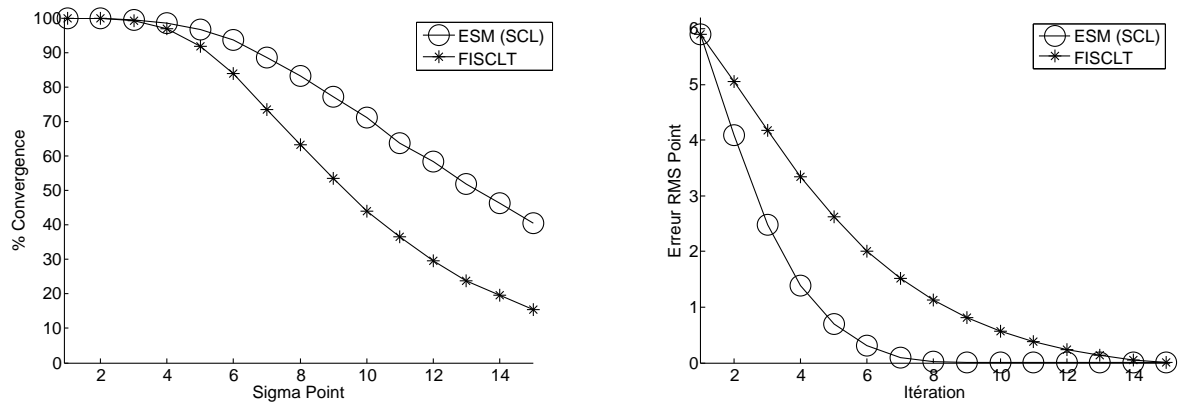


FIGURE E.2 – Fréquence moyenne de convergence et vitesse moyenne de convergence pour les approches FISCLT et ESM. Pour la vitesse de convergence Sigma Point = 6.

Dans un contexte de minimisation par l'algorithme Gauss-Newton, l'approche utilisant une symétrisation du problème par combinaison de deux métriques unidirectionnelles semble donc moins intéressante que l'approche utilisant une symétrisation au niveau des paramétrisations incrémentales.





# Approximation à l'ordre 2 de l'erreur ACL

---

Dans cette partie, nous présentons les propriétés utilisées pour obtenir l'approximation à l'ordre 2 de l'erreur ACL (cf. eq. (5.18)) dans des conditions idéales (non bruitées) :

$$\mathbf{e}_\alpha^{nf}(\mathbf{v}^*) = \mathbf{e}_\alpha^{nf}(\mathbf{0}) + \mathbf{J}_{\text{ESM}}^{nf} \mathbf{v}^* + O(\|\mathbf{v}^*\|^3) \quad (\text{F.1})$$

La démonstration repose sur les lemmes qui suivent.

**Lemme 4** (Développement à l'ordre 2 de l'erreur ACL).

$$\mathbf{e}_\alpha(\mathbf{v}) = \mathbf{e}_\alpha(\mathbf{0}) + \frac{1}{2}(\mathbf{J}_\alpha(\mathbf{0}) + \mathbf{J}_\alpha(\mathbf{v}))\mathbf{v} + O(\|\mathbf{v}\|^3). \quad (\text{F.2})$$

*Démonstration.* En développant  $\mathbf{e}_\alpha(\mathbf{v})$  (5.6) à l'ordre 2 autour de  $\mathbf{v} = \mathbf{0}$ , nous avons pour chaque pixel  $\mathbf{x}_i$  :

$$\mathbf{e}_{\alpha,i}(\mathbf{v}) = \mathbf{e}_{\alpha,i}(\mathbf{0}) + \mathbf{J}_{\alpha,i}(\mathbf{0})\mathbf{v} + \frac{1}{2}\mathbf{v}^t \mathbf{H}_{\alpha,i}(\mathbf{0})\mathbf{v} + O(\|\mathbf{v}\|^3). \quad (\text{F.3})$$

avec  $\mathbf{J}_{\alpha,i}(\mathbf{0})$  la matrice Jacobienne et  $\mathbf{H}_{\alpha,i}(\mathbf{0})$  la matrice Hessienne de  $\mathbf{e}_{\alpha,i}$  en  $\mathbf{0}$ .

Nous avons également :

$$\mathbf{J}_{\alpha,i}(\mathbf{v}) = \mathbf{J}_{\alpha,i}(\mathbf{0}) + \mathbf{v}^t \mathbf{H}_{\alpha,i}(\mathbf{0}) + O(\|\mathbf{v}\|^2). \quad (\text{F.4})$$

Ainsi, en utilisant (F.4) dans (F.3), nous obtenons (F.2). Cette linéarisation est valide dans le cas non-bruité aussi bien que dans les cas bruités.  $\square$

**Lemme 5** (Equivalence des matrices Jacobiennes le long de  $\mathbf{v}^*$ ).

$$\mathbf{J}_I^{nf}((1-\alpha)\mathbf{v}^*)\mathbf{v}^* = \mathbf{J}_T^{nf}(-\alpha\mathbf{v}^*)\mathbf{v}^*. \quad (\text{F.5})$$

*Démonstration.* En utilisant les propriétés des dérivées directionnelles dans la direction  $\mathbf{v}^*$ , l'hypothèse d'égalité des deux images à une compensation géométrique près (5.16) et les propriétés de l'application exponentielle (2.17)(2.18), nous avons :

$$\begin{aligned}
\mathbf{J}_{I,i}^{nf}((1-\alpha)\mathbf{v}^*)\mathbf{v}^* &= \left. \frac{\partial I^{nf}(\mathbf{W}(\boldsymbol{\mu}^k \circ \exp((1-\alpha)\mathbf{v}^* + t\mathbf{v}^*), \mathbf{x}_i))}{\partial t} \right|_{t=0} \\
&= \left. \frac{\partial I^{nf}(\mathbf{W}(\exp(-\alpha\mathbf{v}^* + t\mathbf{v}^*), \mathbf{x}_i))}{\partial t} \right|_{t=0} \\
&= \mathbf{J}_{T,i}^{nf}(-\alpha\mathbf{v}^*) \left. \frac{\partial(-\alpha\mathbf{v}^* + t\mathbf{v}^*)}{\partial t} \right|_{t=0} \\
&= \mathbf{J}_{T,i}^{nf}(-\alpha\mathbf{v}^*)\mathbf{v}^*.
\end{aligned} \tag{F.6}$$

□

**Corollaire.**

$$\mathbf{J}_{\alpha}^{nf}(\mathbf{v}^*)\mathbf{v}^* = \mathbf{J}_I^{nf}((1-\alpha)\mathbf{v}^*)\mathbf{v}^* = \mathbf{J}_T^{nf}(-\alpha\mathbf{v}^*)\mathbf{v}^*, \tag{F.7}$$

avec en particulier :

$$\mathbf{J}_I^{nf}(\mathbf{v}^*)\mathbf{v}^* = \mathbf{J}_T^{nf}(\mathbf{0})\mathbf{v}^* \tag{F.8}$$

$$\mathbf{J}_T^{nf}(-\mathbf{v}^*)\mathbf{v}^* = \mathbf{J}_I^{nf}(\mathbf{0})\mathbf{v}^*. \tag{F.9}$$

*Démonstration.* Nous avons :

$$\mathbf{J}_{\alpha}^{nf}(\mathbf{v}^*)\mathbf{v}^* = (1-\alpha)\mathbf{J}_I^{nf}((1-\alpha)\mathbf{v}^*)\mathbf{v}^* + \alpha\mathbf{J}_T^{nf}(-\alpha\mathbf{v}^*)\mathbf{v}^*. \tag{F.10}$$

En utilisant le Lemme F.5 dans l'équation précédente afin de remplacer  $\mathbf{J}_T^{nf}(-\alpha\mathbf{v}^*)\mathbf{v}^*$ , il vient (F.7). Les cas particuliers sont obtenus avec  $\alpha = 0$  and  $\alpha = 1$ . □

Remarquons que nous proposons ici une extension de l'équivalence (F.8), démontrée par [Benhimane 2007b], sur la matrice Jacobienne générique de l'ACL.

**Lemme 6** (Approximation linéaire de la matrice Jacobienne).

Pour  $\gamma \in [0, 1]$ , nous avons :

$$\mathbf{J}_I(\gamma\mathbf{v}) = \mathbf{J}_I(\mathbf{0}) + \gamma(\mathbf{J}_I(\mathbf{v}) - \mathbf{J}_I(\mathbf{0})) + O(\|\mathbf{v}\|^2). \tag{F.11}$$

*Démonstration.* En développant  $\mathbf{J}_{I,i}(\gamma\mathbf{v})$  à l'ordre 1 autour de  $\mathbf{0}$  nous avons :

$$\mathbf{J}_{I,i}(\gamma\mathbf{v}) - \mathbf{J}_{I,i}(\mathbf{0}) = \gamma\mathbf{v}^t \left. \frac{\partial^2 I(\mathbf{W}(\boldsymbol{\mu}^k \circ \exp(\mathbf{v}), x_i))}{\partial \mathbf{v}^2} \right|_{\mathbf{v}=\mathbf{0}} + O(\|\gamma\mathbf{v}\|^2). \tag{F.12}$$

□

**Lemme 7.**

$$\mathbf{J}_{\alpha}^{nf}(\mathbf{v}^*)\mathbf{v}^* = (\alpha\mathbf{J}_I^{nf}(\mathbf{0}) + (1-\alpha)\mathbf{J}_T^{nf}(\mathbf{0}))\mathbf{v}^* + O(\|\mathbf{v}^*\|^3) \tag{F.13}$$

---

*Démonstration.*

$$\begin{aligned} \mathbf{J}_\alpha^{nf}(\mathbf{v}^*)\mathbf{v}^* &\stackrel{\text{(F.7)}}{=} \mathbf{J}_I^{nf}((1-\alpha)\mathbf{v}^*)\mathbf{v}^* \\ &\stackrel{\text{(F.11)}}{=} (\mathbf{J}_I^{nf}(\mathbf{0}) + (1-\alpha)(\mathbf{J}_I^{nf}(\mathbf{v}^*) - \mathbf{J}_I^{nf}(\mathbf{0})))\mathbf{v}^* + O(\|\mathbf{v}^*\|^3) \\ &\stackrel{\text{(F.8)}}{=} (\alpha\mathbf{J}_I^{nf}(\mathbf{0}) + (1-\alpha)\mathbf{J}_T^{nf}(\mathbf{0}))\mathbf{v}^* + O(\|\mathbf{v}^*\|^3) \end{aligned} \quad (\text{F.14})$$

□



# Bibliographie

- [Alter 2006] F. Alter, Y. Matsushita et X. Tang. *An intensity similarity measure in low-light conditions*. In European Conference on Computer Vision, volume 3954, pages 267–280, 2006.
- [Amberg 2009] B. Amberg, A. Blake et T. Vetter. *On compositional Image Alignment, with an application to Active Appearance Models*. In Conference on Computer Vision and Pattern Recognition, pages 1714–1721, 2009.
- [Averbuch 2003] A. Averbuch et Y. Keller. *Fast Gradient Methods Based Global Motion Estimation for Video Compression*. IEEE Trans. on Circ. and Sys. for Vid. Tech, vol. 13, pages 300–309, 2003.
- [Baker 2001] S. Baker et I. Matthews. *Equivalence and Efficiency of Image Alignment Algorithms*. In IEEE Conference on Computer Vision and Pattern Recognition, pages 1090–1097, Kauai, HI, USA, 2001.
- [Baker 2003a] S. Baker, R. Gross, I. Matthews et T. Ishikawa. *Lucas-Kanade 20 Years On : A Unifying Framework : Part 2*. Rapport technique CMU-RI-TR-03-01, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, February 2003.
- [Baker 2003b] Simon Baker, Ralph Gross, Iain Matthews et Takahiro Ishikawa. *Lucas-Kanade 20 Years On : A Unifying Framework : Part 2*. Rapport technique CMU-RI-TR-03-01, Robotics Institute, Pittsburgh, PA, March 2003.
- [Baker 2004] S. Baker et I. Matthews. *Lucas-Kanade 20 Years On : A Unifying Framework*. International Journal of Computer Vision, vol. 56, no. 3, pages 221–255, 2004.
- [Bartoli 2008] A. Bartoli. *Groupwise Geometric and Photometric Direct Image Registration*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, pages 2098–2108, 2008.
- [Bayro-Corrochano 2007] E. Bayro-Corrochano et J. Ortégón-Aguilar. *Lie algebra approach for tracking and 3D motion estimation using monocular vision*. Image and Vision Computing, vol. 25, pages 907–921, 2007.
- [Benhimane 2004] S. Benhimane et E. Malis. *Real-time image-based tracking of planes using Efficient Second-order Minimization*. In IEEE/RSJ International Conference on Intelligent Robots and Systems, volume 1, pages 943–948, Sendai, Japan, 2004.
- [Benhimane 2006] M. S. Benhimane. *Vers une approche unifiée pour le suivi temps-réel et l’asservissement visuel*. PhD thesis, INRIA Sophia Antipolis, 2006.
- [Benhimane 2007a] S. Benhimane, A. Ladikos, V. Lepetit et N. Navab. *Linear and Quadratic Subsets for Template-Based Tracking*. IEEE Conference on Computer Vision and Pattern Recognition, vol. 0, pages 1–6, 2007.
- [Benhimane 2007b] S. Benhimane et E. Malis. *Homography-based 2D Visual Tracking and Servoing*. International Journal of Robotics Research, vol. 26, no. 7, pages 661–676, July 2007.

- [Black 1993] Michael J. Black et P. Anandan. *A framework for the robust estimation of optical flow*. In International Conference on Computer Vision, pages 231–236, 1993.
- [Black 1996] M. J. Black et A. Rangarajan. *On the unification of line processes, outlier rejection, and robust statistics with applications in early vision*. Int. J. Comput. Vision, vol. 19, no. 1, pages 57–91, 1996.
- [Brooks 2007] R. Brooks, T. Arbel et D. Precup. *Fast image alignment using anytime algorithms*. In International joint conference on Artificial intelligence, pages 2078–2083, San Francisco, CA, USA, 2007. Morgan Kaufmann Publishers Inc.
- [Brooks 2009] R. Brooks et T. Arbel. *Generalizing Inverse Compositional and ESM Image Alignment*. International Journal of Computer Vision, vol. 87, 2009.
- [Brown 1992] L. G. Brown. *A Survey of Image Registration Techniques*. ACM Computing Surveys, vol. 24, pages 325–376, 1992.
- [Buenaposada 2002] J. M. Buenaposada et L. Baumela. *Real-Time Tracking and Estimation of Plane Pose*. International Conference on Pattern Recognition, vol. 2, pages 697–700, 2002.
- [Cachier 2000] P. Cachier et D. Rey. *Symmetrization of the Non-rigid Registration Problem Using Inversion-Invariant Energies : Application to Multiple Sclerosis*. In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 472–481, 2000.
- [Cachier 2003] P. Cachier, E. Bardinet, D. Dormont, X. Pennec et N. Ayache. *Iconic Feature Based Nonrigid Registration : The PASHA Algorithm*. Computer Vision and Image Understanding, vol. 89, no. 2-3, pages 272–298, Feb.-March 2003.
- [Chen 2009] Y. Chen, W. Jin, L. Zhao et F. Li. *A subpixel motion estimation algorithm based on digital correlation for illumination variant and noise image sequences*. Optik - International Journal for Light and Electron Optics, vol. 120, no. 16, pages 835 – 844, 2009.
- [Christensen 1999] G. E. Christensen. *Consistent Linear-Elastic Transformations for Image Matching*. In Information Processing in Medical Imaging, volume 1613, pages 224–237, 1999.
- [Christensen 2001] G. E. Christensen et H. J. Johnson. *Consistent Image Registration*. IEEE Transactions on Medical Imaging, vol. 20, pages 568–582, 2001.
- [Comaniciu 2003] D. Comaniciu, V. Ramesh et P. Meer. *Kernel-Based Object Tracking*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 5, pages 564–575, 2003.
- [Dai 2006] S. Dai, M. Yang, Y. Wu et A. K. Katsaggelos. *Tracking Motion-Blurred Targets in Video*, 2006.
- [Dame 2009] A. Dame et E. Marchand. *Entropy-based visual servoing*. In IEEE International Conference on Robotics and Automation, pages 1970–1976, Piscataway, NJ, USA, 2009. IEEE Press.

- [Dedeoglu 2007] G. Dedeoglu, T. Kanade et S. Baker. *The Asymmetry of Image Registration and Its Application to Face Tracking*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 5, pages 607–623, May 2007.
- [Dowson 2008] N. Dowson et R. Bowden. *Mutual Information for Lucas-Kanade Tracking (MILK) : An Inverse Compositional Formulation*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, pages 180–185, 2008.
- [Drummond 1999a] T. Drummond et R. Cipolla. *Application of Lie Algebras to Visual Servoing*. International Journal of Computer Vision, vol. 37, pages 21–41, 1999.
- [Drummond 1999b] T. Drummond et R. Cipolla. *Visual Tracking and Control using Lie Algebras*. In IEEE Conference on Computer Vision and Pattern Recognition, pages 652–657. IEEE Computer Society Press, 1999.
- [Dufaux 2000] F. Dufaux et J. Konrad. *Efficient, Robust and Fast Global Motion Estimation for Video Coding*. Transactions on Image Processing, vol. 9, pages 497–501, 2000.
- [Evangelidis 2008] G. D. Evangelidis et E. Z. Psarakis. *Parametric Image Alignment Using Enhanced Correlation Coefficient Maximization*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, pages 1858–1865, 2008.
- [Farsiu 2005a] S. Farsiu. *A Fast and Robust Framework for Image Fusion and Enhancement*. PhD thesis, University of California, Santa Cruz, 2005.
- [Farsiu 2005b] S. Farsiu, M. Elad et P. Milanfar. *Constrained, globally optimal, multi-frame motion estimation*. In IEEE Workshop on Statistical Signal Processing, pages 1396–1401, 2005.
- [Faugeras 2001] O. Faugeras, Q.-T. Luong et T. Papadopoulou. *The geometry of multiple images : The laws that govern the formation of images of a scene and some of their applications*. MIT Press, Cambridge, MA, USA, 2001.
- [Fischler 1981] Martin A. Fischler et Robert C. Bolles. *Random Sample Consensus : A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography*. Communications of the ACM, vol. 24, no. 6, pages 381–395, 1981.
- [Gay-Bellile 2007] V. Gay-Bellile, A. Bartoli et P. Sayd. *Feature-driven direct non-rigid image registration*. In British Machine Vision Conference, 2007.
- [Gleicher 1997] M. Gleicher. *Projective registration with difference decomposition*. In IEEE Conference on Computer Vision and Pattern Recognition, page 331, 1997.
- [Guangwei 2008] L. Guangwei, L. Yunpeng, Y. Jian et S. Zelin. *Optimization on Lie Manifolds and Projective Tracking*. Computer Science and Software Engineering, International Conference on, vol. 1, pages 768–771, 2008.
- [Guskov 2006] I. Guskov. *Kernel-based Template Alignment*. In IEEE Conference on Computer Vision and Pattern Recognition, pages 610–617, 2006.
- [Hager 1998] G. D. Hager et P. N. Belhumeur. *Efficient Region Tracking With Parametric Models of Geometry and Illumination*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, no. 10, pages 1025–1039, 1998.



- [Hager 2004] Gregory D. Hager, Maneesh Dewan et Charles V. Stewart. *Multiple Kernel Tracking with SSD*. In IEEE Conference on Computer Vision and Pattern Recognition, pages 790–797, Washington, DC, USA, 2004.
- [Harker 2005] M.J. Harker et P.L. O’Leary. *Computation of homographies*, 2005.
- [Harris 1988] C. Harris et M. Stephens. *A Combined Corner and Edge Detection*. In Proceedings of The Fourth Alvey Vision Conference, pages 147–151, 1988.
- [Hartley 1997] R. I. Hartley. *In Defense of the Eight-Point Algorithm*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, pages 580–593, 1997.
- [Hartley 2004] R. I. Hartley et A. Zisserman. *Multiple view geometry in computer vision*. Cambridge University Press, ISBN : 0521540518, second édition, 2004.
- [Jin 2005] H. Jin, P. Favaro et R. Cipolla. *Visual Tracking in the Presence of Motion Blur*. In IEEE Conference on Computer Vision and Pattern Recognition, volume 2, pages 18–25, Los Alamitos, CA, USA, 2005. IEEE Computer Society.
- [Joshi 2004] S. Joshi, B. Davis, M. Jomier et G. Gerig. *Unbiased diffeomorphic atlas construction for computational anatomy*. Neuroimage, vol. 23, pages 151–160, 2004.
- [Jurie 2002a] F. Jurie et M. Dhome. *Hyperplane approximation for template matching*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 7, pages 996–1000, 2002.
- [Jurie 2002b] Frederic Jurie et Michel Dhome. *Real Time Robust Template Matching*. In British Machine Vision Conference, pages 123–131, 2002.
- [Keller 2004] Y. Keller et A. Averbuch. *Fast motion estimation using bi-directional gradient methods*. IEEE Transactions on Image Processing, vol. 13, no. 8, pages 1042–1054, August 2004.
- [Keller 2008] Y. Keller et A. Averbuch. *Global parametric image alignment via high-order approximation*. Computer Vision and Image Understanding, vol. 109, no. 3, pages 244–259, 2008.
- [Lieberknecht 2009] S. Lieberknecht, S. Benhimane, P. Meier et N. Navab. *A dataset and evaluation methodology for template-based tracking algorithms*. In International Symposium on Mixed and Augmented Reality, pages 145–151, 2009.
- [Lindeberg 1998] T. Lindeberg. *Feature Detection with Automatic Scale Selection*. International Journal of Computer Vision, vol. 30, no. 2, pages 79–116, November 1998.
- [Lowe 2004] D. G. Lowe. *Distinctive Image Features from Scale-Invariant Keypoints*. International Journal of Computer Vision, vol. 60, pages 91–110, 2004.
- [Lucas 1981] B.D. Lucas et T. Kanade. *An Iterative Image Registration Technique with an Application to Stereo Vision*. In International Joint Conference on Artificial Intelligence, pages 674–679, 1981.
- [Maes 1997] F. Maes, A. Collignon, D. Vandermeulen, Guy Marchal et Paul Suetens. *Multimodality Image Registration by Maximization of Mutual Information*. IEEE transactions on Medical Imaging, vol. 16, pages 187–198, 1997.

- [Malis 2005] E. Malis et S. Benhimane. *A unified approach to visual tracking and servoing*. Robotics and Autonomous Systems, vol. 52, pages 39–52, 2005.
- [Malis 2006] E. Malis et E. Marchand. *Experiments with robust estimation techniques in real-time robot vision*. In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, 2006.
- [Malis 2009] E. Malis. Vision-based estimation and robot control. INRIA, 2009.
- [Mann 1997] S. Mann et R. W. Picard. *Video Orbits of the Projective Group : A Simple Approach to Featureless Estimation of Parameters*. Transactions on Image Processing, vol. 6, pages 1281–1295, 1997.
- [Marchand 2005] Éric Marchand et François Chaumette. *Feature tracking for visual servoing purposes*. Robotics and Autonomous Systems, vol. 52, no. 1, pages 53–70, 2005.
- [Matthews 2004a] I. Matthews, T. Ishikawa et S. Baker. *The Template Update Problem*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, pages 810 – 815, 2004.
- [Matthews 2004b] Iain Matthews et Simon Baker. *Active Appearance Models Revisited*. International Journal of Computer Vision, vol. 60, no. 2, pages 135 – 164, November 2004.
- [Mégret 2006] R. Mégret, M. Mikram et Y. Berthoumieu. *Inverse Composition for Multi-kernel Tracking*. In Indian Conference on Computer Vision, Graphics and Image Processing, pages 480–491, 2006.
- [Mei 2008] C. Mei, S. Benhimane, E. Malis et P. Rives. *Efficient Homography-based Tracking and 3D Reconstruction for Single Viewpoint Sensors*. IEEE Transactions on Robotics, 2008.
- [Mikolajczyk 2004] K. Mikolajczyk et C. Schmid. *Scale & Affine Invariant Interest Point Detectors*. International Journal of Computer Vision, vol. 60, no. 1, pages 63–86, 2004.
- [Mikolajczyk 2005] K. Mikolajczyk et C. Schmid. *A performance evaluation of local descriptors*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 10, pages 1615–1630, 2005.
- [Pham 2005] T. Q. Pham, M. Bezuijen, L. J. van Vliet, K. Schutte et C. L. Luengo Hendriks. *Performance of optimal registration estimators*. In proc. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, volume 5817, pages 133–144, may 2005.
- [Pires 2005] B.E. Pires et P.M.Q. Aguiar. *Featureless Global Alignment of Multiple Images*. In International Conference on Image Processing, pages 57–60, 2005.
- [Sabuncu 2009] M. Sabuncu, B. Yeo, K. Van Leemput, T. Vercauteren et P. Golland. *Asymmetric Image-Template Registration*. In International Conference on Medical Image Computing and Computer-Assisted Intervention, volume 12, pages 565–573, 2009.
- [Schreiber 2007] D. Schreiber. *Robust template tracking with drift correction*. Pattern Recognition Letters, vol. 28, no. 12, pages 1483 – 1491, 2007.
- [Shah 2008] C.A. Shah, Y. Sheng et L.C. Smith. *Automated Image Registration Based on Pseudoinvariant Metrics of Dynamic Land-Surface Features*. IEEE Transactions on Geoscience and Remote Sensing, vol. 46, 2008.

- [Shi 1994] J. Shi et C. Tomasi. *Good Features to Track*. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR'94), pages 593 – 600, 1994.
- [Shum 2000] H.-Y. Shum et R. Szeliski. *Construction of panoramic image mosaics with global and local alignment*. International Journal of Computer Vision, vol. 36, no. 2, pages 101–130, 2000.
- [Silveira 2008] G. Silveira, E. Malis et P. Rives. *An efficient direct approach to visual SLAM*. IEEE Transactions on Robotics, 2008.
- [Smith 1997] S. M. Smith et J. M. Brady. *SUSAN-A New Approach to Low Level Image Processing*. International Journal of Computer Vision, pages 45–78, May 1997.
- [Suri 2010] S. Suri et P. Reinartz. *Mutual-Information-Based Registration of TerraSAR-X and Ikonos Imagery in Urban Areas*. IEEE Transactions on Geoscience and Remote Sensing, vol. 48, pages 939–949, 2010.
- [Szeliski 2006] R. Szeliski. *Image alignment and stitching : a tutorial*. Foundations and Trends in Computer Graphics and Vision, vol. 2, no. 1, pages 1–104, 2006.
- [Tagare 2009] H. D. Tagare, D. Groisser et O. Skrinjar. *Symmetric Non-rigid Registration : A Geometric Theory and Some Numerical Techniques*. Journal of Mathematical Imaging and Vision, vol. 34, no. 1, pages 61–88, 2009.
- [TekMate 2000] TekMate. *PhotoFit : <http://www.photofit4panorama.com/>*, 2000.
- [Tomasi 1991] C. Tomasi et T. Kanade. *Detection and Tracking of Point Features*. Rapport technique CMU-CS-91-132, Carnegie Mellon University, April 1991.
- [Vercauteren 2007] T. Vercauteren, X. Pennec, E. Malis, A. Perchant et N. Ayache. *Insight into Efficient Image Registration Techniques and the Demons Algorithm*. In International Conference on Information Processing in Medical Imaging, volume 4584, pages 495–506, 2007.
- [Vercauteren 2009] T. Vercauteren, X. Pennec, A. Perchant et N. Ayache. *Diffeomorphic demons : Efficient non-parametric image registration*. NeuroImage, vol. 45, no. 2, pages 561–572, 2009.
- [Viola 1997] Paul Viola et William M. Wells III. *Alignment by Maximization of Mutual Information*. Int. J. Comput. Vision, vol. 24, no. 2, pages 137–154, 1997.
- [Watkinson 2004] J. Watkinson. *The mpeg handbook*. Elsevier Focal Press, second édition, 2004.
- [Xiao 2008] L. Xiao et P. Li. *Improvement on Mean Shift based tracking using second-order information*. In International Conference on Pattern Recognition, pages 1 –4, 8-11 2008.
- [Y. Ma 2004] J. Kosecka Y. Ma S. Soatto et S. Shankar Sastry. *An invitation to 3-d vision from image to geometric models*. Springer-Verlag, New York, USA, 2004.
- [Zitova 2003] B. Zitova et J. Flusser. *Image Registration Methods : A Survey*. Image and Vision Computing, vol. 21, no. 11, pages 977–1000, October 2003.