

Numéro d'ordre : 2562

# THÈSE

PRÉSENTÉE À

**L'UNIVERSITÉ BORDEAUX I**

ÉCOLE DOCTORALE DES SCIENCES PHYSIQUES ET DE L'INGÉNIEUR

PAR

**Daorolala Alexandre RANDRIANTSOA**

POUR OBTENIR LE GRADE DE

**DOCTEUR**

SPÉCIALITÉ : ÉLECTRONIQUE

\*\*\*\*\*

**ESTIMATION DU MOUVEMENT FONDÉE SUR  
UN SCHÉMA DIRECT ET RÉTROGRADE  
APPLICATION À LA SEGMENTATION**

\*\*\*\*\*

Soutenue le 27 Septembre 2002

Après avis de :

MM. Philippe BOLON,	Professeur, Université de Savoie	Rapporteur
Jack-Gérard POSTAIRE,	Professeur, Université des Sciences et Technologies de Lille	Rapporteur

Devant la commission d'examen formée de :

MM. André TOUBOUL,	Professeur, Université Bordeaux I	Président
Yannick BERTHOUMIEU,	Maître de conférences, ENSEIRB	Rapporteur
Philippe MARCHEGAY,	Professeur, ENSEIRB	Examineur
Philippe BOLON,	Professeur, Université de Savoie	Examineur
Jack-Gérard POSTAIRE,	Professeur, Université des Sciences et Technologies de Lille	Examineur
Mohamed NAJIM,	Professeur, ENSEIRB	Examineur

*À mes parents,  
Acceptant toutes les privations qui vous ont été imposées pour ma longue  
promotion, vous n'avez rien ménagé pour m'encourager à persévérer dans mes efforts,  
Veuillez trouver ici mon affection filiale.*

*À mes frères et sœurs,  
Les tribulations produisent la patience, la patience l'expérience,  
l'expérience l'espérance,*

*À tous mes amis,  
Votre présence en corps, âme et esprit devient un réconfort physique,  
spirituel et moral,  
Je ne suis satisfait sans vous dire merci.*



# Remerciements

*À Monsieur Le Professeur André TOUBOUL, directeur du Laboratoire d'Etudes de l'Intégration des Composants et Systèmes Electroniques (IXL), qui m'a permis d'effectuer cette thèse au sein du Laboratoire IXL,*

*Acceptez mes vifs remerciements pour votre accueil.*

*À Monsieur Le Professeur Philippe MARCHEGAY, directeur de l'Ecole Nationale Supérieure d'Electronique, Informatique et Radiocommunications de Bordeaux (ENSEIRB), mon Directeur de thèse, Qui m'a accueilli dans l'Equipe Synthèse de circuits et d'algorithmes numériques et Qui m'a soutenu dans l'accomplissement de ce travail,*

*Que vous trouviez ici l'assurance de ma profonde gratitude.*

*À Monsieur Yannick BERTHOUMIEU, maître de conférences à l'ENSEIRB et au laboratoire IXL, Qui a proposé et encadré cette thèse avec beaucoup de compétence, avec Qui j'ai plus étroitement travaillé, Qui, tout au long de cette étude, n'a cessé de me conseiller pour mener à bien ce travail, et Qui m'a apporté sa sympathie en m'accordant ses disponibilités au Laboratoire ou ailleurs dans un réel souci de coopération afin que je puisse progresser dans cette étude,*

*Veillez trouver ici l'expression de ma vive reconnaissance.*

*À tous les membres du jury, Monsieur Le Professeur Jack-Gérard POSTAIRE et Monsieur Le Professeur Philippe BOLON, Qui ont accepté d'être les rapporteurs de ce manuscrit, Qui ont diligemment commenté et évalué mon travail, ainsi qu'à Monsieur Le Professeur Mohamed NAJIM, Qui m'a fait l'honneur de participer au jury de soutenance,*

*Veillez accepter l'expression de tous mes remerciements.*

*Je tiens également à remercier les collaborateurs de la société Gaz de France pour leur collaboration à la réalisation de ce travail.*

*Mes remerciements vont également à tous mes collègues du Laboratoire, Qui ont contribué à la réalisation de ce travail par le partage de nombreuses idées.*

*Merci à tous.*



# Table des matières

<b>Liste des figures</b>	<b>xi</b>
<b>Liste des tableaux</b>	<b>xxi</b>
<b>Liste des principales notations</b>	<b>xxiii</b>
<b>Version abrégée</b>	<b>xxv</b>
<b>Abstract</b>	<b>xxvii</b>
<b>Introduction générale</b>	<b>1</b>
1 But .....	3
2 Vue d'ensemble de l'approche retenue .....	4
3 Organisation du manuscrit .....	5
<b>Première partie : État de l'art et contributions</b>	<b>7</b>
<b>Chapitre 1 : Estimation du mouvement</b>	<b>9</b>
1.1 Introduction .....	9
1.2 Quelques notions fondamentales en estimation du mouvement .....	9
1.2.1 Séquence numérique .....	9
1.2.2 Mouvement global, mouvement local .....	10
1.2.3 Champ bidimensionnel de déplacement, flot optique, mouvement apparent 2D .....	10
1.3 Modèle de mouvement .....	11
1.3.1 Description du mouvement 3D et sa projection .....	12
1.3.2 Modèles de mouvement 2D .....	14
1.3.3 Choix d'un modèle approprié .....	17
1.4 Techniques d'estimation du mouvement : rappel .....	18
1.4.1 Approches basées primitives .....	21
1.4.2 Approches basées flot optique .....	21
a) Hypothèse de conservation de la luminance .....	22
b) Hypothèse de cohérence spatiale .....	22
1.5 Estimation du flot optique selon l'approche différentielle .....	22
1.5.1 Équation de contrainte du flot optique .....	23
a) Choix de l'équation de contrainte .....	24
b) Techniques d'estimation des dérivées partielles .....	24
c) Méthode de résolution du système d'équation .....	25
d) Suggestions aux problèmes liés à l'équation du mouvement .....	26
1.5.2 Composante normale du vecteur vitesse .....	26
1.5.3 Méthodes non paramétriques (méthodes basées pixels) .....	27
a) Algorithme de Netravali et Robbins .....	27
b) Algorithme de Walker et Rao .....	29
c) Algorithme de Caffario et Rocca .....	29
d) Discussion .....	29

1.5.4 Méthodes basées sur le modèle paramétrique (méthodes basées régions) .....	30
a) Méthodes indirectes .....	30
b) Méthodes directes .....	30
c) Discussion .....	32
1.5.5 La notion de régularisation .....	32
1.5.6 Méthodes à contraintes globales .....	32
a) Méthode de Horn et Schunck .....	33
b) Mouvement le long des contours : méthode de Hildreth .....	33
c) Méthode de Nagel .....	34
d) Champ de Markov .....	36
1.5.7 Méthodes à contraintes locales .....	36
a) Estimation au sens des moindres carrés .....	36
b) Méthode de Lucas et Kanade .....	37
c) Méthodes différentielles du second ordre .....	38
1.6 Méthodes de transformée (basées énergie et basées phase) .....	40
1.6.1 Filtrage dans le domaine spatio-temporel .....	40
1.6.2 Transformée de Fourier (domaine fréquentiel) .....	42
1.6.3 Mesure de la variation de phase .....	44
1.6.4 Quelques méthodes de transformée .....	45
a) Méthode de Heeger .....	45
b) Méthode de Fleet et Jepson .....	45
1.7 Méthodes corrélatives .....	46
1.7.1 Différentes formes de l'indicateur de corrélation .....	47
1.7.2 Principales mesures de distance .....	47
1.8 Autres problèmes d'estimation du flot optique .....	49
1.8.1 Problème d'occlusion .....	49
1.8.2 Problème d'objet transparent .....	49
1.9 Résumé du chapitre .....	49
<b>Chapitre 2 : Segmentation d'une image</b> .....	<b>51</b>
2.1 Introduction .....	51
2.2 Segmentation spatio-temporelle selon l'approche descendante .....	52
2.3 Segmentation spatio-temporelle selon l'approche ascendante .....	53
2.3.1 Génération de l'ensemble des régions initiales .....	53
2.3.2 Mesure de similarité de régions .....	54
2.3.3 Stratégie de fusion de données .....	56
2.4 Résumé du chapitre .....	57
<b>Conclusion de l'état de l'art et orientation de l'étude</b> .....	<b>59</b>
<b>Deuxième Partie : Estimation du flot optique basée sur un schéma direct-rétrograde</b> .....	<b>61</b>
<b>Chapitre 1 : Schéma direct-rétrograde pour un mouvement de translation</b> .....	<b>63</b>
1.1 Introduction .....	63
1.2 Rappel sur la méthode des moindres carrés, application à l'estimation du mouvement basée sur un schéma direct .....	63
1.2.1 Survol sur la formulation .....	63

1.2.2 Résultats expérimentaux .....	65
a) Séquence de synthèse .....	65
b) Séquence réelle à déplacement synthétique .....	67
1.2.3 Discussion sur la méthode des moindres carrés basée sur un schéma simple ....	68
1.3 Méthode du tenseur : méthodes des moindres carrés totaux .....	69
1.3.1 Principes généraux .....	69
1.3.2 Résultats expérimentaux .....	70
1.3.3 Discussion sur la méthode des moindres carrés totaux basée schéma simple ....	71
1.4 Schéma direct-rétrograde, technique des moindres associée .....	71
1.4.1 Motivation .....	71
1.4.2 Extension de l'équation de contrainte du mouvement : notion de schéma direct-rétrograde .....	71
1.4.3 Résultats expérimentaux .....	73
a) Séquence synthétique : "spirale" .....	74
b) Séquence réelle à déplacement synthétique .....	77
c) Séquence réelle : "taxi" .....	78
1.4.4 Discussion sur l'estimation basée sur un schéma direct-rétrograde et la technique des moindres carrés associée .....	82
1.5 Estimation robuste basée sur un schéma direct-rétrograde .....	83
1.5.1 Description .....	83
1.5.2 Résultats expérimentaux .....	84
1.5.3 Discussion sur l'estimation robuste basée sur un schéma direct- rétrograde ....	86
1.6 Méthode du tenseur basée sur un schéma direct-rétrograde .....	87
1.6.1 Description .....	87
1.6.2 Résultats expérimentaux .....	88
a) Séquence synthétique : "spirale" .....	88
b) Séquence réelle à déplacement synthétique .....	91
1.6.3 Discussion sur la méthode du tenseur robuste basée sur un schéma direct-rétrograde .....	93
1.7 Approche incrémentale à formalisme direct-rétrograde .....	93
1.7.1 Rappel sur l'approche incrémentale .....	93
1.7.2 Approche incrémentale selon le formalisme direct-rétrograde .....	97
1.7.3 Approche incrémentale robuste selon le formalisme direct-rétrograde .....	99
1.7.4 Résultats expérimentaux .....	100
a) Expérimentation sur la séquence "spirale" .....	100
b) Expérimentation sur la séquence réelle à déplacement synthétique .....	101
c) Expérimentation sur la séquence "taxi" .....	105
1.7.5 Discussion sur l'approche incrémentale basée sur un schéma direct-rétrograde	107
1.8 Résumé du chapitre .....	107

## **Chapitre 2 : Formalisme direct-rétrograde pour un modèle de mouvement paramétrique** **109**

2.1 Introduction .....	109
2.2 Rappel sur le formalisme basé simple pour un modèle de mouvement paramétrique .	109
2.2.1 Principes généraux .....	109
2.2.2 Résultats expérimentaux .....	110
2.2.3 Discussion sur la méthode des moindres carrés appliquée au modèle de mouvement .....	110
2.3 Extension du schéma direct-rétrograde au modèle paramétrique .....	111



2.3.1 Description .....	111
2.3.2 Résultats expérimentaux .....	116
a) Analyse des résultats à partir du champ de déplacement .....	117
b) Introduction de la technique robuste dans l'estimation des paramètres du mouvement à base de schéma direct-rétrograde .....	123
c) Analyse des résultats à partir du biais d'estimation .....	124
d) Expérimentation sur une image en rotation .....	127
e) Expérimentation sur la séquence "nuts", mouvement de divergence .....	131
f) Expérimentation sur une séquence "yosemite", flux non homogène .....	135
2.3.3 Discussion sur les différentes approches basées sur un schéma direct-rétrograde pour un modèle de mouvement affine .....	140
2.4 Approche incrémentale et schéma direct-rétrograde pour un modèle paramétrique ...	140
2.5 Extension du schéma direct-rétrograde à $n$ images ( $n > 2$ ) .....	143
2.5.1 Description .....	143
2.5.2 Résultats expérimentaux .....	145
2.5.3 Discussion sur l'approche basée sur un schéma direct-rétrograde à $n$ images ...	145
2.6 Résumé du chapitre .....	146

**Conclusion de la deuxième partie** 149

**Troisième Partie : Segmentation spatio-temporelle associée à l'estimation du mouvement selon le formalisme direct-rétrograde** 151

**Chapitre 1 : Hypothèse markovienne en segmentation spatio-temporelle** 153

1.1 Introduction .....	153
1.2 Description .....	153
1.3 Résultats expérimentaux .....	159
1.4 MRF et estimation du mouvement basée sur un schéma direct-rétrograde .....	159
1.4.1 Description .....	159
1.4.2 Résultats expérimentaux .....	160
1.5 Segmentation et estimation simultanées à l'aide du MRF et du schéma direct-rétrograde .....	162
1.5.1 Description .....	162
1.5.2 Résultats expérimentaux .....	163
1.5.3 Apport du formalisme direct-rétrograde à la segmentation, comparaison avec le formalisme simple .....	164
1.6 Algorithme MRF associé au formalisme direct-rétrograde à $n$ images, application à la segmentation spatio-temporelle .....	165
1.6.1 Description .....	165
1.6.2 Résultats expérimentaux .....	165

**Chapitre 2 : Algorithme K-moyens associé au formalisme direct-rétrograde** 169

2.1 Description .....	169
2.2 Résultats expérimentaux .....	171

<b>Résumé de la troisième partie</b>	<b>173</b>
--------------------------------------	------------

<b>Quatrième Partie : Application à l'estimation du retard en sismique répétitive (projet en liaison avec la société GDF)</b>	<b>175</b>
---	------------

<b>Chapitre 1 : Estimation du retard en sismique répétitive</b>	<b>177</b>
---	------------

1.1 Introduction .....	177
1.2 Survol du principe utilisé .....	177
1.3 Étude et observation des supports numériques fournis .....	180
1.4 Modélisation du signal sismique .....	180
1.5 Différentes méthodes d'estimation du retard .....	181
1.5.1 Méthodes corrélatives .....	181
1.5.2 Méthode basée sur le bispectre ou spectre d'ordre trois .....	183
1.5.3 Méthode de l'intercorrélation généralisée .....	184
1.5.4 Méthode ROTH .....	184
1.5.5 Méthode SCOT .....	185
1.5.6 Méthode PHAT .....	185
1.5.7 Méthode de l'estimateur du maximum de vraisemblance .....	186
1.5.8 Discussion sur les différentes méthodes d'estimation de retard .....	186
1.6 Méthode utilisée par la société GDF .....	187
1.6.1 Description .....	187
1.6.2 Implantation .....	187
1.7 Méthodes dérivatives suivant une direction spatiale 1D .....	189
1.7.1 Méthode dérivative basée sur un mouvement direct des ondes .....	189
1.7.2 Méthode dérivative basée sur un mouvement direct-rétrograde des ondes .....	189
1.8 Méthodes dérivatives dans un espace spatial 2D .....	190
1.8.1 Pré-traitement des données .....	190
a) Objectif de l'opération de pré-traitement .....	190
b) Premier type de filtrage : filtrage bidimensionnel de type triangulaire à angle d'ouverture réglable .....	193
c) Second type de filtrage : filtre gaussien modulé de type elliptique .....	194
1.8.2 Résultats expérimentaux .....	194
a) Profil obtenu en utilisant le premier type de filtrage .....	194
b) Profil obtenu en utilisant le second type de filtrage .....	195

<b>Chapitre 2 : Segmentation spatio-temporelle en sismique répétitive</b>	<b>201</b>
---	------------

2.1 Introduction .....	201
2.2 Principe .....	201
2.3 Résultats préliminaires .....	202
2.4 Développement d'une variante .....	203

<b>Résumé de la quatrième partie</b>	<b>205</b>
--------------------------------------	------------

<b>Conclusion générale</b>	<b>207</b>
----------------------------	------------

1 Résumé de l'accomplissement .....	209
2 Extensions possibles .....	212

<b>Annexe</b>	<b>213</b>
<b>Annexe A : Quelques notions mathématiques utiles en estimation du mouvement</b>	<b>215</b>
A.1 Géométrie projective .....	215
A.2 Angle d'Euler .....	216
A.3 Application de la transformée de Fourier en estimation du mouvement .....	218
a) Théorème de translation .....	218
b) Application à l'estimation du mouvement .....	218
A.4 Analyse des caractéristiques fréquentielles d'une image .....	218
a) Analyse par la transformée de Fourier .....	218
b) Filtres de Gabor .....	220
A.5 Minimisation d'une erreur quadratique au sens des moindres carrés .....	221
A.6 Différents masques utilisés à l'estimation des gradients .....	222
<b>Annexe B : Rappel des outils mathématiques utiles dans le cadre de la segmentation</b>	<b>225</b>
B.1 Classification bayésienne .....	225
B.2 Champ de Markov .....	226
B.3 Classification basée sur l'algorithme K-moyens .....	227
B.4 Formalisme du maximum de vraisemblance .....	228
B.5 Formalisme MDL pour la détermination de la complexité du modèle .....	229
<b>Annexe C : Analyse des erreurs d'estimation associées à la méthode différentielle</b>	<b>231</b>
C.1 Analyse .....	231
a) Erreur due à l'estimation des gradients .....	231
b) Erreur due à l'hypothèse de cohérence .....	232
c) Erreur de propagation caractéristique du système linéaire : moindres carrés ....	234
d) Erreur angulaire .....	236
C.2 Conséquence des erreurs dans la mise en équation : le besoin de contraintes additionnelles .....	237
C.3 Critère de choix du support de calcul .....	238
C.4 Critère de choix des techniques d'estimation du gradient .....	239
<b>Bibliographie</b>	<b>241</b>

# Liste des figures

1.1	Structure lumineuse d'une image .....	10
1.2	Illustration du champ de déplacement bidimensionnel .....	11
1.3	Illustration du flot optique (mouvement apparent 2D) .....	11
1.4	Illustration du mouvement à six degrés de liberté d'un objet rigide dans un espace tridimensionnel .....	12
1.5	Géométrie projective et objet rigide .....	13
1.6	Illustration des différentes manières d'observer un phénomène .....	19
1.7	Différentes manières de former un classement des méthodes d'estimation du mouvement .....	19
1.8	Exemple de classement des différentes méthodes d'estimation du mouvement ...	20
1.9	Un autre exemple de hiérarchie pour classer les différentes méthodes d'estimation du mouvement .....	20
1.10	Droite de contrainte et information locale sur les gradients spatiaux .....	24
1.11	Illustration du problème d'ouverture .....	27
1.12	Comportement de la convergence d'un point vers la droite de contrainte, dans les approches pel-récurrentes .....	28
1.13	Sens d'orientation d'un mouvement dans l'espace temporel .....	41
1.14	Filtres spatio-temporels et équation de contrainte du mouvement pour un cas 1D	42
1.15	Plan ( $P$ ) formé par le spectre d'énergie d'une image en translation et le vecteur de déplacement associé $(v_x, v_y, 1)^T \perp (P)$ .....	43
1.16	Processus de mise en correspondance de bloc et indexation de données .....	48
1.17	Segmentation quadtree d'une image .....	54
2.1	Mouvement suivant les sens causal et anti-causal, formalisme direct-rétrograde ..	63
2.2	Déplacement direct d'un pixel entre deux instants .....	64
2.3	Une trame de la séquence de synthèse "spirale" .....	65
2.4	Résultat sur la séquence "spirale" ; comparaison entre le champ de déplacement estimé et le champ réel, pour la méthode des moindres carrés simple (basée sur un schéma direct) ; séquence bruitée par un bruit gaussien additif ( $SNR = 80 dB$ )	66
2.5	Influence de la taille de la fenêtre de calcul sur la valeur de l'estimation .....	67
2.6	Une trame de la séquence réelle à déplacement synthétique .....	67
2.7	Estimation du mouvement des objets ① et ② de la séquence réelle à déplacement synthétique ; méthode des moindres carrés basée sur un schéma simple ; séquence bruitée par un bruit gaussien additif ( $SNR = 100 dB$ ) .....	68
2.8	Agrandissement du champ estimé de la zone $S$ .....	68
2.9	Champ de déplacement estimé à l'aide de la méthode TLS sur la "spirale", séquence bruitée par un bruit gaussien additif ( $SNR = 80 dB$ ) .....	70
2.10	Déplacements engendrés par les mouvements direct et rétrograde .....	72
2.11	Champ de déplacement estimé à l'aide de l'approche des moindres carrés en schéma direct-rétrograde, avec la fonctionnelle $J_a$ ; séquence synthétique "spirale", bruitée par un bruit gaussien additif ( $SNR = 80 dB$ ) .....	74
2.12	Champ de déplacement estimé à l'aide de l'approche des moindres carrés en schéma direct-rétrograde, avec la fonctionnelle $J_b$ ; séquence synthétique "spirale", bruitée par un bruit gaussien additif ( $SNR = 80 dB$ ) .....	74

2.13	Biais d'estimation de la composante $u$ du vecteur vitesse en fonction du SNR, pour l'approche des moindres carrés (MC) basée sur un schéma simple, approche basée sur un schéma direct-rétrograde selon la fonctionnelle $J_a$ , et approche basée sur un schéma direct-rétrograde selon la fonctionnelle $J_b$ ; séquence "spirale" .....	75
2.14	Biais d'estimation de la composante $v$ du vecteur vitesse en fonction du SNR, pour l'approche moindres carrés basée sur un schéma simple, approche basée sur un schéma direct-rétrograde avec la fonctionnelle $J_a$ , et approche basée sur un schéma direct-rétrograde avec la fonctionnelle $J_b$ ; séquence "spirale" .....	76
2.15	Variance de l'erreur d'estimation pour la composante $u$ du vecteur vitesse en fonction du SNR, pour l'approche moindres carrés basée sur un schéma simple, approche basée sur un schéma direct-rétrograde selon la fonctionnelle $J_a$ , et approche basée sur un schéma direct-rétrograde selon la fonctionnelle $J_b$ ; séquence "spirale" .....	76
2.16	Variance de l'erreur d'estimation pour la composante $v$ du vecteur vitesse en fonction du SNR, pour l'approche moindres carrés basée sur un schéma simple, approche basée sur un schéma direct-rétrograde selon la fonctionnelle $J_a$ , et approche basée sur un schéma direct-rétrograde selon la fonctionnelle $J_b$ ; séquence "spirale" .....	77
2.17	Champ de déplacement estimé sur l'image réelle à déplacement synthétique, avec la méthode des moindres carrés basée sur un schéma direct-rétrograde, en utilisant la fonctionnelle $J_b$ .....	77
2.18	Champ de déplacement en faisant agrandissement de la zone $S$ .....	78
2.19	Séquence réelle dans laquelle quatre objets sont réellement en mouvement; séquence "taxi" .....	78
2.20	Détection du mouvement à l'aide de l'approche des moindres carrés basée sur un schéma simple; amplitude du déplacement; séquence "taxi" .....	79
2.21	Détection du mouvement à l'aide de l'approche des moindres carrés basée sur un schéma direct-rétrograde, avec la fonctionnelle $J_a$ ; amplitude du déplacement; séquence "taxi" .....	80
2.22	Détection du mouvement l'aide de l'approche des moindres carrés selon le formalisme direct-rétrograde, avec la fonctionnelle $J_b$ ; amplitude du déplacement; séquence "taxi" .....	80
2.23	Champ de déplacement de l'objet ① de la séquence "taxi", en utilisant la méthode des moindres carrés simple .....	81
2.24	Champ de déplacement de l'objet ① de la séquence "taxi", en utilisant la méthode des moindres carrés basée sur un schéma direct-rétrograde, avec la fonctionnelle $J_a$ .....	81
2.25	Champ de déplacement de l'objet ① de la séquence "taxi", en utilisant la méthode des moindres carrés basée sur un schéma direct-rétrograde, avec la fonctionnelle $J_b$ .....	82
2.26	Comparaison des biais d'estimation de la composante $u$ du vecteur vitesse sur la séquence "spirale"; méthode des moindres carrés basée sur un schéma direct, celle basée sur un schéma direct-rétrograde selon la fonctionnelle $J_a$ et celle basée sur un schéma direct-rétrograde selon $J_b$ .....	85
2.27	Comparaison des biais d'estimation de la composante $v$ du vecteur vitesse sur la séquence "spirale"; méthode des moindres carrés basée sur un schéma direct, celle basée sur un schéma direct-rétrograde selon la fonctionnelle $J_a$ et celle basée sur un schéma direct-rétrograde selon $J_b$ .....	85

2.28	Comparaison des variances de l'erreur d'estimation pour la composante $u$ du vecteur vitesse, sur la séquence "spirale" ; méthode des moindres carrés basée sur un schéma direct, celle basée sur un schéma direct-rétrograde avec la fonctionnelle $J_a$ et celle basée sur un schéma direct-rétrograde selon $J_b$ .....	86
2.29	Comparaison des variances de l'erreur d'estimation pour la composante $v$ du vecteur vitesse, sur la séquence "spirale" ; méthode des moindres carrés basée sur un schéma direct, celle basée sur un schéma direct-rétrograde avec la fonctionnelle $J_a$ et celle basée sur un schéma direct-rétrograde selon $J_b$ .....	86
2.30	Biais d'estimation pour la composante $u$ du vecteur vitesse de la séquence "spirale" ; méthode TLS robuste basée sur un schéma simple, méthode TLS robuste basée sur un schéma direct-rétrograde ( $J_a$ ) et méthode TLS robuste basée sur un schéma direct-rétrograde ( $J_b$ ) ; résultats au bout de 3 pondérations .....	89
2.31	Biais d'estimation pour la composante $v$ du vecteur vitesse de la séquence "spirale" ; méthode TLS robuste basée sur un schéma simple, méthode TLS robuste basée sur un schéma direct-rétrograde ( $J_a$ ) et méthode TLS robuste basée sur un schéma direct-rétrograde ( $J_b$ ) ; résultats au bout de 3 pondérations .....	89
2.32	Variance de l'erreur d'estimation pour la composante $u$ du vecteur vitesse de la séquence "spirale" ; méthode TLS robuste basée sur un schéma simple, méthode TLS robuste basée sur un schéma direct-rétrograde ( $J_a$ ) et méthode TLS robuste basée sur un schéma direct-rétrograde ( $J_b$ ) ; résultats au bout de 3 pondérations ..	90
2.33	Variance de l'erreur d'estimation pour la composante $v$ du vecteur vitesse de la séquence "spirale" ; méthode TLS robuste basée sur un schéma simple, méthode TLS robuste basée sur un schéma direct-rétrograde ( $J_a$ ) et méthode TLS robuste basée sur un schéma direct-rétrograde ( $J_b$ ) ; résultats au bout de 3 pondérations ..	90
2.34	Détection et estimation du mouvement de l'objet ① de la séquence réelle à déplacement synthétique ; séquence bruitée par un bruit gaussien additif ( $SNR = 100 dB$ ) ; combinaison de deux méthodes des moindres carrés robustes basée sur un schéma simple (TLS+LS) ; déplacement théorique de (-2,1) pixels ; résultats au bout de 3 pondérations .....	91
2.35	Extrait du champ de déplacement obtenu par la combinaison des méthodes LS+TLS simples ; agrandissement de la zone $S$ .....	92
2.36	Détection et estimation du mouvement de l'objet ① de la séquence réelle à déplacement synthétique ; séquence bruitée par un bruit gaussien additif ( $SNR = 100 dB$ ) ; combinaison de deux méthodes LS+TLS robustes basées sur un schéma direct-rétrograde ; déplacement théorique de (-2,1) pixels ; résultats au bout de 3 pondérations .....	92
2.37	Extrait du champ de déplacement obtenu par la combinaison des méthodes LS+TLS basées sur un schéma direct-rétrograde ; agrandissement de la zone $S$ ..	93
2.38	Interprétation de l'ECM .....	94
2.39	Interprétation de l'ECM décalée .....	95
2.40	Approche incrémentale basée sur un mouvement direct-rétrograde .....	98
2.41	Biais d'estimation de la composante $u$ de la vitesse pour l'approche MC robuste basée schéma simple, l'approche incrémentale basée schéma simple, l'approche MC basée schéma direct-rétrograde et l'approche incrémentale basée schéma direct-rétrograde ( $J_b$ ) .....	102
2.42	Biais d'estimation de la composante $v$ de la vitesse pour l'approche MC robuste basée schéma simple, l'approche incrémentale basée schéma simple, l'approche MC basée schéma direct-rétrograde et l'approche incrémentale basée schéma direct-rétrograde ( $J_b$ ) .....	102

2.43	Champ de déplacement de l'objet ① de la séquence réelle à déplacement synthétique ; séquence bruitée par un bruit gaussien additif ( $SNR = 100 dB$ ) ; résultat de la méthode des moindres carrés robuste, basée sur un schéma simple .	103
2.44	Champ de déplacement de l'objet ① de la séquence réelle à déplacement synthétique ; séquence bruitée par un bruit gaussien additif ( $SNR = 100 dB$ ) ; résultat de l'approche incrémentale pondérée, basée sur un schéma simple .....	103
2.45	Champ de déplacement de l'objet ① de la séquence réelle à déplacement synthétique ; séquence bruitée par un bruit gaussien additif ( $SNR = 100 dB$ ) ; méthode des moindres carrés robuste, basée sur un schéma direct-rétrograde ( $J_b$ )	104
2.46	Champ de déplacement de l'objet ① de la séquence réelle à déplacement synthétique ; séquence bruitée par un bruit gaussien additif ( $SNR = 100 dB$ ) ; résultat de l'approche incrémentale pondérée, basée sur un schéma direct-rétrograde ( $J_b$ ) .....	104
2.47	Champ de déplacement des objets ① et ② de la séquence "taxi", à l'aide de la méthode des moindres carrés basée sur un schéma simple .....	105
2.48	Champ de déplacement des objets ① et ② de la séquence "taxi", à l'aide de l'approche incrémentale basée sur un schéma simple .....	105
2.49	Champ de déplacement des objets ① et ② de la séquence "taxi", à l'aide de la méthode des moindres carrés basée sur un schéma direct-rétrograde ( $J_b$ ).....	106
2.50	Champ de déplacement des objets ① et ② de la séquence "taxi", à l'aide de l'approche incrémentale basée sur un schéma direct-rétrograde ( $J_b$ ) .....	106
2.51	Champ de déplacement de la séquence "spirale" pour un modèle affine ; méthode des moindres carrés basée sur un schéma simple, $SNR = 90 dB$ .....	111
2.52	Formalisme direct-rétrograde pour un modèle affine .....	111
2.53	Champ de déplacement pour la séquence "spirale", bruitée par un bruit gaussien additif ( $SNR = 90 dB$ ) ; le modèle de mouvement utilisé est de type affine dont les paramètres théoriques sont: $t_x = -1$ , $t_y = 1$ , $k = 0,1$ et $\theta = \pi/10$ ; résultat de la méthode des moindres carrés basée sur un schéma direct .....	117
2.54	Champ de déplacement pour la séquence "spirale", bruitée par un bruit gaussien additif ( $SNR = 90 dB$ ) ; le modèle de mouvement utilisé est de type affine dont les paramètres théoriques sont: $t_x = -1$ , $t_y = 1$ , $k = 0,1$ et $\theta = \pi/10$ ; résultat de la méthode des moindres carrés basée sur un schéma direct-rétrograde, avec la fonctionnelle $J_a$ , variante n°1 .....	118
2.55	Champ de déplacement pour la séquence "spirale", bruitée par un bruit gaussien additif ( $SNR = 90 dB$ ) ; le modèle de mouvement utilisé est de type affine dont les paramètres théoriques sont: $t_x = -1$ , $t_y = 1$ , $k = 0,1$ et $\theta = \pi/10$ ; résultat de la méthode des moindres carrés basée sur un schéma direct-rétrograde, avec la fonctionnelle $J_b$ , variante n°2 .....	119
2.56	Champ de déplacement pour la séquence "spirale", bruitée par un bruit gaussien additif ( $SNR = 90 dB$ ) ; le modèle de mouvement utilisé est de type affine dont les paramètres théoriques sont: $t_x = -1$ , $t_y = 1$ , $k = 0,1$ et $\theta = \pi/10$ ; résultat de la méthode des moindres carrés basée sur un schéma direct-rétrograde, avec la fonctionnelle $J_b$ , variante n°6 .....	119
2.57	Champ de déplacement pour la séquence "spirale", bruitée par un bruit gaussien additif ( $SNR = 90 dB$ ) ; résultat de la méthode des moindres carrés basée sur schéma simple ; le modèle de mouvement utilisé est de type affine dont les paramètres théoriques sont : $t_x = 0$ , $t_y = 0$ , $k = 0,3$ et $\theta = \pi/3$ .....	120

2.58	Champ de déplacement pour la séquence "spirale", bruitée par un bruit gaussien additif ( $SNR = 90\text{ dB}$ ), pour un modèle affine de forte rotation et divergence ; paramètres : $t_x = 0$ , $t_y = 0$ , $k = 0,3$ et $\theta = \pi/3$ ; résultat de la méthode des moindres carrés basée sur un schéma direct-rétrograde sans terme non linéaire (variante n°2) .....	120
2.59	Champ de déplacement pour la séquence "spirale", bruitée par un bruit gaussien additif ( $SNR = 90\text{ dB}$ ), pour un modèle affine de forte rotation et divergence ; résultat de la méthode des moindres carrés basée sur un schéma direct-rétrograde, variante n°4 ; paramètres : $t_x = 0$ , $t_y = 0$ , $k = 0,3$ et $\theta = \pi/3$ .....	121
2.60	Champ de déplacement pour la séquence "spirale", bruitée par un bruit gaussien additif ( $SNR = 90\text{ dB}$ ), pour un modèle affine dont les paramètres sont : $t_x = 4$ , $t_y = -3$ , $k = 0,3$ , $\theta = \pi/8$ ; résultat de la méthode des moindres carrés à base de flot direct .....	122
2.61	Champ de déplacement pour la séquence "spirale", bruitée par un bruit gaussien additif ( $SNR = 90\text{ dB}$ ), pour un modèle affine dont les paramètres sont : $t_x = 4$ , $t_y = -3$ , $k = 0,3$ , $\theta = \pi/8$ ; résultat de la méthode des moindres carrés à base de flot direct-rétrograde, sans prise en compte du terme non linéaire ; variante n°2 ..	122
2.62	Champ de déplacement pour la séquence "spirale", bruitée par un bruit gaussien additif ( $SNR = 90\text{ dB}$ ), pour un modèle affine dont les paramètres sont : $t_x = 4$ , $t_y = -3$ , $k = 0,3$ , $\theta = \pi/8$ ; résultat de la méthode des moindres carrés à base de flot direct-rétrograde, sans prise en compte du terme non linéaire selon la variante n°6 .....	123
2.63	Champ de déplacement réel, résultat de l'approche des moindres carrés pondérés (MCP) basée schéma simple et déplacement estimé à l'aide de l'approche MCP à flot direct-rétrograde ( $J_b$ ) ; le nombre d'itérations est égal à 2 .....	123
2.64	Champ de déplacement pour la séquence "spirale", pour les trois méthodes testées, avec les paramètres de mouvement suivants : $t_x = 2$ , $t_y = -1$ , $k = 0,2$ et $\theta = \pi/10$ .....	124
2.65	Évolution de l'erreur sur la composante de translation suivant $x$ pour les 3 méthodes testées ; courbes tracées en fonction du SNR ; séquence "spirale" .....	125
2.66	Évolution de l'erreur sur la composante de translation suivant $y$ pour les 3 méthodes testées ; courbes tracées en fonction du SNR ; séquence "spirale" .....	125
2.67	Évolution de l'erreur sur le paramètre de rotation pour les 3 méthodes testées ; courbes tracées en fonction du SNR ; séquence "spirale" .....	126
2.68	Évolution de l'erreur sur le paramètre de divergence pour les 3 méthodes testées ; courbes tracées en fonction du SNR ; séquence "spirale" .....	126
2.69	Une image en rotation autour d'un axe vertical, séquence "cube en rotation" .....	127
2.70	Champ de déplacement estimé avec la méthode des moindres carrés basée mouvement direct, selon un modèle à quatre paramètres ; séquence "cube en rotation" .....	128
2.71	Erreur de reconstruction du niveau de gris ; différence temporelle entre l'image $t$ et l'image interpolée $t + Te$ , interpolation avec les déplacements estimés à l'aide de la méthode des moindres carrés basée schéma simple ; séquence "cube en rotation" .....	128
2.72	Champ de déplacement estimé avec la méthode des moindres carrés basée sur un schéma direct-rétrograde, sans prise en compte du terme non linéaire selon la variante n°2 ; modèle à quatre paramètres, séquence "cube en rotation" ..	129



2.73	Erreur de reconstruction du niveau de gris ; différence temporelle entre l'image $t$ et l'image interpolée $t + Te$ , interpolation avec les déplacements estimés à l'aide de la méthode des moindres carrés basée schéma direct-rétrograde ; sans traitement du terme non linéaire ; séquence "cube en rotation" .....	129
2.74	Champ de déplacement estimé avec la méthode des moindres carrés basée sur un schéma direct-rétrograde, avec traitement du terme non linéaire selon la variante n°6 ; modèle à quatre paramètres ; séquence "cube en rotation" .....	130
2.75	Erreur de reconstruction du niveau de gris ; différence temporelle entre l'image $t$ et l'image interpolée $t + Te$ , interpolation avec les déplacements estimés à l'aide de la méthode des moindres carrés basée schéma direct-rétrograde, avec traitement du terme non linéaire ; séquence "cube en rotation" .....	130
2.76	Une image de la séquence "nuts" présentant un mouvement de divergence, source : <a href="http://www.inrialpes.fr/movi/pub/images/sequences/g3_vp_tz_s2">www.inrialpes.fr/movi/pub/images/sequences/g3_vp_tz_s2</a> .....	132
2.77	Estimation du flot optique dû au mouvement de la caméra ; champ de déplacement estimé à l'aide de la méthode des moindres carrés basée sur un schéma simple, modèle à quatre paramètres ; séquence "nuts" .....	132
2.78	Erreur de reconstruction du niveau de gris ; différence temporelle entre l'image $t$ et l'image interpolée $t + Te$ , interpolation avec les déplacements estimés à l'aide de la méthode des moindres carrés basée sur un schéma simple ; séquence "nuts" .....	133
2.79	Estimation du flot optique dû au mouvement de la caméra ; champ de déplacement estimé à l'aide de la méthode des moindres carrés basée sur un schéma direct-rétrograde, sans traitement du terme non linéaire selon la variante n°2 ; modèle à quatre paramètres ; séquence "nuts" .....	133
2.80	Erreur de reconstruction du niveau de gris ; différence temporelle entre l'image $t$ et l'image interpolée $t + Te$ , interpolation avec les déplacements estimés à l'aide de la méthode des moindres carrés basée sur un schéma direct-rétrograde, sans traitement du terme non linéaire ; séquence "nuts" .....	134
2.81	Estimation du flot optique dû au mouvement de la caméra ; champ de déplacement estimé à l'aide de la méthode des moindres carrés basée sur un schéma direct-rétrograde, avec traitement du terme non linéaire selon la variante n°6 ; modèle à quatre paramètres ; séquence "nuts" .....	134
2.82	Erreur de reconstruction du niveau de gris ; différence temporelle entre l'image $t$ et l'image interpolée $t + Te$ , interpolation avec les déplacements estimés à l'aide de la méthode des moindres carrés basée sur un schéma direct-rétrograde, avec traitement du terme non linéaire ; séquence "nuts" .....	135
2.83	Une image de la séquence "yosemite" présentant un mouvement global (mouvement de la caméra) et des mouvements locaux (déplacements de l'orage) .....	136
2.84	Champ de déplacement estimé avec la méthode des moindres carrés basée sur un schéma simple ; modèle à quatre paramètres ; séquence "yosemite" .....	136
2.85	Erreur de reconstruction du niveau de gris, différence temporelle entre l'image $t$ et l'image interpolée $t + Te$ , interpolation avec les déplacements estimés à l'aide de la méthode des moindres carrés basée sur un schéma simple ; séquence "yosemite" .....	137
2.86	Champ de déplacement estimé avec la méthode des moindres carrés basée sur un schéma direct-rétrograde ; terme non linéaire négligeable selon la variante n°2 ; modèle à quatre paramètres ; séquence "yosemite" .....	137

2.87	Erreur de reconstruction du niveau de gris ; différence temporelle entre l'image $t$ et l'image interpolée $t + Te$ , interpolation avec les déplacements estimés à l'aide de la méthode des moindres carrés basée sur un schéma direct-rétrograde, sans prise en compte du terme non linéaire ; séquence "yosemite" .....	138
2.88	Champ de déplacement estimé avec la méthode des moindres carrés basée sur un schéma direct-rétrograde ; terme non linéaire traité selon la variante n°6 ; modèle à quatre paramètres ; séquence "yosemite" .....	138
2.89	Erreur de reconstruction du niveau de gris ; différence temporelle entre l'image $t$ et l'image interpolée $t + Te$ , interpolation avec les déplacements estimés à l'aide de la méthode des moindres carrés basée sur un schéma direct-rétrograde, avec prise en compte du terme non linéaire ; séquence "yosemite" .....	139
2.90	Exemple d'estimation du flot optique à partir des trois images .....	144
2.91	Formalisme direct-rétrograde sur trois images .....	145
2.92	Champ de déplacement obtenu avec l'approche basée direct-rétrograde, en utilisant 5 images successives .....	146
3.1	Ensemble $\omega_p$ de voisinages de $p$ ; le système de voisinage noté $\Omega$ est formé de partie $\omega_p$ .....	154
3.2	Différentes configurations possibles de cliques .....	154
3.3	Segmentation au sens des niveaux de gris ; fond : <i>classe 0</i> ; la lune de l'arc : <i>classe 1</i> .....	155
3.4	Voisinage spatio-temporel et cliques binaires .....	158
3.5	Approche markovienne appliquée à la segmentation spatio-temporelle .....	158
3.6	Segmentation spatio-temporelle de la séquence "taxi", approche markovienne ; $\beta_p = 30$ , $\beta_b = 5$ , $\beta_f = 80$ , <i>seuil</i> = 4, $\mu = 15$ , $\sigma^2 = 20$ ; résultat au bout de 5 itérations .....	159
3.7	Segmentation selon le MRF associé à l'estimation du mouvement (formalisme basé sur un schéma direct-rétrograde) .....	161
3.8	Segmentation markovienne associée à l'estimation basée schéma direct-rétrograde ; $\beta_p = 30$ , $\beta_b = 5$ , $\beta_f = 80$ , <i>seuil</i> = 0.4, $\sigma^2 = 20$ ; résultat au bout de 5 itérations .....	162
3.9	Segmentation et estimation simultanées suivant l'approche markovienne couplée avec l'estimation du mouvement basée sur un schéma direct-rétrograde ; $\beta_p = 30$ , $\beta_b = 5$ , $\beta_f = 80$ , <i>seuil</i> = 0.4, $\sigma^2 = 20$ ; résultat au bout de 5 itérations .....	163
3.10	Segmentation et estimation simultanées suivant l'approche MRF couplée avec l'estimation du mouvement basée schéma simple ; $\beta_p = 30$ , $\beta_b = 5$ , $\beta_f = 80$ , <i>seuil</i> = 0.4, $\sigma^2 = 20$ ; résultat au bout de 5 itérations .....	164
3.11	Comparaison de deux approches en terme de convergence (approche basée sur un schéma simple et celle basée sur un schéma direct-rétrograde) ; nombre de classes en fonction du nombre d'itérations .....	165
3.12	Segmentation spatio-temporelle basée sur l'algorithme MRF et le formalisme direct-rétrograde à $n$ images .....	166
3.13	Calcul des gradients à partir de 5 images .....	167
3.14	Segmentation suivant le MRF associé à l'estimation du mouvement basée sur un schéma direct-rétrograde à $n$ images avec $n = 5$ , $\beta_p = 30$ , $\beta_b = 5$ , $\beta_f = 80$ , <i>seuil</i> = 0.4, $\sigma^2 = 20$ ; résultat au bout de 5 itérations .....	167

3.15	Les $k$ plans de résidus ; le $1^{er}$ plan est constitué par les résidus associés aux $m \times n$ points et calculés avec le déplacement $\vec{d}_1 \dots$ ; de même, le $k^{ème}$ plan représente les résidus aux $m \times n$ points et calculés avec le déplacement $\vec{d}_k \dots$	170
3.16	Algorithme K-moyens associé à l'estimation du mouvement basée sur un schéma direct-rétrograde ; les régions initiales sont initialisées à l'aide d'un ensemble de déplacements aléatoires	171
3.17	Image segmentée à l'aide de l'algorithme de K-moyens couplé avec une estimation robuste basée sur un schéma direct-rétrograde ; les régions initiales sont initialisées à l'aide de l'estimation à plusieurs niveaux de déplacement et cela, par utilisation de différents supports spatiaux	172
4.1	Méthode de prospection sismique réflexion classique	178
4.2	Principe utilisé par la société GDF	178
4.3	Visualisation des matrices des traces sans gaz et avec gaz	179
4.4	Visualisation d'une trace (n°700) sur les deux images (avant et après injection du gaz) pour une profondeur allant de 1 à 850	179
4.5	Visualisation du retard entre les deux images (avant et après injection du gaz), pour la trace n°700 et pour une profondeur entre 400 et 500 (zone où se trouve le réservoir)	180
4.6	Implantation des méthodes corrélatives	187
4.7	Algorithme d'estimation du retard pour une méthode corrélative	188
4.8	Formulation basée sur un mouvement direct des ondes	189
4.9	Formulation basée sur un mouvement direct-rétrograde des ondes ; présentation du déplacement direct	189
4.10	Formulation basée sur un mouvement direct-rétrograde des ondes ; présentation du déplacement rétrograde	190
4.11	Image sismique d'origine (sans filtrage) du profil MO1	191
4.12	Estimation du retard sur l'image sismique d'origine, sans pré-traitement des données ; composante suivant l'axe vertical (profondeur) du vecteur estimé à l'aide de la méthode dérivateur 2D et avec un schéma direct-rétrograde ( $J_b$ ) ; support de calcul : 81 pixels en profondeur et 21 pixels en nombre de traces	191
4.13	Profil cumulé du retard pour une estimation sans pré-traitement des données	192
4.14	Spectre 2D en module associé à l'image sismique du profil MO1	192
4.15	Gabarit du filtre bidimensionnel à angle d'ouverture réglable	193
4.16	Profil du retard estimé à l'aide de la méthode des moindres carrés à mouvement direct-rétrograde ; les données sont filtrées en utilisant le premier type de filtrage avec un angle d'ouverture de $2\pi/9$ , profil MO3	195
4.17	Profil du retard estimé à l'aide de la méthode des moindres carrés, basée sur un mouvement direct des ondes ; profil MO1_93_94 ; les données sont filtrées en utilisant le deuxième type de filtrage	196
4.18	Profil du retard estimé à l'aide de la méthode des moindres carrés, basée sur un mouvement direct des ondes ; profil MO2_93_94 ; les données sont filtrées en utilisant le deuxième type de filtrage	196
4.19	Profil du retard estimé à l'aide de la méthode des moindres carrés, basée sur un mouvement direct des ondes ; profil MO3_93_94 ; les données sont filtrées en utilisant le second type de filtrage	197
4.20	Profil du retard estimé à l'aide de la méthode des moindres carrés, basée sur un mouvement direct-rétrograde des ondes ; profil MO1_93_94 ; les données sont filtrées en utilisant le second type de filtrage	198

4.21	Profil du retard estimé à l'aide de la méthode des moindres carrés, basée sur un schéma direct-rétrograde des ondes ; profil MO2_93_94 ; les données sont filtrées en utilisant le second type de filtrage .....	198
4.22	Profil du retard estimé à l'aide de la méthode des moindres carrés, basée sur un mouvement direct-rétrograde des ondes ; profil MO3_93_94 ; les données sont filtrées en utilisant le second type de filtrage .....	199
4.23	Algorithme de classification en utilisant le principe des nuées dynamiques .....	202
4.24	Résultats de la segmentation spatio-temporelle en utilisant le principe des nuées dynamiques .....	202
4.25	Représentation des matrices de poids associés au calcul des déplacements .....	203
4.26	Processus du blocage des couches par décision sur la stabilité de la couverture spatiale de chacune des couches .....	204
4.27	Segmentation spatio-temporelle pour le profil MO2 .....	204
A.1	Modèle de géométrie projective .....	215
A.2	Modèle simplifié de géométrie projective .....	216
A.3	Angles d'Euler .....	217
C.1	Représentation graphique de l'écart entre le gradient spatial et le déplacement ...	233
C.2	Représentation graphique des droites de contraintes et le point d'intersection associé à la solution réelle du système .....	237
C.3	Plusieurs points d'intersection, effet des différentes sources d'erreurs lors de la résolution du système d'équations de contrainte du mouvement .....	238



# Liste des tableaux

1.1	Modèles de mouvement 2D et leur transformation associée .....	18
2.1	Comparaison des performances des différentes méthodes étudiées, pour un mouvement de translation .....	108
2.2	Erreur de reconstruction et variance de l'erreur pour les trois approches testées, expérimentation sur la séquence "cube en rotation" .....	131
2.3	Erreur de reconstruction et variance de l'erreur pour les trois approches testées ; expérimentation sur la séquence "nuts" .....	135
2.4	Erreur d'estimation et variance de l'erreur pour les trois approches testées, expérimentation sur la séquence "yosemite" .....	139
2.5	Évaluation des différentes approches pour l'estimation de paramètres du mouvement .....	141
2.6	Comparaison des performances des différentes méthodes pour un modèle de mouvement .....	147
2.7	Comparaison des approches des moindres carrés, basée sur un schéma simple et basée sur un schéma direct-rétrograde .....	150
2.8	Tableau comparatif des différentes méthodes d'estimation des paramètres de mouvement, méthodes basées sur le formalisme direct-rétrograde .....	150
4.1	Comparaison des différentes méthodes d'intercorrélation .....	186
A.1	Approximation des gradients spatiaux et temporels .....	223



## Liste des principales notations

$t$	: instant
$dt$	: temps infinitésimal
$T_e$	: temps d'échantillonnage
$p(x, y)$	: position d'un point $p$
$I(p, t)$	: luminance du point $p$ à l'instant $t$
$t_x, t_y$	: paramètres de translation suivant $x$ et suivant $y$
$k$	: paramètre de divergence
$\theta$	: paramètre de rotation
$\vec{d} = (dx, dy)^T$	: vecteur déplacement au point $(x, y)$
$\vec{V} = (u, v)^T$	: vecteur vitesse
$\Theta = (t_x, k, \theta, t_y)^T$	: vecteur de paramètres du mouvement
$\vec{\nabla}$	: vecteur gradient
$\vec{d}_d$	: déplacement direct
$\vec{d}_r$	: déplacement rétrograde
$\varepsilon_d$	: erreur de prédiction associée au déplacement direct
$\varepsilon_r$	: erreur de prédiction associée au déplacement rétrograde
$J$	: fonctionnelle de l'erreur à minimiser par une méthode d'optimisation
$\rho$	: résidus





## Version abrégée

Les télécommunications numériques sont aujourd'hui en pleine évolution grâce aux progrès des technologies de l'intégration électronique et de l'algorithmique. En effet, les moyens quotidiens de communication sont à la veille de grands bouleversements avec l'intégration d'algorithmes de traitements de l'image et du son offrant une grande souplesse dans la manipulation de ces sources d'information. Dans le cadre des actions scientifiques engagées par le groupe Télécommunication du laboratoire de microélectronique IXL, le développement d'algorithmes et de circuits numériques dédiés aux traitements vidéo est un axe récent.

L'intégration de traitements dédiés au signal vidéo est une problématique importante dans de nombreux domaines d'application comme la vidéosurveillance, la compression et l'édition vidéo. En particulier, la gestion temporelle des objets présents dans la scène permet de concevoir des systèmes automatiques de gestion de l'information. Au sein du laboratoire, nous abordons deux grands volets. Le premier est relatif à la création de nouveaux algorithmes dédiés à la modélisation et à l'estimation des primitives permettant de caractériser les objets en mouvement. Le second traite de la problématique inverse à savoir la mise à jour des zones statiques de la scène afin d'extraire les zones mobiles.

Le travail développé dans cette thèse porte sur l'étude des méthodes d'estimation du mouvement basées sur des modèles, dans le contexte de l'analyse de séquence d'images. Plus précisément, le but de ce manuscrit est le développement d'algorithmes d'estimation et de segmentation du mouvement, qui tendent à déterminer les changements relatifs de position des objets visibles dans une scène. Le but d'une telle technique est de décrire le mouvement des régions de la scène, sans connaissance a priori des mouvements, ni des objets, ni de la caméra. De nombreuses techniques ont été développées dans ce cadre. Chacune d'entre elles a ses avantages et ses inconvénients. Un problème majeur de l'estimation du mouvement est l'insuffisance des contraintes pour rendre l'estimation robuste. Dans notre cas, nous proposons un formalisme qui permet de renforcer les contraintes dans la mise en équation du mouvement. Cette proposition s'intitule : "Estimation du mouvement fondée sur un schéma direct et rétrograde". Le formalisme est construit à partir de l'information spatio-temporelle causale et anti-causale (mouvements direct et rétrograde) de deux trames consécutives.

Le but de notre technique est d'imaginer une abstraction à travers certaines hypothèses : les techniques d'estimation du mouvement se basent sur la comparaison de deux ou plusieurs trames. Si un objet change d'intensité le long de sa trajectoire et si la luminance associée à cet objet évolue de façon non modélisable, cette comparaison devient impossible. Bien que des algorithmes estimant la variation de luminance aient été proposés comme règle générale, seule la variation proportionnelle de luminance est prise en compte dans les techniques d'estimation du mouvement. Par la suite, nous prenons comme hypothèse qu'il n'y a pas de variation de luminance et que la variation d'intensité est due uniquement aux mouvements direct et/ou rétrograde de pixels. La modélisation mathématique de cette hypothèse permet de définir un critère à minimiser par une méthode d'optimisation. L'estimation du mouvement découle de la minimisation de ce critère. Différents types de solutions sont explorés. Ces approches sont basées sur la technique des moindres carrés à laquelle est associé le schéma "direct-rétrograde".

De nombreuses expérimentations ont été effectuées afin d'évaluer la qualité des approches proposées. Des séquences réelles et synthétiques ont été utilisées. Les résultats ont permis

d'illustrer la différence entre l'approche utilisant un schéma simple et celle utilisant le schéma direct-rétrograde. Dans un premier temps, le formalisme direct-rétrograde est testé sur un mouvement de translation. Le second volet dédié à l'estimation consiste en l'extension du formalisme proposé sur un modèle plus général. Il s'agit du modèle paramétrique. Ce modèle est une description plus proche du cas réel car plus riche dans les phénomènes reproduits. A la différence des méthodes précédentes, les paramètres de rotation et de divergence sont pris en compte, et cela dans les deux sens. La modélisation en schéma direct-rétrograde est plus complexe que celle en schéma simple. Dans ce second volet, nous déclinons différentes méthodes dédiées à l'estimation des paramètres du mouvement selon le schéma direct-rétrograde. Les expérimentations sont faites sur des séquences de synthèse et réelles. Nous avons focalisé nos tests sur l'évaluation du schéma proposé par rapport au schéma classique.

Conformément au problème de segmentation spatio-temporelle, nous décrivons des méthodes s'appuyant sur l'estimation du mouvement fondée sur un schéma direct-rétrograde. Dans ce cadre, nous nous intéressons à l'application du formalisme intégré dans le processus de segmentation spatio-temporelle. Deux catégories de méthode sont évoquées. La première est l'application du schéma direct-rétrograde à la segmentation markovienne. Le principe est d'associer le processus markovien à l'estimation basée sur un schéma direct-rétrograde. La deuxième catégorie est basée sur la classification utilisant les nuées dynamiques. Pour cela, nous avons développé une méthode de segmentation qui consiste à associer la méthode K-moyens et le formalisme direct-rétrograde. Les tests ont été effectués sur différents types de données. Les résultats sont satisfaisants par rapport aux méthodes basées uniquement sur un schéma simple.

Le dernier axe de notre travail est consacré à l'application sur des données réelles. Cette application est réalisée dans le contexte de l'analyse en imagerie sismique fournie par la société GDF. Les résultats sont nettement meilleurs, aussi bien sur l'estimation que sur la segmentation.

# Abstract

Today, digital telecommunications are booming, thanks to the great advances made in the technologies of electronic integration and of algorithmic. Indeed, the everyday means of communication are about to know great upheavals with the integration of image and sound processing algorithms, allowing a compression of information. Within the context of the scientific actions engaged by the Telecom group of the laboratory of micro-electronics IXL, the study of the consequences of the implementation of the numerical algorithms on the structure of the electronic circuits is a recent axis.

The implementation of the video signal processing is a significant problem in many applications like the video surveillance, compression and video edit. In particular, the temporal management of the objects in the scene makes it possible to design automatic systems of management of information. Within the laboratory, we approach two large shutters. The first relates to the creation of new algorithms dedicated to the modeling and estimation of the primitives, to characterize the moving objects. The second task is to process the reverse problem like the update of the static areas in order to extract the mobile areas.

The work developed in this thesis concerns the motion estimation approaches based on models, in the context of the sequence analysis. More precisely, the goal of this manuscript is the development of motion estimation and segmentation algorithms, which rely on the changes of image brightness. The goal of such a technique is to describe the motion of the areas of the scene without knowledge a priori of the motion of the objects and the camera. For this subject, several approaches have been developed. Each of them has its drawbacks and its disadvantages. The main problem of the motion estimation is the insufficient of constraint for provide a robust estimate. In our case, we will try to develop a formalism which allows to add some constraints in the motion equation. This proposition is called "Motion estimation based on forward and backward scheme". The formalism is carried out in starting from the causal and anti-causal spatio-temporal information (forward and backward motions) of two consecutive frames.

The goal of our techniques is to imagine an abstraction through some assumptions : the motion estimation techniques are based on the comparison of two or more frames. If an object changes intensity along its trajectory and if the brightness associated with this object evolves in manner no modeling, this comparison becomes impossible. Although algorithms estimating the variation of brightness were proposed as general rule, only the proportionate change of brightness is taken into account in the motion estimation techniques. Thereafter, we take as assumption that there is no variation of brightness. So, all variation of intensity is due only to the forward and/or backward displacements of pixels. The mathematical modeling of this assumption allow to define a criterion, to be minimised by an optimisation method. The motion estimation rises from the minimisation of this criterion. Some types of solutions are explored. These approaches are based on the least squares technique which is associated with the forward-backward formalism.

Many experiments were carried out in order to evaluate the quality of the suggested approaches. Real and synthetic sequences were used. The results allow to illustrate the difference between the approach using a simple displacement and that using the forward-backward formalism. Initially, the forward-backward formalism is tested on a translation motion. The second shutter consists of the extension of the suggested formalism on a general model. It is about the parametric model. This model is a description closer to the real case. With the difference of the preceding methods, the parameters of rotation and divergence are

taken into account, and that in the forward and backward senses. Modeling in forward-backward context is more complex than that with simple motion. In this chapter, we describe various methods dedicated to estimation of the motion parameters, according to the forward-backward formalism. The experiments are made on synthetic and real sequences. We focused our tests on the evaluation of the suggested formalism compared to the traditional formalism.

In accordance with the problem of spatio-temporal segmentation, we describe methods using motion estimation based on the forward-backward formalism. The results are satisfactory compared to the methods using only a simple motion. Within this framework, we are interested in the application of the formalism, integrated in the process of spatio-temporal segmentation. Two categories of methods are evoked. The first is the application of the forward-backward formalism to the markovian segmentation. The principle is to associate the markovian process to the estimation based forward-backward flow. The second category is based on classification using the dynamic clouds. For that, we developed a method of segmentation in which, we associate the K-means algorithm and the forward-backward formalism. The tests were carried out on various types of data.

The last axis of our work is dedicated to the application on real data. This application was carried out in the context of the analysis in seismic imagery, provided by the GDF company. The results are satisfactory, as well on the estimation as on the segmentation.

# **Introduction générale**



## 1 But

Ces dernières années ont été le théâtre de nombreux bouleversements technologiques dans le domaine des systèmes dédiés au traitement de l'information. Le secteur des télécommunications est sans nul doute celui qui a connu le plus de mutations. A titre d'illustration, nous sommes passés d'une simple information sonore à une information au contenu plus riche regroupant du texte, du son, de l'image et de la vidéo. La manipulation de l'information multimédia tend à se généraliser. De par l'évolution et les progrès des moyens techniques, les systèmes dédiés à l'information audio ont aujourd'hui atteint une certaine maturité. Pour l'information visuelle, les systèmes sont encore, pour certains, à l'état d'émergence. Ce retard tient surtout à la complexité et au volume de ce type d'information. L'information visuelle est en général constituée d'une séquence d'images numérisées. Chaque image est représentée par un ensemble de pixels. Cette représentation provient des contraintes techniques imposées lors de la capture. Il s'en suit logiquement que les pixels, pris individuellement, n'ont pas de sens direct par rapport à la scène. Certains pixels regroupés peuvent amener à une interprétation tandis que d'autres peuvent entraîner le système de traitement vers une mauvaise analyse (effet d'éclairage, occlusion partielle d'objet ...). Du fait de cette indétermination, le système d'analyse doit se fonder sur des primitives qui ont une signification sémantique. De nombreuses primitives ont été proposées dans la littérature tels que les bords, les textures ou le mouvement. A cet effet, une bonne adéquation entre l'exploitation et la représentation est atteinte lorsque les primitives coïncident avec les objets réels de la scène. Cela permet de décomposer entièrement l'information visuelle en composantes ayant un fort potentiel d'exploitation.

Dans le cadre de l'analyse d'une séquence d'images, les objets peuvent être caractérisés par leurs cohérences spatiale et temporelle. Par conséquent, l'analyse dynamique peut fournir des primitives intéressantes pour la compréhension de la scène. Le mouvement apparent apparaît notamment comme une caractéristique ayant un fort potentiel discriminant. De ce fait, l'estimation de cette primitive est un enjeu crucial pour de nombreuses applications nécessitant la connaissance de l'évolution du contenu de la scène. En outre, cette primitive présente alors un intérêt fondamental pour le développement d'applications fondées sur la poursuite des objets. Le suivi de chaque objet permet d'interagir avec le contenu de la séquence vidéo et permet d'envisager un prolongement dans la gestion des bases de données multimédia. Les techniques de compression vidéo sont elles aussi en attente de moyens performants de segmentation spatio-temporelle. Les dernières générations de codeur MPEG-4 et 7 placent en effet la notion de compensation de mouvement sur le plan de l'objet vidéo. La détection et l'estimation du mouvement sont donc utilisées dans de nombreuses applications ayant pour support l'image numérique comme la télévision numérique, la vidéo surveillance ou dans des secteurs exploratoires comme le biomédical ou la sismique. L'objet de l'étude présentée dans ce document concerne l'évaluation du mouvement apparent des objets de la scène à partir de l'information pixel. Dans cette optique, nous essayons de décrire les mouvements des objets sans connaissance a priori ni de la scène ni du mouvement de la caméra. Une telle problématique débouche forcément sur une volonté de fournir une forme segmentée de l'image en un certain nombre de régions. Si les régions segmentées coïncident avec les objets réels de la scène, comme dans le cas d'objets rigides, les problèmes de discontinuité de mouvements et de l'intensité lumineuse ne se posent pas. La portée sémantique des régions issues de l'analyse du mouvement n'est pas aussi directe pour des objets plus complexes. La description d'un mouvement d'une région n'est alors qu'une étape intermédiaire dans le processus de compréhension de la scène.

Ce mémoire traite du développement d'algorithmes permettant de détecter et d'estimer le mouvement à travers l'évolution spatio-temporelle des objets. Le problème traité se limite à la phase d'estimation du mouvement et de segmentation spatio-temporelle sur une suite limitée



de trames successives, allant de deux à cinq. Pour la phase d'estimation, notre étude se situe dans le cadre des techniques d'estimation du flot optique. L'objectif majeur de notre étude est de permettre la décomposition de chaque trame vidéo en un ensemble d'objets marginalisés par leur mouvement.

## 2 Vue d'ensemble de l'approche retenue

Les thèmes abordés dans ce mémoire concernent successivement l'estimation du mouvement et la segmentation spatio-temporelle. Dans le cadre de l'estimation du mouvement, nous nous plaçons dans un premier temps dans le domaine de l'estimation du flot optique. Conformément à la théorie de l'estimation, le flot optique est calculé de façon optimale en intégrant un support d'observation sur-dimensionné afin de réduire l'effet du bruit associé à l'observation. Pour le mouvement, ce principe est valide en intégrant plus de données spatio-temporelles qu'il ne serait nécessaire si les données n'étaient pas soumises au bruit. De nombreuses techniques issues de ce principe ont été proposées dans la littérature, notamment dans la communauté "vision par ordinateur". Bon nombre d'entre elles traitent de mouvements modélisés par une paramétrisation linéaire débouchant sur une solution au sens des moindres carrés. L'inconvénient majeur d'une intégration globale de tous les points d'observation dans un procédé d'estimation simple est que les points aberrants sont inclus inévitablement dans ce processus. Ces informations erronées conduisent alors le système des moindres carrés vers une mauvaise solution. Les données aberrantes et celles en adéquation avec la modélisation sont en effet intégrées sans discernement. Face à ce problème, des techniques dites robustes ont été très récemment proposées pour fournir une méthodologie de rejet ou de régularisation. Dans ce travail, nous argumentons la nécessité de mettre en œuvre, dans le cas de l'estimation du mouvement, des techniques permettant une régularisation de l'observation. Cette nécessité est inhérente à la problématique du mouvement apparent qui est souvent le résultat d'une projection spatiale 3D-2D conditionnée par une source lumineuse plus ou moins ponctuelle.

Afin de rendre plus efficaces les méthodes d'estimation du mouvement, nous proposons une approche basée sur le mouvement direct et le mouvement rétrograde. L'approche retenue s'appuie sur une approche dérivative utilisant comme base l'équation de contrainte du mouvement. L'objectif de cette combinaison est d'introduire une régularisation permettant d'augmenter les performances des algorithmes d'estimation. De nombreux auteurs ont précisé le fait que l'utilisation d'une moyenne des estimés des gradients spatiaux permet d'améliorer les résultats. Grâce à notre formulation utilisant les mouvements dans les deux sens temporels, nous généralisons cette approche. En outre, nous déclinons différentes procédures permettant d'améliorer les performances. Nous citerons l'extension des méthodes robustes et incrémentales. Le concept utilisé est, en premier lieu, introduit pour une estimation d'un mouvement modélisé par deux paramètres. Nous traitons dans un deuxième temps des modèles de mouvement plus complexes.

Pour ce qui est de la segmentation spatio-temporelle, l'objectif est de décomposer la scène en régions homogènes au sens du mouvement. Les objets que nous étudions, dans la phase de segmentation, sont des objets compacts. Nous associons donc la notion d'objet à un ensemble de pixels connexes. Le mouvement de tous les pixels de chaque objet suit un modèle dont la complexité est choisie a priori. L'étiquetage de chaque pixel est déduit de la description du mouvement en fonction des paramètres du modèle. Les objets peuvent également être caractérisés par une certaine cohérence spatiale, bien que celle-ci soit moins importante que la cohérence temporelle. Ainsi, les objets seront segmentés en tant qu'entités homogènes à l'échelle spatio-temporelle.

L'étiquetage par le flot optique ne mène pas directement à la classification des pixels parce que l'information fournie par chaque pixel n'est pas individuellement suffisante pour

segmenter convenablement l'image. Des stratégies doivent être développées pour classifier conjointement les pixels suivant leur cohérence spatio-temporelle. Les procédés de segmentation proposés sont dérivés de certains algorithmes de classification traditionnels auxquels nous associons l'estimation du mouvement mentionnée.

### 3 Organisation du manuscrit

Le contenu de ce rapport peut être résumé comme suit :

\*La première partie présente une analyse des méthodes proposées dans la littérature. Cette partie a un double objectif ; tout d'abord faire un rappel sur les modèles de mouvement, ensuite, mettre en évidence les caractéristiques des principales techniques liées à l'estimation et à la segmentation du mouvement. Cette partie est elle-même divisée en deux chapitres :

- Le premier chapitre est consacré à un balayage assez large des techniques d'estimation du mouvement. Ainsi, un survol notamment des techniques d'estimation du flot optique est présenté.

- Le deuxième chapitre consiste en un rappel des techniques de segmentation des images numériques. Cette partie permet d'introduire et de justifier à première vue le choix de l'approche de segmentation que nous avons retenue dans ce mémoire.

\*Dans la seconde partie, nous présentons une solution au problème d'estimation du mouvement, qui utilise la méthode différentielle. L'estimation s'effectue de manière statistique de type moindres carrés à partir de l'information locale. Pour évaluer l'intérêt du schéma direct-rétrograde, nous avons illustré les extensions possibles dans le cadre des moindres carrés. La solution se caractérise par la formulation utilisée lors de la mise en équation du mouvement. Cette formulation s'articule autour de la prise en compte à la fois des mouvements direct et rétrograde. La notion clé introduite dans ce chapitre consiste à renforcer les contraintes dans l'équation du mouvement. Le formalisme ainsi adopté s'appelle : approche basée sur un schéma direct-rétrograde. Ce formalisme est utilisé tout au long de ce travail. Nous divisons ici le thème d'étude en deux chapitres :

- Le premier chapitre décrit de manière détaillée l'étude des différentes méthodes d'estimation du mouvement, basées sur le formalisme proposé et cela dans le contexte du mouvement de translation. Pour toutes les méthodes proposées, nous effectuons une étude descriptive et nous procédons ensuite aux expérimentations. Les résultats de simulation sont analysés et comparés avec ceux d'autres méthodes existantes.

- Le deuxième chapitre présente une étude sur l'extension du formalisme proposé au modèle de mouvement plus complexe.

\*La troisième partie traite de la segmentation spatio-temporelle. Cette étude est motivée par la nécessité d'employer, d'une façon optimale, le formalisme basé mouvement direct-rétrograde pour réduire au minimum l'effet des informations erronées. Il s'agit d'une procédure permettant de résoudre le problème d'initialisation, ainsi que le problème d'affectation, dans le cadre de la classification. Pour pouvoir situer l'apport du schéma direct-rétrograde dans la segmentation, nous avons sélectionné deux approches de classification différentes, ayant chacune sa spécificité :

- La première approche, présentée au premier chapitre, est plus proche de l'aspect pixel dans lequel un ensemble de cliques est considéré ; il s'agit du modèle markovien.

- La deuxième approche s'articule plutôt autour de l'aspect région ; il s'agit de la classification utilisant le principe des nuées dynamiques. Cette approche est traitée dans le second chapitre. Nous utilisons ici l'algorithme K-moyens.

Dans chacun de ces cadres, l'estimation du mouvement basée sur un schéma direct-rétrograde est introduite et c'est l'idée que nous allons développer.

\*La quatrième partie propose une application concrète dans le cadre des données sismiques. Il s'agit d'une estimation du temps de retard de pénétration du gaz naturel. Ce projet est mené en liaison avec la Société Gaz de France. L'objectif est d'évaluer la présence et la densité de gaz existant dans un réservoir naturel. Deux chapitres seront ainsi abordés.

- Le premier chapitre est consacré aux méthodes d'estimation du retard. Dans un premier temps, une approche regroupant les méthodes d'intercorrélation généralisée est présentée. Cette étude fait l'objet d'une analyse des méthodes utilisées par la société GDF. Enfin, une approche basée sur l'équation différentielle du mouvement est présentée. Nous faisons intervenir ici le formalisme basé mouvement direct-rétrograde.

- Le second chapitre traite de la segmentation dans le cadre des images sismiques. L'approche proposée s'inspire de la méthode des nuées dynamiques

\*Enfin, nous concluons cette étude par une argumentation quant aux principales contributions qu'elle apporte, par une discussion des problématiques non résolues, ainsi que par les améliorations et extensions suggérées.

**Première partie**

**État de l'art et contributions**



# Chapitre 1

## Estimation du mouvement

### 1.1 Introduction

Une primitive importante dans l'analyse d'une séquence vidéo est caractérisée par l'estimation du mouvement des objets contenus dans une scène. Cette phase d'estimation est dépendante de la modélisation sélectionnée pour rendre compte de l'évolution spatiale des objets. L'obtention de cette primitive s'appuie sur l'observation de plusieurs trames consécutives. Cette observation contient en effet des objets communs ayant changé de position relative. Ce principe peut être engendré par deux configurations caractérisées par des objets qui bougent devant la caméra ou par un mouvement de caméra. Ces deux configurations peuvent se combiner, l'évolution temporelle engendrée est appelée de façon générique : mouvement.

En revanche, si l'information de mouvement est présente parmi les informations d'une image donnée, elle ne l'est que de manière sous-jacente au travers des variations temporelles de la luminance (mouvement apparent 2D ou flot optique).

En terme d'intérêt, vu la richesse de la caractéristique du mouvement comme source d'information, l'estimation du mouvement est essentielle pour de nombreuses applications. Elle est la primitive utilisée pour contrôler les phases de poursuite et de reconnaissance dans le cas d'un système de vidéo surveillance. Les techniques de codage vidéo l'utilisent pour augmenter significativement leur performance de compression. Les codeurs de première génération comme les séries MPEG et H-26x s'appuient sur cette primitive en terme d'appariement de blocs. Ceux de la seconde génération comme MPEG-7 intègrent une compréhension sémantique de la scène avec notamment comme descripteur, le mouvement.

Le présent chapitre est consacré à la présentation de l'existant en terme d'estimation. Nous supposons que l'image est segmentée en régions représentant le support de l'estimation. La section 1.2 rappelle les concepts de base associés à l'observation du mouvement. Dans la section 1.3 sont illustrés les différents modèles du mouvement allant de la simple translation à des modèles plus complexes. Les sections 1.4, 1.5, 1.6 et 1.7 présentent les différentes techniques d'estimation du mouvement et les problèmes qui leurs sont associés. Enfin, une conclusion sur ce chapitre est dressée dans la section 1.9.

### 1.2 Quelques notions fondamentales en estimation du mouvement

#### 1.2.1 Séquence numérique.

Dans une séquence vidéo numérique, deux trames (images) successives présentent de nombreuses similarités caractérisant le fond et les objets en mouvement. Ces trames sont composées d'un même ensemble d'objets. Une trame est une représentation planaire d'une scène ou d'un objet, généralement situé dans un espace tridimensionnel. Une image numérique est constituée d'un ensemble de points élémentaires appelés pixels. Le couple de réels  $(x, y)$  représente la position spatiale d'un pixel, et la valeur  $I(x, y)$  son niveau de gris. Le niveau de gris d'un pixel est une grandeur proportionnelle à l'intensité du signal réfléchi par ce pixel lorsqu'il est radié notamment par une onde électromagnétique. La structure lumineuse d'une image est constituée par un ensemble de motifs brillants.

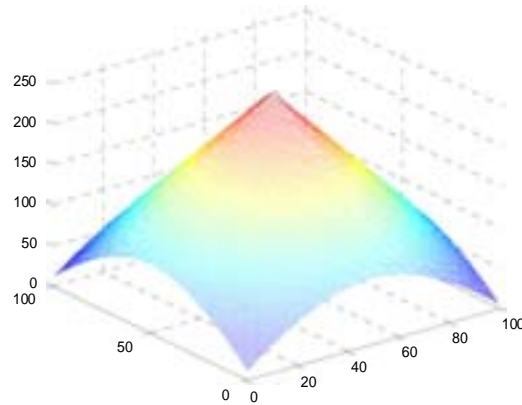


Figure 1.1 : Structure lumineuse d'une image.

La figure 1.1 montre une représentation en intensité de la structure lumineuse d'une image. La luminance d'un pixel, aussi appelée niveau de gris, est une densité d'énergie ( $W/m^2$ ).

Une image numérique est à la base un tableau bidimensionnel, dont les valeurs représentent le niveau de gris des pixels se trouvant aux coordonnées correspondantes.

### 1.2.2 Mouvement global, mouvement local

Un mouvement global est un mouvement tel que tous les pixels de l'image se déplacent d'une manière homogène. Le mouvement local décrit individuellement le mouvement d'un ensemble de points de l'image. En général, ces deux spécifications ne restent qu'une définition, car du point de vue structurel, aucune distinction n'est faite entre mouvement global et local. Pour mieux comprendre ces définitions, plaçons-nous par exemple dans le contexte de l'analyse d'une image toute entière. Imaginons que nous observons une séquence vidéo en un point arbitraire. Le point est le plus petit élément que nous puissions observer sur un élément d'image. Dans le temps, l'intensité des pixels est susceptible de varier. Trois causes peuvent être la source d'une telle variation. La première est due au mouvement global ou mouvement de la caméra. Même si aucun mouvement n'apparaît dans cette scène, le mouvement de la caméra produit un mouvement global de la scène capturée. La seconde cause est le mouvement intrinsèque des objets dans la scène. Celle-ci apparaît comme un mouvement local car elle n'affecte pas l'image entière. Enfin, la troisième raison est la variation de la luminance. Si la condition d'éclairage change pendant que la séquence est capturée par la caméra, les intensités du pixel varient. Nous pouvons donc dire que le mouvement local résume l'aspect de chaque objet de l'image et le mouvement global est affecté à l'image toute entière. Ce qui n'est pas toujours valable car dans un contexte d'analyse d'objet (contenu dans une scène), le mouvement global correspond à la transformation subie par l'objet lui-même (déformations globales de la forme) et le mouvement local correspond à l'information pertinente (déformations résiduelles de la forme) [ODO98].

### 1.2.3 Champ bidimensionnel de déplacement, flot optique, mouvement apparent 2D

Un mouvement 3D correspond au déplacement entre deux instants d'un objet dans un espace 3D. On appelle champ de déplacement, l'ensemble des vecteurs associés à chaque point de l'objet, décrivant le mouvement des pixels entre deux instants successifs. Comme le montre la figure 1.2, le champ bidimensionnel de déplacement est obtenu par la projection sur un plan image des déplacements relatifs entre l'objet et l'observateur.

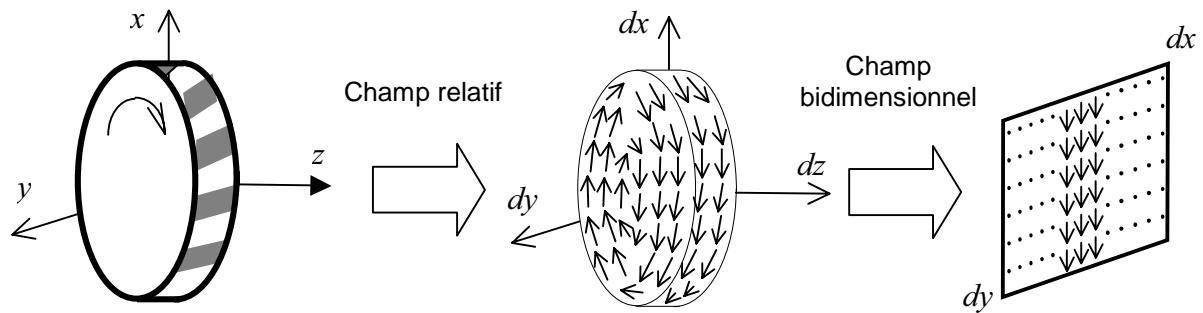


Figure 1.2 : Illustration du champ de déplacement bidimensionnel.

Dans une séquence vidéo numérique, une scène 3D est ramenée à une image issue d'une projection sur un repère 2D. Dès lors, le champ de déplacement résultant de la variation spatio-temporelle de la luminance d'une image 2D n'est qu'apparent grâce aux effets qualifiés d'illusion d'optique. Reprenons l'exemple d'une roue qui tourne autour de son axe (figure 1.3). Sur le plan image  $x$ - $y$ , son mouvement réel de rotation est perçu comme un mouvement de translation des bandes ornant cette roue. Il s'agit d'un mouvement apparent (perçu) 2D qui s'appelle aussi flot optique.

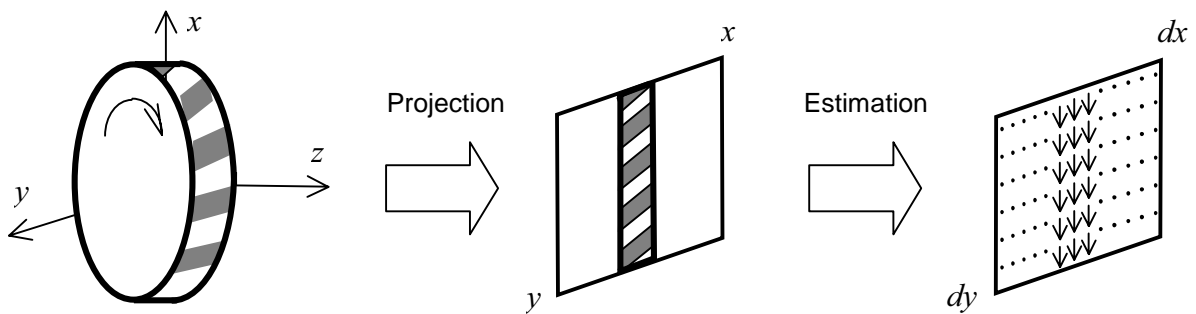


Figure 1.3 : Illustration du flot optique (mouvement apparent 2D).

Le flot optique (mouvement apparent 2D) et le champ bidimensionnel de déplacement sont deux aspects différents, mais dans le domaine du traitement numérique des images, ils ne sont pas distingués. Nous signalons ici que notre étude concerne le mouvement apparent 2D et l'expression mouvement devrait être entendue comme flot optique.

### 1.3 Modèle de mouvement

En général, l'estimation du mouvement nécessite la définition des deux points suivants :

- Mettre en place un modèle de mouvement, c'est une opération qui consiste à décrire le mouvement par un ensemble de paramètres reliés aux paramètres de la caméra.
- Choisir la méthode pour estimer les paramètres du modèle de mouvement utilisé.

Nous allons décrire les différents types de modèles de mouvement 2D applicables au sein de l'estimation et de la segmentation spatio-temporelle. Un modèle de mouvement englobe certaines caractéristiques de telle sorte que le mouvement d'un voisinage puisse être représenté. Ainsi, le modèle régit le lien entre les mouvements de chaque point d'une région, en fonction de leur position.



### 1.3.1 Description du mouvement 3D et sa projection

Observons le mouvement d'un objet rigide dans un espace. Tous les points doivent s'y déplacer entièrement d'une même manière par rapport à un repère d'observation. En effet, le mouvement 3D de chaque point entre deux instants  $t$  et  $t + dt$  peut être représenté par une même transformation à partir de l'instant  $t$ . D'après la figure 1.4, l'utilisation d'un tenseur cinématique permet de décomposer cette transformation en six mouvements élémentaires : trois translations et trois rotations.

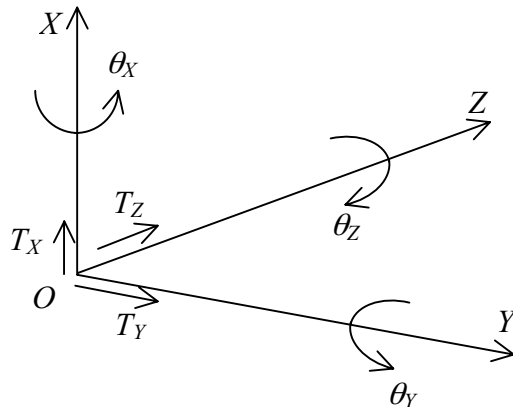


Figure 1.4 : Illustration du mouvement à six degrés de liberté d'un objet rigide dans un espace tridimensionnel.

La projection d'une structure 3D sur le plan image (figure 1.5) peut se formuler de deux manières : en terme de déplacement ou en terme de vitesse.

La transformation affine qui définit la relation entre les coordonnées des points aux deux instants successifs est définie par [BAL82] :

$$P' = \vec{T} + RP \quad (1)$$

où  $P(X, Y, Z)$  et  $P'(X', Y', Z')$  sont les deux positions respectives aux instants  $t$  et  $t + dt$  par rapport au centre de rotation  $O(X_o, Y_o, Z_o)$  et  $\vec{T} = (T_x, T_y, T_z)^T$  le vecteur de translation. En général, les angles de rotation sont supposés faibles. Selon cette hypothèse (cf. annexe A), la matrice de rotation  $R$  est approximativement :

$$R = \begin{pmatrix} 1 & -\theta_z & \theta_y \\ \theta_z & 1 & -\theta_x \\ -\theta_y & \theta_x & 1 \end{pmatrix} \quad (2)$$

Soit :

$$\begin{pmatrix} X' - X_o \\ Y' - Y_o \\ Z' - Z_o \end{pmatrix} = \begin{pmatrix} T_x \\ T_y \\ T_z \end{pmatrix} + \begin{pmatrix} 1 & -\theta_z & \theta_y \\ \theta_z & 1 & -\theta_x \\ -\theta_y & \theta_x & 1 \end{pmatrix} \begin{pmatrix} X - X_o \\ Y - Y_o \\ Z - Z_o \end{pmatrix} \quad (3)$$

Une théorie sur la géométrie projective est présentée en annexe A. En introduisant cette théorie dans l'équation (3) et en prenant la valeur de la distance focale égale à l'unité de mesure, nous avons :

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \frac{x - \theta_z y + \theta_y + T_x / Z}{1 - \theta_y x + \theta_x y + T_z / Z} \\ \frac{\theta_z x + y - \theta_x + T_y / Z}{1 - \theta_y x + \theta_x y + T_z / Z} \end{pmatrix} \quad (4)$$

où  $(x, y)$  et  $(x', y')$  sont respectivement les coordonnées de  $p$  et  $p'$ , projection respective de  $P(X, Y, Z)$  et  $P'(X', Y', Z')$  sur le plan image. Nous obtenons ainsi l'expression qui définit le déplacement d'une image plane par transformation d'une structure 3D :

$$\begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} x' - x \\ y' - y \end{pmatrix} = \begin{pmatrix} \frac{x - \theta_z y + \theta_y + T_x / Z}{1 - \theta_y x + \theta_x y + T_z / Z} \\ \frac{\theta_z x + y - \theta_x + T_y / Z}{1 - \theta_y x + \theta_x y + T_z / Z} \end{pmatrix} - \begin{pmatrix} x \\ y \end{pmatrix} \quad (5)$$

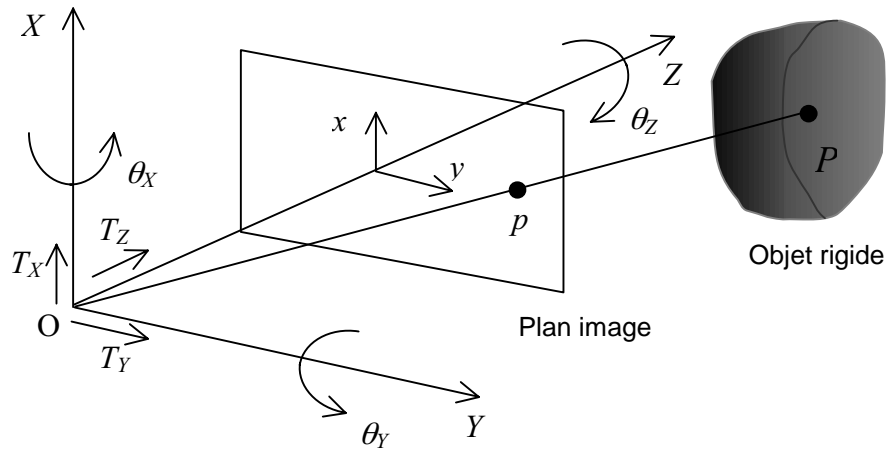


Figure 1.5 : Géométrie projective et objet rigide.

En terme de vitesse instantanée, l'expression de la transformation 3D s'obtient par dérivation de l'expression du déplacement entre deux points successifs ( $P$  et  $P'$ ) par rapport au temps. Dans un espace 3D, le déplacement est exprimé par :

$$\begin{pmatrix} X' - X \\ Y' - Y \\ Z' - Z \end{pmatrix} = \begin{pmatrix} T_x \\ T_y \\ T_z \end{pmatrix} + \begin{pmatrix} 1 & -\theta_z & \theta_y \\ \theta_z & 1 & -\theta_x \\ -\theta_y & \theta_x & 1 \end{pmatrix} \begin{pmatrix} X - X_o \\ Y - Y_o \\ Z - Z_o \end{pmatrix} - \begin{pmatrix} X - X_o \\ Y - Y_o \\ Z - Z_o \end{pmatrix} \quad (6)$$

En introduisant le terme  $\omega = \frac{d\theta}{dt}$ , la transformation 3D d'un objet rigide en terme de vitesse est donnée par la relation :

$$\begin{pmatrix} \frac{dX}{dt} \\ \frac{dY}{dt} \\ \frac{dZ}{dt} \end{pmatrix} = \begin{pmatrix} V_x \\ V_y \\ V_z \end{pmatrix} + \begin{pmatrix} 1 & -\omega_z & \omega_y \\ \omega_z & 1 & -\omega_x \\ -\omega_y & \omega_x & 1 \end{pmatrix} \begin{pmatrix} X - X_o \\ Y - Y_o \\ Z - Z_o \end{pmatrix} - \begin{pmatrix} X - X_o \\ Y - Y_o \\ Z - Z_o \end{pmatrix} \quad (7)$$

Cette transformation peut s'écrire sous la forme :

$$\dot{P} = \vec{V} + (\dot{R} - I)P \quad (8)$$

où  $\dot{P} = (\partial X / \partial t, \partial Y / \partial t, \partial Z / \partial t)^T$  représente la vitesse instantanée du point  $P$ . Ce point est référencé par rapport à l'origine  $O$ . La quantité  $\vec{V} = (V_x, V_y, V_z)^T$  désigne le vecteur de translation (la vitesse linéaire). La matrice  $\dot{R}$  est analogue à la matrice de transformation (2), sauf qu'ici, elle est exprimée en fonction des composantes  $\omega$  de la vitesse angulaire. Le terme  $I$  correspond à la matrice identité d'ordre 3.

De l'équation (7) et de la théorie de la géométrie projective (équations A5 et A6 avec une distance focale égale à une unité de mesure), nous obtenons l'expression du vecteur vitesse suivante :

$$\begin{pmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{pmatrix} = \begin{pmatrix} \frac{x - \theta_z y + \theta_y + T_x / Z}{1 - \theta_y x + \theta_x y + T_z / Z} \\ \frac{\theta_z x + y - \theta_x + T_y / Z}{1 - \theta_y x + \theta_x y + T_z / Z} \end{pmatrix} - \begin{pmatrix} x \\ y \end{pmatrix} \quad (9)$$

où  $(x, y)$  est référencé par rapport à  $(x_o, y_o)$ . Les équations (5) et (9) sont équivalentes. La première est exprimée en terme de déplacement tandis que la seconde l'est en terme de vitesse. L'équation (9) peut se présenter sous la forme :

$$\begin{pmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{pmatrix} = \begin{pmatrix} \frac{-\omega_x xy + \omega_y (x^2 + 1) - \omega_z y + \frac{V_x - V_z x}{Z}}{1 - \omega_y x + \omega_x y + \frac{V_z}{Z}} \\ \frac{-\omega_x (y^2 + 1) + \omega_y xy + \omega_z x + \frac{V_y - V_z y}{Z}}{1 - \omega_y x + \omega_x y + \frac{V_z}{Z}} \end{pmatrix} \quad (10)$$

Les équations (9) et (10) définissent le lien entre une structure du mouvement 3D et le vecteur vitesse.

### 1.3.2 Modèles de mouvement 2D

Nous avons vu précédemment que le mouvement 2D est issu de la transformation projective du mouvement d'un objet 3D sur une image plane. Les modèles 2D les plus généraux sont obtenus à partir des modèles de mouvement 3D, modèles de surface 3D et modèles de transformation projective de la caméra.

En général, dans le codage vidéo, les angles de rotation et la variation relative de profondeur entre deux trames consécutives sont faibles. Par la suite, pour décrire un modèle de mouvement 2D, nous supposons que ces deux conditions sont satisfaites.

D'après les équations (5) et (9), la description 2D d'un objet rigide s'avère difficile à cause de la dépendance en profondeur ( $Z$ ). En effet, l'estimation exacte de la profondeur est une opération délicate d'autant plus que cette grandeur n'est pas nécessairement utile lors d'un codage vidéo. Pour faire ressortir les modèles de mouvement à partir des équations précédentes, nous allons modéliser les termes de profondeur en fonction des coordonnées  $x, y$ . Admettons que chaque objet 3D de la scène puisse se décomposer en un ensemble de surfaces. Plusieurs types de modélisation de surface sont envisageables : nous n'allons tenir compte que des trois équations suivantes :

- Le premier modèle ne prend pas en compte les termes du second ordre. Il s'agit d'une facette définie par :

$$\alpha X + \beta Y + \gamma Z = 1 \quad (11)$$

ou encore :

$$\alpha x + \beta y + \gamma = \frac{1}{Z} \quad (12)$$

- Le second modèle consiste à considérer une surface parabolique définie par l'équation :

$$\alpha X + \beta Y + \gamma Z + \varepsilon (0^2) = 1 \quad (13)$$

- Dans le troisième, il s'agit de l'utilisation d'une transformation linéaire associée à une annulation des composantes hautes fréquences. Nous allons considérer séparément ces trois cas.

**1<sup>er</sup> cas** : Pour simplifier, nous commençons d'abord par le cas où chaque objet est décomposable en un ensemble de facettes, suivant l'équation (12). En substituant cette relation dans l'équation (5), nous avons l'expression du déplacement d'une surface plane :

$$\begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} \frac{(\theta_Y + \gamma T_X) + (1 + \alpha T_X)x + (\beta T_X - \theta_Z)y}{(1 + \gamma T_Z) + (\alpha T_Z - \theta_Y)x + (\theta_X + \beta T_Z)y} \\ \frac{(\gamma T_Y - \theta_X) + (\theta_Z + \alpha T_Y)x + y + (1 + \beta T_Y)T_Y}{(1 + \gamma T_Z) + (\alpha T_Z - \theta_Y)x + (\theta_X + \beta T_Z)y} \end{pmatrix} - \begin{pmatrix} x \\ y \end{pmatrix} \quad (14)$$

L'équation (14) a une forme générale :

$$\begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} \frac{a_1 + a_2x + a_3y}{1 + a_4x + b_4y} \\ \frac{b_1 + b_2x + b_3y}{1 + a_4x + b_4y} \end{pmatrix} - \begin{pmatrix} x \\ y \end{pmatrix} \quad (15)$$

avec :

$$\begin{aligned} a_1 &= \frac{\theta_Y + cT_X}{1 + cT_Z}, a_2 = \frac{1 + aT_X}{1 + cT_Z}, a_3 = \frac{bT_X - \theta_Z}{1 + cT_Z}, b_1 = \frac{cT_Y - \theta_X}{1 + cT_Z}, \\ b_2 &= \frac{\theta_Z + aT_Y}{1 + cT_Z}, b_3 = \frac{1 + bT_Y}{1 + cT_Z}, a_4 = \frac{aT_Z - \theta_Y}{1 + cT_Z}, b_4 = \frac{\theta_X + bT_Z}{1 + cT_Z} \end{aligned} \quad (16)$$

L'équation ((15) est un descripteur général d'un champ de déplacement dans le plan image, dérivé d'un champ de déplacement d'une facette sous la projection perspective d'un objet 3D. C'est un modèle à 8 paramètres dont les coefficients sont déterminés par projection perspective.

Une formulation analogue en terme de vitesse sera décrite par la suite. En général, les objets sont éloignés de la caméra et par conséquent, la variation relative de la profondeur  $Z$  est faible. C'est à dire  $\Delta Z/Z \ll 1$  avec  $\Delta Z = Z' - Z$ . La relation (A5) nous permet d'écrire :

$$x' = \frac{X'}{Z'} = \frac{X + \Delta X}{Z + \Delta Z} \quad (17)$$

ou encore :

$$x' \approx \frac{X + \Delta X}{Z} \left(1 - \frac{\Delta Z}{Z}\right) \approx \frac{X'}{Z} - \frac{X}{Z} \frac{\Delta Z}{Z} + O(2) \quad (18)$$

En négligeant les termes du second ordre, nous avons :

$$x' \approx \frac{X'}{Z} - \frac{X}{Z} \frac{\Delta Z}{Z} \quad (19)$$

L'expression de  $X'$  et  $\Delta Z$  dans (6) reportée dans (19) donne :

$$x' = x + \left[ -\theta_X xy + \theta_Y (x^2 + 1) - \theta_Z y \right] + \frac{T_X - T_Z x}{Z} \quad (20)$$

De même pour  $y'$ , nous avons :

$$y' = y + \left[ -\theta_Y (y^2 + 1) + \theta_X xy + \theta_Z x \right] + \frac{T_Y - T_Z y}{Z} \quad (21)$$

Il est à rappeler que lorsque nous raisonnons en terme de déplacement, les notations  $T$  et  $\theta$  sont utilisées pour désigner les composantes de la translation et la rotation. Quand il s'agit du vecteur vitesse, nous utilisons les notations  $V$  et  $\omega$ . Les expressions (20) et (21) permettent de

simplifier l'équation (5) (en terme de déplacement) ainsi que l'équation (10) (en terme de vitesse). L'équation (10) se simplifie alors de la manière suivante :

$$\begin{pmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{pmatrix} = \begin{pmatrix} -xy\omega_x + (1+x^2)\omega_y - y\omega_z + \frac{V_x - V_z x}{Z} \\ -(1+y^2)\omega_x + \omega_y xy + \omega_z x + \frac{V_y - V_z y}{Z} \end{pmatrix} \quad (22)$$

En substituant la relation (12) dans la relation (22), la vitesse d'une surface plane peut être décrite par :

$$\begin{pmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{pmatrix} = \begin{pmatrix} (\omega_y + \gamma V_x) + (\alpha V_x - \gamma V_z)x + (\beta V_x - \omega_z)y + (-\omega_x - \beta V_z)xy + (\omega_y - \alpha V_z)x^2 \\ (-\omega_x + \gamma V_y) + (\alpha V_y + \omega_z)x + (\beta V_y - \gamma V_z)y + (\omega_y - \alpha V_z)xy + (-\omega_x - \beta V_z)y^2 \end{pmatrix} \quad (23)$$

L'équation (23) peut s'écrire sous une forme générale :

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} a_1 + a_2 x + a_3 y + a_4 xy + b_4 x^2 \\ b_1 + b_2 x + b_3 y + b_4 xy + a_4 y^2 \end{pmatrix} \quad (24)$$

avec :

$$\begin{cases} a_1 = \omega_y + \gamma V_x \\ a_2 = \alpha V_x - \gamma V_z \\ a_3 = \beta V_x - \omega_z \\ b_1 = -\omega_x + \gamma V_y \\ b_2 = \alpha V_y + \omega_z \\ b_3 = \beta V_y - \gamma V_z \\ a_4 = -\omega_x - \beta V_z \\ b_4 = \omega_y - \alpha V_z \\ u = \partial x / \partial t \\ v = \partial y / \partial t \end{cases} \quad (25)$$

Par la suite, nous gardons les notations  $u$  et  $v$  pour les composantes du vecteur vitesse dans un espace 2D. L'équation (24) est un descripteur général d'un vecteur vitesse dans le plan image. Ce modèle est également à 8 paramètres. Il permet de décrire le mouvement de translation, rotation et déformation des objets plans et rigides.

Pour une transformation par projection orthogonale, nous avons  $x = X/Z \ll 1$ ,  $y = Y/Z \ll 1$  et les termes du second ordre sont ainsi négligeables. Le descripteur simplifié ainsi obtenu s'appelle modèle affine, composé uniquement de 6 paramètres dont les termes sont linéaires et constants. Ce modèle est une version simplifiée du modèle précédent.

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} a_1 + a_2 x + a_3 y \\ b_1 + b_2 x + b_3 y \end{pmatrix} \quad (26)$$

Le modèle affine peut être encore simplifié pour obtenir un modèle de translation à 2 paramètres. Ce descripteur n'utilise que les termes constants. Tous les termes linéaires en  $x$  et en  $y$  sont nuls.

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \quad (27)$$

**2<sup>ème</sup> cas** : Prenons maintenant le modèle de surface parabolique, suivant l'équation (13). Chaque objet est alors décomposable en un ensemble de surfaces paraboliques. Cette considération permet de couvrir une plus large classe d'objets autres que les objets plans. Les termes du second ordre de l'équation (13) ne sont plus négligés. Sous une projection orthogonale ( $x = X/Z \ll 1$  et  $y = Y/Z \ll 1$ ), un modèle plus complexe à 12 paramètres sera obtenu [DIE91] :

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} a_1 + a_2x + a_3y + a_4x^2 + a_5y^2 + a_6xy \\ b_1 + b_2x + b_3y + b_4x^2 + b_5y^2 + b_6xy \end{pmatrix} \quad (28)$$

Ce modèle s'appelle modèle quadratique, utilisé pour décrire la structure et le mouvement d'une surface parabolique. En revanche, il ne permet pas de décrire les surfaces très prononcées comme celles d'un cube.

**3<sup>ème</sup> cas** : D'après la structure lumineuse d'une image, les champs de déplacement sont fortement corrélés suivant l'échelle spatiale. Par conséquent, il est probable que de tels champs puissent être efficacement représentés en utilisant une transformation linéaire associée à une annulation des composantes hautes fréquences. A titre d'exemple, une transformation polynomiale [STI92, STI94, STI99, SAN94, KAR97] décrite au tableau 1.1 généralise la plupart des modèles paramétriques linéaires et quadratiques. Pour  $K = \{(0,0), (0,1), (1,0)\}$ , la description polynomiale est ramenée à un modèle affine. De même, le modèle quadratique est obtenu pour  $K = \{(0,0), (0,1), (1,0), (0,2), (1,1), (2,0)\}$ . Ce modèle généralisé permet de décrire des mouvements complexes quand le nombre de paramètres est suffisamment adapté à cette complexité. L'expression de la transformation polynomiale est donnée par :

$$\begin{pmatrix} u \\ v \end{pmatrix} = \sum_{i=0}^n \sum_{j=0}^i \begin{pmatrix} a_{ij} \\ b_{ij} \end{pmatrix} x^{i-j} y^j \quad (29)$$

où  $a_{ij}$  et  $b_{ij}$  sont les paramètres du mouvement et  $n$  est l'ordre de la transformation.

Il est possible d'envisager de très nombreux modèles de mouvement. Le tableau 1.1 présente quelques modèles usuels [STI99, NIC92].

### 1.3.3 Choix d'un modèle approprié

Les modèles de mouvement peuvent être classés en trois catégories [ANA93] : modèles non paramétriques, modèles quasi paramétriques et modèles totalement paramétriques.

Les modèles non paramétriques reposent sur le champ dense du mouvement local. La plupart des méthodes d'estimation dont nous allons parler dans cette étude utilisent la catégorie des modèles non paramétriques. Un vecteur déplacement 2D (naturellement translationnel) est attribué à chaque pixel.

Les modèles paramétriques décrivent le mouvement d'une région à l'aide d'un ensemble de paramètres. Cette catégorie s'appelle méthodes basées régions ou paramétriques.

Les modèles quasi paramétriques spécifient le mouvement aussi bien par un ensemble de paramètres que par un champ local, tandis que les modèles totalement paramétriques décrivent complètement le mouvement par un ensemble de paramètres. Pour mieux comprendre ces deux aspects, prenons l'exemple des modèles détaillés dans le paragraphe 1.3.2. Pour établir les équations des modèles de mouvement 2D, nous avons remarqué que, dans un premier temps, nous sommes partis de l'équation de transformation 3D-2D. Dans la mesure où chaque pixel ou chaque voisinage local est considéré indépendamment, nous pouvons utiliser ces équations pour décrire le mouvement d'un pixel ou d'un voisinage local. Nous avons eu jusqu'à ce niveau des modèles quasi paramétriques. Après avoir établi les équations de transformation 3D-2D, nous avons utilisé un modèle de surface pour décrire une

scène 3D, pour finalement avoir des modèles paramétriques complets (affine, linéaire, quadratique, polynomial). Les modèles totalement paramétriques sont donnés sur les quatre dernières lignes du tableau 1.1.

Modèle 2D		Modèle 3D		Modèle de projection (caméra)
Modèle	Équation du modèle	Modèle de surface	Mouvement 3D	
Translation	$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}$	Arbitraire	Translation 3D rigide	Orthogonale
Affine	$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} + \begin{pmatrix} a_2 & a_3 \\ b_2 & b_3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$	Facette	Affine 3D	Orthogonale
Linéaire projectif	$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} a_1 + a_2x + a_3y + a_4xy + b_4x^2 \\ b_1 + b_2x + b_3y + b_4xy + a_4y^2 \end{pmatrix}$ ou $\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} a_1 + a_2x + a_3y \\ 1 + a_4x + b_4y \\ b_1 + b_2x + b_3y \\ 1 + a_4x + b_4y \end{pmatrix} - \begin{pmatrix} x \\ y \end{pmatrix}$	Facette	Affine 3D	Perspective
Quadratique	$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} a_1 + a_2x + a_3y + a_4x^2 + a_5y^2 + a_6xy \\ b_1 + b_2x + b_3y + b_4x^2 + b_5y^2 + b_6xy \end{pmatrix}$	Parabolique	Affine 3D	Orthogonale
Polynomial	$\begin{pmatrix} u \\ v \end{pmatrix} = \sum_{i=0}^n \sum_{j=0}^i \begin{pmatrix} a_{ij} \\ b_{ij} \end{pmatrix} x^{i-j} y^j$	Surface lissée par une transformation linéaire (K)		Arbitraire

Tableau 1.1 : Modèles de mouvement 2D et leur transformation associée.

## 1.4 Techniques d'estimation du mouvement : rappel

Une large variété de méthodes a été proposée par différents auteurs pour les techniques d'estimation du mouvement [TEK95, STI99, BAR92, AGG88]. Avant de présenter les différentes méthodes d'estimation du mouvement, nous allons considérer une métaphore à partir de l'image en perspective présentée en figure 1.6. Si on la regarde de face, deux points de vue sont possibles :

- On peut penser que c'est un cône rectangulaire qui se rétrécit.
- En revanche, en partant de la surface intérieure, on peut se dire que c'est un tube conique dont la base s'élargit.

De même, dans de nombreuses références, chaque auteur a sa façon de classer les différentes méthodes d'estimation du mouvement. Comme indiqué sur la figure 1.7, chaque observateur peut avoir des modes de classement différents, selon la façon dont il observe le phénomène. Il n'y a pas de règle générale qui définit la hiérarchie d'un classement. Un classement peut être défini de la manière suivante : on part d'une rubrique donnée (par exemple du haut sur la figure 1.7) et les autres rubriques pourront devenir des sous-classes de cette rubrique principale. Cela correspond à l'exemple de la figure 1.8.

Mais certains auteurs peuvent avoir un autre raisonnement en procédant en sens inverse, c'est-à-dire en définissant comme catégorie principale l'une des sous-rubriques précédentes et en plaçant les autres rubriques au niveau inférieur. C'est le cas de l'exemple de la figure 1.9.

En outre, sans tenir compte d'une hiérarchie dans le classement, nous verrons par la suite que certaines méthodes sont liées entre elles.

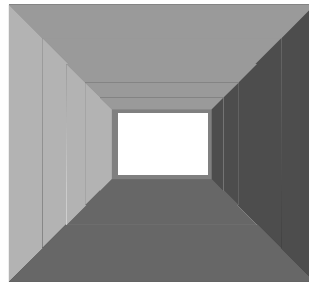


Figure 1.6 : Illustration des différentes manières d'observer un phénomène.

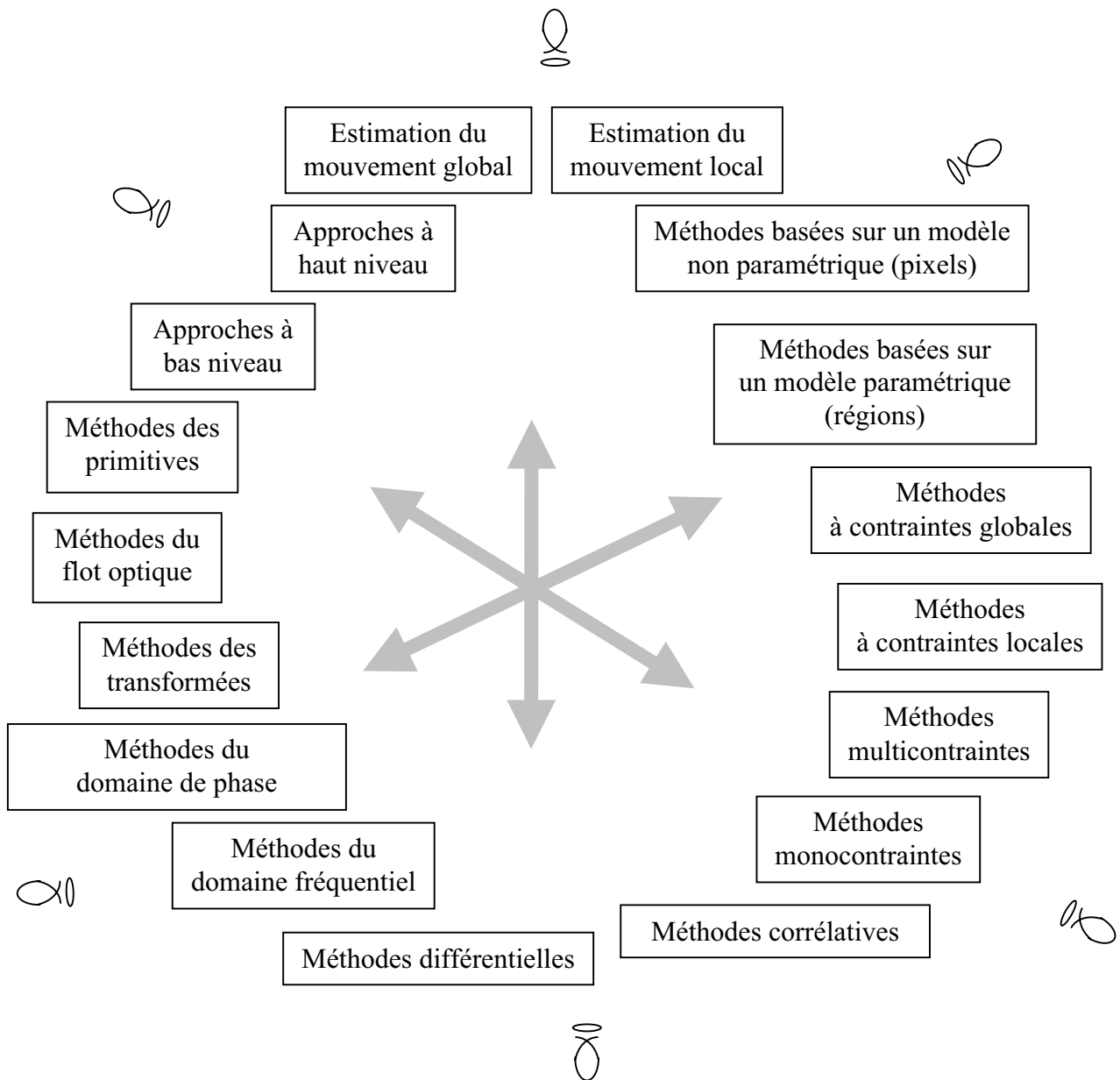


Figure 1.7 : Différentes manières de former un classement des méthodes d'estimation du mouvement.



Dans ce mémoire, nous n'allons pas préciser comment classer les différentes méthodes d'estimation du mouvement, mais nous allons plutôt rappeler quelques techniques d'estimation du mouvement et éventuellement, la catégorie dans laquelle chaque méthode peut se situer. A titre d'exemple, nous pouvons dire que les méthodes d'estimation du mouvement peuvent se classer en deux catégories (figure 1.8) :

- Approche de bas niveau dans laquelle l'estimation du mouvement s'effectue sur tous les pixels de l'image : méthodes des transformées globales, méthodes différentielles, corrélation.
- Approche de haut niveau dans laquelle l'estimation ne s'effectue que sur les primitives extraites de l'image (régions, contours, ...) : méthodes différentielles, méthodes d'appariement de régions ou contours (corrélation), ...

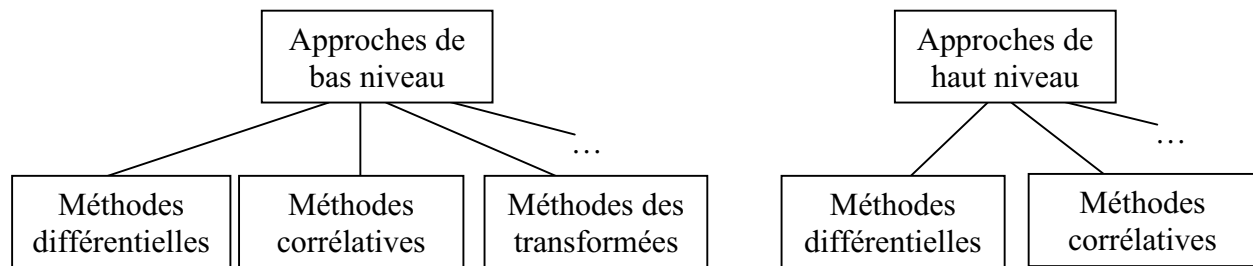


Figure 1.8 : Exemple de classement des différentes méthodes d'estimation du mouvement.

Mais une autre formulation est possible, en disant que les méthodes d'estimation du mouvement peuvent être classées en trois catégories (figure 1.9) :

- Méthodes différentielles : approches basées pixels (de bas niveau), approches basées primitives (haut niveau), ...
- Méthodes de transformée.
- Méthodes de corrélation .

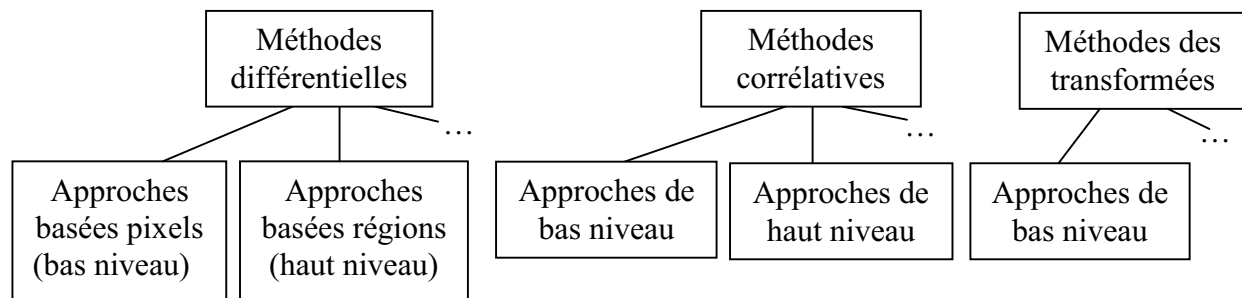


Figure 1.9 : Un autre exemple de hiérarchie pour classer les différentes méthodes d'estimation du mouvement.

Sur ces deux illustrations, nous pouvons remarquer qu'une méthode différentielle peut être à la fois de bas niveau et de haut niveau. Inversement, une estimation de bas niveau peut être différentielle ou basée sur une autre approche.

De même, nous verrons par la suite que dans les méthodes multicontraîntes, on peut classer des méthodes fréquentielles, méthodes différentielles, etc., et inversement, les méthodes différentielles peuvent être classées en multicontraîntes (plusieurs équations de contrainte) et en monocontraîntes (avec ajout d'un terme de régularisation).

Il n'y a pas de règle définie pour classer hiérarchiquement les différentes méthodes d'estimation du flot optique. Dans ce mémoire, nous citons les différentes méthodes qui peuvent exister sans faire de classement hiérarchique.

Nous allons commencer par décomposer les méthodes d'estimation [AGG88] en deux groupes : l'approche basée sur l'analyse de primitives caractérisant l'objet et l'approche basée flot optique.

Pour la première famille, le principe s'appuie sur l'extraction de caractéristiques bidimensionnelles 2D, relativement sensibles mais extrêmement discriminantes, associées aux primitives 3D dans l'image. Dans la scène 3D, ces primitives peuvent être des bornes d'occlusion, des bornes délimitant des changements de la réflectivité externe, les coins, etc. Dans l'image 2D, ces primitives peuvent se présenter sous forme de points, lignes et courbes. Cette première catégorie de méthodes exige qu'une correspondance soit établie entre l'ensemble des caractéristiques extraites et un autre groupe d'images adjacentes. Le mouvement est déterminé en estimant le déplacement de ces éléments d'une trame à l'autre. La condition requise dans cette méthode présente un inconvénient majeur car les techniques basées sur la correspondance de formes ne sont pas valables dans certains cas.

La deuxième famille n'exige pas la sélection de caractéristiques particulières. Son principe est basé sur une analyse de la variation de la luminance entre deux ou plusieurs trames successives. C'est vers cette approche que notre étude sera orientée. Elle sera plus longuement développée par la suite.

#### **1.4.1 Approches basées primitives**

Dans ce cadre, nous allons citer trois approches différentes [AGG88].

- Formulation directe dans laquelle les contraintes sont exprimées implicitement à l'aide d'une hypothèse de mouvement rigide (la distance 3D entre deux formes sur un objet rigide reste constante après le mouvement de l'objet ou de la caméra). Ces contraintes aboutissent à un système d'équations non linéaire. Deux ou trois prises de vue sont utilisées. Les déplacements observés aux dispositifs d'image 2D sont utilisés pour résoudre ce système afin de déterminer finalement les paramètres du mouvement des objets dans la scène.
- Formulation typique dans laquelle la contrainte de mouvement d'objet rigide est explicitement évoquée.
- Utilisation d'une longue séquence d'images.

Dans cette catégorie de méthodes, les difficultés principales sont d'établir une correspondance entre les primitives des images planes, de formuler rigoureusement le problème en s'appuyant sur le fait que l'objet visualisé soit soumis à un mouvement rigide, et de développer un algorithme itératif à la fois stable et précis.

#### **1.4.2 Approches basées flot optique**

La deuxième catégorie consiste à déterminer le flot optique ou le champ de déplacement bidimensionnel des intensités lumineuses dans le plan image (voir définition du flot optique et du champ bidimensionnel dans le paragraphe 1.2.3). Une telle approche n'exige pas la sélection de dispositifs particuliers. L'estimation du mouvement résume la variation instantanée de luminance d'une image. Toutes les méthodes dans cette deuxième catégorie s'appuient sur certaines hypothèses communes et sur des informations additionnelles. Parmi les hypothèses existantes, nous citons ci-dessous les deux hypothèses les plus notables et les plus utilisées dans la plupart des algorithmes. L'estimation basée flot optique n'est pas tributaire de l'analyse préalable de correspondance entre les images successives. La mesure du mouvement nécessite une étape minimale d'abstraction mathématique. Le but de cette abstraction est d'introduire des simplifications à travers certaines hypothèses.

### a) Hypothèse de conservation de la luminance

Comme évoqué précédemment, les techniques d'estimation du mouvement sont basées sur la comparaison de deux ou plusieurs trames. Si un objet change d'intensité le long de sa trajectoire et si la luminance associée à cet objet évolue de façon non modélisable, cette comparaison devient impossible. Bien que des algorithmes estimant la variation de la luminance aient été proposés [MOL91, NEG93, TRE94], comme règle générale, seule la variation proportionnelle de la luminance est prise en compte dans les techniques d'estimation du mouvement. Par la suite, nous prenons comme hypothèse qu'il n'y a pas de variation de la luminance et que la variation de l'intensité est due uniquement au mouvement global et au mouvement local. De plus, les variations perçues du signal ne doivent pas avoir d'autres causes que le mouvement. La modélisation mathématique de cette conservation de la luminance, appelée également hypothèse d'invariance de la luminance, est décrite par la relation :

$$I(x + dx, y + dy, t + dt) = I(x, y, t) \quad (30)$$

où  $I(x, y, t)$  désigne l'intensité lumineuse à la position  $(x, y)$  et à l'instant  $t$ . Le vecteur  $(dx, dy)^T$  désigne le déplacement pendant l'intervalle  $dt$ .

Malgré sa simplicité, l'équation (30) n'est pas linéaire. Son expression plus générale est de la forme :

$$I(x + u(x, y).dt, y + v(x, y).dt, t + dt) = I(x, y, t) \quad (31)$$

où  $u$  et  $v$  sont les composantes du vecteur vitesse.

### b) Hypothèse de cohérence spatiale

L'hypothèse de conservation de la luminance n'est pas suffisante car le problème soulevé par l'équation (30) est mal conditionné. Considérons en l'occurrence le cas d'un pixel qui se déplace, son mouvement est forcément une translation. Nous avons donc deux inconnues décrites par une seule équation. Le système d'équations à résoudre est donc sous-déterminé. Le calcul du flot optique subit cette ambiguïté qui s'appelle problème d'ouverture (aperture problem). Nous le détaillerons dans le paragraphe 1.5.2.

Pour rendre possible l'estimation du flot optique, de nombreuses contributions ont consisté à ajouter une ou plusieurs contraintes. La régularisation commune aux techniques d'estimation s'appuie sur la cohérence spatiale. Cette contrainte consiste à supposer que les pixels d'un voisinage spatial se propagent conformément aux mêmes paramètres de déplacement.

L'approche basée flot optique peut être divisée en trois classes primaires [FER99] :

- Méthodes différentielles ;
- Méthodes de transformée ;
- Méthodes corrélatives.

Dans les sections suivantes, nous examinerons ces différentes techniques et nous essayerons d'orienter notre choix sur la technique la plus appropriée pour notre application.

## 1.5 Estimation du flot optique selon l'approche différentielle

De nombreuses méthodes ont été proposées pour l'estimation du flot optique. Dans [STI99, ORK96, BAR94] nous trouvons un état de l'art sur quelques techniques classiques existantes [HOR81, HOR86, NAG86, VER 89, FLE90, NAG92, VER90, HEI93, COH93, OTT94, PRO94]. Le but de ce paragraphe est d'établir les différentes notions pouvant être utilisées dans l'estimation du flot optique. D'une manière générale, une technique d'estimation peut être élaborée à partir d'une ou plusieurs théories combinées.

### 1.5.1 Équation de contrainte du flot optique

Selon que l'on travaille dans le domaine temporel ou fréquentiel, l'équation de contrainte du flot optique peut se présenter différemment. Dans le domaine temporel, notre cadre d'intérêt, l'équation de contrainte du flot optique est une équation différentielle qui relie les gradients spatio-temporels aux composantes du vecteur vitesse. Voilà pourquoi elle s'appelle également équation de contrainte spatio-temporelle du flot optique. Cette équation est la base de toutes les techniques différentielles, appelées également techniques basées gradients. Dans ces techniques, les vecteurs vitesses sont déterminés à partir des gradients spatio-temporels des intensités lumineuses ou des versions filtrées de l'image (en utilisant des filtres passe bas ou passe bande). Le domaine de l'image est supposé continu, donc dérivable dans l'espace spatio-temporel. Considérons un point  $p(x, y)$  qui a effectué un déplacement  $\vec{d} = (dx, dy)^T$  pendant l'intervalle  $dt$ . Admettons l'hypothèse de conservation de la luminance, c'est-à-dire :

$$I(p + \vec{d}, t + dt) = I(p, t) \quad (32)$$

et le flot localement constant, c'est-à-dire :

$$\begin{cases} \frac{dx}{dt} = u \text{ (une constante)} \\ \frac{dy}{dt} = v \text{ (une constante)} \end{cases} \quad (33)$$

En effectuant le développement de Taylor au premier ordre du premier terme de l'équation (32), nous avons :

$$u \frac{\partial I}{\partial x} + v \frac{\partial I}{\partial y} + \frac{\partial I}{\partial t} = 0 \quad (34)$$

ou bien :

$$\vec{\nabla} I(p, t) \vec{V} + \frac{\partial I(p, t)}{\partial t} = 0 \quad (35)$$

La quantité  $p(x, y)$  désigne les coordonnées d'un pixel à l'instant  $t$  et  $\vec{V} = (u, v)^T$  le vecteur vitesse.

L'équation (34) ou (35) caractérise le flot optique par conservation de la structure lumineuse. Elle s'appelle équation du flot optique (EFO). L'équation (32) nous conduit à écrire la relation suivante :

$$I(p + dp, t + dt) - I(p, t) = 0 \quad (36)$$

ou bien :

$$I(x + dx, y + dy, t + dt) - I(x, y, t) = 0 \quad (37)$$

L'équation (37) définit un terme d'erreur qui s'appelle différence inter-pixel déplacé<sup>1</sup>. La somme des erreurs de tous les points de l'image s'appelle différence inter-image déplacée<sup>2</sup>.

La contrainte du flot optique peut être présentée sous deux formes :

- La première consiste en l'utilisation de l'équation du flot optique (34). Ce cas se retrouve souvent dans les méthodes différentielles. Ainsi, l'équation du flot optique peut aussi être appelée équation de contrainte du flot optique ou équation de contrainte du mouvement (ECFO ou bien ECM).

- Quant à la deuxième forme, la contrainte peut être imposée en minimisant la DPD ou la DFD (37). Ce point est la base des méthodes de corrélation et pel-récurrentes.

Dans ce mémoire, chaque fois que nous parlons de l'équation du flot optique, nous utilisons, par abus de langage, l'expression "équation de contrainte du flot optique" ou "équation de

<sup>1</sup> DPD : Displaced pixel difference

<sup>2</sup> DFD : Displaced frame difference

contrainte du mouvement". La raison en est simple : dans le domaine de l'estimation du mouvement, l'équation du flot optique est utilisée en terme de contrainte.

D'une manière plus formelle, l'hypothèse d'invariance de la luminance peut se traduire par :

$$\frac{d}{dt}I(p,t) = 0 \quad (38)$$

et toute équation de contrainte du mouvement plus généralisée découle de cette dernière expression.

La figure 1.10 représente un espace 2D associé aux axes  $u$  et  $v$ , qui s'appelle espace de vitesse. Tous les estimés qui obéissent à l'équation de contrainte se situent sur une ligne que nous appelons droite de contrainte.

Nous avons vu que les méthodes différentielles sont basées sur l'équation de contrainte du mouvement, donc sur l'utilisation des dérivées spatio-temporelles. Les méthodes différentielles peuvent différer essentiellement par :

- L'expression des équations de contraintes utilisées dans un critère à minimiser ;
- La méthode de résolution du système d'équations ;
- Le choix de la technique d'estimation des dérivées partielles.

On emploie généralement des équations de conservation de grandeurs physiques telles que la luminosité, l'énergie, etc., qui s'apparentent à des équations de transport ou de mécanique des fluides. On ajoute parfois des équations de contrainte liées spécifiquement au problème à résoudre.

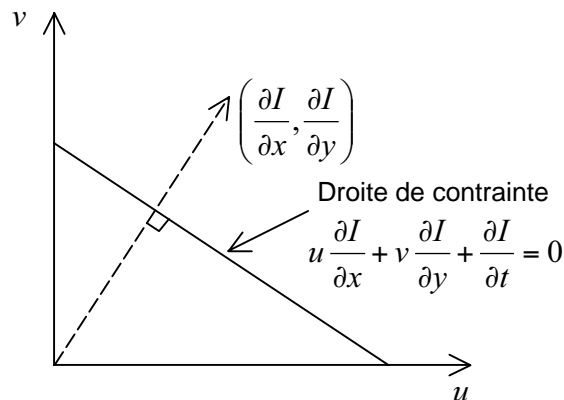


Figure 1.10 : Droite de contrainte et information locale sur les gradients spatiaux.

### a) Choix de l'équation de contrainte

Plusieurs formes d'équations de contrainte existent :

- Équation de contrainte spatio-temporelle ;
- Équation de contrainte fréquentielle ;
- Équation à multiples contraintes ou avec terme de régularisation.

### b) Techniques d'estimation des dérivées partielles

Certaines méthodes sont basées sur le développement de Taylor ou sur une approche convolutive de type Canny-Deriche.

\*Pour les filtres dérivateurs basés sur le développement de Taylor, la généralisation du principe est donnée par l'approche développée par Savitzky et Golay. Leur méthode permet, par une modélisation polynomiale, d'obtenir les coefficients du filtre pour des ordres différents ( $L = 2N + 1$  dans le développement qui suit). Cette modélisation s'effectue de la manière suivante :

- Soit un polynôme  $f(n)$  tel que :

$$f(n) = \sum_{i=0}^M a_i n^i = [1 \quad n \quad \cdots \quad n^M] [a_0 \quad a_1 \quad \cdots \quad a_M]^T \quad (39)$$

Ce polynôme peut s'exprimer sous la forme du produit matriciel suivant :

$$[1 \quad n \quad \cdots \quad n^M] [a_0 \quad a_1 \quad \cdots \quad a_M]^T = x(n)^T a \quad (40)$$

- Soit le filtrage linéaire ayant comme expression :

$$g_n = \sum_{k=-N}^N c_k \tilde{f}_{n-k} \quad (41)$$

- Les coefficients du polynôme au sens des moindres carrés connaissant une observation bruitée de la séquence  $f(n)$ , notée  $\tilde{f}(n)$ , peuvent être identifiés par :

$$[a_0 \quad a_1 \quad \cdots \quad a_M]^T = (A^T A)^{-1} A^T \tilde{f} \quad (42)$$

La matrice  $A$  est donnée par la forme :

$$A = [x(-N) \quad x(-N+1) \quad \cdots \quad x(N)]^T \quad (43)$$

L'expression du vecteur  $\tilde{f}$  est donnée par :

$$\tilde{f} = [\tilde{f}(-N) \quad \tilde{f}(-N+1) \quad \cdots \quad \tilde{f}(N)]^T \quad (44)$$

Pour la dérivation, nous remarquons que les deux modes opératoires (polynôme et convolution) sont linéaires, nous avons donc  $g_0 = a_1$ . En outre, nous avons  $c_n = a_1$  lorsque  $\tilde{f}$  est remplacé par le vecteur unité  $\vec{e}(n) = [0 \quad 0 \quad \cdots \quad 1 \quad \cdots \quad 0]^T$  pour  $-N \leq n \leq N$ . La solution est alors donnée par la relation :

$$c_n = \sum_{m=0}^M \left\{ (A^T A)^{-1} \right\}_{1,m} n^m \quad (45)$$

- A titre d'illustration, pour calculer les dérivées orthogonales, Roberts utilise des masques de deux par deux pixels.

- De même, Prewitt [BOL95] utilise des masques de  $3 \times 3$  pixels.

- Sobel [BOL95] utilise des masques légèrement différents de ceux de Prewitt en effectuant un moyennage gaussien dans une direction et une différentiation dans l'autre direction.

- D'autres formes de masques existent. En l'occurrence, dans cette étude, nous utilisons parfois les masques notés MDIF [BOL95].

Un tableau récapitulatif des masques utilisés est présenté en annexe A. Le choix de ces filtres de gradients dans la présente étude peut être justifié par le fait que nous envisageons une application qui a besoin d'un nombre restreint d'images (algorithme temps réel). En outre, l'application que nous avons étudiée, développée dans la quatrième partie de ce mémoire, concerne l'analyse d'images sismiques. Dans cette application, l'estimation du déplacement ne peut s'appuyer que sur deux trames consécutives. Cette contrainte explique la limitation sur la forme discrète des gradients spatiaux et temporels à des masques simples et non récursifs.

\*Quant à l'approche convolutive, seule l'approche de type masque gaussien de réponse impulsionnelle finie est envisagée. Le filtre de Deriche [DER87] étant le plus souvent associé à une forme récursive causale et anticausale, il est mal adapté à un problème de flot continu. Cependant, toutes ces approches peuvent être considérées.

### c) Méthode de résolution du système d'équation

Comme il y a deux inconnues ( $u$  et  $v$ ) pour une seule équation, l'équation de contrainte du flot optique est sous-déterminée. Il existe donc une infinité de flots optiques qui vérifient une

même équation. Mathématiquement, l'équation du flot optique est mal posée. Un problème d'estimation est bien posé si la solution est unique et si sa résolution numérique offre une garantie de stabilité. Le problème est mal posé dès lors que l'exigence d'unicité n'est pas remplie et il est mal conditionné si la solution n'est pas stable.

#### d) Suggestions aux problèmes liés à l'équation du mouvement

Il y a trois façons de traiter la limitation introduite par  $a$ ,  $b$  et  $c$  :

- Retrouver seulement la composante normale du mouvement. Il est possible de reconstruire le champ en le modélisant sur un voisinage .

- Utilisation des algorithmes moyennant l'observation sur l'étendue d'un voisinage, afin de renforcer l'équation de contrainte (techniques multipoints, multicontraintes). Dans le cas des voisinages dégénérés, cette approche ne pourrait pas être efficace. Nous verrons dans les exemples suivants que Lucas et Kanade [LUC81] ont proposé d'intégrer l'équation du flot optique sur un voisinage local défini par une fonction de fenêtre afin d'obtenir une solution au sens des moindres carrés. Dans [BIG91] et [JAH92], est proposée une approche qui estime le mouvement local à partir des vecteurs propres de la matrice des gradients spatio-temporels.

- Ajout de termes de régularisation pour renforcer les contraintes. Comme conséquence, il n'y a aucune garantie que le champ obtenu corresponde au champ réel. La difficulté réside dans le choix de la contrainte qui conduit à un estimé le plus précis possible par rapport à la réalité. Horn et Schunck [HOR81] étaient parmi les premiers à proposer une telle contrainte.

L'équation de contrainte conduit à un calcul local ou global sur toute l'image ou sur une région. Cette résolution peut être directe ou itérative et fournit un champ dense du champ de vélocité (flot optique). L'étude de ce champ de vitesse permet de retrouver tous les mouvements (translation, rotation, déformation) à partir d'un post-traitement adapté.

Cependant, pour être applicables, ces méthodes imposent des hypothèses assez contraignantes sur le signal observé : en particulier, les variations perçues du signal doivent être dues uniquement au mouvement. De plus, pour être observable, un mouvement doit engendrer une variation significative du signal. Leur complexité varie beaucoup en fonction du formalisme utilisé.

### 1.5.2 Composante normale du vecteur vitesse

Parmi les vecteurs du déplacement, la composante normale du vecteur vitesse est la seule que nous pouvons déterminer directement. C'est la composante du flot suivant le gradient spatial.

$$\vec{V}_n = - \frac{\frac{\partial I(p,t)}{\partial t}}{\|\vec{\nabla} I(p,t)\|} \vec{n} \quad (46)$$

$\vec{n}$  désigne le vecteur unitaire normal au contour dont l'expression est :

$$\vec{n} = \frac{\vec{\nabla} I(p,t)}{\|\vec{\nabla} I(p,t)\|} \quad (47)$$

Maintenant, nous essayons d'analyser ce problème en étendant le phénomène à un certain nombre de points contenus dans une petite ouverture  $\Omega$ . Si dans cette fenêtre, on ne voit qu'une portion de contour de l'objet, le seul déplacement estimable est celui qui est perpendiculaire au contour (figure 1.11). Ce phénomène met en évidence la limitation de calcul du flot optique. Une telle ambiguïté s'appelle problème d'ouverture (aperture problem) [MAR82, HOR81, NAG83]. Ce terme résume le critère associé au choix de la région d'analyse.

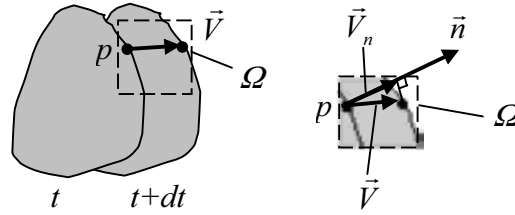


Figure 1.11 : Illustration du problème d'ouverture.

### 1.5.3 Méthodes non paramétriques (méthodes basées pixels)

Cette famille de méthodes cherche à estimer indépendamment le mouvement de chaque pixel. Contrairement aux méthodes basées régions, la notion de modèle de mouvement n'a plus d'utilité ici, car entre deux trames successives, le mouvement d'un pixel est essentiellement une translation. Les méthodes fournissent un champ de déplacement dense.

Toutes les techniques qui permettent d'estimer localement le mouvement de chaque pixel peuvent être classées dans la catégorie des méthodes non paramétriques.

- Les techniques de régularisation peuvent être locales ou globales et consistent à exploiter la similarité du déplacement d'un point avec celui de ses voisins [HOR81, MAR87, POG87]. Nous en parlerons dans la section consacrée aux moyens de régularisation.

- Les techniques récursives minimisent une fonction du carré de la DPD [NET79, CAF83]. Nous allons rappeler ici quelques méthodes récursives.

#### a) Algorithme de Netravali et Robbins

La méthode pel-réursive introduite par Netravali et Robbins [NET79, ROB83] utilise la méthode de descente du gradient. Le terme correctif est une fonction du carré de la DPD. A chaque itération, l'estimé courant est mis à jour suivant l'expression :

$$\begin{cases} u^{(k+1)} = u^{(k)} - \frac{\varepsilon}{2} \frac{\partial}{\partial u} DPD^2 \\ v^{(k+1)} = v^{(k)} - \frac{\varepsilon}{2} \frac{\partial}{\partial v} DPD^2 \end{cases} \quad (48)$$

$(u^{(k)}, v^{(k)})$  représente la valeur de l'estimé à la  $k^{\text{ème}}$  itération,  $\varepsilon$  un terme de gain de convergence.

Après avoir développé la DPD, l'équation (48) peut s'écrire sous la forme :

$$\begin{cases} u^{(k+1)} = u^{(k)} - \varepsilon DPD \frac{\partial I}{\partial x} \\ v^{(k+1)} = v^{(k)} - \varepsilon DPD \frac{\partial I}{\partial y} \end{cases} \quad (49)$$

L'équation (49) exprime le vecteur vitesse d'un pixel donné en corrigeant à chaque itération la DPD.

Le développement suivant consiste à remplacer la DPD par l'ECM. En utilisant l'équation (34), la relation (49) peut être exprimée par :

$$\begin{cases} u^{(k+1)} = u^{(k)} - \varepsilon \left( u \frac{\partial I}{\partial x} + v \frac{\partial I}{\partial y} + \frac{\partial I}{\partial t} \right) \frac{\partial I}{\partial x} \\ v^{(k+1)} = v^{(k)} - \varepsilon \left( u \frac{\partial I}{\partial x} + v \frac{\partial I}{\partial y} + \frac{\partial I}{\partial t} \right) \frac{\partial I}{\partial y} \end{cases} \quad (50)$$



En introduisant le laplacien  $\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2$  dans le second terme du second membre de la relation (50), nous avons :

$$\begin{cases} u^{(k+1)} = u^{(k)} - \varepsilon \Delta u \left[ \left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2 \right] \\ v^{(k+1)} = v^{(k)} - \varepsilon \Delta v \left[ \left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2 \right] \end{cases} \quad (51)$$

avec :

$$\begin{cases} \Delta u = \frac{\frac{\partial I}{\partial x} \left( u \frac{\partial I}{\partial x} + v \frac{\partial I}{\partial y} + \frac{\partial I}{\partial t} \right)}{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2} \\ \Delta v = \frac{\frac{\partial I}{\partial y} \left( u \frac{\partial I}{\partial x} + v \frac{\partial I}{\partial y} + \frac{\partial I}{\partial t} \right)}{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2} \end{cases} \quad (52)$$

De l'équation (51), nous pouvons déduire que les méthodes pel-récurrentes cherchent à minimiser l'écart  $\Delta u$  et  $\Delta v$  pour que la solution  $\vec{V} = (u, v)^T$  à la  $k^{\text{ème}}$  itération s'approche de la droite de contrainte (figure 1.12). Les éléments  $I_x = \partial I / \partial x$ ,  $I_y = \partial I / \partial y$  et  $I_t = \partial I / \partial t$  désignent respectivement les gradients suivant  $x$ ,  $y$  et  $t$ .

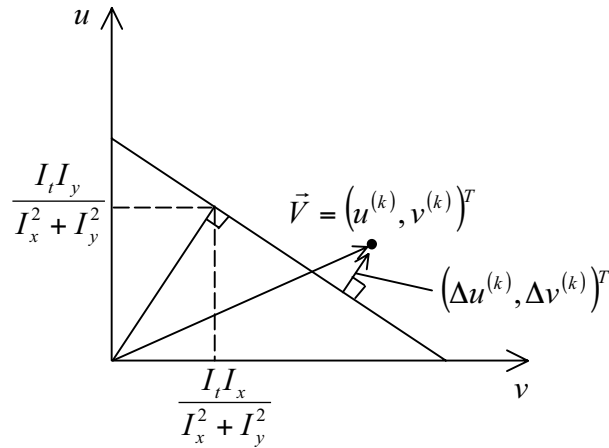


Figure 1.12 : Comportement de la convergence d'un point vers la droite de contrainte, dans les approches pel-récurrentes.

La convergence et la vitesse de convergence de l'algorithme de Netravali et Robins dépendent du choix de la valeur du paramètre  $\varepsilon$  et la valeur initiale de l'estimé. Plus  $\varepsilon$  est faible plus le nombre d'itérations sera grand. En revanche, une valeur forte pour  $\varepsilon$  pourrait produire un phénomène oscillatoire. On peut prendre comme valeur initiale de l'itération le déplacement calculé sur le pixel précédent. Cela permet d'introduire implicitement une contrainte de similarité sur le mouvement du pixel.

### b) Algorithme de Walker et Rao

Dans [WAL84], le facteur de convergence  $\varepsilon$  est adaptatif. Le principe d'adaptation s'appuie sur les deux phénomènes suivants :

- Dans un voisinage où  $\|\bar{\nabla}I\|$  est important,  $\varepsilon$  devrait être faible si la DPD est faible. Donc, une faible mise à jour sera suffisante.

- Dans une région uniforme où  $\|\bar{\nabla}I\|$  est faible, un faible facteur de convergence sera suffisant quand la DPD est grande.

Ces deux motivations conduisent à remplacer le facteur de convergence par l'inverse du module du gradient spatial. Soit :

$$\varepsilon = \frac{1}{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2} \quad (53)$$

Les composantes du vecteur vitesse deviennent :

$$\begin{cases} u^{(k+1)} = u^{(k)} - \Delta u \\ v^{(k+1)} = v^{(k)} - \Delta v \end{cases} \quad (54)$$

### c) Algorithme de Caffario et Rocca

Remarquons que dans la relation (53), si les gradients spatiaux sont faibles, l'itération risque de diverger à cause de la division par zéro. Pour éviter ce phénomène, on peut borner le terme correcteur en introduisant un facteur supplémentaire  $\alpha$  [CAF83]. Soit :

$$\varepsilon = \frac{1}{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2 + \alpha^2} \quad (55)$$

Les composantes du vecteur vitesse sont exprimées par :

$$\begin{cases} u^{(k+1)} = u^{(k)} - \frac{1}{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2 + \alpha^2} \left( u \frac{\partial I}{\partial x} + v \frac{\partial I}{\partial y} + \frac{\partial I}{\partial t} \right) \frac{\partial I}{\partial x} \\ v^{(k+1)} = v^{(k)} - \frac{1}{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2 + \alpha^2} \left( u \frac{\partial I}{\partial x} + v \frac{\partial I}{\partial y} + \frac{\partial I}{\partial t} \right) \frac{\partial I}{\partial y} \end{cases} \quad (56)$$

La valeur typique de  $\alpha$  est égale à 10 [TEK95]. Cette approche utilise le même principe que celle proposée par Horn et Schunck [HOR81] (cf. paragraphe 1.5.6.a).

### d) Discussion

Les méthodes pel-récurrentes présentent cependant plusieurs inconvénients par rapport aux autres méthodes différentielles. Il faut tout d'abord souligner la faiblesse de la contrainte appliquée qui impose une similarité des mouvements de proche en proche. Elle débouche sur un algorithme d'estimation très locale qui fournit des champs de vecteurs extrêmement perturbés. De plus, comme nous l'avons démontré, la DPD est une estimation très discrète donc peu sensible aux variations d'intensité dans la direction du mouvement. Par conséquent, les termes correcteurs  $\Delta u$  et  $\Delta v$  qui contrôlent la convergence sont peu fiables et risquent de conduire l'algorithme vers des solutions erronées. Par ailleurs Walter [WAL96] qui a testé la méthode [CAF83] sur les séquences sismiques, n'a pas pu obtenir de résultats probants.

### 1.5.4 Méthodes basées sur le modèle paramétrique (méthodes basées régions)

Dans cette famille de techniques, l'image est segmentée en régions de contours arbitraires. Un modèle de mouvement paramétrique est associé à chaque région. Ce modèle sert de contrainte de mouvement pour les pixels contenus dans la région. Dans cette famille, deux approches différentes peuvent être évoquées : méthodes indirectes et directes.

#### a) Méthodes indirectes

Pour les méthodes indirectes, les paramètres du mouvement ne sont pas calculés directement à partir de l'information luminance mais plutôt selon les deux étapes suivantes :

- Estimation du champ dense pour chaque région, sans tenir compte du modèle paramétrique. En d'autres termes, on estime d'abord le champ dense du mouvement à l'aide d'une méthode non paramétrique. Cette première phase peut être réalisée à l'aide des différentes techniques : méthodes corrélatives, méthodes différentielles, ...

- Les vecteurs ainsi obtenus sont modélisés par un ensemble de paramètres, en utilisant une technique d'optimisation.

L'appellation indirecte vient de la présence de l'étape intermédiaire. Cette approche s'appelle également technique régressive.

A titre d'illustration, nous allons présenter une méthode indirecte permettant d'estimer des paramètres du mouvement selon les deux étapes suivantes :

\**Étape n°1* : Nous estimons le champ dense du vecteur vitesse en supposant que le mouvement est uniforme et de translation dans une fenêtre  $\Omega$ . Un critère à minimiser peut être défini par :

$$J = \sum_{\Omega} \left( u \frac{\partial I}{\partial x} + v \frac{\partial I}{\partial y} + \frac{\partial I}{\partial t} \right)^2 \quad (57)$$

La minimisation de l'équation (57) au sens des moindres carrés donne un champ dense  $(u, v)$ .

\**Étape n°2* : Nous identifions le champ dense à un modèle de mouvement. En partant de l'estimé ci-dessus et en admettant une région prédéfinie  $S$  (ou segmentée au sens du mouvement), les paramètres de cette région peuvent être déterminés.

Désignons par  $u(P, \Theta)$  et  $v(P, \Theta)$  les modèles complets du vecteur vitesse pour un point  $P(x, y)$ . La quantité  $\Theta$  représente le vecteur de paramètres du mouvement. La minimisation des expressions suivantes au sens des moindres carrés permet d'obtenir les paramètres du mouvement :

$$\min_{\Theta} \sum_S [u - u(P, \Theta)]^2 \quad (58)$$

et :

$$\min_{\Theta} \sum_S [v - v(P, \Theta)]^2 \quad (59)$$

Les coefficients des paramètres du mouvement sont calculés en égalisant à zéro les dérivées partielles respectivement à  $\Theta_i$ ,  $i = 1, \dots, m$  ; ( $m$  étant le nombre de paramètres du modèle utilisé). Cette approche est sensible aux mauvais estimés initiaux. La précision obtenue par les techniques indirectes dépend fortement des valeurs introduites initialement par l'estimation dense du vecteur de mouvement.

#### b) Méthodes directes

Les méthodes directes estiment directement les paramètres du modèle utilisé. On part d'un critère basé sur l'information de luminance et la minimisation de ce critère conduit à

déterminer directement les paramètres du mouvement. Ce modèle sert de contrainte de mouvement pour les pixels contenus dans la région. L'équation du modèle est intégrée dans une fonctionnelle. La plupart des fonctionnelles sont établies à partir de la DFD. Le critère à minimiser est plus généralement défini par :

$$\min_{\Theta} \sum_S [I(x, y, t) - I(x + dx(\Theta), y + dy(\Theta), t + dt)]^2 \quad (60)$$

La quantité  $\Theta$  désigne le vecteur de paramètres du mouvement et  $[dx(\Theta), dy(\Theta)]^T$  le modèle de mouvement.

De nombreuses méthodes ont été proposées pour minimiser ce critère. Nous pouvons utiliser, entre autres, les méthodes différentielles ou les méthodes de corrélation basées sur un schéma itératif. Les méthodes itératives associées aux techniques paramétriques s'appellent également : techniques de région récursive. L'itération est de la forme :

$$\Theta^{(k+1)} = \Theta^{(k)} - \Delta\Theta^{(k)} \quad (61)$$

où  $\Delta\Theta^{(k)}$  est le terme correctif à la  $k^{\text{ème}}$  itération.

A titre d'illustration, la méthode suivante décrit une estimation des paramètres du mouvement selon un procédé direct :

On estime le flot optique en utilisant le modèle affine (voir tableau 1.1) et la méthode différentielle. Ce modèle est défini par :

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} a_1 + a_2x + a_3y \\ b_1 + b_2x + b_3y \end{pmatrix} \quad (62)$$

En transformant la relation (60) sous une forme d'équation de contrainte du mouvement et en tenant compte des relations  $u = dx(\Theta)/dt$ ,  $v = dy(\Theta)/dt$ , le critère à minimiser peut être défini par :

$$J = \sum_{\Omega} \left( (a_1 + a_2x + a_3y) \frac{\partial I}{\partial x} + (b_1 + b_2x + b_3y) \frac{\partial I}{\partial y} + \frac{\partial I}{\partial t} \right)^2 \quad (63)$$

La somme s'étend sur un voisinage  $\Omega$  de taille  $n$ . La minimisation de la fonctionnelle  $J$  par la méthode des moindres carrés donne :

$$\Theta = -(A^T A)^{-1} A^T B \quad (64)$$

La quantité  $\Theta$  représente le vecteur de paramètres du mouvement. Son expression est donnée par la forme :

$$\Theta = (a_1, a_2, a_3, b_1, b_2, b_3)^T \quad (65)$$

La matrice  $A$  contient des coefficients relatifs aux gradients spatiaux. Elle est définie par :

$$A = \begin{bmatrix} I_x(1) & x_1 I_x(1) & y_1 I_x(1) & I_y(1) & x_1 I_y(1) & y_1 I_y(1) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ I_x(n) & x_n I_x(n) & y_n I_x(n) & I_y(n) & x_n I_y(n) & y_n I_y(n) \end{bmatrix} \quad (66)$$

Pour les gradients temporels, nous avons :

$$B = \begin{bmatrix} I_t(1) \\ \vdots \\ I_t(n) \end{bmatrix} \quad (67)$$

Les éléments  $I_x(p) = \partial I(p)/\partial x$ ,  $I_y(p) = \partial I(p)/\partial y$  et  $I_t(p) = \partial I(p)/\partial t$  désignent respectivement les gradients suivant  $x$ ,  $y$  et  $t$ , au point  $p$ .

## c) Discussion

Nous pouvons observer que l'utilisation du modèle paramétrique permet implicitement de contraindre l'estimation du champ de déplacement et l'ajout d'une contrainte additionnelle n'est généralement pas fait. En revanche, dans la première étape de l'approche indirecte, plusieurs auteurs ajoutent des contraintes additionnelles telles que le terme de régularisation, terme de pondération, etc.. Nous notons au passage que dans le présent travail, nous avons cherché un compromis entre ces deux aspects, c'est-à-dire que nous appliquons une méthode paramétrique directe mais également avec ajout de contrainte.

### 1.5.5 La notion de régularisation

Le problème d'estimation du flot optique à partir de la conservation de la luminance peut être ramené au problème de minimisation suivant :

Trouver à chaque pixel  $(x,y)$  le déplacement  $d = (dx, dy)^T$  qui minimise le critère :

$$J = \sum_x \sum_y [I(x, y, t) - I(x + dx, y + dy, t + dt)]^2 \quad (68)$$

C'est un problème de minimisation classique sur lequel on peut employer un algorithme de descente suivant la plus grande pente en chaque point. Cependant, à cause du problème d'ouverture, la solution n'est pas unique. La théorie de régularisation consiste à reformuler le problème en introduisant une information a priori (une contrainte) dont le but est d'assurer l'existence, l'unicité et la stabilité de la solution.

Il existe deux types de régularisation :

- Régularisation à contrainte globale : cette catégorie de méthode ajoute une contrainte globale qui s'exerce sur l'intégralité des déplacements de la région traitée. Nous verrons dans le paragraphe suivant quelques méthodes de régularisation globale telles que la méthode de Horn et Schunck, la méthode de Hildreth et la méthode de Nagel.
- Régularisation à contrainte locale : cette catégorie de méthode utilise l'information du vecteur normal dans des voisinages locaux et s'appuie sur la résolution au sens des moindres carrés pour estimer le déplacement convenable. Par exemple, Lucas et Kanade utilisent un terme de régularisation locale en affectant un poids à chaque gradient des pixels d'un voisinage considéré. Cette méthode est connue sous l'appellation résolution au sens des moindres carrés pondérés.

### 1.5.6 Méthodes à contraintes globales

Cette famille de méthode utilise une fonctionnelle à minimiser sur la totalité des déplacements de l'image. Plus formellement, ces techniques introduisent une contrainte spatiale sur le champ de vélocité, pour fournir un champ dense. Les techniques à contraintes globales sont généralement plus robustes et plus performantes par rapport aux techniques à contraintes locales.

Nous notons au passage qu'il ne faut pas confondre les deux notions présentées dans 1.5.5 et 1.5.6 :

- Pour la régularisation à contrainte globale, c'est l'ajout de contrainte globale qui s'exerce sur l'intégralité des déplacements de l'image ou sur la totalité d'une région.
- Quant à l'estimation à contrainte globale, c'est la minimisation du critère qui s'effectue sur la totalité des déplacements de l'image.

Ces notions sont différentes mais peuvent être associées. En l'occurrence, dans une estimation à contrainte globale, on peut introduire une technique de régularisation à contrainte globale mais très rarement une régularisation à contrainte locale.

### a) La méthode de Horn et Schunck

Cette méthode peut être classée dans des catégories diverses : domaine spatio-temporel, méthode différentielle, méthode de régularisation, méthode non paramétrique, etc.

L'approche permettant de contrôler le mauvais conditionnement de l'équation du flot optique, consiste à contraindre les estimés en combinant cette équation à un terme de lissage global. C'est une méthode introduite par Horn et Schunck [HOR81]. Elle minimise simultanément l'équation de contrainte et les variations spatiales du champ de vitesse. La variation pixel à pixel du champ de vitesse peut être quantifiée par la somme des carrés des normes de gradients spatiaux, de chaque composante du vecteur vitesse. Avec cette contrainte, le vecteur vitesse est obtenu par minimisation du critère suivant :

$$J = \iint_{\Omega} \left[ \left( u \frac{\partial I}{\partial x} + v \frac{\partial I}{\partial y} + \frac{\partial I}{\partial t} \right)^2 + \alpha^2 \left( \|\bar{\nabla} u(x, y)\|^2 + \|\bar{\nabla} v(x, y)\|^2 \right) \right] dx dy \quad (69)$$

ou encore :

$$J = \iint_{\Omega} \left[ \left( u \frac{\partial I}{\partial x} + v \frac{\partial I}{\partial y} + \frac{\partial I}{\partial t} \right)^2 + \alpha^2 \left( \left( \frac{\partial u}{\partial x} \right)^2 + \left( \frac{\partial u}{\partial y} \right)^2 + \left( \frac{\partial v}{\partial x} \right)^2 + \left( \frac{\partial v}{\partial y} \right)^2 \right) \right] dx dy \quad (70)$$

$\Omega$  désigne un support continu de l'image et  $\alpha^2$  le facteur de pondération qui contrôle la robustesse de la contrainte de lissage. Ce facteur est prépondérant dans les zones où les gradients de la luminance sont faibles.

Pour trouver les valeurs appropriées de  $u$  et  $v$ , on minimise la fonctionnelle  $J$  par des calculs de variations, par introduction d'une estimation discrète du laplacien et finalement par une procédure itérative de Gauss-Seidel [HOR81]. La solution ainsi obtenue est de la forme :

$$\begin{cases} u^{(k+1)} = \bar{u}^{(k)} - \frac{\frac{\partial I}{\partial x} \bar{u}^{(k)} + \frac{\partial I}{\partial y} \bar{v}^{(k)} + \frac{\partial I}{\partial t}}{\alpha^2 + \left( \frac{\partial I}{\partial x} \right)^2 + \left( \frac{\partial I}{\partial y} \right)^2} \\ v^{(k+1)} = \bar{v}^{(k)} - \frac{\frac{\partial I}{\partial x} \bar{u}^{(k)} + \frac{\partial I}{\partial y} \bar{v}^{(k)} + \frac{\partial I}{\partial t}}{\alpha^2 + \left( \frac{\partial I}{\partial x} \right)^2 + \left( \frac{\partial I}{\partial y} \right)^2} \end{cases} \quad (71)$$

avec :

$$\begin{aligned} \bar{u}(i, j) = & \frac{1}{6} [u(i-1, j) + u(i, j+1) + u(i+1, j) + u(i, j-1)] \\ & + \frac{1}{12} [u(i-1, j-1) + u(i-1, j+1) + u(i+1, j+1) + u(i+1, j-1)] \end{aligned} \quad (72)$$

On applique la même règle pour le calcul de  $\bar{v}(i, j)$ .

### b) Mouvement le long des contours : méthode de Hildreth

Hildreth a proposé une méthode de régularisation qui cherche la vitesse du contour de l'objet, en appliquant la méthode de Horn et Schunck dans un espace curviligne. L'équation de gradient et la contrainte de lissage sont appliquées uniquement sur les vecteurs vitesse projetés sur les normales aux contours de l'image [HIL84, GON90].

$$J = \alpha \int_{\text{contour}} \frac{1}{\sqrt{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2}} \left( u \frac{\partial I}{\partial x} + v \frac{\partial I}{\partial y} + \frac{\partial I}{\partial t} \right)^2 dp + \int_{\text{contour}} \left[ \left( \frac{\partial u}{\partial p} \right)^2 + \left( \frac{\partial v}{\partial p} \right)^2 \right] dp \quad (73)$$

La quantité  $p$  désigne un point arbitraire du contour,  $\alpha$  est une constante de pondération. Sous forme discrète, la fonctionnelle (73) peut s'écrire :

$$J = \sum_{i=2}^{n-1} \left\{ [u(i) - u(i-1)]^2 + [u(i) - u(i+1)]^2 + [v(i) - v(i-1)]^2 + [v(i) - v(i+1)]^2 \right\} + \alpha \sum_{i=1}^n \frac{1}{\sqrt{I_x^2 + I_y^2}} [I_x(i)u(i) + I_y(i)v(i) + I_t(i)]^2 \quad (74)$$

Les éléments  $I_x$ ,  $I_y$  et  $I_t$  représentent respectivement les gradients spatiaux suivant  $x$ ,  $y$  et le gradient temporel ;  $n$  est le nombre de points du contour.

La minimisation de la fonctionnelle  $J$  s'effectue en calculant ses dérivées partielles respectivement par rapport aux  $u(i)$ ,  $v(i)$  avec  $i=1 \dots n$  et en les égalant à zéro. Nous obtenons alors un système de  $2n$  équations. Soit pour  $i=1$  à  $n$  :

$$\begin{cases} [4 + 2\alpha [N_x(i)]^2] u(i) - 2u(i+1) - 2u(i-1) + 2\alpha N_x(i)N_y(i)v(i) = 2\alpha N_t(i)N_x(i) \\ [4 + 2\alpha [N_y(i)]^2] v(i) - 2v(i+1) - 2v(i-1) + 2\alpha N_x(i)N_y(i)u(i) = 2\alpha N_t(i)N_y(i) \end{cases} \quad (75)$$

avec :

$$N_x = \frac{I_x}{\sqrt{I_x^2 + I_y^2}}, N_y = \frac{I_y}{\sqrt{I_x^2 + I_y^2}} \text{ et } N_t = \frac{I_t}{\sqrt{I_x^2 + I_y^2}} \quad (76)$$

La résolution du système s'effectue à l'aide de la méthode des gradients conjugués. Cette méthode fournit un champ de vecteur vitesse peu dense et se heurte au problème de correspondance entre les contours et les objets mobiles. Ce n'est pas une méthode de bas niveau car elle nécessite l'extraction préalable des contours ; elle n'est donc citée ici que pour mémoire.

### c) Méthode de Nagel

A la différence de la méthode de Horn et Schunck, cette méthode intègre une contrainte de régularité orientée [NAG83, NAG87, NAG86]. Cette contrainte évite au lissage d'être imposé au travers d'une variation brusque des gradients. Nagel a modifié la contrainte de lissage en introduisant une matrice de poids qui dépend des variations de niveau de gris de telle sorte que la contrainte de lissage soit retenue seulement pour le flot optique perpendiculaire aux transitions importantes du niveau de gris. Le problème est formulé par la minimisation de la fonctionnelle suivante [NAG83] :

$$J = \iint_{\Omega} \left\{ \left( u \frac{\partial I}{\partial x} + v \frac{\partial I}{\partial y} + \frac{\partial I}{\partial t} \right)^2 + \alpha^2 \text{trace} \left[ \left( \vec{\nabla} V \right)^T W \left( \vec{\nabla} V \right) \right] \right\} dx dy \quad (77)$$

La quantité  $\vec{V} = (u, v)$  désigne le vecteur vitesse. La matrice de pondération  $W$  est donnée par l'expression suivante [NAG87] :

$$W = \frac{F + \begin{bmatrix} \gamma & 0 \\ 0 & \gamma \end{bmatrix}}{\text{trace} \left( F + \begin{bmatrix} \gamma & 0 \\ 0 & \gamma \end{bmatrix} \right)} \quad (78)$$

La quantité  $\gamma$  est une constante et  $F$  a comme expression :

$$F = \begin{bmatrix} I_y \\ -I_x \end{bmatrix} \begin{bmatrix} I_y \\ -I_x \end{bmatrix}^T \quad (79)$$

Le développement de l'équation (77) donne [NAG87] :

$$J = \iint_{\Omega} \left\{ \left( u \frac{\partial I}{\partial x} + v \frac{\partial I}{\partial y} + \frac{\partial I}{\partial t} \right)^2 + \frac{\alpha^2}{\|\vec{\nabla}I\|^2 + 2\gamma} \left[ \left( \frac{\partial u}{\partial x} I_y - \frac{\partial u}{\partial y} I_x \right)^2 + \left( \frac{\partial v}{\partial x} I_y - \frac{\partial v}{\partial y} I_x \right)^2 + \gamma \left( \left( \frac{\partial u}{\partial x} \right)^2 + \left( \frac{\partial u}{\partial y} \right)^2 + \left( \frac{\partial v}{\partial x} \right)^2 + \left( \frac{\partial v}{\partial y} \right)^2 \right) \right] \right\} dx dy \quad (80)$$

Si  $\gamma=0$ , le deuxième terme de la fonctionnelle  $J$  représente le produit scalaire du vecteur  $\vec{\nabla}V$  et du vecteur normal au gradient. Cela veut dire que la minimisation de  $J$ , par rapport au couple  $(u, v)$ , atténue les variations du flot  $\vec{\nabla}V$  dans la direction perpendiculaire au gradient.

En utilisant la méthode itérative de Gauss-Seidel, la solution peut être exprimée par :

$$\begin{cases} u^{(k+1)} = \xi(u^{(k)}) - \frac{I_x [I_x \xi(u^{(k)}) + I_y \xi(v^{(k)}) + I_t]}{I_x^2 + I_y^2 + \alpha^2} \\ v^{(k+1)} = \xi(v^{(k)}) - \frac{I_y [I_x \xi(u^{(k)}) + I_y \xi(v^{(k)}) + I_t]}{I_x^2 + I_y^2 + \alpha^2} \end{cases} \quad (81)$$

Les expressions de  $\xi(u)$  et  $\xi(v)$  à la  $k^{\text{ème}}$  itération sont données par :

$$\begin{cases} \xi(u^{(k)}) = \bar{u}^{(k)} - 2I_x I_y \frac{\partial^2 u^{(k)}}{\partial x \partial y} - q^T \frac{\partial u^{(k)}}{\partial x} \\ \xi(v^{(k)}) = \bar{v}^{(k)} - 2I_x I_y \frac{\partial^2 v^{(k)}}{\partial x \partial y} - q^T \frac{\partial v^{(k)}}{\partial y} \end{cases} \quad (82)$$

avec :

$$q = \frac{1}{I_x^2 + I_y^2 + 2\gamma} \nabla I^T \left( \begin{bmatrix} I_{yy} & -I_{xy} \\ -I_{xy} & I_{xx} \end{bmatrix} + 2 \begin{bmatrix} I_{xx} & I_{xy} \\ -I_{xy} & I_{yy} \end{bmatrix} W \right) \quad (83)$$

Les éléments  $\bar{u}$  et  $\bar{v}$  représentent les valeurs moyennes de  $u$  et  $v$  dans un voisinage local. Dans l'équation (83), les notations suivantes sont introduites :  $I_{xx} = \partial^2 I / \partial x^2$ ,  $I_{yy} = \partial^2 I / \partial y^2$ ,  $I_{xy} = \partial^2 I / \partial x \partial y$  et  $I_{yx} = \partial^2 I / \partial y \partial x$ .

Les trois méthodes précédentes, Horn et Schunck, Hildreth, Nagel, introduisent une contrainte de régularité de la variation du flot optique. Dans le cas de la méthode de Horn et Schunck, cette contrainte est appliquée inconditionnellement sur toute l'image. Dans le cas de la méthode de Nagel, on utilise une contrainte d'orientation locale pour diriger la propagation de contrainte de régularité. La méthode de Nagel évite donc en grande partie le problème de la continuité forcée induite par Horn et Schunck. En revanche, dans [BAR94], nous pouvons constater que la méthode de Horn et Schunck permet d'obtenir des résultats plus probants, pour une image préalablement lissée par un filtre spatio-temporel de type Gaussien avec un écart type de 1.5 pixel.



### d) Champ de Markov

Dans les méthodes de régularisation précédentes, nous avons vu que la solution pour traiter le mauvais conditionnement de l'ECM consiste à introduire un terme de régularisation qui permet d'avoir un compromis entre un pixel et ses pixels voisins. La régularisation introduite par une hypothèse markovienne<sup>3</sup> (MRF) découle de ce même principe, en précisant que le mouvement d'un pixel n'est influencé que par celui de ses voisins [GEM84, PER91]. La fonction de densité de la probabilité conditionnelle d'un point dépend uniquement des valeurs du champ aléatoire d'un voisinage de ce point. La probabilité pour qu'un point prenne une valeur, sachant la valeur des points de son voisinage, est égale à la probabilité en ce point sachant celle de tous les autres points. Le théorème de Hammersley-Clifford donne l'équivalence, dans un processus de Markov, entre une distribution de probabilité et une distribution de Gibbs d'une fonction d'énergie [TEK95]. Cette énergie est la somme des fonctions de potentiels exprimant les interactions locales qui, dans le cas du mouvement, sont représentées par les gradients de la vitesse  $\partial u/\partial x$ ,  $\partial u/\partial y$ ,  $\partial v/\partial x$ ,  $\partial v/\partial y$ .

Appliquée à la contrainte du flot optique, cette approche conduit à une formulation du type de celle de Horn et Schunck, dans laquelle il faut minimiser sur toute l'image la fonctionnelle  $J$ , définie par :

$$J = \iint_{\Omega} \left[ \left( u \frac{\partial I}{\partial x} + v \frac{\partial I}{\partial y} + \frac{\partial I}{\partial t} \right)^2 + \Phi \left( \left( \frac{\partial u}{\partial x} \right)^2 \right) + \Phi \left( \left( \frac{\partial u}{\partial y} \right)^2 \right) + \Phi \left( \left( \frac{\partial v}{\partial x} \right)^2 \right) + \Phi \left( \left( \frac{\partial v}{\partial y} \right)^2 \right) \right] dx dy \quad (84)$$

La fonction de distribution de probabilité  $\Phi$ , telle que  $0 \leq \Phi \leq 1$ , permet de ne propager la contrainte de régularité que si celle-ci est en-dessous d'un certain seuil. La régularité est évaluée a posteriori dans un schéma itératif.

L'intérêt de cette modélisation est la prise en compte des interactions locales entre les pixels tout en ayant une approche globale. C'est une approche stochastique dont le but est de prendre en compte les discontinuités du champ de vélocité.

### 1.5.7 Méthodes à contraintes locales

Les méthodes à contraintes locales estiment le flot optique en résolvant un groupe d'équations du flot optique, comme étant un système d'équations définies sur un voisinage local (petite région). La contrainte peut agir différemment d'un pixel à un autre d'une manière indépendante. Généralement, ces méthodes font appel à des équations multicontraintes.

#### a) Estimation au sens des moindres carrés

Cette méthode détermine le vecteur vitesse en minimisant une fonction de somme quadratique des erreurs sur un voisinage local  $\Omega$ . Le critère à minimiser est de la forme :

$$J = \sum_{\Omega} \left( \vec{\nabla} I \cdot \vec{V} + \frac{\partial I}{\partial t} \right)^2 \quad (85)$$

Pour estimer le vecteur  $\vec{V} = (u, v)^T$  au sens des moindres carrés [GIO85, SIM93], nous calculons la dérivée suivant  $V$  de cette fonctionnelle quadratique. Nous avons ainsi :

$$\frac{\partial}{\partial V} J = 2 \sum_{\Omega} \vec{\nabla} I \left( \vec{\nabla} I^T \cdot \vec{V} + \frac{\partial I}{\partial t} \right) \quad (86)$$

Soit :

$$\frac{\partial}{\partial V} J = 2(M\vec{V} + N) \quad (87)$$

<sup>3</sup> Markov random field

avec :

$$M = \sum_{\Omega} \vec{\nabla} I \cdot \vec{\nabla} I^T = \begin{bmatrix} \sum_{\Omega} \left( \frac{\partial I}{\partial x} \right)^2 & \sum_{\Omega} \left( \frac{\partial I}{\partial x} \right) \left( \frac{\partial I}{\partial y} \right) \\ \sum_{\Omega} \left( \frac{\partial I}{\partial x} \right) \left( \frac{\partial I}{\partial y} \right) & \sum_{\Omega} \left( \frac{\partial I}{\partial y} \right)^2 \end{bmatrix} \quad (88)$$

et :

$$N = \begin{bmatrix} \sum_{\Omega} \left( \frac{\partial I}{\partial x} \right) \left( \frac{\partial I}{\partial t} \right) \\ \sum_{\Omega} \left( \frac{\partial I}{\partial y} \right) \left( \frac{\partial I}{\partial t} \right) \end{bmatrix} \quad (89)$$

En égalant la dérivée de la fonctionnelle  $J$  à zéro, nous obtenons l'estimé  $\vec{V}$  sous la forme :

$$\vec{V} = -M^{-1}N \quad (90)$$

En introduisant la notion du pseudo-inverse d'une matrice (voir dans  $b$ ), nous obtenons le même résultat.

### b) Méthode de Lucas et Kanade

Lucas et Kanade appliquent, à l'équation de contrainte locale, une résolution au sens des moindres carrés pondérés et cela avec un modèle constant de déplacement [LUC81]. Le critère à minimiser est de la forme :

$$J = \sum_{\Omega} [\alpha(p)]^2 \left( u \frac{\partial I}{\partial x} + v \frac{\partial I}{\partial y} + \frac{\partial I}{\partial t} \right)^2 \quad (91)$$

Le terme  $\Omega$  représente un voisinage spatial de points  $p(x, y)$ , et  $\alpha(p)$  le coefficient de pondération qui met plus d'influence sur les contraintes au centre du voisinage qu'à la périphérie. La solution au sens de moindres carrés de (91) est donnée par :

$$A^T W^2 A \vec{V} = -A^T W^2 B \quad (92)$$

où :

$$A = \begin{bmatrix} I_x(1) & I_y(1) \\ \vdots & \vdots \\ I_x(n) & I_y(n) \end{bmatrix} \quad (93)$$

et :

$$B = \begin{bmatrix} I_t(1) \\ \vdots \\ I_t(n) \end{bmatrix} \quad (94)$$

Les composantes de la matrice des poids sont exprimées par la forme :

$$W = \text{diag}[\alpha(1) \quad \dots \quad \alpha(n)] \quad (95)$$

La variable  $n$  étant le nombre de points du voisinage  $\Omega$ . Étant donné que  $A^T W^2 A$  est une matrice carrée, si elle n'est pas singulière, la solution peut se trouver sous la forme :

$$\vec{V} = -(A^T W^2 A)^{-1} A^T W^2 B \quad (96)$$

Dans le cas où la matrice de pondération est égale à une matrice identité, nous revenons à la méthode des moindres carrés classique.

En multipliant le second membre de l'équation (91) par  $\|\vec{\nabla}I(p,t)\|^2$  et en le divisant par cette même quantité, nous avons une fonctionnelle exprimée en fonction de la vitesse normale :

$$J(v_x, v_y) = \sum_{\Omega} [\alpha(p)]^2 [w(p)]^2 [\vec{V} \cdot \vec{n} + V_n]^2 \quad (97)$$

avec  $[w(p)]^2 = \|\vec{\nabla}I(p,t)\|^2$ . Les expressions du vecteur normal  $\vec{n}$  et de la vitesse normale  $V_n$  ont été présentées dans le paragraphe 1.5.2. Le coefficient  $[w(p)]^2$  se comporte comme une valeur de confiance de l'estimation du champ normal.

Dans [BAR94], nous trouvons un exemple de valeur du voisinage  $\Omega$  ( $5 \times 5$  pixels), caractérisé par une matrice de pondération gaussienne séparable  $W^2$ , dont le noyau 1D est [0.0625 0.25 0.375 0.25 0.0625].

L'application de l'équation de contrainte, non seulement sur un point, mais plutôt sur l'étendue d'un voisinage permet de lever l'ambiguïté sur l'équation de contrainte de mouvement. Cela dit, ce choix ne garantit pas que le système d'équation soit soluble, c'est-à-dire que  $[A^T W^2 A]$  soit non singulière. La technique des moindres carrés nécessite un test de conditionnement sur la singularité de cette matrice. Le champ estimé est donc a priori moins dense que celui estimé par les méthodes de régularisation ; les problèmes d'ouverture et d'existence des gradients peuvent subsister dans la technique des moindres carrés. Pour palier à ces problèmes, nous essayons d'établir, dans ce manuscrit, une formulation qui consiste à renforcer les contraintes dans la résolution au sens des moindres carrés.

### c) Méthodes différentielles du second ordre

L'équation de contrainte du flot optique est une équation différentielle du premier ordre qui est basée sur l'hypothèse de conservation de la luminance.

Les méthodes différentielles du second ordre s'appuient sur l'observation de l'intensité lumineuse ainsi que sur ses dérivées premières qui sont conservées lors du déplacement. Ces conditions impliquent :

$$\begin{cases} I(x, y, t) - I(x + dx, y + dy, t + dt) = 0 \\ \frac{\partial}{\partial x} I(x, y, t) - \frac{\partial}{\partial x} I(x + dx, y + dy, t + dt) = 0 \\ \frac{\partial}{\partial y} I(x, y, t) - \frac{\partial}{\partial y} I(x + dx, y + dy, t + dt) = 0 \\ \frac{\partial}{\partial t} I(x, y, t) - \frac{\partial}{\partial t} I(x + dx, y + dy, t + dt) = 0 \end{cases} \quad (98)$$

Comme avec l'équation du mouvement à simple contrainte, le flot optique  $\vec{V}$  est implicitement considéré localement constant et le développement de Taylor du premier ordre de chaque ligne nous donne :

$$\begin{cases} uI_x + vI_y + I_t = 0 \\ uI_{xx} + vI_{yx} + I_{xt} = 0 \\ uI_{xy} + vI_{yy} + I_{yt} = 0 \\ uI_{xt} + vI_{yt} + I_{tt} = 0 \end{cases} \quad (99)$$

$I_x$ ,  $I_y$  et  $I_t$  représentent respectivement les gradients suivant  $x$ ,  $y$  et  $t$ . Les termes de dérivation seconde sont désignés par :  $I_{xx} = \partial I_x / \partial x$ ,  $I_{yy} = \partial I_y / \partial y$ ,  $I_{xy} = \partial I_x / \partial y$  et  $I_{yx} = \partial I_y / \partial x$ .

Toutes les méthodes qui découlent d'un tel système d'équations sont dites multicontraîntes. Nous pouvons en citer quelques unes telles que la méthode de Haralick et Lee, la méthode de Tretyak et Pastor, la méthode de Campani et Verri et la méthode de Tistarelli [HAR83, TRE84, CAM90, TIS94]. A la différence des méthodes monocontraîntes (une seule équation de contrainte), le système multicontraînt est ici surdéterminé.

L'application de la pseudo-inverse d'une matrice permet de trouver la solution sous la forme :

$$\vec{V} = -(A^T A)^{-1} A^T B \quad (100)$$

avec :

$$A = \begin{bmatrix} I_x & I_y \\ I_{xx} & I_{yx} \\ I_{xy} & I_{yy} \\ I_{xt} & I_{yt} \end{bmatrix} \quad (101)$$

et :

$$B = \begin{bmatrix} I_t \\ I_{tx} \\ I_{ty} \\ I_{tt} \end{bmatrix} \quad (102)$$

Nous pouvons remarquer une similarité entre l'équation (100) et la solution au sens des moindres carrés (90) si le produit matriciel  $(A^T A)^{-1}$  est directement calculé comme suit :

$$(A^T A)^{-1} = \begin{bmatrix} \sum (I_x^2 + I_{xx}^2 + I_{xy}^2 + I_{xt}^2) & \sum (I_x I_y + I_{xx} I_{yx} + I_{xy} I_{yy} + I_{xt} I_{yt}) \\ \sum (I_x I_y + I_{xx} I_{yx} + I_{xy} I_{yy} + I_{xt} I_{yt}) & \sum (I_y^2 + I_{yx}^2 + I_{yy}^2 + I_{yt}^2) \end{bmatrix} \quad (103)$$

et  $A^T B$  par :

$$A^T B = - \begin{bmatrix} \sum (I_x I_t + I_{xx} I_{tx} + I_{xy} I_{ty} + I_{xt} I_{tt}) \\ \sum (I_y I_t + I_{yx} I_{tx} + I_{yy} I_{ty} + I_{yt} I_{tt}) \end{bmatrix} \quad (104)$$

Si l'on n'utilise que les équations du second ordre, les équations de contrainte résultantes sont de la forme:

$$\begin{cases} uI_{xx} + vI_{yx} + I_{xt} = 0 \\ uI_{xy} + vI_{yy} + I_{yt} = 0 \\ uI_{xt} + vI_{yt} + I_{tt} = 0 \end{cases} \quad (105)$$

Ce système se résout à l'aide de la pseudo-inverse d'une matrice. Si nous supposons que les gradients spatiaux sont invariants au mouvement, c'est à dire  $\frac{\partial}{\partial t} \vec{\nabla} I(p, t) = 0$ , nous avons le système d'équations suivant :

$$\begin{cases} uI_x + vI_y + I_t = 0 \\ uI_{xx} + vI_{yx} + I_{tx} = 0 \\ uI_{xy} + vI_{yy} + I_{ty} = 0 \end{cases} \quad (106)$$

Trois équations de contrainte formant un système surdéterminé sont obtenues, et cela pour deux inconnues. Le vecteur vitesse  $\vec{V} = (u, v)^T$  peut être obtenu par la technique pseudo-inverse.

Dans certaines méthodes, l'équation de contrainte est différenciée par rapport à  $x$  et  $y$ , comme ci-dessus, mais on n'utilise que les deux dernières équations. Les méthodes différentielles du second ordre utilisent ici la matrice du Hessian de  $I$  [TRE84, CAM90, TIS94].

$$H = \begin{bmatrix} I_{xx} & I_{xy} \\ I_{yx} & I_{yy} \end{bmatrix} \quad (107)$$

Comme avec l'équation de contrainte du mouvement appliquée dans le domaine fréquentiel (voir méthodes de transformée), les méthodes multicontraîntes peuvent s'appliquer dans d'autres domaines. La forme générale du système est :

$$\begin{cases} f_{xx}u + f_{yx}v + f_{tx} = 0 \\ f_{xy}u + f_{yy}v + f_{ty} = 0 \end{cases} \quad (108)$$

avec  $f$  étant une grandeur dont le gradient est stationnaire dans le temps. C'est à dire  $\frac{\partial}{\partial t} \vec{\nabla} f = 0$ . Cette grandeur peut représenter une luminance, un contraste, une fréquence, etc.

Les méthodes multicontraîntes présentent divers avantages par rapport aux méthodes monocontraîntes. Le calcul du flot optique s'effectue avec haute résolution spatiale, vu que le nombre de contraintes se multiplie. En revanche, comme dans toute résolution du système surdéterminé ou déterminé, ces méthodes présentent des inconvénients. Si une ou quelques contraintes sont mal conditionnées, la résolution risque de présenter une mauvaise fiabilité. Une technique de conditionnement sur la matrice du Hessian peut se faire pour garantir la stabilité de son inversion [NAG86].

## 1.6 Méthodes de transformée (basée énergie et basée phase)

Ces méthodes consistent à transformer l'énergie spatio-temporelle de l'image sous une autre forme d'énergie qui résume les propriétés dynamiques [ADE85, HEE88, BIG91]. Elles sont basées sur l'observation des effets d'un mouvement 2D de l'image sur ses composantes temporelles ou sur ses composantes fréquentielles. En l'occurrence, l'énergie pour un nombre de triplets fréquence spatiale et temporelle peut être extraite par filtrage de l'énergie spatio-temporelle. Les trois principes fondamentaux des méthodes de transformée sont présentés dans les paragraphes qui suivent.

### 1.6.1 Filtrage dans le domaine spatio-temporel

Dans cette approche, l'estimation du mouvement est obtenue par filtrage dans les domaines spatiaux et temporel. Les images sont transformées à l'aide des filtres orientés spatio-temporellement. Considérons par exemple le mouvement d'un objet rectangulaire suivant [HEE96].

La figure 1.13.a représente le sens de déplacement de l'objet dans le sens spatial. Dans l'espace spatio-temporel, la propagation de l'image est obtenue en empilant une trame après l'autre. Nous avons ainsi un cube selon la figure 1.13.b. Sur la figure 1.13.c, nous voyons une découpe en tranche du cube suivant le plan  $x-t$ . L'orientation de cette coupe correspond à une composante du vecteur vitesse  $\vec{V} = (u, v)^T$ . La composante  $u$  de la vitesse est donnée par la pente de cette coupe. La composante  $v$  se calcule de la même manière.

Le mouvement peut être considéré comme une orientation dans l'espace spatio-temporel. Un filtre orienté spatio-temporellement peut être utilisé pour l'estimer [ADE86, ADE86, HEE88]. En effet, les images sont transformées par des filtres linéaires orientés spatialement et

temporellement [ADE86]. Considérons, par exemple, un filtre spatio-temporel de type gaussien. Le gradient suivant  $x$  est défini par :

$$\frac{\partial I(x, y, t)}{\partial x} = \frac{\partial}{\partial x} [g(x, y, t) * I(x, y, t)] = g_x(x, y, t) * I(x, y, t) \quad (109)$$

où  $*$  désigne le produit de convolution et  $g_x$  l'opérateur différentiel gaussien suivant  $x$ .

En faisant la même opération pour  $\partial I/\partial y$  et  $\partial I/\partial t$ , l'équation de contrainte du flot optique peut s'écrire :

$$(ug_x + vg_y + g_t) * I(x, y, t) = 0 \quad (110)$$

Les éléments  $g_x$ ,  $g_y$  et  $g_t$  sont les filtres dérivatifs spatio-temporels.

Pour une vitesse constante  $\vec{V} = (u, v)^T$ , l'équation (110) est une convolution d'un filtre spatio-temporel,  $(ug_x + vg_y + g_t)$ , avec la séquence d'images  $I(x, y, t)$ . Ce filtre est orienté spatio-temporellement et il est composé d'une somme pondérée de trois filtres dérivatifs  $g_x$ ,  $g_y$  et  $g_t$ . L'équation de contrainte (110) résume le fait qu'il y a certaines orientations pour lesquelles la réponse du filtre est nulle. Le principe est d'utiliser plusieurs filtres et par la suite, la comparaison des énergies contenues dans les images filtrées permet de déterminer les amplitudes et la direction des mouvements. L'illustration de ce principe, pour l'exemple précédent, est faite dans la figure 1.14. L'équation de contrainte se traduit ici par la figure 1.14.a qui montre que sur certaines orientations spatiales, le gradient directionnel est nul. Sur la figure 1.14.b, nous voyons le mouvement d'une portion de l'image, de gauche à droite, et le filtre spatio-temporel orienté perpendiculairement à l'orientation spatio-temporelle. La convolution des deux est nulle.

Ces techniques n'opèrent pas sur le mouvement local. Elles nécessitent de nombreux filtrages coûteux en temps de calcul. De plus, d'après la figure 1.14.b, il faut plusieurs séquences pour former l'espace spatio-temporel ainsi que le filtrage temporel. Par conséquent, cette technique a un coût élevé en terme de mémoire.

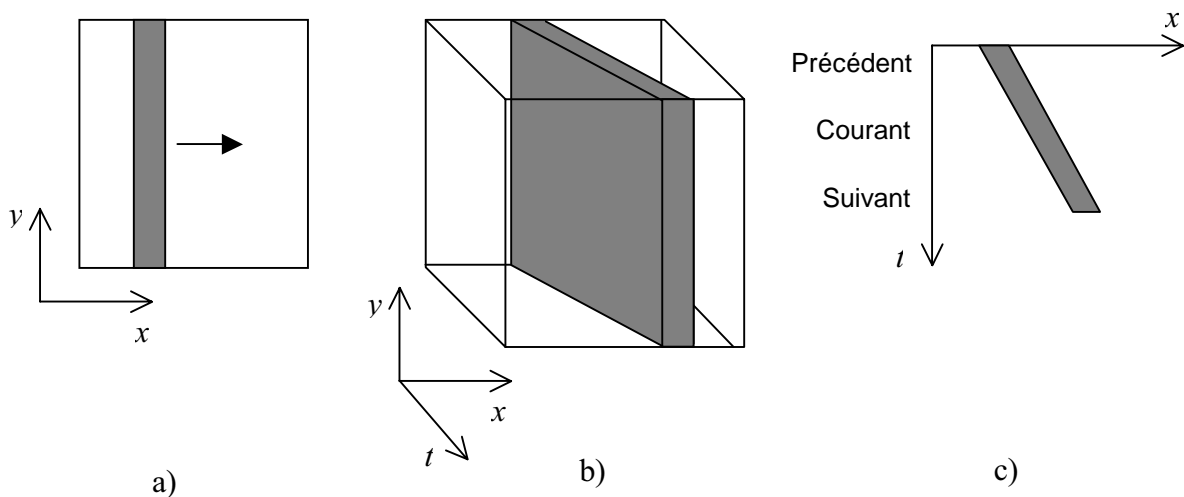


Figure 1.13 : Sens d'orientation d'un mouvement dans l'espace temporel.

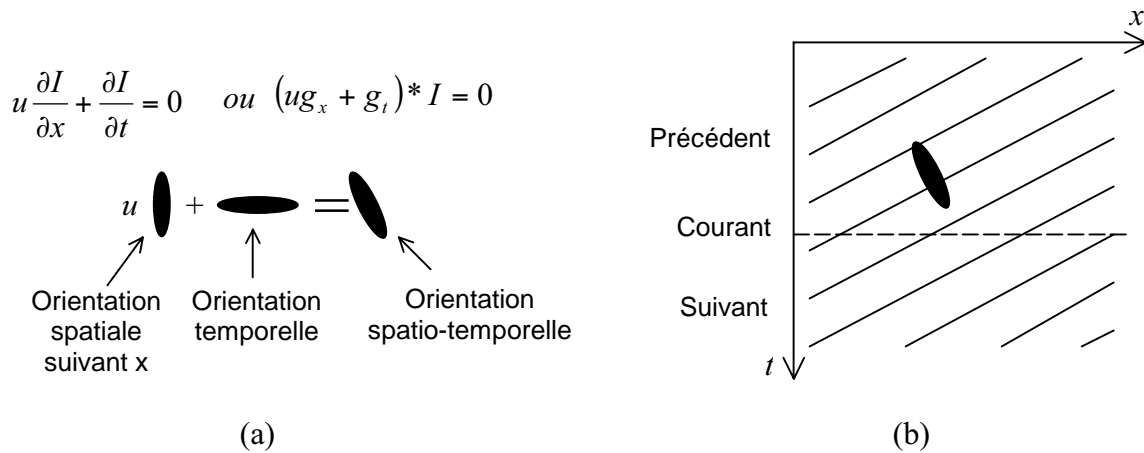


Figure 1.14 : Filtres spatio-temporels et équation de contrainte du mouvement pour un cas 1D.

### 1.6.2 Transformée de Fourier (domaine fréquentiel)

L'équation de contrainte du mouvement a aussi son interprétation dans le domaine fréquentiel. Nous verrons par la suite que dans le domaine fréquentiel, le spectre d'énergie d'un mouvement de translation 2D repose sur un plan. Une translation spatiale se traduit par une inclinaison de ce plan de spectre tridimensionnel. Le but est alors de trouver un plan sur lequel se concentre l'énergie du spectre de Fourier.

Dans un premier temps, nous allons rappeler l'équation du plan du spectre dans le domaine fréquentiel pour un modèle de translation 2D. Une expression de ce résultat peut être obtenue à première vue en considérant séparément chaque composante de la fréquence spatio-temporelle. La fréquence spatiale d'une fonction d'intensité est exprimée en nombre de cycles par pixel et sa fréquence temporelle est exprimée en cycles par trame. Plus précisément, un signal sinusoïdal 1D peut être exprimé en terme de fréquence spatiale et vitesse comme :

$$s(x, t) = s_0 \sin[2\pi f_x (x - ut)] \quad (111)$$

La quantité  $f_x$  désigne la fréquence spatiale,  $u$  la vitesse et  $s_0$  son amplitude maximale. Ce signal peut aussi être exprimé en terme de fréquence spatiale  $f_x$  et fréquence temporelle  $\omega_t = 2\pi f_t$  comme :

$$s(x, t) = s_0 \sin(2\pi f_x x + \omega_t t) \quad (112)$$

D'après (111) et (112), nous avons la relation suivante qui donne la vitesse (distance par unité de temps ou pixels par trame), en fonction de la fréquence temporelle et fréquence spatiale :

$$u = -\frac{f_t}{f_x} \quad (113)$$

Considérons maintenant un modèle spatial 1D qui a plusieurs composantes fréquentielles (image constituée par des bandes parallèles) et qui se déplace à une vitesse  $u$ . Chaque composante fréquentielle possède une fréquence spatiale  $f_x$  et une fréquence temporelle  $f_t$  mais toutes les composantes de Fourier donnent la même vitesse. Alors, chaque composante satisfait l'équation suivante :

$$f_x u + f_t = 0 \quad (114)$$

Par analogie, un modèle de translation 2D (arbitrairement texture) dans le plan image occupe un plan dans le domaine de fréquence spatio-temporelle. L'équation de ce plan est donnée par :

$$f_x u + f_y v + f_t = 0 \quad (115)$$

Ayant obtenu l'équation du plan, nous allons maintenant voir comment s'interprète l'équation de contrainte dans le domaine fréquentiel. Le principe se base sur le théorème de translation relatif à la transformée de Fourier (cf. annexe A). En effet, la transformée de Fourier d'un modèle de translation 2D tel que :

$$I(p, t) = I(p - \vec{v}t, 0) \quad (116)$$

peut s'exprimer sous la forme :

$$\hat{I}(f, f_t) = \hat{I}_0(f) \cdot \delta(\vec{v} \cdot \vec{f} + f_t) \quad (117)$$

où  $\hat{I}_0(f)$  est la transformée de Fourier 2D de  $I(p, 0)$ ,  $\delta$  la distribution de Dirac,  $f_t$  la fréquence temporelle,  $\vec{f} = (f_x, f_y)^T$  la fréquence spatiale et  $\vec{v} = (u, v)^T$  le vecteur vitesse. Cette équation définit une contrainte dans l'espace fréquentiel de sorte que toute l'énergie d'un modèle de translation réside sur un plan passant par l'origine. Plus précisément, l'équivalent de l'équation de contrainte de mouvement dans le domaine fréquentiel est donné par la relation (117), qui peut également s'écrire sous la forme :

$$\hat{I}(f_x, f_y, f_t) = \hat{I}_0(f) \cdot \delta(uf_x + vf_y + f_t) \quad (118)$$

D'après la propriété de la distribution de Dirac, l'équation (118) montre que le spectre d'énergie d'une image en translation présente des valeurs non nulles uniquement sur un plan  $P$  passant par l'origine dont l'équation est donnée par (115). L'équation (115) se traduit par un produit scalaire de deux vecteurs à 3 dimensions :  $(u, v, 1)^T$  et  $(f_x, f_y, f_t)^T$ . Seules les valeurs non nulles de  $\hat{I}_0(f)$  sont concentrées sur le plan orthogonal à  $(u, v, 1)^T$ . Sur la figure 1.15, on voit le spectre d'énergie d'une image en translation qui représente un plan  $(P)$  passant par l'origine. Cette translation  $(u, v, 1)^T$  est un vecteur orthogonal au plan  $(P)$ . L'estimation du flot optique par les méthodes fréquentielles consiste donc à déterminer le vecteur vitesse  $(u, v, 1)^T$  orthogonal au plan  $(P)$  ajustant au mieux le spectre d'énergie [PAU99]. L'équation (115) est l'équivalent de l'équation de contrainte du flot optique dans le domaine fréquentiel.

L'extension de cette technique aux modèles complexes est décrite dans [DUC94]. La généralisation de l'hypothèse de conservation de la luminance pour un modèle quelconque se traduit par l'équation (38), de la forme :  $\partial I(p, t) / \partial t = O$ . Cette équation génère une équation différentielle correspondant à une transformation donnée. En appliquant la transformée de Fourier à cette relation, nous aurons un résultat analogue au cas particulier précédent. Pour un modèle à  $n$  paramètres par exemple, l'opération de dérivation conduit à un produit scalaire de deux vecteurs à  $n$ -dimensions. Le premier vecteur contient les paramètres du mouvement. Le second est concentré sur un hyperplan, qui passe par l'origine et est orthogonal au vecteur des paramètres du mouvement.

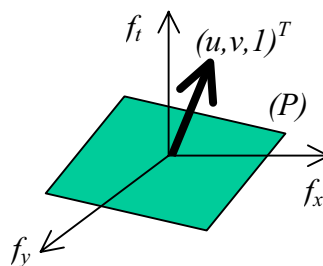


Figure 1.15 : Plan  $(P)$  formé par le spectre d'énergie d'une image en translation et le vecteur de déplacement associé  $(u, v, 1)^T \perp (P)$ .



### 1.6.3 Mesure de la variation de phase

On peut démontrer que la phase de la transformée d'un signal contient les informations essentielles [JAH97]. Une image peut encore être reconnue quand l'information sur l'amplitude est perdue, mais pas quand la phase est perdue. Une illumination globale modifie l'amplitude du signal mais pas sa phase.

La présente méthode s'appuie sur l'hypothèse de la conservation de phase. La transformation consiste à chercher la réponse en terme de phase pour les fréquences spatio-temporelles. Les images sont transformées par des filtres passe-bande et la phase sera donnée par les réponses du filtre [FLE90]. Une translation spatiale correspond à un décalage de phase entre deux images successives. La mesure de la variation de phase permet de déduire le mouvement de translation.

Pour mieux illustrer le principe d'estimation du mouvement basée sur la phase, nous considérons une onde plane 1D qui est exprimée en terme de phase par la relation :

$$s(x, t) = s_0 \exp[-i\phi(x, t)] \quad (119)$$

La quantité  $s_0$  est l'amplitude du signal  $s$  et  $\phi(x, t)$  sa phase. Par analogie avec l'équation (111), nous avons :

$$s(x, t) = s_0 \exp[-i2\pi f_x(x - ut)] \quad (120)$$

L'équation (120) montre que la phase d'une onde plane varie d'une façon linéaire dans le temps et dans l'espace, soit :

$$s(x, t) = s_0 \exp[-i(kx + \omega_t t)] \quad (121)$$

Les éléments  $k$  et  $\omega_t$  sont respectivement le nombre d'ondes et la pulsation. Comme dans la relation (113), ces deux grandeurs sont reliées à la vitesse, suivant la relation :

$$u = -\frac{\omega_t}{k} \quad (122)$$

Le calcul des dérivées temporelles et spatiales de la phase produit le nombre d'ondes et la fréquence temporelle de la structure périodique du mouvement :

$$\nabla_{xt}\phi = \begin{bmatrix} \frac{\partial \phi}{\partial x} \\ \frac{\partial \phi}{\partial t} \end{bmatrix} = \begin{bmatrix} k \\ \omega_t \end{bmatrix} \quad (123)$$

La vitesse est alors donnée par :

$$u = -\frac{\frac{\partial \phi}{\partial t}}{\frac{\partial \phi}{\partial x}} \quad (124)$$

Cette formule ressemble à l'estimation basée sur l'équation spatio-temporelle du flot optique. En raison des discontinuités, il n'est pas recommandé de calculer directement les dérivées partielles à partir de la fonction de phase. En revanche, il est possible d'estimer les gradients spatio-temporels à partir de la sortie d'une paire de filtres quadratiques [JAH97]. Si nous désignons la paire de filtres quadratiques par  $q_+(x, t)$  et  $q_-(x, t)$ , les gradients spatio-temporels de la phase sont exprimés par :

$$\nabla_{xt}\phi(x, t) = \frac{q_+(x, t)\nabla_{xt}q_-(x, t) - q_-(x, t)\nabla_{xt}q_+(x, t)}{q_+^2(x, t) + q_-^2(x, t)} \quad (125)$$

A partir des équations (123, 124, 125), l'expression de la vitesse basée sur la mesure de phase est donnée par :

$$u = \frac{q_+(x,t) \frac{\partial q_-(x,t)}{\partial t} - q_-(x,t) \frac{\partial q_+(x,t)}{\partial t}}{q_+(x,t) \frac{\partial q_-(x,t)}{\partial x} - q_-(x,t) \frac{\partial q_+(x,t)}{\partial x}} \quad (126)$$

### 1.6.4 Quelques méthodes de transformée

#### a) Méthode de Heeger

Heeger utilise les filtres de Gabor (cf. annexe A) dans le domaine spatio-temporel. Dans cette méthode le problème de flot optique est formulé comme étant un ajustement au sens des moindres carrés du spectre d'énergie d'une image en utilisant une famille de Filtres de Gabor. Pour chaque résolution spatiale, Heeger échantillonne le spectre local avec un banc de 12 filtres de Gabor, accordés sur différentes orientations spatio-temporelles [HEE88]. Pour un mouvement de translation, les réponses de ces filtres  $F_i$ ,  $1 \leq i \leq 12$ , sont concentrées sur une surface plane, située dans un espace fréquentiel. Chaque filtre donne une réponse non linéaire qui est fonction de  $u$  et  $v$ . Les solutions au sens des moindres carrés pour  $u$  et  $v$  sont obtenues en utilisant une méthode d'optimisation. Le principe est le suivant :

- On établit l'expression théorique du spectre d'énergie  $R(u, v, f_x, f_y, f_t)$  à partir de la convolution d'un filtre de Gabor accordé à la fréquence donnée  $(f_x, f_y, f_t)^T$  par un bruit blanc subissant une translation connue  $(u, v)^T$  [HEE87] :

$$R(u, v, f_x, f_y, f_t) = \frac{k}{8\pi\sqrt{\alpha}} \exp\left[\frac{-4\pi\sigma_x^2\sigma_y^2\sigma_t^2(uf_x + vf_y + f_t)^2}{\alpha}\right] \quad (127)$$

avec :

$$\alpha = (u\sigma_x\sigma_t)^2 + (v\sigma_y\sigma_t)^2 + (\sigma_x\sigma_y)^2 \quad (128)$$

Les éléments  $\sigma_x$ ,  $\sigma_y$  et  $\sigma_t$  désignent les écart-types de la fenêtre spatio-temporelle gaussienne du filtre et  $k$  une constante de normalisation (énergie due à l'effet de contraste de l'image locale).

- Soit  $M_i$  le spectre d'énergie expérimental, issu de la convolution de chaque filtre de Gabor  $F_i$ ,  $1 \leq i \leq 12$ , par l'image à traiter.

- On compare la mesure expérimentale et la mesure théorique. Cette comparaison consiste en la minimisation d'un critère défini par la différence de l'ensemble des mesures expérimentales et l'ensemble des valeurs théoriques :

$$J = \sum_{i=1}^{12} \left[ M_i - \bar{M}_i \frac{R_i(v_x, v_y, f_x^i, f_y^i, f_t^i)}{\bar{R}_i(v_x, v_y)} \right]^2 \quad (129)$$

avec  $\bar{M}_i = \sum_j M_j$  et  $\bar{R}_i = \sum_j R_j$  tel que  $j$  désigne les filtres ayant la même orientation spatiale.

Ces deux termes servent de facteur de normalisation pour éviter les problèmes liés aux images dans lesquelles le contraste varie en fonction de la direction.

Pour déterminer le vecteur vitesse, de nombreuses méthodes sont applicables à l'équation (129) [BAR92, PEL96].

#### b) Méthode de Fleet et Jepson

Cette méthode est basée sur la mesure de phase. Fleet et Jepson [FLE90, BAR94] utilisent aussi des filtres de Gabor, mais ils ont considéré la phase des réponses complexes, tandis que Heeger a considéré leurs modules (c'est-à-dire le spectre d'énergie locale). L'évolution

temporelle des contours de phase constante est utilisée pour définir le mouvement de translation de l'image. On détermine la composante du vecteur vitesse instantanée, normale au contour de phase, à l'aide de la réponse des filtres passe bande. La réponse complexe de chaque filtre peut s'écrire ainsi :

$$R(p, t) = \rho(p, t)e^{i\phi(p, t)} \quad (130)$$

Les termes  $\rho(p, t)$  et  $\phi(p, t)$  désignent respectivement l'amplitude et la phase de  $R$ .

Dans la méthode différentielle, nous avons vu que l'équation de contrainte du mouvement se déduit de l'hypothèse de conservation de la luminance. Par analogie, en considérant que nous avons la conservation de phase dans le temps, l'équation de contrainte dans le domaine de phase est donnée par :

$$\vec{\nabla}\phi(p, t)\vec{V} + \frac{\partial\phi(p, t)}{\partial t} = 0 \quad (131)$$

où  $p = (x, y)$  et  $\vec{\nabla}\phi(p, t) = [\phi_x(p, t), \phi_y(p, t)]^T$ .

Le vecteur vitesse dans la direction normale au contour de phase est donné par :

$$\vec{V}_n = -\frac{\frac{\partial\phi(p, t)}{\partial t}}{\|\vec{\nabla}\phi(p, t)\|} \vec{n} \quad (132)$$

L'élément  $\vec{n}$  désigne le vecteur normal au gradient dont l'expression est :

$$\vec{n} = \frac{\vec{\nabla}\phi(p, t)}{\|\vec{\nabla}\phi(p, t)\|} \quad (133)$$

L'opération de dérivation sur la phase est décrite de la manière suivante :

$$\vec{\nabla}\phi(p, t) = \frac{\text{Im}[R^*(p, t)\vec{\nabla}R(p, t)]}{|R(p, t)|^2} \quad (134)$$

La quantité  $R^*$  désigne le complexe conjugué de  $R$ ,  $\vec{\nabla}R$  le gradient de  $R$ . La notation  $Im$  désigne la partie imaginaire du nombre complexe.

Pour le cas des scènes 3D dans lesquelles le mouvement de translation subit de faibles déviations, il peut être démontré que les réponses en phase d'un filtre passe-bande sont plus stables que celles en amplitude. L'utilisation de la technique basée sur les phases peut ainsi être efficace [FLE93, BAR94]. En revanche, l'instabilité des pixels voisins pourrait rendre la phase instable (singularité de phase). Un critère de stabilité, qui est fonction de la fréquence instantanée et de l'amplitude de la réponse d'un filtre, est défini de la manière suivante [FLE92] :

$$\|\vec{\nabla}\log R(p, t) - i(f_x, f_y, f_t)\| \leq \sigma_k \tau \quad (135)$$

Le triplet  $(f_x, f_y, f_t)$  désigne la fréquence spatio-temporelle d'accord du filtre,  $\sigma_k$  l'écart-type du spectre d'amplitude isotrope du filtre et  $\tau$  un seuil qui peut être utilisé pour rejeter les composantes peu fiables du vecteur vitesse.

## 1.7 Méthodes corrélatives

Dans les méthode différentielles, le calcul des gradients souffre d'une sensibilité aux bruits. Les résultats peuvent ainsi être biaisés. Pour avoir des résultats corrects, il faut de bons estimateurs de dérivée. La présente solution consiste à utiliser les méthodes basées sur la mise en correspondance des blocs. L'image est divisée en plusieurs blocs de  $m \times n$  pixels. Le déplacement est défini comme étant le décalage spatial  $\vec{d} = (dx, dy)^T$  qui produit le meilleur appariement entre les blocs d'images aux instants successifs. Trouver la meilleure

correspondance revient à maximiser une fonction de corrélation sur la valeur de  $\vec{d}$ , ou à minimiser une mesure de distance. Cette corrélation et cette mesure de distance peuvent avoir plusieurs formes :

### 1.7.1 Différentes formes de l'indicateur de corrélation

On peut chercher le sommet dans la fonction de corrélation croisée. Cette fonction peut avoir plusieurs formes :

- La corrélation directe :

$$F_d(x, y, t) = \sum_{j=-n/2}^{n/2} \sum_{i=-m/2}^{m/2} g[I(x+i, y+j, t), I(x+dx+i, y+dy+j, t+dt)] \quad (136)$$

$$= W(x, y) * g[I(x, y, t), I(x+dx, y+dy, t+dt)]$$

Le point  $(x, y)$  est un point arbitraire du bloc de l'image de référence,  $I(x+i, y+j, t)$  son niveau de gris et  $I(x+dx+i, y+dy+j, t+dt)$  est celui du bloc candidat,  $g$  présente la fonction de corrélation entre les deux niveaux de gris. Par exemple, la corrélation peut être mesurée par la covariance si nous pouvons ignorer le fait que les erreurs, dans les estimés de la valeur de  $I$  dans un voisinage de points, ne sont pas indépendantes.  $W$  désigne la fonction de fenêtre 2D discrète. Cette approche est utilisable dans le cadre de l'hypothèse de conservation de la luminance.

- La corrélation normalisée par rapport aux valeurs moyennes :

$$F_m(x, y, d) = \sum_{j=-n/2}^{n/2} \sum_{i=-m/2}^{m/2} g \left\{ I(x+i, y+j, t) - \bar{I}(x+i, y+j, t), \right. \quad (137)$$

$$\left. [I(x+dx+i, y+dy+j, t+dt) - \bar{I}(x+dx+i, y+dy+j, t+dt)] \right\}$$

La quantité  $\bar{I}$  correspond à la valeur moyenne de la fonction  $I$  calculée sur la fenêtre  $m \times n$ . Cette approche peut être applicable dans le cas où l'intensité varie au cours du temps.

- La corrélation normalisée utilisant les écart-types :

$$F_o(x, y, d) = \sum_{j=-n/2}^{n/2} \sum_{i=-m/2}^{m/2} \frac{1}{\sigma_{I(x,y,t)} \sigma_{I(x+dx,y+dy,t+dt)}} g \left\{ I(x+i, y+j, t) - \bar{I}(x+i, y+j, t), \right. \quad (138)$$

$$\left. [I(x+dx+i, y+dy+j, t+dt) - \bar{I}(x+dx+i, y+dy+j, t+dt)] \right\}$$

Le coefficient  $\sigma_{I(x,y,t)}$  représente l'écart-type de  $I$  sur la fenêtre  $m \times n$ . Cette approche est applicable dans le cas où le gain du capteur varie spatialement dans l'image. Ainsi, on utilise l'écart type pour compenser la non linéarité de la fonction de transfert du capteur entre le flux lumineux incident et le niveau de gris du capteur.

### 1.7.2 Principales mesures de distance

De nombreuses méthodes de mesure de distance sont envisagées [PUM87, MUS85, NAM96] :

- La différence absolue moyenne (DAM) :

$$DAM(x, y, d) = \sum_{j=-n/2}^{n/2} \sum_{i=-m/2}^{m/2} |I(x+i, y+j, t) - I(x+dx+i, y+dy+j, t+dt)| \quad (139)$$

- L'erreur quadratique moyenne (EQM) :

$$EQM(x, y, d) = \frac{1}{mn} \sum_{j=-n/2}^{n/2} \sum_{i=-m/2}^{m/2} [I(x+i, y+j, t) - I(x+dx+i, y+dy+j, t+dt)]^2 \quad (140)$$

- L'erreur quadratique moyenne normalisée :

$$EQMN(x, y, d) = \frac{\sum_{j=-n/2}^{n/2} \sum_{i=-m/2}^{m/2} [I(x+i, y+j, t) - I(x+dx+i, y+dy+j, t+dt)]^2}{\sum_{j=-n/2}^{n/2} \sum_{i=-m/2}^{m/2} I(x+i, y+j, t)^2} \quad (141)$$

A titre d'exemple, considérons la figure 1.16 qui présente un processus de mise en correspondance de blocs et l'indexation de données [NAM96]. Prenons comme critère de recherche la minimisation de la distance DAM, définie à l'équation (139). Soit :

$$DAM(x, y, d) = \sum_{j=1}^n \sum_{i=1}^m |I(x+i, y+j, t) - I(x+dx+i, y+dy+j, t+dt)| \quad (142)$$

avec :

$$\begin{cases} -a \leq dx < +a \\ -b \leq dy < +b \end{cases} \quad (143)$$

La recherche ne s'effectue pas depuis le premier pixel de l'image mais plutôt depuis le premier pixel d'une zone d'intérêt, qui s'appelle fenêtre de recherche. Si l'équation (143) définit les bornes de calcul, nous aurons une fenêtre de taille  $2a \times 2b$ . Les quantités  $m$  et  $n$  désignent respectivement la hauteur et la largeur du bloc de référence,  $I(x, y, t)$  la luminance du pixel arbitraire du bloc de référence et  $I(x+dx, y+dy, t+dt)$  celle de l'un des blocs candidats, situés dans la fenêtre de recherche, avec un déplacement de  $\vec{d} = (dx, dy)^T$ .

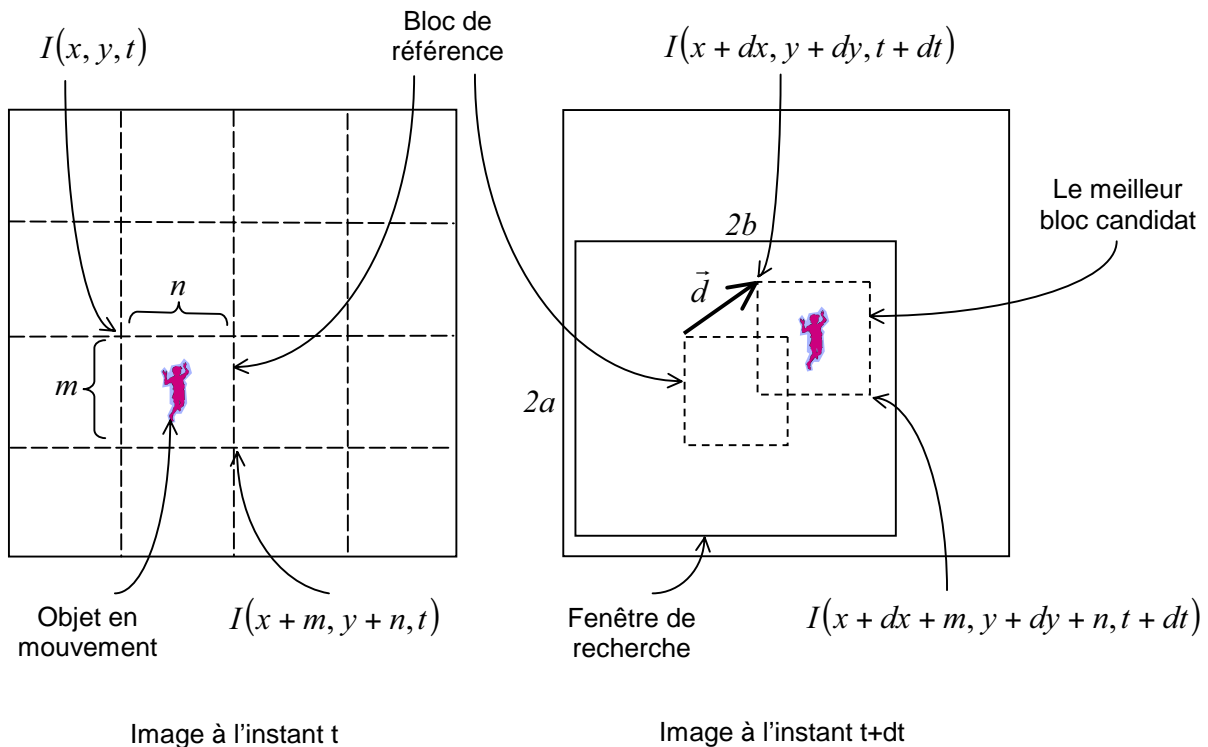


Figure 1.16 : Processus de mise en correspondance de blocs et indexation de données.

## 1.8 Autres problèmes d'estimation du flot optique

### 1.8.1 Problème d'occlusion

Par définition, le champ de déplacement établit une correspondance entre les points de deux ou plusieurs trames. En effet, il est difficile d'estimer correctement le déplacement des zones d'occlusion car il n'existe pas de zone de correspondance d'une trame à l'autre.

Dans [DEP92], une méthode qui consiste à ajouter un champ d'occlusion est présentée. Ce champ entre en interaction à la fois avec le champ de discontinuité et le modèle structurel. Le but est d'annuler l'hypothèse d'invariance et de l'intensité dans les zones occultées. Par une modélisation semblable au champ de discontinuité, une fonction de potentiel, un système de voisinage et ses cliques ont été définis heuristiquement, afin de favoriser la création des zones d'occlusion au voisinage des discontinuités de mouvement. Driessen et Biemond [DRI91] ont présenté dans un travail indépendant un modèle très semblable à celui de [DEP92]. Heitz et Bouthemy [HEI90] modélisent aussi le champ de discontinuité par le champ markovien pour tenir compte des discontinuités de mouvement. Mais une nouvelle technique a été proposée pour le traitement des occlusions ; elle utilise les gradients spatio-temporels et les informations obtenues d'un détecteur de contours photométriques dans un test d'hypothèse afin de détecter les régions d'occlusion où la contrainte spatio-temporelle n'est pas valide. Le champ de déplacement y subira un lissage spécial à partir de l'information des voisins.

Bien que des méthodes résolvant le problème d'occlusion aient été proposées [ADE90, IRA94], nous resterons sur les techniques qui ne tiennent pas compte du cas d'occlusion.

### 1.8.2 Problème d'objet transparent

Chaque pixel de l'image doit correspondre à un point de la surface des objets de la scène. Dans une scène tridimensionnelle qui contient un objet transparent, cette condition n'est pas vérifiée. Il est difficile d'estimer le déplacement des objets transparents.

## 1.9 Résumé du chapitre

Nous avons rappelé dans ce chapitre quelques notions sur la théorie de l'estimation du mouvement. Nous avons notamment présenté deux familles d'estimation : les approches basées primitives et les approches basées flot optique. Au sein de ces deux familles, la notion du flot optique a été développée plus profondément. Plusieurs techniques d'estimation du flot optique ont donc été présentées. On peut observer que certaines techniques opèrent dans le domaine spatio-temporel, d'autres dans le domaine fréquentiel.

Dans le domaine spatio-temporel, nous avons rappelé l'équation de contrainte spatio-temporelle du flot optique qui relie les composantes du vecteur vitesse aux gradients spatio-temporels. Les approches qui découlent de cette équation s'appellent méthodes différentielles. Ces dernières ont été développées davantage dans ce premier chapitre. Par la suite, nous ne retiendrons que cette technique en raison de sa facilité de mise en œuvre et parce que l'on peut facilement modéliser le déplacement. En outre, comme nous envisageons une application qui nécessite un nombre d'images restreint (deux images), cela justifie également le choix de la technique différentielle car elle peut s'appuyer uniquement sur deux trames consécutives.

Différents problèmes liés à l'estimation du flot optique ont été également présentés. Ces problèmes montrent l'insuffisance de contrainte dans le critère à minimiser. Par conséquent, l'intervention d'une ou plusieurs techniques de régularisation est nécessaire. Nous avons présenté diverses techniques de régularisation.



# Chapitre 2

## Segmentation d'une image

### 2.1 Introduction

Le but de l'opération est de segmenter une image en régions homogènes (au sens des niveaux de gris, du mouvement, de la texture, etc.).

Cette partition peut être obtenue en s'appuyant sur :

- La détection des bords des différentes régions : il s'agit d'une segmentation basée contours (exemple : Snake) ;
- La détection de tous les pixels de chacune des régions : il s'agit de la segmentation basée régions (exemple : Champ de Markov).

Dans cette étude, nous allons nous intéresser tout particulièrement à la segmentation orientée régions.

Dans le chapitre précédent, nous avons souvent parlé de la notion de région. Une séquence d'images peut contenir des mouvements désordonnés, définis au niveau du pixel, ou plusieurs régions ayant chacune un mouvement propre. A titre d'illustration, si nous nous intéressons au déplacement d'une facette, comme la totalité de l'image ne sera pas concernée par le déplacement, il est intéressant de limiter l'étude au support de la facette. Ce support spatial engendre alors une région. En multipliant le nombre de régions, le problème majeur est de déterminer dans une séquence les différents modèles (associés à chaque région) et leur support spatial.

Afin de pouvoir travailler avec la notion de région, le concept doit être défini au sens mathématique. Autrement dit, un modèle doit être utilisé pour décrire une région. Ce chapitre est consacré aux différentes techniques de segmentation et à la description du modèle associé.

Dans le cadre de la segmentation spatio-temporelle, nous pouvons constater un lien implicite entre l'estimation et la segmentation. D'une part, comme l'estimation du mouvement dépend du support spatial, une bonne segmentation doit être faite au préalable pour pouvoir estimer les paramètres du modèle de mouvement. D'autre part, comme les objets mobiles sont définis à partir d'une représentation du mouvement, une bonne estimation est exigée pour donner une bonne segmentation. Ce point revient au dilemme de 'poule & œuf'. Les deux considérations sont envisageables.

Généralement, la définition d'un objet ne dépend pas uniquement de l'information du mouvement mais aussi des caractéristiques spatiales. En l'occurrence, les contours d'intensité coïncident avec les contours de mouvement. On peut bien envisager une segmentation basée intensité pour l'utiliser par la suite comme une segmentation initiale au sens du mouvement. En revanche, les objets peuvent être séquentiellement extraits en déterminant successivement les mouvements dominants. Cette considération réside en une hypothèse de cohérence temporelle.

Face aux exigences de la segmentation spatio-temporelle, une simple utilisation des ressources disponibles dans l'image ne suffit pas, il faut aussi prendre en compte la structure de la scène pour satisfaire un ensemble de contraintes. Ces dernières sont introduites par 'la connaissance a priori', terme utilisé dans le contexte de la segmentation.

Pour avoir une bonne segmentation, les conditions suivantes doivent être satisfaites :

- La segmentation doit être complète (c'est-à-dire chaque pixel doit être dans une classe) ;
- Les pixels appartenant à la même région doivent être connexes ;
- Les régions doivent être disjointes.



Dans les deux paragraphes suivants, nous décrivons le principe des approches descendante et ascendante.

## 2.2 Segmentation spatio-temporelle selon l'approche descendante

Dans cette méthode, les objets sont extraits séquentiellement en déterminant itérativement les mouvements dominants successifs. L'approche consiste d'abord à manipuler l'image entière comme support d'estimation pour obtenir un mouvement dominant. Il est probable que plusieurs mouvements soient présents dans ce support mais l'estimation fournit un modèle global de mouvement. Les pixels qui ne sont pas conformes au mouvement global sont détectés, afin de former deux classes de partition. Un modèle de mouvement peut être de nouveau estimé sur les pixels non conformes et une autre classe sera trouvée, et ainsi de suite.

La méthode descendante est confrontée à deux principaux problèmes. Ce sont notamment le problème généralisé d'ouverture [JEP93, BLA93] et le procédé de détection/élimination des données erronées. Le premier problème est lié à l'estimation du mouvement dominant. Le mouvement dominant doit en effet être estimé correctement malgré le fait que de multiples mouvements peuvent coexister dans le support d'estimation. En présence d'une telle difficulté, la procédure d'estimation du mouvement peut fournir un déplacement imprécis. En supposant que l'évaluation de mouvement a été correctement exécutée, la segmentation d'objet se fonde sur le traitement de détection/élimination. La sélection des points erronés doit tenir compte des ambiguïtés ou des erreurs inhérentes au type d'information utilisé.

En ce qui concerne le procédé d'évaluation du mouvement dominant, Burt [BUR91] préconise que le procédé soit rendu plus robuste en utilisant une décomposition pyramidale. Cette décomposition permet de classer indépendamment les différents mouvements existant dans la scène. En présence du mouvement dominant, le procédé d'évaluation se concentre sur ce mouvement, tout en ignorant l'entrée des autres. L'approche pyramidale a été employée par les techniques descendantes proposées dans [BUR89, BER90, IRA92]. Afin d'améliorer le procédé d'évaluation, Rousseeuw et de Leroy [ROU87] préconisent l'utilisation des estimateurs résistants. Ce sont généralement les estimateurs robustes. Ils ont la capacité de dériver le mouvement dominant sans être influencé par les autres mouvements [AYE93a, ODO94, SCH94].

Le deuxième problème est le paradigme de détection/élimination des données non significatives. Ce processus est responsable de la détermination des zones qui se déplacent en conformité avec le mouvement dominant et celles qui ne le sont pas. Les pixels en déplacement selon un mouvement dominant sont groupés en un objet simple. Le défi consiste à déterminer quels pixels possèdent cette caractéristique. À cet effet, la plupart des techniques se fondent seulement sur l'information temporelle pour accomplir cette tâche [PEL90, IRA92, AYE93a, ODO94]. La trame précédente est déformée en utilisant le mouvement dominant estimé et le résultat est comparé à la trame courante. Les points aberrants sont définis comme une zone correspondant à de grandes erreurs de prédiction. En revanche, l'utilisation de l'information temporelle peut s'avérer être incertaine au niveau du pixel. C'est particulièrement le cas dans des régions à faible gradient. Pour surmonter ce problème, quelques techniques intègrent les erreurs résiduelles sur de larges régions spatio-temporelles afin de rendre la détection plus robuste [AYE93b]. En outre, l'utilisation des informations spatiales permet de définir avec précision les contours des objets [ODO94, DUC95, GU96a]. Pour résumer, l'approche descendante présente certains avantages. Elle permet à l'utilisateur d'arrêter l'opération quand un nombre prédéfini d'objets est atteint. Cela peut être utilisé dans des applications où le nombre d'objets qui composent la scène est connu. Cette approche peut aussi être applicable dans la segmentation où on s'intéresse uniquement aux régions associées aux déplacements les plus importants [IRA92]. Dans la méthode descendante, les algorithmes sont caractérisés par leur simplicité et leur coût de calcul faible. En revanche, cette approche

présente différents inconvénients. La segmentation peut être biaisée quand l'hypothèse du mouvement dominant dans la scène est fautive. Du point de vue algorithmique, les méthodes descendantes ne traitent pas les différents mouvements d'une manière équivalente. Chaque mouvement est défini sur la base des mouvements dominants précédemment extraits. Cette approche séquentielle conduit à une hiérarchie parmi les différents mouvements et parmi les différents objets. Finalement, une partition complète n'est pas assurée. L'extraction successive des mouvements dominants peut conduire à une situation où les points isolés ne sont représentatifs d'aucun objet (points aberrants). En conséquence, aucune représentation d'image basée sur la segmentation spatio-temporelle ne pourrait être obtenue correctement.

### 2.3 Segmentation spatio-temporelle selon l'approche ascendante

Cette approche se base sur la stratégie de fusion de régions. Dans une première phase, un ensemble de régions initiales est extrait. L'approche ascendante consiste à fusionner les objets pour obtenir des objets mobiles. La fusion est effectuée selon une règle de décision. Cette dernière peut être définie par une mesure de similarité spatio-temporelle. On suppose que chaque objet peut être naturellement défini à travers la relation spatio-temporelle entre les régions qui le composent. Avec l'approche ascendante, les différents objets peuvent être extraits d'une manière simultanée. C'est en contraste avec l'approche descendante dans laquelle les objets sont extraits séquentiellement.

L'approche ascendante peut être décomposée en trois étapes :

- Choix de l'ensemble des régions initiales ;
- Définition de la mesure de similarité ;
- Stratégie de fusion de données.

#### 2.3.1 Génération de l'ensemble des régions initiales

Il existe plusieurs solutions pour sélectionner l'ensemble des régions initiales. Les régions peuvent être arbitrairement définies suite à une segmentation par quadtree. Cette opération se base uniquement sur l'information temporelle. Nous allons décrire très brièvement son principe. C'est une segmentation issue d'une partition récursive de l'image en quatre blocs rectangulaires. Au départ, l'image est divisée en blocs qui forment l'ensemble des régions initiales et leurs paramètres du mouvement sont estimés (figure 1.17). Les blocs dont la DFD est supérieure à un seuil sont divisés en quatre blocs égaux et ainsi de suite. L'opération s'arrête quand la limite de la taille d'un bloc est atteinte. En raison de sa simplicité de représentation, cette procédure de partition est souvent utilisée en segmentation d'image, puisqu'il suffit de donner les différentes phases de découpes utilisées pour définir la segmentation [SAM80, NIC92, GEL96].

En revanche, ces régions n'incarnent pas la teneur de l'image courante. En particulier, leur délimitation spatiale ne tient pas compte de l'information spatiale actuelle dans la scène. C'est-à-dire que la fusion de ces régions ne permet pas forcément d'avoir des régions cohérentes à l'échelle spatio-temporelle. Par conséquent, les régions initiales devraient être, autant que possible, homogènes au sens du mouvement.

Pour pallier ce problème, deux considérations sont envisageables :

\*La première consiste à considérer un cas limite où chaque pixel est défini comme région initiale [BOU 93, ADI 85]. Comme le pixel est l'élément le plus élémentaire d'une image, son mouvement est cohérent dans une échelle spatio-temporelle. Le problème est que les pixels sont des éléments manipulés résultant de la numérisation d'une image et n'ont intrinsèquement aucune signification. Or, l'ensemble des régions initiales obéit à la propriété de connexité d'objet. C'est-à-dire que chaque objet est formé par des zones dont les pixels

sont naturellement liés les uns aux autres. Nous risquons d'obtenir l'isolement de n'importe quel pixel ou d'une petite zone.

\*D'autres techniques proposent une définition des régions initiales qui sont visuellement significatives et spatio-temporellement homogènes. Plusieurs critères peuvent être utilisés pour l'homogénéité spatio-temporelle :

- Les régions peuvent être classées selon leur homogénéité spatiale. Dans ce cadre, nous pouvons procéder au seuillage du niveau de gris [HOR86]. Les pixels ayant un niveau de gris supérieur à un seuil sont mis à "1" sur la position correspondante de l'image binaire et ceux qui sont en dessous sont mis à "0".

- Dans [AGG81, SAL95], les régions sont classées selon leur homogénéité spatiale, à travers une segmentation statique. Le groupement initial est ainsi caractérisé par l'homogénéité de texture ou de contours. L'information temporelle peut aussi être utilisée. En l'occurrence, dans [WAN94], l'ensemble des régions spatiales est obtenu à l'aide d'une segmentation arbitraire par quadtree qui s'appuie en même temps sur l'information mouvement pour l'affiner. Dans [DUF95], le groupement initial est généré à partir des informations spatiale et temporelle. A partir d'une segmentation initiale, une région sur laquelle le mouvement est mal compensé sera partagée en deux. Ce partage s'effectue à l'aide d'une segmentation statique.

Une remarque générale peut être faite quand on considère les techniques proposées pour générer des régions initiales plus larges qu'un pixel. Généralement parlant, l'extraction de régions se base sur une approche descendante [BOU87]. A partir d'un premier estimé de régions, elles sont itérativement fractionnées si elles ne sont pas conformes au critère de cohérence. Les régions résultantes définissent une sur-segmentation de la scène à partir de laquelle les objets sont construits. C'est pourquoi l'approche ascendante est dite "approche fusionner – diviser".

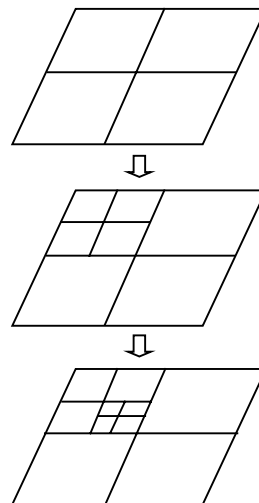


Figure 1.17 : Segmentation quadtree d'une image.

### 2.3.2 Mesure de similarité de régions

Plusieurs techniques ont été élaborées pour identifier les régions homogènes dans une image ou pour détecter la frontière entre les régions. Le critère d'homogénéité se réfère à un certain nombre d'attributs, tels que l'intensité, la couleur et la texture [LEC89].

La mesure de similarité définit le degré d'affinité spatio-temporelle existant entre deux ou plusieurs régions. Selon cette mesure, la vraisemblance que les régions fassent partie d'un seul objet peut être évaluée. La mesure de similarité fournit les informations sur lesquelles se

base la procédure de fusion de régions. La définition de la mesure de similarité dépend considérablement du type d'information qui est utilisé. Par exemple, la mesure de similarité peut seulement se baser sur l'information spatiale [SAL94, PAR94, SAL95]. Les objets résultants montrent une cohérence spatiale. Ils sont également susceptibles d'être homogènes au sens du mouvement [MAR95]. Cependant, dans un même objet, les régions avec des caractéristiques spatiales très différentes peuvent coexister. Par exemple, quand nous observons la tête d'une personne, la peau du visage est spatialement très différente des cheveux, bien qu'ils puissent se déplacer de la même manière. A partir de cet exemple, il apparaît clairement qu'une mesure de similarité basée uniquement sur l'information spatiale aura comme conséquence une sur-segmentation des objets de la scène. Pour éviter cette imperfection, la mesure de similarité doit utiliser l'information temporelle. Ceci exige une représentation de mouvement qui, pour une région donnée, permet une comparaison entre différents mouvements. Dans ce contexte, l'approche paramétrique est privilégiée [ANA93]. L'information de mouvement est représentée par un ensemble de paramètres (selon 1.3). Un modèle peut s'étendre du modèle de translation au modèle polynomial.

En évaluant la similarité entre régions, la représentation paramétrique de mouvement a été utilisée dans plusieurs techniques. A titre d'exemple, nous allons donner deux classes d'approche. La première utilise seulement l'information temporelle disponible dans l'espace de paramètres du mouvement. Les méthodes de Dufaux et autres [DUF95], et Wang et Adelson [WAN94a] définissent la mesure de similarité de régions à partir d'un critère de distance entre les ensembles de paramètres du mouvement. Cette approche est évidemment sensible à l'erreur d'estimation du mouvement. De plus, selon la scène et le modèle de mouvement choisis, des flots optiques semblables peuvent être représentés par des ensembles très différents de paramètres du mouvement [ADI89]. En d'autres termes, le paramétrage du mouvement peut ne pas avoir une seule solution et l'hypothèse qu'un jeu de paramètres représente l'information entière de mouvement peut en effet être erronée. Ceci implique que deux régions qui se déplacent d'une même manière peuvent avoir des paramètres du mouvement très différents.

Une autre alternative sur l'utilisation du paramètre temporel est basée sur des statistiques des distributions résiduelles obtenues par la compensation de mouvement. En comparant la région  $A$  à la région  $B$ , le mouvement du dernier est appliqué à l'ancien. La distribution résiduelle résultante permet d'évaluer comment la région  $A$  se déplace de la même manière que la région  $B$ . Par sa nature, la distribution résiduelle résultante est représentative de l'erreur existant entre le mouvement paramétrique et le vrai déplacement. L'information dans la distribution résiduelle peut être utilisée dans une analyse statistique. Par exemple, la mesure de similarité proposée par Adiv [ADI85] est la variance de la distribution résiduelle. Deux régions sont fusionnées si la variance de la région nouvellement formée est semblable à celle des régions individuelles avant la fusion. Cependant, l'utilisation d'une statistique simple produit aussi une réduction de la dimension de distribution. Cela peut conduire à une mauvaise décision de fusion puisque toutes les informations présentes dans la distribution ne sont pas examinées. A titre d'exemple, l'information temporelle qui est à la fois présente dans la distribution résiduelle et dans la représentation paramétrique, peut être utilisée. Ceci est réalisé dans le cadre du formalisme MDL<sup>4</sup>, technique basée sur la minimisation du coût de description globale [RIS78]. C'est un paradigme qui dérive directement des mots d'Occam au quatorzième siècle : "les entités ne doivent pas être multipliées au-delà de la nécessité". La mesure de similarité entre deux régions est définie comme la réduction idéale de codage. En d'autres termes, la similarité représente le nombre de bits qui serait sauvegardé si les deux régions étaient fusionnées.

---

<sup>4</sup> Minimum description length

La définition de la similarité de régions est un problème principal dans la segmentation d'une image. Comme commentaire général, on peut indiquer que toutes les informations disponibles doivent être mises à contribution afin de définir d'une manière robuste les objets présents dans la scène. La mesure de similarité doit exploiter les informations spatiale et temporelle [THO80, ADI85, GU96b]. En particulier, Thompson propose de fusionner les régions à base de critère de contraste, modifié par l'information temporelle [THO80]. Cependant, les deux types d'information ne sont pas combinés dans une seule mesure de similarité. Ils sont utilisés séparément, en s'appuyant plus particulièrement sur l'information spatiale. En conclusion, nous soulignons que le but de la segmentation est d'obtenir une représentation sémantique de la scène.

### 2.3.3 Stratégie de fusion de données

La mesure de similarité de régions permet d'évaluer si deux régions ou plus font partie du même objet. Afin de mettre en évidence le présent procédé de fusion, une stratégie de fusion de région doit être définie. Elle est responsable de l'obtention de la meilleure utilisation de l'information fournie par la mesure de similarité. Cette stratégie devrait également être robuste par rapport à l'erreur et l'ambiguïté qui peuvent apparaître dans la définition de la similarité de régions. En conclusion, le volume des calculs directs doit demeurer raisonnable. Deux types de stratégies peuvent être définis. Le premier type utilise un processus en deux étapes [WAN94b, AYE95a]. Les mouvements décrivant la scène sont d'abord déterminés, puis les objets sont déduits en fusionnant les régions correspondant à chaque mouvement. Le deuxième type de stratégie de fusion détermine simultanément les mouvements et les objets [BOU93].

Le premier type de stratégie de fusion est effectué en deux étapes. Dans la première étape, le nombre de mouvements décrivant la scène est déterminé. Dans la deuxième étape, ceux-ci sont employés pour déterminer quelles régions devraient être fusionnées ensemble.

Dans [WAN94b], chaque région est assignée à des paramètres du mouvement affines. Ces derniers sont calculés par une méthode régressive de genre moindres carrés pondérés. Un algorithme de groupement de K-moyens<sup>5</sup> [KAU90] est alors appliqué dans l'espace de paramètres du mouvement. Par conséquent, cet aspect fournit un ensemble de représentations de modèles de mouvement dans la scène. Cet ensemble est employé pour fusionner les régions qui couvrent le même objet mobile. Un autre exemple représentatif d'une stratégie à deux étapes est proposé par Ayer et Sawhney [AYE95a]. Ils proposent un critère MDL qui vise à réduire au minimum le coût de l'ensemble de mouvements. Le critère est constitué de deux parties, le coût de codage des paramètres du mouvement et le coût de codage de la DFD correspondante. Basé sur cela, un algorithme itératif de "expectation-maximisation" (EM) dérive l'ensemble des mouvements optimums. Après avoir déterminé ces mouvements, les régions sont fusionnées en conséquence. Ceci est exécuté en se basant sur l'évaluation de maximum a posteriori (MAP).

A la différence des techniques à deux étapes décrites ci-dessus, Bouthemy et François [BOU93] utilisent une stratégie de fusion qui détermine simultanément l'ensemble de mouvements et la fusion de régions. La segmentation est exprimée comme un problème de relaxation basé sur la modélisation par champ markovien. En raison de l'équivalence entre MRF et distribution de Gibbs [GEM84], le modèle ramène le problème à une minimisation de fonctions des voisinages locaux. Par conséquent, la probabilité  $P$  pour que le système soit dans un état particulier est donnée par :

---

<sup>5</sup> K-mean clustering

$$P = \frac{1}{Z} e^{-U/T} \quad (144)$$

Le dénominateur  $Z$  est une constante de normalisation (également connue sous le nom de fonction de partition),  $T$  est une constante (correspondant à la température en mécanique statistique), et  $U$  est une fonction d'énergie (ou fonction objective) qui peut être minimisée par une méthode de relaxation déterministe. L'algorithme K-moyens peut aussi être utilisé pour la classification des objets. L'approche markovienne et l'algorithme K-moyens seront exploités dans la présente étude pour la segmentation spatio-temporelle. Les théories concernant ces deux approches ainsi que d'autres outils mathématiques à la segmentation sont présentés en annexe B.

## 2.4 Résumé du chapitre

Dans ce chapitre, nous avons présenté deux stratégies principales pour la segmentation d'une image. La première s'appelle approche descendante. Dans cette approche, les objets sont extraits séquentiellement par une détermination successive des mouvements dominants. Les pixels relatifs au mouvement dominant courant sont supposés appartenir à un objet et classés comme des points significatifs. Les autres pixels sont classés comme des points aberrants. Ceux-ci sont pris en compte dans la prochaine itération afin d'estimer le mouvement dominant suivant et les nouveaux points significatifs.

La seconde s'appelle approche ascendante. Dans cette méthode, les objets sont extraits d'une manière simultanée. Dans une première phase, un ensemble de régions initiales sont définies. Ces régions sont ensuite fusionnées en utilisant une règle de décision. On peut prendre ici la mesure de similarité spatio-temporelle comme critère de décision.



## Conclusion de l'état de l'art et orientation de l'étude

Cette partie fait un rappel de quelques notions sur l'estimation et la segmentation du mouvement. Les méthodes d'estimation basées primitives et flot optique ont été évoquées.

Pour les approches basées flot optique, cadre d'étude de ce manuscrit, nous avons principalement retenu quelques notions de la théorie d'estimation :

- L'équation différentielle du mouvement qui est la base de toute méthode différentielle. Dans le domaine temporel, cette équation relie les composantes du vecteur vitesse aux gradients spatio-temporels. Elle s'appelle ainsi équation de contrainte spatio-temporelle du flot optique ou tout simplement équation de contrainte du flot optique.

- La composante normale du vecteur vitesse, le premier élément évident qui peut être déterminé à partir de l'équation du flot optique.

- Les méthodes non paramétriques qui consistent à estimer le mouvement pixel par pixel. L'équation du flot optique peut être utilisée dans ces méthodes pour définir le critère à minimiser par une méthode d'optimisation. Nous sommes ainsi dans le cadre des méthodes différentielles et non paramétriques.

- Les méthodes paramétriques dans lesquelles le mouvement d'une région est décrit par un modèle. Ces approches peuvent aussi être fondées sur la minimisation d'un critère qui est fondé sur l'équation différentielle du mouvement.

- La notion de régularisation qui permet de renforcer les contraintes de l'équation du mouvement.

- Les méthodes à contraintes globales ou à contraintes locales. Tout dépend si la minimisation du critère s'effectue sur la totalité de l'image ou uniquement sur un voisinage local (un support spatial de faible taille).

- Les méthodes de transformée qui opèrent dans le domaine fréquentiel ou dans le domaine des phases.

- Les méthodes corrélatives.

Par la suite, nous ne retiendrons que la technique différentielle à contrainte locale en raison de son optimalité, sa facilité de mise en œuvre, et parce que l'on peut facilement modéliser le déplacement. En outre, comme l'approche différentielle peut s'appuyer même sur deux trames seulement, nous l'utilisons comme technique de base pour développer un algorithme destiné aux applications ayant besoin d'un nombre restreint d'images. L'objectif est d'élaborer une approche performante sans augmenter le coût de calcul, ni en terme de mémoire, ni en terme de rapidité.

Le but recherché est de renforcer les contraintes de régularisation pour rendre les estimés plus robustes. Certaines approches, évoquées dans notre étude, s'appuient également sur l'utilisation du modèle de mouvement.

Pour la segmentation, nous avons globalement présenté deux familles. La segmentation selon l'approche descendante et celle selon l'approche ascendante.

Dans le cadre de la segmentation spatio-temporelle, l'approche descendante consiste à extraire successivement les objets à l'aide des mouvements dominants. Les pixels associés au mouvement dominant courant sont classés comme un objet. Les autres seront considérés dans l'itération suivante pour estimer à nouveau le prochain mouvement dominant. Une règle de décision doit être prise pour la détection et le rejet des points aberrants. A ce titre, on peut utiliser une erreur de prédiction. Certaines techniques utilisent d'autres mesures telles que le résidu sur une large région spatio-temporelle.

Pour l'approche descendante, les objets sont segmentés d'une manière simultanée. On part d'un ensemble de régions initiales, générées par une méthode quelconque telle que la segmentation par quadtree. Ensuite, une règle de décision est définie pour prendre en charge



la mesure de similarité. Finalement, une stratégie de fusion est utilisée pour procéder à la segmentation.

Nous pouvons constater un lien implicite entre l'estimation et la segmentation. C'est la raison pour laquelle notre étude sur la segmentation spatio-temporelle se base fortement sur l'estimation du mouvement. Les techniques de segmentation que nous allons présenter sont une dichotomie entre les approches descendante et ascendante. Les méthodes que nous proposons sont fondées sur une modélisation stochastique de type champ aléatoire de Markov et sur des approches issues de la classification utilisant l'algorithme de type K-moyens. Ces algorithmes vont être associés à l'estimation du mouvement utilisant un formalisme qui minimise une erreur quadratique.

## **Deuxième partie**

### **Estimation du flot optique basée sur un schéma direct-rétrograde**



# Chapitre 1

## Schéma direct-rétrograde pour un mouvement de translation

### 1.1 Introduction

Dans ce chapitre, nous allons nous intéresser à l'extension de l'approche différentielle, basée sur l'équation du flot optique. Notre approche s'appuie sur la prise en compte du mouvement direct et du mouvement rétrograde d'un objet dans une séquence d'images. L'idée principale de l'approche exploite une cohérence du mouvement suivant le sens causal et anti-causal, dans la zone d'estimation. Comme nous pouvons l'observer sur la figure 2.1, pour une même position spatiale, nous prolongeons la modélisation du mouvement. L'introduction de cette nouvelle équation de contrainte se justifie d'autant plus que la phase d'estimation augmente la validité de la modélisation à travers le support spatial pour l'estimation.

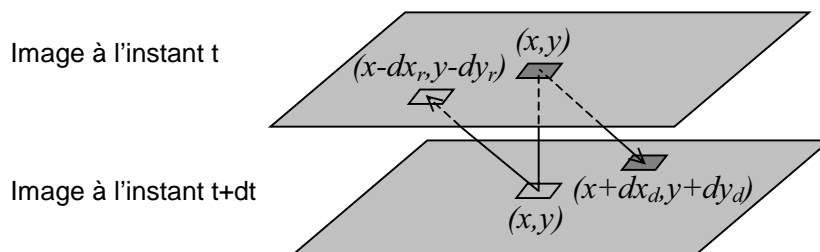


Figure 2.1 : Mouvement suivant les sens causal et anti-causal, formalisme direct-rétrograde.

Dans ce chapitre, pour un modèle de translation, nous allons étudier différentes méthodes d'estimation liées au formalisme proposé. Les composantes du vecteur déplacement sont obtenues de manière statistique (moindres carrés) à partir de l'information locale (méthode à contrainte locale). L'extension est assez simple à gérer dans des approches de type robuste ou incrémentale [BLA92, BUR84].

Pour les modèles de mouvement plus complexes, points traités dans le second chapitre, nous étudions l'extension des différentes méthodes de résolution. Nous montrons notamment que cette extension génère des équations non linéaires. Après une linéarisation simple, mais qui n'est valide que dans le cas de faibles paramètres, nous introduisons des approches incrémentales qui prennent en compte le terme non-linéaire.

### 1.2 Rappel sur la méthode des moindres carrés<sup>6</sup>, application à l'estimation du mouvement basée sur un schéma direct

#### 1.2.1 Survol sur la formulation

D'après la littérature [HOR81], le processus d'estimation du flot optique s'articule généralement autour d'une hypothèse de conservation de la luminance lors du déplacement (voir première partie, chapitre 1, paragraphe 1.4.2). Cette hypothèse nous permet de générer un critère basé sur la minimisation de l'erreur suivante :

$$\varepsilon = I(x, y, t) - I(x + dx, y + dy, t + dt) \quad (145)$$

<sup>6</sup> LS : Least squares

Le terme  $I(x, y, t)$  désigne la luminance à une position spatiale  $(x, y)$  et à l'instant  $t$ . Nous considérons ici un modèle de translation dans le sens direct du temps, noté  $(dx, dy)$ . La figure 2.2 illustre le déplacement d'un point entre les instants  $t$  et  $t + dt$ .

Afin de mettre en œuvre une méthode d'optimisation, de nombreuses méthodes s'appuient sur la fonction d'erreur (145) pour définir un critère à minimiser. Dans [STI99], une fonctionnelle basée sur le principe classique de cohérence du mouvement dans une fenêtre d'observation  $\Omega$ , est définie par :

$$J = \sum_{\Omega} \varepsilon^2 \quad (146)$$

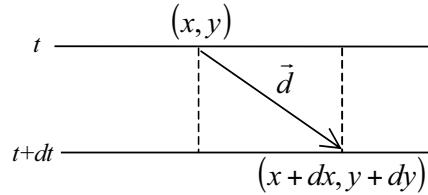


Figure 2.2 : Déplacement direct d'un pixel entre deux instants.

En faisant le développement de Taylor au premier ordre de l'équation (145), nous obtenons l'équation de contrainte du mouvement :

$$\frac{\partial I(x, y, t)}{\partial x} dx + \frac{\partial I(x, y, t)}{\partial y} dy + \frac{\partial I(x, y, t)}{\partial t} dt = 0 \quad (147)$$

Par la suite, les dérivées partielles spatiales et temporelles à une position  $(x, y)$  sont notées :

$$I'_x = \frac{\partial I(x, y, t)}{\partial x}, \quad I'_y = \frac{\partial I(x, y, t)}{\partial y} \quad \text{et} \quad I'_t = \frac{\partial I(x, y, t)}{\partial t}; \quad \text{il s'agit d'une notation discrète.}$$

Les composantes du vecteur déplacement sont obtenues de manière statistique à partir de l'information locale. Il s'agit de la méthode des moindres carrés dont font partie les méthodes à contrainte locale. Ainsi, le positionnement en discret du problème consiste à considérer que, localement sur une fenêtre  $\Omega$  centrée sur le pixel, le mouvement est uniforme. Les  $n$  pixels de  $\Omega$  doivent vérifier l'équation (147) avec le même couple  $(dx, dy)$ . Le principe consiste alors à chercher le vecteur  $(u, v)^T$  qui minimise le critère suivant :

$$J = \sum_{i=1}^n (u I'_x + v I'_y + I'_t)^2 \quad (148)$$

Les quantités  $u = dx/dt$  et  $v = dy/dt$  sont les composantes du vecteur vitesse. La minimisation de ce critère (cf. annexe A) aboutit à l'équation matricielle suivante :

$$(u, v)^T = -(A^T A)^{-1} A^T B \quad (149)$$

La matrice  $A$  correspond au parcours sur les gradients spatiaux et  $B$  la matrice des gradients temporels, avec :

$$A = \begin{bmatrix} I'_x(1) & I'_y(1) \\ \vdots & \vdots \\ I'_x(n) & I'_y(n) \end{bmatrix} \quad (150)$$

et :

$$B = \begin{bmatrix} I'_t(1) \\ \vdots \\ I'_t(n) \end{bmatrix} \quad (151)$$

Selon l'équation (149), nous pouvons noter que l'estimation du flot optique est conditionnée par l'approximation choisie pour l'estimation des dérivées partielles. En annexe C, nous présentons quelques caractéristiques engendrant une détérioration des performances des méthodes différentielles.

### 1.2.2 Résultats expérimentaux

L'analyse des résultats de simulation se déroule en deux étapes. Nous appliquons, en premier lieu, la méthode de résolution sur une séquence synthétique dont nous contrôlons tous les paramètres. Le deuxième test est effectué sur une image réelle à déplacement synthétique. Nous tenons à préciser que nous désignons parfois la méthode des moindres carrés basée sur un schéma simple par méthode des moindres carrés simple ou classique.

#### a) Séquence de synthèse

Afin de générer une forme différentiable, nous avons utilisé une spirale décrite par l'équation:

$$I_i(x,y) = P \cdot \left[ \arctg \left( \frac{x - \frac{m}{2}}{y - \frac{n}{2}} + K \right) + \frac{\pi}{2} \sqrt{\left(x - \frac{m}{2}\right)^2 + \left(y - \frac{n}{2}\right)^2} \right] \quad (152)$$

où  $m$  et  $n$  définissent la taille de l'image. Le paramètre  $K$  permet de faire varier la forme de la spirale obtenue. Le coefficient  $P$  est le facteur de proportionnalité qui permet de régler le niveau de gris. Une trame de la séquence "spirale" est représentée à la figure 2.3. Un bruit blanc gaussien peut être ajouté à la séquence. Il s'agit du bruit additif. Nous appliquons à cette séquence une translation de  $(-2,1)$  pixels .

Pour l'approximation des gradients spatiaux et temporels, nous avons testé plusieurs types de filtres en prenant la convention décrite en annexe A. Lors de ces tests, nous n'avons pas constaté une grosse différence entre les résultats des filtres de Sobel et ceux des filtres de Prewitt. Dans les expérimentations qui suivent, nous utilisons les masques SOBEL pour les séquences qui nécessitent très peu de filtrages et les masques MDIF pour celles qui exigent davantage de filtrages spatiaux. Le filtre temporel MOYEN\_t est pris pour les applications nécessitant un bon filtrage temporel.

Le flot optique estimé sur la spirale et le flot réel sont représentés sur la figure 2.4.

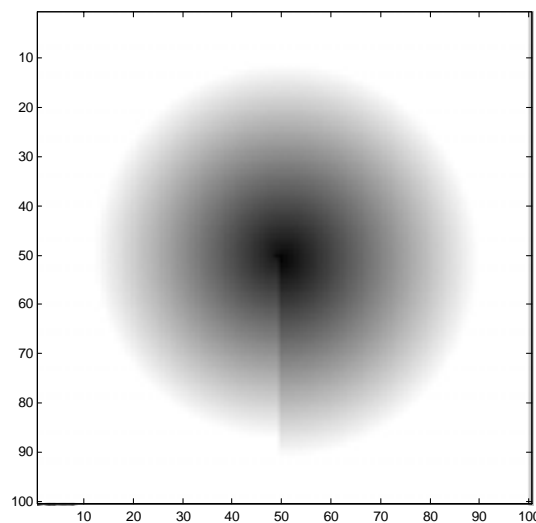


Figure 2.3 : Une trame de la séquence de synthèse "spirale".

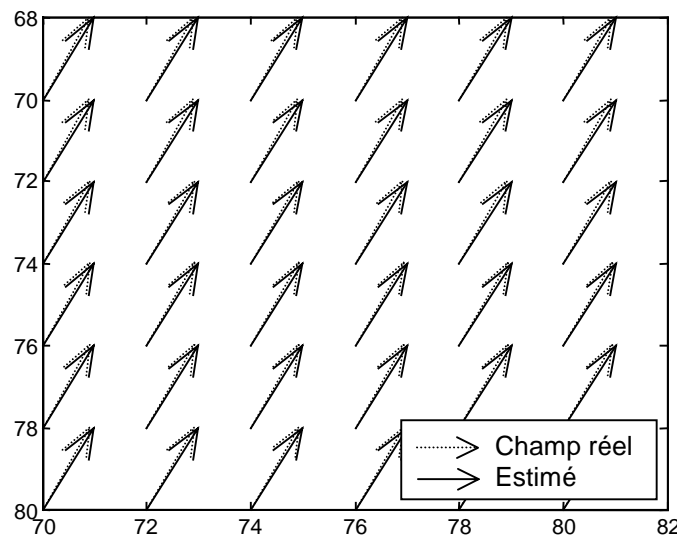


Figure 2.4 : Résultat sur la séquence "spirale" ; comparaison entre le champ de déplacement estimé et le champ réel, pour la méthode des moindres carrés simple (basée sur un schéma direct) ; séquence bruitée par un bruit gaussien additif ( $SNR = 80 \text{ dB}$ ).

La figure 2.4 montre un exemple de champ de mouvement de translation de la séquence spirale. La séquence est bruitée par un bruit additif gaussien avec un  $SNR = 80 \text{ dB}$ . Le support d'estimation est un carré de  $13 \times 13$  pixels. Les gradients spatiaux sont approchés à l'aide des masques SOBEL (cf. annexe A). Le masque MOYEN\_t est utilisé pour les gradients temporels. Nous rappelons que le choix de ce masque se justifie par le fait que notre algorithme est élaboré pour une application ayant besoin d'un nombre restreint d'images. L'estimation du déplacement ne peut s'appuyer que sur deux trames consécutives. Cette contrainte explique la raison de la limitation sur la forme discrète des gradients spatiaux et temporels à des masques simples et non récursifs.

Sur cette figure, nous remarquons la présence d'un décalage entre l'estimé et le champ réel. Ce décalage a différentes origines que nous présentons en annexe C.

Nous analysons maintenant l'effet de la taille du support de calcul, utilisé dans la résolution au sens des moindres carrés, sur les résultats d'estimation. La figure 2.5 montre les résultats obtenus sur un point quelconque de la spirale dans le cas où nous faisons croître la taille de la fenêtre de calcul  $\Omega$ .

Comme nous travaillons sur une image d'équation bien définie, les dérivées partielles peuvent être approchées correctement. Par conséquent, le système d'équations peut être bien conditionné, même seulement avec un faible nombre de points (support de calcul de faible taille). D'après la figure 2.5, la précision est atteinte à partir d'un support de  $5 \times 5$  pixels.

Dans la plupart des tests, nous utilisons une fenêtre de  $13 \times 13$  pixels. Nous verrons en annexe C que l'important n'est pas tant d'avoir un support très grand mais plutôt d'avoir un support adapté au conditionnement du système à résoudre.

En revanche, la taille du support de calcul peut avoir une influence sur le choix du modèle de mouvement ; en effet plus la région considérée est petite, moins les paramètres linéaires (rotation, divergence, déformation) seront significatifs par rapport aux paramètres constants (translation). En particulier, quand une région se réduit à un pixel, les paramètres linéaires n'ont plus aucun sens car le déplacement d'un point est décrit essentiellement par un vecteur de translation. Donc, pour un modèle paramétrique, plus on utilise une grande région, plus il sera nécessaire d'utiliser un modèle complexe.

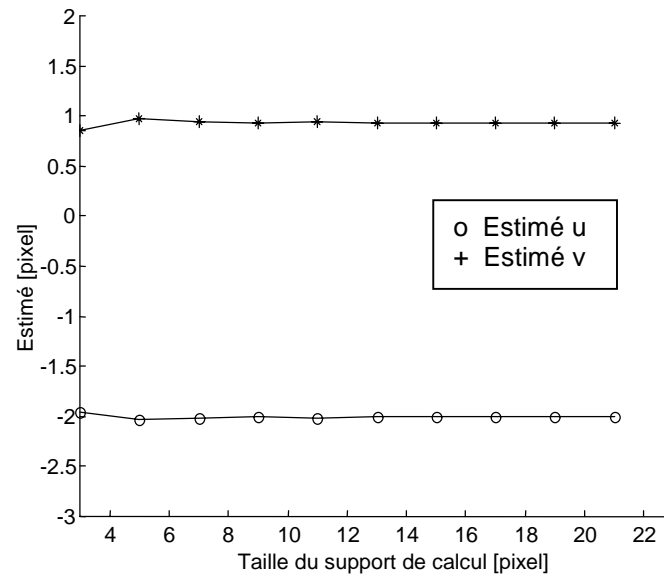


Figure 2.5 : Influence de la taille de la fenêtre de calcul sur la valeur de l'estimation.

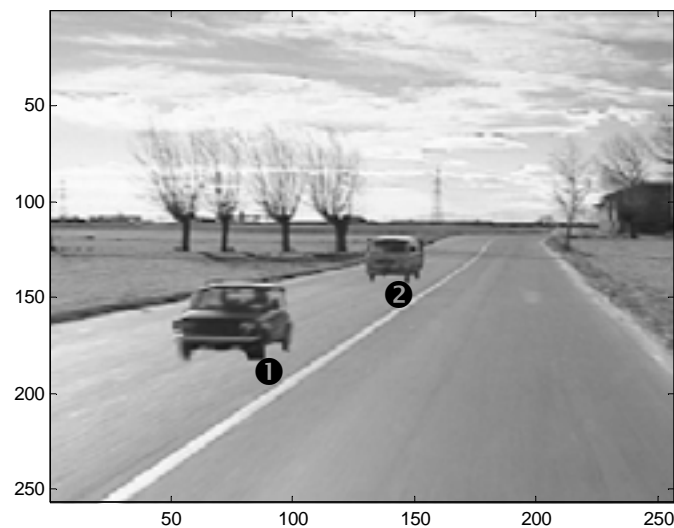


Figure 2.6 : Une trame de la séquence réelle à déplacement synthétique.

### b) Séquence réelle à déplacement synthétique

La deuxième expérimentation est effectuée sur une séquence réelle dont les objets ont été déplacés de façon synthétique. Les objets sont deux véhicules, numérotés 1 et 2, dont les valeurs des déplacements sont respectivement égales à  $(2, -1)$  et  $(5, -3)$  pixels. Les gradients spatiaux sont estimés par convolution du masque MDIF (cf. annexe A) avec la matrice des niveaux de gris. Le gradient temporel est obtenu par une différence temporelle, moyennée par le masque MOYEN<sub>t</sub>. Le support de calcul utilisé par la résolution au sens des moindres carrés est égal à  $13 \times 13$  pixels. L'image est représentée à la figure 2.6. La séquence est bruitée par un bruit additif gaussien avec un  $SNR = 100 \text{ dB}$ . La figure 2.7 illustre la détection du mouvement de deux objets. Une analyse fine au sein du champ de déplacement est effectuée à la figure 2.8.



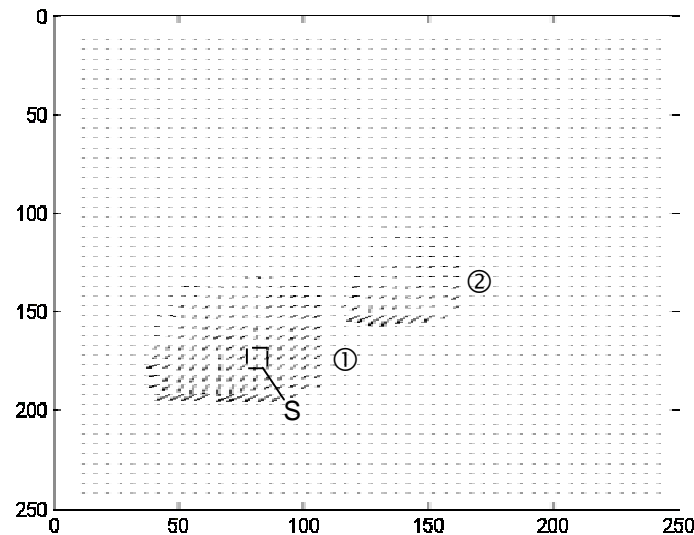


Figure 2.7 : Estimation du mouvement des objets ① et ② de la séquence réelle à déplacement synthétique ; méthode des moindres carrés basée sur un schéma simple ; séquence bruitée par un bruit gaussien additif avec ( $SNR = 100 \text{ dB}$ ).

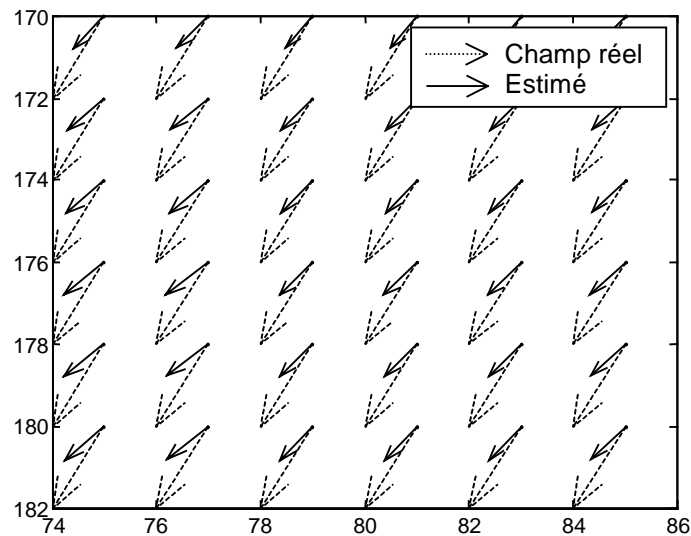


Figure 2.8 : Agrandissement du champ estimé de la zone S.

L'utilisation de cette séquence réelle à déplacement synthétique permet de voir la limite de validité de l'approche des moindres carrés basée sur un schéma simple. D'après la figure 2.8, le biais est plus fort, comparé à celui obtenu avec l'image synthétique. Le bruit est la première cause de cette dégradation. L'utilisation de l'approximation des gradients discrets dans l'algorithme d'estimation engendre des résultats médiocres. L'utilisation de l'information uniquement dans le sens direct ne suffit pas pour avoir de bons résultats.

### 1.2.3 Discussion sur la méthode des moindres carrés basée sur un schéma simple

Plusieurs tests ont été effectués dans le cadre du mouvement de translation. Une image synthétique et une image réelle ont été utilisées. Dans les deux cas, nous remarquons la présence de biais. Plusieurs causes sont à l'origine du décalage entre le champ réel et l'estimé.

Cette analyse est présentée en annexe C. L'influence de la taille du support de calcul sur les résultats a été analysée. Sur une image synthétique, le système commence à être bien conditionné, seulement avec un très faible nombre de points. Sur une image quasi-réelle, nous avons remarqué que les résultats commencent à se stabiliser à partir d'un support de calcul plus grand. Le fait d'augmenter infiniment ce support ne résout pas le problème car la stabilité dépend de la proportion entre les points aberrants et le nombre total des points du support. Une analyse sur ce point est présentée en annexe C. Différentes sources sont à l'origine d'une erreur d'estimation. Dans le cadre des méthodes différentielles, nous pouvons citer les principales sources suivantes :

- Erreur due à l'estimation des gradients ;
- Erreur due à l'hypothèse de cohérence ;
- Erreur de propagation caractéristique du système linéaire (par exemple la résolution au sens des moindres carrés) ;
- Erreur due au mauvais choix du support de calcul.

Une étude analytique de ces différentes sources d'erreur est aussi présentée en annexe C.

Lors des expérimentations, nous avons testé différents masques des gradients (masques présentés en annexe A) mais le fait de choisir un masque plutôt qu'un autre n'entraîne pas en lui-même de gain significatif.

### 1.3 Méthode du tenseur : méthode des moindres carrés totaux<sup>7</sup>

#### 1.3.1 Principes généraux

Dans la méthode du tenseur, on suppose que les deux membres de l'équation linéaire sont responsables des erreurs. C'est à dire que l'équation  $A\vec{V} \approx -B$  s'exprime sous la forme :

$$(A + \Delta A)\vec{V} = -(B + \Delta B) \quad (153)$$

Les quantités  $\Delta A$  et  $\Delta B$  présentent les erreurs entre les données théoriques, notées  $\hat{A}$  et  $\hat{B}$ , et les données approchées  $A$  et  $B$ . Le terme  $\vec{V}$  désigne le vecteur vitesse à estimer.

$$[\Delta A \ \Delta B] = [\hat{A} \ \hat{B}] - [A \ B] \quad (154)$$

Selon la relation (155), l'approche TLS consiste à trouver une solution en minimisant la norme de Frobenius des erreurs totales  $\|[\hat{A} \ \hat{B}] - [A \ B]\|_F$ .

$$\min \| \Delta A \ \Delta B \|_F \text{ tel que } (B + \Delta B) \in (A + \Delta A) \quad (155)$$

Si  $[A \ B]$  résout ce problème, tout vecteur  $\vec{V}$  vérifiant  $A\vec{V} = -B$  est appelé solution de la méthode du tenseur. La minimisation des erreurs totales peut s'effectuer en minimisant la fonctionnelle suivante :

$$\min(J) = \min_{\Delta A, \Delta B, \vec{V}} \| [\Delta A \ \Delta B] \|_F^2 \quad (156)$$

tel que :

$$(A + \Delta A)\vec{V} = -(B + \Delta B) \quad (157)$$

L'indice  $F$  désigne la norme de Frobenius. La matrice des erreurs  $[\Delta A \ \Delta B]$  a la même structure que la matrice des données  $[A \ B]$ . La relation (156) peut être reformulée ainsi :

$$\min(J) = \min_{\Delta A, \Delta B, \vec{V}} \left\{ \text{tr} \left( [A \ B]^T [A \ B] \right) \right\} \quad (158)$$

Dans un contexte d'estimation du mouvement, les expressions de  $A$  et  $B$  sont données par les équations (150) et (151). La procédure standard de résolution du problème du tenseur

<sup>7</sup> TLS : Total least squares

implique une décomposition en éléments propres. La décomposition en éléments propres [TOM98] de la matrice des données est décrite par l'expression suivante :

$$[A \ B]^T [A \ B] = U \Sigma_{3 \times 3} U^T \quad (159)$$

Dans le cadre d'un modèle de translation,  $\Sigma$  est une matrice  $3 \times 3$  des valeurs propres non nulles. Son expression est de la forme :

$$\Sigma = \text{diag}(\lambda_1, \lambda_2, \lambda_3) \quad (160)$$

La matrice  $U$  est constituée par les vecteurs propres  $\vec{e}_j, j = 1 \dots 3$ , associés aux valeurs propres  $\lambda_j$ . D'après la technique de résolution TLS, le vecteur qui minimise la fonctionnelle  $J$  peut être déterminé selon l'expression suivante :

$$\vec{v} = -\frac{1}{e_k(3)} \begin{bmatrix} e_k(1) \\ e_k(2) \end{bmatrix} \quad (161)$$

où  $\{k \leq 3, \lambda_k = \min(\lambda_1, \lambda_2, \lambda_3)\}$ . Les éléments  $e_k(i), i = 1 \dots 3$  correspondent aux composantes du vecteur propre associé à l'élément propre plus petit.

L'algorithme proposé nous indique que l'obtention d'un résultat fiable nécessite une surdétermination du système d'équations. Aussi, afin d'être moins contraints par les hypothèses de validité, nous pouvons introduire des contraintes additionnelles et un schéma itératif ; c'est l'objectif de notre étude. Avant d'évoquer les autres algorithmes permettant d'améliorer les résultats, nous présentons les résultats d'expérimentation associés à l'approche du tenseur

### 1.3.2 Résultats expérimentaux

Dans cette expérimentation, un déplacement de  $(-2,1)$  pixels est affecté à la séquence spirale. La séquence est bruitée par un bruit blanc gaussien additif de  $80 \text{ db}$ . Les gradients spatiaux sont approchés en utilisant le masque SOBEL (cf. annexe A). Quant au gradient temporel, il est estimé en appliquant le masque MOYEN\_t à la différence temporelle. Le support d'estimation utilisé est de taille  $13 \times 13$  pixels. Le champ de déplacement sur la séquence "spirale", en utilisant la méthode du tenseur, est représenté à la figure 2.9.

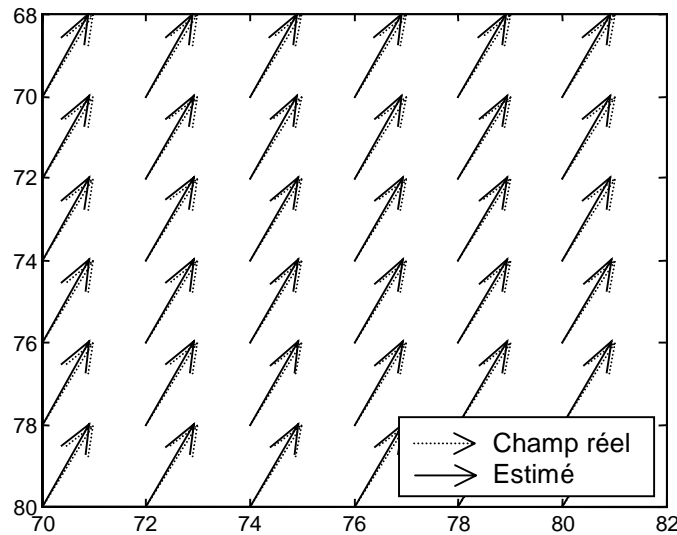


Figure 2.9 : Champ de déplacement estimé à l'aide de la méthode TLS sur la "spirale", séquence bruitée par un bruit gaussien additif ( $SNR = 80 \text{ dB}$ ).

Comme avec la méthode des moindres carrés classiques, nous constatons une erreur d'estimation avec la méthode du tenseur. Aucune différence notable ne peut être faite sur les figures 2.4 et 2.9. Dans ce test, l'apport de la méthode du tenseur sur la précision de l'estimation, par rapport à la méthode des moindres carrés classique n'est pas clairement démontré. Pour une séquence de synthèse, les deux méthodes donnent un champ de déplacement correct par rapport au champ réel. Certains auteurs précisent que dans certains cas, les résultats donnés par la méthode du tenseur sont meilleurs. Nous avons remarqué lors de nos tests que cette méthode n'était pas stable. C'est le cas avec la séquence réelle que nous avons prise. En revanche, quand elle converge correctement, elle produit de meilleurs résultats. C'est en l'occurrence le cas avec la séquence synthétique.

### 1.3.3 Discussion sur la méthode des moindres carrés totaux basée schéma simple

Jusqu'ici, nous remarquons au travers des résultats, la présence d'un léger décalage entre le déplacement réel et l'estimé, dans le cas de l'image synthétique. Tant que nous restons dans le cadre du mouvement de translation, le formalisme basé schéma simple permet d'avoir une estimation correcte. L'allure des vecteurs d'estimation est proche de celle du champ théorique. Cette validité n'est plus justifiée dès que nous passons au modèle paramétrique de mouvement. A titre d'illustration, nous verrons plus tard qu'un mouvement composé d'une translation, d'une rotation et d'une divergence est difficile à estimer par les méthodes basées schéma simple.

Les résultats sont encore plus décevants dans le cas des images réelles ; le champ de déplacement est fortement biaisé. Le biais d'estimation a généralement différentes sources ; sans être exhaustif, nous pouvons citer l'approximation des gradients, les ruptures des modèles de conservation de la luminance et de continuité. En annexe C, nous rappelons les différentes sources d'erreurs présentes dans l'estimation du mouvement.

## 1.4 Schéma direct-rétrograde, technique des moindres carrés associée

### 1.4.1 Motivation

Les paragraphes précédents ayant mis en évidence les limites de la méthode basée sur une équation à simple contrainte, pour l'estimation du flot optique sur une région, il est nécessaire de mettre en place une technique de régularisation. Nous avons vu également dans le paragraphe 1.5.4.a de la première partie que la précision obtenue par les techniques indirectes dépend fortement des valeurs introduites initialement par l'estimation du vecteur de mouvement. Le premier but du formalisme direct-rétrograde est d'introduire implicitement une contrainte de régularisation qui rend stable l'inversion de la matrice des gradients. Sans omettre les autres apports que peut présenter le présent formalisme, il permet au moins de résoudre les imperfections ci-mentionnées. Le principe est que, dans le contexte des méthodes différentielles, nous proposons une régularisation de la fonction d'erreur en prenant en compte à la fois le mouvement direct et le mouvement rétrograde.

### 1.4.2 Extension de l'équation de contrainte du mouvement : notion de schéma direct-rétrograde

Cette formulation est une extension de la méthode des moindres carrés basée schéma simple, qui regarde les informations dans le sens ascendant du temps. L'idée principale est de considérer, pour une même position spatiale  $(x,y)$ , les informations causale et anti-causale du mouvement. Cette considération conduit à écrire deux termes d'erreur entre deux trames successives [RAN00a]. En regardant sur la figure 2.10, l'hypothèse de conservation de la luminance fait naître les deux équations suivantes :

- L'erreur directe définie par la DPD entre la trame  $t$  et celle  $t + dt$  :

$$\varepsilon_d = I(x + dx_d, y + dy_d, t + dt) - I(x, y, t) \quad (162)$$

- L'erreur rétrograde entre  $t + dt$  et  $t$  :

$$\varepsilon_r = I(x - dx_r, y - dy_r, t) - I(x, y, t + dt) \quad (163)$$

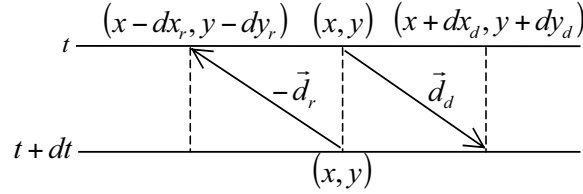


Figure 2.10 : Déplacements engendrés par les mouvements direct et rétrograde.

Dans le formalisme direct-rétrograde, nous faisons intervenir les informations dans les deux sens temporels.

Le développement de Taylor au premier ordre appliqué aux deux erreurs donne les équations suivantes :

$$(ECM_d) : \frac{\partial I(x, y, t)}{\partial x} dx_d + \frac{\partial I(x, y, t)}{\partial y} dy_d + \frac{\partial I(x, y, t)}{\partial t} dt = 0 \quad (164)$$

$$(ECM_r) : \frac{\partial I(x, y, t + dt)}{\partial x} dx_r + \frac{\partial I(x, y, t + dt)}{\partial y} dy_r + \frac{\partial I(x, y, t + dt)}{\partial t} dt = 0 \quad (165)$$

Ces équations s'appellent équations de contrainte des mouvements direct et rétrograde.

Suite à l'introduction de l'erreur rétrograde  $\varepsilon_r$ , une nouvelle fonction coût à minimiser, qui manipule à la fois les informations dans les deux sens temporels, sera définie. En utilisant les équations (162) et (163), la fonctionnelle des erreurs peut être formulée de deux manières :

\*La première consiste à former indépendamment le critère relatif aux mouvements direct et rétrograde. La fonctionnelle globale correspond à la somme des termes quadratiques des deux erreurs, soit :

$$J_a = \sum_{\Omega} [(\varepsilon_d)^2 + (\varepsilon_r)^2] \quad (166)$$

En passant par les équations de contrainte du mouvement, cette fonctionnelle s'écrit :

$$J_a = \sum_{\Omega} \left[ \left( I_x^t u_d + I_y^t v_d + I_t^t \right)^2 + \left( I_x^{t+Te} u_r + I_y^{t+Te} v_r + I_t^{t+Te} \right)^2 \right] \quad (167)$$

où  $u_d = dx_d/dt$ ,  $v_d = dy_d/dt$ ,  $u_r = dx_r/dt$  et  $v_r = dy_r/dt$ .

En notation discrète, nous remplaçons  $t + dt$  par  $t + Te$ . Nous désignons ainsi par  $I_x^{t+Te}$  et  $I_y^{t+Te}$  les gradients spatiaux suivant  $x$  et  $y$  pour l'image à  $t + dt$  et par  $I_t^{t+Te}$  le gradient temporel.

Dans un contexte de mouvement de translation, malgré la présence de deux notations différentes (direct et rétrograde), la minimisation du critère au sens des moindres carrés ne fournit pas en conséquence deux déplacements distincts. En effet, nous cherchons un déplacement commun qui soit issu de la minimisation du critère  $J_a$ . La résolution au sens des moindres carrés donne :

$$(u, v)^T = -(A_1^T A_1 + A_2^T A_2)^{-1} (A_1^T B_2 + A_2^T B_2) \quad (168)$$

Les deux matrices des gradients spatiaux sont définies par :

$$A_1 = \begin{bmatrix} I_x^t(1) & I_y^t(1) \\ \vdots & \vdots \\ I_x^t(n) & I_y^t(n) \end{bmatrix} \text{ et } A_2 = \begin{bmatrix} I_x^{t+Te}(1) & I_y^{t+Te}(1) \\ \vdots & \vdots \\ I_x^{t+Te}(n) & I_y^{t+Te}(n) \end{bmatrix} \quad (169)$$

Pour les gradients temporels, nous avons :

$$B_1 = \begin{bmatrix} I_t^t(1) \\ \vdots \\ I_t^t(n) \end{bmatrix} \text{ et } B_2 = \begin{bmatrix} I_t^{t+Te}(1) \\ \vdots \\ I_t^{t+Te}(n) \end{bmatrix} \quad (170)$$

Nous rappelons que  $n$  désigne le nombre de pixels dans la fenêtre d'estimation  $\Omega$ .

\*Quant à la deuxième formulation, le principe est le suivant : dans le contexte d'un mouvement de translation, une large contrainte d'uniformité doit être introduite. Une technique pour réaliser cette condition est d'utiliser une fonctionnelle en admettant que l'erreur totale est donnée par la somme algébrique de deux erreurs :

$$J_b = \sum_{\Omega} (\varepsilon_d + \varepsilon_r)^2 \quad (171)$$

En introduisant les équations de contraintes, la fonctionnelle devient :

$$J_b = \sum_{\Omega} \left[ (I_x^t + I_x^{t+Te})u + (I_y^t + I_y^{t+Te})v + (I_t^t + I_t^{t+Te}) \right]^2 \quad (172)$$

La minimisation de ce critère aboutit à une solution au sens des moindres carrés de la forme :

$$(u, v)^T = -[(A_1 + A_2)^T (A_1 + A_2)]^{-1} (A_1 + A_2)^T (B_1 + B_2) \quad (173)$$

Nous rappelons au passage que le formalisme direct-rétrograde fournit un déplacement commun aux deux informations. C'est pour cela que nous aboutissons finalement à une notation unique  $(u, v)^T$  pour l'estimation du vecteur vitesse.

En tant qu'estimation de base, nous réduisons volontairement le support temporel sur deux images. Ceci a pour but de minimiser le coût de calcul en terme de mémoire et en terme de rapidité. L'extension de la méthode à plus de deux images sera détaillée plus tard. Pour l'estimation sur deux images, les gradients temporels aux instants  $t$  et  $t + dt$  sont pris égaux.

Quand on regarde l'équation (172), on a l'impression que l'expression du critère  $J_b$  est similaire à celle du critère utilisé dans l'approche classique (basée schéma simple), sauf qu'ici, les gradients sont calculés en prenant leur moyenne entre les deux trames ( $t$  et  $t + dt$ ). Autrement dit, lorsqu'on utilise la moyenne des gradients spatiaux entre les images  $t$  et  $t + dt$  (approximation des gradients proposée par Horn et Schunck [HOR81]), la formulation basée sur un schéma simple ressemble à celle basée sur un schéma direct-rétrograde. En revanche, nous verrons par la suite que la méthodologie que nous présentons ici ne consiste pas en une reformulation implicite de l'utilisation des gradients selon Horn et Schunck. A titre d'indication, nous verrons plus tard qu'avec l'utilisation de la présente formulation dans l'approche incrémentale, l'ambiguïté entre la dérivée proposée par Horn et Schunck et le formalisme direct-rétrograde ne se présente plus. En outre, l'extension de ce formalisme au modèle de mouvement permet de justifier qu'il ne s'agit pas d'une simple utilisation des gradients moyens.

Les avantages, les inconvénients et les commentaires liés à ces deux formulations ( $J_a$  et  $J_b$ ) seront dressés par la suite à l'aide de résultats de simulation.

### 1.4.3 Résultats expérimentaux

Nous essayons de comparer la méthode des moindres carrés simple et celle basée sur un schéma direct-rétrograde. La première comparaison sera effectuée sur l'image synthétique de la spirale. Deux autres images réelles seront aussi considérées.

### a) Séquence synthétique : "spirale"

Les paramètres de translation utilisés sont les mêmes que ceux utilisés dans la méthode des moindres carrés simple. Quant à l'approche basée sur un schéma direct-rétrograde, nous considérons en premier lieu le formalisme associé à la fonctionnelle  $J_a$ . Les gradients spatiaux sont estimés en utilisant le masque SOBEL. La différence inter-image convoluée avec le masque MOYEN\_t donne le gradient temporel. L'image est bruitée avec un bruit additif gaussien de 80 dB. L'estimation s'effectue en utilisant un support de calcul de  $13 \times 13$  pixels. La figure 2.11 représente le champ de déplacement, superposé avec le champ réel.

Nous allons considérer maintenant le cas de la fonctionnelle  $J_b$ . Les paramètres d'estimation sont inchangés. Les champs de déplacement réel et estimé sont donnés sur la figure 2.12.

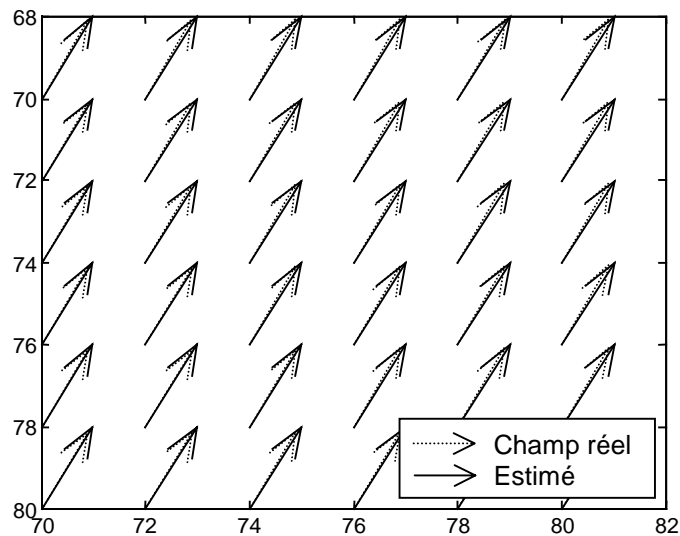


Figure 2.11 : Champ de déplacement estimé à l'aide de l'approche des moindres carrés en schéma direct-rétrograde, avec la fonctionnelle  $J_a$  ; séquence synthétique "spirale", bruitée par un bruit gaussien additif ( $SNR = 80 \text{ dB}$ ).

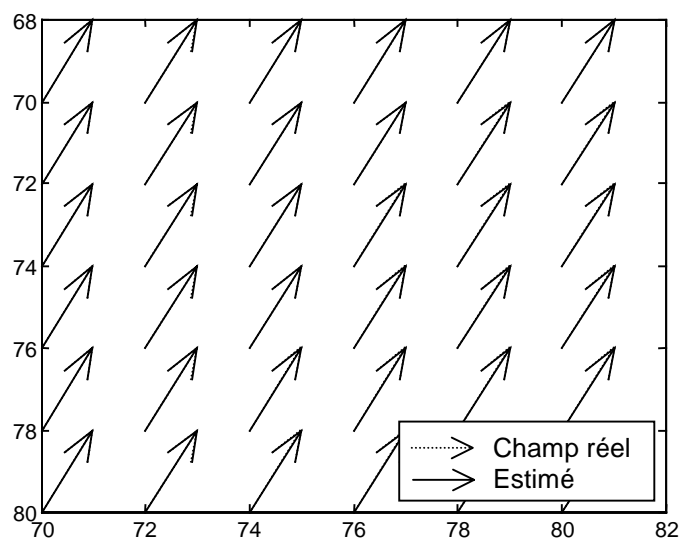


Figure 2.12 : Champ de déplacement estimé à l'aide de l'approche des moindres carrés en schéma direct-rétrograde, avec la fonctionnelle  $J_b$  ; séquence synthétique "spirale", bruitée par un bruit gaussien additif ( $SNR = 80 \text{ dB}$ ).

L'objectif de la présente expérimentation est d'abord de comparer le formalisme basé schéma direct-rétrograde, qu'il soit selon  $J_a$  (figure 2.11) ou selon  $J_b$  (figure 2.12), avec le formalisme basé flot simple (figure 2.4). Ce test permet également de comparer entre eux les résultats associés à la fonctionnelles  $J_a$  et ceux associés à  $J_b$ .

Tout d'abord, en ce qui concerne la comparaison entre le formalisme simple et les deux formalismes direct-rétrograde, les résultats des figures 2.11 et 2.12 permettent de juger les performances du schéma proposé. Contrairement à la figure 2.4, les champs estimés sur ces deux figures sont visuellement plus proches des valeurs réelles.

Les résultats suivants reposent sur l'analyse des biais. La courbe permet, pour chacune des approches, d'évaluer l'écart entre la valeur réelle et l'estimé. Elle permet également de comparer entre eux les résultats obtenus pour les différents formalismes. Chaque point représente le biais moyen évalué sur un ensemble de pixels. Nous notons que le résultat donné dans toute expérimentation correspond à la moyenne des résultats d'un lot de tests (plusieurs boucles de tests). Les figures 2.13 et 2.14 représentent les biais d'estimation du déplacement (vecteur vitesse) suivant  $x$  et celui suivant  $y$ , pour un SNR allant de 40 dB à 120 dB. Sur les figures 2.15 et 2.16, nous voyons la variance de l'erreur d'estimation correspondant aux deux composantes de la vitesse.

L'erreur d'estimation se stabilise à partir d'un SNR de 60 dB. L'erreur associée au formalisme direct-rétrograde ( $J_a$  ou  $J_b$ ) reste inférieure à celle du formalisme simple.

Quant aux formalismes  $J_a$  et  $J_b$ , nous pouvons dire que la méthode utilisant le critère  $J_a$  est plus sensible aux bruits que celle de  $J_b$ . Cet écart peut venir du fait que le critère  $J_b$  introduit implicitement une contrainte d'uniformité tandis qu'avec  $J_a$ , l'optimisation sur les deux déplacements peut être faite indépendamment l'une de l'autre.

L'analyse de la variance de l'erreur d'estimation permet de constater que pour un SNR allant de 40 dB à 60 dB, la variance varie rapidement. Au delà de 60 dB, elle évolue lentement avec une variation inférieure à 0,5%. Sur ces courbes, les variances associées au formalisme basé sur un schéma direct-rétrograde ( $J_a$  ou  $J_b$ ) sont faibles par rapport à celles du formalisme classique (basé schéma simple).

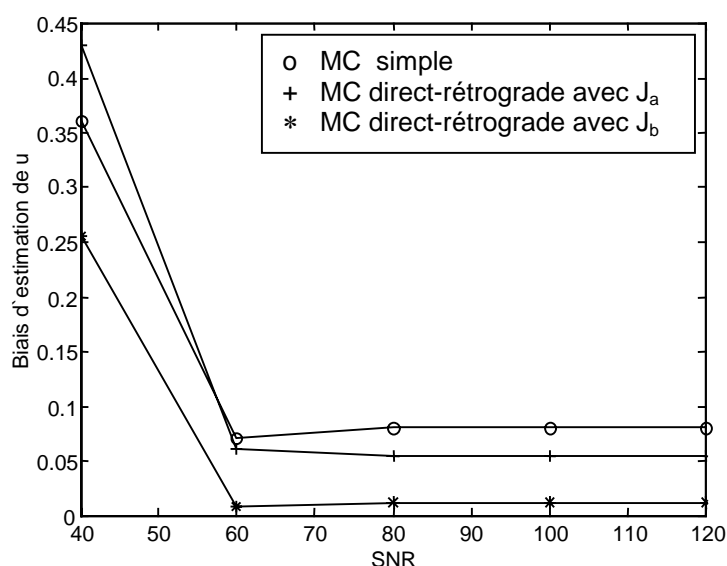


Figure 2.13 : Biais d'estimation de la composante  $u$  du vecteur vitesse en fonction du SNR, pour l'approche moindres carrés (MC) basée sur un schéma simple, approche basée sur un schéma direct-rétrograde selon la fonctionnelle  $J_a$ , et approche basée sur un schéma direct-rétrograde selon la fonctionnelle  $J_b$ ; séquence "spirale".



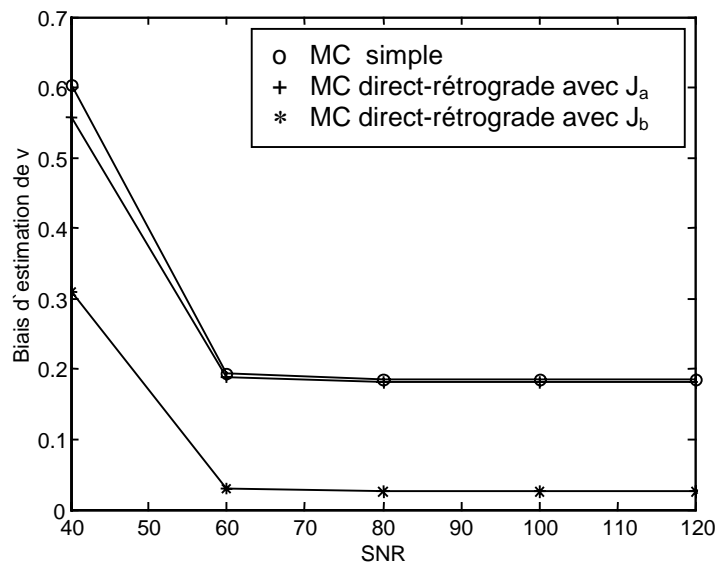


Figure 2.14 : Biais d'estimation de la composante  $v$  du vecteur vitesse en fonction du SNR, pour l'approche moindres carrés basée sur un schéma simple, approche basée sur un schéma direct-rétrograde avec la fonctionnelle  $J_a$ , et approche basée sur un schéma direct-rétrograde avec la fonctionnelle  $J_b$ ; séquence "spirale".

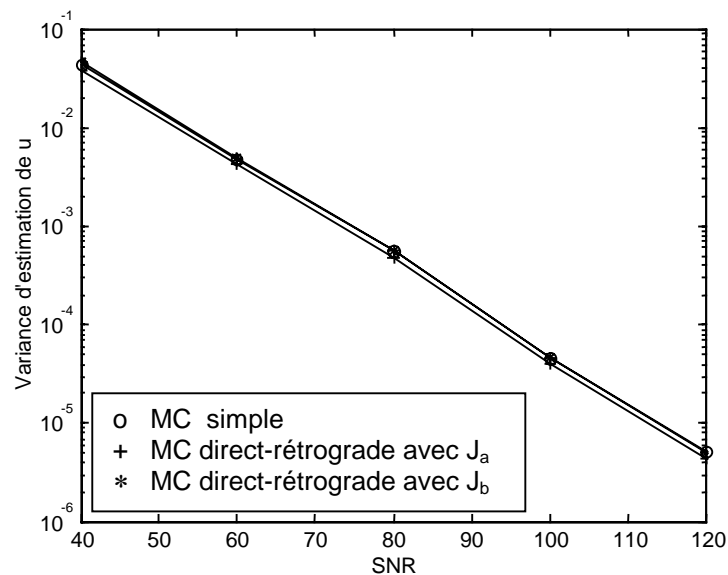


Figure 2.15 : Variance de l'erreur d'estimation pour la composante  $u$  du vecteur vitesse en fonction du SNR, pour l'approche moindres carrés basée sur un schéma simple, approche basée sur un schéma direct-rétrograde selon la fonctionnelle  $J_a$ , et approche basée sur un schéma direct-rétrograde selon la fonctionnelle  $J_b$ ; séquence "spirale".

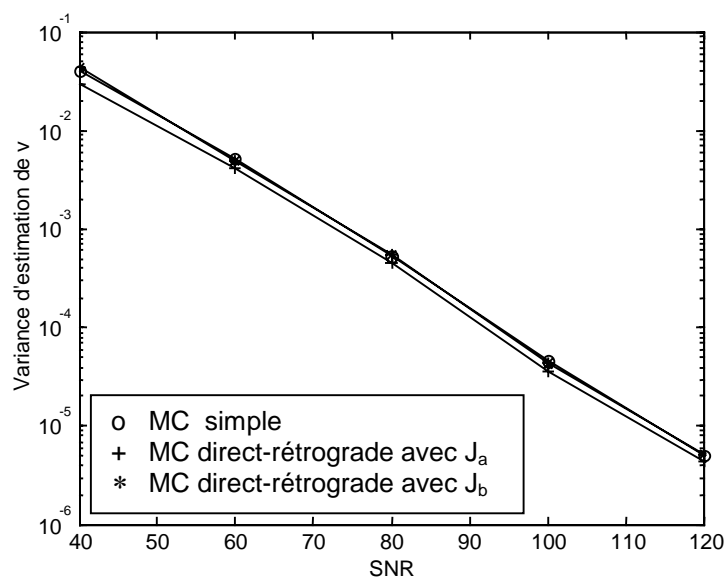


Figure 2.16 : Variance de l'erreur d'estimation pour la composante  $v$  du vecteur vitesse en fonction du SNR, pour l'approche moindres carrés basée sur un schéma simple, approche basée sur un schéma direct-rétrograde selon la fonctionnelle  $J_a$ , et approche basée sur un schéma direct-rétrograde selon la fonctionnelle  $J_b$  ; séquence "spirale".

### b) Séquence réelle à déplacement synthétique

La présente expérimentation est effectuée sur l'image réelle présentée à la figure 2.6 ; une translation de  $(-2,1)$  pixels est affectée à l'objet ①. La séquence est bruitée par un bruit blanc gaussien additif de  $100 \text{ dB}$ . Les gradients spatiaux sont calculés à l'aide de l'opérateur MDIF. Le gradient temporel est obtenu par la différence temporelle, convoluée avec le masque MOYEN\_t. Le choix de ces deux masques est justifié par leur meilleur filtrage. L'estimation est effectuée à l'aide d'un support de calcul de  $13 \times 13$  pixels. Les figures 2.17 et 2.18 présentent le champ de déplacement estimé, comparé au champ réel, et cela en utilisant la fonctionnelle  $J_b$ .

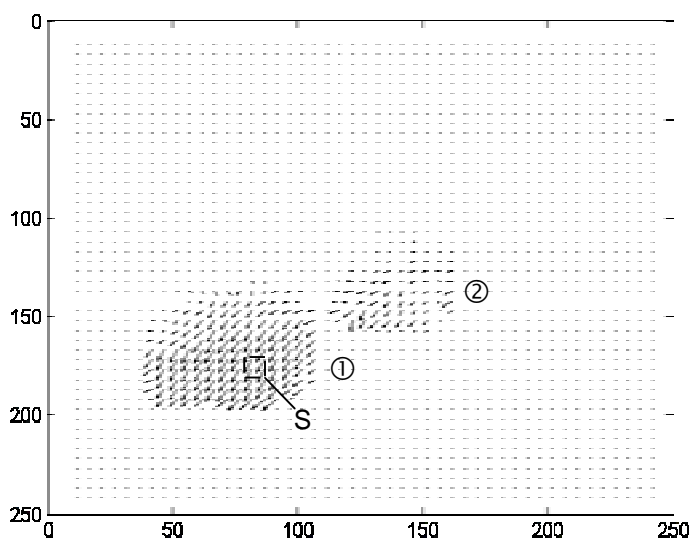


Figure 2.17 : Champ de déplacement estimé sur l'image réelle à déplacement synthétique, avec la méthode des moindres carrés basée sur un schéma direct-rétrograde, en utilisant la fonctionnelle  $J_b$ .

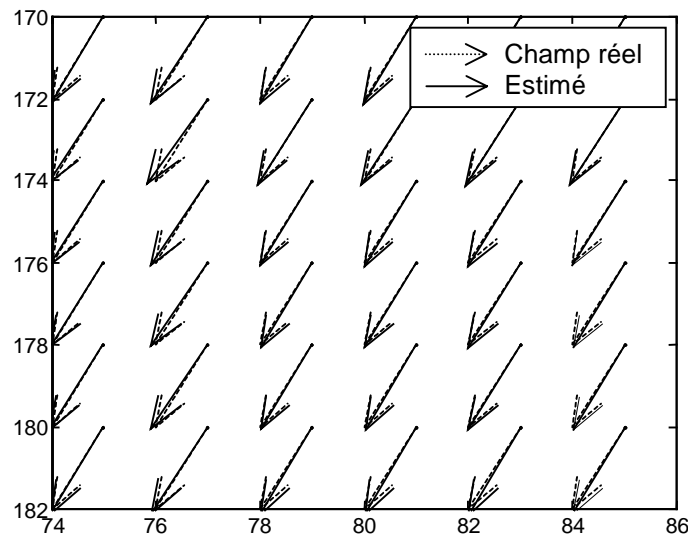


Figure 2.18 : Champ de déplacement en faisant un agrandissement de la zone  $S$ .

La différence entre l'approche des moindres carrés basée sur un schéma simple et celle basée sur un schéma direct-rétrograde peut se constater en revenant sur la figure 2.8. La comparaison des figures 2.18 et 2.8 permet d'observer l'apport du formalisme direct-rétrograde sur l'orientation et l'amplitude du déplacement. L'utilisation de cette image réelle à mouvement synthétique permet de valider les performances du schéma proposé. Le champ estimé et le déplacement réel sont quasiment superposés. Les résultats sont positifs au niveau des valeurs quantitatives.

### c) Séquence réelle : "taxi"

Dans ce test, nous avons choisi d'illustrer le comportement du schéma direct-rétrograde sur une image plus complexe. Il s'agit d'une séquence dont quatre objets sont mobiles avec des déplacements différents (figure 2.19).



Figure 2.19 : Séquence réelle dans laquelle quatre objets sont réellement en mouvement ; séquence "taxi".

La détection du mouvement est effectuée sur cette séquence. Les résultats sont donnés sous la forme d'amplitude des déplacements, issue des trois approches différentes :

- L'amplitude représentée à la figure 2.20 est calculée à partir de la méthode des moindres carrés classiques (c'est-à-dire en tenant compte uniquement du mouvement direct).

- Le deuxième test est consacré à l'approche des moindres carrés à flot direct-rétrograde, avec la fonctionnelle  $J_a$ . Nous rappelons que cette formulation correspond à la somme des termes quadratiques des erreurs (directe et rétrograde). Le résultat est présenté à la figure 2.21. Dans cette approche, pour l'estimation du gradient temporel, nous avons fait intervenir une troisième trame. Ainsi, dans l'équation du mouvement rétrograde, le gradient temporel est calculé à partir de la différence temporelle entre la deuxième trame et la troisième. La raison est que, comme nous l'avons vu dans l'expérimentation  $a$ , cette formulation est sensible au bruit ; l'introduction d'une troisième trame permet un bon filtrage à l'approximation du gradient temporel.

- Le troisième test correspond à la fonctionnelle  $J_b$ , c'est à dire à la méthode des moindres carrés dont le critère à minimiser est formulé par le carré de la somme de deux erreurs. La figure 2.23 illustre l'amplitude du déplacement obtenu.

Pour la résolution au sens des moindres carrés, un voisinage spatial de taille  $13 \times 13$  pixels est utilisé. L'approximation des gradients spatiaux s'effectue en utilisant le masque SOBEL. Les gradients temporels sont estimés par la différence temporelle, convoluée avec le masque MOYEN\_t.

La comparaison de ces trois figures (2.20, 2.21 et 2.22) est ici peu significative. Elles sont caractérisées par l'amplitude des déplacements et l'efficacité de détection des objets. Nous pouvons néanmoins remarquer que la détection est nettement meilleure dans les deux cas du formalisme basé flot direct-rétrograde.

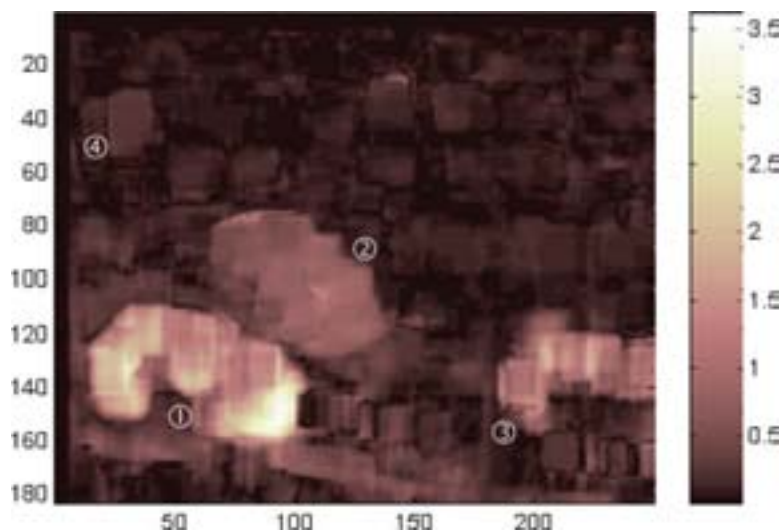


Figure 2.20 : Détection du mouvement à l'aide de l'approche des moindres carrés basée sur un schéma simple ; amplitude du déplacement ; séquence "taxi".

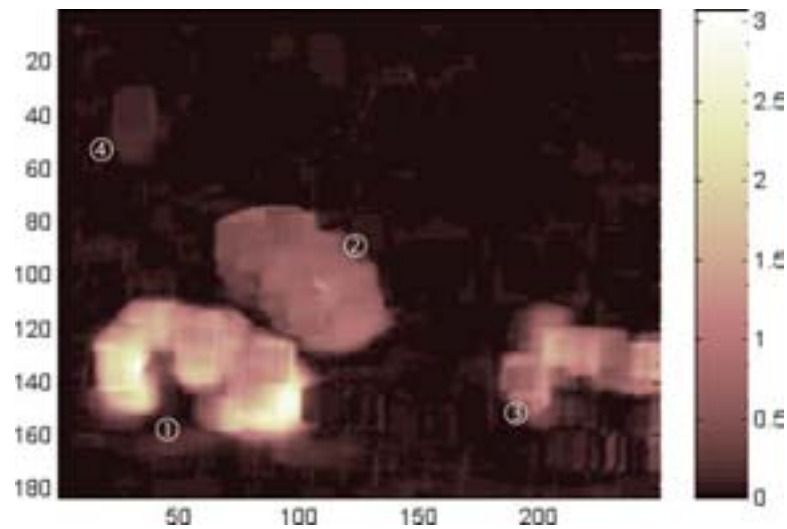


Figure 2.21 : Détection du mouvement à l'aide de l'approche des moindres carrés basée sur un schéma direct-rétrograde, avec la fonctionnelle  $J_a$  ; amplitude du déplacement ; séquence "taxi".

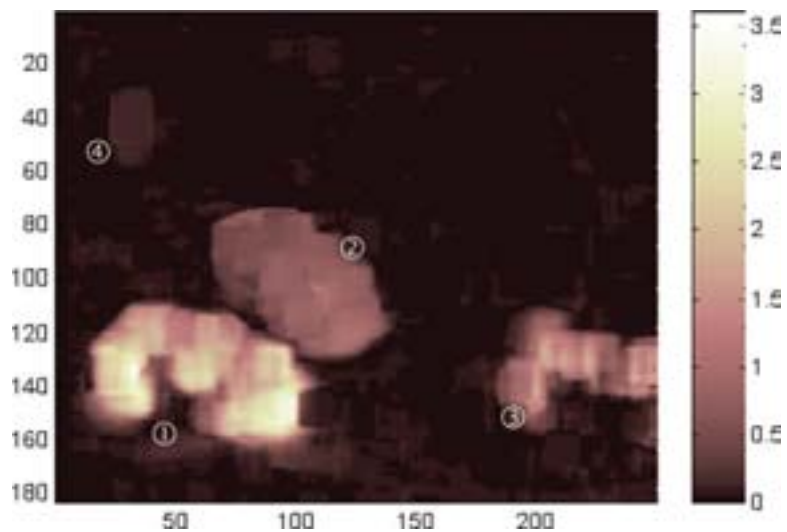


Figure 2.22 : Détection du mouvement à l'aide de l'approche des moindres carrés selon le formalisme direct-rétrograde, avec la fonctionnelle  $J_b$  ; amplitude du déplacement ; séquence "taxi".

Pour pouvoir analyser de près les résultats associés à ces trois cas de figure, nous allons tracer le champ de déplacement. Prenons l'exemple de l'objet ① ; les résultats suivants correspondent au flot optique estimé sur cet objet :

- Le champ représenté à la figure 2.23 est calculé à partir de la méthode des moindres carrés simple (c'est à dire en prenant en compte uniquement le mouvement direct).
- Le deuxième test est consacré à l'approche basée sur un schéma direct-rétrograde, et cela avec la fonctionnelle  $J_a$  ; le résultat est illustré sur la figure 2.24.

- Le troisième test correspond à l'approche basée sur un schéma direct-rétrograde, associée à la fonctionnelle  $J_b$ . La figure 2.25 montre le champ de déplacement obtenu.

Quantitativement, une différence entre les trois figures précédentes n'apparaît pas de manière évidente. En revanche, une observation qualitative peut en être tirée. A titre d'illustration, elles présentent une différence au niveau de l'orientation du champ de déplacement. A première vue, les champs des deux dernières figures sont moins désordonnés et plus significatifs par rapport au mouvement réel de l'objet.

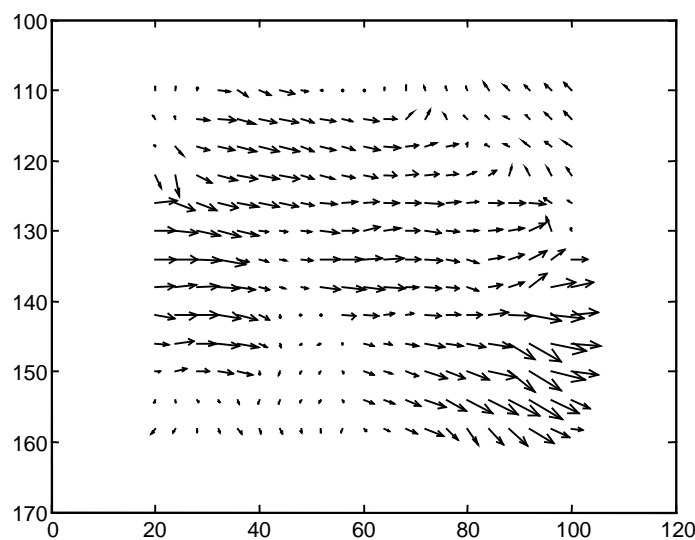


Figure 2.23 : Champ de déplacement de l'objet ① de la séquence "taxi", en utilisant la méthode des moindres carrés simple.

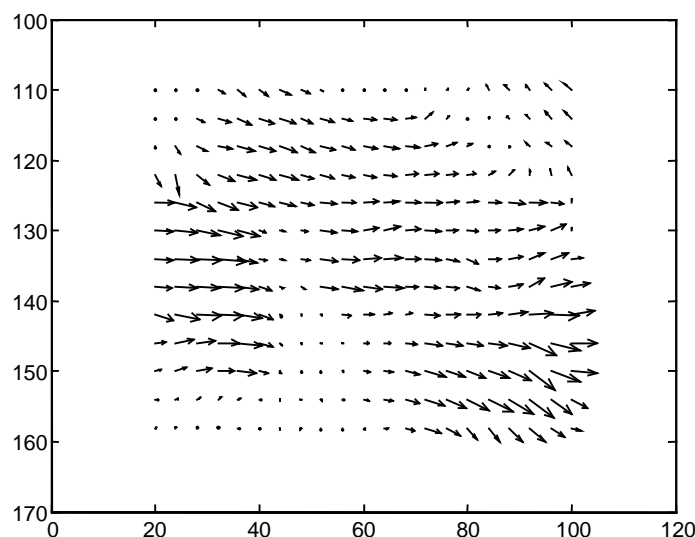


Figure 2.24 : Champ de déplacement de l'objet ① de la séquence "taxi", en utilisant la méthode des moindres carrés basée sur un schéma direct-rétrograde, avec la fonctionnelle  $J_a$ .

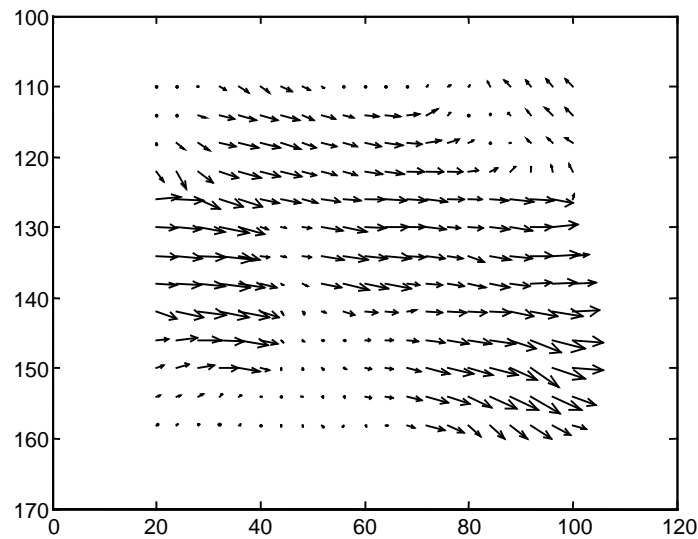


Figure 2.25 : Champ de déplacement de l'objet ① de la séquence "taxi", en utilisant la méthode des moindres carrés basée sur un schéma direct-rétrograde, avec la fonctionnelle  $J_b$ .

#### 1.4.4 Discussion sur l'estimation basée sur un schéma direct-rétrograde et la technique des moindres carrés associée

Les méthodes présentées dans ce paragraphe s'inspirent à la fois de la méthode différentielle et de la résolution au sens des moindres carrés. Nous avons considéré le cas d'un mouvement de translation. Dans ce cadre, pour renforcer les contraintes du mouvement, nous avons introduit en sus de l'information dans le sens temporel direct, l'information dans le sens rétrograde. Ceci a permis de générer un nouveau formalisme au niveau du critère à minimiser par une méthode d'optimisation. La minimisation a été effectuée à l'aide de la technique des moindres carrés.

Pour comparer entre eux les résultats obtenus en utilisant des méthodes des moindres carrés de formalisme différent (simple, direct-rétrograde), nous avons choisi de tracer le biais et le champ de déplacement. Dans le formalisme classique (basé sur un schéma simple), l'erreur d'estimation est notable. L'application du formalisme à mouvement direct-rétrograde permet de diminuer cette erreur. Nous constatons d'après les résultats que l'utilisation du formalisme basé sur un schéma direct-rétrograde permet d'obtenir une erreur nettement plus faible sur les séquences réelle et synthétique.

De même, en terme d'affinement, le résultat est satisfaisant par rapport au résultat obtenu avec le formalisme classique. Ces résultats démontrent aussi l'utilité de l'introduction de l'information rétrograde dans l'équation de contrainte du mouvement.

La dernière étape consiste maintenant à introduire le formalisme basé sur un mouvement direct-rétrograde dans un schéma itératif. Dans cette optique, plusieurs voies sont encore envisageables ; à savoir une procédure itérative par pondération des données et l'approche incrémentale. Le paragraphe suivant est consacré à l'estimation robuste basée sur un schéma direct-rétrograde, en pondérant les données.

## 1.5 Estimation robuste basée sur un schéma direct-rétrograde<sup>8</sup>

### 1.5.1 Description

L'approche robuste en question consiste en la pondération de données utilisées dans l'algorithme d'estimation s'appuyant sur le schéma direct-rétrograde. Le but de la pondération est de collecter une quantité suffisante d'informations pour rendre l'estimation plus robuste et de rejeter les informations erronées. Nous utilisons ainsi le M-estimateur [HOL77, MEE91] qui consiste à minimiser une fonction de résidus évaluée à partir du déplacement estimé et, par la suite, à calculer les poids attribués aux pixels. Les poids sont calculés différemment pour les deux fonctionnelles.

\* Considérons en premier lieu la fonctionnelle à somme des termes quadratiques des erreurs ( $J_a$ ), la fonction d'erreur résiduelle est exprimée par :

$$\sum_{\Omega} [\rho(\varepsilon_d) + \rho(\varepsilon_r)] = \sum_{\Omega} \left( \frac{1}{2} \alpha_d \varepsilon_d^2 + \frac{1}{2} \alpha_r \varepsilon_r^2 \right) \quad (174)$$

où  $\alpha_d$  (respectivement  $\alpha_r$ ) désigne le poids attribué à chaque point  $(x, y)$  et  $\varepsilon_d$  (respectivement  $\varepsilon_r$ ) l'erreur résiduelle dans le sens temporel direct (respectivement rétrograde). Les expressions de ces deux erreurs sont données par les équations (162, 163).

La fonction résiduelle est minimale quand sa dérivée est nulle, c'est à dire :

$$\frac{\partial \rho(\varepsilon)}{\partial d} = 0 \quad (175)$$

Toujours dans le cas de la fonctionnelle  $J_a$ , nous avons deux ensembles de poids distincts [ODO94, RAN00b].

Le poids calculé en un point arbitraire, relatif au mouvement direct, est exprimé par :

$$\alpha_d = \frac{1}{\varepsilon_d} \frac{\partial \rho(\varepsilon_d)}{\partial \varepsilon_d} \quad (176)$$

Les coefficients  $\alpha_d$ , pour chaque point  $(x, y)$  du support  $\Omega$ , constituent les composantes de la matrice des poids  $W_1$ .

Le poids calculé en un point, associé au mouvement rétrograde, est exprimé par :

$$\alpha_r = \frac{1}{\varepsilon_r} \frac{\partial \rho(\varepsilon_r)}{\partial \varepsilon_r} \quad (177)$$

Ces coefficients constituent les composantes de la matrice de poids  $W_2$ .

Comme le passage de l'équation (166) à l'équation (168), la minimisation de la fonctionnelle (174) permet d'avoir les composantes du vecteur vitesse. La solution donnée par l'approche robuste directe-rétrograde au sens des moindres carrés<sup>9</sup> est exprimée par :

$$(u, v)^T = -\left( A_1^T W_1 A_1 + A_2^T W_2 A_2 \right)^{-1} \left( A_1^T W_1 B_1 + A_2^T W_2 B_2 \right) \quad (178)$$

Les matrices diagonales  $W_1$  et  $W_2$  sont construites avec l'ensemble des poids, avec :

$$W_1 = \text{diag}[\alpha_d(1), \dots, \alpha_d(n)] \quad (179)$$

et :

$$W_2 = \text{diag}[\alpha_r(1), \dots, \alpha_r(n)] \quad (180)$$

La quantité  $n$  désigne le nombre de pixels dans le support  $\Omega$ .

<sup>8</sup> Forward-backward robust estimation

<sup>9</sup> Robust forward-backward least squares



\*Quant à la deuxième fonctionnelle  $J_b$ , nous avons :

$$\sum_{\Omega} \rho(\varepsilon) = \sum_{\Omega} \frac{1}{2} \alpha (\varepsilon_d + \varepsilon_r)^2 \quad (181)$$

où  $\alpha$  désigne le poids attribué à chaque point  $(x, y)$  et  $\varepsilon$  l'erreur résiduelle.

Le poids en tout point  $(x, y)$  est déterminé par :

$$\alpha = \frac{1}{\varepsilon} \frac{\partial \rho(\varepsilon)}{\partial \varepsilon} \quad (182)$$

Comme le passage de l'équation (171) à l'équation (173), la solution est obtenue par minimisation de la fonctionnelle (181). Cette solution est de la forme :

$$(u, v)^T = -[(A_1 + A_2)^T W (A_1 + A_2)]^{-1} (A_1 + A_2)^T W (B_1 + B_2) \quad (183)$$

où  $W$  désigne la matrice de poids communs pour les deux gradients, avec :

$$W = \text{diag}[\alpha(1), \dots, \alpha(n)] \quad (184)$$

Pour l'expression de  $\rho$ , nous avons pris en compte la fonction de Tukey. Son expression est de la forme :

$$\rho(\varepsilon) = \begin{cases} \frac{C^2}{6} \left[ 1 - \left( 1 - \left( \frac{\varepsilon}{C} \right)^2 \right)^3 \right] & \text{si } |\varepsilon| \leq C \\ \frac{C^2}{6} & \text{si } |\varepsilon| > C \end{cases} \quad (185)$$

où :

$$C = 1,48 \cdot P \cdot \text{Med} \left\{ \varepsilon(i) - \text{Med}[\varepsilon(j)] \right\}_{i=1 \dots n, j=1 \dots n} \quad (186)$$

Le paramètre  $P$  est le facteur de proportionnalité qui assure le meilleur seuil  $C$  dans le cas où le bruit est gaussien. Le terme *Med* fait référence à la fonction médiane.

Une telle pondération est aussi bien applicable dans la méthode du tenseur basée mouvement direct-rétrograde que nous allons voir plus tard.

### 1.5.2 Résultats expérimentaux

Ce test porte sur des résultats issus de l'estimation robuste associée au formalisme direct-rétrograde. Dans un premier temps, le champ est calculé sur une zone de la spirale. Les gradients sont pondérés itérativement, tout en minimisant l'erreur résiduelle en un point considéré. Au bout d'un certain nombre d'itérations préalablement définies, nous arrêtons le processus. De nombreuses simulations nous ont permis de conclure que, pour avoir un meilleur seuillage  $C$ , la valeur du paramètre  $P$  est fixée à 4.7.

L'amélioration apportée par l'approche des moindres carrés pondérés, à flot direct-rétrograde, peut se mesurer en comparant les biais d'estimation et les variances de l'erreur d'estimation. Les figures 2.26 et 2.27 montrent les biais d'estimation pour les composantes du vecteur vitesse, et cela, pour les approches robustes à flot simple et à flot direct-rétrograde. Le biais pour un bruit donné est calculé par la moyenne des biais sur un ensemble de points. Cette valeur résulte également de la moyenne des résultats d'un lot de tests (plusieurs boucles de test). L'estimation est stable à partir d'un rapport signal sur bruit de 80 dB. En comparant les trois courbes de chacune de ces figures, nous pouvons constater que l'approche robuste à flot simple est plus biaisée par rapport aux deux approches robustes à flot direct-rétrograde ( $J_a$  et  $J_b$ ). La courbe associée au formalisme  $J_b$  présente une meilleure performance.

Quant aux variances, elles sont représentées par les figures 2.28 et 2.29. Aucune différence notable ne peut être ici remarquée sur les trois approches.

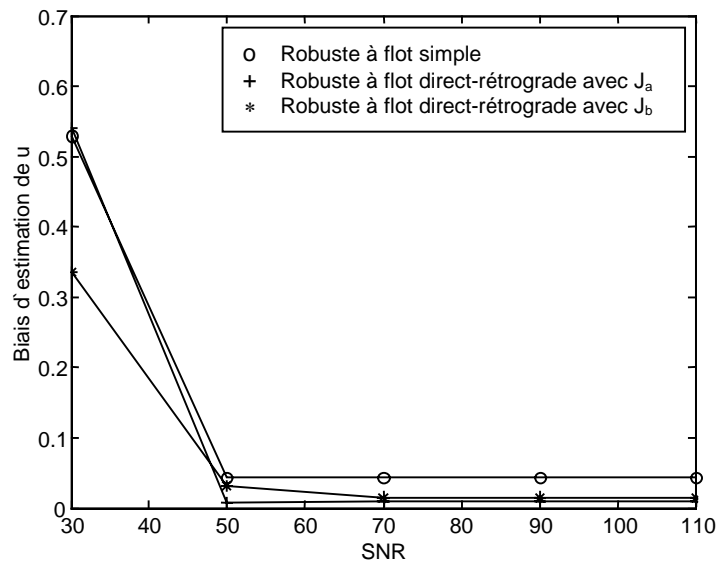


Figure 2.26 : Comparaison des biais d'estimation de la composante  $u$  du vecteur vitesse sur la séquence "spirale" ; méthode des moindres carrés basée sur un schéma direct, celle basée sur un schéma direct-rétrograde selon la fonctionnelle  $J_a$  et celle basée sur un schéma direct-rétrograde selon  $J_b$ .

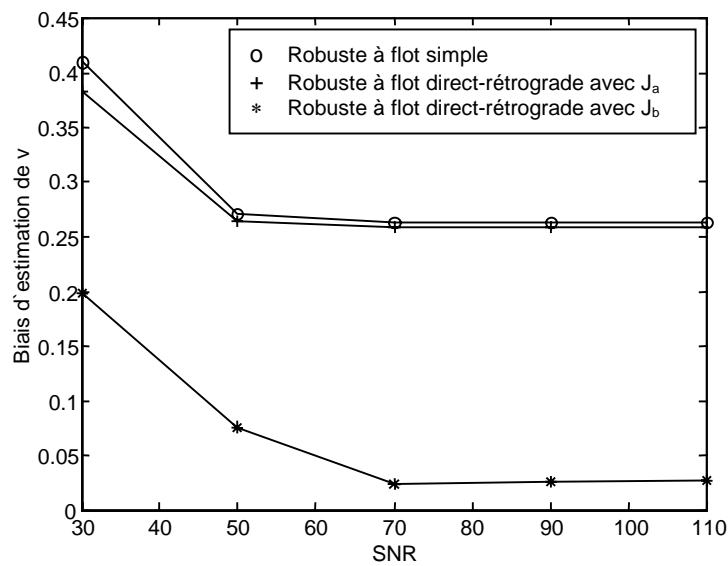


Figure 2.27 : Comparaison des biais d'estimation de la composante  $v$  du vecteur vitesse sur la séquence "spirale" ; méthode des moindres carrés basée sur un schéma direct, celle basée sur un schéma direct-rétrograde selon la fonctionnelle  $J_a$  et celle basée sur un schéma direct-rétrograde selon  $J_b$ .

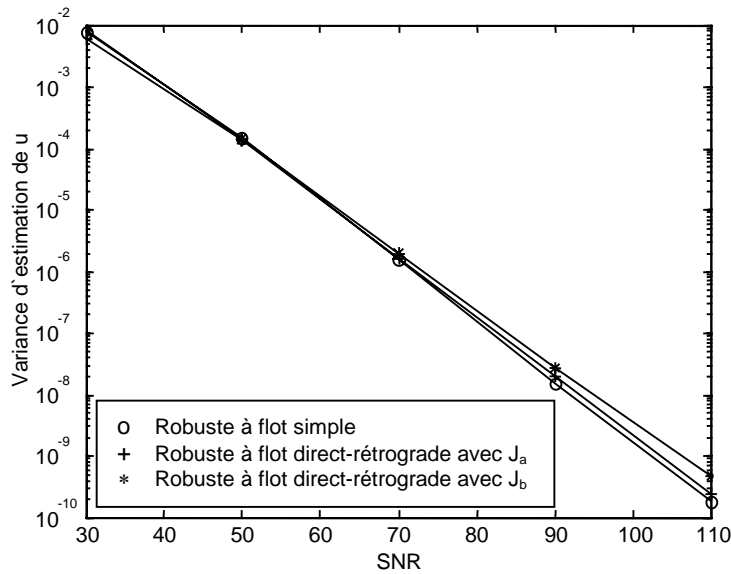


Figure 2.28 : Comparaison des variances de l'erreur d'estimation pour la composante  $u$  du vecteur vitesse, sur la séquence "spirale" ; méthode des moindres carrés basée sur un schéma direct, celle basée sur un schéma direct-rétrograde avec la fonctionnelle  $J_a$  et celle basée sur un schéma direct-rétrograde selon  $J_b$ .

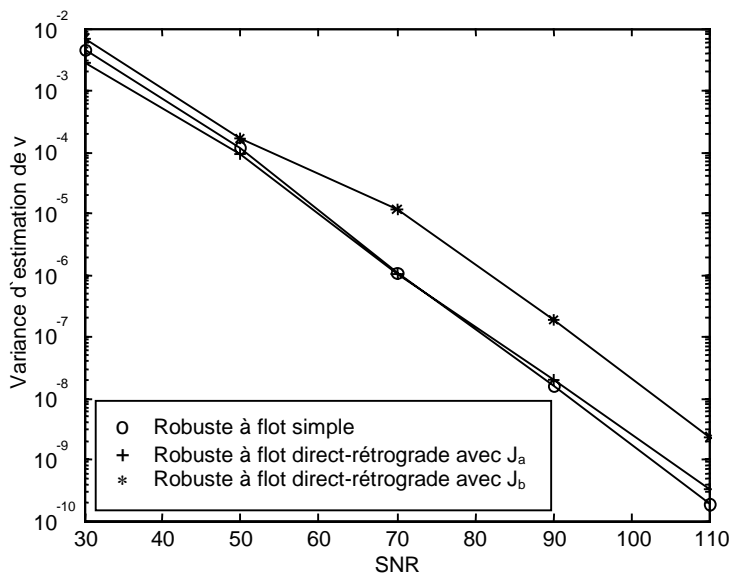


Figure 2.29 : Comparaison des variances de l'erreur d'estimation pour la composante  $v$  du vecteur vitesse, sur la séquence "spirale" ; méthode des moindres carrés basée sur un schéma direct, celle basée sur un schéma direct-rétrograde avec la fonctionnelle  $J_a$  et celle basée sur un schéma direct-rétrograde selon  $J_b$ .

### 1.5.3 Discussion sur l'estimation robuste basée sur un schéma direct-rétrograde

Dans ce paragraphe, nous avons testé la validité du formalisme direct-rétrograde dans le cadre de l'estimation robuste. Les résultats mettent en évidence cette validité. Les méthodes testées s'inspirent de la technique robuste M-estimateur. Nous tenons à rappeler que le but de la pondération était de rejeter les points aberrants dont l'erreur résiduelle est importante. Pour optimiser les résultats, il est possible d'augmenter le nombre d'itérations mais nous avons remarqué que très peu d'itérations (2 itérations) suffisent pour avoir des résultats stables.

## 1.6 Méthode du tenseur basée sur un schéma direct-rétrograde<sup>10</sup>

### 1.6.1 Description

Dans la méthode du tenseur classique (paragraphe 1.3), on ne prend en compte que les erreurs totales formulées suivant un seul sens temporel (déplacement direct). Nous rappelons que ces erreurs sont composées de celle de la matrice des gradients spatiaux et celle de la matrice des gradients temporels. A la différence de l'approche du tenseur classique, celle à base de schéma direct-rétrograde consiste à prendre en compte les erreurs des gradients spatiaux et temporel, à la fois selon le mouvement direct et selon le mouvement rétrograde [RAN00b]. La mise en équation sous forme matricielle des erreurs directe et rétrograde peut être formulée comme suit :

- Pour l'erreur directe, on a :

$$\begin{bmatrix} A_1 + \Delta A_1 & B_1 + \Delta B_1 \end{bmatrix} \begin{bmatrix} \vec{V}_d \\ 1 \end{bmatrix} = 0 \quad (187)$$

- Quant à l'erreur rétrograde, on a :

$$\begin{bmatrix} A_2 + \Delta A_2 & B_2 + \Delta B_2 \end{bmatrix} \begin{bmatrix} \vec{V}_r \\ 1 \end{bmatrix} = 0 \quad (188)$$

Les expressions de  $A_1$  et  $A_2$  sont données par l'équation (169). Celles de  $B_1$  et  $B_2$  sont présentées par l'équation (170). Comme dans la résolution au sens des moindres carrés de l'approche basée sur un schéma direct-rétrograde, deux formulations de fonctionnelle sont possibles à partir des équations (187) et (188) :

\*La première formulation consiste à créer indépendamment deux termes quadratiques. Le premier résume l'erreur directe et le second est associé à l'erreur rétrograde. La somme de ces deux termes donne le critère global à minimiser par la méthode du tenseur. C'est l'équivalent de la fonctionnelle à somme des termes quadratiques des erreurs. L'approche direct-rétrograde associée à la méthode du tenseur cherche donc à minimiser l'erreur :

$$\min(J_a) = \min \left\{ \left\| W_1^{1/2} \begin{bmatrix} \Delta A_1 & \Delta B_1 \end{bmatrix} \right\|_F^2 + \left\| W_2^{1/2} \begin{bmatrix} \Delta A_2 & \Delta B_2 \end{bmatrix} \right\|_F^2 \right\} \quad (189)$$

tel que :

$$(A_1 + \Delta A_1) \vec{V} = -(B_1 + \Delta B_1) \quad \text{et} \quad (A_2 + \Delta A_2) \vec{V} = -(B_2 + \Delta B_2) \quad (190)$$

Les matrices diagonales  $W_1$  et  $W_2$  forment la structure des poids. La procédure de minimisation de la fonctionnelle peut se traduire par :

$$\min(J_a) = \min \left\{ \text{tr} \left( \begin{bmatrix} A_1 & B_1 \end{bmatrix}^T W_1 \begin{bmatrix} A_1 & B_1 \end{bmatrix} \right) + \text{tr} \left( \begin{bmatrix} A_2 & B_2 \end{bmatrix}^T W_2 \begin{bmatrix} A_2 & B_2 \end{bmatrix} \right) \right\} \quad (191)$$

La décomposition en éléments propres des matrices de données est décrite comme suit :

$$\begin{bmatrix} A_1 & B_1 \end{bmatrix}^T W_1 \begin{bmatrix} A_1 & B_1 \end{bmatrix} + \begin{bmatrix} A_2 & B_2 \end{bmatrix}^T W_2 \begin{bmatrix} A_2 & B_2 \end{bmatrix} = U \Sigma_{3 \times 3} U^T \quad (192)$$

L'élément  $\Sigma$  correspond à une matrice diagonale des valeurs propres non nulles, résultat de la décomposition en éléments propres de la matrice des données (tout ce qui est en premier membre de l'équation (192)).

$$\Sigma = \text{diag}(\lambda_1, \lambda_2, \lambda_3) \quad (193)$$

La matrice  $U$  est constituée par les vecteurs propres  $\vec{e}_j, j = 1 \dots 3$ , associés aux valeurs propres  $\lambda_j$ .

<sup>10</sup> FBTLs : Forward backward total least squares

Selon la résolution au sens des moindres carrés totaux [HOL77], la solution de la méthode du tenseur à mouvement direct-rétrograde, pour un modèle de translation, peut être exprimée par :

$$\vec{V} = -\frac{1}{e_p(3)} \begin{bmatrix} e_p(1) \\ e_p(2) \end{bmatrix} \quad (194)$$

où  $\vec{e}_p$  est le vecteur propre associé au minimum des valeurs propres.

Comme on est dans un contexte de modèle de translation, la minimisation du critère  $J_a$  ne fournit pas en conséquence deux vecteurs vitesse différents. Le terme  $\vec{V}$  est une notation qui désigne l'estimé commun issu de cette minimisation.

\*La deuxième fonctionnelle consiste à manipuler une erreur globale qui est la somme algébrique de deux erreurs. Ainsi, le critère à minimiser s'obtient à l'aide de la norme de Frobenius, résultante de l'erreur globale. Nous avons donc la fonctionnelle définie par le carré de la somme de cette norme :

$$\min(J_b) = \min \left\{ \left\| W^{1/2} \left( \begin{bmatrix} \Delta A_1 & \Delta B_1 \end{bmatrix} + \begin{bmatrix} \Delta A_2 & \Delta B_2 \end{bmatrix} \right) \right\|_F^2 \right\} \quad (195)$$

avec :

$$\left( A_1 + A_2 + \Delta A_1 + \Delta A_2 \right) \vec{V} = - \left( B_1 + B_2 + \Delta B_1 + \Delta B_2 \right) \quad (196)$$

$\begin{bmatrix} \Delta A_1 & \Delta B_1 \end{bmatrix}$  et  $\begin{bmatrix} \Delta A_2 & \Delta B_2 \end{bmatrix}$  ont respectivement la même structure que  $\begin{bmatrix} A_1 & B_1 \end{bmatrix}$  et  $\begin{bmatrix} A_2 & B_2 \end{bmatrix}$ . Comme avec la première fonctionnelle, la minimisation décrite dans (195) conduit à écrire :

$$\min(J_b) = \min \operatorname{tr} \left\{ \left( \begin{bmatrix} A_1 & B_1 \end{bmatrix} + \begin{bmatrix} A_2 & B_2 \end{bmatrix} \right)^T W \left( \begin{bmatrix} A_1 & B_1 \end{bmatrix} + \begin{bmatrix} A_2 & B_2 \end{bmatrix} \right) \right\} \quad (197)$$

La décomposition en éléments propres de la matrice de données s'écrit :

$$\left( \begin{bmatrix} A_1 & B_1 \end{bmatrix} + \begin{bmatrix} A_2 & B_2 \end{bmatrix} \right)^T W \left( \begin{bmatrix} A_1 & B_1 \end{bmatrix} + \begin{bmatrix} A_2 & B_2 \end{bmatrix} \right) = U \Sigma_{3 \times 3} U^T \quad (198)$$

La solution peut être obtenue de la même manière que l'équation (194).

## 1.6.2 Résultats expérimentaux

### a) Séquence synthétique : "spirale"

Ce paragraphe décrit les résultats obtenus sur la séquence "spirale". L'expérimentation permet d'évaluer la fiabilité du formalisme direct-rétrograde pour l'approche du tenseur. Les résultats sont présentés sous la forme des courbes de biais d'estimation (figures 2.30 à 2.33). Le support de calcul fait la même taille que celui utilisé dans l'approche du tenseur basée sur un schéma simple ( $13 \times 13$  pixels). La séquence est bruitée par un bruit gaussien additif.

Chaque point représente le biais moyen relevé sur un ensemble de pixels. Le calcul s'effectue en faisant dérouler plusieurs tests et en prenant la moyenne des valeurs estimées sur le lot de tests. Comme le montrent les figures 2.30 et 2.31, les courbes des erreurs d'estimation ne sont pas stables pour des bruits dont le SNR est inférieur à  $70 \text{ dB}$ . On remarque la présence d'un point critique à  $50 \text{ dB}$ . L'intervalle de confiance ne peut être pris qu'à partir d'un SNR de  $70 \text{ dB}$ . On constate que les erreurs produites par le formalisme direct-rétrograde ( $J_a$  ou  $J_b$ ) restent inférieures à celles du formalisme classique.

En ce qui concerne la comparaison entre eux de deux critères  $J_a$  et  $J_b$ , on remarque que les résultats obtenus avec la fonctionnelle  $J_b$  restent meilleurs.

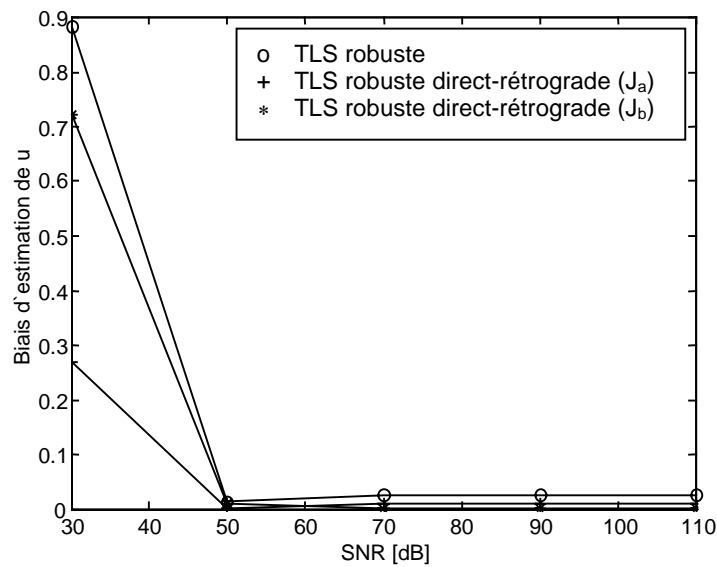


Figure 2.30 : Biais d'estimation pour la composante  $u$  du vecteur vitesse de la séquence "spirale" ; méthode TLS robuste basée sur un schéma simple, méthode TLS robuste basée sur un schéma direct-rétrograde ( $J_a$ ) et méthode TLS robuste basée sur un schéma direct-rétrograde ( $J_b$ ) ; résultats au bout de 3 pondérations.

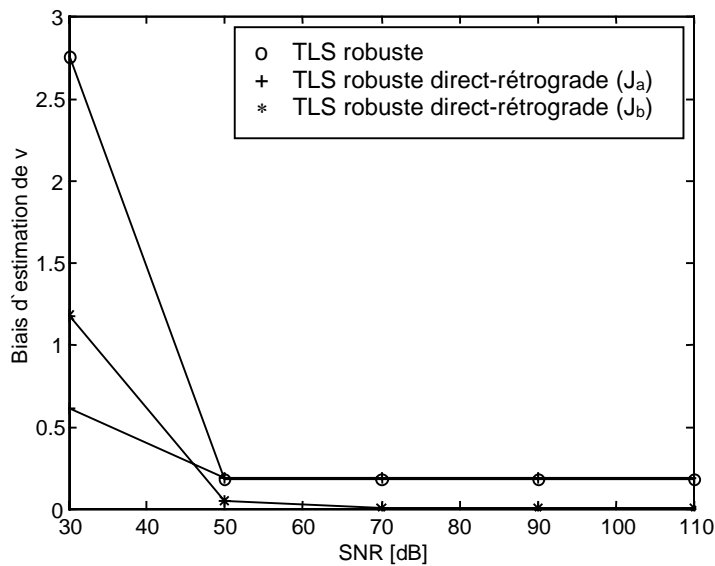


Figure 2.31 : Biais d'estimation pour la composante  $v$  du vecteur vitesse de la séquence "spirale" ; méthode TLS robuste basée sur un schéma simple, méthode TLS robuste basée sur un schéma direct-rétrograde ( $J_a$ ) et méthode TLS robuste basée sur un schéma direct-rétrograde ( $J_b$ ) ; résultats au bout de 3 pondérations.

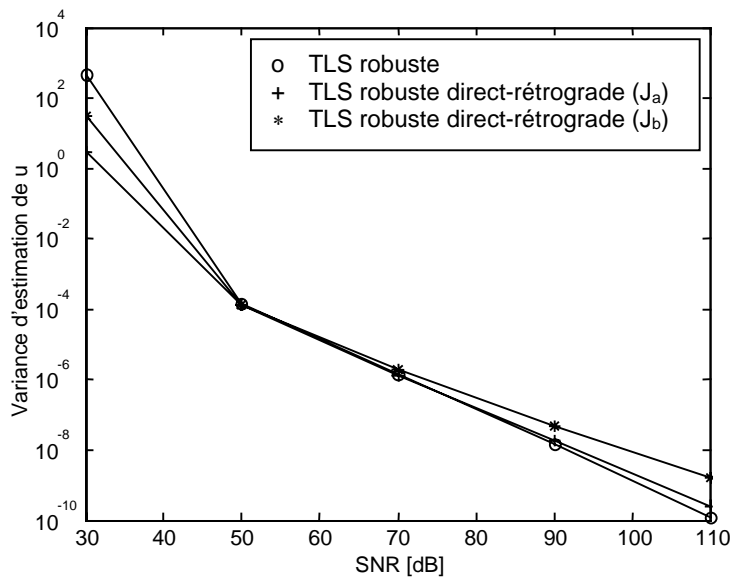


Figure 2.32 : Variance de l'erreur d'estimation pour la composante  $u$  du vecteur vitesse de la séquence "spirale" ; méthode TLS robuste basée sur un schéma simple, méthode TLS robuste basée sur un schéma direct-rétrograde ( $J_a$ ) et méthode TLS robuste basée sur un schéma direct-rétrograde ( $J_b$ ) ; résultats au bout de 3 pondérations.

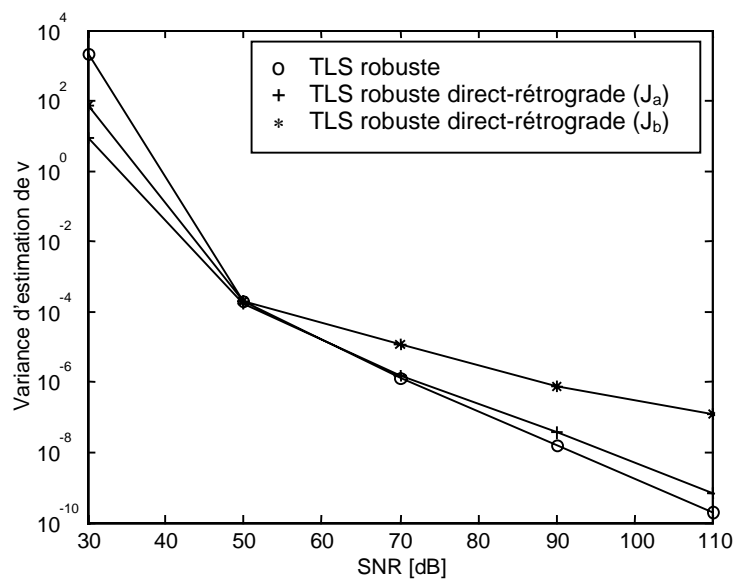


Figure 2.33 : Variance de l'erreur d'estimation pour la composante  $v$  du vecteur vitesse de la séquence "spirale" ; méthode TLS robuste basée sur un schéma simple, méthode TLS robuste basée sur un schéma direct-rétrograde ( $J_a$ ) et méthode TLS robuste basée sur un schéma direct-rétrograde ( $J_b$ ) ; résultats au bout de 3 pondérations.

### b) Séquence réelle à déplacement synthétique

Pour examiner les performances de l'approche des moindres carrés totaux à base de schéma direct-rétrograde, nous proposons des résultats pour lesquels une observation qualitative peut être faite. L'expérimentation est effectuée sur une séquence réelle à déplacement synthétique. Nous rappelons qu'une translation de  $(-2,1)$  pixels a été affectée à l'objet à tester. Un bruit blanc gaussien de  $100\text{ dB}$  a aussi été rajouté à la séquence.

Afin de réaliser un maximum de convergence, nous combinons simultanément les méthodes LS et TLS pondérées dans un même algorithme. Pour cela, l'algorithme s'initialise avec la méthode TLS pondérée et quand au bout d'un certain nombre d'itérations, la solution a tendance à diverger, nous faisons intervenir la méthode LS pondérée. Cette technique permet de révéler l'instabilité de l'approche TLS.

La figure 2.34 présente la détection et l'estimation du mouvement de l'objet  $\textcircled{1}$ . Un agrandissement la zone  $S$  est reporté à la figure 2.35. Les deux méthodes LS et TLS sont ici à base de schéma direct.

La même procédure de test, avec l'approche basée sur un schéma direct-rétrograde, a été appliquée sur la même séquence. L'algorithme s'opère avec la méthode TLS robuste à base de flot direct-rétrograde. Nous avons pris dans ce test la deuxième fonctionnelle ( $J_b$ ). En cas de divergence, l'opération est basculée en LS robuste à flot direct-rétrograde. La figure 2.36 correspond à la détection et l'estimation du mouvement de l'objet  $\textcircled{1}$ . L'agrandissement de la zone  $S$  est représenté à la figure 2.37.

On peut voir sur la figure 2.35 que les déplacements sont fortement biaisés. L'utilisation de la technique de pondération dans les approches LS et TLS à flot direct ne suffit pas pour avoir des résultats robustes. Sur la figure 2.37, on peut constater une forte amélioration. L'utilisation du formalisme basé sur un schéma direct-rétrograde avec les méthodes LS et TLS pondérées, permet de rendre les résultats stables.

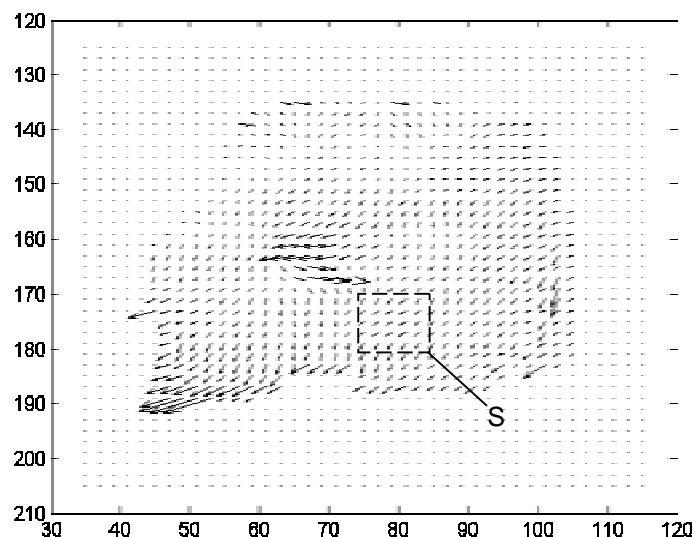


Figure 2.34 : Détection et estimation du mouvement de l'objet  $\textcircled{1}$  de la séquence réelle à déplacement synthétique ; séquence bruitée par un bruit gaussien additif ( $SNR = 100\text{ dB}$ ) ; combinaison de deux méthodes des moindres carrés robustes basées sur un schéma simple (TLS+LS) ; déplacement théorique de  $(-2,1)$  pixels ; résultats au bout de 3 pondérations.



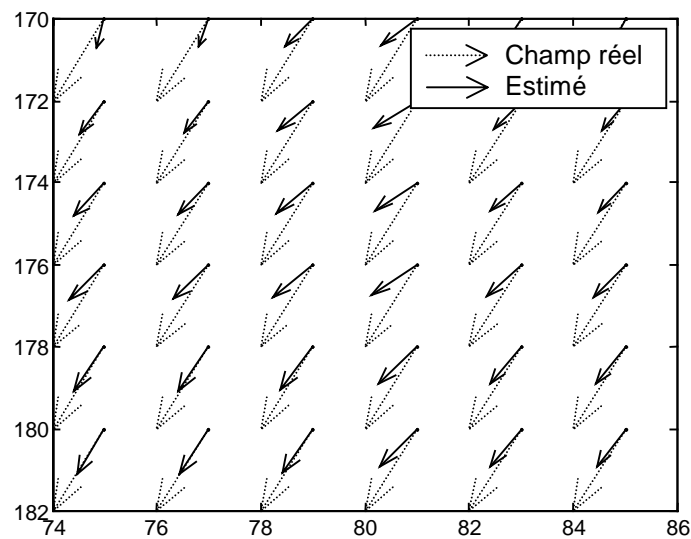


Figure 2.35 : Extrait du champ de déplacement obtenu par la combinaison des méthodes LS+TLS simples ; agrandissement de la zone  $S$ .

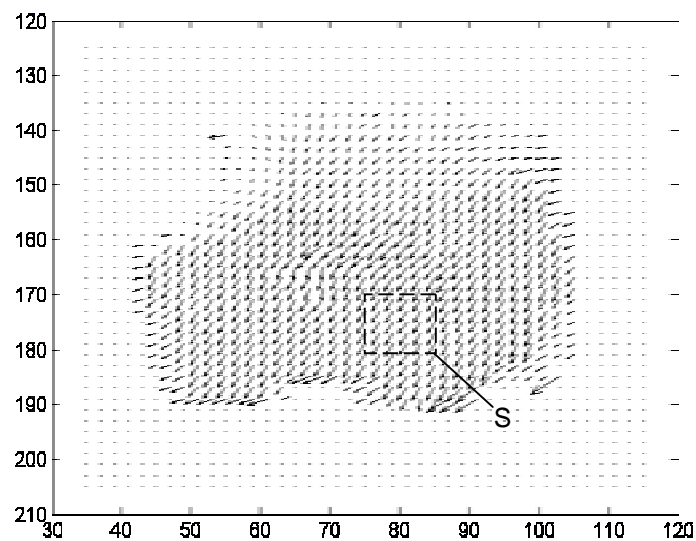


Figure 2.36 : Détection et estimation de l'objet  $\textcircled{1}$  de la séquence réelle à déplacement synthétique ; séquence bruitée par un bruit gaussien additif ( $SNR = 100 \text{ dB}$ ) ; combinaison de deux méthodes LS+TLS robustes basées sur un schéma direct-rétrograde ; déplacement théorique de  $(-2,1)$  pixels ; résultats au bout de 3 pondérations .

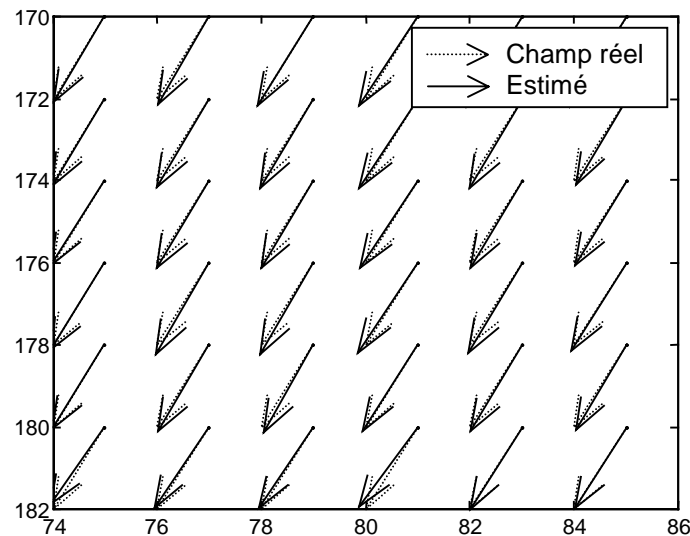


Figure 2.37 : Extrait du champ de déplacement obtenu par la combinaison des méthodes LS+TLS basées sur un schéma direct-rétrograde ; agrandissement de la zone  $S$ .

### 1.6.3 Discussion sur la méthode du tenseur robuste basée sur un schéma direct-rétrograde

Nous avons testé dans ce paragraphe, la méthode du tenseur (moindres carrés totaux<sup>11</sup>) associée au formalisme direct-rétrograde. Une technique de pondération des gradients est intégrée dans l'algorithme. Pour montrer l'intérêt du formalisme proposé dans le contexte de l'approche du tenseur, nous avons essayé de l'appliquer, en premier lieu, sur une image synthétique. A partir des résultats tracés, nous pouvons conclure statistiquement que la prise en compte de ce formalisme permet de réduire nettement le biais d'estimation.

Pour valider le test avec d'autres séquences, nous avons pris une séquence réelle à déplacement synthétique. Les résultats obtenus attestent de l'importance de ce formalisme. Une perspective afin d'améliorer les résultats serait donc d'utiliser trois images successives. L'ajout d'une troisième image permettrait alors de limiter les défauts inhérents aux techniques d'approximation des gradients temporels.

Une autre perspective consiste à introduire un critère de stabilité en analysant les valeurs propres de la matrice des données. Cette technique contribue au critère de validation des résultats de la méthode du tenseur [JAH97].

## 1.7 Approche incrémentale à formalisme direct-rétrograde

### 1.7.1 Rappel sur l'approche incrémentale

L'approche incrémentale s'appuie également sur l'hypothèse de conservation de la luminance. Nous rappelons ici l'expression qui traduit cette hypothèse pour un déplacement théorique :

$$I(x, y, t) = I(x + \hat{dx}, y + \hat{dy}, t + dt) \quad (199)$$

Comment peut-on interpréter cette équation ?

<sup>11</sup> Total least squares

a) Cette équation veut dire que si on fait décaler l'intensité  $I(x + \hat{d}x, y + \hat{d}y, t + dt)$  d'une quantité  $-\hat{d}$ , l'intensité ainsi obtenue, notée  $I_{\text{décalée}}(t + dt)$ , coïncide avec l'intensité  $I(x, y, t)$ , notée  $I(t)$ .

b) Supposons qu'on ait estimé un déplacement  $d$ . Si l'estimé est bon ( $d = \hat{d}$ ), on doit tomber sur l'image  $I(t)$  lorsqu'on fait décaler l'intensité  $I(t + dt)$  d'une même quantité  $-d$ . Dans le cas contraire, c'est-à-dire  $d \neq \hat{d}$ , l'intensité  $I_{\text{décalée}}(t + dt)$  ne coïncide pas avec  $I(t)$ . Cela veut dire que le déplacement estimé présente une erreur  $\delta d$  par rapport au déplacement théorique. On a :

$$\hat{d} = d + \delta d \quad (200)$$

c) Si on continue le raisonnement, rien n'empêche de dire que si on fait décaler l'image  $I_{\text{décalée}}(t + dt)$  d'une quantité  $-\delta d$ , on doit tomber sur l'image  $I(t)$ .

Il est évident que dans b), on utilise l'intensité  $I(t + dt)$  et l'intensité  $I(t)$  pour le calcul de  $d$ . Par analogie, pour le calcul de  $\delta d$  dans c), on utilise l'intensité déplacée  $I_{\text{décalée}}(t + dt)$  et l'intensité  $I(t)$ .

L'interprétation de l'équation de contrainte du mouvement est représentée à la figure 2.38 et nous interprétons l'équation de contrainte déplacée sur la figure 2.39. L'axe vertical représente la fonction de luminosité et celui horizontal la position d'un pixel  $p(x, y)$ .

Sur la figure 2.39, nous voyons que pour un déplacement de forte amplitude, la résolution de l'équation de contrainte peut fournir un résultat incorrect. En effet, quand le déplacement est faible, l'écart entre  $C$  et  $G$  devient négligeable. Ce qui n'est pas forcément vrai pour un déplacement important car dans ce cas, l'erreur est notable. Le rattrapage de l'erreur s'effectue en introduisant un incrément  $\delta d$ . Par analogie avec le calcul de  $d$ , la quantité  $\delta d$  est aussi exploitable par une formulation des moindres carrés, en utilisant l'intensité décalée  $I_{\text{décalée}}(t + dt)$ . Sa mise en équation est décrite ci-dessous. La solution proposée permet d'atteindre un déplacement de l'ordre d'un sous-multiple de pixel.

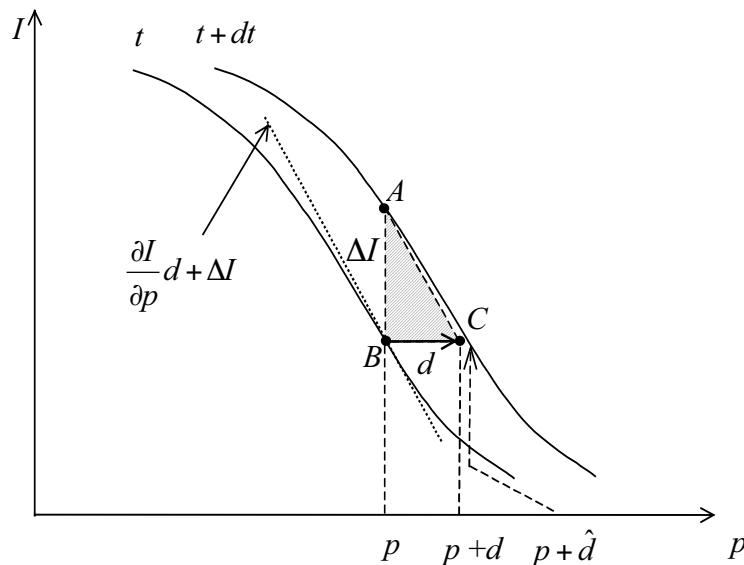


Figure 2.38 : Interprétation de l'ECM.

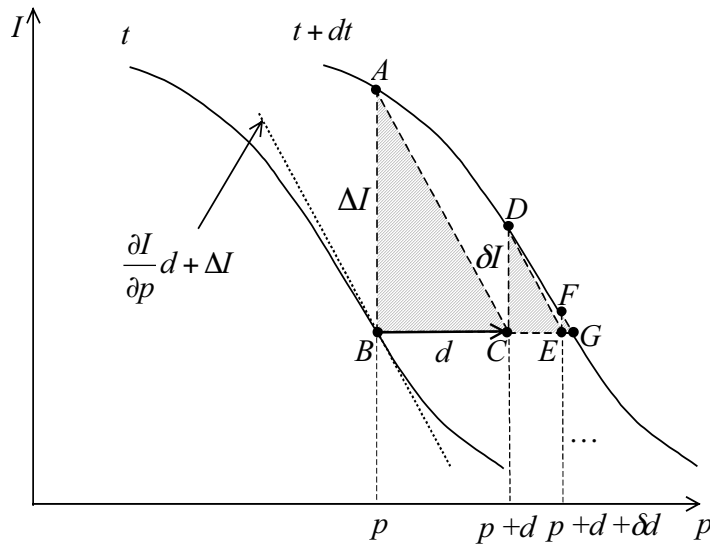


Figure 2.39 : Interprétation de l'ECM décalée.

La particularité de l'approche incrémentale se trouve dans l'utilisation de cette intensité déplacée. Le but est de rattraper l'écart  $\delta d$  d'une manière itérative en introduisant le dernier déplacement estimé dans l'intensité  $I_{\text{décalée}}(t + dt)$ . Soit :

$$I_{\text{décalée}}(t + dt) = I(p + \vec{d} + \delta \vec{d}, t + dt) \quad (201)$$

A la  $k^{\text{ème}}$  itération,  $\delta \vec{d}^{(k)}$  est estimé en utilisant l'image  $I(t)$  et l'image  $I_{\text{décalée}}(t + dt)$  avec un décalage de  $\vec{d}^{(k)}$ . Dans l'approche incrémentale, nous avons besoin d'atteindre un sous-pixel plus exactement. En utilisant la méthode des moindres carrés pour estimer le déplacement, nous décrivons la méthode comme suit :

Un déplacement initial est estimé à l'aide de la méthode des moindres carrés. Le critère à minimiser s'écrit :

$$J = \sum_{\Omega} [I(x + dx, y + dy, t + dt) - I(x, y, t)]^2 \quad (202)$$

En terme de vitesse, le développement de Taylor à l'ordre un de l'équation (202) donne :

$$J = \sum_{\Omega} \left[ \vec{\nabla} I^t(x, y) \cdot \vec{V} + \frac{\partial I^t}{\partial t} \right]^2 \quad (203)$$

$\vec{\nabla} I^t(x, y)$  désigne les gradients spatiaux de l'image à l'instant  $t$ . Cette quantité correspond à la tangente de l'angle  $C$  du triangle  $ABC$  ou bien le coefficient de la droite de contrainte (tangente en  $x$  de  $I(t)$ ). La quantité  $\vec{V} = (dx/dt, dy/dt)^T$  est le vecteur vitesse. La minimisation de  $J$  au sens des moindres carrés donne :

$$\vec{V} = -(A^T A)^{-1} A^T B \quad (204)$$

La matrice  $A$  contient les gradients spatiaux et  $B$  le gradients temporels (c'est la quantité  $\Delta I$ ). Jusqu'à ce niveau, notre raisonnement se situe autour du triangle  $ABC$ . L'étape suivante consiste en une amélioration du résultat, en passant au triangle  $DCE$ . Le déplacement est optimisé en l'affinant par un terme de correction.

$$\vec{d}^{(k+1)} = \vec{d}^{(k)} + \delta \vec{d}^{(k)} \quad (205)$$

Les grandeurs  $\vec{d}^{(k)}$  et  $\vec{V}^{(k)}$  sont liées par le temps  $dt$ . Elles sont équivalentes pour un intervalle de temps égal à l'unité de mesure. Il en est de même pour  $\delta \vec{d}^{(k)}$  et  $\delta \vec{V}^{(k)}$ .

Le critère (202) peut s'écrire :

$$J^{(k)} = \sum_{\Omega} \left[ I(x + dx^{(k)} + \delta dx^{(k)}, y + dy^{(k)} + \delta dy^{(k)}, t + dt) - I(x, y, t) \right]^2 \quad (206)$$

Comme avec l'équation (203), le développement de l'équation (206) par rapport à  $\delta \vec{d}$  permet d'écrire l'équation suivante :

$$J^{(k)} = \sum_{\Omega} \left\{ \vec{\nabla} I^{t+dt}(x + dx^{(k)}, y + dy^{(k)}) \delta \vec{V}^{(k)} + \left[ I(x + dx^{(k)}, y + dy^{(k)}, t + dt) - I(x, y, t) \right] \right\}^2 \quad (207)$$

La quantité  $\vec{\nabla} I^{t+dt}(x + dx^{(k)}, y + dy^{(k)})$  désigne les gradients spatiaux de l'image déplacée à l'instant  $t + dt$ . Cette quantité correspond à la tangente de l'angle  $E$  du triangle  $DCE$  ; elle n'est que la tangente en  $x + d$  de la fonction  $I_{t+dt}$ . La quantité  $\delta \vec{V}$  désigne le terme d'incrément en vecteur vitesse, associé au terme  $\delta \vec{d}$ .

Comme les composantes du déplacement  $(dx, dy)^T$  ne sont pas nécessairement à valeur entière, il est nécessaire d'interpoler les gradients  $\vec{\nabla} I^{t+dt}(x + dx^{(k)}, y + dy^{(k)})$  et la fonction d'intensité  $I(x + dx^{(k)}, y + dy^{(k)}, t + dt)$ . Pour cela, il existe plusieurs techniques : interpolation par le plus proche voisin, interpolation bilinéaire, interpolation bi-cubique, interpolation par Splines, B-Splines. Dans notre étude, nous avons retenu l'approche bilinéaire à cause de sa simplicité, et parce que les autres approches n'amènent pas de gain significatif.

Nous pouvons remarquer que l'estimé  $\vec{V} = (dx/dt, dy/dt)^T$  à la  $k^{\text{ème}}$  itération contribue au calcul du terme correcteur  $\delta \vec{V}$ . A cet effet, à chaque itération, la position de l'intensité  $I(t + dt)$  est incrémentée (ou décrémentée) pour être en concordance avec l'intensité  $I(t)$ . La résolution de l'équation (207) par rapport à la variable  $\delta \vec{V}$  donne :

$$\delta \vec{V}^{(k)} = -(C^T C)^{-1} C^T D \quad (208)$$

La matrice  $C$  des gradients spatiaux  $\vec{\nabla} I^{t+dt}(x + dx^{(k)}, y + dy^{(k)})$  est donnée par :

$$C = \begin{bmatrix} I_x^{t+Te}(x_1 + dx^{(k)}, y_1 + dy^{(k)}) & I_y^{t+Te}(x_1 + dx^{(k)}, y_1 + dy^{(k)}) \\ \vdots & \vdots \\ I_x^{t+Te}(x_n + dx^{(k)}, y_n + dy^{(k)}) & I_y^{t+Te}(x_n + dx^{(k)}, y_n + dy^{(k)}) \end{bmatrix} \quad (209)$$

Nous rappelons que la notation utilisée dans l'expression de  $C$  est une notation discrète. C'est pourquoi nous remplaçons  $t + dt$  par  $t + Te$ .

La matrice  $D$  correspond à la quantité  $\delta I$  ; son expression à la  $k^{\text{ème}}$  itération est de la forme :

$$D = \begin{bmatrix} I(x_1 + dx^{(k)}, y_1 + dy^{(k)}, t + Te) - I(x_1, y_1, t) \\ \vdots \\ I(x_n + dx^{(k)}, y_n + dy^{(k)}, t + Te) - I(x_n, y_n, t) \end{bmatrix} \quad (210)$$

La première itération sur le calcul de  $\delta \vec{d}$  (ou  $\delta \vec{V}$ ) s'achève à ce niveau. Le raisonnement est situé au niveau du triangle  $DCE$ . Le passage au triangle  $EFG$  fait l'objet de l'itération suivante et l'opération se poursuit jusqu'à ce qu'on atteigne un triangle infinitésimal.

Nous tenons à préciser que le calcul de  $\vec{d}^{(k)}$  (ou  $\vec{V}^{(k)}$ ) et de  $\delta \vec{d}^{(k)}$  (ou  $\delta \vec{V}^{(k)}$ ) peut être accompagné d'une méthode de pondération. Nous en parlerons dans le formalisme direct-rétrograde de l'approche incrémentale.

### 1.7.2 Approche incrémentale selon le formalisme direct-rétrograde

Ce paragraphe décrit les différentes possibilités de l'exploitation de l'approche incrémentale associée au mouvement direct-rétrograde. Le principe est plus ou moins le même que le précédent sauf qu'ici, la correction opère dans les deux sens temporels [RAN00c]. L'un est par rapport au déplacement direct et l'autre par rapport au déplacement rétrograde.

\* Nous allons commencer par la fonctionnelle des erreurs à somme des termes quadratiques. Les différentes étapes d'estimation sont illustrées à la figure 2.40.

A la  $k^{\text{ème}}$  itération, les matrices de données sont exprimées comme suit :

-  $C_1$  exprime les gradients spatiaux de l'image  $t$ , décalée conformément au rattrapage de l'information suivant le sens descendant du temps (rétrograde) :

$$C_1 = \begin{bmatrix} I_x^t(x_1 - dx^{(k)}, y_1 - dy^{(k)}) & I_y^t(x_1 - dx^{(k)}, y_1 - dy^{(k)}) \\ \vdots & \vdots \\ I_x^t(x_n - dx^{(k)}, y_n - dy^{(k)}) & I_y^t(x_n - dx^{(k)}, y_n - dy^{(k)}) \end{bmatrix} \quad (211)$$

-  $C_2$  exprime les gradients spatiaux de l'image  $t + dt$ , décalée conformément au rattrapage de l'information suivant le sens ascendant du temps (direct) :

$$C_2 = \begin{bmatrix} I_x^{t+Te}(x_1 + dx^{(k)}, y_1 + dy^{(k)}) & I_y^{t+Te}(x_1 + dx^{(k)}, y_1 + dy^{(k)}) \\ \vdots & \vdots \\ I_x^{t+Te}(x_n + dx^{(k)}, y_n + dy^{(k)}) & I_y^{t+Te}(x_n + dx^{(k)}, y_n + dy^{(k)}) \end{bmatrix} \quad (212)$$

La notation  $t + dt$  est remplacée par  $t + Te$  en notation discrète des gradients.

- La matrice  $D_1$  correspond à la DFD associée au mouvement rétrograde :

$$D_1 = \begin{bmatrix} I(x_1 - dx^{(k)}, y_1 - dy^{(k)}, t) - I(x_1, y_1, t + dt) \\ \vdots \\ I(x_n - dx^{(k)}, y_n - dy^{(k)}, t) - I(x_n, y_n, t + dt) \end{bmatrix} \quad (213)$$

- La matrice  $D_2$  contient les composantes de la DFD associée au mouvement direct :

$$D_2 = \begin{bmatrix} I(x_1 + dx^{(k)}, y_1 + dy^{(k)}, t + dt) - I(x_1, y_1, t) \\ \vdots \\ I(x_n + dx^{(k)}, y_n + dy^{(k)}, t + dt) - I(x_n, y_n, t) \end{bmatrix} \quad (214)$$

Nous rappelons que les composantes du déplacement  $(dx, dy)^T$  ne sont pas nécessairement à valeur entière, il est donc nécessaire d'interpoler ces matrices de données. L'interpolation bilinéaire est utilisée dans notre approche.

Comme nous l'avons précisé au passage, l'application du formalisme direct-rétrograde ne fournit pas en conséquence deux déplacements différents, mais plutôt un déplacement commun à ces deux informations. C'est pour cela que nous avons utilisé une notation unique  $\vec{\delta d}$  pour le terme d'incrément.

\* Les mêmes étapes d'estimation peuvent aussi être applicables à la fonctionnelle  $J_b$ . Nous soulignons les points suivants :

- On commence par l'estimation du déplacement initial. La méthode des moindres carrés basée sur un schéma direct-rétrograde peut être utilisée en minimisant le critère suivant :

$$J_b = \sum_{\Omega} \left\{ [I(x + dx_d, y + dy_d, t + dt) - I(x, y, t)] + [I(x + dx_r, y + dy_r, t) - I(x, y, t + dt)] \right\}^2 \quad (215)$$

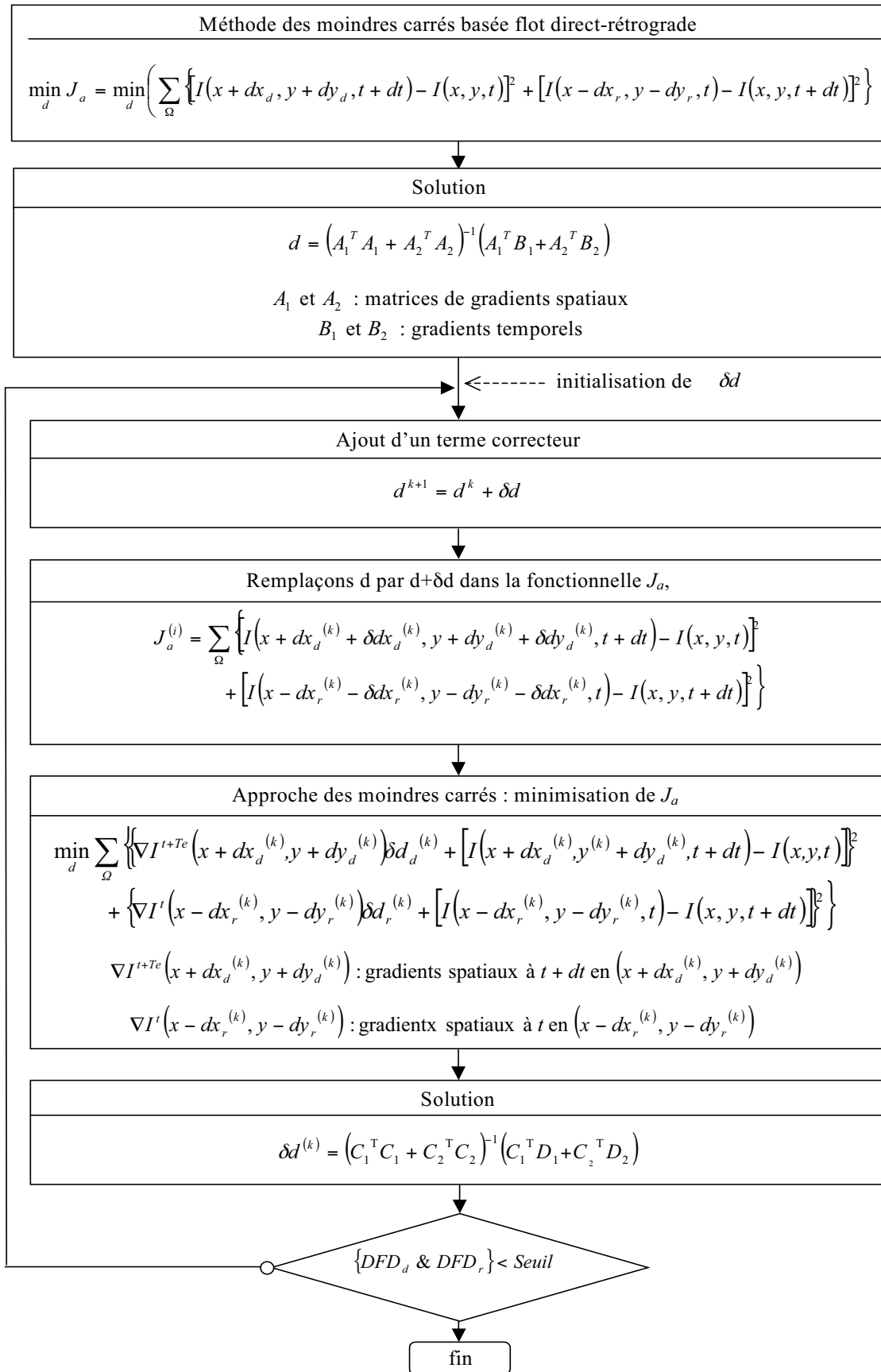


Figure 2.40 : Approche incrémentale basée sur un mouvement direct-rétrograde.

- La solution est de la forme :

$$(u, v)^T = -[(A_1 + A_2)^T (A_1 + A_2)]^{-1} [(A_1 + A_2)(B_1 + B_2)] \quad (216)$$

$A_1$  et  $A_2$  sont respectivement les matrices de gradients spatiaux de l'image à l'instant  $t$  et l'image à l'instant  $t + dt$  ;  $B_1$  et  $B_2$  sont les gradients temporels. Leurs expressions sont données par les relations (169, 170).

- Optimisation du résultat à l'aide d'un procédé itératif :

$$\vec{d}^{(k+1)} = \vec{d}^{(k)} + \delta\vec{d}^{(k)} \quad (217)$$

Les grandeurs  $\vec{d}^{(k)} = (dx^{(k)}, dy^{(k)})$  et  $\vec{V}^{(k)} = (u^{(k)}, v^{(k)})$  sont liées par le terme  $dt$ . Elles sont équivalentes pour un intervalle de temps égal à l'unité de mesure. Il en est de même pour  $\delta\vec{d}^{(k)} = (\delta dx^{(k)}, \delta dy^{(k)})$  et  $\delta\vec{V}^{(k)} = (\delta u^{(k)}, \delta v^{(k)})$ .

- En remplaçant  $\vec{d}$  par  $\vec{d} + \delta\vec{d}$ , la fonctionnelle à minimiser peut s'écrire :

$$J_b = \sum_{\Omega} \left\{ I(x + dx_d^{(k)} + \delta dx_d^{(k)}, y + dy_d^{(k)} + \delta dy_d^{(k)}, t + dt) - I(x, y, t) \right\} \\ + \left\{ I(x - dx_r^{(k)} - \delta dx_r^{(k)}, y - dy_r^{(k)} - \delta dy_r^{(k)}, t) - I(x, y, t + dt) \right\}^2 \quad (218)$$

- Après avoir effectué un développement de Taylor au premier ordre par rapport à la quantité  $\delta\vec{d}^{(k)}$ , la fonctionnelle  $J_b$  à la  $k^{\text{ème}}$  itération peut s'écrire sous la forme :

$$J_b^{(k)} = \sum_{\Omega} \left\{ \vec{\nabla} I^{t+dt}(x + dx_d^{(k)}, y + dy_d^{(k)}) + \vec{\nabla} I^t(x - dx_r^{(k)}, y - dy_r^{(k)}) \right\} \delta\vec{V}^{(k)} \\ + \left[ I(x + dx_d^{(k)}, y + dy_d^{(k)}, t + dt) - I(x, y, t) \right] + \left[ I(x - dx_r^{(k)}, y - dy_r^{(k)}, t) - I(x, y, t + dt) \right]^2 \quad (219)$$

- La résolution de cette fonctionnelle par rapport à  $\delta\vec{V}$  et au sens des moindres carrés donne :

$$\delta\vec{V}^{(k)} = -[(C_1 + C_2)^T (C_1 + C_2)]^{-1} [(C_1 + C_2)^T (D_1 + D_2)] \quad (220)$$

Les matrices des données à la  $k^{\text{ème}}$  itération ont les mêmes expressions que les précédentes.

### 1.7.3 Approche incrémentale robuste selon le formalisme direct-rétrograde

Dans cette méthode, une technique robuste est intervenue à la fois sur l'estimation du déplacement  $\vec{d}$  et celle du terme correctif  $\delta\vec{d}$ .

\* Considérons le cas de la fonctionnelle à somme quadratique des erreurs ( $J_a$ ). Les étapes à suivre sont les suivantes :

- *Étape n°1* : estimation du déplacement initial. Le résultat peut être obtenu en utilisant l'approche incrémentale basée sur un schéma direct-rétrograde.

- *Étape n°2* : pondération des matrices de gradients pour améliorer l'estimé initial. On calcule le résidu des erreurs selon la procédure décrite au paragraphe 1.5.1, avec :

$$\rho(\varepsilon_d) = [I(x + dx_d, y + dy_d, t + dt) - I(x, y, t)]^2 \quad (221)$$

et :

$$\rho(\varepsilon_r) = [I(x - dx_r, y - dy_r, t) - I(x, y, t + dt)]^2 \quad (222)$$

Le résultat issu de la première pondération s'écrit :

$$(u, v)^T = -(A_1^T W_1 A_1 + A_2^T W_2 A_2)^{-1} (A_1^T W_1 B_1 + A_2^T W_2 B_2) \quad (223)$$

La structure des matrices de poids  $W$  a été présentée dans la méthode des moindres carrés robuste à formalisme direct-rétrograde.

- *Étape n°3* : procédé itératif de l'approche incrémentale ; le détail de cette étape a été présenté dans le paragraphe précédent.



- *Étape n°4* : introduction d'une deuxième pondération pour améliorer le terme correctif  $\delta\vec{d}$ . Pour le calcul des poids, nous préférons réécrire ici la fonction d'erreur résiduelle associée au critère utilisé :

$$\sum_{\Omega} [\rho(\varepsilon'_d) + \rho(\varepsilon'_r)] = \sum_{\Omega} \left( \frac{1}{2} \alpha'_d \varepsilon'^2_d + \frac{1}{2} \alpha'_r \varepsilon'^2_r \right) \quad (224)$$

Nous utilisons la notation «'» pour signifier que les notations sont liées au terme correctif  $\delta\vec{d}$  mais non pas par rapport à  $\vec{d}$ .

A une itération donnée, l'expression de l'erreur résiduelle associée au terme de correction  $\delta\vec{d}_d$ , suivant le sens temporel direct, est donnée par :

$$\rho(\varepsilon'_d) = I(x + dx_d + \delta dx_d, y + dy_d + \delta dy_d, t + dt) - I(x, y, t) \quad (225)$$

tandis que celle de la correction  $\delta\vec{d}_r$  suivant le sens temporel rétrograde est :

$$\rho(\varepsilon'_r) = I(x - dx_r - \delta dx_r, y - dy_r - \delta dy_r, t) - I(x, y, t + dt) \quad (226)$$

Le poids attribué à chaque point  $(x + dx_d, y + dy_d, t + dt)$  est calculé selon l'équation :

$$\alpha'_d = \frac{1}{\varepsilon'_d} \frac{\partial \rho(\varepsilon'_d)}{\partial \varepsilon'_d} \quad (227)$$

Le poids attribué à chaque point  $(x - dx_r, y - dy_r, t)$  peut s'exprimer par :

$$\alpha'_r = \frac{1}{\varepsilon'_r} \frac{\partial \rho(\varepsilon'_r)}{\partial \varepsilon'_r} \quad (228)$$

L'expression du terme de correction en utilisant la pondération de données est donnée par :

$$(\delta u, \delta v)^T = -\left( C_1^T W'_1 C_1 + C_2^T W'_2 C_2 \right)^{-1} \left( C_1^T W'_1 D_1 + C_2^T W'_2 D_2 \right) \quad (229)$$

Les matrices diagonales des poids sont données par les deux relations suivantes :

$$W'_1 = \text{diag}[\alpha'_d(1), \dots, \alpha'_d(n)] \quad (230)$$

et :

$$W'_2 = \text{diag}[\alpha'_r(1), \dots, \alpha'_r(n)] \quad (231)$$

- *Étape n°5* : retour à l'étape n°3 jusqu'à convergence.

Compte-tenu du bruit et des erreurs de quantification, la DFD sera rarement nulle. Le critère d'arrêt sera alors choisi en introduisant un seuil en deçà duquel l'estimation sera supposée correcte. Nous avons alors le critère d'arrêt suivant :

$$\{\rho(\varepsilon_d) \& \rho(\varepsilon_r)\} < \text{Seuil} \quad (232)$$

Il est également possible de définir le critère d'arrêt en fixant un nombre maximal d'itérations ou une combinaison de ces deux critères.

\*On procède de la même manière pour la deuxième fonctionnelle (terme quadratique de la somme algébrique de deux erreurs). Nous n'allons plus présenter le détail de calcul pour cette formulation.

#### 1.7.4 Résultats expérimentaux

Les performances de l'approche incrémentale basée sur un schéma direct-rétrograde ont été testées sur trois séquences différentes. Il s'agit de la séquence synthétique (la spirale) et les deux séquences réelles.

##### a) Expérimentation sur la séquence "spirale"

Un déplacement de (-2.5 ; 3) est affecté à la spirale. Pour pouvoir mettre en œuvre l'approche incrémentale, nous avons pris une valeur de déplacement décimale et plus importante que

celle utilisée dans les expérimentations précédentes. Les résultats sont ici donnés sous la forme des biais en fonction du SNR ; les courbes correspondantes sont représentées aux figures 2.41 et 2.42. Chaque point d'une courbe représente le biais moyen, valeur moyenne des biais qui sont calculés sur un ensemble de points et sur une série de tests (un lot de tests).

Les paramètres d'estimation sont les suivants :

- Support de calcul de taille  $13 \times 13$  pixels pour la résolution au sens des moindres carrés ;
- Masque MDIF pour le calcul des gradients spatiaux ;
- Masque Mask\_t4 pour le gradient temporel ;
- Nombre d'itération égal à 2 pour la pondération des données ;
- Le paramètre  $P$  est égal à 4.7 pour le calcul du seuil de pondération ;
- Au niveau de l'algorithme incrémental, l'itération s'arrête quand la valeur du terme de correction passe au dessous d'un seuil.

Sur les figures 2.41 et 2.42, on constate que les erreurs sur les deux composantes du déplacement deviennent stables à partir d'une valeur de SNR égale à  $60 \text{ dB}$ . Dans cet intervalle stable, une différence importante peut être constatée au niveau des amplitudes des biais. Les résultats obtenus avec la méthode des moindres carrés pondérés, basée sur un schéma simple, sont fortement biaisés. En effet, comme l'amplitude du déplacement est assez importante, la minimisation du critère utilisé ne permet pas d'atteindre une précision de l'ordre d'une fraction de pixel. L'utilisation de l'équation de contrainte de base (sans le terme de correction) n'est pas suffisante car l'optimisation s'effectue d'une manière exhaustive. Il en découle que l'estimation n'est pas traitée de façon satisfaisante.

L'utilisation de l'approche incrémentale est ici très significative (courbes MCP incrémentale simple). En effet, l'introduction du terme de correction dans le critère à minimiser permet d'atteindre une valeur de déplacement de l'ordre de sous-pixel, d'où une précision accrue.

Le résultat est encore meilleur lorsqu'on associe l'algorithme incrémental au formalisme direct-rétrograde (courbes MCP incrémentale direct-rétrograde). L'application du terme correcteur dans les deux sens temporels a permis de diminuer le biais d'estimation. Nous pouvons remarquer que cette dernière approche est plus précise comparée à l'approche direct-rétrograde non incrémentale (courbe MCP).

### **b) Expérimentation sur la séquence réelle à déplacement synthétique**

Les résultats sont ici donnés sous la forme de champ de vecteur déplacement. Considérons l'objet ① de la séquence présentée à la figure 2.6. La séquence est bruitée par un bruit gaussien additif de  $SNR = 100 \text{ dB}$ . Les estimations du champ de déplacement sur cette séquence, et cela pour différentes méthodes, sont illustrées comme suit :

- Figure 2.43 : résultat de la méthode des moindres carrés robuste, basée schéma simple ;
- Figure 2.44 : résultat de l'approche incrémentale pondérée et basée schéma simple ;
- Figure 2.45 : résultat de la méthode des moindres carrés robuste, basée schéma direct-rétrograde ( $J_b$ ) ;
- Figure 2.46 : résultat de l'approche incrémentale basée schéma direct-rétrograde ( $J_b$ ).

Nous avons utilisé un support d'estimation de taille  $13 \times 13$  pixels .

La figure 2.43 montre que l'algorithme des moindres carrés basé sur un schéma simple ne parvient à estimer correctement le mouvement qu'en l'associant avec une approche incrémentale (figure 2.44). La figure 2.45 montre clairement la capacité de la méthode basée sur un schéma direct-rétrograde à obtenir des résultats corrects, que ce soit en terme d'orientation ou en terme d'amplitude. Le résultat est nettement meilleur en utilisant l'approche incrémentale à formalisme direct-rétrograde (figure 2.46). Avec ce dernier, l'algorithme parvient à estimer un champ le plus proche possible du champ réel.

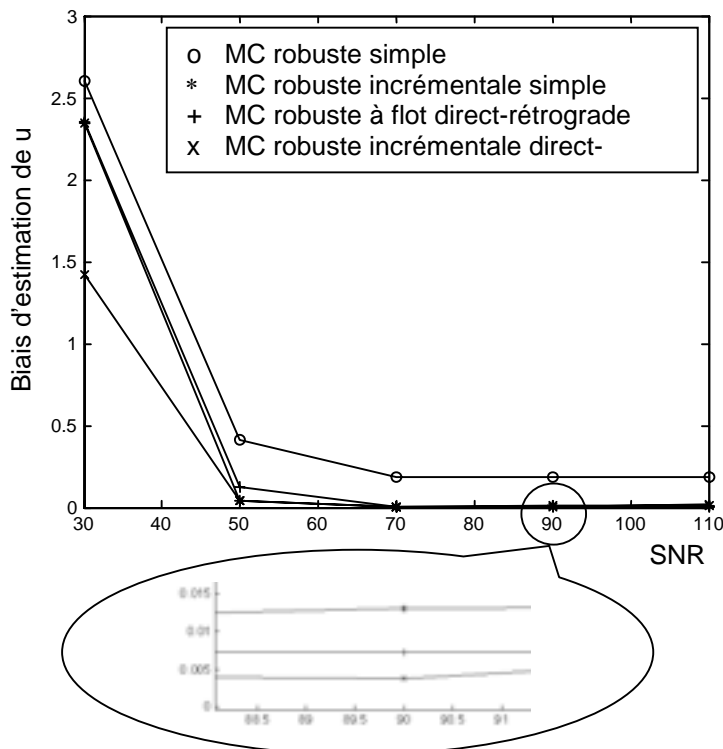


Figure 2.41 : Biais d'estimation de la composante  $u$  de la vitesse pour l'approche MC robuste basée schéma simple, l'approche incrémentale basée schéma simple, l'approche MC basée schéma direct-rétrograde et l'approche incrémentale basée schéma direct-rétrograde ( $J_b$ ).

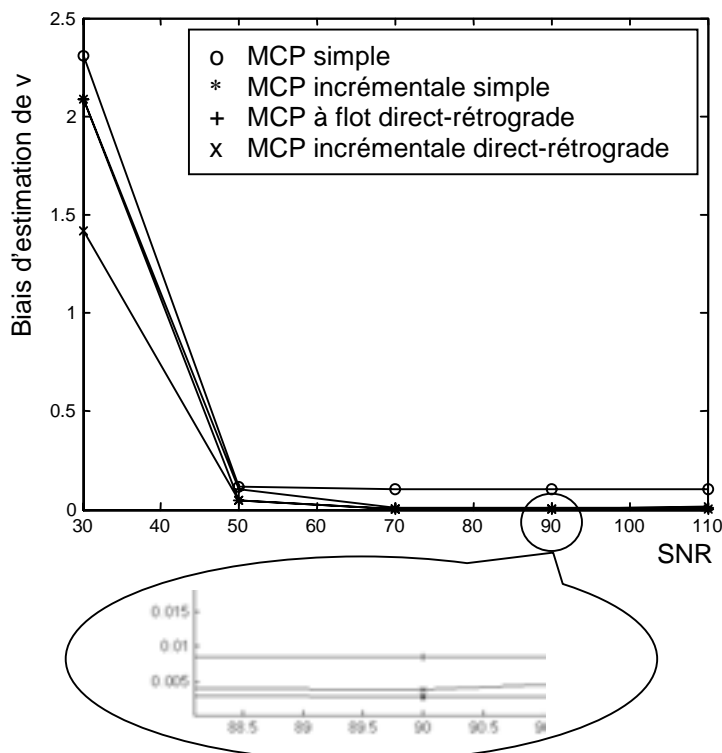


Figure 2.42 : Biais d'estimation de la composante  $v$  de la vitesse pour l'approche MC robuste basée schéma simple, l'approche incrémentale basée schéma simple, l'approche MC basée schéma direct-rétrograde et l'approche incrémentale basée schéma direct-rétrograde ( $J_b$ ).

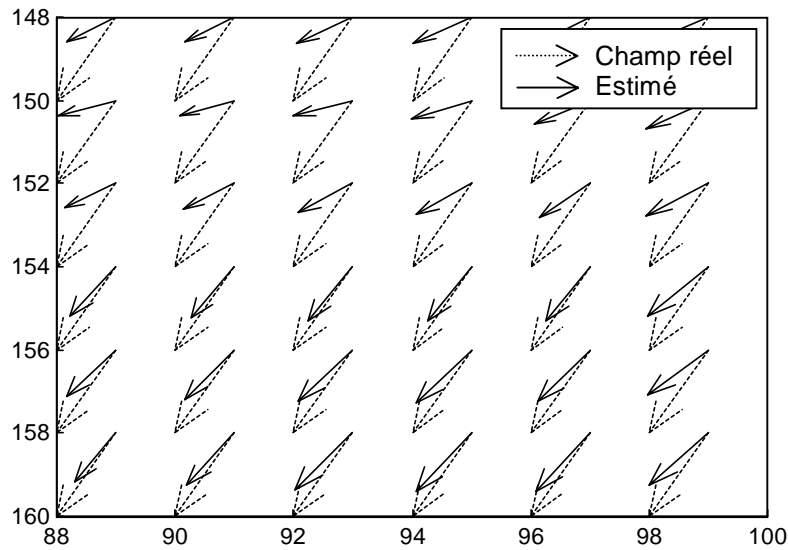


Figure 2.43 : Champ de déplacement de l'objet ① de la séquence réelle à déplacement synthétique ; séquence bruitée par un bruit gaussien additif ( $SNR = 100 \text{ dB}$ ) ; résultat de la méthode des moindres carrés robuste, basée sur un schéma simple.

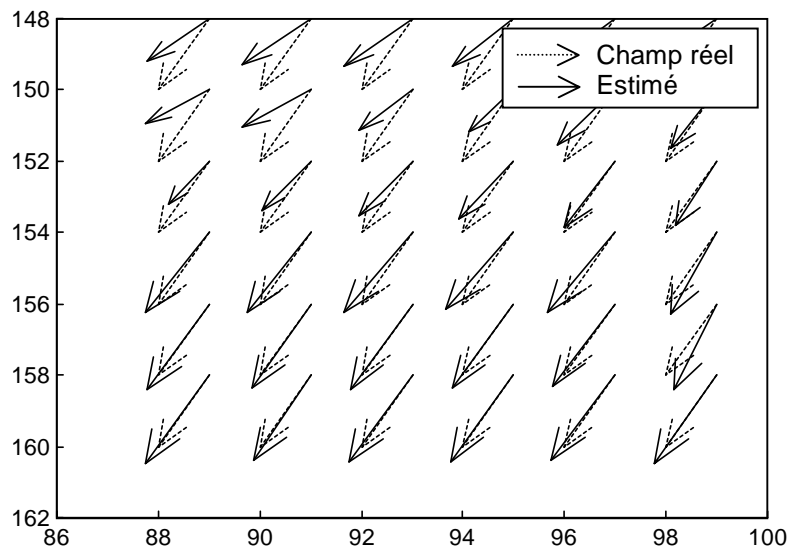


Figure 2.44 : Champ de déplacement de l'objet ① de la séquence réelle à déplacement synthétique ; séquence bruitée par un bruit gaussien additif ( $SNR = 100 \text{ dB}$ ) ; résultat de l'approche incrémentale pondérée, basée sur un schéma simple.

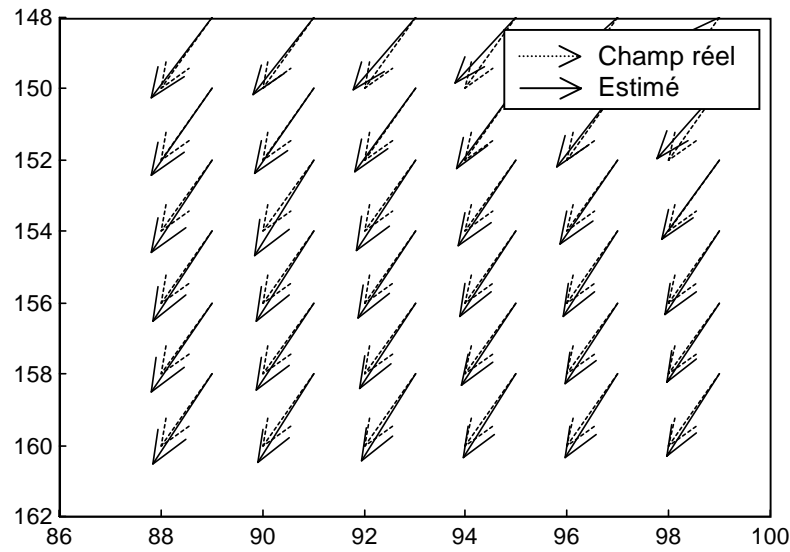


Figure 2.45 : Champ de déplacement de l'objet ① de la séquence réelle à déplacement synthétique ; séquence bruitée par un bruit gaussien additif ( $SNR = 100 \text{ dB}$ ) ; méthode des moindres carrés robuste, basée sur un schéma direct-rétrograde ( $J_b$ ).

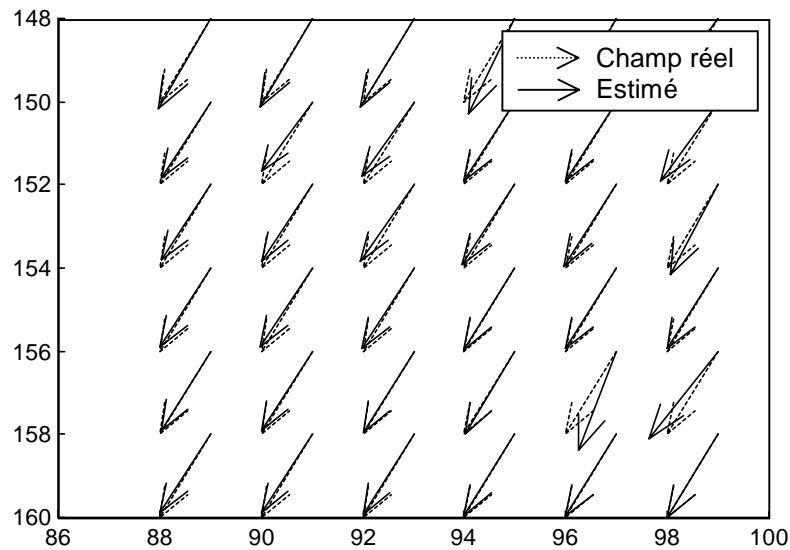


Figure 2.46 : Champ de déplacement de l'objet ① de la séquence réelle à déplacement synthétique ; séquence bruitée par un bruit gaussien additif ( $SNR = 100 \text{ dB}$ ) ; résultat de l'approche incrémentale pondérée, basée sur un schéma direct-rétrograde ( $J_b$ ).

**c) Expérimentation sur la séquence "taxi"**

L'expérimentation suivante porte sur le test de ces quatre approches sur la séquence taxi. Du point de vue quantitatif, une comparaison semble difficile entre les figures 2.47, 2.48, 2.49 et 2.50.

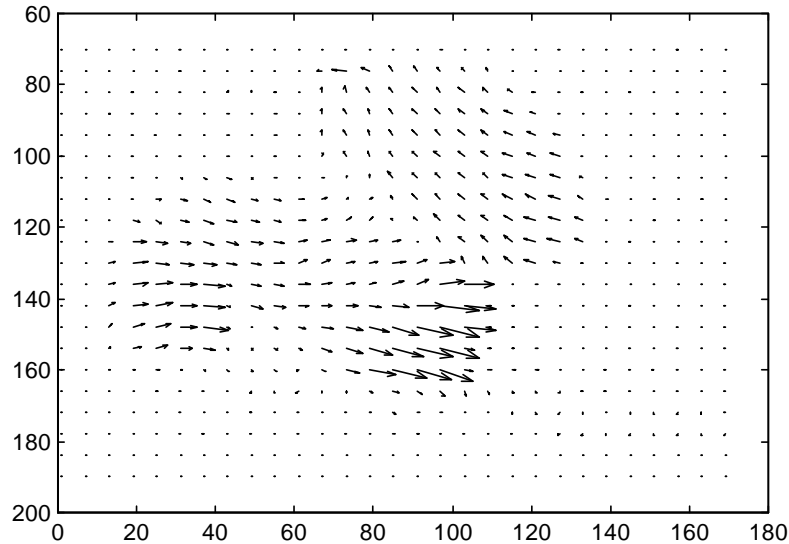


Figure 2.47 : Champ de déplacement des objets ① et ② de la séquence "taxi", à l'aide de la méthode des moindres carrés basée sur un simple.

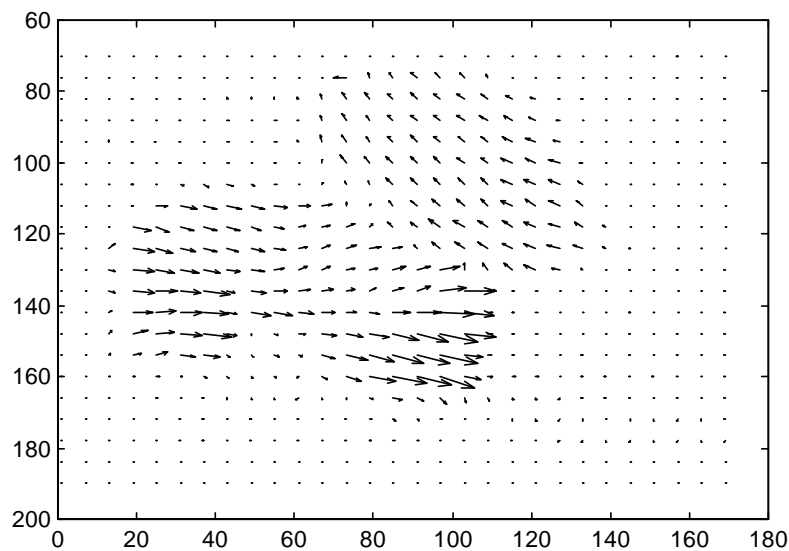


Figure 2.48 : Champ de déplacement des objets ① et ② de la séquence "taxi", à l'aide de l'approche incrémentale basée sur un schéma simple.

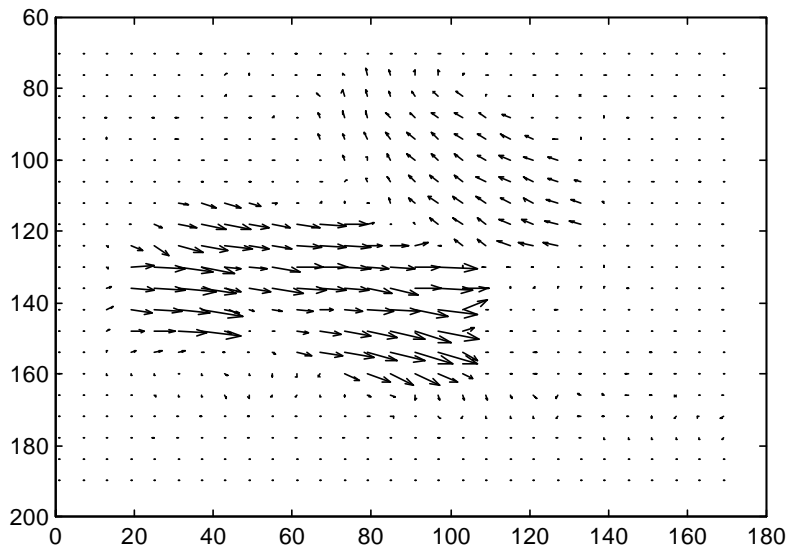


Figure 2.49 : Champ de déplacement des objets ① et ② de la séquence "taxi", à l'aide de la méthode des moindres carrés basée sur un schéma direct-rétrograde ( $J_b$ ).

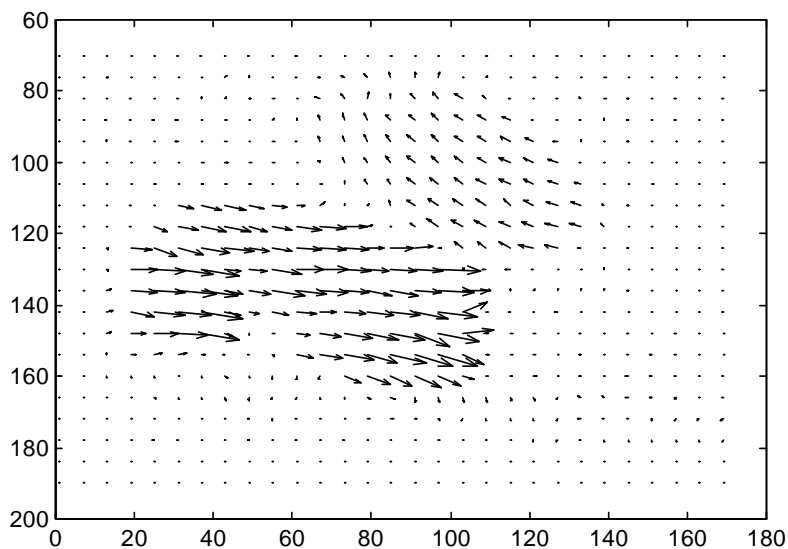


Figure 2.50 : Champ de déplacement des objets ① et ② de la séquence "taxi", à l'aide de l'approche incrémentale basée sur un schéma direct-rétrograde ( $J_b$ ).

Nous ne nous intéressons qu'à la comparaison visuelle. Nous pouvons déduire de ces figures les remarques suivantes :

- L'application de l'algorithme incrémental permet une amélioration de l'uniformité du champ de déplacement (comparaison entre les figures 2.47 et 2.48),
- Des améliorations ont été apportées notamment en introduisant le formalisme direct-rétrograde et cela, même sans utilisation de l'algorithme incrémental (comparaison entre les figures 2.47 et 2.49). Le champ est plus dense avec le formalisme proposé.
- L'association de l'approche incrémentale avec le formalisme direct-rétrograde permet d'avoir une meilleure performance (figure 2.50). Nous rappelons en effet que la particularité de cette technique est qu'on peut appliquer le terme de correction dans les deux sens temporels.

### 1.7.5 Discussion sur l'approche incrémentale basée sur un schéma direct-rétrograde

Cette méthode d'estimation est une généralisation de l'algorithme récursif à base de formalisme direct-rétrograde. Les résultats montrent qu'avec l'approche des moindres carrés basée sur un schéma simple, les déplacements de forte amplitude ne peuvent pas fréquemment être correctement estimés. L'utilisation de l'approche incrémentale permet d'estimer plus efficacement le mouvement des objets. Les courbes des biais montrent qu'il est possible d'obtenir une bonne estimation quantitative seulement à partir d'une certaine valeur du SNR (60 dB). D'autre part, l'utilisation de l'approche incrémentale avec le formalisme direct-rétrograde donne de meilleurs résultats.

L'approche incrémentale présente pourtant différentes limites :

- Si la fonction d'erreur à minimiser n'est pas monotone, il existe généralement plusieurs minima locaux et si ces minima se trouvent entre la valeur initiale et un minimum absolu, alors il n'est pas possible de déterminer seulement le minimum local le plus proche. Le comportement de l'algorithme dépend dans une large mesure du problème d'initialisation.
- Le coût temps d'estimation peut varier, tout dépend du nombre d'itérations.

## 1.8 Résumé du chapitre

Dans ce chapitre, nous avons présenté les solutions apportées par le formalisme direct-rétrograde à l'estimation du mouvement. Un rappel de la méthode des moindres carrés a été présenté. L'analyse des résultats de cette méthode permet de conclure que l'équation de contrainte du mouvement, prise uniquement dans un sens temporel, n'est pas suffisante pour avoir une bonne estimation. Nous avons ainsi introduit une stratégie d'estimation en reformulant l'équation de contrainte du mouvement. Cette formulation s'appuie sur la considération à la fois des mouvements direct et rétrograde dans la mise en équation. Le formalisme obtenu a permis de satisfaire le critère à minimiser. De nombreuses techniques ont été présentées pour minimiser ce critère. Nous avons d'abord présenté une estimation s'appuyant sur la méthode des moindres carrés pour la résolution du problème. Les résultats obtenus sur les séquences d'images ont montré que l'on réduit effectivement le biais d'estimation sans que cela nécessite un coût de calcul élevé. Le champ obtenu est beaucoup plus dense par rapport au résultat du formalisme classique (basé sur un schéma simple).

En revanche, les images présentant des discontinuités de mouvement ne sont pas traitées de façon optimale. Pour cela, il est nécessaire d'introduire un critère permettant de rejeter les données erronées. Une solution consiste donc à utiliser un estimateur robuste. Cette technique s'appelle estimation robuste basée sur un schéma direct-rétrograde. De nombreuses expérimentations ont permis de valider les performances de cette technique.

Malgré cette robustesse, le problème du déplacement de forte amplitude peut subsister. Une suggestion possible est d'utiliser un algorithme permettant d'introduire une correction récursive. Ainsi, l'approche incrémentale a été choisie. Cette approche consiste à généraliser la méthode des moindres carrés en introduisant un terme correctif. Cette idée a pour objectif de minimiser le critère des erreurs d'une manière optimale (au lieu de minimiser ce critère d'une manière exhaustive comme le font les méthodes précédentes). Les résultats obtenus ont montré que l'estimation a permis d'atteindre un déplacement avec une précision accrue.

L'usage de la technique incrémentale permet d'estimer des déplacements plus importants que ceux susceptibles d'être estimés par les approches classiques, tandis que l'estimation robuste permet de limiter les erreurs d'estimation dues à des points aberrants ou bien encore à la présence de zones d'occultation. Lors de nos tests sur l'approche incrémentale, nous sommes allés jusqu'à une amplitude de déplacement d'environ 8 pixels.

Pour conclure, les simulations sur diverses séquences d'images synthétique et réelle ont montré un taux de réussite satisfaisant et un résultat compatible avec notre objectif : la



segmentation spatio-temporelle. Les résultats justifient alors l'avantage du formalisme direct-rétrograde. Le tableau 2.1 présente une synthèse des méthodes examinées dans ce chapitre.

<b>Méthodes</b>	<b>Formalisme</b>	<b>Récurusif</b>	<b>Avantages</b>	<b>Inconvénients</b>
<b>LS<sup>12</sup></b>	Simple	Non	Simplicité	Imprécision
	Direct-rétrograde	Non	Précision et stabilité	
<b>TLS<sup>13</sup></b>	Simple	Non	Précision	Instabilité
	Direct-rétrograde	Non	Précision	Instabilité
<b>LS robuste</b>	Simple	Oui	Insensibilité aux points erronés	Besoin d'un réglage de seuil
	Direct-rétrograde	Oui	Insensibilité et précision aux points erronés	Besoin d'un réglage de seuil
<b>TLS robuste</b>	Simple	Oui	Précision	Instabilité
	Direct-rétrograde	Oui	Précision	Instabilité
<b>Incrémentale</b>	Simple	Oui	Estimation des grands déplacements	Coût de calcul élevé
	Direct-rétrograde	Oui	Estimation des grands déplacements	Coût de calcul élevé

Tableau 2.1 : Comparaison des performances des différentes méthodes étudiées, pour un mouvement de translation.

<sup>12</sup> Least squares

<sup>13</sup> Total least squares

## Chapitre 2

# Formalisme direct-rétrograde pour un modèle de mouvement paramétrique

### 2.1 Introduction

Ce chapitre traite de l'estimation des paramètres du modèle de mouvement affine. Nous avons vu lors de la première partie différents modèles de mouvement. L'intérêt du modèle paramétrique est de pouvoir modéliser des mouvements plus évolués qu'une simple translation (zoom, rotation, déformation linéaire). De plus, la description du mouvement à l'aide de ce modèle peut être vue comme une bonne approximation du mouvement apparent d'une facette plane dans un espace 3D.

Dans cette étude, nous allons considérer le modèle à quatre paramètres. A ce titre, on suppose que le mouvement d'un objet est composé d'un mouvement de translation (deux paramètres), une rotation (un paramètre) et une divergence (un paramètre). Le paramètre de déformation ne sera pas traité dans notre cas.

L'objet principal de ce chapitre est de montrer l'intérêt du formalisme basé schéma direct-rétrograde pour estimer les différents paramètres du mouvement. Pour cela, une nouvelle méthode d'estimation des paramètres du mouvement sera mise en œuvre, en utilisant le modèle paramétrique mis en évidence au chapitre 1 de la première partie.

### 2.2 Rappel sur le formalisme basé schéma simple pour un modèle de mouvement paramétrique

#### 2.2.1 Principes généraux

Comme il a été mentionné à la section 1.3 de la première partie, les paramètres d'un modèle de mouvement permettent de décrire sa nature. Nous allons reformuler l'exemple pris dans le paragraphe 1.5.7.b de la première partie, en utilisant un modèle affine. Il s'agit du mouvement de translation combiné avec une rotation et une divergence. Toute transformation appliquée à un pixel de l'image s'exprime de la manière suivante :

$$\begin{pmatrix} x' - x_G \\ y' - y_G \end{pmatrix} = T + M \begin{pmatrix} x - x_o \\ y - y_o \end{pmatrix} \quad (233)$$

Le choix du centre de rotation, des coordonnées  $x_o$  et  $y_o$ , est arbitraire. Ce point peut être rapporté au centre de gravité de l'objet. La matrice de translation  $T$  est exprimée par :

$$T = \begin{pmatrix} t_x \\ t_y \end{pmatrix} \quad (234)$$

La matrice  $M$  est constituée par l'angle de rotation  $\theta$  et le facteur de divergence  $k$ .

$$M = \begin{bmatrix} \cos \theta + k & -\sin \theta \\ \sin \theta & \cos \theta + k \end{bmatrix} \quad (235)$$

Exprimons le vecteur de déplacement en fonction de ces paramètres :

$$\begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} x' - x \\ y' - y \end{pmatrix} = T + (M - I) \begin{pmatrix} x - x_o \\ y - y_o \end{pmatrix} \quad (236)$$

où  $I$  est la matrice identité. Une simplification, en considérant que l'angle de rotation est faible, permet d'écrire :

$$\begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} t_x \\ t_y \end{pmatrix} + \begin{bmatrix} k & -\theta \\ \theta & k \end{bmatrix} \begin{pmatrix} x - x_o \\ y - y_o \end{pmatrix} \quad (237)$$

Pour alléger l'écriture, nous supposons que le centre de rotation coïncide avec l'origine. Nous avons donc  $x_o = 0$  et  $y_o = 0$  et l'erreur à minimiser est de la forme :

$$J = \sum_Q \left[ (t_x + kx - \theta y) I'_x + (t_y + \theta x + ky) I'_y + I'_t \right]^2 \quad (238)$$

La minimisation de cette fonctionnelle au sens des moindres carrés donne :

$$\begin{pmatrix} t_x, k, \theta, t_y \end{pmatrix}^T = -(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{B} \quad (239)$$

avec :

$$\mathbf{A} = \begin{bmatrix} I'_x(1) & x_1 I'_x(1) + y_1 I'_y(1) & x_1 I'_y(1) - y_1 I'_x(1) & I'_y(1) \\ \vdots & \vdots & \vdots & \vdots \\ I'_x(n) & x_n I'_x(n) + y_n I'_y(n) & x_n I'_y(n) - y_n I'_x(n) & I'_y(n) \end{bmatrix} \quad (240)$$

et :

$$\mathbf{B} = \begin{bmatrix} I'_t(1) \\ \vdots \\ I'_t(n) \end{bmatrix} \quad (241)$$

$I'_x$ ,  $I'_y$  et  $I'_t$  désignent respectivement les gradients suivant  $x$ ,  $y$  et  $t$  en un pixel donné.

### 2.2.2 Résultats expérimentaux

Cette expérimentation porte sur l'estimation des paramètres  $(t_x, t_y, k, \theta)$ . Dans un voisinage local de  $13 \times 13$  pixels, le mouvement des pixels est modélisé par ces quatre paramètres. L'estimation est faite dans un premier temps, sur une image synthétique. L'image a subi un mouvement de translation de  $(1,1)$ , une rotation de  $\pi/20$  par rapport à son centre de gravité et une divergence de  $0.1$ . La séquence est bruitée par un bruit gaussien additif, avec un SNR de  $90 \text{ dB}$ . L'estimation des gradients spatiaux est effectuée à l'aide du masque MDIF. Le masque `Mask_t4` appliqué à la différence temporelle donne le gradient temporel. L'estimation fournit quatre paramètres tels que deux translations, une rotation et une divergence. Le champ des déplacements est reconstruit à partir de ces paramètres. Le résultat est représenté à la figure 2.51. Une erreur d'estimation est associée au calcul de chaque paramètre et cette erreur se répercute sur la reconstruction des composantes du vecteur déplacement. Il en découle que les vecteurs de déplacement sont légèrement biaisés. La comparaison de ce biais avec celui du formalisme direct-rétrograde sera présentée dans la section 2.3

### 2.2.3 Discussion sur la méthode des moindres carrés simple appliquée au modèle de mouvement

Dans le cadre du mouvement paramétrique, la totalité de la zone en mouvement est biaisée d'une manière non uniforme. Sur certaines zones, le biais est faible mais sa valeur peut atteindre une valeur intolérable par rapport à la précision voulue. Le formalisme basé schéma simple n'est pas suffisant pour retrouver les paramètres du mouvement.

Le but du paragraphe suivant est d'utiliser un formalisme permettant d'optimiser les résultats d'une manière efficace. Cette approche consiste à appliquer le schéma direct-rétrograde à l'équation paramétrique du mouvement.

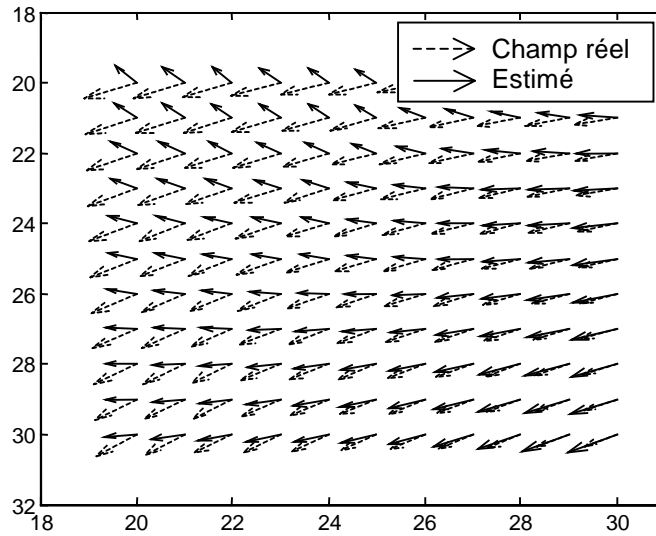


Figure 2.51 : Champ de déplacement de la séquence "spirale" pour un modèle affine ; méthode des moindres carrés basée sur un schéma simple,  $SNR = 90 \text{ dB}$ .

## 2.3 Extension du schéma direct-rétrograde au modèle paramétrique

### 2.3.1 Description

Nous avons vu dans la section précédente le développement de la méthode des moindres carrés sur des modèles paramétriques. Le but du présent paragraphe est d'étendre le schéma direct-rétrograde à un modèle de mouvement affine. Considérons les deux mouvements représentés à la figure 2.52.

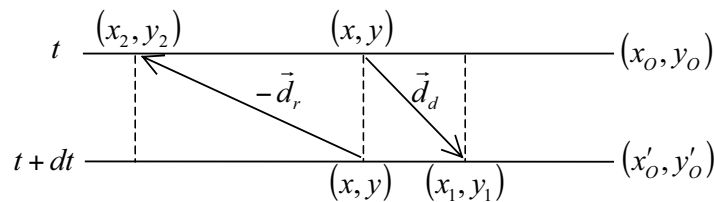


Figure 2.52 : Formalisme direct-rétrograde pour un modèle affine.

L'équation de transformation le long du déplacement direct s'écrit :

$$\begin{pmatrix} dx_d \\ dy_d \end{pmatrix} = \bar{T} + (M - I) \begin{pmatrix} x - x_o \\ y - y_o \end{pmatrix} \quad (242)$$

Pour le déplacement rétrograde, on a :

$$\begin{pmatrix} dx_r \\ dy_r \end{pmatrix} = \bar{T} + (M - I) \begin{pmatrix} x_2 - x_o \\ y_2 - y_o \end{pmatrix} \quad (243)$$

On cherche à exprimer  $(dx_r, dy_r)$  en fonction des paramètres du mouvement mais aussi en fonction de  $x$  et  $y$ . Soit :

$$\begin{pmatrix} dx_r \\ dy_r \end{pmatrix} = \bar{T} + (M - I) \begin{pmatrix} x - dx_r - x_o \\ y - dy_r - y_o \end{pmatrix} \quad (244)$$

En groupant les termes du déplacement dans le premier membre, nous avons :

$$(dx_r, dy_r)^T = M^{-1}\vec{T} + M^{-1}(M - I) \begin{pmatrix} x - x_o \\ y - y_o \end{pmatrix} \quad (245)$$

avec :

$$M^{-1} = \frac{1}{(k+1)^2 + \theta^2} \begin{bmatrix} k+1 & \theta \\ -\theta & k+1 \end{bmatrix} \quad (246)$$

En faisant apparaître l'expression du déplacement direct, nous pouvons écrire :

$$(dx_r, dy_r)^T = \vec{T} + (M - I)p + NL(d, p) \quad (247)$$

$p = \begin{pmatrix} x - x_o \\ y - y_o \end{pmatrix}$  désigne les coordonnées d'un pixel  $p$ . Le terme non linéaire  $NL(d, p)$  est une fonction des paramètres du mouvement et des coordonnées du pixel  $p$ .

$$NL(\Theta, p) = (M^{-1} - I) \left[ \vec{T} + (M - I)p \right] \quad (248)$$

Par rapport à la valeur du déplacement direct, nous avons :

$$\vec{d}_r = \vec{d}_d + NL(\Theta, p) \quad (249)$$

$\vec{d}_d = (dx_d, dy_d)^T$  représente le déplacement direct et  $\vec{d}_r = (dx_r, dy_r)^T$  le déplacement rétrograde.

La formulation des critères à base des déplacements direct et rétrograde s'avère compliquée du fait de la présence de ce terme non linéaire. La difficulté de la formulation réside surtout dans la restitution de ce terme. Plusieurs approches peuvent être développées afin de traiter ce problème.

*Variante n°1* : La première considération consiste en une linéarisation simple en ignorant le terme non linéaire. Cette hypothèse est faite afin de résoudre facilement le système d'équations. Cette considération n'est valable que pour des paramètres de rotation et de divergence très faibles par rapport à l'unité. L'approximation imposée peut être justifiée par les résultats de simulation. Nous verrons que jusqu'à certaines valeurs des paramètres du mouvement, l'utilisation de l'approche basée sur un schéma direct-rétrograde est encore plus intéressante que l'approche classique. Pour la fonctionnelle à somme des termes quadratiques, le critère s'écrit :

$$J_a = \sum_{\Omega} \left\{ (t_x + kx - \theta y)I'_x + (t_y + \theta x + ky)I'_y + I'_t \right\} + \left\{ (t_x + kx - \theta y)I_x^{t+T_e} + (t_y + \theta x + ky)I_y^{t+T_e} + I_t^{t+T_e} \right\} \quad (250)$$

Pour faciliter l'écriture, nous exprimons directement la fonctionnelle sous forme d'une norme euclidienne. Soit :

$$J_a = \|A_1\Theta + B_1\|_2^2 + \|A_2\Theta + B_2\|_2^2 \quad (251)$$

Dans cette expression, les matrices des données sont exprimées comme suit :

$$A_1 = \begin{bmatrix} I'_x(1) & x_1 I'_x(1) + y_1 I'_y(1) & x_1 I'_y(1) - y_1 I'_x(1) & I'_y(1) \\ \vdots & \vdots & \vdots & \vdots \\ I'_x(n) & x_n I'_x(n) + y_n I'_y(n) & x_n I'_y(n) - y_n I'_x(n) & I'_y(n) \end{bmatrix} \quad (252)$$

et :

$$A_2 = \begin{bmatrix} I_x^{t+T_e}(1) & x_1 I_x^{t+T_e}(1) + y_1 I_y^{t+T_e}(1) & x_1 I_y^{t+T_e}(1) - y_1 I_x^{t+T_e}(1) & I_y^{t+T_e}(1) \\ \vdots & \vdots & \vdots & \vdots \\ I_x^{t+T_e}(n) & x_n I_x^{t+T_e}(n) + y_n I_y^{t+T_e}(n) & x_n I_y^{t+T_e}(n) - y_n I_x^{t+T_e}(n) & I_y^{t+T_e}(n) \end{bmatrix} \quad (253)$$

Les éléments  $I_x^{t+Te}$  et  $I_y^{t+Te}$  désignent respectivement les gradients spatiaux discrets suivant  $x$  et suivant  $y$  d'un pixel, calculés sur l'image à  $t + dt$  ;  $B_1$  et  $B_2$  sont les gradients temporels que nous avons défini au chapitre 1 ; le terme  $\Theta = (t_x, k, \theta, t_y)^T$  représente le vecteur de paramètres du mouvement.

En minimisant l'expression (251) au sens des moindres carrés, le vecteur de paramètres du mouvement résultant est donné par :

$$\Theta = -(A_1^T A_1 + A_2^T A_2)^{-1} (A_1^T B_1 + A_2^T B_2) \quad (254)$$

*Variante n°2* : Quant à la deuxième réflexion, elle est semblable à la première, sauf qu'ici, on utilise la deuxième fonctionnelle (terme quadratique de la somme algébrique des erreurs). Le critère à minimiser s'écrit :

$$J_b = \|(A_1 + A_2)\Theta + (B_1 + B_2)\|_2^2 \quad (255)$$

La solution est donnée par :

$$\Theta = -[(A_1 + A_2)^T (A_1 + A_2)]^{-1} [(A_1 + A_2)^T (B_1 + B_2)] \quad (256)$$

*Variante n°3* : La troisième réflexion consiste à prendre en compte le terme non linéaire. D'après l'équation (248), ce terme peut être défini comme étant une fonction linéaire du déplacement ; son expression est donnée par :

$$NL(\Theta, p) = Q\vec{d} \quad (257)$$

où l'inconnue  $Q$  est une matrice carrée  $2 \times 2$  dont l'expression générale est :

$$Q = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} \quad (258)$$

D'après les équations (248) et (257), la matrice  $Q$  est définie par la relation suivante :

$$Q = (M^{-1} - I) \quad (259)$$

Nous déterminons les coefficients  $q_{11}$ ,  $q_{12}$ ,  $q_{21}$  et  $q_{22}$  d'une manière itérative, à partir des paramètres  $k$  et  $\theta$ , tel qu'à la  $i^{\text{ème}}$  itération, les valeurs de ces coefficients soient exprimées par :

$$\begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}^{(i+1)} = (M^{-1} - I)^i \quad (260)$$

Avant de définir le critère à minimiser, nous rappelons les deux fonctions d'erreur qui interviennent dans la formulation :

- Pour l'erreur suivant le sens temporel direct, nous avons :

$$\varepsilon_d = I(x + dx_d, y + dy_d, t + dt) - I(x, y, t) \quad (261)$$

- Pour le sens temporel rétrograde, nous avons :

$$\varepsilon_r = I(x - dx_r, y - dy_r, t) - I(x, y, t + dt) \quad (262)$$

En utilisant la fonctionnelle des erreurs à somme des termes quadratiques et d'après les relations (249, 257), nous avons :

$$J_a = \sum_{\Omega} \left\{ \left[ \vec{\nabla} I^t(x, y) \vec{V} + \frac{\partial I^t}{\partial t} \right]^2 + \left[ \vec{\nabla} I^{t+dt}(x, y) (\vec{V} - Q\vec{V}) + \frac{\partial I^{t+dt}}{\partial t} \right]^2 \right\} \quad (263)$$

Nous rappelons que la notation  $\vec{\nabla} I^t$  représente les gradients spatiaux de l'image à  $t$ . Pour le gradient temporel de l'image à  $t$ , la notation  $\partial I^t / \partial t$  est utilisée. Une notation unique  $\vec{V}$  est utilisée pour désigner le vecteur vitesse à estimer. En introduisant la norme Euclidienne et en utilisant une notation discrète, la relation (263) peut être exprimée par :

$$J_a = \|A_1 \Theta + B_1\|_2^2 + \|A_2^* \Theta + B_2\|_2^2 \quad (264)$$

La nouvelle matrice  $A_2^*$  fait la différence entre cette approche et les approches qui ne traitent pas le terme non linéaire. Sa forme est donnée par :

$$A_2^* = A_2 - A_2' - A_2'' \quad (265)$$

avec :

$$A_2' = \begin{bmatrix} q_{11}I_x^{t+Te}(1) & (q_{11}x_1 + q_{12}y_1)I_x^{t+Te}(1) & (-q_{11}y_1 + q_{12}x_1)I_x^{t+Te}(1) & q_{12}I_x^{t+Te}(1) \\ \vdots & \vdots & \vdots & \vdots \\ q_{11}I_x^{t+Te}(n) & (q_{11}x_n + q_{12}y_n)I_x^{t+Te}(n) & (-q_{11}y_n + q_{12}x_n)I_x^{t+Te}(n) & q_{12}I_x^{t+Te}(n) \end{bmatrix} \quad (266)$$

et :

$$A_2'' = \begin{bmatrix} q_{21}I_y^{t+Te}(1) & (q_{21}x_1 + q_{22}y_1)I_y^{t+Te}(1) & (-q_{21}y_1 + q_{22}x_1)I_y^{t+Te}(1) & q_{22}I_y^{t+Te}(1) \\ \vdots & \vdots & \vdots & \vdots \\ q_{21}I_y^{t+Te}(n) & (q_{21}x_n + q_{22}y_n)I_y^{t+Te}(n) & (-q_{21}y_n + q_{22}x_n)I_y^{t+Te}(n) & q_{22}I_y^{t+Te}(n) \end{bmatrix} \quad (267)$$

Les composantes de  $A_2$  sont les mêmes que celles utilisées dans la variante n°1. La minimisation au sens des moindres carrés de cette fonctionnelle donne la solution :

$$\Theta = -\left(A_1^T A_1 + A_2^{*T} A_2^*\right)^{-1} \left(A_1^T B_1 + A_2^{*T} B_2\right) \quad (268)$$

Au départ, les coefficients  $q_{ij}$  sont initialisés à zéro, c'est-à-dire que le terme non linéaire est ignoré. Donc, la matrice  $A_2^*$  prend la valeur de  $A_2$ . On estime ainsi un ensemble de paramètres du mouvement  $\Theta = (t_x, k, \theta, t_y)^T$  selon la relation (239). A partir de ces paramètres, on ré-estime les coefficients  $(q_{11}, q_{12}, q_{21}, q_{22})$  selon la relation (260). Les coefficients ainsi obtenus sont introduits dans les matrices (266, 267). On résout à nouveau l'équation (268) et l'opération se poursuit jusqu'à convergence.

Pour définir le critère d'arrêt de cet algorithme itératif, il est possible de définir soit un nombre d'itérations maximal, soit un seuil d'erreur en-dessous duquel on considère l'estimation comme correcte. On peut également utiliser la combinaison de ces deux critères.

La technique proposée permet de s'affranchir des problèmes posés par le terme non linéaire. Nous rappelons que ce terme est particulièrement apparu lors de l'introduction du mouvement rétrograde dans la mise en équation. Il est à rappeler que si on ne prend en compte que la formulation basée sur un schéma simple, le problème de non linéarité ne se présente pas dans la mise en équation. Pour le schéma direct-rétrograde, le traitement de ce terme permet de traiter de façon équilibrée les deux informations (celle dans le sens temporel direct et celle dans le sens temporel rétrograde). En effet, si on néglige le terme non linéaire, l'équation de contrainte obtenue ressemble à une forme classique qui s'est basée uniquement sur un simple mouvement direct ; cette équation favorise l'estimation du déplacement direct. Le seul moyen pour mieux considérer le déplacement rétrograde est de garder la forme non linéaire. La prise en compte de ce terme permet à l'algorithme d'estimation de trouver un compromis entre l'estimation du déplacement direct et celle du déplacement rétrograde.

*Variante n°4* : On peut procéder de la même manière pour la deuxième formulation (fonctionnelle définie par le carré de la somme algébrique de deux erreurs). Soit :

$$J_b = \left\| (A_1 + A_2^*)\Theta + (B_1 + B_2) \right\|_2^2 \quad (269)$$

La solution est donnée par :

$$\Theta = -\left[ (A_1 + A_2^*)^T (A_1 + A_2^*) \right]^{-1} (A_1 + A_2^*)^T (B_1 + B_2) \quad (270)$$

*Variante n°5* : Nous pouvons remarquer que le traitement du terme non linéaire suivant la troisième réflexion ne permet pas de linéariser complètement la contrainte. En effet, sur

l'équation (263), l'introduction du produit  $Q\vec{V}$  (ou  $Q\vec{d}$ ) fait naître les matrices  $A'_2$  et  $A''_2$ . L'ajout de ces matrices à la matrice initiale  $A_2$  a permis d'introduire une contrainte de linéarisation, mais celle-ci n'est que sous une forme implicite. La présente considération vise à linéariser complètement la contrainte. Le principe est d'appliquer le développement de Taylor au premier ordre à la quantité  $I(x - dx_r, y - dy_r, t)$  et de remplacer la quantité  $d_r$  par  $d_d - NL$ , nous avons ainsi :

$$I(x - dx_d - NL_x, y - dy_d - NL_y, t) \approx I(x - NL_x, y - NL_y, t + dt) - dx \frac{\partial}{\partial x} I(x - NL_x, y - NL_y, t + dt) - dy \frac{\partial}{\partial y} I(x - NL_x, y - NL_y, t + dt) - \frac{\partial}{\partial t} I(x - NL_x, y - NL_y, t + dt) \quad (271)$$

En introduisant l'équation (271) dans l'expression de l'erreur (262), nous obtenons l'équation de contrainte associée au mouvement rétrograde :

$$dx \frac{\partial}{\partial x} I(x - NL_x, y - NL_y, t + dt) + dy \frac{\partial}{\partial y} I(x - NL_x, y - NL_y, t + dt) + \frac{\partial}{\partial t} I(x - NL_x, y - NL_y, t + dt) + I(x, y, t + dt) - I(x - NL_x, y - NL_y, t + dt) = 0 \quad (272)$$

Nous utilisons ici une notation générale  $\vec{d} = (dx, dy)^T$  ou  $\vec{V} = (dx/dt, dy/dt)^T$  pour le mouvement direct. La fonctionnelle des erreurs s'écrit :

$$J_a = \sum_{\Omega} \left\{ \left[ \vec{\nabla} I'(x, y) \vec{V} + \frac{\partial I}{\partial t}(x, y, t) \right]^2 + \left[ \vec{\nabla} I^{t+dt}(x - NL_x, y - NL_y) \vec{V} + \frac{\partial}{\partial t} I(x - NL_x, y - NL_y, t + dt) + I(x, y, t + dt) - I(x - NL_x, y - NL_y, t + dt) \right]^2 \right\} \quad (273)$$

Il est à rappeler que, selon l'équation (257), le terme  $NL$  correspond au produit  $Q\vec{d}$ . La quantité  $\frac{\partial}{\partial t} I(x - NL_x, y - NL_y, t + dt) + I(x, y, t + dt) - I(x - NL_x, y - NL_y, t + dt)$  prend la place de la dérivée temporelle de l'image  $t + dt$ . Le calcul du gradient  $\vec{\nabla} I^{t+dt}(x + NL_x, y + NL_y)$  peut s'effectuer par interpolation. L'estimation est faite de manière itérative en introduisant le nouveau terme  $NL$  à chaque étape d'itération. En introduisant une notation discrète, la solution globale s'obtient en minimisant l'énergie suivante :

$$J_a = \|A_1 \Theta + B_1\|_2^2 + \|A_2^{\#} \Theta + B_2^{\#}\|_2^2 \quad (274)$$

La solution est de la forme :

$$\Theta = -\left(A_1^T A_1 + A_2^{\#T} A_2^{\#}\right)^{-1} \left(A_1^T B_1 + A_2^{\#T} B_2^{\#}\right) \quad (275)$$

$A_2^{\#}$  est une matrice de gradients spatiaux, estimés sur les pixels de l'image à  $t + dt$ , décalés du terme  $NL$ . Ses composantes sont exprimées par :

$$A_2^{\#} = \begin{bmatrix} I_x^{t+Te}(x_1 - NL_{x_1}, y_1 - NL_{y_1}) & x_1 I_x^{t+Te}(x_1 - NL_{x_1}, y_1 - NL_{y_1}) + y_1 I_y^{t+Te}(x_n - NL_{x_n}, y_n - NL_{y_n}) \\ \vdots & \vdots \\ I_x^{t+Te}(x_n - NL_{x_n}, y_n - NL_{y_n}) & x_n I_x^{t+Te}(x_n - NL_{x_n}, y_n - NL_{y_n}) + y_n I_y^{t+Te}(x_n - NL_{x_n}, y_n - NL_{y_n}) \end{bmatrix} \quad (276)$$

$$\begin{bmatrix} x_1 I_y^{t+Te}(x_1 - NL_{x_1}, y_1 - NL_{y_1}) - y_1 I_x^{t+Te}(x_1 - NL_{x_1}, y_1 - NL_{y_1}) & I_y^{t+Te}(x_1 - NL_{x_1}, y_1 - NL_{y_1}) \\ \vdots & \vdots \\ x_n I_y^{t+Te}(x_n - NL_{x_n}, y_n - NL_{y_n}) - y_n I_x^{t+Te}(x_n - NL_{x_n}, y_n - NL_{y_n}) & I_y^{t+Te}(x_n - NL_{x_n}, y_n - NL_{y_n}) \end{bmatrix}$$



Nous notons au passage que la notation  $I_x^{t+Te}$  représente les gradients spatiaux discrets suivant  $x$  de l'image à  $t + dt$ . Compte tenu du terme  $NL$ , les composantes de la matrice des gradients temporels de l'image  $t + dt$  sont modifiées. Son terme général s'écrit :

$$B_2^\# = \begin{bmatrix} I_t^{t+Te}(1) + I(x_1, y_1, t + Te) - I(x_1 - NL_{x_1}, y_1 - NL_{y_1}, t + Te) \\ \vdots \\ I_t^{t+Te}(n) + I(x_n, y_n, t + Te) - I(x_n - NL_{x_n}, y_n - NL_{y_n}, t + Te) \end{bmatrix} \quad (277)$$

Cette méthode permet de lever le problème de la non linéarité en apportant une correction sur les valeurs des gradients. Le calcul s'effectue d'une manière récursive. La phase d'initialisation peut être traitée en utilisant la variante n°1 ou la variante n°2. Le déplacement initial obtenu sera utilisé pour calculer le terme non linéaire et ainsi, pour interpoler les valeurs des gradients. Ces gradients servent à leur tour à estimer à nouveau le déplacement et la procédure continue jusqu'à convergence. Le critère d'arrêt peut être défini en comparant l'écart entre deux résidus successifs à un seuil.

*Variante n°6* : La présente approche utilise le même principe que la variante n°5, sauf qu'ici, on utilise la fonctionnelle  $J_b$ . L'estimation consiste à minimiser l'énergie suivante :

$$J_b = \sum_{\Omega} \left\{ \left[ \vec{\nabla} I^t(x, y) + \nabla I^{t+dt}(x - NL_x, y - NL_y) \right] \vec{V} + \frac{\partial}{\partial t} I(x, y, t) + \frac{\partial}{\partial t} I(x - NL_x, y - NL_y, t + dt) + I(x, y, t + dt) - I(x - NL_x, y - NL_y, t + dt) \right\}^2 \quad (278)$$

En passant par la notation discrète, l'expression de cette énergie en norme euclidienne est donnée par :

$$J_b = \left\| (A_1 + A_2^\#) \Theta + (B_1 + B_2^\#) \right\|_2^2 \quad (279)$$

La minimisation de cette énergie au sens des moindres carrés aboutit à une estimation des paramètres du mouvement.

$$\Theta = - \left[ (A_1 + A_2^\#)^T (A_1 + A_2^\#) \right]^{-1} (A_1 + A_2^\#)^T (B_1 + B_2^\#) \quad (280)$$

Cet algorithme peut également se dérouler d'une manière récursive. Le déplacement dernièrement calculé est utilisé pour calculer le terme  $NL$ , qui à son tour, sera réintroduit dans l'expression  $J_b$ .

Dans toute méthode d'estimation du déplacement, la prépondérance des problèmes liés au calcul du flot optique (problème d'ouverture, déplacement trop important, existence du gradient, ...) varie en fonction de la méthode de résolution utilisée [TEK95]. Le formalisme direct-rétrograde appliqué à un modèle de mouvement affine est sensible au problème des paramètres du mouvement trop importants.

Les réflexions ci-mentionnées peuvent être accompagnées d'un estimateur robuste. Pour ce faire, on reconstruit la valeur du déplacement à partir des paramètres estimés. Ce déplacement permet de déterminer l'erreur résiduelle en un point donné. Le poids en ce point peut être calculé à l'aide d'une fonction M-estimateur de cette erreur. Le vecteur  $\Theta = (t_x, k, \theta, t_y)^T$  est estimé en pondérant les données  $A_1$ ,  $A_2^\#$ ,  $B_1$  et  $B_2^\#$ .

Nous présentons les résultats de test de ces approches dans le paragraphe suivant.

### 2.3.2 Résultats expérimentaux

L'analyse des résultats expérimentaux se déroule en trois volets :

- Le premier volet consiste à analyser visuellement les champs de déplacement obtenus par ces méthodes. Dans ce cadre, nous effectuons différents tests permettant d'analyser le

comportement des algorithmes proposés en fonction de la valeur des paramètres du mouvement. Trois séries de résultats sont ainsi présentées. Tout d'abord, les figures 2.53 à 2.56 montrent les résultats obtenus avec un modèle affine à quatre paramètres (deux translations + une rotation + une divergence). La deuxième série de résultats permet d'analyser le comportement des algorithmes sur un mouvement de forte rotation et de forte divergence (figures 2.57 à 2.59). Dans la dernière série de tests, nous présentons des résultats en augmentant les valeurs de translation, rotation et divergence (figures 2.60 à 2.62).

- Le deuxième volet est consacré au test de validité de ces algorithmes dans un contexte robuste.

- Le dernier volet permet d'avoir une analyse plus profonde à travers le biais associé à chaque paramètre du mouvement.

### a) Analyse des résultats à partir du champ de déplacement

#### • *Mouvement affine composé d'une translation, d'une rotation et d'une divergence*

Le mouvement de la spirale est composé de deux composantes de translation  $(-1,1)$ , un coefficient de divergence de 0.1 et un angle de rotation de  $\pi/10$ . Un bruit blanc gaussien de 90 dB a été rajouté à l'image. Les gradients spatiaux sont estimés en utilisant le masque MDIF. Le gradient temporel se détermine avec la différence temporelle convoluée avec le masque Mask\_t4. Avant d'analyser la performance de l'approche basée schéma direct-rétrograde, pour les différentes variantes présentées, nous allons d'abord tracer le champ produit par l'approche basée sur un schéma simple. Pour ce faire, on estime d'abord les quatre paramètres du mouvement selon l'équation (239) et on reconstruit après le champ de déplacement à partir de ces paramètres. La figure 2.53 montre clairement l'insuffisance de contrainte pour obtenir des résultats précis. Plus on s'éloigne du centre de rotation, de coordonnées  $(50,50)$ , plus les vecteurs estimés s'écartent des déplacements réels.

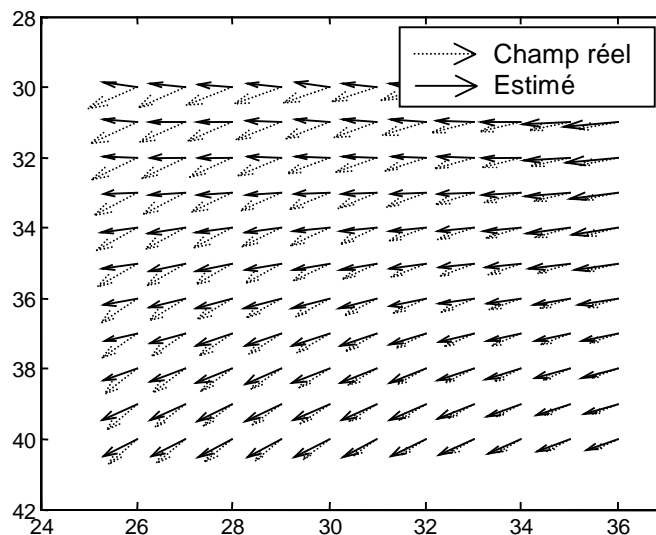


Figure 2.53 : Champ de déplacement pour la séquence "spirale", bruitée par un bruit gaussien additif ( $SNR = 90 dB$ ) ; le modèle de mouvement utilisé est de type affine dont les paramètres théoriques sont :  $t_x = -1$ ,  $t_y = 1$ ,  $k = 0,1$  et  $\theta = \pi/10$  ; résultat de la méthode des moindres carrés basée sur un schéma direct.

Quant aux résultats avec l'approche moindres carrés basée sur un schéma direct-rétrograde, nous allons dans un premier temps considérer le cas qui ne traite pas le terme non linéaire. La figure 2.54 correspond à la formulation utilisant la fonctionnelle  $J_a$ . Le champ de déplacement est plus ou moins bien ordonné sur une certaine zone. Par rapport au résultat précédent, l'application du formalisme direct-rétrograde permet ici de réduire considérablement le biais d'estimation.

Le test suivant est relatif à la deuxième réflexion (fonctionnelle  $J_b$ ). Le champ estimé correspond approximativement au champ réel (figure 2.55).

Ces résultats montrent aussi que l'utilisation de la méthode des moindres carrés, pour l'estimation des paramètres du mouvement, reste valable et cela en appliquant le formalisme direct-rétrograde.

Le dernier test permet d'analyser les résultats obtenus en prenant en compte la forme non linéaire. L'idée est ici de partir d'un jeu de paramètres initial, en l'occurrence le résultat de l'approche basée sur un schéma simple, pour extraire les éléments initiaux de la matrice  $Q$ . A partir de ces éléments, on calcule les composantes du terme non linéaire  $(NL_x, NL_y)$ . Ces composantes seront introduites dans les matrices des données (notées généralement par  $A_{t+1}^\#$  et  $B_{t+1}^\#$ ) pour estimer à nouveau un jeu de paramètres du mouvement. L'opération se poursuit jusqu'à trois itérations (nombre d'itérations maximal). Nous avons testé ici la variante n°6. Sur ce nouveau champ (figure 2.56), l'écart entre le champ réel et l'estimé se réduit visiblement. Par contre, les paramètres obtenus n'ont a priori aucune raison de correspondre à une erreur de reconstruction minimale mais ils permettent en tout cas de reconstruire un champ correct.

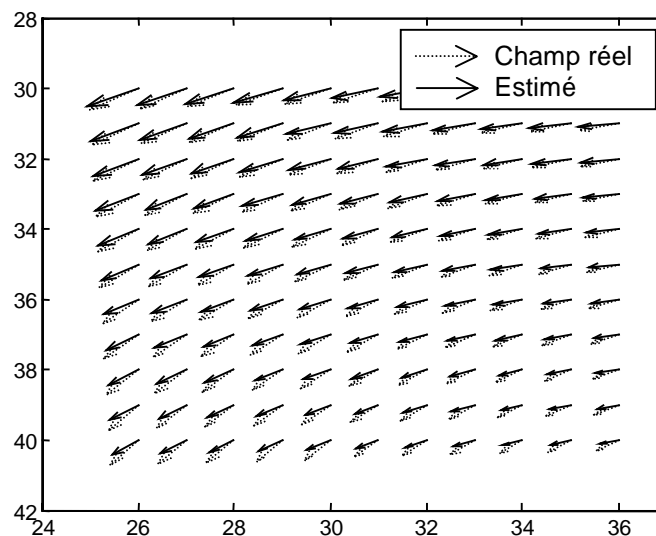


Figure 2.54 : Champ de déplacement pour la séquence "spirale", bruitée par un bruit gaussien additif ( $SNR = 90 \text{ dB}$ ) ; le modèle de mouvement utilisé est de type affine dont les paramètres théoriques sont :  $t_x = -1$ ,  $t_y = 1$ ,  $k = 0,1$  et  $\theta = \pi/10$  ; résultat de la méthode des moindres carrés basée sur un schéma direct-rétrograde, avec la fonctionnelle  $J_a$ , variante n°1.

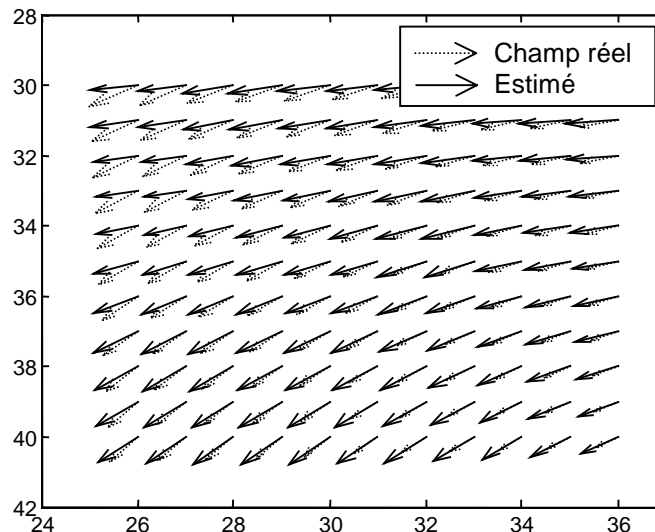


Figure 2.55 : Champ de déplacement pour la séquence "spirale", bruitée par un bruit gaussien additif ( $SNR = 90 \text{ dB}$ ) ; le modèle de mouvement utilisé est de type affine dont les paramètres théoriques sont :  $t_x = -1$ ,  $t_y = 1$ ,  $k = 0,1$  et  $\theta = \pi/10$  ; résultat de la méthode des moindres carrés basée sur un schéma direct-rétrograde, avec la fonctionnelle  $J_b$ , variante n°2.

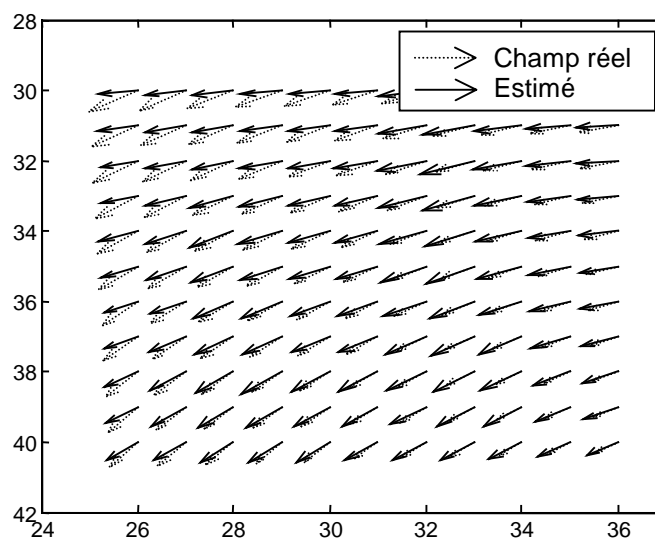


Figure 2.56 : Champ de déplacement pour la séquence "spirale", bruitée par un bruit gaussien additif ( $SNR = 90 \text{ dB}$ ) ; le modèle de mouvement utilisé est de type affine dont les paramètres théoriques sont :  $t_x = -1$ ,  $t_y = 1$ ,  $k = 0,1$  et  $\theta = \pi/10$  ; résultat de la méthode des moindres carrés basée sur un schéma direct-rétrograde, avec la fonctionnelle  $J_b$ , variante n°6.

- **Mouvement composé d'une forte rotation et une forte divergence**

Le test présenté ici est un cas particulier du mouvement, composé uniquement de fortes rotation et divergence. Une translation nulle, un coefficient de divergence de 0.3 et un angle de rotation  $\pi/3$  ont été attribués à la spirale. Un bruit blanc gaussien additif de 90 dB a été rajouté à l'image. Les gradients spatiaux sont estimés en utilisant le masque MDIF. Le gradient temporel est déterminé par la différence temporelle convoluée avec le masque Mask\_t4. Afin de comparer entre eux les résultats obtenus avec les différents algorithmes,

nous allons commencer par tracer les résultats de la méthode des moindres carrés basée sur un schéma simple. Les vecteurs de déplacement sont illustrés sur la figure 2.57.

Le second résultat est associé à la méthode des moindres carrés basée sur un schéma direct-rétrograde, et cela avec l'équation simplifiée sans le terme non linéaire (variante n°2). Le champ obtenu est représenté sur la figure 2.58. Comparée à la figure 2.57, on ne voit pas l'apport du formalisme direct-rétrograde sur la figure 2.58. Ce résultat montre la limitation de la variante n°2 sur l'estimation des composantes de rotation et de divergence. Le fait d'ignorer le terme non linéaire n'aboutit pas à une solution fiable. La prise en considération de ce terme est nécessaire pour un mouvement de rotation et de divergence dominantes, nous allons ainsi en tenir compte dans le troisième test (figure 2.59).

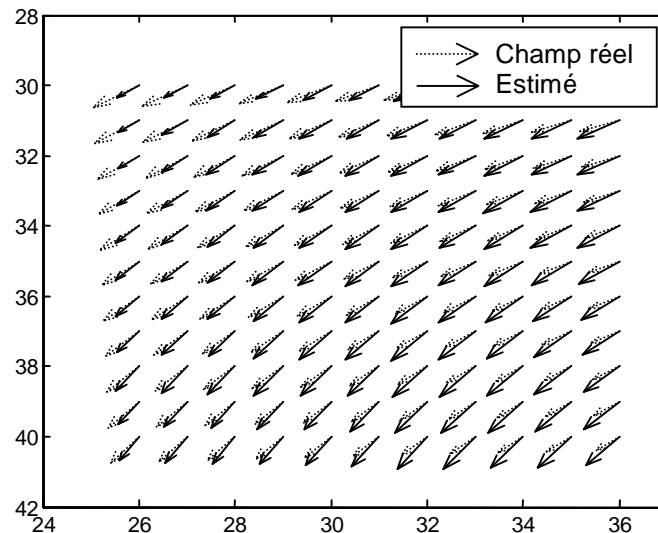


Figure 2.57 : Champ de déplacement pour la séquence "spirale", bruitée par un bruit gaussien additif ( $SNR = 90 \text{ dB}$ ) ; résultat de la méthode des moindres carrés basée sur schéma simple ; le modèle de mouvement utilisé est de type affine dont les paramètres théoriques sont :  $t_x = 0$ ,  $t_y = 0$ ,  $k = 0,3$  et  $\theta = \pi/3$ .

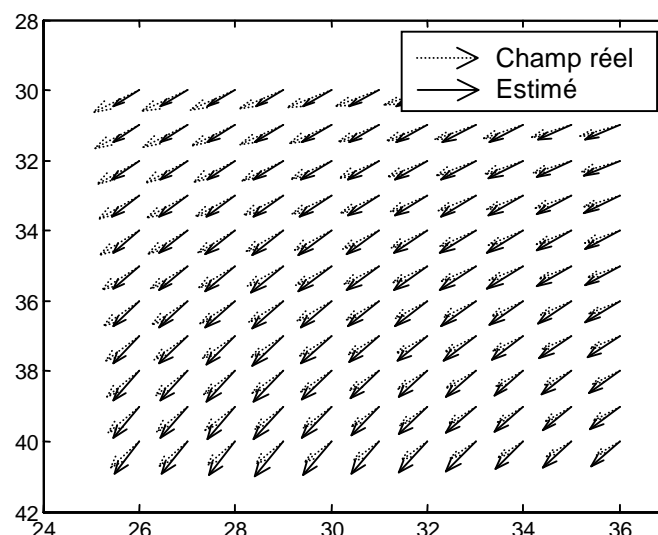


Figure 2.58 : Champ de déplacement pour la séquence "spirale", bruitée par un bruit gaussien additif ( $SNR = 90 \text{ dB}$ ), pour un modèle affine de forte rotation et divergence ; paramètres :  $t_x = 0$ ,  $t_y = 0$ ,  $k = 0,3$  et  $\theta = \pi/3$  ; résultat de la méthode des moindres carrés basée sur un schéma direct-rétrograde sans terme non linéaire (variante n°2).

Comme nous l'avons remarqué, le terme non linéaire est fortement lié aux paramètres de rotation et de divergence, notre attention porte sur l'un des algorithmes qui traite la non linéarité. La variante n°4 est testée ici. Le champ de déplacement est représenté à la figure 2.59.

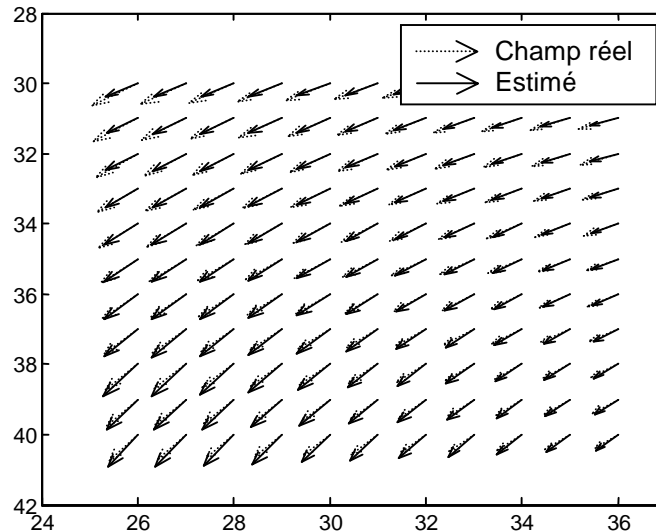


Figure 2.59 : Champ de déplacement pour la séquence "spirale", bruitée par un bruit gaussien additif ( $SNR = 90 \text{ dB}$ ), pour un modèle affine de forte rotation et divergence ; résultat de la méthode des moindres carrés basée sur un schéma direct-rétrograde, variante n°4 ; paramètres :  $t_x = 0$ ,  $t_y = 0$ ,  $k = 0,3$  et  $\theta = \pi/3$ .

- **Analyse des performances des algorithmes quand on augmente les valeurs des paramètres**

Les tests suivants permettent d'évaluer l'apport de la prise en compte du terme non linéaire à l'approche basée sur un schéma direct-rétrograde. Dans ce contexte, nous essayons d'augmenter les valeurs des paramètres du mouvement et nous analysons ainsi le comportement des algorithmes face à ce test.

Pour mettre en évidence les performances de l'approche basée modèle paramétrique et basée formalisme direct-rétrograde et tout cela avec la forme non linéaire, nous avons pris un déplacement avec des paramètres importants. Quantitativement, lorsqu'on augmente la valeur des paramètres du mouvement, la méthode des moindres carrés simple (schéma simple) et la méthode des moindres carrés à base de schéma direct-rétrograde (sans traitement du terme non linéaire) sont limitées à fournir des résultats fiables. En effet, pour des paramètres du mouvement trop importants, l'erreur introduite par le terme non linéaire devient prépondérante et l'hypothèse qui consiste à ignorer ce terme n'est plus valable, d'où une contradiction avec la négligence de ce terme. D'après les figures 2.60 et 2.61, les champs sont éloignés des valeurs théoriques. Pour avoir des champs plus précis, nous avons procédé au traitement du terme non linéaire selon la variante n°6. Le résultat du test est présenté à la figure 2.62.

Ces expérimentations illustrent clairement l'importance de la prise en compte du terme non linéaire dans la méthode des moindres carrés basée sur un schéma direct-rétrograde et cela d'autant plus que les valeurs des paramètres du mouvement sont importantes. Nous avons montré ici la nécessité de mettre en œuvre une stratégie de linéarisation (variante n°6) permettant de gérer le problème de la non linéarité.

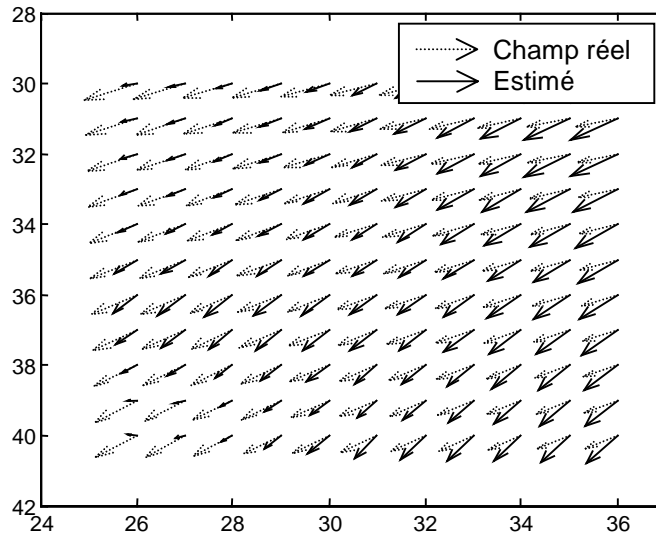


Figure 2.60 : Champ de déplacement pour la séquence "spirale", bruitée par un bruit gaussien additif ( $SNR = 90 \text{ dB}$ ), pour un modèle affine dont les paramètres sont :  $t_x = 4$ ,  $t_y = -3$ ,  $k = 0.3$ ,  $\theta = \pi/8$  ; résultat de la méthode des moindres carrés à base de flot direct.

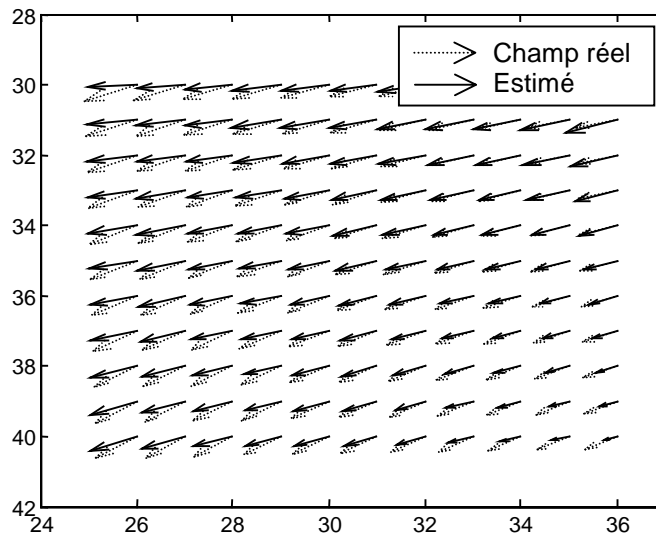


Figure 2.61 : Champ de déplacement pour la séquence "spirale", bruitée par un bruit gaussien additif ( $SNR = 90 \text{ dB}$ ), pour un modèle affine dont les paramètres sont :  $t_x = 4$ ,  $t_y = -3$ ,  $k = 0.3$ ,  $\theta = \pi/8$  ; résultat de la méthode des moindres carrés à base de flot direct-rétrograde, sans prise en compte du terme non linéaire ; variante n°2.

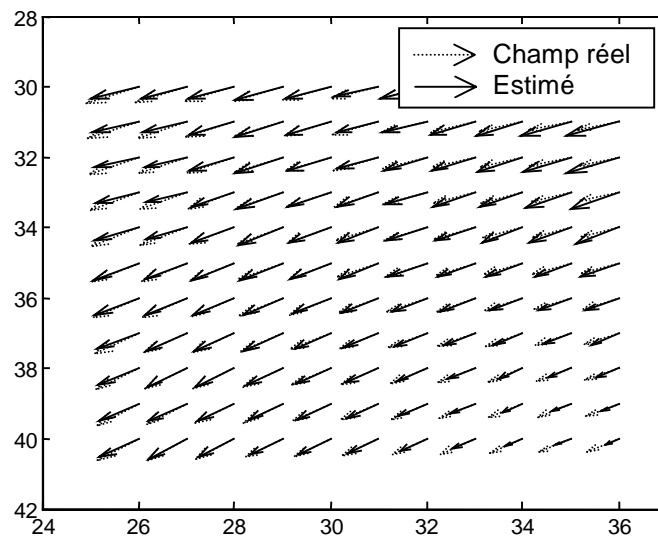


Figure 2.62 : Champ de déplacement pour la séquence "spirale", bruitée par un bruit gaussien additif ( $SNR = 90 \text{ dB}$ ), pour un modèle affine dont les paramètres sont :  $t_x = 4$ ,  $t_y = -3$ ,  $k = 0.3$ ,  $\theta = \pi/8$  ; résultat de la méthode des moindres carrés à base de flot direct-rétrograde, sans prise en compte du terme non linéaire selon la variante n°6.

#### b) Introduction de la technique robuste dans l'estimation des paramètres du mouvement à base de schéma direct-rétrograde

La méthode d'estimation combinant une stratégie de pondération et un formalisme direct-rétrograde permet de résoudre le problème lié aux points aberrants d'une région donnée. Dans ce manuscrit, l'estimateur robuste (M-estimateur) est souvent évoqué. Le test suivant permet d'associer la méthode des moindres carrés direct-rétrograde, selon la variante n°2, à un estimateur robuste. Le test a été effectué sur une zone présentant des discontinuités de mouvement. Les résultats sont reportés à la figure 2.63.

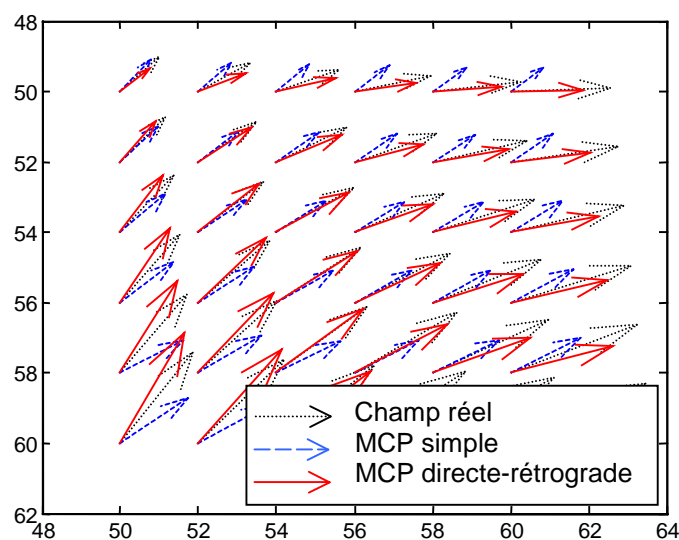


Figure 2.63 : Champ de déplacement réel, résultat de l'approche des moindres carrés pondérés (MCP) basée schéma simple et déplacement estimé à l'aide de l'approche MCP à flot direct-rétrograde ( $J_b$ ) ; le nombre d'itérations est égal à 2.



La figure 2.63 montre la comparaison des champs suivants :

- Le champ théorique ;
- Le champ de déplacement obtenu à l'aide de la méthode des moindres carrés pondérés, à base de flot direct ;
- Le champ obtenu à l'aide de la méthode des moindres carrés pondérés, à base de flot direct-rétrograde et selon la variante n°2.

La qualité visuelle des champs estimés est moins bonne dans le cas de l'approche à flot direct.

### c) Analyse des résultats à partir du biais d'estimation

Afin de mieux analyser les résultats, nous passons à un test comparatif des biais pour les approches suivantes :

- L'approche des moindres carrés simple,
- L'approche basée sur un schéma direct- rétrograde à modèle linéaire (sans prise en compte du terme non linéaire),
- L'approche basée sur un schéma direct-rétrograde à modèle non linéaire (avec prise en compte du terme non linéaire).

Pour les deux dernières méthodes, nous utilisons la formulation avec le critère  $J_b$ . Les trois méthodes sont testées à partir de paramètres de simulation égaux. La comparaison est effectuée au travers du biais et de la variance de l'erreur d'estimation. Avant de passer à l'analyse des biais, nous affichons simultanément sur une même figure (figure 2.64) les flots fournis par ces méthodes, et cela pour les paramètres du mouvement suivants :  $t_x = 2$ ,  $t_y = -1$ ,  $k = 0,2$  et  $\theta = \pi/10$ .

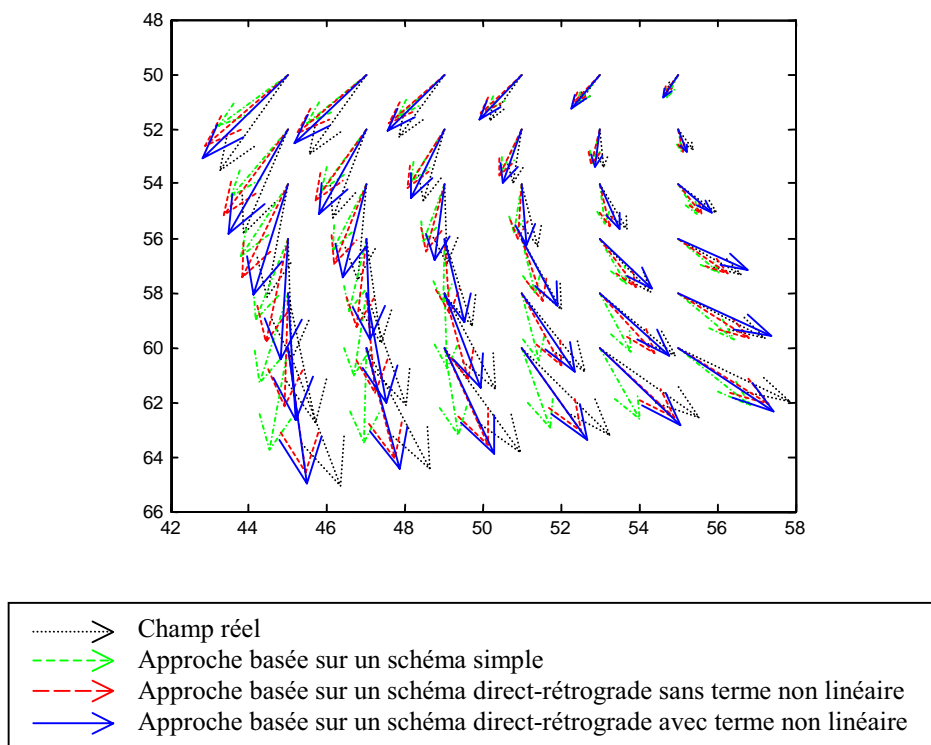


Figure 2.64 : Champ de déplacement pour la séquence "spirale", pour les trois méthodes testées, avec les paramètres du mouvement suivants :  $t_x = 2$ ,  $t_y = -1$ ,  $k = 0,2$  et  $\theta = \pi/10$ .

Nous traçons ici, pour les trois méthodes, le biais associé à chaque paramètre en fonction du bruit (figures 2.65 à 2.68) . Les biais sont calculés en prenant la moyenne des biais sur un ensemble de points. Une série de tests est ainsi effectuée et le résultat final résulte de la moyenne des résultats du lot de tests. L'intervalle de confiance pour les biais d'estimation se trouve entre les valeurs 40 et 100  $dB$  .

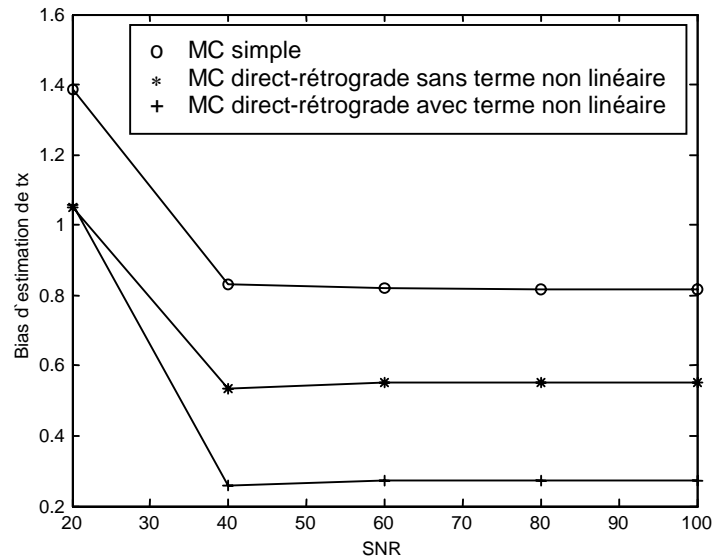


Figure 2.65 : Évolution de l'erreur sur la composante de translation suivant  $x$  pour les 3 méthodes testées ; courbes tracées en fonction du SNR ; séquence "spirale".

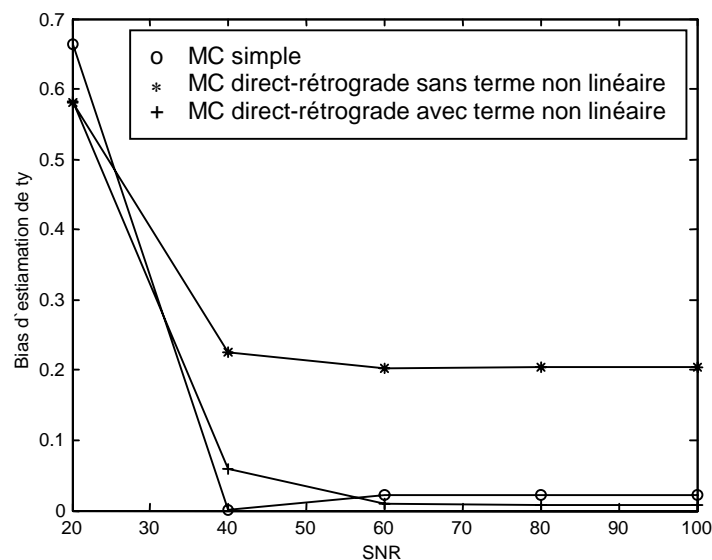


Figure 2.66 : Évolution de l'erreur sur la composante de translation suivant  $y$  pour les 3 méthodes testées ; courbes tracées en fonction du SNR ; séquence "spirale".

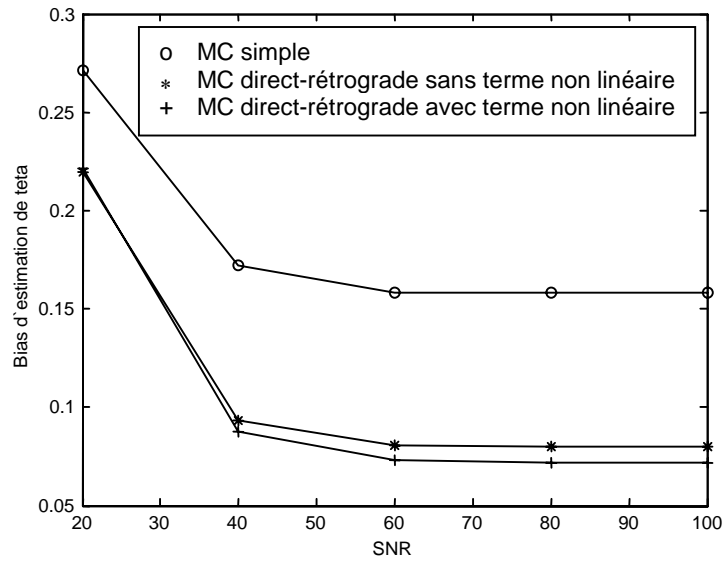


Figure 2.67 : Évolution de l'erreur sur le paramètre de rotation pour les 3 méthodes testées ; courbes tracées en fonction du SNR ; séquence "spirale".

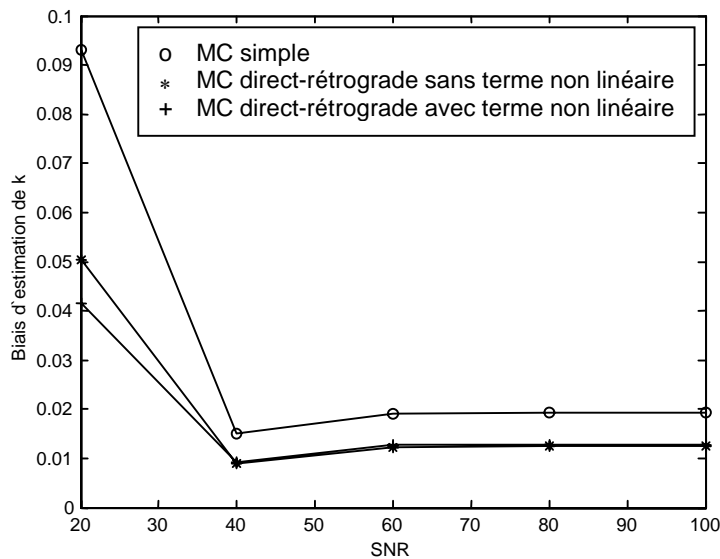


Figure 2.68 : Évolution de l'erreur sur le paramètre de divergence pour les 3 méthodes testées ; courbes tracées en fonction du SNR ; séquence "spirale".

Le flot calculé à faible bruit ou en l'absence de bruit est très proche du flot réel. Sur les données très bruitées, soit pour un bruit additif gaussien au dessous de  $60 \text{ dB}$ , toutes les méthodes produisent des flots extrêmement perturbés. Il est montré que tous les paramètres n'ont pas la même sensibilité au bruit. Les deux approches basées flot direct-rétrograde présentent des biais d'estimation faibles par rapport à ceux de l'approche basée sur un schéma simple. Pour l'approche basée sur un schéma direct-rétrograde, l'intérêt de la prise en compte du terme non linéaire est illustré par ces figures. Cette approche fournit des résultats plus proches des paramètres de rotation et divergence réels. Pour une meilleure estimation du modèle paramétrique, nous recommandons l'utilisation de ce formalisme.

Ces résultats montrent clairement l'incapacité de la méthode des moindres carrés classiques à obtenir ici des résultats corrects, que ce soit en terme d'estimation des paramètres du mouvement ou en terme de reconstruction du champ de déplacement. Les résultats sont donc fortement perturbés par les erreurs de sources diverses. Les courbes de biais montrent qu'il est possible de réduire l'erreur d'estimation du déplacement, seulement si l'on utilise l'approche à base de flot direct-rétrograde, que les paramètres de translation soient faibles ou plus importants. En revanche, cette approche est plus sensible au bruit en estimation du paramètre de rotation. Dans l'étude théorique, il est déjà précisé que la validité de l'approche directe-rétrograde est limitée par la valeur des paramètres de rotation et de divergence. En effet, le terme non linéaire reste prépondérant pour un paramètre de rotation ou de divergence important. L'hypothèse qui consiste à ignorer le terme non linéaire n'est donc plus valable, d'où l'utilité du traitement de ce terme.

Les séries de tests suivantes décrivent les résultats obtenus sur des séquences plus complexes. Plusieurs cas de mouvements de nature différente seront considérés.

#### d) Expérimentation sur une image en rotation

Nous choisissons, dans cette expérimentation, d'estimer un mouvement de rotation dominant. Il s'agit d'un mouvement local d'un objet en perspective autour d'un axe vertical. L'image est représentée à la figure 2.69. Les résultats sont analysés à partir du champ de déplacement et de l'erreur de reconstruction.

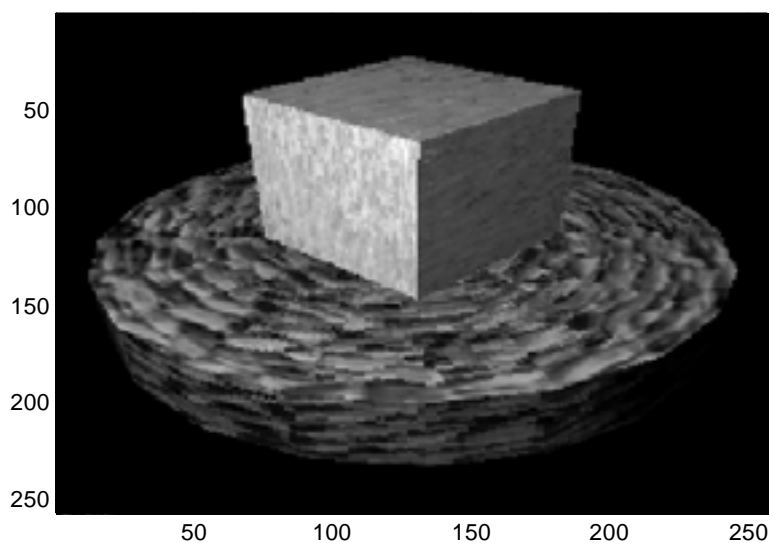


Figure 2.69 : Une image en rotation autour d'un axe vertical, séquence "cube en rotation".

Les résultats représentés par les figures 2.70 et 2.71 sont associés à l'algorithme des moindres carrés basée sur un schéma simple. Le champ estimé à la figure 2.70 montre que l'objet en question est parfaitement identifié. Les vecteurs de déplacement décrivent qualitativement le mouvement réel observé. Quantitativement, l'interprétation du résultat peut être déduite de la reconstruction de l'image  $t$  à partir de l'image  $t + Te$ . Pour ce faire, on interpole d'abord l'image  $I(x + dx, y + dy, t + Te)$  et ensuite, on calcule l'erreur en faisant la différence entre l'image interpolée et l'image d'origine  $t$ . L'image de l'erreur de reconstruction obtenue est présentée à la figure 2.71.

D'après cette figure, les erreurs sont nettement visibles sur certaines zones. Différents facteurs peuvent être à l'origine de ces erreurs :

- Une mauvaise interpolation ;
- Une mauvaise estimation du déplacement.

Sans tenir compte du premier facteur, on peut dire que les erreurs mises en évidence sur la figure 2.71 proviennent de l'erreur d'estimation avec l'utilisation de l'approche basée schéma simple. Cette interprétation ne restera valable qu'après comparaison de ce premier résultat avec les résultats de deux autres approches. La comparaison quantitative des erreurs de reconstruction pour les trois méthodes est présentée au tableau 2.2.

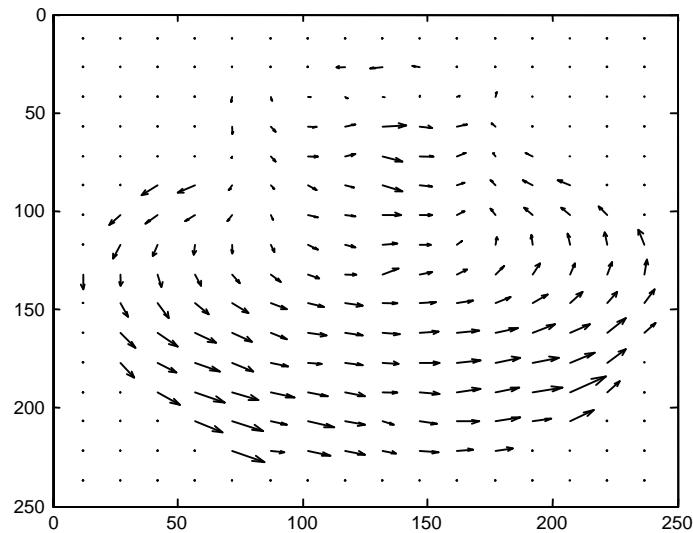


Figure 2.70 : Champ de déplacement estimé avec la méthode des moindres carrés basée mouvement direct, selon un modèle à quatre paramètres ; séquence "cube en rotation".

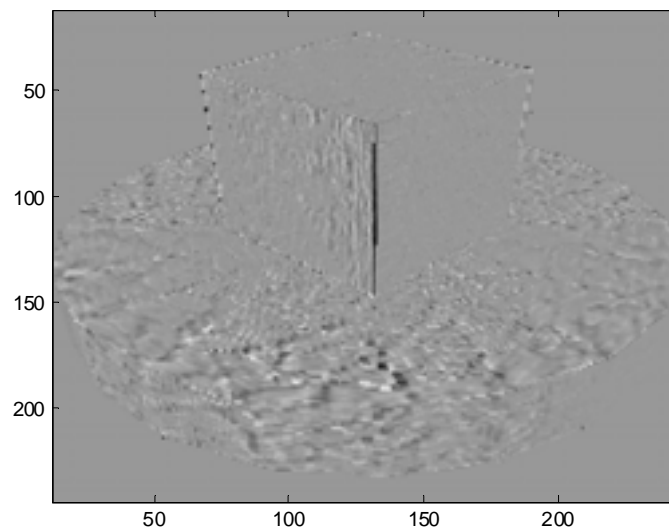


Figure 2.71 : Erreur de reconstruction du niveau de gris ; différence temporelle entre l'image  $t$  et l'image interpolée  $t + Te$ , interpolation avec les déplacements estimés à l'aide de la méthode des moindres carrés basée schéma simple ; séquence "cube en rotation".

La figure 2.72 représente le champ de déplacement associé à l'approche basée sur un schéma direct-rétrograde et cela sans traitement du terme non linéaire. A la différence du résultat de l'approche basée sur un schéma simple, le champ rotationnel est ici qualitativement régulier. L'aspect quantitatif des résultats peut être interprété à l'aide de la figure 2.73. Selon cette

figure, la qualité visuelle de l'image reconstruite est meilleure par rapport à l'erreur précédente (figure 2.71). Nous pouvons donc en déduire à première vue que la présente méthode est très efficace. L'aspect quantitatif de l'erreur d'estimation et de la variance de l'erreur est illustré au tableau 2.2.

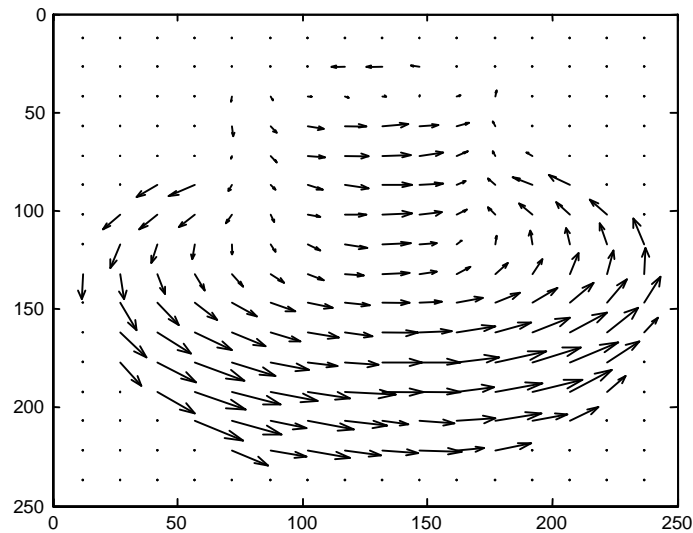


Figure 2.72 : Champ de déplacement estimé avec la méthode des moindres carrés basée sur un schéma direct-rétrograde, sans prise en compte du terme non linéaire selon la variante n°2 ; modèle à quatre paramètres, séquence "cube en rotation".

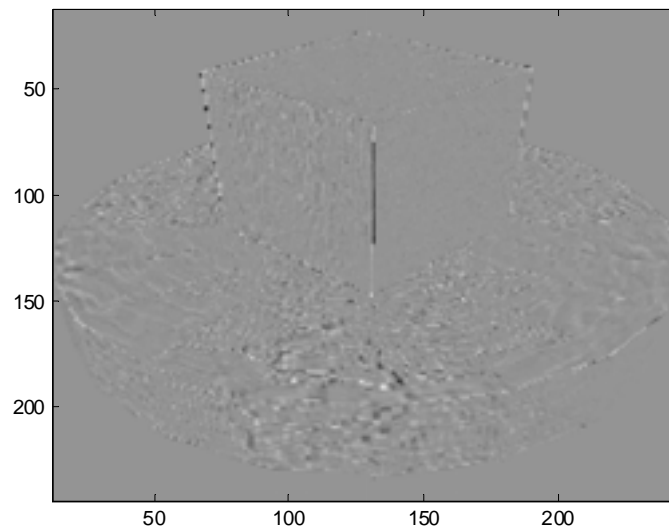


Figure 2.73 : Erreur de reconstruction du niveau de gris ; différence temporelle entre l'image  $t$  et l'image interpolée  $t + Te$ , interpolation avec les déplacements estimés à l'aide de la méthode des moindres carrés basée schéma direct-rétrograde ; sans traitement du terme non linéaire ; séquence "cube en rotation".

La troisième expérimentation consiste à évaluer les performances de l'approche basée schéma direct-rétrograde dans laquelle est traité le terme non linéaire. Les figures 2.74 et 2.75 représentent respectivement le champ de déplacement et l'erreur de reconstruction. Comme avec les deux méthodes précédentes, la figure 2.74 montre que l'on détecte effectivement les variations d'illuminations sur l'objet en mouvement et cela sans dégradation du champ de déplacement. En ce qui concerne la figure 2.75, sa comparaison avec les deux résultats précédents permet de constater visuellement que l'erreur résiduelle est nettement réduite.

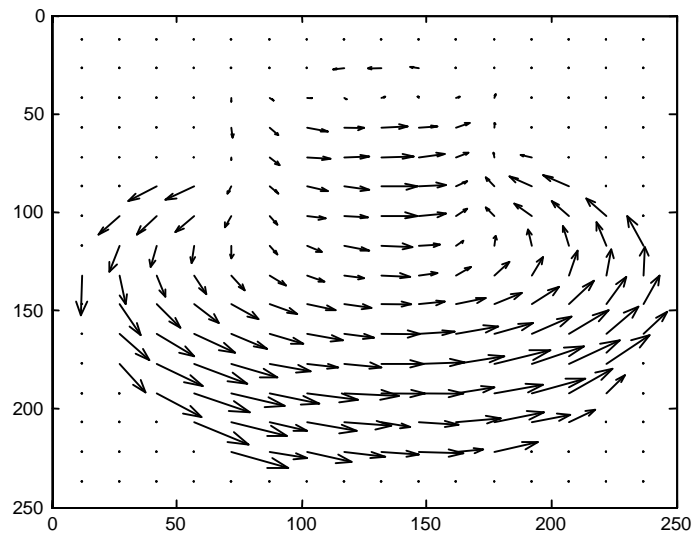


Figure 2.74 : Champ de déplacement estimé avec la méthode des moindres carrés basée sur un schéma direct-rétrograde, avec traitement du terme non linéaire selon la variante n°6 ; modèle à quatre paramètres ; séquence "cube en rotation".

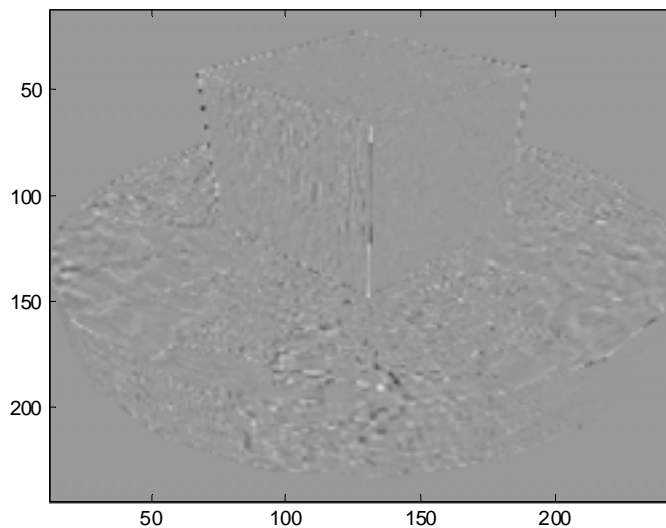


Figure 2.75 : Erreur de reconstruction du niveau de gris ; différence temporelle entre l'image  $t$  et l'image interpolée  $t + Te$ , interpolation avec les déplacements estimés à l'aide de la méthode des moindres carrés basée schéma direct-rétrograde, avec traitement du terme non linéaire ; séquence "cube en rotation".

Le tableau 2.2 fait la comparaison des résultats obtenus. L'erreur indiquée dans ce tableau est la moyenne des erreurs de reconstruction sur toute l'image. De même, la variance présente la moyenne des valeurs estimées sur la totalité de l'image.

On peut observer que l'erreur est plus faible pour l'approche basée sur un schéma direct-rétrograde dans laquelle le terme non linéaire est pris en compte (résultat de la dernière ligne du tableau). La variance de l'erreur est plus ou moins du même ordre de grandeur pour les deux approches basées sur le formalisme direct-rétrograde. Les résultats montrent que le fait de prendre en compte et de traiter la non linéarité permet d'estimer le paramètre de rotation d'une manière précise. Les résultats sont nettement meilleurs (faible biais et faible variance) pour les approches basées schéma direct-rétrograde, comparés à ceux de l'approche basée sur un schéma simple.

Méthodes \ Caractéristiques	Erreur de reconstruction	Variance de l'erreur de reconstruction
MC simple	-0.1652	84.5502
MC direct-rétrograde à équation linéaire	-0.0148	43.9701
MC direct-rétrograde à équation non linéaire	4.0744e-004	45.4962

Tableau 2.2 : Erreur de reconstruction et variance de l'erreur pour les trois approches testées, expérimentation sur la séquence "cube en rotation".

Dans cette expérimentation, nous avons choisi une séquence ayant un mouvement de rotation dominant. Les simulations ont montré un taux de réussite satisfaisant avec les deux méthodes basées flot direct-rétrograde (sans et avec traitement de la non linéarité). En particulier, la méthode basée sur un schéma direct-rétrograde avec prise en compte du terme non linéaire fournit quantitativement un résultat satisfaisant en terme d'erreur de reconstruction. Dans les expérimentations suivantes, nous continuons de tester cette performance sur le paramètre de divergence.

#### e) Expérimentation sur la séquence "nuts", mouvement de divergence

Dans la séquence suivante (figure 2.76), les images sont également prises avec les mêmes conditions d'environnement mais avec un mouvement de translation de la caméra le long de l'axe optique (effet de divergence entre chaque trame).

Les résultats associés à la méthode basée sur un schéma simple sont présentés aux figures 2.77 et 2.78. Quant à l'approche basée sur un schéma direct-rétrograde avec modèle linéaire (terme non linéaire négligé), le champ de déplacement et l'image d'erreur de reconstruction sont illustrés aux figures 2.79 et 2.80.

Le dernier test est consacré à la méthode basée sur un schéma direct-rétrograde avec modèle non linéaire (terme non linéaire pris en compte). Les résultats sont illustrés aux figures 2.81 et 2.82.

La comparaison de ces méthodes est effectuée au tableau 2.3. Selon ce tableau, les résultats obtenus présentent les caractéristiques principales suivantes :

- L'erreur et la variance de l'erreur associées à l'approche basée sur un schéma simple sont plus importantes ;
- Quant aux deux approches basées schéma direct-rétrograde, celle qui traite de la non linéarité apporte une amélioration en terme de biais, mais elle est largement moins bonne en



terme de variance. Entre les deux approches, les améliorations obtenues ne sont bien sur importantes, en absolu, que pour l'approche traitant la non linéarité.

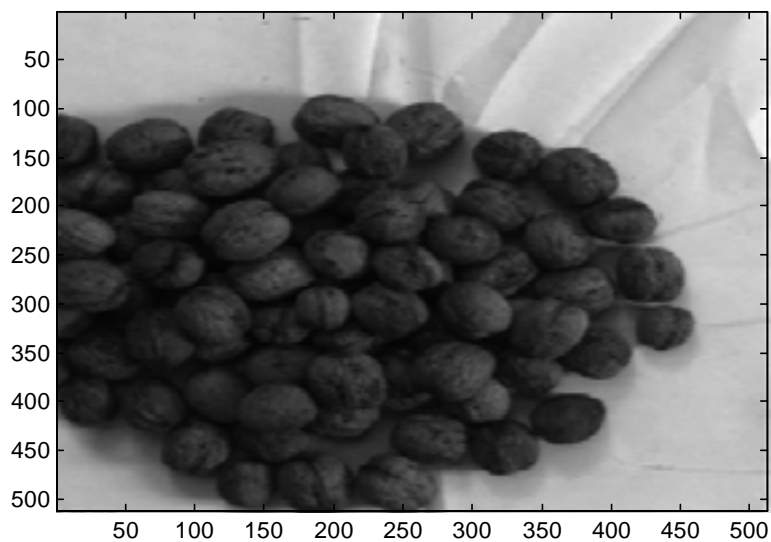


Figure 2.76 : Une image de la séquence "nuts" présentant un mouvement de divergence, source : [www.inrialpes.fr/movi/pub/images/sequences/g3\\_vp\\_tz\\_s2](http://www.inrialpes.fr/movi/pub/images/sequences/g3_vp_tz_s2).

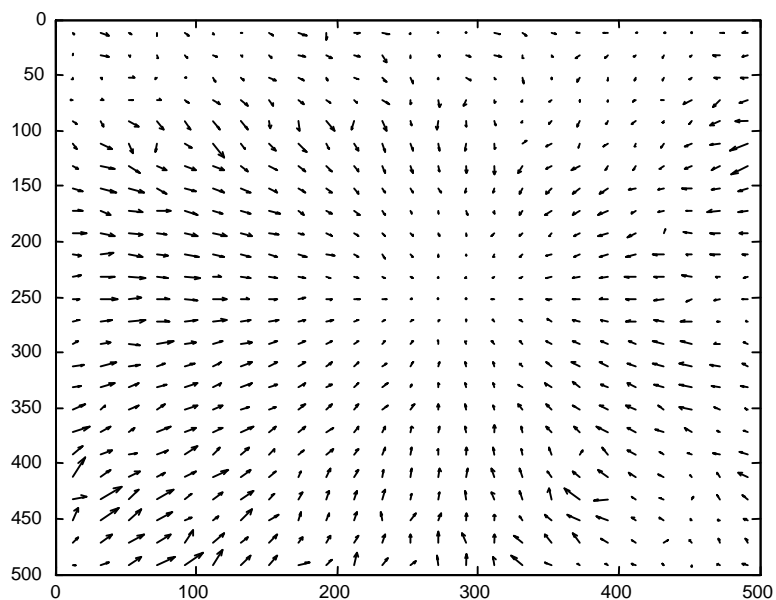


Figure 2.77 : Estimation du flot optique dû au mouvement de la caméra ; champ de déplacement estimé à l'aide de la méthode des moindres carrés basée sur un schéma simple, modèle à quatre paramètres ; séquence "nuts".

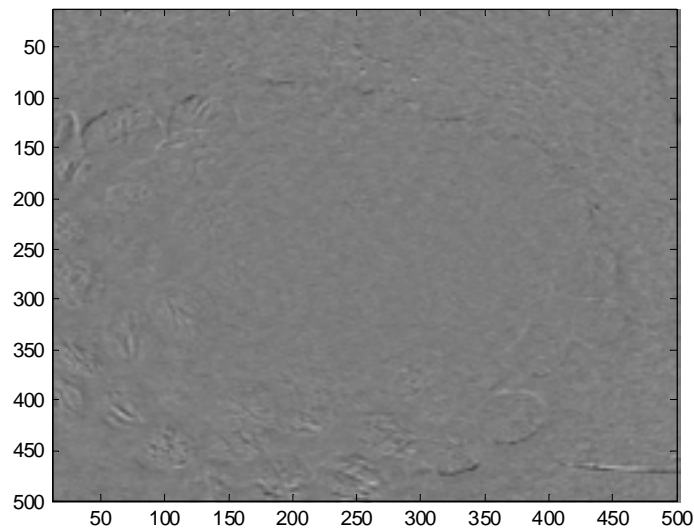


Figure 2.78 : Erreur de reconstruction du niveau de gris ; différence temporelle entre l'image  $t$  et l'image interpolée  $t + Te$ , interpolation avec les déplacements estimés à l'aide de la méthode des moindres carrés basée sur un schéma simple ; séquence "nuts".

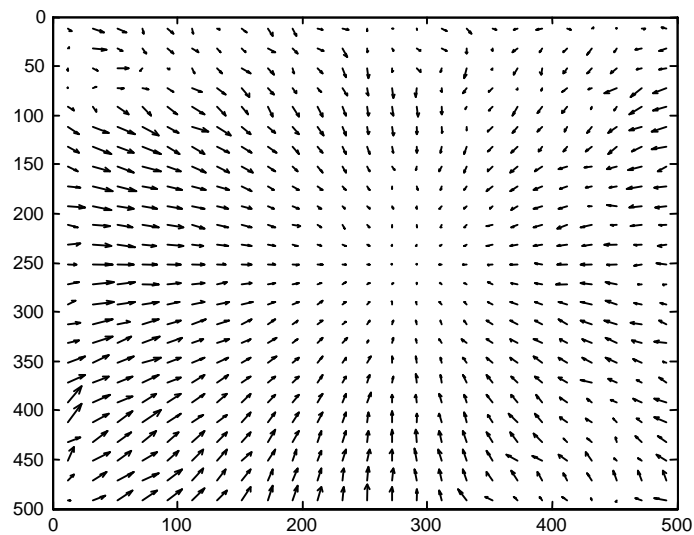


Figure 2.79 : Estimation du flot optique dû au mouvement de la caméra ; champ de déplacement estimé à l'aide de la méthode des moindres carrés basée sur un schéma direct-rétrograde, sans traitement du terme non linéaire selon la variante n°2 ; modèle à quatre paramètres ; séquence "nuts".

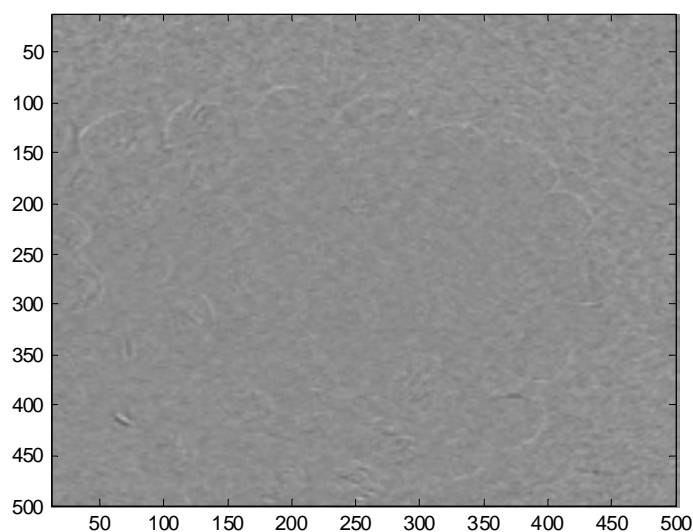


Figure 2.80 : Erreur de reconstruction du niveau de gris ; différence temporelle entre l'image  $t$  et l'image interpolée  $t + Te$ , interpolation avec les déplacements estimés à l'aide de la méthode des moindres carrés basée sur un schéma direct-rétrograde, sans traitement du terme non linéaire ; séquence "nuts".

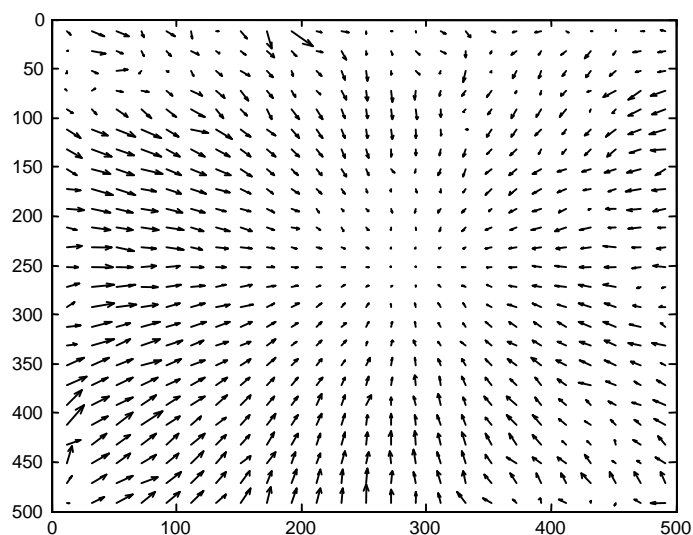


Figure 2.81 : Estimation du flot optique dû au mouvement de la caméra ; champ de déplacement estimé à l'aide de la méthode des moindres carrés basée sur un schéma direct-rétrograde, avec traitement du terme non linéaire selon la variante n°6 ; modèle à quatre paramètres ; séquence "nuts".

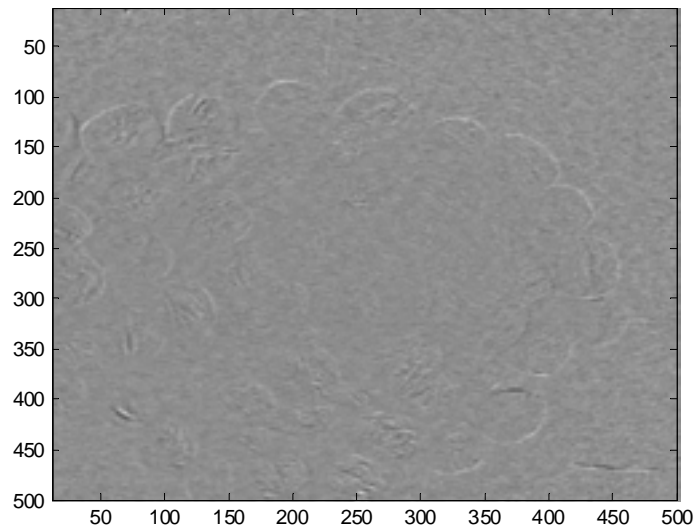


Figure 2.82 : Erreur de reconstruction du niveau de gris ; différence temporelle entre l'image  $t$  et l'image interpolée  $t + Te$ , interpolation avec les déplacements estimés à l'aide de la méthode des moindres carrés basée sur un schéma direct-rétrograde, avec traitement du terme non linéaire ; séquence "nuts".

Méthodes \ Caractéristiques	Erreur de reconstruction	Variance de l'erreur de reconstruction
MC simple	-0.1982	5.0371
MC direct-rétrograde à équation linéaire	-0.0914	3.2294
MC direct-rétrograde à équation non linéaire	-0.0665	3.9907

Tableau 2.3 : Erreur de reconstruction et variance de l'erreur pour les trois approches testées ; expérimentation sur la séquence "nuts".

#### f) Expérimentation sur la séquence "yosemite", flux non homogène

Nous choisissons, dans cette expérience, d'estimer le mouvement d'une séquence comportant à la fois un mouvement global et un mouvement local. La séquence s'appelle "yosemite". L'image est présentée par la figure 2.83.

Les figures 2.84 et 2.85 illustrent les résultats d'estimation associés à la méthode basée sur un schéma simple. La première figure montre le champ de déplacement de l'ensemble de l'image. On peut y observer un mouvement local qui résume le déplacement de l'orage par rapport à la montagne. Il s'agit des vecteurs de déplacement de la partie supérieure de l'image. Le reste du champ présente le mouvement global de l'image. La figure 2.85 montre l'erreur de reconstruction, différence entre l'image à  $t$  et celle à  $t + Te$ , cette dernière interpolée avec les déplacements estimés. On peut observer que sur certaines zones, des couches résiduelles sont nettement visibles, notamment sur les régions où l'amplitude de déplacement est importante. Il s'agit de la partie inférieure gauche de l'image. Sur le tracé du champ de déplacement, ces parasites apparaissent sous une forme de mouvements désorganisés.

Les figures 2.86 et 2.87 illustrent les résultats d'estimation associés à la méthode basée sur un schéma direct-rétrograde et cela, sans traitement du terme non linéaire. Le champ de déplacement est présenté par la figure 2.86. Par rapport au résultat précédent, on aperçoit visuellement que les parasites disparaissent sous l'effet des contraintes additionnelles. L'image 2.87 montre l'erreur de reconstruction obtenue par interpolation de l'image à  $t + Te$  avec les déplacements estimés.

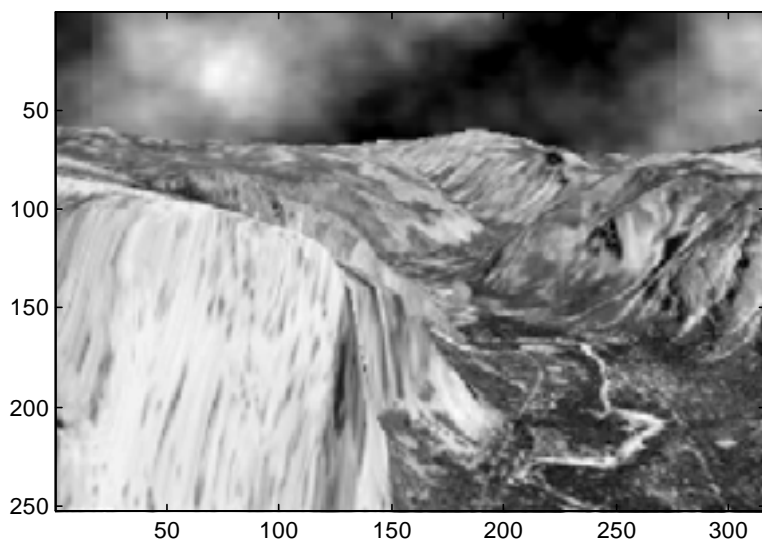


Figure 2.83 : Une image de la séquence "yosemite" présentant un mouvement global (mouvement de la caméra) et des mouvements locaux (déplacements de l'orage).

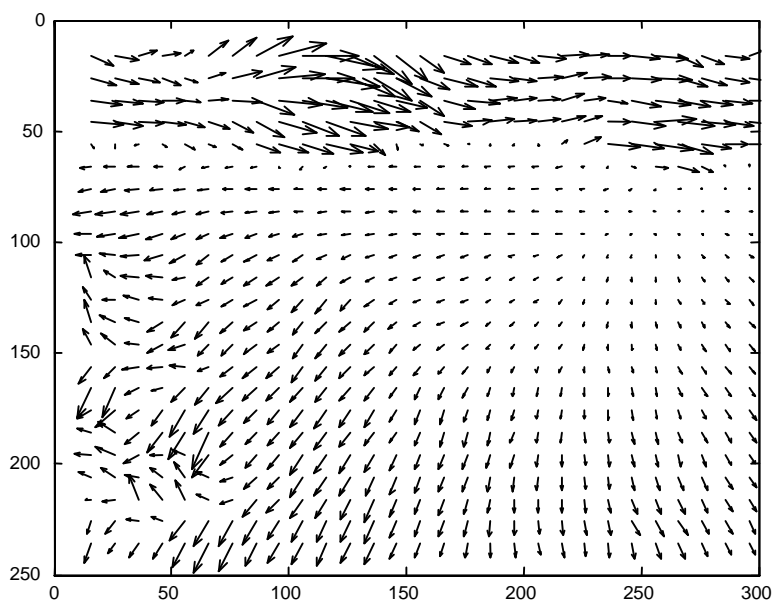


Figure 2.84 : Champ de déplacement estimé avec la méthode des moindres carrés basée sur un schéma simple ; modèle à quatre paramètres ; séquence "yosemite".

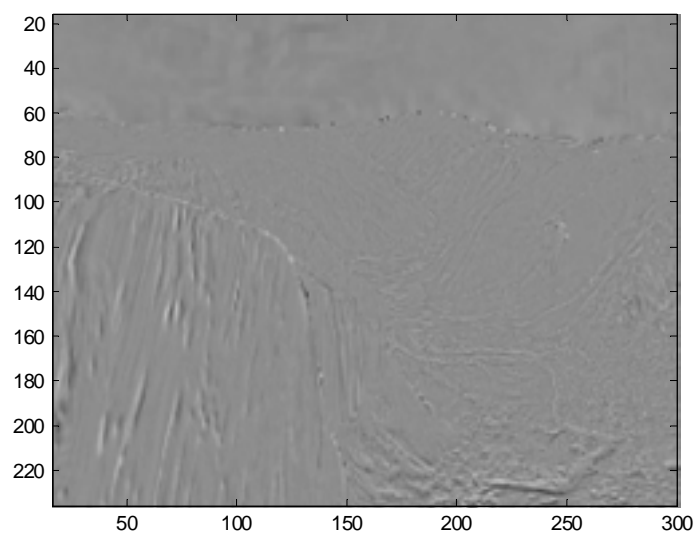


Figure 2.85 : Erreur de reconstruction du niveau de gris, différence temporelle entre l'image  $t$  et l'image interpolée  $t + Te$ , interpolation avec les déplacements estimés à l'aide de la méthode des moindres carrés basée sur un schéma simple ; séquence "yosemite".

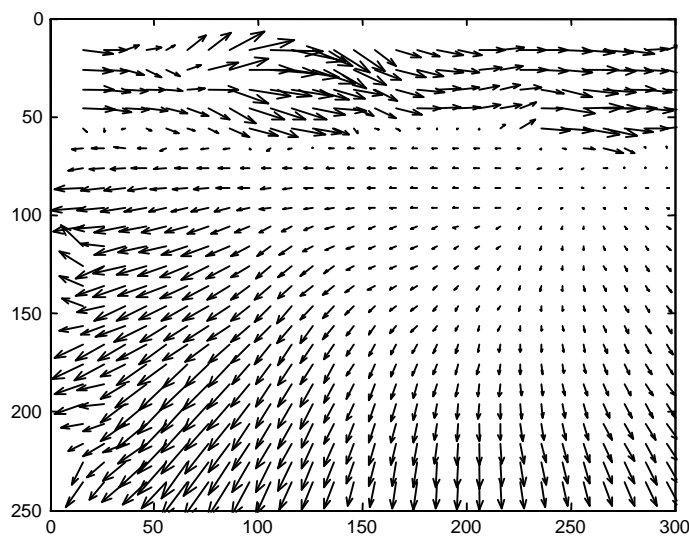


Figure 2.86 : Champ de déplacement estimé avec la méthode des moindres carrés basée sur un schéma direct-rétrograde ; terme non linéaire négligeable selon la variante n°2 ; modèle à quatre paramètres ; séquence "yosemite".

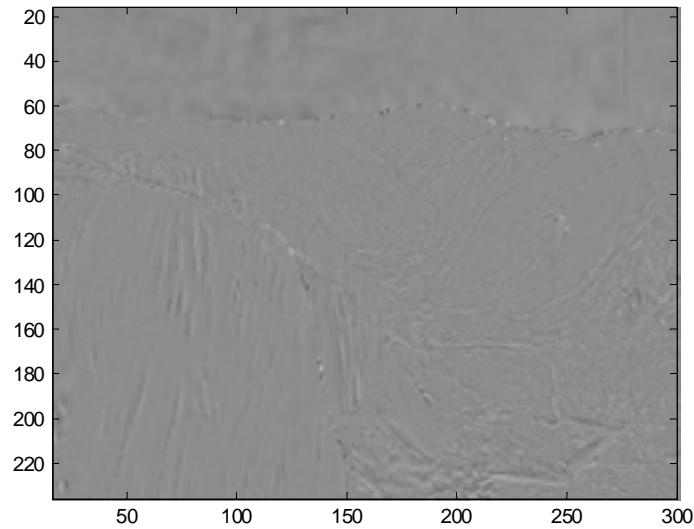


Figure 2.87 : Erreur de reconstruction du niveau de gris ; différence temporelle entre l'image  $t$  et l'image interpolée  $t + Te$ , interpolation avec les déplacements estimés à l'aide de la méthode des moindres carrés basée sur un schéma direct-rétrograde, sans prise en compte du terme non linéaire ; séquence "yosemite".

En ce qui concerne la méthode basée sur un schéma direct-rétrograde dans laquelle le terme non linéaire est pris en compte, les résultats sont illustrés aux figures 2.88 et 2.89. Le champ de déplacement est présenté à la première figure. La comparaison des figures 2.88 et 2.86 ne permet pas de constater visuellement une différence entre les résultats des deux approches basées flot direct-rétrograde proposées. La synthèse des résultats du tableau 2.4 permet d'effectuer une comparaison quantitative. La figure 2.89 correspond à l'image de l'erreur de reconstruction.

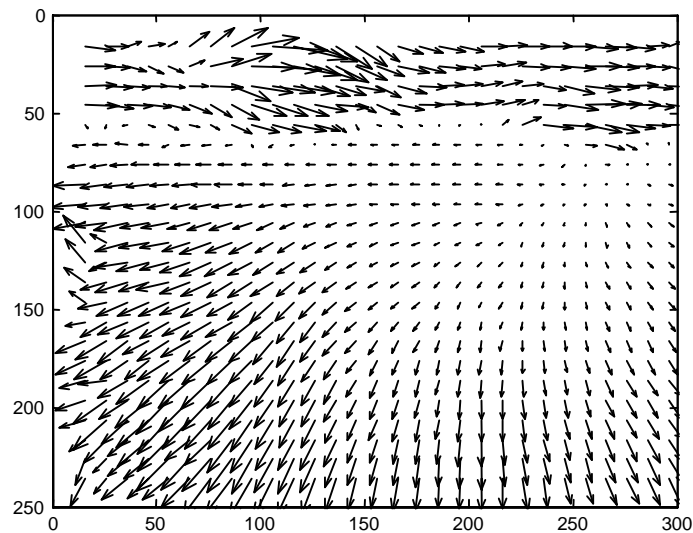


Figure 2.88 : Champ de déplacement estimé avec la méthode des moindres carrés basée sur un schéma direct-rétrograde ; terme non linéaire traité selon la variante n°6 ; modèle à quatre paramètres ; séquence "yosemite".

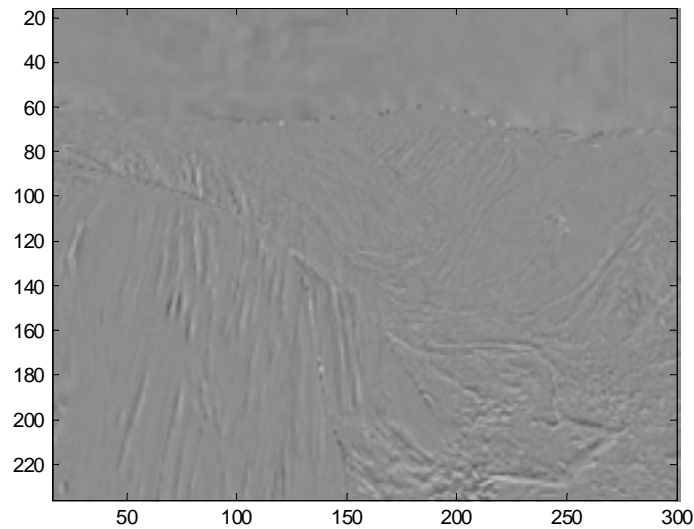


Figure 2.89 : Erreur de reconstruction du niveau de gris ; différence temporelle entre l'image  $t$  et l'image interpolée  $t + Te$ , interpolation avec les déplacements estimés à l'aide de la méthode des moindres carrés basée sur un schéma direct-rétrograde, avec prise en compte du terme non linéaire ; séquence "yosemite".

L'analyse quantitative de l'erreur et de la variance de l'erreur, pour les trois approches, est présentée dans le tableau 2.4. D'après ce tableau, l'approche direct-rétrograde dans laquelle on néglige le terme non linéaire s'avère être très efficace. On peut observer expérimentalement ici que la prise en compte du terme non linéaire dans le formalisme direct-rétrograde n'est pas adaptée pour cette séquence. En fait, l'inefficacité de cette approche vient de la nature de la séquence utilisée. En effet, cette séquence présente des discontinuités en terme de mouvement, et l'approximation itérative des composantes non linéaire semble susceptible d'être instable au cours de l'estimation. Le fait d'ignorer la non linéarité n'est donc pas pénalisant et permet de valider l'approche associée sur une séquence complexe.

Méthodes \ Caractéristiques	Erreur de reconstruction	Variance de l'erreur de reconstruction
MC simple	-0.0564	68.0292
MC direct-rétrograde à équation linéaire	0.0040	36.0257
MC direct-rétrograde à équation non linéaire	-0.0072	72.7914

Tableau 2.4 : Erreur d'estimation et variance de l'erreur pour les trois approches testées, expérimentation sur la séquence "yosemite".

Pour conclure, nous avons analysé lors de cette expérimentation la différence entre les résultats de la méthode basée sur un schéma simple et ceux des méthodes basées flot direct-rétrograde. Avec l'approche basée sur un schéma simple, nous constatons la présence des mouvements désordonnés. Nous avons proposé les méthodes basées flot direct-rétrograde pour remédier à cet artefact. Les résultats des méthodes proposées montrent que qualitativement, les mouvements estimés décrivent correctement les mouvements réels observés.



En ce qui concerne la différence entre le traitement ou non du terme non linéaire, nous avons constaté que l'algorithme qui prend en charge la modélisation du terme non linéaire (variante n°6) fournit de meilleurs résultats du point de vue du sens d'orientation du champ. En revanche, le modèle ignorant le terme non linéaire présente un intérêt au niveau de l'amplitude de déplacement et cela au moins pour les composantes du vecteur translation. L'interprétation des résultats numériques n'est donc pas évidente pour savoir laquelle de ces deux propositions est meilleure.

### 2.3.3 Discussion sur les différentes approches basées sur un schéma direct-rétrograde pour un modèle de mouvement affine

Dans ce paragraphe, nous avons fait un tour d'horizon des méthodes basées sur un schéma direct-rétrograde, et cela dans le contexte de mouvement affine. L'introduction du modèle de mouvement 2D a permis de mettre en œuvre une méthode différentielle, basée sur un schéma direct-rétrograde, permettant d'obtenir une estimation globale du mouvement, ainsi que l'estimation locale des régions par rapport au fond. Plusieurs approches, dérivées du formalisme direct-rétrograde, ont été envisagées. Les résultats montrent la validité du formalisme proposé pour l'estimation des paramètres de translation, de rotation et de divergence. La méthode d'estimation décrite dans ce paragraphe est une généralisation de l'algorithme d'estimation du mouvement à base de flot direct-rétrograde.

Les enseignements tirés de ces simulations sont les suivants :

- Par rapport à l'approche basée sur un schéma simple, les approches à base de formalisme direct-rétrograde fonctionnent correctement et restent plus robustes au bruit ;
- Le formalisme direct-rétrograde à modèle linéaire (terme non linéaire négligé) est plus adapté pour l'estimation des paramètres de translation ; en effet, la négligence du terme non linéaire permet d'introduire dans le critère à minimiser une contrainte d'uniformité ;
- Le formalisme direct-rétrograde à modèle non linéaire fournit une bonne estimation du mouvement à forte rotation et forte divergence. En effet, la prise en compte du terme non linéaire a tendance à favoriser les types de mouvement plus complexes (rotation, divergence)
- Un avantage important du formalisme proposé est sa simplicité par rapport à d'autres algorithmes de régularisation qui, parfois, exigent un calcul fastidieux.

Le procédé de pondération de données reste applicable dans toutes les méthodes proposées. Un récapitulatif des résultats est illustré sur le tableau 2.5.

Nous remarquons à travers de nombreuses expérimentations que, si nous utilisons uniquement deux images successives pour le calcul des gradients temporels, l'approche basée sur un schéma direct-rétrograde et formulée avec la fonctionnelle  $J_a$  est moins stable que celle associée à  $J_b$ . Cela s'explique par le fait que d'après l'expression de la fonctionnelle  $J_b$ , celle-ci permet d'introduire implicitement une contrainte d'uniformité entre les deux informations (directe et rétrograde). En revanche, l'utilisation d'une troisième image pour le calcul des gradients temporels à  $t + T_e$ , permet d'obtenir des résultats stables avec la fonctionnelle  $J_a$ .

## 2.4 Approche incrémentale et schéma direct-rétrograde pour un modèle paramétrique

Cette présentation concerne l'application de l'approche incrémentale à l'équation paramétrique du modèle de mouvement, et cela selon le formalisme direct-rétrograde.

D'après l'équation (249), le déplacement rétrograde est égal au déplacement direct auquel on ajoute un terme non linéaire. L'application d'un incrément sur le déplacement rétrograde implique un terme d'incrément sur la non-linéarité. En effet, nous avons :

$$\Delta d_r = \Delta[d_d + NL(\Theta, P)] \quad (281)$$

Paramètres \ Méthodes		MC simple (basée mouvement direct)	MC direct-rétrograde sans prise en compte du terme non linéaire	MC direct-rétrograde avec traitement du terme non linéaire
Mouvement à translation dominant et à faible rotation et/ou divergence	$t_x$	Peu adaptée	Très adaptée	Adaptée
	$t_y$	Peu adaptée	Très adaptée	Adaptée
	$\theta$	Adaptée	Adaptée	Adaptée
	$K$	Adaptée	Adaptée	Adaptée
	<b>Champ reconstruit</b>	Moins significatif	Plus cohérent	Peu significatif
Mouvement de forte rotation et/ou divergence et de faible translation	$t_x$	Peu adaptée	Adaptée	Adaptée
	$t_y$	Peu adaptée	Adaptée	Adaptée
	$\theta$	Peu adaptée	Peu adaptée	Très adaptée
	$K$	Peu adaptée	Peu adaptée	Très adaptée
	<b>Champ reconstruit</b>	Moins significatif	Moins significatif	Plus cohérent
Mouvement de très faible rotation et/ou divergence et/ou translation	$t_x$	Peu adaptée	Adaptée	Peu adaptée
	$t_y$	Peu adaptée	Adaptée	Peu adaptée
	$\theta$	Adaptée	Adaptée	Adaptée
	$K$	Adaptée	Adaptée	Adaptée
	<b>Champ reconstruit</b>	Moins significatif	Plus cohérent	Peu significatif

Tableau 2.5 : Évaluation des différentes approches pour l'estimation de paramètres du mouvement.

Par conséquent, nous aurons deux termes correctifs à estimer dans une même phase d'estimation. L'un correspond à l'incrément du déplacement et l'autre correspond à l'incrément du terme non linéaire.

$$\delta d^{(i+1)} = d^{(i)} + \delta d^{(i)} \quad (282)$$

$$NL^{(i+1)} = NL^{(i)} + \delta NL^{(i)} \quad (283)$$

Dans cette étude, nous n'allons pas traiter l'incrément sur la quantité non linéaire. La méthode décrite ci-après consiste à corriger le déplacement par un processus incrémental sans pour autant rajouter un terme correctif sur la quantité non linéaire.

\*Nous considérons dans un premier temps la fonctionnelle à somme des termes quadratiques des erreurs. Les étapes du traitement sont les suivantes :

- Estimation des paramètres initiaux à l'aide d'une des approches décrites au paragraphe 2.3.

- Calcul des paramètres correctifs  $\delta\Theta = (\delta t_x, \delta k, \delta\theta, \delta t_y)^T$  en minimisant la fonctionnelle suivante :

$$J_a = \sum_{\Omega} \left\{ \left[ \begin{pmatrix} \delta t_x + \delta kx - \delta\theta y \\ \delta t_y + \delta\theta x + \delta ky \end{pmatrix} \bar{\nabla} I^{t+dt}(x + dx_d, y + dy_d) + [I(x + dx_d, y + dy_d, t + dt) - I(x, y, t)] \right]^2 + \left[ \begin{pmatrix} \delta t_x + \delta kx - \delta\theta y \\ \delta t_y + \delta\theta x + \delta ky \end{pmatrix} \bar{\nabla} I^t(x - dx_r, y - dy_r) + [I(x - dx_r, y - dy_r, t) - I(x, y, t + dt)] \right]^2 \right\} \quad (284)$$

Le vecteur  $\vec{d}_d = (dx_d, dy_d)^T$  correspond au déplacement direct estimé dans la phase initiale. Quant au déplacement rétrograde  $\vec{d}_r = (dx_r, dy_r)^T$ , sa valeur vaut  $\vec{d}_r = \vec{d}_d$  ou  $\vec{d}_r = \vec{d}_d + NL$  selon qu'on choisisse l'approche sans ou avec prise en compte du terme non linéaire (c.f paragraphe 2.3). La minimisation du critère (284), par rapport aux variables  $(\delta t_x, \delta k, \delta\theta, \delta t_y)$  et au sens des moindres carrés, donne le résultat suivant :

$$\delta\Theta = -(C_1^T C_1 + C_2^T C_2)^{-1} (C_1^T D_1 + C_2^T D_2) \quad (285)$$

Les matrices de données sont présentées de la manière suivante :

$C_1$  est une matrice  $4 \times n$  qui exprime les gradients spatiaux de l'image  $t$ , décalée selon le rattrapage de l'information dans le sens descendant du temps :

$$C_1 = \begin{bmatrix} I'_x(x_1 - dx_r, y_1 - dy_r) & x_1 I'_x(x_1 - dx_r, y_1 - dy_r) + y_1 I'_y(x_1 - dx_r, y_1 - dy_r) & & \\ \vdots & \vdots & & \\ I'_x(x_n - dx_r, y_n - dy_r) & x_n I'_x(x_n - dx_r, y_n - dy_r) + y_n I'_y(x_n - dx_r, y_n - dy_r) & & \\ x_1 I'_y(x_1 - dx_r, y_1 - dy_r) - y_1 I'_x(x_1 - dx_r, y_1 - dy_r) & I'_y(x_1 - dx_r, y_1 - dy_r) & & \\ \vdots & \vdots & & \\ x_n I'_y(x_n - dx_r, y_n - dy_r) - y_n I'_x(x_n - dx_r, y_n - dy_r) & I'_y(x_n - dx_r, y_n - dy_r) & & \end{bmatrix} \quad (286)$$

$C_2$  exprime les gradients spatiaux de l'image  $t + dt$  décalée selon le rattrapage de l'information dans le sens ascendant du temps :

$$C_2 = \begin{bmatrix} I_x^{t+dt}(x_1 + dx_d, y_1 + dy_d) & x_1 I_x^{t+dt}(x_1 + dx_d, y_1 + dy_d) + y_1 I_y^{t+dt}(x_1 + dx_d, y_1 + dy_d) & & \\ \vdots & \vdots & & \\ I_x^{t+dt}(x_n + dx_d, y_n + dy_d) & x_n I_x^{t+dt}(x_n + dx_d, y_n + dy_d) + y_n I_y^{t+dt}(x_n + dx_d, y_n + dy_d) & & \\ x_1 I_y^{t+dt}(x_1 + dx_d, y_1 + dy_d) - y_1 I_x^{t+dt}(x_1 + dx_d, y_1 + dy_d) & I_y^{t+dt}(x_1 + dx_d, y_1 + dy_d) & & \\ \vdots & \vdots & & \\ x_n I_y^{t+dt}(x_n + dx_d, y_n + dy_d) - y_n I_x^{t+dt}(x_n + dx_d, y_n + dy_d) & I_y^{t+dt}(x_n + dx_d, y_n + dy_d) & & \end{bmatrix} \quad (287)$$

Les matrices  $D_1$  et  $D_2$  correspondent à la DFD associée respectivement aux mouvements rétrograde et direct. L'expression de  $D_1$  ressemble à l'équation (213) dans laquelle le déplacement est associé à  $\vec{d}_d = (dx_d, dy_d)^T$ . Quant à  $D_2$ , sa valeur est donnée par la relation (214) dans laquelle est introduit le déplacement  $\vec{d}_r = (dx_r, dy_r)^T$ .

- Calcul des nouveaux paramètres selon l'équation (à la  $i^{\text{ème}}$  itération) :

$$\Theta^{(i+1)} = \Theta^{(i)} + \delta\Theta^{(i)} \quad (288)$$

- Retour à la deuxième étape pour aboutir à un processus itératif.

\* Quant à la deuxième fonctionnelle, les étapes de calcul sont présentées comme suit :

- Estimation des paramètres initiaux suivant l'une des méthodes décrites au paragraphe 2.3.

- Calcul des paramètres correctifs  $\delta\Theta = (\delta t_x, \delta k, \delta\theta, \delta t_y)^T$  en minimisant la fonctionnelle de l'équation (289).

$$J_b = \sum_{\Omega} \left\{ \left[ \begin{array}{c} \delta t_x + \delta kx - \delta\theta y \\ \delta t_y + \delta\theta x + \delta ky \end{array} \right] \left[ \nabla \bar{I}^{t+dt}(x + dx_d, y + dy_d) + \nabla \bar{I}^t(x - dx_r, y - dy_r) \right] + \right. \\ \left. \left[ I(x + dx_d, y + dy_d, t + dt) - I(x, y, t) \right] + \left[ I(x - dx_r, y - dy_r, t) - I(x, y, t + dt) \right] \right\}^2 \quad (289)$$

Le vecteur  $\vec{d}_d = (dx_d, dy_d)^T$  correspond au déplacement direct estimé dans la phase initiale. Quant au déplacement rétrograde  $\vec{d}_r = (dx_r, dy_r)^T$ , il peut être estimé avec ou sans considération du terme non linéaire.

La minimisation de la fonctionnelle (289) au sens des moindres carrés, par rapport aux variables  $(\delta t_x, \delta\theta, \delta k, \delta t_y)$ , donne le résultat suivant :

$$\delta\Theta = -[(C_1 + C_2)^T (C_1 + C_2)]^{-1} [(C_1 + C_2)^T (D_1 + D_2)] \quad (290)$$

Les matrices  $C_1$ ,  $C_2$ ,  $D_1$  et  $D_2$  sont calculées de la même manière que l'approche précédente.

- Calcul des nouveaux paramètres, à la  $i^{ème}$  itération, selon l'équation suivante :

$$\Theta^{(i+1)} = \Theta^{(i)} + \delta\Theta^{(i)} \quad (291)$$

- Retour à la deuxième étape pour passer à la prochaine itération.

Le critère d'arrêt pour ces deux approches peut être défini soit en fixant le nombre d'itérations maximal, soit en utilisant un seuil des DFD en dessous duquel l'estimation est supposée bonne.

## 2.5 Extension du schéma direct-rétrograde à $n$ images ( $n > 2$ )

### 2.5.1 Description

De nombreuses méthodes ont été proposées pour estimer un mouvement à partir de  $n$  images. A titre d'illustration, certains auteurs ont pris la valeur de  $n$  égale à trois. Les termes d'erreur suivants ont été définis :

- L'erreur directe est définie entre les trames  $I(p, t)$  et  $I(p + \vec{d}_d, t + dt)$  ;

- L'erreur rétrograde est envisagée entre les trames  $I(p, t)$  et  $I(p - \vec{d}_r, t - dt)$ .

Nous rappelons brièvement une version simplifiée de cette approche en mode incrémental. Selon la figure 2.90, l'hypothèse de conservation de la luminance conduit à écrire l'expression de l'erreur suivante :

$$\varepsilon_d = I(p, t) - I(p + \vec{d}_d, t + dt) \quad (292)$$

Le déplacement  $\vec{d}_d$  est défini par :

$$\vec{d}_d = X(p)\Theta dt \quad (293)$$

$\Theta = [a_0 \ a_1 \ a_2 \ a_3 \ a_4 \ a_5 \ a_6 \ a_7]^T$  est un vecteur des paramètres d'un modèle linéaire projectif (cf. tableau 1.1). La quantité  $X(p)$  est définie par :

$$X(p) = \begin{bmatrix} x^2 & xy & x & y & 1 & 0 & 0 & 0 \\ xy & y^2 & 0 & 0 & 0 & x & y & 1 \end{bmatrix} \quad (294)$$

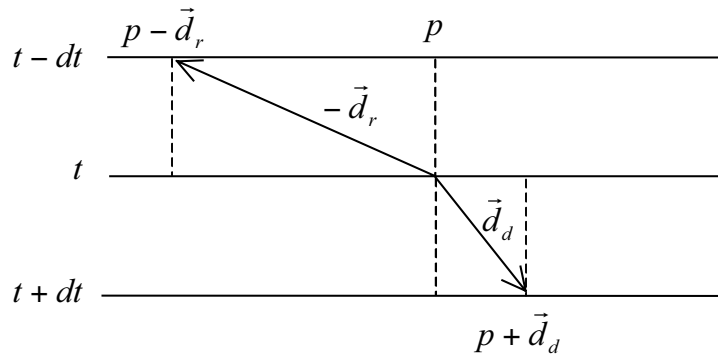


Figure 2.90 : Exemple d'estimation du flot optique à partir des trois images.

En introduisant l'approche incrémentale telle que :

$$\vec{d}_d^{(i+1)} = \vec{d}_d^{(i)} + \delta\vec{d}^{(i)} \quad (295)$$

et  $\delta\vec{d}^{(i)} = [X(p)\Theta^{(i+1)} - X(p)\Theta^{(i)}]dt$ , le développement de l'erreur par rapport au terme infinitésimal  $\delta\vec{d}$  nous conduit à écrire l'expression de l'erreur directe suivante :

$$\varepsilon_d \approx [I(p, t) - I(p + \vec{d}_d^{(i)}, t + dt)] - \vec{\nabla}I'(p)[X(p)\Theta^{(i+1)} - X(p)\Theta^{(i)}] \quad (296)$$

En supposant que le mouvement est constant dans l'intervalle de temps entre deux trames, centré sur la trame courant  $I(p, t)$ , l'erreur rétrograde peut être exprimée par :

$$\varepsilon_r \approx [I(p, t) - I(p - \vec{d}_r^{(i)})] + \vec{\nabla}I'(p)[X(p)\Theta^{(i+1)} - X(p)\Theta^{(i)}] \quad (297)$$

La fonctionnelle à minimiser est définie comme suit :

$$J = \sum_{\Omega} \varepsilon_d^2 + \sum_{\Omega} \varepsilon_r^2 \quad (298)$$

La minimisation de cette fonctionnelle par un procédé itératif aboutit au résultat suivant :

$$\Theta^{(k+1)} = \frac{1}{2} \left\{ \sum_{\Omega} X(p)^T \vec{\nabla}I'(p) \cdot \vec{\nabla}I'^{(i+1)}(p)^T X(p) \right\}^{-1} \times \left\{ \sum_{\Omega} X(p)^T \vec{\nabla}I'(p) \left[ 2\vec{\nabla}I'(p)^T [X(p)\Theta^{(i)}] - I(p, t + dt) + I(p, t - dt) \right] \right\} \quad (299)$$

Que peut-on constater de ce formalisme ?

- Entre  $t$  et  $t + dt$ , uniquement le mouvement direct est considéré ;
- Entre  $t - dt$  et  $t$ , seul le mouvement rétrograde est pris en compte ;
- Ces déplacements sont tous décrits par un même jeu de paramètres.

Dans ce paragraphe, nous allons établir une version étendue du formalisme direct-rétrograde. L'étude sur deux instants successifs ( $t$  et  $t + dt$ ) a été traitée lors des paragraphes précédents. La même philosophie sera appliquée ici. L'idée est d'essayer de manipuler le maximum d'informations en appliquant également le même formalisme entre les instants  $t - dt$  et  $t$ . Enfin, nous avons deux mouvements directs et deux mouvements rétrogrades (figure 2.91).

Le principe est basé sur l'hypothèse d'invariance de la luminance de chaque point, lors de son déplacement. Théoriquement, les deux DFD suivantes peuvent être définies :

- L'erreur relative au déplacement direct entre  $t - dt$  et  $t$  :

$$\varepsilon_d = I(p, t - dt) - I(p + \vec{d}_d, t) \quad (300)$$

- L'erreur relative au déplacement rétrograde entre  $t - dt$  et  $t$  :

$$\varepsilon_r = I(p, t) - I(p - \vec{d}_r, t - dt) \quad (301)$$

- L'erreur relative au déplacement direct entre  $t$  et  $t + dt$  :

$$\varepsilon'_d = I(p, t) - I(p + \vec{d}'_d, t + dt) \quad (302)$$

- L'erreur relative au déplacement rétrograde entre  $t$  et  $t + dt$  :

$$\varepsilon'_r = I(p, t + dt) - I(p - \vec{d}'_r, t) \quad (303)$$

En supposant que dans l'intervalle  $[t - dt, t + dt]$ , tous les pixels d'un voisinage considéré subissent une même transformation, les quatre modèles du mouvement utilisent un même jeu de paramètres sans pour autant être égaux. Le critère à minimiser peut s'exprimer de plusieurs manières. Nous retiendrons les formulations suivantes :

- La première fonctionnelle est définie par la somme des termes quadratiques des erreurs :

$$J_a = \sum_{\Omega} \varepsilon_d^2 + \sum_{\Omega} \varepsilon_r^2 + \sum_{\Omega} \varepsilon'_d{}^2 + \sum_{\Omega} \varepsilon'_r{}^2 \quad (304)$$

- La deuxième fonctionnelle correspond au carré de la somme algébrique des erreurs :

$$J_b = \sum_{\Omega} (\varepsilon_d + \varepsilon_r + \varepsilon'_d + \varepsilon'_r)^2 \quad (305)$$

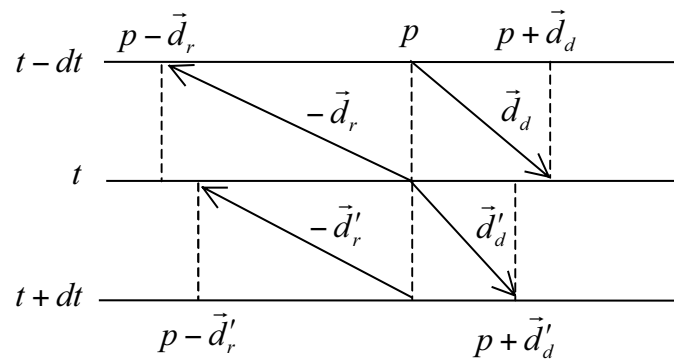


Figure 2.91 : Formalisme direct-rétrograde sur trois images.

### 2.5.2 Résultats expérimentaux

Ce test porte sur des résultats obtenus sur la séquence taxi pour  $n$  égal à 5. Sur la figure 2.92, le résultat de l'estimation du mouvement des objets ① et ② est affiché. L'approche basée sur un schéma direct-rétrograde à  $n$  images, avec un modèle de translation, est utilisée. Le mouvement de rotation de l'objet ② est plus prononcé ; ce qui est cohérent avec le mouvement réel de l'objet. Le champ de déplacement montre ainsi que l'utilisation d'une séquence à plus de deux images contribue à l'efficacité de l'estimation.

### 2.5.3 Discussion sur l'approche basée sur un schéma direct-rétrograde à $n$ images

Nous avons développé une extension du formalisme basé mouvement direct-rétrograde pour  $n$  images ( $n > 2$ ). L'utilisation de plusieurs trames permet de mieux estimer les gradients temporels et ainsi, de renforcer les contraintes dans le critère à minimiser. Cette extension introduit également le filtrage temporel et permet donc de lisser le champ de déplacement. L'utilisation d'une séquence à plus de deux images est conseillée pour une meilleure détection des objets en mouvement. En effet, la robustesse contre les fluctuations temporelles aléatoires de mouvement est améliorée en tenant compte d'un plus grand intervalle temporel (en augmentant le nombre d'images).

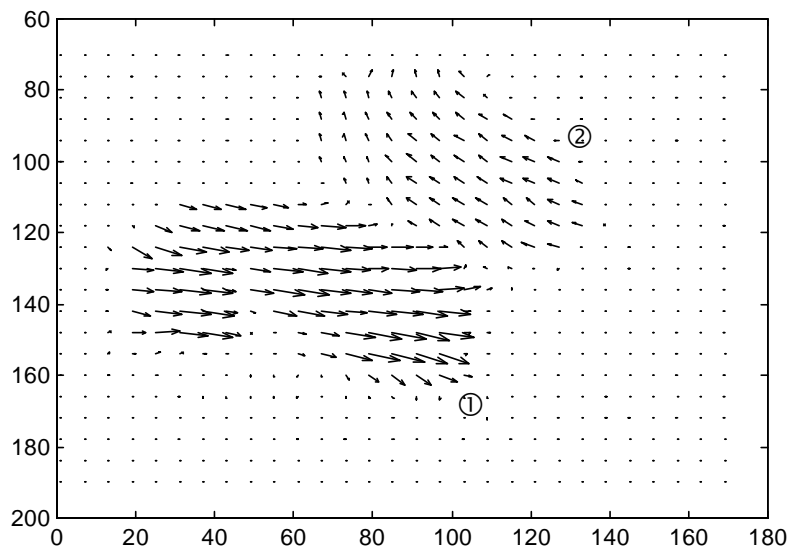


Figure 2.92 : Champ de déplacement obtenu avec l'approche basée direct-rétrograde, en utilisant 5 images successives.

## 2.6 Résumé du chapitre

Dans ce chapitre, nous avons présenté différents schémas de description du formalisme direct-rétrograde en modèle paramétrique. Pour le mouvement direct, sa modélisation ressemble à un modèle classique à quatre paramètres (un vecteur de translation et une matrice de transformation). Quant au mouvement rétrograde, sa modélisation a permis d'avoir une équation qui n'est que la transformation inverse du mouvement direct.

En partant de ces deux informations, la mise en équation du flot optique a abouti à une fonctionnelle non linéaire. Différentes approches sont retenues pour la minimisation du critère obtenu. En résumé, les approches présentées peuvent être classées en deux familles :

- La première famille consiste à mettre la fonctionnelle sous une forme linéaire, en ignorant les termes non linéaires. Toutes les formes quadratiques dans le critère à minimiser sont négligées et on se retrouve avec un modèle classique dans lequel les composantes du vecteur déplacement sont une fonction linéaire des paramètres.
- La seconde famille consiste à considérer le critère à minimiser tel qu'il est défini. Ce critère est défini par des termes linéaires et des termes quadratiques. Nous avons proposé une modélisation de la non linéarité et estimer les paramètres de ce modèle ainsi que ceux du mouvement d'une manière itérative.

Toutes les méthodes résumées par ces familles se différencient, soit selon le type de fonctionnelle utilisée ( $J_a$  ou  $J_b$ ), soit selon la méthode du traitement de la non linéarité. Pour le traitement de la non linéarité, une première approche est proposée dans les variantes n°3 et n°4 ; une autre approche est présentée dans les variantes n°5 et n°6.

Différentes expérimentations ont été effectuées sur diverses séquences. Dans ce contexte, nous avons essayé d'évaluer les performances de chacune des méthodes. Les tests ont montré que le fait de négliger le terme non linéaire a permis d'introduire une contrainte d'uniformité dans la définition du critère à minimiser. Les approches associées à cette condition sont favorisées en estimation des composantes de translation. Ces approches sont avantageuses pour les séquences à translation dominante.

Cette étude permet également de voir l'intérêt de la fonctionnelle non linéaire. Les approches utilisant cette fonctionnelle sont efficaces sur l'estimation du paramètre de rotation. Il est donc nécessaire de traiter les termes non linéaires (au lieu de les négliger) car cela permet d'avoir une bonne précision sur l'estimation des paramètres du mouvement.

L'approche incrémentale associée au modèle paramétrique est aussi développée dans ce chapitre. Le terme de correction est appliquée uniquement sur le déplacement mais pas sur le terme non linéaire.

Enfin, une extension du formalisme direct-rétrograde sur  $n$  images et cela avec un modèle paramétrique, est présentée. Les résultats sont intéressants notamment pour la segmentation spatio-temporelle.

Le tableau 2.6 fait la synthèse des formalismes utilisés dans ce chapitre.

<b>Formalisme</b>	<b>Récurif</b>	<b>Avantages</b>	<b>Inconvénients</b>
Simple	Non	Simplicité	Imprécis
Direct-rétrograde en négligeant le terme non linéaire	Non	- Simplicité - Précis pour l'estimation de la translation - Intéressant pour les mouvements de forte translation - Stabilité	Non valable pour une forte divergence ou pour une forte rotation
Direct-rétrograde avec prise en compte du terme non linéaire	Oui	- Précis sur l'estimation du paramètre de rotation - Intéressant pour les mouvements à rotation dominante - Stabilité	Peu précis pour le cas d'un mouvement de forte translation
Direct-rétrograde et approche incrémentale	Oui	- Précis	Coût de calcul important
Direct-rétrograde à $n$ images ( $n > 2$ )	Non	- Permet un lissage temporel - Intéressant pour la segmentation spatio-temporelle	Amplitude du déplacement imprécis

Tableau 2.6 : Comparaison des performances des différentes méthodes pour un modèle de mouvement.





## Conclusion de la deuxième partie

Dans cette partie, nous avons présenté une méthode d'estimation du mouvement basée sur les gradients spatio-temporels. Notre étude concerne plus particulièrement le cadre du flot optique. Les mouvements sont modélisés comme une transformation globale 2D. Ces modèles sont utilisés pour minimiser une fonction d'erreur définie directement à partir des intensités lumineuses.

Après avoir introduit la notion de mouvements direct et rétrograde, un nouveau système d'équation de contrainte du mouvement a été établi. Nous venons ainsi de présenter une approche d'estimation du flot optique que nous avons nommée "approche basée sur un schéma direct-rétrograde". En s'appuyant sur l'hypothèse de conservation de la luminance et de l'uniformité de mouvement pour tous les pixels d'un voisinage donné, un critère à minimiser par une méthode d'optimisation a été défini. Le principe général est fondé sur la minimisation de ce critère dont une fonction d'erreur liée aux mouvements direct et rétrograde.

Le formalisme proposé s'avère être plus performant et apporte une solution aux principaux problèmes liés à l'équation de contrainte du mouvement. La prise en compte des informations dans les deux sens temporels permet de renforcer les contraintes de l'équation du flot optique. Divers estimateurs au sens des moindres carrés tels que la méthode LS, l'approche TLS, l'estimateur robuste et l'approche incrémentale ont été présentés.

Le premier chapitre a été consacré à l'estimation du mouvement dans un contexte translationnel. Nous avons d'abord présenté les approches LS et TLS pour un mouvement de translation. Pour ces deux approches, les résultats obtenus avec le formalisme direct-rétrograde sont généralement plus significatifs, en amplitude et en orientation, par rapport à ceux du formalisme basé sur un schéma simple. En revanche, l'estimation sur une image présentant une discontinuité spatiale n'est pas très efficace.

Nous avons alors proposé une estimation robuste afin d'améliorer les résultats. Une fonction de type M-estimateur a été utilisée. L'estimateur robuste basé formalisme direct-rétrograde permet une meilleure séparation entre les informations pertinentes et les points aberrants. Comparé aux résultats de l'estimateur robuste classique, l'estimateur proposé conduit à de meilleurs résultats pour des images présentant une discontinuité.

Quant au problème de déplacement trop important, il peut être résolu en utilisant l'approche incrémentale manipulant les mouvements direct et rétrograde. La particularité du formalisme direct-rétrograde appliqué à l'approche incrémentale est de pouvoir appliquer la correction dans les deux sens temporels : incrément direct et incrément rétrograde.

Les intérêts du formalisme basé sur un schéma direct-rétrograde sont présentés au tableau 2.7 et comparés aux intérêts du formalisme basé sur un schéma simple.

Le second chapitre est consacré à l'estimation du mouvement associé à un modèle paramétrique. Nous avons montré que la prise en compte des mouvements direct et rétrograde conduit à la génération d'une équation paramétrique non linéaire. Le formalisme proposé aboutit donc à un modèle non linéaire. Deux considérations possibles ont été retenues. La première consiste à ignorer les termes linéaires et se résume donc à une équation qui est similaire à celle du modèle de mouvement classique (équation basée sur un schéma direct). La seconde considération tient compte intégralement de l'équation du modèle y compris des termes non linéaires.

Le premier volet de l'étude est orienté vers le modèle linéarisé. Diverses techniques y ont été développées. Le reste de notre étude porte davantage sur le modèle non linéaire. Différentes variantes ont ainsi été proposées pour le traitement des termes non linéaires. Grâce aux

différentes approches de modélisation de la non linéarité, nous avons pu tester la validité du formalisme direct-rétrograde pour un modèle paramétrique de mouvement. Des extensions possibles telles que l'estimateur robuste des paramètres du mouvement et l'approche incrémentale d'un modèle de mouvement ont été aussi développées. Le tableau 2.8 présente une synthèse des résultats obtenus.

Méthodes	Approche simple	Approche basée sur un schéma direct-rétrograde
LS <sup>14</sup>	<ul style="list-style-type: none"> <li>- Résultat moyen</li> <li>- Résultat moyen pour un modèle de mouvement</li> <li>- Résultat correcte pour une forte rotation et/ou divergence</li> <li>- Non respect des discontinuités</li> </ul>	<ul style="list-style-type: none"> <li>- Avantageux pour un mouvement de translation</li> <li>- Avantageux pour un modèle paramétrique à faible rotation et/ou divergence</li> <li>- Non valable pour une forte rotation et/ou forte divergence</li> <li>- Traitement du problème de forte rotation et/ou forte divergence par modélisation du terme non linéaire</li> </ul>
TLS <sup>15</sup>	<ul style="list-style-type: none"> <li>- Bon résultat lorsqu'il y a convergence</li> <li>- Méthode d'optimisation instable</li> </ul>	<ul style="list-style-type: none"> <li>- Bon résultat lorsqu'il y a convergence</li> <li>- Non stable</li> </ul>
Robuste	<ul style="list-style-type: none"> <li>- Approche itérative</li> <li>- Avantageux pour une image présentant une discontinuité</li> <li>- Réglage du seuil difficile à manipuler</li> </ul>	<ul style="list-style-type: none"> <li>- Approche itérative</li> <li>- Avantageux pour une image présentant une discontinuité</li> <li>- Offre plus de possibilité sur le calcul de poids</li> <li>- Réglage du seuil difficile à manipuler</li> </ul>
Incrémentale	<ul style="list-style-type: none"> <li>- Approche itérative</li> <li>- Traitement du déplacement trop important</li> <li>- Estimation précise (de l'ordre de sub-pixel)</li> </ul>	<ul style="list-style-type: none"> <li>- Approche itérative</li> <li>- Traitement du déplacement trop important dans les deux sens (direct et rétrograde)</li> <li>- Estimation précise (de l'ordre de sub-pixel)</li> </ul>

Tableau 2.7 : Comparaison des approches des moindres carrés, basée sur un schéma simple et basée sur un schéma direct-rétrograde.

Approches	$t_x$	$t_y$	$k$	$\theta$
LS simple	Moyen	Moyen	Moyen	Moyen
FBLS <sup>16</sup> , modèle linéaire,	Bien	Bien	Moyen	Moyen
FBLS, modèle non linéaire	Moyen	Moyen	Bien	Bien
FB, incrémentale	Bien	Bien	Bien	Bien

Tableau 2.8 : Tableau comparatif des différentes méthodes d'estimation des paramètres du mouvement, méthodes basées sur le formalisme direct-rétrograde.

<sup>14</sup> Least squares

<sup>15</sup> Total least squares

<sup>16</sup> Forward backward least squares, méthode des moindres carrés basée sur un schéma direct-rétrograde

## **Troisième partie**

### **Segmentation spatio-temporelle associée à l'estimation du mouvement selon le formalisme direct-rétrograde**



# Chapitre 1

## Hypothèse markovienne en segmentation spatio-temporelle

### 1.1 Introduction

En général, la segmentation spatio-temporelle se développe à partir de deux trames consécutives. Dans ce contexte, plusieurs facteurs peuvent altérer l'interprétation de la scène. Nous en donnons quelques exemples : une mauvaise estimation du mouvement, le bruit et le manque d'information décisive. La segmentation spatio-temporelle peut être non seulement incomplète mais souvent erronée.

Un autre problème est relatif à l'instabilité de la procédure de segmentation suivant l'axe temporel. Admettons qu'aucun changement notable de la scène n'apparaisse entre deux trames consécutives. Ces trames figurent la même configuration d'objets (coordonnées spatio-temporelles, niveau de gris, ...). Cette constance est en contraste avec la définition de l'évolution temporelle des objets constituant la scène.

Pour rendre la segmentation plus robuste et plus cohérente dans le temps, certains auteurs ont proposé des méthodes utilisant plusieurs trames.

Dans la méthode de segmentation suivante, nous essayons de faire davantage intervenir l'estimation basée sur un schéma direct-rétrograde. La classification repose sur trois trames successives. La méthode proposée s'appuie sur la modélisation par champ aléatoire de Markov. Nous rappelons, dans un premier temps, le principe de la segmentation markovienne. Ensuite, nous procédons à une segmentation dans laquelle la modélisation précédente est couplée avec l'estimation basée sur un schéma direct-rétrograde. Une série de résultats est présentée. Les tests permettent de comparer la segmentation markovienne classique aux approches développées avec le schéma direct-rétrograde.

### 1.2 Description

Soit une image définie sur une maille dont les sommets constituent un ensemble  $S$  de sites  $p$  :

$$S = \{p = (i, j); 1 \leq i \leq m, 1 \leq j \leq n\} \quad (306)$$

A chaque site  $p$  (ou pixel) est associé une variable aléatoire  $y_p$  qui représente le niveau de gris. L'image de  $m \times n$  pixels est représentée par un vecteur de variable aléatoire  $Y_p$  tel que :

$$Y_p = \{y_p, y_p = (0, \dots, 255); p \in S\} \quad (307)$$

Un système de voisinage  $\Omega = \{\omega_p, p \in S\}$ , est formé de partie  $\omega_p$  de  $S$ , vérifiant les propriétés suivantes :

$$\Omega = \{\omega_p, p \in \omega_r \Leftrightarrow r \in \omega_p, p \notin \omega_p\} \quad (308)$$

L'ensemble  $\omega_p$  est appelé le voisinage de  $p$ ,  $r$  est dit voisin de  $p$  si  $r \in \omega_p$ .

Un champ aléatoire  $Y$  est dit champ markovien associé à ce système de voisinage si et seulement si :

$$P_{Y_p|Y_r}(y_p|y_r, r \neq p) = P_{Y_p|Y_r}(y_p|y_r, r \in \omega_p) \quad (309)$$

$P$  est une fonction de distribution de probabilité.

Une clique est un sous-ensemble de voisinage. Si  $c$  est une clique relative au système de voisinage  $\Omega$ , alors :

- $c$  peut être un singleton,
- $c$  peut être d'ordre  $n, n \geq 2$  et deux pixels quelconques de  $c$  sont voisins au sens de  $\Omega$ ,
- l'ordre d'une clique correspond au nombre de ses éléments.

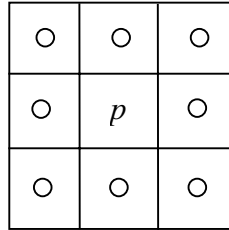


Figure 3.1 : Ensemble  $\omega_p$  de voisinages de  $p$  ; le système de voisinage noté  $\Omega$  est formé de partie  $\omega_p$ .

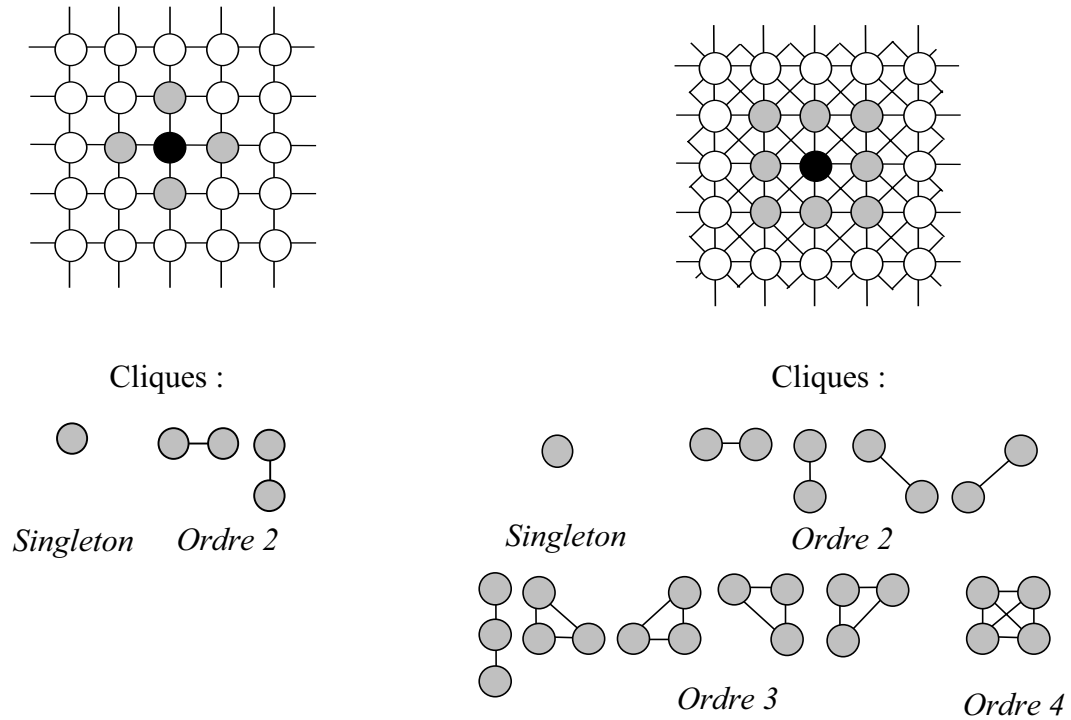


Figure 3.2 : Différentes configurations possibles de cliques.

Le théorème de Hammersley-Clifford démontre qu'un champ aléatoire  $Y$  est un champ markovien, associé au système de voisinage  $\Omega$ , si et seulement si :

$$\forall y, P_Y(y) = \frac{1}{Z} \exp[-U(y)] \tag{310}$$

$P_Y(y)$  est une mesure de Gibbs. La quantité  $U$  définit une fonction appelée énergie, exprimée par :

$$U(y) = \sum_{c \in C} V_c(y) \tag{311}$$

Les fonctions  $V_C$  sont appelées fonctions potentielles, et ne dépendent que des éléments de  $y$  qui appartiennent à la clique  $c$  ( $C$  est l'ensemble des cliques). Une clique peut être notée par :

- $c = \langle p \rangle$  : singleton,

- $c = \langle p, r \rangle$  : clique d'ordre 2,
- $c = \langle p, r_1, r_2 \rangle$  : clique d'ordre 3,

et l'ensemble  $C$  peut se présenter sous la forme :

$$C = \{ \langle p \rangle, \langle p, r \rangle, \langle p, r_1, r_2 \rangle, \dots \} \quad (312)$$

La quantité  $Z$  est une fonction de partition qui sert comme une constante de normalisation de la probabilité  $P_Y(y)$ .

$$Z = \sum_{y \in \Lambda} \exp[-U(y)] \quad (313)$$

A titre d'illustration, prenons une fonction  $V_C$  de la forme :

$$V_C(y) = [1 - \delta(y_p, y_r)] \quad (314)$$

La segmentation markovienne consiste à classifier un nuage de points en sous-nuage  $X$  (figure 3.3). En effet, à chaque pixel  $p$  (ou pixel) de  $X$  est associée une variable aléatoire  $x_p$  dont la valeur correspond à une des  $k$  classes.

$$X = \{x_p \in \{e_0, \dots, e_k\}\} \quad (315)$$

$\{e_0, \dots, e_k\}$  représente les  $k$  classes ou zones homogènes de l'image. L'image  $m \times n$  est donc représentée par un couple de vecteurs de variable aléatoire  $(X, Y)$ .

Si on suppose que  $X$  est un champ markovien, on a :

$$P_X(x) = \frac{1}{Z} \exp\left[-\sum_{c \in C} V_C(x)\right] \quad (316)$$

A titre d'illustration,  $k = 2$  classes et  $V_C(x) = \beta_{p,r} [1 - \delta(x_p, x_r)]$  avec  $\beta_{p,r} = \beta_1, \beta_2, \beta_3, \beta_4$  associé respectivement à la clique horizontale, verticale, diagonale droite et diagonale gauche.

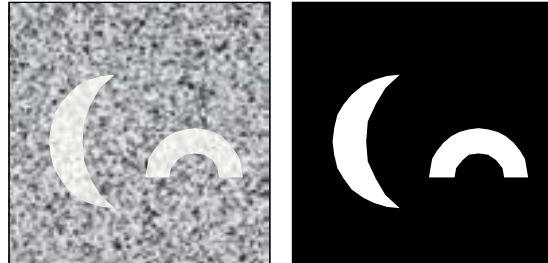


Figure 3.3 : Segmentation au sens des niveaux de gris ;  
fond : classe 0 ;  
la lune et l'arc : classe 1.

L'approche bayésienne-markovienne essaie de maximiser la probabilité conditionnelle (a posteriori) suivante :

$$P_{X|Y}(x, y) = \frac{P_{Y|X}(y|x)P_X(x)}{P_Y(y)} \quad (317)$$

Le dénominateur  $P_Y(y)$  est une distribution constante par rapport aux étiquettes car la réalisation du champ des observations (l'image) est connue ;  $P_X(x)$  est la distribution markovienne ;  $P_{Y|X}(y, x)$  est un modèle de dégradation de l'image. La maximisation de l'équation (317) peut se ramener à la maximisation de  $P_{Y|X}(y|x)P_X(x)$ . Supposons que  $P_{Y_p|X_p}(y_p, c_0)$  et  $P_{Y_p|X_p}(y_p, c_1)$  soient connus. Considérons le cas particulier suivant :

$$P_{Y_p|X_p}(y_p, c_0) = N(\mu_0, \sigma_0) \quad (318)$$



$$P_{y_p|x_p}(y_p, c_1) = N(\mu_1, \sigma_1) \quad (319)$$

Supposons également que :

$$P_{Y|X}(y, x) = \prod_{p \in S} P_{y_p|x_p}(y_p|x_p) \quad (320)$$

c'est à dire que le bruit sur chacune des classes soit blanc.

D'après les équations (316, 317, 320), un estimé  $x_{MAP}$  peut être déterminé par :

$$x_{MAP} = \arg \max_x \left\{ \frac{1}{Z} \exp \left[ - \left( - \ln \prod_{p \in S} P_{y_p|x_p}(y_p|x_p) + \sum_{c \in C} V_C(x_p) \right) \right] \right\} \quad (321)$$

Soit, après développement :

$$x_{MAP} = \arg \max_x \left\{ \frac{1}{Z} \exp \left[ - \left( \sum_{p \in S} V(x_p, y_p) + \sum_{c \in C} V_C(x_p) \right) \right] \right\} \quad (322)$$

avec :

$$V(x_p, y_p) = - \ln P_{y_p|x_p}(y_p|x_p) \quad (323)$$

Cette maximisation peut s'effectuer en minimisant l'exposant de l'exponentielle. On a :

$$x_{MAP} = \arg \min_x \left\{ \sum_{p \in S} V(x_p, y_p) + \sum_{c \in C} V_C(x_p) \right\} \quad (324)$$

Pour  $V_C(x_p) = \beta_c [1 - \delta(x_p, x_r)]$  et en tenant compte de l'équation  $P_{y_p|x_p}(y_p, c_p) = N(\mu_p, \sigma_p)$ , nous obtenons pour des cliques d'ordre 2 :

$$x_{MAP} = \arg \min_x \left\{ \sum_{p \in S} \left[ \ln \left( \sqrt{2\pi} \sigma_p + \frac{y_p - \mu_p}{2\sigma_p^2} \right) \right] + \sum_{c \in C} \beta_c [1 - \delta(x_p, x_r)] \right\} \quad (325)$$

Dans cette équation, la première somme correspond à un terme d'attache aux données. Ce terme modélise le lien entre les étiquettes (les classes) et les observations (les niveaux de gris). Quant à la deuxième somme, elle correspond à un terme d'énergie dont l'expression est fonction des potentiels. Ce terme caractérise et modélise les propriétés du champ d'étiquette que l'on recherche. Il est aussi appelé terme de régularisation. Trouver  $x_{MAP}$  est ensuite un problème classique d'optimisation.

Dans cette étude, nous nous intéressons à l'application du champ markovien dans la détection du mouvement. Nous faisons l'hypothèse que la valeur du niveau de gris du pixel considéré ne dépend que d'un voisinage réduit du pixel. Le processus markovien est complètement défini par :

- un état initial ;
- un ensemble de probabilités entre états (matrice de transition du processus).

Chaque état du processus correspond à une distribution particulière des niveaux de gris sur le voisinage. Avant d'introduire l'équation de cette distribution, nous allons rappeler quelques définitions concernant l'étiquetage et l'observation d'un pixel dans le cadre de la segmentation du flot optique [DUM99, GEL96, KER95].

A un instant  $t$  donné, chaque pixel  $p$  d'une image  $S$  est étiqueté par une valeur :

$$l_p = \begin{cases} 1, & \text{si } p \in \text{objet en mouvement} \\ 0, & \text{si } p \in \text{fond fixe} \end{cases} \quad (326)$$

D'après l'hypothèse de conservation de la luminance, l'information de mouvement d'un pixel  $p$  dépend du changement temporel de son intensité  $I(p, t)$ . Nous pouvons ainsi définir une observation  $o_p$  :

$$o_p = |I(p, t + dt) - I(p, t)| \quad (327)$$

Soit  $l = \{l_p, p \in S\}$  une réalisation particulière d'un champ d'étiquettes  $L$  sur l'image  $S$ . Soit  $o = \{o_p, p \in S\}$  une réalisation d'un champ d'observation  $O$ .

Pour un voisinage spatio-temporel donné (figure 3.4),  $(O, L)$  est modélisé comme un champ aléatoire de Markov. En effet, la configuration la plus probable du champ d'étiquettes  $L$ , pour une réalisation de champ d'observation  $O$  donnée, est obtenue en utilisant l'équivalence entre la distribution de Gibbs et le MRF.

$$\Pr[L = l / O = o] \sim e^{-U(o, l)} \quad (328)$$

Le champ d'étiquettes optimal  $l_{opt}$  peut être obtenu selon le critère de maximum a posteriori (MAP). Maximiser cette probabilité a posteriori revient à trouver le minimum de la fonction d'énergie  $U(o, l)$ .

$$l_{opt} = \arg \min_{c \in C} U(o, l) \quad (329)$$

La quantité  $C$  désigne l'ensemble de réalisations possibles de  $L$ . Dans notre cas selon la figure 3.4, l'ensemble  $C$  contient les dix éléments de combinaisons possibles de cliques  $c = \langle p, r \rangle$ . La fonction d'énergie  $U(o, l)$  est constituée par une somme de deux termes.

$$U(o, l) = U_p(l) + U_o(l) \quad (330)$$

Le premier terme permet d'introduire une régularisation spatio-temporelle. Il se décompose en une somme des potentiels locaux, notés  $V_c(l_p, l_r)$ .

$$U_p(l) = \sum_{c \in C} V_c(l_p, l_r) \quad (331)$$

Les potentiels locaux sont utilisés afin de former des contraintes d'homogénéité dans le modèle a priori (c'est à dire de donner des avantages aux réalisations où des voisins ont une même étiquette). Un potentiel élémentaire pour une clique  $c = \langle p, r \rangle$  est défini par :

$$V_c(l_p, l_r) = \begin{cases} -\beta < 0, & \text{si } l_p = l_r \\ +\beta > 0, & \text{si } l_p \neq l_r \end{cases} \quad (332)$$

Trois valeurs différentes de  $\beta$  peuvent être considérées :  $\beta_p$  pour les cliques spatiales,  $\beta_b$  pour la clique spatio-temporelle de l'image courante-précédente (backward) et  $\beta_f$  pour l'image courante-suivante (forward).

Le deuxième terme réalise l'adéquation aux données. Il correspond à une observation d'énergie qui relie les étiquettes et les observations. Son expression est définie de la manière suivante :

$$U_o(l) = \frac{1}{2\sigma^2} \sum_{p \in S} [o_p - \psi(l_p)]^2 \quad (333)$$

où  $[o_p - \psi(l_p)]$  est supposé avoir une distribution gaussienne centrée de variance  $\sigma^2$ .

La quantité  $\Psi$  quant à elle permet de modéliser le comportement de l'observation,

$$\psi(l_p) = \begin{cases} 0, & \text{si } l_p = 0 \\ \mu > 0, & \text{si } l_p = 1 \end{cases} \quad (334)$$

Si un pixel appartient au fond de l'image, il n'y aura pas de changement temporel entre deux trames consécutives. Par conséquent, l'observation sera toujours nulle. Si l'observation appartient à un objet mobile, l'observation est supposée positive en prenant la valeur  $\mu$ , qui correspond à la valeur moyenne des observations non nulles.

La minimisation de la fonction d'énergie  $U(o, l)$  peut s'effectuer d'une manière itérative. L'algorithme de minimisation est décrit à la figure 3.5.

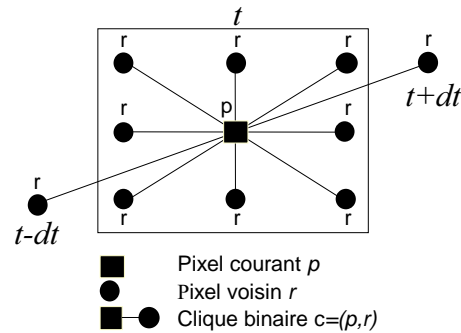


Figure 3.4 : Voisinage spatio-temporel et cliques binaires.

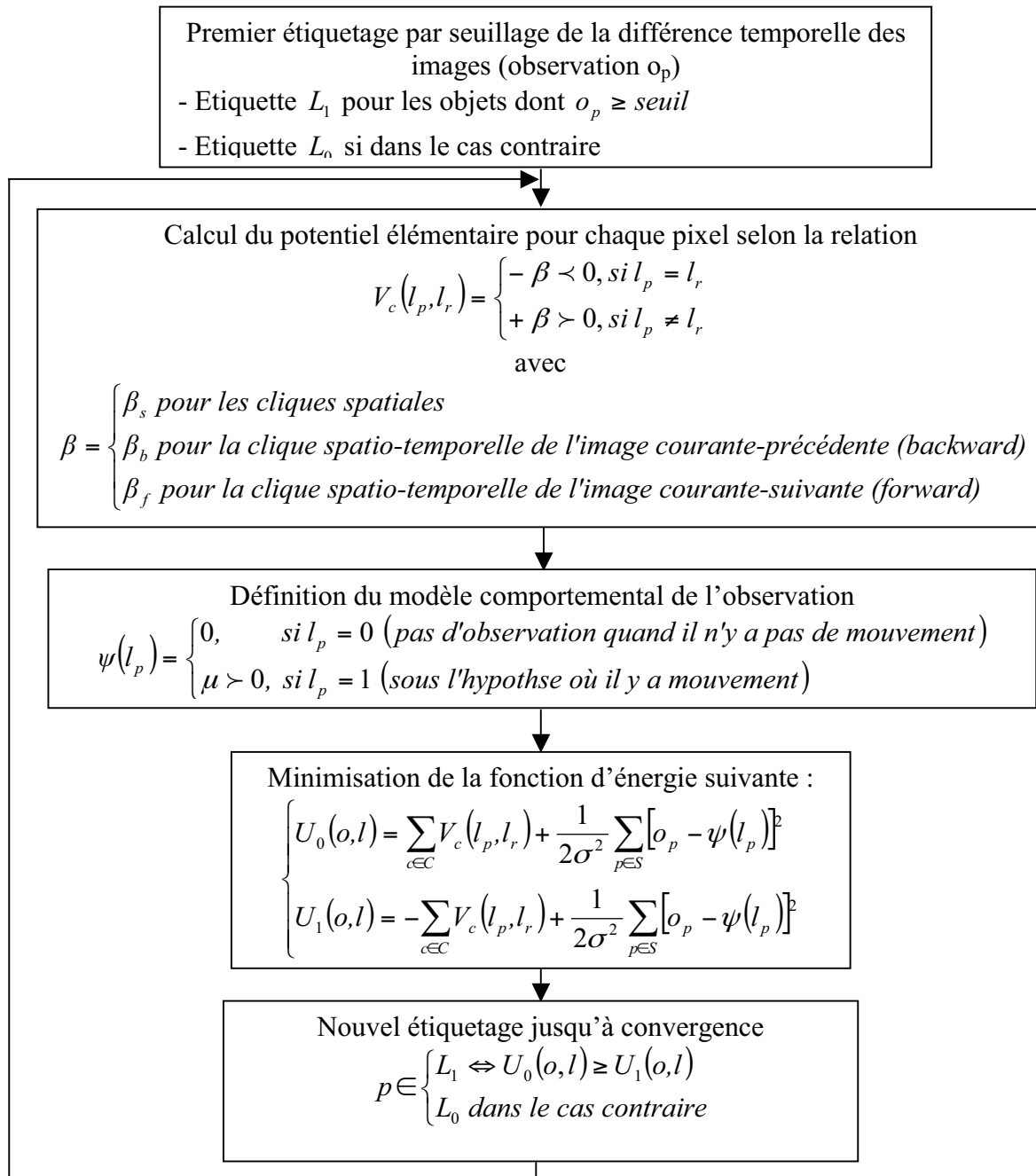


Figure 3.5 : Approche markovienne appliquée à la segmentation spatio-temporelle.

### 1.3 Résultats expérimentaux

La figure 3.6 montre le résultat obtenu en utilisant un processus markovien. On peut observer que la segmentation est assez bruitée, et que des zones d'occultation apparaissent. Ceci est dû au fait que d'abord, l'initialisation est effectuée à défaut sur l'image elle-même et que, de plus, le terme d'adéquation aux données est mal adapté pour la prise de décision. L'algorithme utilisé ne parvient donc pas à segmenter correctement la séquence. Un bon réglage des paramètres de segmentation permettrait d'éliminer les zones aberrantes mais cela au détriment de la forme des objets réels. Dans le paragraphe suivant, nous proposons une amélioration en faisant un lien entre l'estimation que nous avons proposée et les techniques de segmentation à base de champs de Markov.

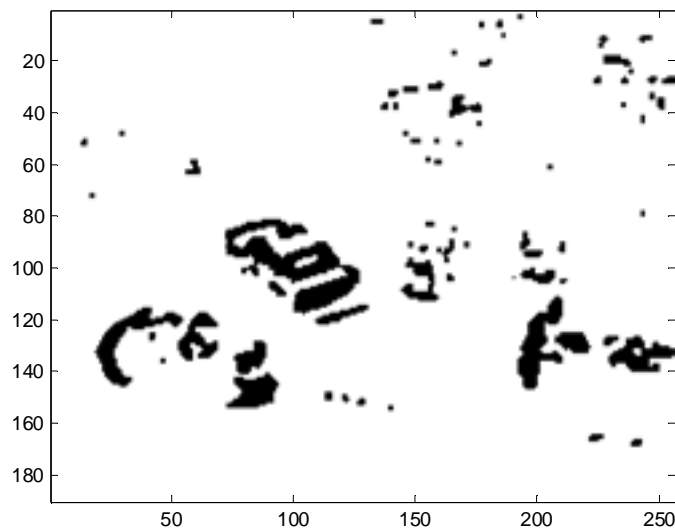


Figure 3.6 : Segmentation spatio-temporelle de la séquence "taxi", approche markovienne ;  $\beta_p = 30$ ,  $\beta_b = 5$ ,  $\beta_f = 80$ ,  $seuil = 4$ ,  $\mu = 15$ ,  $\sigma^2 = 20$  ; résultat au bout de 5 itérations.

### 1.4 MRF et estimation du mouvement basée sur un schéma direct-rétrograde

Le paragraphe précédent a mis en évidence les limitations du modèle markovien pour la segmentation spatio-temporelle. Nous proposons une amélioration en faisant un lien entre l'estimation du mouvement et la segmentation markovienne. Ce lien nous permettra de voir quels sont les paramètres, dans le contexte markovien, sur lesquels on peut influencer, et de donner un sens aux paramètres du mouvement dans le cadre de la segmentation. Le principe de base est de mettre en place une phase d'estimation supplémentaire et une règle d'affectation plus adéquate dans la modélisation markovienne.

#### 1.4.1 Description

Dans cette estimation, nous nous attachons à proposer une méthode qui exploite le résidu associé à l'équation de contrainte du mouvement comme critère d'attache aux données (le terme  $U_o(I)$ ). Nous désirons introduire dans l'approche markovienne, la quantité d'estimation du flot optique fondée sur le formalisme direct-rétrograde. Le principe est décrit comme suit :

\* Estimation du mouvement par une approche basée schéma direct-rétrograde ;

\* Étape de pré-étiquetage pour séparer les pixels qui sont susceptibles d'être en déplacement. Cette initialisation peut se faire, soit par un seuillage de la différence temporelle des trames, soit par un seuillage du champ de déplacement.

\* Le terme d'hypothèse gaussienne  $U_o(l)$  est remplacé par le résidu de l'équation du mouvement. Afin de proposer une fonction d'énergie à minimiser, nous introduisons deux types de résidus représentant les deux hypothèses respectivement "pas mouvement" et "mouvement". Nous avons pour une image notée  $S$  :

$$\forall p(x, y) \in S, Res_0(p) = \begin{cases} \frac{1}{2\sigma^2} \left[ \frac{\partial I(p)}{\partial t} \right]^2, & \text{si } \sqrt{dx^2 + dy^2} \geq \text{seuil} \\ 0, & \text{dans le cas contraire} \end{cases} \quad (335)$$

$$\forall p(x, y) \in S, Res_1(p) = \begin{cases} \frac{1}{2\sigma^2} \left[ \frac{\partial I(p)}{\partial x} dx + \frac{\partial I(p)}{\partial y} dy + \frac{\partial I(p)}{\partial t} \right]^2, & \text{si } \sqrt{dx^2 + dy^2} \geq \text{seuil} \\ 0, & \text{dans le cas contraire} \end{cases} \quad (336)$$

L'élément  $I$  représente le niveau de gris de l'image,  $\sigma^2$  la variance utilisée dans la fonction de distribution gaussienne, expression standard de  $U_o(l)$ .

De ces deux relations, on peut dire que pour les pixels qui appartiennent à une zone en déplacement, nous avons :  $Res_0 \geq Res_1$  et inversement dans le cas contraire. Nous utilisons alors ces deux résidus dans la fonction d'énergie  $U(o, l)$ .

Le détail de l'algorithme est montré par la figure 3.7 [IZQ00]. Dans cet algorithme, l'étiquetage est réalisé de manière récursive dans la boucle après comparaison entre les fonctions d'énergie relative aux deux hypothèses. La variation du comportement au cours des itérations est réalisée par le cumul des étiquettes du voisinage.

## 1.4.2 Résultats expérimentaux

Dans cette expérimentation, l'algorithme est initialisé en séparant les objets selon que leurs déplacements soient supérieurs ou inférieurs à un seuil. Le résultat est présenté à la figure 3.8. A première vue, ce résultat montre la participation de l'estimation du mouvement à la segmentation. L'algorithme utilisé parvient à segmenter correctement l'image. L'introduction de l'estimation du mouvement basée sur un schéma direct-rétrograde dans l'algorithme MRF permet d'éliminer les régions aberrantes. Cette figure montre tout particulièrement l'intérêt de l'estimation basée sur un schéma direct-rétrograde pour la segmentation spatio-temporelle. On peut noter également que via ce parallèle effectué avec le modèle markovien, on peut avoir une explication plus rigoureuse de l'usage des informations mouvement.

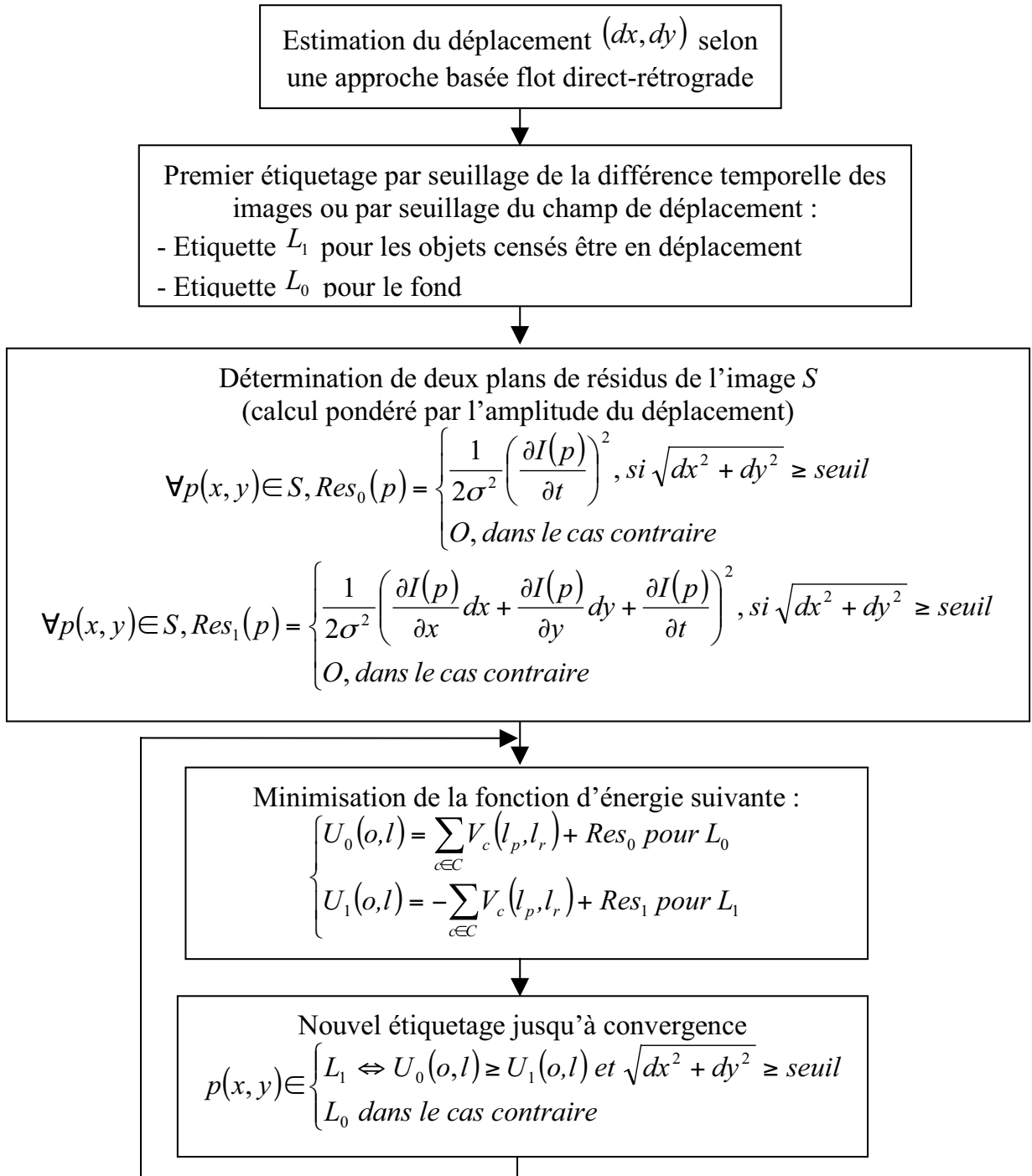


Figure 3.7 : Segmentation selon le MRF associée à l'estimation du mouvement (formalisme basé sur un schéma direct-rétrograde).

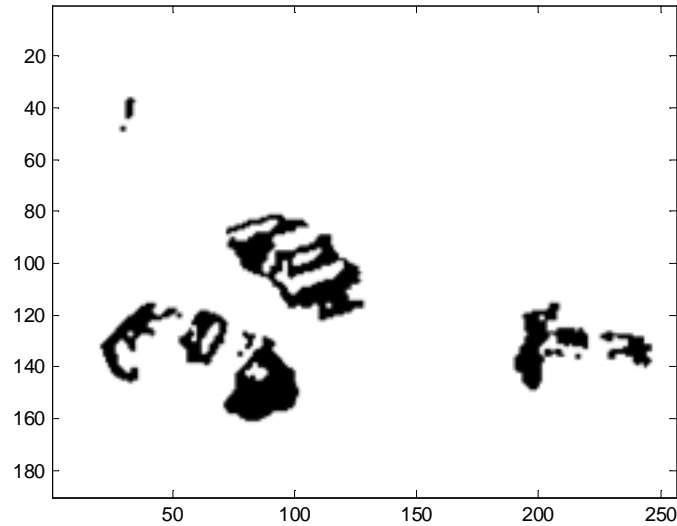


Figure 3.8 : Segmentation markovienne associée à l'estimation basée schéma direct-rétrograde;  $\beta_p = 30$ ,  $\beta_b = 5$ ,  $\beta_f = 80$ ,  $seuil = 0.4$ ,  $\sigma^2 = 20$ ; résultat au bout de 5 itérations.

## 1.5 Segmentation et estimation simultanées à l'aide du MRF et du schéma direct-rétrograde

### 1.5.1 Description

Dans ce paragraphe, nous allons essayer d'établir un algorithme qui permet à la fois de segmenter l'image et de générer un champ de déplacement. Les étiquettes, qui sont locales, et les paramètres du mouvement, qui sont globaux, sont estimés alternativement et itérativement. Une étiquette supplémentaire est considérée afin de déterminer les nouvelles régions qui apparaissent. La convergence est réalisée quand le nombre de sites dont l'étiquette a été changée chute au-dessous d'un seuil. Le déroulement de l'opération est détaillé comme suit :

\*Premier étiquetage en passant par toutes les étapes décrites dans le paragraphe 1.4.

\*Recherche du nombre d'objets labellisés. L'objet ayant comme label 0 correspond au fond.

\*Estimation du mouvement de chaque objet labellisé mis à part le fond. Une méthode des moindres carrés, associée à un formalisme direct-rétrograde peut être utilisée. Le support de calcul pour chaque objet est pris directement à la taille de chacun, résultat de la pré-segmentation à la première étape.

\*Ré-étiquetage de l'image  $S$  en partant de l'image segmentée initialement à l'étape 1, et en utilisant les résidus suivants :

- Plan de résidus sous l'hypothèse "pas mouvement" :

$$\forall p(x, y) \in S, \text{Re } s_0(p) = \frac{1}{2\sigma^2} \left[ \frac{\partial I(p)}{\partial t} \right]^2 \& \sqrt{dx^2 + dy^2} \geq \text{seuil} \quad (337)$$

- Plan de résidus sous l'hypothèse "mouvement" :

$$\forall p(x, y) \in S, \text{Re } s_1(p) = \frac{1}{2\sigma^2} \left[ \frac{\partial I(p)}{\partial x} dx + \frac{\partial I(p)}{\partial y} dy + \frac{\partial I(p)}{\partial t} \right]^2 \& \sqrt{dx^2 + dy^2} \geq \text{seuil} \quad (338)$$

\*La fonction à minimiser pour déterminer les étiquettes actuelles est de la forme :

$$U(o, l) = \sum_{c \in C} V_c(l_p, l_r) + \begin{cases} \text{Res}_0 & \text{pour les étiquettes à 0 de l'étape 1} \\ \text{Res}_1 & \text{pour les étiquettes à 1 de l'étape 1} \end{cases} \quad (339)$$

La minimisation de cette fonction d'énergie peut se faire d'une manière itérative depuis la deuxième étape.

### Critère de validation

L'amélioration apportée par une nouvelle estimation/segmentation, à une itération donnée se situe à trois niveaux :

- Réduction immédiate de l'erreur résiduelle.
- Une meilleure division-fusion de régions doit avoir lieu lors du rebouclage sur la deuxième étape.
- Une meilleure estimation du mouvement peut être obtenue lorsque nous repassons à la troisième étape.

### Critère d'arrêt

Le processus itératif s'arrêtera lorsqu'il n'y aura plus de changement possible. Dans notre cas, le processus s'arrête quand le niveau de classification (nombre de classes) reste stable. On peut renforcer ce critère en ajoutant une autre variante qui consiste à définir le nombre d'itérations maximal.

### 1.5.2 Résultats expérimentaux

Contrairement à l'algorithme présenté dans le paragraphe 1.4, l'estimation du mouvement est utilisée, non pas seulement pour initialiser l'approche markovienne, mais surtout en interaction avec cette approche. En d'autres termes, nous utilisons l'estimation du mouvement pour mieux segmenter l'image et inversement, les régions ainsi segmentées seront à leur tour réutilisées pour mieux ré-estimer le déplacement. En effet, dans la méthode des moindres carrés, on utilise les régions issues de la segmentation pour optimiser le choix de la fenêtre de calcul. En outre, le calcul des poids peut évoluer suivant la forme et la structure des dernières régions segmentées. L'algorithme s'effectue d'une manière itérative. Comme on peut le remarquer à la figure 3.9, la technique proposée fournit de meilleurs résultats. Les quatre objets sont bien mis en évidence sans présence des régions parasites.

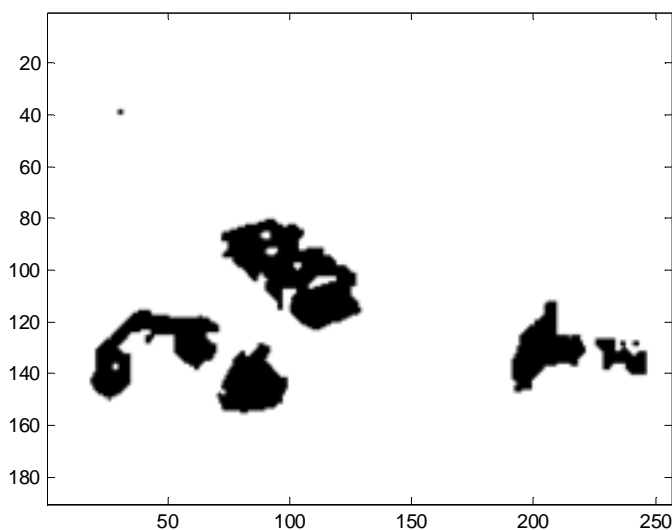


Figure 3.9 : Segmentation et estimation simultanées suivant l'approche markovienne couplée avec l'estimation du mouvement basée sur un schéma direct-rétrograde ;  $\beta_p = 30$  ,  $\beta_b = 5$  ,  $\beta_f = 80$  , *seuil* = 0.4 ,  $\sigma^2 = 20$  ; résultat au bout de 5 itérations.



### 1.5.3 Apport du formalisme direct-rétrograde à la segmentation, comparaison avec le formalisme simple

Jusqu'ici, nous avons mis en évidence l'intérêt de l'estimation du mouvement dans le cadre de la segmentation spatio-temporelle. Ce paragraphe a pour but de comparer les résultats obtenus avec l'estimation basée sur un schéma simple et l'estimation basée sur un schéma direct-rétrograde. La figure 3.10 montre la segmentation obtenue en utilisant une estimation basée schéma simple. Les expérimentations associées au formalisme simple et celui direct-rétrograde (figures 3.10 et 3.9) ont été effectuées avec les mêmes paramètres. Nous pouvons constater la différence entre ces deux résultats. Dans le cas de la segmentation avec estimation basée sur un schéma simple (figure 3.10), d'autres régions aberrantes sont apparues alors que le piéton est effacé.

Une autre analyse qui consiste à évaluer la vitesse de convergence de ces deux approches est illustrée par la figure 3.11. Ces courbes représentent le niveau de classification (nombre de classes) pour chaque itération. Au départ, l'image est divisée en un ensemble de régions en utilisant une méthode de seuillage du déplacement. L'algorithme converge rapidement car les résultats sont stables au bout de quelques itérations seulement.

On constate qu'au départ, pour une même valeur du seuil, l'approche basée sur un schéma direct-rétrograde génère des régions initiales plus connexes. A la fin du processus, le niveau de classification donné par cette approche est plus adapté à la séquence.

Un critère d'arrêt peut être envisagé soit en définissant un nombre d'itérations maximal, soit en effectuant un test s'il n'y a plus de changement possible au niveau du nombre de classes.

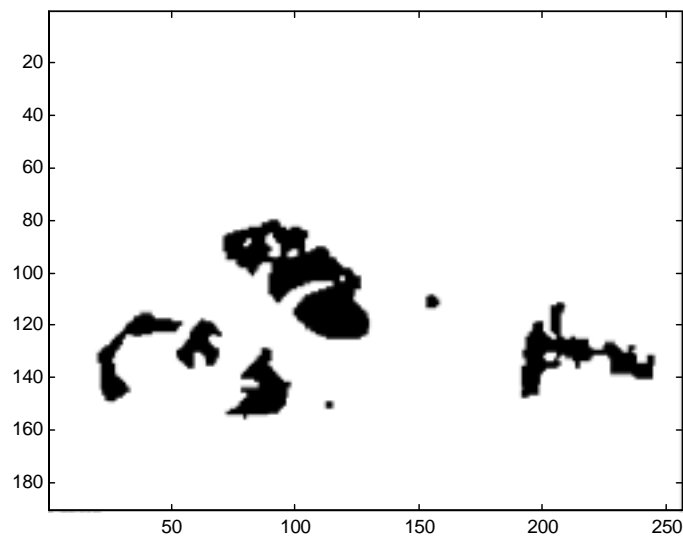


Figure 3.10 : Segmentation et estimation simultanées suivant l'approche MRF couplée avec l'estimation du mouvement basée schéma simple ;  $\beta_p = 30$  ;  $\beta_b = 5$  ,  $\beta_f = 80$ ,  $seuil = 0.4$ ,  $\sigma^2 = 20$  ; résultat au bout de 5 itérations.

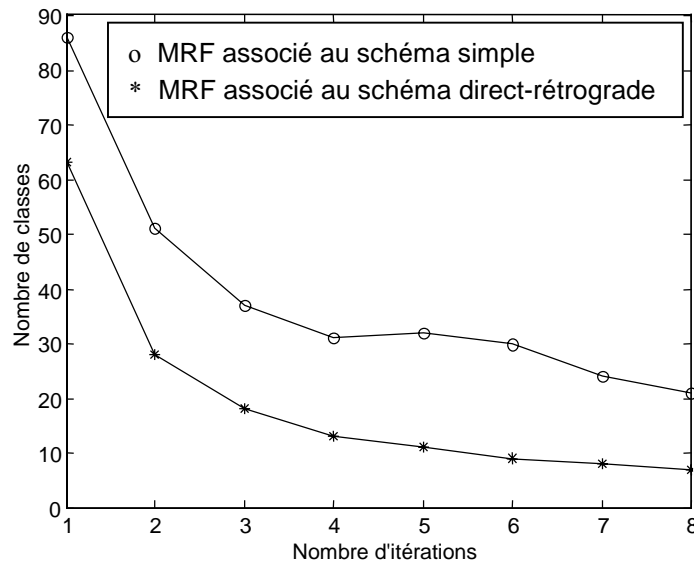


Figure 3.11 : Comparaison de deux approches en terme de convergence (approche basée sur un schéma simple et celle basée sur un schéma direct-rétrograde) ; nombre de classes en fonction du nombre d'itérations.

## 1.6 Algorithme MRF associé au formalisme direct-rétrograde à $n$ images, application à la segmentation spatio-temporelle

### 1.6.1 Description

Afin d'évaluer les performances du formalisme direct-rétrograde à  $n$  images dans le cadre de la segmentation spatio-temporelle, nous l'associons à l'algorithme MRF. La méthode est décrite par la figure 3.12

### 1.6.2 Résultats expérimentaux

Le présent test porte sur des résultats issus de l'association du MRF et l'approche basée sur un schéma direct-rétrograde à 5 images. La figure 3.13 illustre les différents gradients utilisés dans l'estimation. Les éléments  $I_x$ ,  $I_y$  et  $I_t$  désignent respectivement les gradients spatiaux suivant  $x$ ,  $y$  et le gradient temporel.

Pour l'estimation du mouvement, nous définissons le critère à minimiser à partir des valeurs des gradients présentées à la figure 3.13. Le voisinage spatio-temporel et les cliques binaires sont définis à partir de 3 images centrales ( $I_2$ ,  $I_3$  et  $I_4$ ). L'image segmentée est présentée à la figure 3.14. On peut observer que cette technique permet d'avoir une bonne segmentation des objets, notamment au niveau du piéton.

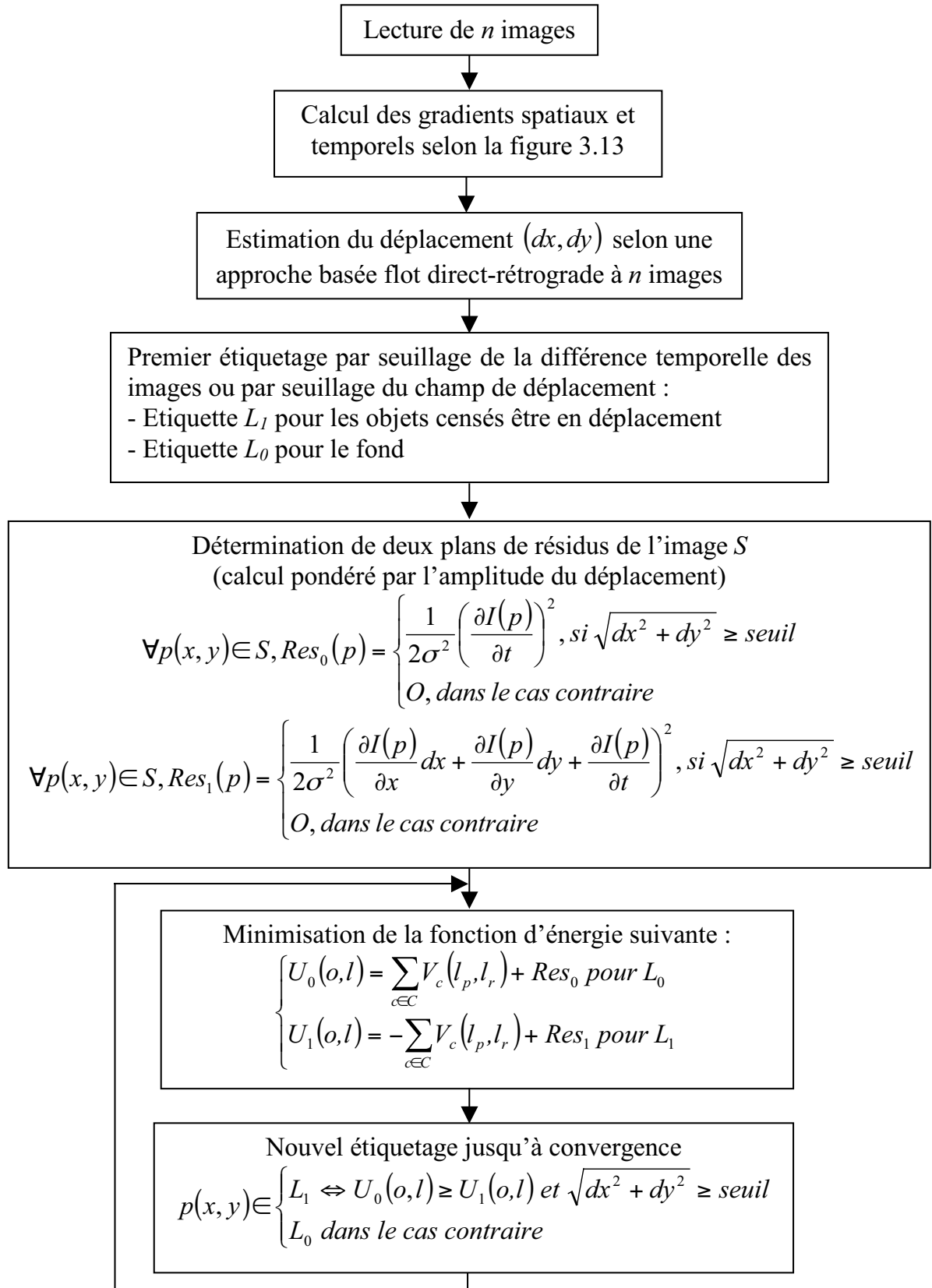


Figure 3.12 : Segmentation spatio-temporelle basée sur l'algorithme MRF et le formalisme direct-rétrograde à  $n$  images.

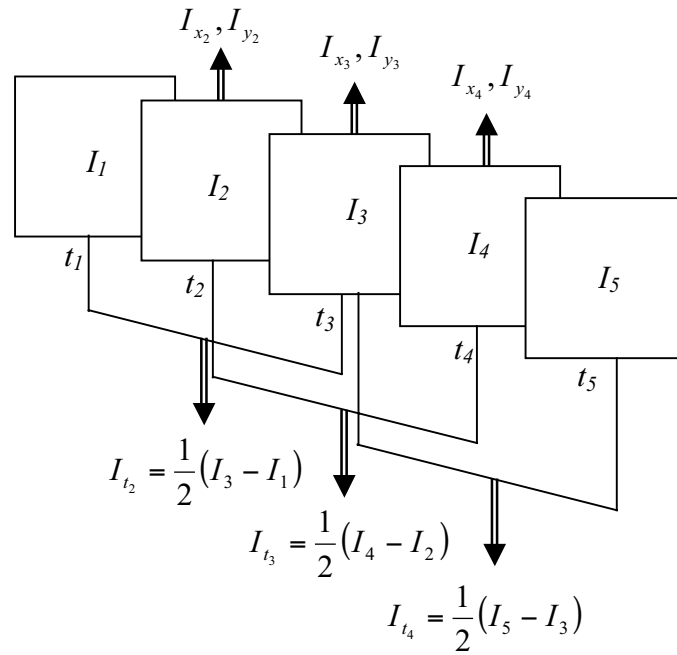
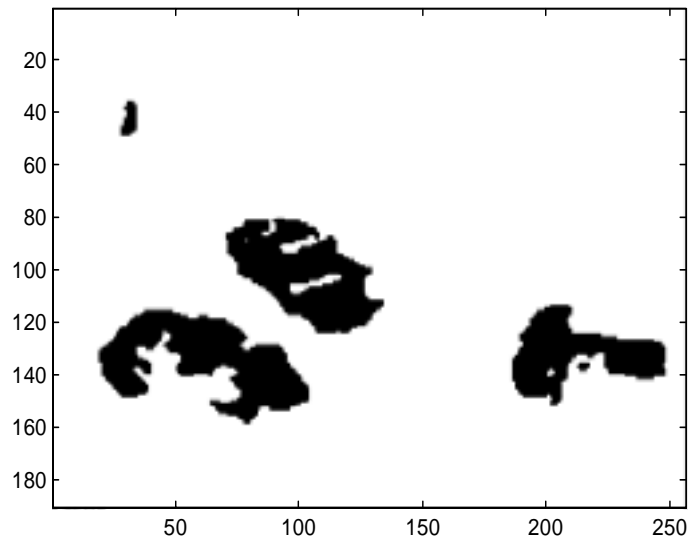


Figure 3.13 : Calcul des gradients à partir de 5 images.

Figure 3.14 : Segmentation suivant le MRF associé à l'estimation du mouvement basée sur un schéma direct-rétrograde à  $n$  images avec  $n = 5$ ,  $\beta_p = 30$ ,  $\beta_b = 5$ ,  $\beta_f = 80$ ,  $seuil = 0.4$ ,  $\sigma^2 = 20$  ; résultat au bout de 5 itérations.



## Chapitre 2

# Algorithme K-moyens associé au formalisme direct-rétrograde

### 2.1 Description

Dans ce paragraphe, nous proposons une méthode de classification qui s'inspire de l'algorithme K-moyens. Le principe général de cet algorithme est mentionné en annexe B. Dans la présente étude, nous associons l'algorithme K-moyens au formalisme direct-rétrograde. Considérons une image formée d'un nuage  $S$  de  $m \times n$  points. Ce nuage est défini par :

$$S = \{I(p), \forall p \in [1, m \times n]\} \quad (340)$$

$I(p)$  désigne le niveau de gris en un point  $p$ . L'objectif est de partitionner l'image nommée  $S$  en  $k$  sous-nuages  $S_k$ . Pour ce faire, le principe suivant est retenu :

Étant précisé dans le principe général de cet algorithme (cf. annexe B), le problème posé se réduit à la détermination des représentants d'une classe  $k$ . Nous considérons comme représentant le vecteur de déplacement. La phase d'affectation d'un pixel à l'un des sous-nuages est la comparaison de la métrique en fonction du représentant de la classe. Dans notre application, l'erreur résiduelle issue de la contrainte de mouvement est considérée comme métrique. Il est alors nécessaire d'introduire les relations suivantes :

- Admettons que l'amplitude du déplacement  $d$  en tout point  $p$  de l'image se trouve entre deux valeurs  $d_{min}$  et  $d_{max}$  et que son orientation  $\alpha$  se situe entre deux angles  $\alpha_{min}$  et  $\alpha_{max}$ . Nous admettons également que les amplitudes et les orientations de déplacement s'échelonnent sur  $k$  valeurs différentes. Un champ de déplacement, noté  $D$ , peut ainsi être caractérisé par :

$$D = \left\{ \vec{d}, \forall p \in [1, m \times n], d_1 \leq \|\vec{d}\| \leq d_k \text{ et } \alpha_1 \leq \alpha \leq \alpha_k \right\} \quad (341)$$

avec  $d_{min} \leq d_1$ ,  $d_k \leq d_{max}$ ,  $\alpha_{min} \leq \alpha_1$  et  $\alpha_k \leq \alpha_{max}$ .

Pour un pixel  $p$  donné, nous définissons la valeur de l'amplitude du déplacement par :  $\|\vec{d}\| = \sqrt{dx^2 + dy^2}$  et son orientation par :  $\alpha = \arctan(dx/dy)$ . Dans cette étude, au lieu de raisonner en terme d'amplitude et d'orientation du vecteur déplacement, nous utilisons les composantes  $dx$  et  $dy$ . Cela permet d'exprimer le résidu en fonction de ces composantes ; le résidu sera par la suite utilisé pour la métrique.

- Soit le résidu en point  $p(x, y)$ , associé à un déplacement quelconque  $\vec{d} = (dx, dy)^T$  :

$$Res(\vec{d}) = \frac{\partial I(p)}{\partial x} dx + \frac{\partial I(p)}{\partial y} dy + \frac{\partial I(p)}{\partial t} \quad (342)$$

Comme nous disposons de  $k$  valeurs de déplacement différentes, nous pouvons donc calculer  $k$  plans différents pour la matrice de résidus.

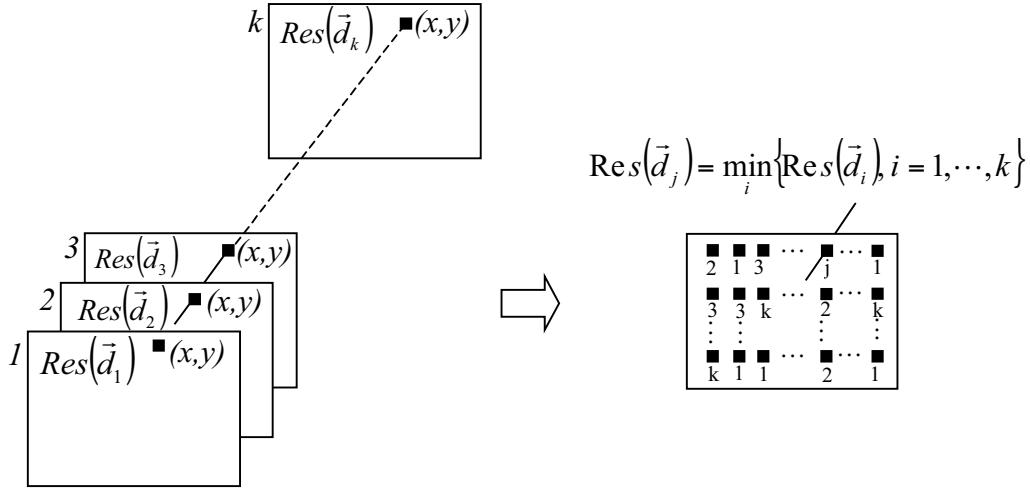


Figure 3.15 : Les  $k$  plans de résidus ; le  $1^{er}$  plan est constitué par les résidus associés aux  $m \times n$  points et calculés avec le déplacement  $\vec{d}_1 \dots$ ; de même, le  $k^{ème}$  plan représente les résidus aux  $m \times n$  points et calculés avec le déplacement  $\vec{d}_k$ .

Soit  $\vec{d}_i, i \in [1 : k]$ , les différents vecteurs déplacement. La règle de décision pour qu'un point  $p(x, y)$  soit affecté au sous-nuage  $S_j, j \in [1 : k]$ , est que :

$$p(x, y) \in S_j \Leftrightarrow \left\{ Res(p, \vec{d}_j) = \min_{d_i} \left\{ Res(p, \vec{d}_i), \forall i = 1, \dots, k \right\} \right\} \quad (343)$$

L'algorithme K-moyens correspond à une minimisation itérative du critère  $J$ , tel que :

$$J(\vec{d}_1, \dots, \vec{d}_k) = \sum_{p_i \in S_j} \left[ Res(p_i, \vec{d}_j) \right]^2, 1 \leq j \leq k \quad (344)$$

Les différentes étapes de l'algorithme sont les suivantes :

\* Choix du nombre de classes  $k$  ;

\* Initialisation des valeurs du déplacement ;

$$k \text{ classes : } \vec{d}_j^{(0)}, j \in [1 : k].$$

\* Calcul des  $k$  matrices de résidus associées aux  $k$  déplacements ;

$$\begin{array}{l} \text{pour } j = 1, \dots, k \\ \left[ Res(\vec{d}_j) = \begin{bmatrix} \frac{\partial I(1,1,t)}{\partial x} dx_j + \frac{\partial I(1,1,t)}{\partial y} dy_j + \frac{\partial I(1,1,t)}{\partial t} & \dots & \frac{\partial I(1,n,t)}{\partial x} dx_j + \frac{\partial I(1,n,t)}{\partial y} dy_j + \frac{\partial I(1,n,t)}{\partial t} \\ \vdots & \ddots & \vdots \\ \frac{\partial I(m,1,t)}{\partial x} dx_j + \frac{\partial I(m,1,t)}{\partial y} dy_j + \frac{\partial I(m,1,t)}{\partial t} & \dots & \frac{\partial I(m,n,t)}{\partial x} dx_j + \frac{\partial I(m,n,t)}{\partial y} dy_j + \frac{\partial I(m,n,t)}{\partial t} \end{bmatrix} \right. \\ \text{fin } j. \end{array}$$

\* Affectation des points aux sous-nuages  $S_j, j \in [1 : k]$  ( $k$  classes) ;

$$\begin{array}{l} \text{pour } x = 1, \dots, m \\ \left[ \text{pour } y = 1, \dots, n \right. \\ \left. \left[ p(x, y) \in S_j \Leftrightarrow \left\{ Res(p, \vec{d}_j) = \min_{d_i} \left\{ Res(p, \vec{d}_i), \forall i = 1, \dots, k \right\} \right\} \right. \right. \\ \left. \text{fin } y ; \right. \\ \left. \text{fin } x. \right. \end{array}$$

- \* Ré-estimation des déplacements en utilisant l'approche directe-rétrograde pondérée ;
- Calcul des poids en partant du plan  $\min\{Res\}$  : résidus associés à tous les points des sous-nuages  $S_j : j \in [1:k]$  ;
- Estimation du nouveau champ de déplacement en utilisant la méthode des moindres carrés robustes, basée sur un schéma direct-rétrograde. Le critère à minimiser est de la forme :

$$J(\vec{d}_1, \dots, \vec{d}_k) = \sum_{i \in S_j} [Res(p_i, \vec{d}_j)]^2, \quad 1 \leq j \leq k \quad (345)$$

A chaque itération, nous obtenons  $k$  nouvelles valeurs de déplacement :  $\vec{d}_j, j \in [1:k]$  ;

\* Retour à la troisième étape jusqu'à convergence.

Le critère d'arrêt peut être défini en fixant un nombre d'itérations maximal.

## 2.2 Résultats expérimentaux

L'algorithme ci-mentionné est testé sur la séquence "taxi" (figure 3.16). Le champ de déplacement initial a été choisi d'une manière aléatoire. Les phases de segmentation sont décrites comme suit :

- Choix du nombre de classes ; on veut par exemple obtenir cinq niveaux de classification différents, soit  $k = 5$  ;
- Génération d'un champ de déplacement aléatoire ; nous avons généré cinq amplitudes de déplacement différentes.
- Calcul des résidus pour chacune des valeurs des déplacements.
- Initialisation des sous-nuages (cinq) selon la règle de décision décrite à la quatrième étape.
- Application d'un algorithme de gestion de pondération à partir des sous-nuages.
- Stockage des poids en cumulant leurs valeurs à chaque itération.
- Ré-estimation des déplacements en appliquant la méthode des moindres carrés robuste, basée sur un schéma direct-rétrograde ; si les poids d'un sous-nuage sont tous nuls, un déplacement aléatoire sera ré-affecté à ce sous-nuage.
- Retour à la détermination des plans de résidus et l'opération se poursuit d'une manière itérative.

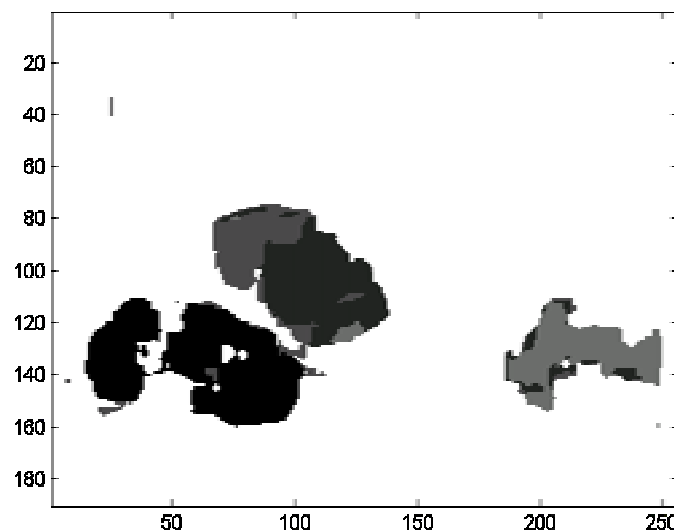


Figure 3.16 : Algorithme K-moyens associé à l'estimation du mouvement basée sur un schéma direct-rétrograde ; les régions initiales sont initialisées à l'aide d'un ensemble de déplacements aléatoires.



Le résultat complémentaire suivant porte sur une carte de segmentation issue d'une modification apportée dans l'algorithme précédent. Au lieu d'initialiser le champ de déplacement par un générateur aléatoire, nous partons des champs issus des cinq cartes d'estimation différentes. En fait, les cinq cartes d'estimation sont obtenues en utilisant cinq supports de calcul différents dans l'algorithme d'estimation. La version modifiée est décrite comme suit :

- Choix du nombre de classes, soit  $k = 5$ .
- Estimation du mouvement : l'objectif est ici d'avoir cinq ( $k = 5$ ) champs de déplacement d'amplitudes différentes. A titre d'illustration, pour l'image  $190 \times 256$  pixels, nous appliquons cinq fois un algorithme d'estimation du mouvement avec un support de calcul de taille différente pour chaque algorithme. Ceci permet d'avoir 5 matrices de déplacements. Nous avons ici utilisé une approche des moindres carrés basée sur un schéma direct-rétrograde.
- Calcul des résidus (cinq plans de résidus) ;
- Initialisation des sous-nuages selon la règle de décision décrite à la quatrième étape ;
- Application d'un algorithme de gestion de pondération à partir des sous-nuages ;
- Stockage des poids en cumulant leur valeur à chaque itération ;
- Ré-estimation des déplacements en appliquant la méthode des moindres carrés robuste basée sur un schéma direct-rétrograde; si les poids d'un sous-nuage sont tous nuls, ce sous-nuage est censé être éliminé et son déplacement sera ainsi affecté à la valeur nulle ;
- Retour à la détermination des plans de résidus et ainsi de suite d'une manière itérative.

La figure 3.17 représente les régions pour lesquelles, les valeurs des poids sont cumulées au bout d'un certain nombre d'itérations, et sont inférieures à un seuil.

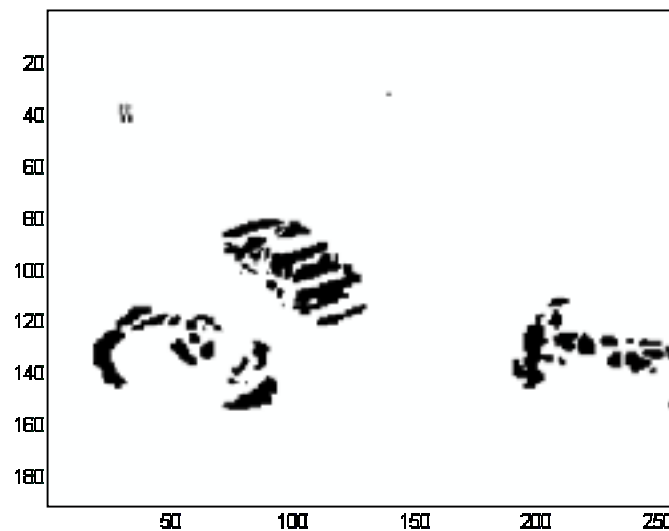


Figure 3.17 : Image segmentée à l'aide de l'algorithme de K-moyens couplé avec une estimation robuste basée sur un schéma direct-rétrograde ; les régions initiales sont initialisées à l'aide de l'estimation à plusieurs niveaux de déplacement et cela, par utilisation de différents supports spatiaux.

Nous pouvons remarquer que l'algorithme modifié parvient à segmenter correctement les quatre objets. Ce résultat montre les performances de l'algorithme proposé. Quand cet algorithme est convenablement optimisé, avec un bon choix des différents paramètres (nombre de classe, critère de convergence) et avec une phase d'initialisation correcte (bonne estimation du mouvement), il peut produire des résultats comparables à ceux de MRF.

## Résumé de la troisième partie

Nous avons montré comment le formalisme basé mouvement direct-rétrograde peut être utilisé dans la segmentation spatio-temporelle. Dans cette partie, on s'est attaché à optimiser, dans les algorithmes de segmentation, la phase d'initialisation et la règle d'affectation d'un pixel à une région. Nous rappelons que l'objectif est d'extraire les objets en terme de mouvement mais pas selon leur niveau de gris. Rendre la phase d'initialisation optimale, consiste à extraire les objets, en tant qu'objets au sens strict du mouvement. Cette phase d'initialisation dépend fortement des résultats d'estimation. Quant à la règle d'affectation,, l'optimisation consiste à introduire les résidus dans les règles de décision. Les résidus sont une fonction du déplacement et par conséquent, l'estimation du mouvement joue un rôle très important dans cette phase.

Deux familles de technique ont été ainsi présentées. Ces techniques intègrent l'estimation du mouvement basée sur un schéma direct-rétrograde. Elles partent d'un ensemble de régions initiales sur lesquelles une opération de groupement ou de partition sera réalisée afin de segmenter les objets constituant l'image.

Pour la première famille, la classification provient d'algorithme d'estimation du mouvement et de la modélisation markovienne. Plus précisément, le principe utilisé consiste à intégrer les informations mouvements dans la phase d'initialisation et dans le terme d'adéquation aux données du processus markovien. A la différence de la segmentation markovienne classique, la fonction d'énergie est, dans notre cas, basée sur l'estimation et la validation de paramètres du mouvement. Ce critère de validation repose sur l'évaluation de la fonction de résidu associée à un pixel. Quelques méthodes qui découlent de ce principe ont été présentées. L'objectif principal est d'améliorer la prise de décision sur l'affectation d'un pixel à une région. Un bon choix de l'ensemble des régions initiales est exigé. C'est l'un des apports essentiels de l'intégration de l'information mouvement. En effet, la phase d'initialisation peut s'effectuer par un seuillage des amplitudes des déplacements. Les différentes techniques proposées ont été testées sur une séquence vidéo. L'association du champ markovien à l'estimation du mouvement permet d'avoir une bonne segmentation. La segmentation obtenue est plus significative au niveau de la structure de la scène. L'élimination des régions parasites est mieux contrôlée.

Quant à la deuxième famille de méthodes, elle fait l'objet d'une segmentation spatio-temporelle à base de l'algorithme K-moyens et s'accompagne de l'estimation du mouvement. Nous avons considéré comme représentant d'une classe K le vecteur de mouvement. La phase d'affectation d'un pixel à une région est la comparaison directe de l'erreur d'estimation résiduelle en fonction du représentant de la classe. Le nombre de classes est choisi comme étant une constante. L'ensemble de régions initiales peut être généré en partant d'un résultat d'estimation du mouvement. Les résultats expérimentaux ont montré les performances de l'approche proposée à la segmentation spatio-temporelle.

En résumé, la motivation principale de cette étude est de montrer la participation de l'estimation du mouvement à la segmentation. Nous avons montré notamment l'importance de la prise en compte du formalisme direct-rétrograde dans la segmentation spatio-temporelle. L'extension du formalisme à  $n$  images permet d'augmenter encore plus les performances.



## **Quatrième partie**

### **Application à l'estimation du retard en sismique répétitive (projet en liaison avec la société GDF)**



# Chapitre 1

## Estimation du retard en sismique répétitive

### 1.1 Introduction

Dans la présente partie, nous abordons la problématique du calcul des temps de retard. Il s'agit d'une estimation du retard en sismique répétitive. Cette étude est menée en liaison avec le centre de Recherche et Développement de la société GDF. Dans ce cadre, GDF a sélectionné un site géographique pour injecter du gaz dans des réservoirs naturels. Plusieurs mesures ont été effectuées sur ces réservoirs et l'analyse ainsi utilisée s'appelle "sismique répétitive" (figure 4.1).

La présente expérimentation permet de localiser la présence de gaz dans un réservoir ainsi que d'évaluer sa densité. Le principe utilisé (figure 4.2) est décrit comme suit :

- En premier lieu, une mesure sur un site vide de gaz est effectuée afin d'avoir un signal de référence ;
- Ensuite, une série d'injections de gaz naturel sur le même site permet d'obtenir des signaux répétés ;
- Finalement, nous estimons le retard  $d$  entre un signal de référence et un signal répété.

Les informations sur le retard estimé seront utiles pour GDF afin d'évaluer la densité de gaz. Dans ce contexte, nous nous chargeons de l'analyse du retard en sismique répétitive en nous préoccupant notamment du filtrage des images et de l'estimation du retard.

Dans les expérimentations suivantes, deux familles d'analyse sont présentées :

- La première est associée à la méthode différentielle monodimensionnelle ;
- Quant à la deuxième approche, nous appliquons des algorithmes basés sur l'approche basée sur un schéma direct-rétrograde et cela, en 2D.

### 1.2 Survol du principe utilisé

Ce paragraphe préliminaire a pour objectif d'introduire, tout d'abord, les notions élémentaires nécessaires à une meilleure compréhension de la prospection sismique.

GDF a implanté son premier site de stockage à Beynes en 1956. Actuellement en France, de nombreux sites sont en exploitation. L'objectif est d'assurer l'ajustement saisonnier des consommations à l'approvisionnement constant de gaz. Cette variation peut venir des fluctuations de température au cours de l'année (plus de 30% du gaz vendu en France est utilisé pour le chauffage).

L'acquisition des données s'effectue à l'aide de la méthode dite de sismique de réflexion. Le but de la prospection sismique est d'obtenir des informations liées à la nature structurale et lithologique du sous-sol. L'exploration sismique consiste à générer des ondes sismiques et à mesurer certaines grandeurs dont le temps mis par ces ondes pour se propager, de la source émettrice à une série de géophones (capteurs acoustiques). Son principe est illustré à la figure 4.1. La source est ici constituée par un camion vibrant ou charge explosive. On peut obtenir une coupe du sous-sol en mesurant les instants d'arrivée des événements réfléchis. Les temps de trajet de chacune de ces ondes dépendront essentiellement des propriétés physiques des structures traversées. Pour chaque tir, un géophone enregistre une trace sismique. Un ensemble de  $N_c$  enregistrements dans une direction donnée, où  $N_c$  est le nombre de capteurs, constitue ce qu'on appelle une section sismique ou image sismique (figure 4.2).

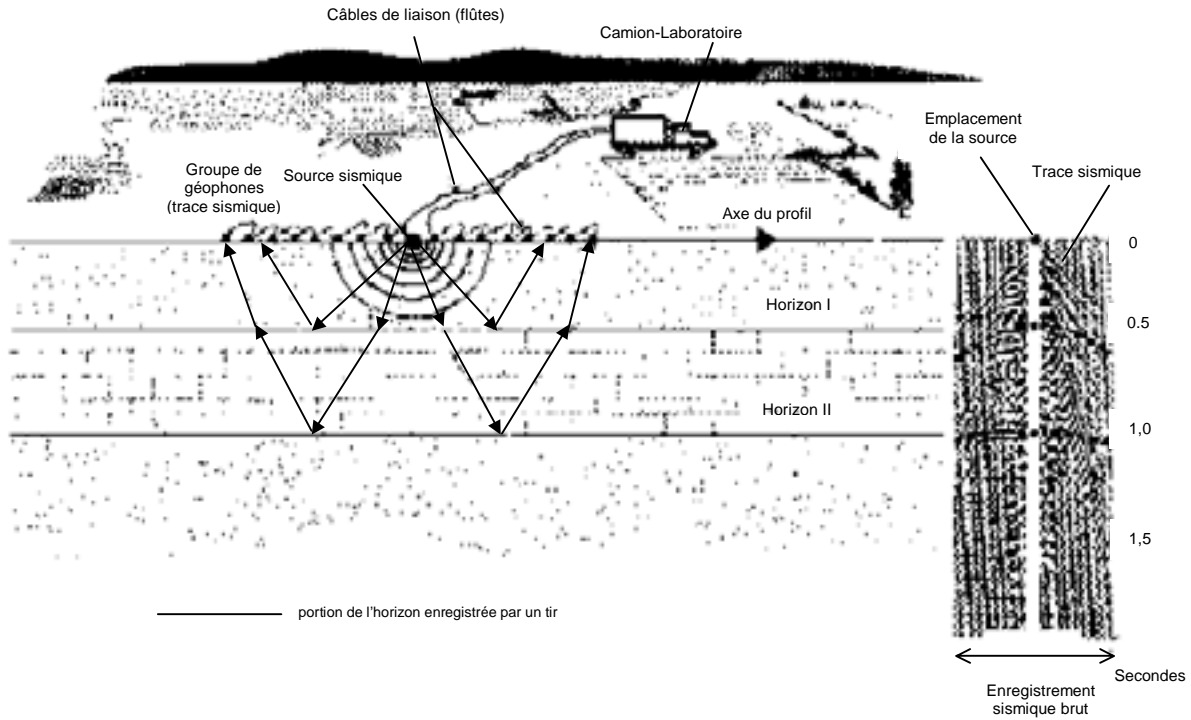


Figure 4.1 : Méthode de prospection sismique réflexion classique.

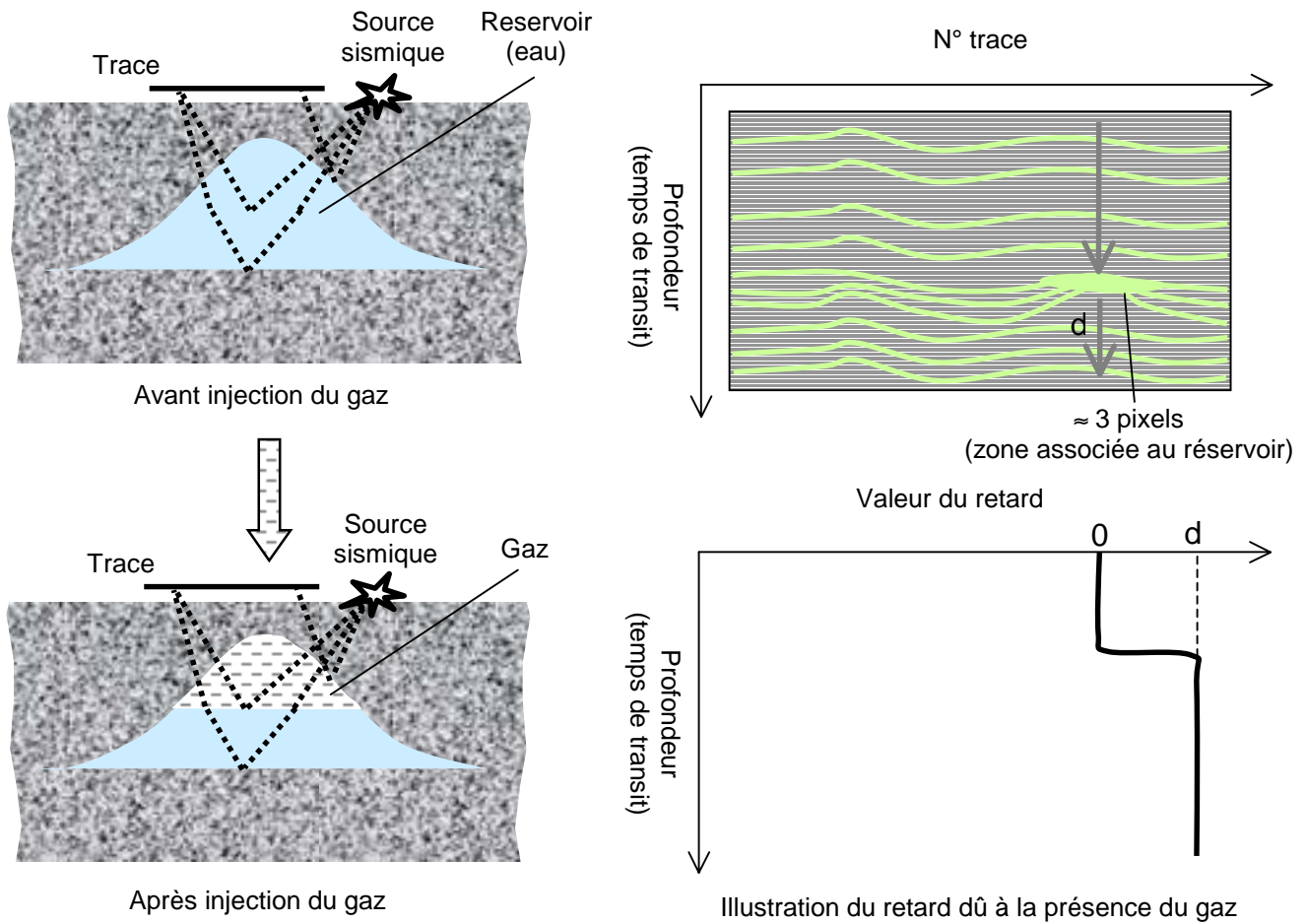


Figure 4.2 : Principe utilisé par la société GDF

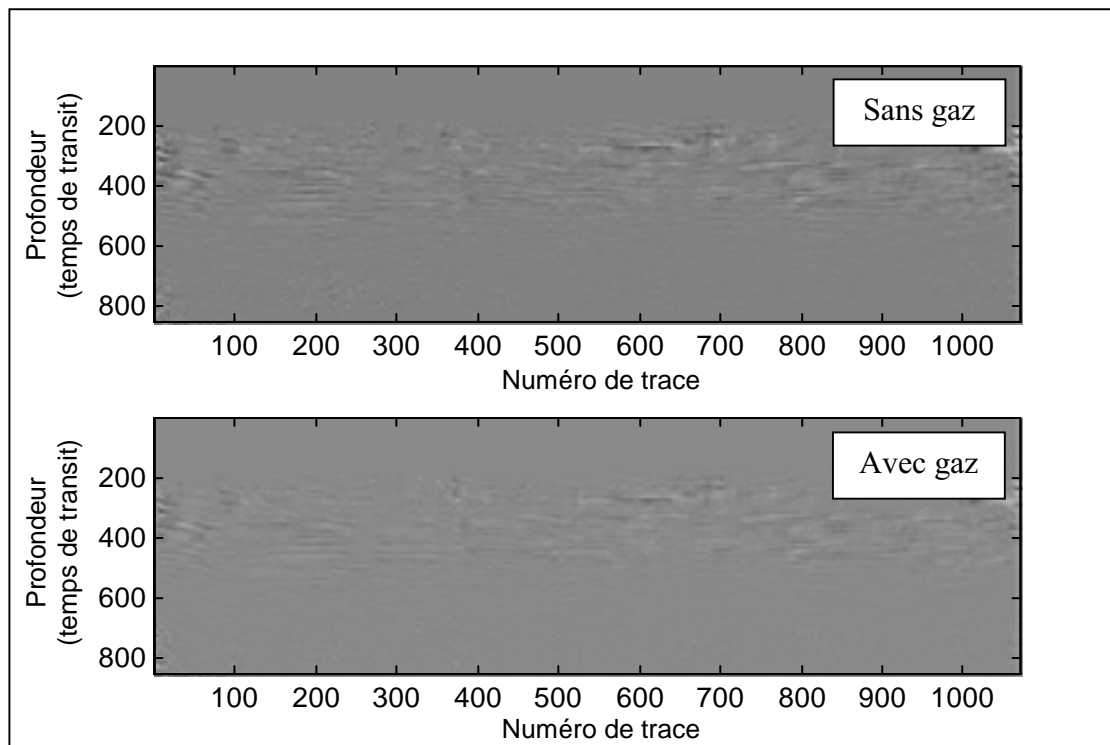


Figure 4.3 : Visualisation des matrices des traces sans gaz et avec gaz.

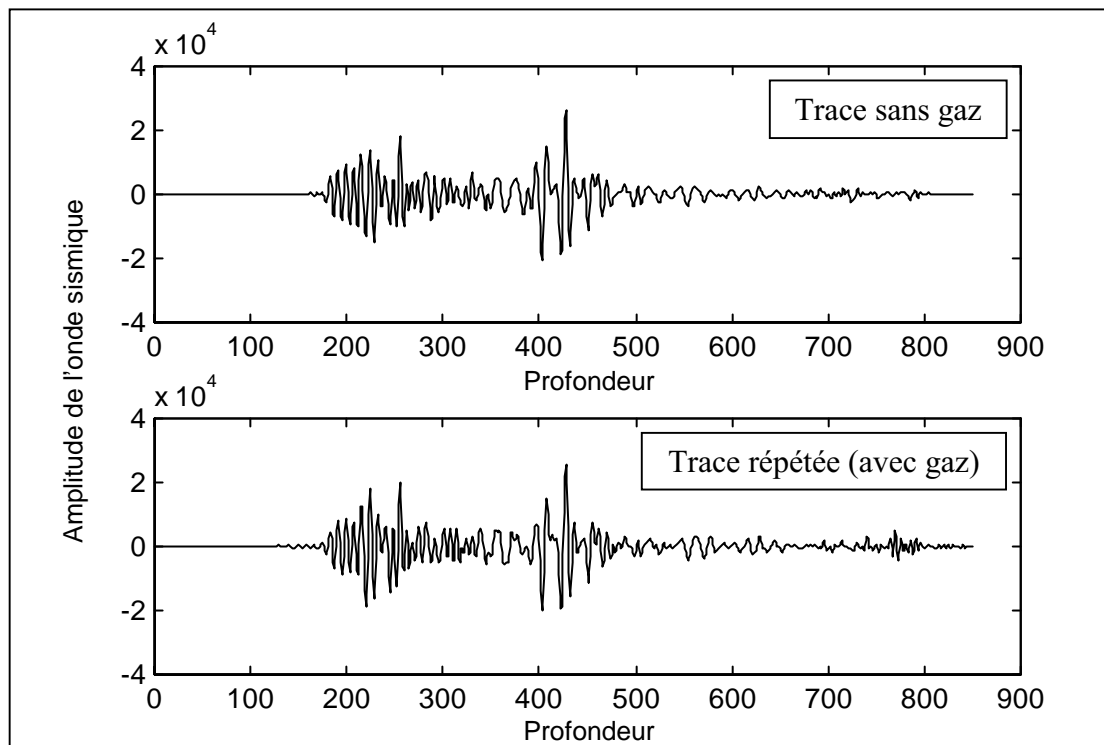


Figure 4.4 : Visualisation d'une trace (n°700) sur les deux images (avant et après injection du gaz) pour une profondeur allant de 1 à 850.



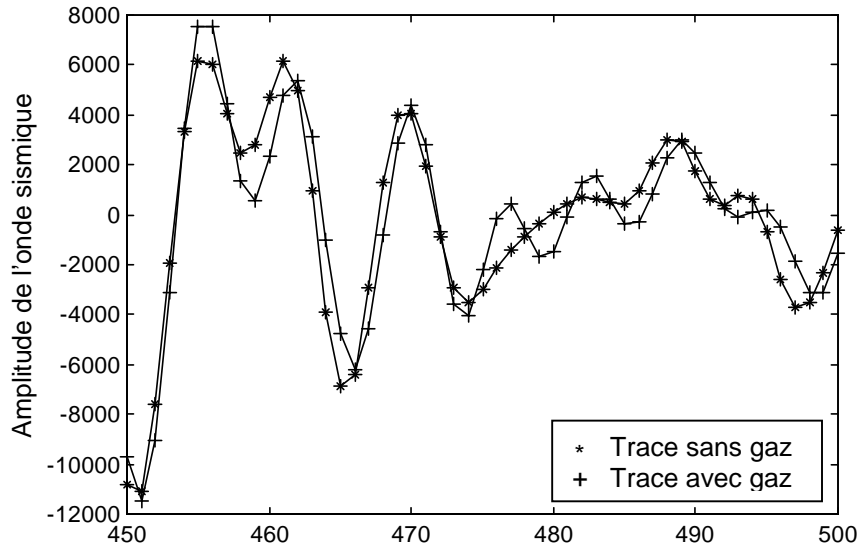


Figure 4.5 : Visualisation du retard entre les deux images (avant et après injection du gaz), pour la trace n°700 et pour une profondeur entre 450 et 500 (zone où se trouve le réservoir).

### 1.3 Étude et observation des supports numériques fournis

Les données sont enregistrées sous un format SEG Y (Society of Exploration Geophysicists Y). Chaque image est représentée sous la forme d'une matrice de  $1170 \times 850$ . Sur la figure 4.3, nous présentons une image sismique sans gaz et avec gaz. La figure 4.4 représente l'allure des traces relevées sur les images (sans gaz et avec gaz). La figure 4.5 illustre le retard entre deux images pour une trace donnée.

### 1.4 Modélisation du signal sismique

Dans les deux premières parties, la notion de modèle a été évoquée à chaque fois qu'un problème était traité. Dans le traitement des images sismiques, il est aussi nécessaire de recourir à un modèle afin d'en accroître la lisibilité et d'en faciliter l'interprétation. Pour cela, un modèle convolutif de représentation des données issues de la géophysique interne peut être suggéré [ROB67]. Le principe est le suivant : pour un tir, le  $i^{\text{ème}}$  géophone enregistre la  $i^{\text{ème}}$  trace sismique, le signal reçu  $r_i(t)$  sur ce capteur résultera de la superposition des  $N_0$  ondes qui auront été détectées à son niveau. Ce signal est supposé être perturbé par un bruit aléatoire, noté  $n_i$ , dû aux différents éléments d'acquisition. Tout cela se traduit par l'équation suivante :

$$r_i(t) = \sum_{j=1}^{N_0} a_{i,j}(t) * s_j(t) + n_i(t) \quad \forall i = 1..N_c \quad (346)$$

La gain  $a_{i,j}$  décrit la propagation de la  $j^{\text{ème}}$  onde sismique jusqu'au  $i^{\text{ème}}$  capteur et  $s_j$  caractérise l'amplitude de la  $j^{\text{ème}}$  onde. Le terme  $n_i(t)$  représente le bruit additif sur le  $i^{\text{ème}}$  capteur, qu'on suppose aléatoire, indépendant des signaux sources et centré. La quantité  $N_0$  désigne le nombre d'ondes sismiques reçues sur le  $i^{\text{ème}}$  capteur ;  $N_c$  est le nombre total de capteurs.

Si la réponse du sous-sol est idéale [MAR97, YIL87], le signal  $r_i(t)$  reçu sur le  $i^{\text{ème}}$  capteur est :

$$r_i(t) = \sum_{j=1}^{N_0} r_j(t) * a_{i,j}(t) s_j(t - d_{j,i}) + n_i(t) \quad \forall i = 1..N_c \quad (347)$$

où  $d_{j,i}$  est le retard en temps entre la  $j^{\text{ème}}$  onde et le  $i^{\text{ème}}$  capteur.

Si on considère qu'on a seulement une onde, alors le signal reçu est :

$$r_i(t) = a_i s(t - d_i) + n_i(t) \quad \forall i = 1..N_c \quad (348)$$

## 1.5 Différentes méthodes d'estimation du retard

L'estimation conventionnelle du temps de retard en sismique est basée sur la corrélation croisée, fonction d'ordre deux, qui mesure les similitudes entre deux mesures dans le domaine du temps. Dans le cas où les sources de bruit seraient gaussiennes et corrélées, la crête de corrélation peut être sensiblement modifiée résultant d'un estimateur de temps de retard imprécis. Une solution alternative est l'estimation du retard à haute résolution dans le domaine d'ordre supérieur.

Nous allons présenter les théories sur les différentes méthodes d'estimation du retard. Certains détails évoqués ci-après sont extraits d'un rapport de contrat entre la société GDF, le laboratoire IXL et l'ADERA [BER00].

### 1.5.1 Méthodes corrélatives

Dans ces méthodes, l'estimation du retard entre deux traces se base sur le principe suivant :

On part de deux signaux reçus  $r_1(t)$  et  $r_2(t)$ , définis selon l'équation (349). Soient :

$$\begin{cases} r_1(t) = s(t) + n_1(t) \\ r_2(t) = As(t-d) + n_2(t) \end{cases} \quad (349)$$

où  $s(t)$  est le signal source,  $d$  est le retard à estimer, et  $n_1(t)$  et  $n_2(t)$  sont des sources de bruit. La quantité  $A$  caractérise le gain relatif d'amplitude. Par la suite, sa valeur est prise égale à l'unité.

L'objectif est d'estimer le retard  $d$  à partir d'observations finies de  $r_1(t)$  et  $r_2(t)$ . Ce problème se rencontre dans plusieurs domaines comme celui du radar, du sonar, de la bio médecine, des communications hertziennes ou des applications géophysiques. La première approche consiste à calculer la fonction d'intercorrélation  $R_{r_1 r_2}(\tau)$  entre  $r_1(t)$  et  $r_2(t)$ . L'intercorrélation [MAX87] entre deux signaux aléatoires  $r_1(t)$  et  $r_2(t)$  est donnée par la relation :

$$R_{r_1 r_2}(\tau) = E[r_1^*(t) r_2(t + \tau)] \quad (350)$$

Le sigle \* désigne la transformée de Fourier discrète complexe conjugué. On peut démontrer que cette équation se traduit par la formule suivante :

$$R_{r_1 r_2}(\tau) = AR_{ss}(\tau - d) + R_{n_1, n_2}(\tau) \quad (351)$$

où  $R_{ss}(\tau)$  représente la fonction d'autocorrélation de la source  $s(t)$  et  $R_{n_1, n_2}(\tau)$  est l'intercorrélation entre les deux processus de bruit.

Dans le cas où les sources de bruit sont des processus aléatoires stationnaires, décorrelés entre eux et décorrelés de  $s(t)$ , de moyenne nulle, la fonction d'intercorrélation  $R_{r_1 r_2}(\tau)$  entre  $r_1(t)$  et  $r_2(t)$  s'écrit :

$$R_{r_1 r_2}(\tau) = R_{ss}(\tau - d) \quad (352)$$

Puisque  $R_{ss}(0) \geq R_{ss}(\tau)$ , la fonction  $R_{r_1 r_2}(\tau)$  sera maximale en  $d$ . La valeur du retard  $d$  peut donc être estimée en choisissant la valeur de  $\tau$  qui maximise la fonction d'intercorrélation.

$$d = \arg \max_{\tau} [R_{r_1 r_2}(\tau)] \quad (353)$$

La méthode d'intercorrélation classique présente des inconvénients majeurs. Si le signal  $s(t)$  est à bande étroite, sa fonction d'autocorrélation va être étalée dans le temps. L'estimateur du

maximum de vraisemblance, qui permet d'avoir accès à la borne de Cramer-Rao<sup>17</sup>, sera ainsi difficile à mettre en œuvre. De plus, s'il existe plusieurs retards et si  $R_{ss}(\tau)$  a un support plus grand que l'écart entre les retards, on ne peut plus distinguer les différents maxima dans  $R_{r_1 r_2}(\tau)$ . On dit alors que les différentes répliques de  $s(t)$  se recouvrent. Pour améliorer les performances de cette méthode, plusieurs démarches ont été entreprises.

La première porte le nom d'intercorrélation généralisée. Elle consiste à pondérer l'interspectre  $S_{r_1 r_2}(f)$  de  $r_1(t)$  et  $r_2(t)$  par une fonction de la fréquence  $w(f)$  et à repasser dans le plan temporel par transformée de Fourier inverse, d'où le nom d'intercorrélation généralisée. L'intervention de  $w(f)$  permet de réduire le support de  $R_{ss}(\tau)$ , ce qui a pour conséquence d'augmenter la résolution temporelle de la méthode. Cependant, les améliorations apportées par les différentes expressions de  $w(f)$  sont contrebalancées par une plus grande sensibilité de la méthode vis-à-vis du rapport du signal sur bruit<sup>18</sup>. Le choix de  $w(f)$  résulte donc d'un compromis entre précision et stabilité.

Parmi toutes les expressions de  $w(f)$ , l'une d'entre elles permet de réaliser l'estimateur du maximum de vraisemblance<sup>19</sup>, ce qui garantit une variance minimale sur l'estimation des retards. Les autres méthodes offrent des avantages différents, comme la robustesse vis-à-vis des sources de bruits.

Ces méthodes souffrent malheureusement des erreurs introduites :

- Soit par l'estimation des fonctions qui interviennent dans le calcul de  $w(f)$  (densité spectrale de puissance des signaux, interspectre, fonction de cohérence);
- Soit par les passages successifs entre le plan temporel et le plan fréquentiel par transformée de Fourier.

Ces erreurs d'estimation ont tendance à se répercuter sur l'estimation de l'intercorrélation généralisée dont les performances se dégradent de façon conséquente. Des pics ou des maxima parasites peuvent apparaître. Ils sont dus :

- Soit à la forme du signal  $s(t)$  ou à celle de sa fonction d'intercorrélation  $R_{ss}(\tau)$  ;
- Soit au faible rapport signal sur bruit ;
- Soit aux sources de bruits  $n_1(t)$  et  $n_2(t)$  qui ne sont pas forcément décorrélées ;
- Soit aux effets liés à la durée des observations des signaux.

Pour pallier les inconvénients liés à la nature même des sources de bruits, une autre solution a été introduite. Elle consiste en l'utilisation des statistiques d'ordre supérieur tels que l'approche basée sur les moments d'ordre trois ou les bispectres. Les propriétés fondamentales des statistiques d'ordre supérieur sont :

- Les cumulants d'ordre supérieur ou égal à trois de processus aléatoires gaussiens sont tous identiquement nuls ;
- Les signaux contenant l'information ne sont pas forcément des processus gaussiens suivant les applications ;
- Les HOS<sup>20</sup> conservent l'information contenue dans la phase des processus non gaussiens.

L'utilisation de cette approche permet de supprimer les effets des sources de bruit gaussiennes tout en gardant l'information sur la phase apportée par le retard dans le plan fréquentiel. En revanche, les HOS ne présentent pas de critère d'optimalité minimisé et la complexité calculatoire des algorithmes de calcul est beaucoup trop élevée à l'heure actuelle pour envisager une implantation temps réel.

<sup>17</sup> CRLB : Cramer-Rao Lower Bound

<sup>18</sup> SNR : Signal to Noise Ratio

<sup>19</sup> ML : Maximum Likelihood

<sup>20</sup> HOS : Higher Order Statistics

### 1.5.2 Méthode basée sur le bispectre ou spectre d'ordre trois.

Ce paragraphe décrit une méthode basée sur les moments d'ordre trois ou les bispectres. La transformée de Fourier d'une corrélation d'ordre trois est souvent appelée bispectre (mais pour éviter les confusions des terminologies dans le domaine temporel et fréquentiel, il est peut-être préférable d'utiliser les termes de spectre d'ordre trois). Supposons que  $r_1(n)$  est la version échantillonnée du signal de référence  $r_1(t)$  et  $r_2(n)$  celle du signal retardé  $r_2(t)$ . Si la source  $s(n)$  est considérée comme un processus aléatoire stationnaire non gaussien de moyenne nulle et les sources de bruit comme des processus aléatoires stationnaires gaussiens de moyenne nulle, statistiquement indépendant de  $s(n)$ , la corrélation d'ordre trois entre le signal  $r_1(t)$  et  $r_1(t)$  est notée :

$$R_{r_1 r_1}(\tau, \rho) = E[r_1(n)r_1(n+\tau)r_1(n+\rho)] = E[s(n)s(n+\tau)s(n+\rho)] = R_{sss}(\tau, \rho) \quad (354)$$

De même entre  $r_2(t)$  et  $r_2(t)$ , on a :

$$R_{r_2 r_2}(\tau, \rho) = E[r_2(n)r_2(n+\tau)r_2(n+\rho)] = E[s(n)s(n+\tau)s(n+\rho)] = R_{sss}(\tau, \rho) \quad (355)$$

Et puis finalement entre  $r_1(t)$  et  $r_2(t)$ , on a :

$$R_{r_1 r_2}(\tau, \rho) = E[r_1(n)r_2(n+\tau)r_1(n+\rho)] = E[s(n)s(n-d+\tau)s(n+\rho)] = R_{sss}(\tau-d, \rho) \quad (356)$$

La quantité  $R_{sss}(\tau, \rho)$  définit la bicorrélation pour le signal émis  $s(n)$  :

$$R_{sss}(\tau, \rho) = E[s(n)s(n+\tau)s(n+\rho)] \quad (357)$$

Les spectres d'ordre trois sont, dans le cas des signaux échantillonnés :

$$B_{r_1 r_1}(w_1, w_2) = \sum_{\tau=-\infty}^{+\infty} \sum_{\rho=-\infty}^{+\infty} R_{r_1 r_1}(\tau, \rho) \quad (358)$$

$$B_{r_2 r_2}(w_1, w_2) = \sum_{\tau=-\infty}^{+\infty} \sum_{\rho=-\infty}^{+\infty} R_{r_2 r_2}(\tau, \rho) \quad (359)$$

$$B_{r_1 r_2}(w_1, w_2) = \sum_{\tau=-\infty}^{+\infty} \sum_{\rho=-\infty}^{+\infty} R_{r_1 r_2}(\tau, \rho) \quad (360)$$

On peut calculer directement ces spectres d'ordre trois à partir des transformées de Fourier de différentes réalisations :

$$B_{r_1 r_1}(w_1, w_2) = TF2D[R_{r_1 r_1}(\tau, \rho)] = B_{sss}(w_1, w_2) \quad (361)$$

$$B_{r_2 r_2}(w_1, w_2) = TF2D[R_{r_2 r_2}(\tau, \rho)] = B_{sss}(w_1, w_2) \quad (362)$$

$$B_{r_1 r_2}(w_1, w_2) = TF2D[R_{r_1 r_2}(\tau, \rho)] = B_{sss}(w_1, w_2)e^{jw_1 d} \quad (363)$$

En conséquence, le temps de retard apparaît dans le bispectre comme une fonction de phase. Suivant que l'on fasse intervenir le bispectre  $B_{r_2 r_2}(w_1, w_2)$  ou non, plusieurs méthodes peuvent être adoptées pour calculer le retard  $d$ . Une de ces méthodes est le calcul du signal d'hogramme [SAT77] :

$$I(w_1, w_2) = \frac{B_{r_1 r_2}(w_1, w_2)}{B_{sss}(w_1, w_2)} \quad (364)$$

Étant donné les expressions des spectres d'ordre trois, on obtient :

$$I(w_1, w_2) = e^{jw_1 d} \quad (365)$$

ainsi que la fonction bispectre-corrélation  $T(\tau)$  [NIK88], définie par :

$$T(\tau) = \iint I(w_1, w_2) e^{-jw_1 \tau} dw_1 dw_2 \quad (366)$$

On peut déterminer le temps de retard entre les deux signaux comme étant la valeur  $\tau$  qui maximise la fonction  $T(\tau)$  :

$$d = \arg \max_{\tau} [T(\tau)] \quad (367)$$

### 1.5.3 Méthode de l'intercorrélation généralisée

Un pré-filtrage des signaux avant de définir la corrélation, mène à la méthode nommée intercorrélation généralisée.

Cette approche suppose que les signaux  $s(t)$ ,  $n_1(t)$  et  $n_2(t)$  soient des processus aléatoires réels, conjointement stationnaires et décorrelés entre eux. Ces conditions permettent d'écrire les relations suivantes :

$$\begin{cases} S_{r_1} (f) = S_{ss} (f) + S_{n_1 n_1} (f) \\ S_{r_2} (f) = S_{ss} (f) + S_{n_2 n_2} (f) \\ S_{r_1 r_2} (f) = S_{ss} (f) e^{-j2\pi f d} \end{cases} \quad (368)$$

La fonction d'intercorrélation généralisée  $R_{r_1 r_2}^G(\tau)$  est définie par la relation suivante :

$$R_{r_1 r_2}^G(\tau) = \int S_{r_1 r_2} (f) w(f) e^{j2\pi f \tau} df \quad (369)$$

La valeur du retard  $d$  correspond à la valeur de  $\tau$  qui maximise cette fonction. Soit :

$$d = \arg \max_{\tau} [R_{r_1 r_2}^G(\tau)] \quad (370)$$

Suivant l'expression de  $w(f)$ , diverses méthodes (ROTH, SCOT, PATH, Hannan-Thomson) [CAR87] sont suggérées pour pondérer la fonction d'intercorrélation. L'idée est de faire en sorte que  $R_{ss}(\tau)$  soit beaucoup plus fin aux alentours de  $d$  pour éviter toute confusion avec d'autres maxima parasites dus à la forme du signal ou aux sources de bruits. Cependant, plus on s'efforce d'obtenir des pics étroits, plus on risque d'avoir des pics parasites. D'un autre coté, plus les pics sont larges et moins ils sont précis. Ainsi, le choix de la fonction  $w(f)$  résulte d'un compromis entre stabilité et précision. En pratique, on ne peut avoir accès qu'aux estimateurs de  $S_{r_1 r_2}(f)$  et  $w(f)$ .

Nous donnons, dans ce qui suit, une description des différentes méthodes suivant l'expression de  $w(f)$ .

### 1.5.4 Méthode ROTH

Dans cette méthode, l'expression de  $w(f)$  est la suivante :

$$w_r(f) = \frac{1}{S_{r_1} (f)} \quad (371)$$

La fonction d'intercorrélation devient donc :

$$R_{r_1 r_2}^G(\tau) = \int \frac{S_{r_1 r_2} (f)}{S_{r_1} (f)} e^{j2\pi f \tau} df \quad (372)$$

En introduisant les relations du système (368) dans la relation (372), nous obtenons :

$$R_{r_1 r_2}^G(\tau) = \int \frac{S_{ss} (f) e^{-j2\pi f d}}{S_{ss} (f) + S_{n_1 n_1} (f)} e^{j2\pi f \tau} df \quad (373)$$

Finalement, nous avons :

$$R_{r_1 r_2}^G(\tau) = \delta(\tau - d) * \int \frac{S_{ss} (f)}{S_{ss} (f) + S_{n_1 n_1} (f)} e^{j2\pi f \tau} df \quad (374)$$

L'avantage de cette méthode est d'atténuer les bandes fréquentielles où  $S_{n_1 n_1}(f)$  est grand.

### 1.5.5 Méthode SCOT<sup>21</sup>

Les erreurs d'estimation de  $S_{r_1 r_2}(f)$  peuvent être dues aux plages de fréquences où  $S_{n_2 n_2}(f)$  est grand, de même qu'aux plages où  $S_{n_1 n_1}(f)$  est grand. Ainsi, on ne peut raisonnablement choisir entre  $w_r(f) = \frac{1}{S_{r_1 r_1}(f)}$  et  $w_r(f) = \frac{1}{S_{r_2 r_2}(f)}$ . Pour la méthode SCOT, on préfère choisir le compromis suivant :

$$w_s(f) = \frac{1}{\sqrt{S_{r_1 r_1}(f) S_{r_2 r_2}(f)}} \quad (375)$$

qui donne :

$$R_{xy}^G(\tau) = \int \frac{S_{r_1 r_2}(f)}{\sqrt{S_{r_1 r_1}(f) S_{r_2 r_2}(f)}} e^{j2\pi f \tau} df \quad (376)$$

Cette méthode assure un compromis pour éviter les effets néfastes induits par les bandes de fréquences où  $S_{n_1 n_1}(f)$  ou  $S_{n_2 n_2}(f)$  sont grands par rapport aux signaux utiles.

Si les deux sources de bruit ne sont pas nulles, la méthode SCOT apporte alors les mêmes perturbations que la méthode ROTH, à savoir un étalement de la fonction de Dirac autour de  $d$ . L'étalement persiste parce que l'on n'arrive pas à blanchir correctement la densité interspectrale de puissance  $S_{r_1 r_2}(f)$ .

### 1.5.6 Méthode PHAT<sup>22</sup>

Contrairement à d'autres techniques, cette approche n'exige pas la modélisation des caractéristiques statiques de la source du signal et du bruit. L'approche PHAT est indépendante des caractéristiques du signal source à moins que les signaux soient strictement à bande étroite.

Dans les méthodes précédentes, on n'évite pas un élément du pic en  $\tau = d$ . On définit alors une autre méthode avec la fonction suivante :

$$w_p(f) = \frac{1}{|S_{r_1 r_2}(f)|} \quad (377)$$

Dans ce cas, si  $n_1(t)$  et  $n_2(t)$  sont décorrélés, la fonction d'intercorrélacion devient :

$$R_{r_1 r_2}^G(\tau) = \int \frac{S_{r_1 r_2}(f)}{|S_{r_1 r_2}(f)|} e^{j2\pi f \tau} df \quad (378)$$

or on a :

$$\frac{S_{r_1 r_2}(f)}{|S_{r_1 r_2}(f)|} = e^{-j2\pi f d} \quad (379)$$

La fonction d'intercorrélacion s'écrit finalement :

$$R_{r_1 r_2}^G(\tau) = \delta(\tau - d) \quad (380)$$

La méthode PHAT a été développée comme une technique ad hoc pour répondre au problème posé. Théoriquement, cette méthode ne souffre pas des étalements de pics comme les autres. Cependant, si les estimations des densités interspectrales de puissance ne sont pas exactes, les erreurs introduites sont conséquentes. Un autre inconvénient vient du fait que l'interspectre  $S_{r_1 r_2}(f)$  est pondéré par l'inverse de  $S_{ss}(f)$ . Si  $S_{ss}(f)$  est nul dans une certaine bande, alors la phase  $\theta(f)$  est indéterminée dans cette bande.

<sup>21</sup> Smoothed coherence transform

<sup>22</sup> Phase transform

### 1.5.7 Méthode de l'estimateur du maximum de vraisemblance

On montre que la fonction  $w(f)$ , introduite par Hannan et Thomson, permet de mettre en œuvre l'estimateur du ML. La fonction de pondération est ici de la forme :

$$w_{ML}(f) = \frac{1}{|S_{r_1}(f)| \left[ 1 - |\gamma_{r_1 r_2}(f)|^2 \right]} \quad (381)$$

La quantité  $\gamma_{r_1 r_2}$  est une fonction de cohérence.

Parmi les estimateurs que nous venons de présenter, celui du maximum de vraisemblance présente l'intérêt particulier d'atteindre le CRLB. Malheureusement, cet indicateur des performances des méthodes d'estimation n'a de sens que lorsque l'estimation se révèle exempte d'ambiguïté, c'est-à-dire lorsqu'il n'y a pas de risque de confondre le pic de l'intercorrélacion qui doit donner l'estimation du retard et un pic parasite. Cela peut arriver dans plusieurs cas, notamment quand la fonction d'autocorrélacion du signal d'entrée présente plusieurs pics rapprochés, quand le rapport signal sur bruit est faible ou quand les erreurs d'estimation sont grandes. Dans ces cas, la CRLB ne donne qu'une estimation très optimiste des performances de l'estimateur du ML et il faut bien souvent recevoir ces mesures à la baisse.

### 1.5.8 Discussion sur les différentes méthodes d'estimation du retard

Diverses méthodes ont été suggérées pour estimer le temps de retard. Les avantages et les limitations de chaque technique sont liés à la nature des sources de signal et de bruit. Le choix de telle ou telle technique pour l'estimation du retard entre deux signaux doit se faire suivant la connaissance du problème que l'on a a priori. Suivant la distribution du signal retardé  $s(t)$  et ses caractéristiques statistiques, suivant que les sources de bruit soient gaussiennes ou non, blanches ou colorées, corrélées ou décorrélées, les différentes techniques sont plus ou moins adaptées.

L'utilisation des statistiques d'ordre supérieur permet de réaliser une extraction plus efficace des signaux pollués par un bruit gaussien et aussi de conduire à une analyse plus fine des signaux. En général, les techniques développées à l'ordre deux peuvent être transposées à des ordres supérieurs. Il reste le désavantage lié au temps de calcul plus grand, d'où un coût élevé. Généralement, on choisit l'estimateur du maximum de vraisemblance. Mais nous avons montré que les performances de cet estimateur n'avaient de sens que lorsque l'estimation se faisait dans de bonnes conditions. Il convient alors de choisir un estimateur plus robuste vis-à-vis du bruit comme l'estimateur de ROTH ou de SCOT.

METHODES		AVANTAGES	INCONVENIENTS
Intercorrélacion simple		Simplicité	Dépend de $R_{ss}(\tau)$ Variance non minimale
Intercorrélacion généralisée	ROTH	Robuste par rapport au bruit $n_1(t)$	Variance non minimale
	SCOT	Robuste par rapport aux bruits $n_1(t)$ et $n_2(t)$	Variance non minimale
	PHAT	Théoriquement ad hoc	Très sensible aux erreurs d'estimation sur $S_{r_1 r_2}(f)$
	ML	CRLB	

Tableau 4.1 : Comparaison des différentes méthodes d'intercorrélacion.

## 1.6 Méthode utilisée par la société GDF

### 1.6.1 Description

Nous rappelons très brièvement la méthode introduite par la société GDF. Cette méthode utilise le domaine spectral, qui s'appuie sur la relation suivante :

$$S_{r_1 r_2}(f) = S_{ss}(f) e^{-j2\pi f d} \quad (382)$$

Si l'on suppose que la densité spectrale de puissance du signal test  $s(t)$  est blanche, on peut estimer le retard en se basant sur le terme de l'exponentielle et en moyennant ce terme sur toutes les fréquences disponibles, soit :

$$d = \frac{1}{N} \sum_{i=1}^N \frac{\arg[S_{r_1 r_2}(f_i)]}{2\pi f_i} \quad (383)$$

où  $N$  est le nombre de points de l'interspectre  $S_{r_1 r_2}(f)$ .

Cette méthode suppose que l'on a préalablement blanchi l'interspectre  $S_{r_1 r_2}(f)$  par rapport au signal test  $s(t)$  et que l'on a supprimé les plages de fréquences où la densité spectrale de puissance de  $s(t)$  est nulle. Cette opération peut être utilisée en sélectionnant, dans un premier temps, une plage de fréquences pour laquelle seront conservées les valeurs de l'interspectre. Une fois que cette plage est sélectionnée, on ne conserve que les valeurs de l'interspectre qui correspondent à une valeur suffisamment élevée de la fonction de cohérence entre les signaux  $r_1(t)$  et  $r_2(t)$ .

Le grand avantage de cette méthode est de pouvoir disposer de plusieurs échantillons pour effectuer l'estimation. Comparativement aux précédentes méthodes fondées sur une technique de détection de maximum, cette méthode est plus robuste. L'inconvénient majeur de cette méthode repose sur le fait qu'il faut filtrer empiriquement les échantillons qui vont servir au calcul du retard.

### 1.6.2 Implantation

Les premières approches ont tenté de mettre en œuvre directement les méthodes basées sur l'intercorrélation généralisée. Étant donné que les résultats sur des retards artificiels n'ont pas été concluants, il a été décidé de coupler ces approches avec la méthode utilisée par GDF. Ainsi, les méthodes corrélatives ont été implantées de la façon suivante (figure 4.6) :

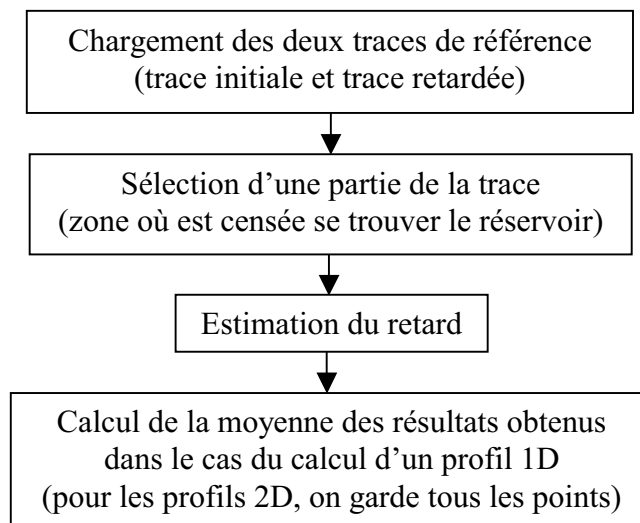


Figure 4.6 : Implantation des méthodes corrélatives.



Toutes les méthodes fonctionnent de la façon indiquée par la figure 4.7.

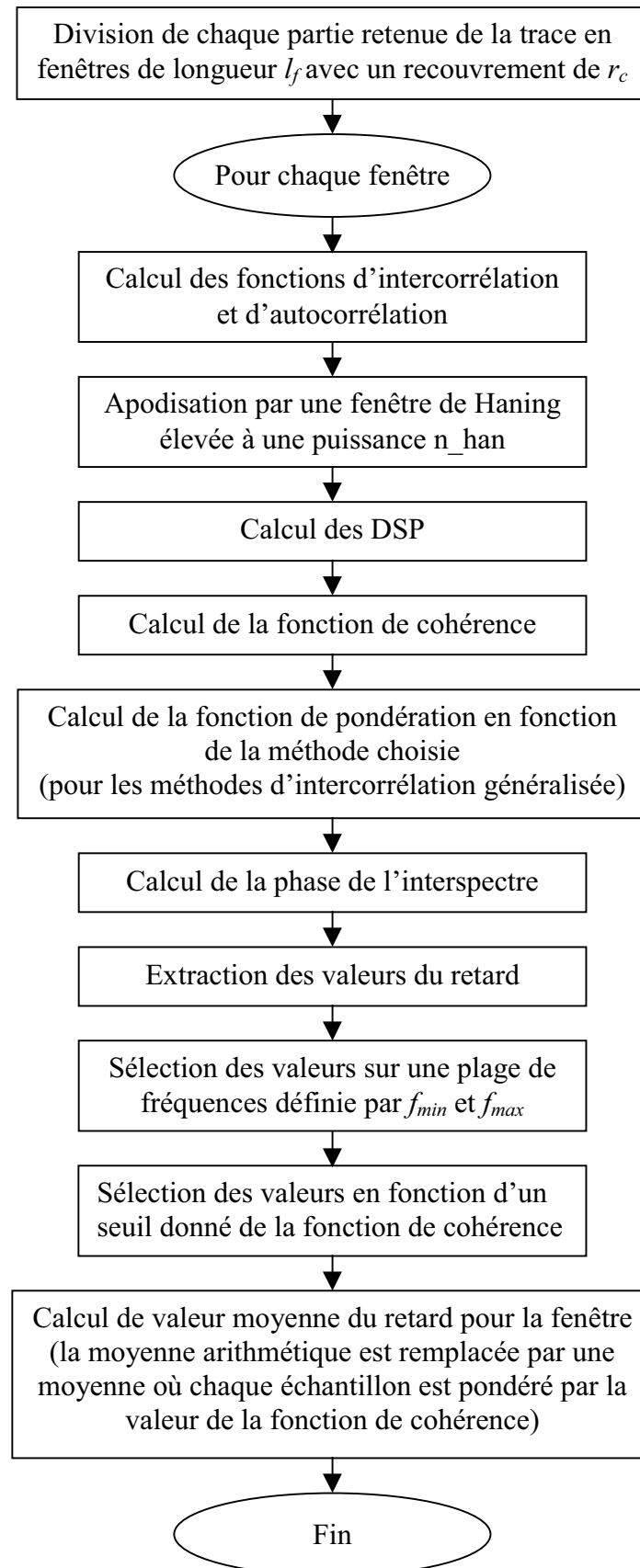


Figure 4.7 : Algorithme d'estimation du retard pour une méthode corrélative.

## 1.7 Méthodes dérivatives suivant une direction spatiale 1D

### 1.7.1 Méthode dérivative basée sur un mouvement direct des ondes

L'estimation du temps de retard repose sur l'exploitation d'une équation différentielle à base de déplacement direct (figure 4.8).

Le déplacement entre deux ondes introduit une déformation du signal. Cette déformation peut se décrire par une relation sous forme différentielle. C'est l'équation de contrainte dans un repère spatial monodimensionnel. Le temps d'échantillonnage est noté par  $T_e$ .

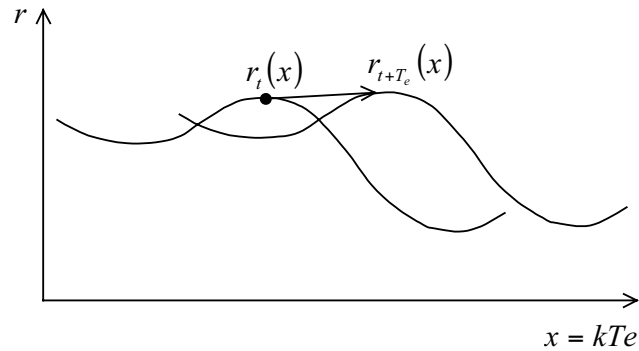


Figure 4.8 : Formulation basée sur un mouvement direct des ondes.

Nous n'allons plus détailler les différentes étapes du développement pour arriver à l'équation de contrainte du mouvement suivant :

$$\frac{\partial r}{\partial x} d + \frac{\partial r}{\partial t} T_e = 0 \quad (384)$$

La méthode de résolution est la même que celle en 2D. Nous n'allons plus présenter le détail du calcul. La solution de cette équation au sens des moindres carrés est de la même forme que l'équation (149). La matrice des gradients spatiaux est ici une matrice colonne.

Pour les données sismiques, l'estimé  $d$  correspond à une valeur de retard pour une fenêtre d'observation donnée. En réalisant le déplacement de la fenêtre, nous pouvons générer une estimation le long de la trace et ceci pour tout le profil.

### 1.7.2 Méthode dérivative basée sur un mouvement direct-rétrograde des ondes

On considère ici le sens direct et le sens rétrograde du mouvement entre deux ondes. Le déplacement est considéré uniquement dans un repère spatial monodimensionnel. Le déplacement dans le sens direct est représenté à la figure 4.9. Celui dans le sens rétrograde est montré à la figure 4.10.

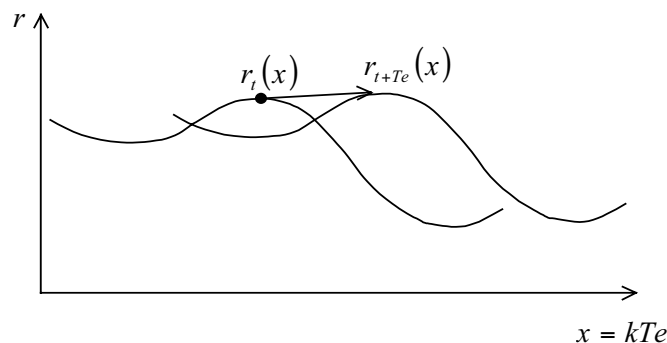


Figure 4.9 : Formulation basée sur un mouvement direct-rétrograde des ondes ; présentation du déplacement direct.

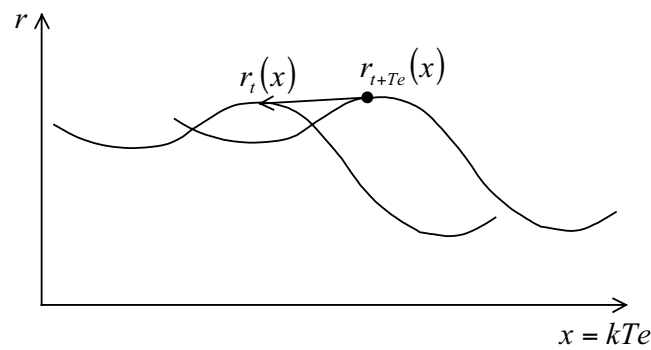


Figure 4.10 : Formulation basée sur un mouvement direct-rétrograde des ondes ; présentation du déplacement rétrograde.

Comme nous avons pu le constater dans les méthodes des moindres carrés basées schéma direct-rétrograde, deux solutions peuvent être considérées selon qu'il s'agisse de la minimisation de la fonctionnelle constituée de la somme de termes quadratiques reflétant les deux erreurs (critère  $J_a$ ), ou la minimisation de la fonctionnelle constituée d'un terme quadratique de la somme algébrique des erreurs (critère  $J_b$ ). Les équations (168, 173) donnent les solutions. Comme l'opération se déroule dans un espace 1D, les matrices des gradients spatiaux, approchés sur l'onde à l'instant  $t$  ou ceux estimés sur l'onde à  $t + Te$ , sont des matrices colonnes.

## 1.8 Méthodes dérivatives dans un espace spatial 2D

L'objet de ce paragraphe n'est pas de refaire la description des méthodes différentielles en 2D car cette description a déjà été présentée antérieurement. Nous tenons seulement à signaler que dans la présente expérimentation, compte-tenu de la structure de données, la résolution du système au sens des moindres carrés exige un support de calcul très grand. Le choix d'un support de calcul plus large permet d'introduire implicitement un filtre de lissage spatial. De plus, comme le sens du déplacement est essentiellement vertical, le support de calcul utilisé est de forme rectangulaire. Or, même si l'utilisation d'un tel support de calcul permet d'obtenir une meilleure estimation, il introduit une non homogénéité des résultats et une perte d'information, en particulier une diminution de l'amplitude des retards (due au phénomène de lissage). Pour pallier ces pertes d'informations, nous avons décidé de procéder à un pré-traitement des données. Il s'agit d'un filtrage amont des images, c'est-à-dire un filtrage des traces initiales. Le paragraphe préliminaire suivant décrit cette opération.

### 1.8.1 Pré-traitement des données

#### a) Objectif de l'opération de pré-traitement

Nous essayons d'analyser les résultats obtenus avec les données d'un profil (profil M01, référence utilisée par la société GDF) dont l'image sismique d'origine est présentée à la figure 4.11.

Sans précaution de pré-traitement des données, les méthodes d'estimation du retard donnent typiquement des résultats similaires à ceux présentés par la figure 4.12. Ce test a été effectué avec les données du profil M01 sur la plage [300-500] en profondeur et [100-900] en nombre de trace. Nous utilisons ici la méthode des moindres carrés à flot direct-rétrograde. Les paramètres d'estimation sont les suivants :

- Méthode des moindres carrés à formalisme direct-rétrograde, sans pondération de données ;
- Masque MDIF pour les gradients spatiaux ;
- Masque Mask\_t4 pour le gradient temporel ;
- Support de calcul :  $81 \times 21$  pixels .

La figure 4.13 présente le profil cumulé du retard.

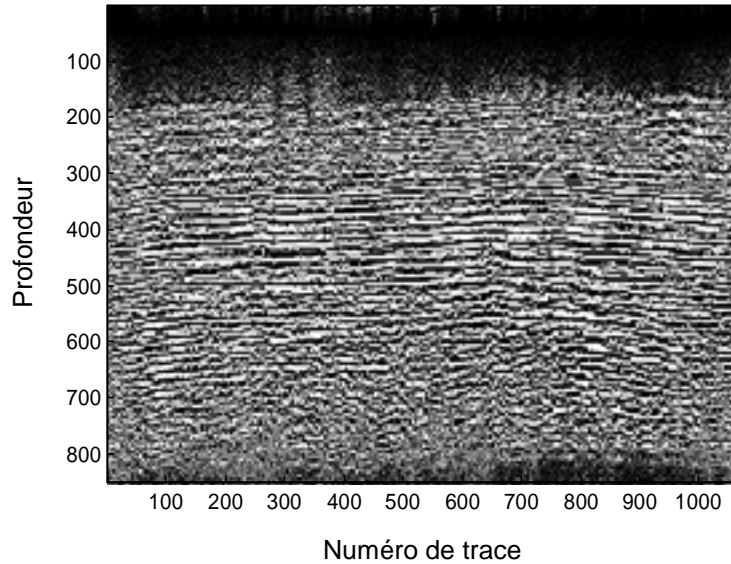


Figure 4.11 : Image sismique d'origine (sans filtrage) du profil M01.

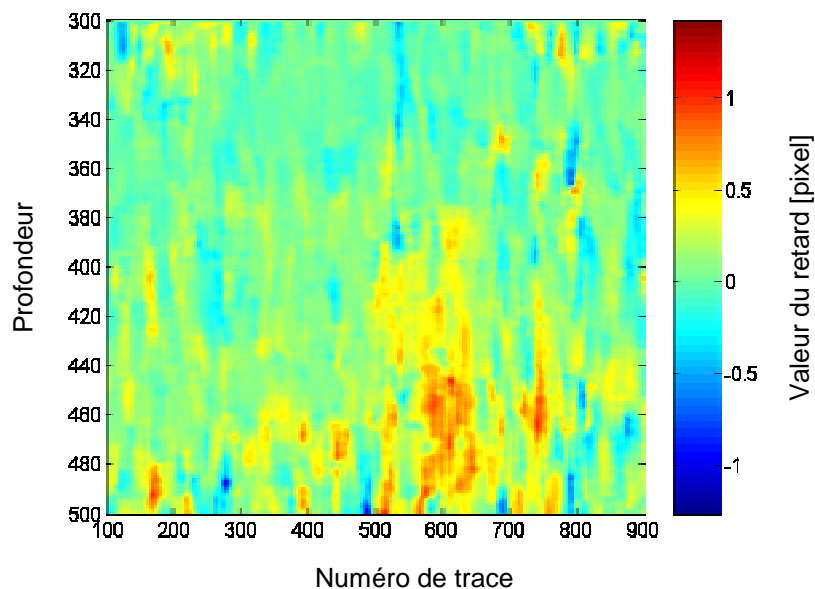


Figure 4.12 : Estimation du retard sur l'image sismique d'origine, sans pré-traitement des données ; composante suivant l'axe vertical (profondeur) du vecteur estimé à l'aide de la méthode dérivateur 2D et avec un schéma direct-rétrograde ( $J_b$ ) ; support de calcul : 81 pixels en profondeur et 21 pixels en nombre de traces.

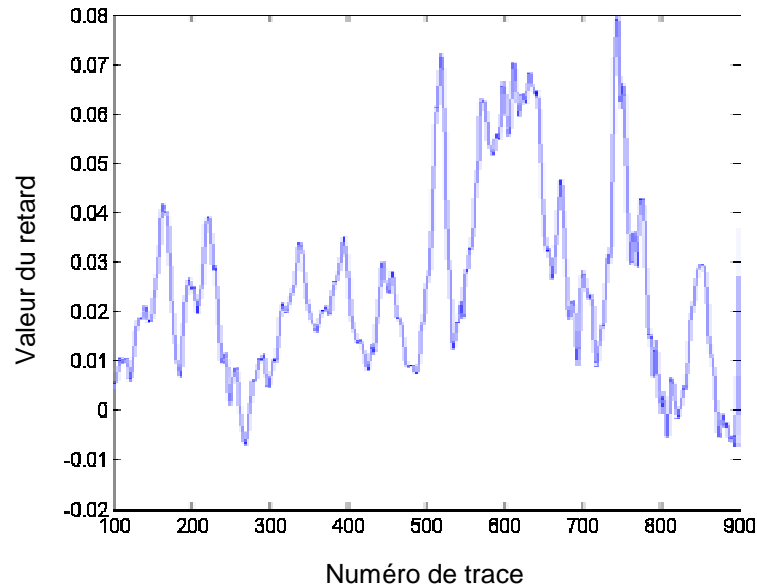


Figure 4.13 : Profil cumulé du retard pour une estimation sans pré-traitement des données.

Le résultat obtenu présente un niveau de perturbation élevé. Pour trouver la trace de ces perturbations, nous pouvons visualiser le spectre 2D en module associé à l'image. La figure 4.14 donne le résultat obtenu. D'après la forme du spectre, une opération de filtrage doit être envisagée. Cette opération doit répondre à deux objectifs principaux, qui sont l'élimination du bruit et le rehaussement des transitions inter-régionales. Dans la logique de notre processus d'analyse qui se décompose en trois phases que sont le filtrage, l'estimation du temps de retard ainsi que la segmentation, nous avons recherché des méthodes de filtrage 2D permettant de lisser le profil du retard. Nous envisageons les deux types de filtrage suivants :

- Filtre bidimensionnel à angle d'ouverture réglable ;
- Filtre gaussien modulé de type elliptique.

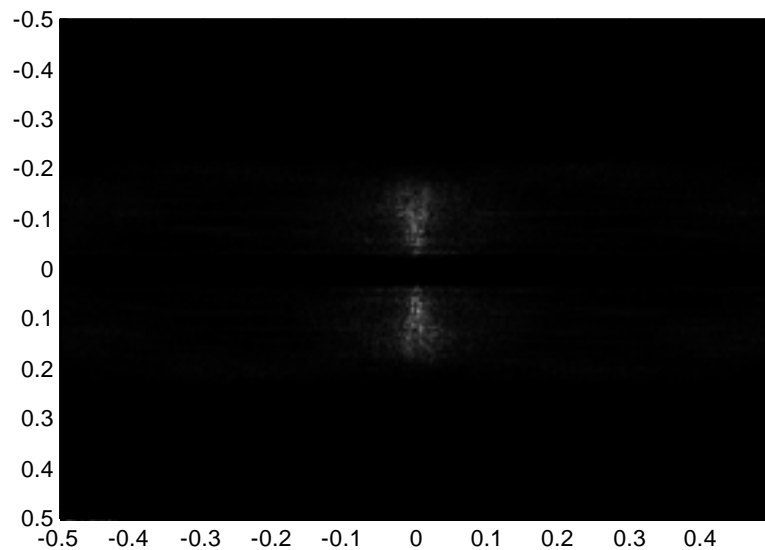


Figure 4.14 : Spectre 2D en module associé à l'image sismique du profil MO1.

**b) Premier type de filtrage : filtre bidimensionnel de type triangulaire à angle d'ouverture réglable**

Il s'agit ici de mettre en œuvre un filtre monodimensionnel, et d'en régler la sélectivité à l'aide de la valeur de son angle d'ouverture. Ce filtre est basé sur les statistiques ordonnées qui trient les échantillons du support de filtrage selon un critère de pavage numérique. Le support de filtrage est présenté à la figure 4.15. Cet opérateur élimine de manière plus efficace le bruit impulsionnel et sa sortie constitue une estimation optimale pour des distributions de type gaussien.

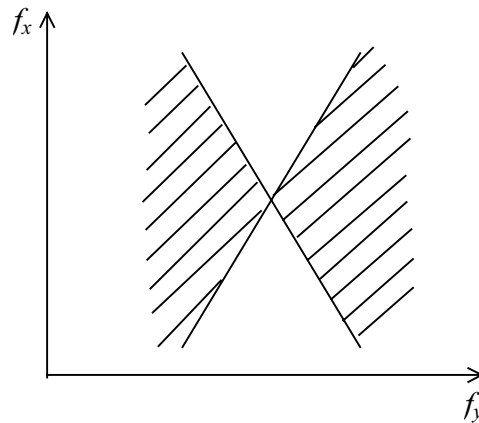


Figure 4.15 : Gabarit du filtre bidimensionnel à angle d'ouverture réglable.

Le pavage numérique est évalué sur la totalité de l'image en partant d'une fenêtre constituée par les composantes :

$$z(f_x, f_y) = a \tan \left| \frac{f_x}{f_y} \right| \quad (385)$$

La distribution du filtre est ajustée localement par un modèle gaussien selon l'équation suivante :

$$G(f_x, f_y) = K(f_x, f_y) \exp \left[ - \frac{\left( \frac{f_y - \frac{f_{y_{\max}}}{2}}{\frac{f_{y_{\max}}}{2}} \right)^2}{2(\tan \theta)^2} \right]^2 \quad (386)$$

Le terme  $\theta$  désigne l'angle d'ouverture du filtre. La quantité  $K$  conditionne le pavage, sa valeur est donnée par :

$$K(f_x, f_y) = \begin{cases} 0, & \text{si } z(f_x, f_y) > \theta \\ 1, & \text{si } z(f_x, f_y) \leq \theta \end{cases} \quad (387)$$

La transformée de Fourier 2D de l'image est multipliée avec la fonction  $G$ . La transformée inverse du résultat de cette opération permet d'obtenir le résultat filtré. Ce filtre élimine le bruit haute fréquence correspondant à un pendage caractérisé par l'ouverture  $\theta$ , et permet en effet d'améliorer l'estimation du temps de retard entre deux traces. En contrepartie, le type de filtre utilisé n'est pas capable d'effectuer un lissage systématique. La raison est que, pour

certaines profils, la fenêtre de lissage de ce filtre n'est pas favorable à la forme du spectre. La proposition suivante vise à adapter le type de filtre selon la forme du spectre.

### c) Second type de filtrage : filtre gaussien modulé de type elliptique

La forme du filtre à mettre en œuvre est de type elliptique. La modulation a pour objectif de centrer l'ellipse sur chacun des lobes du spectre. Afin de générer la forme elliptique, nous avons choisi un filtre dont la réponse impulsionnelle est guidée par l'équation monodimensionnelle suivante :

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma_x} \exp\left(-\frac{x^2}{2\sigma_x^2}\right) \quad (388)$$

L'extension de ce type de filtre en 2D est donnée par la relation :

$$g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2}\right) \quad (389)$$

L'intérêt de conserver la séparation entre les deux dimensions spatiales se caractérise par la possibilité de générer des filtres ayant des réponses en fréquences elliptiques ou circulaires. La translation fréquentielle est donnée par le produit caractérisé par l'équation suivante :

$$h(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2}\right) \cos(2\pi f_0 x) \quad (390)$$

La fréquence  $f_0$  permet de centrer l'action du filtrage sur la zone fréquentielle choisie ;  $\sigma_x$  et  $\sigma_y$  désignent les écarts types.

## 1.8.2 Résultats expérimentaux

Cette partie concerne uniquement les résultats de simulation obtenus avec les méthodes différentielles 2D. Les méthodes 1D ne seront pas présentées. En effet, une méthode monodimensionnelle n'est qu'un cas particulier des méthodes bi-dimensionnelles. Nous avons donc, dans certains tests, utilisé un support de calcul rectangulaire, pour se rapporter à l'estimation monodimensionnelle.

Le premier volet du test concerne les résultats issus du premier type de filtrage. Les méthodes des moindres carrés à base de flot direct et celles à base de flot direct-rétrograde sont ainsi toutes testées.

Nous abordons ensuite les expérimentations sur différents profils, en utilisant le second type de filtrage. De nombreux résultats de la méthode des moindres carrés à base de schéma direct-rétrograde sont illustrés dans les paragraphes suivants.

### a) Profil obtenu en utilisant le premier type de filtrage

- *Méthode des moindres carrés à flot direct-rétrograde*

La figure 4.16 représente le profil des retards estimé à partir de la méthode des moindres carrés basée schéma direct-rétrograde. Ce profil s'appelle MO3 (référence employée par la société GDF). Le filtre utilisé est de type bidimensionnel à angle d'ouverture réglable. Ce test a été effectué avec un angle d'ouverture de  $2\pi/9$ . Dans cette expérimentation, nous commençons à mettre en évidence les améliorations apportées par la méthode dérivative.

Nous pouvons aussi remarquer l'impact du filtrage en amont sur les résultats. Le profil des retards est moins bruité. En revanche, le type de filtrage utilisé ne permet pas d'obtenir un profil net. D'autre part, les résultats numériques ne sont pas très concluants.

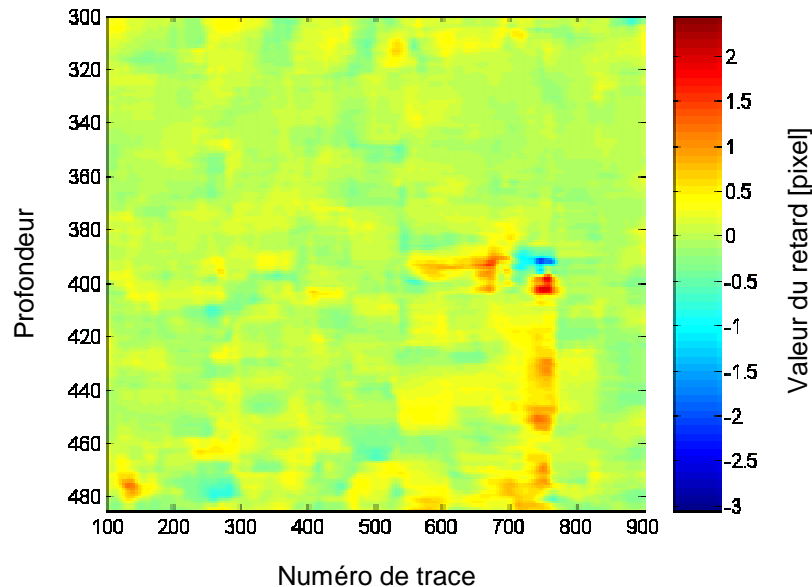


Figure 4.16 : Profil du retard estimé à l'aide de la méthode des moindres carrés à mouvement direct-rétrograde ; les données sont filtrées en utilisant le premier type de filtrage avec un angle d'ouverture de  $2\pi/9$ , profil MO3.

#### b) Profil obtenu en utilisant le second type de filtrage

Les expérimentations suivantes consistent à évaluer les performances des approches dérivatives, notamment celles basées sur un mouvement direct-rétrograde des ondes. Le filtre utilisé est un filtre gaussien de type ellipse modulée avec une taille de  $m \times n$ . Ce filtre est calibré avec les valeurs des paramètres suivantes :  $m = 60$  points/trace ;  $n = 60$  traces,  $\sigma_x = 8$  ;  $\sigma_y = 6$  et  $f_0 = 0.08$ .

Les tests ont été élaborés sur trois profils différents :

- Profil appelé MO1 sur lequel ont été effectués deux relevés successifs en 1993 et 1994 ;
- Profil MO2 en 1993 et en 1994 ;
- Profil MO3 toujours avec les mêmes moments de relevé.

Nous présentons les résultats correspondant à chaque profil, et cela, en approche dérivative simple et en approche dérivative basée sur un mouvement direct-rétrograde des ondes.

- **Méthode des moindres carrés simple (basée sur un mouvement direct des ondes)**

Les figures 4.17, 4.18 et 4.19 représentent le profil obtenu avec l'approche des moindres carrés simple, toujours dans le cadre du filtrage gaussien modulé de type elliptique. Ces résultats montrent que la zone de réservoir est nettement localisée. Ils sont conformes à ceux qui ont été trouvés avec les autres méthodes d'estimation du retard y compris celle de GDF. Le saut d'amplitude à partir d'une certaine profondeur (environ égale à 400) marque le passage au niveau du réservoir. Théoriquement, selon la figure 4.2, l'amplitude du retard reste constant après le passage au réservoir. Le profil des retards est de bien meilleure qualité tant au niveau numérique qu'au niveau du découpage.



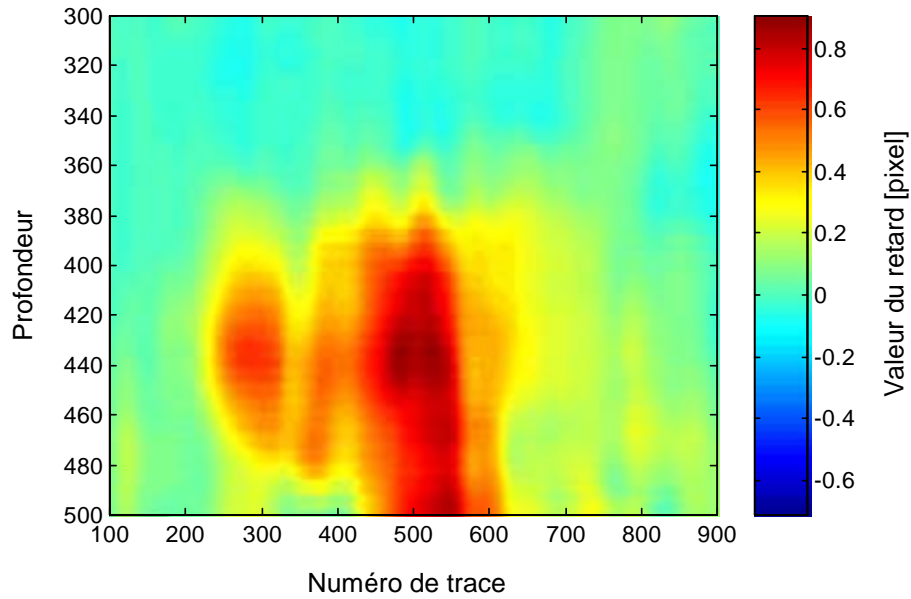


Figure 4.17 : Profil du retard estimé à l'aide de la méthode des moindres carrés, basée sur un mouvement direct des ondes ; profil MO1\_93\_94 ; les données sont filtrées en utilisant le deuxième type de filtrage.

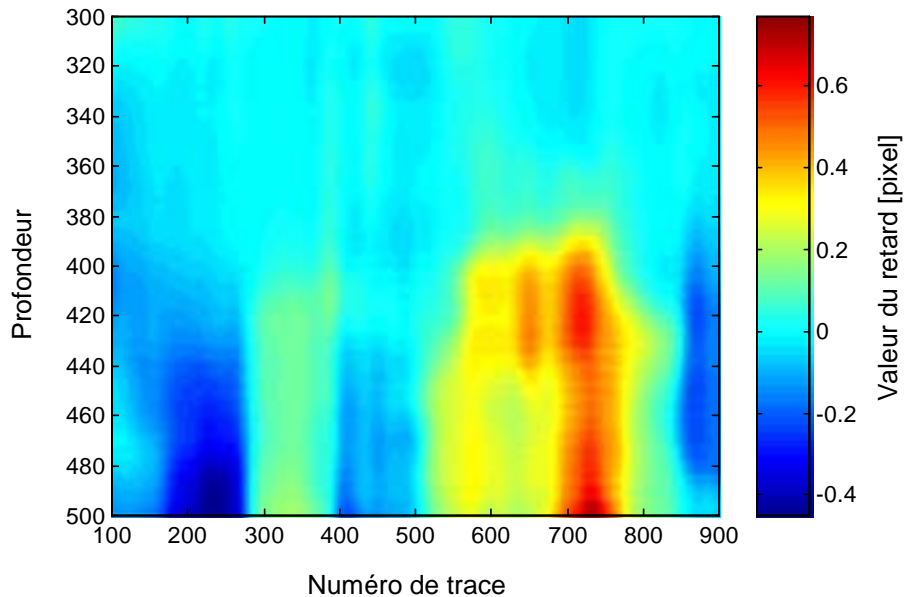


Figure 4.18 : Profil du retard estimé à l'aide de la méthode des moindres carrés, basée sur un mouvement direct des ondes ; profil MO2\_93\_94 ; les données sont filtrées en utilisant le deuxième type de filtrage.

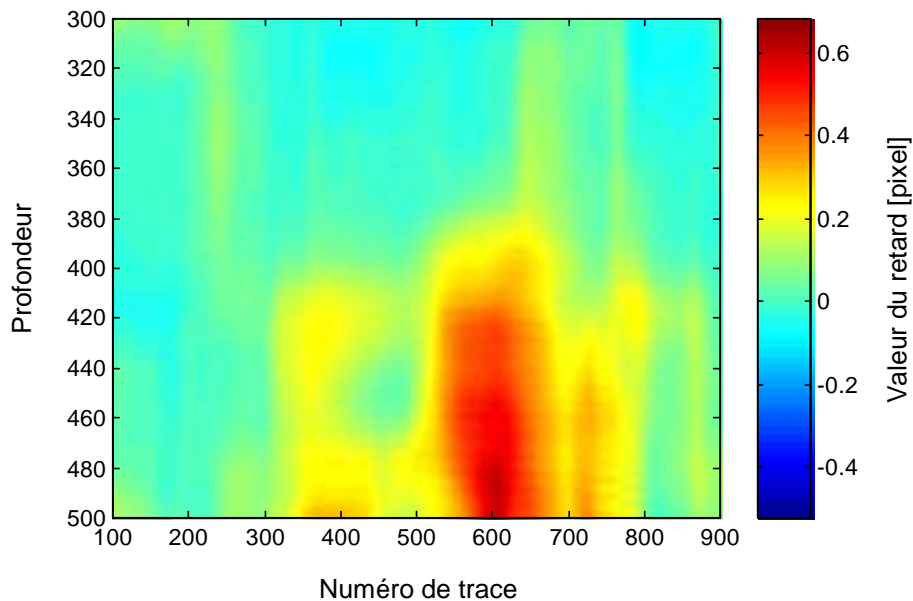


Figure 4.19 : Profil du retard estimé à l'aide de la méthode des moindres carrés, basée sur un mouvement direct des ondes ; profil MO3\_93\_94 ; les données sont filtrées en utilisant le second type de filtrage.

- ***Approche des moindres carrés basée sur un mouvement direct-rétrograde des ondes***

Les expérimentations suivantes ont trait au test de l'approche des moindres carrés basée sur un schéma direct-rétrograde et cela, avec le second type de filtrage. Les temps de retard pour les trois profils sont présentés respectivement par les figures 4.20, 4.21 et 4.22.

En comparant ces trois figures aux trois précédentes, les deux approches (simple, direct-rétrograde) donnent, sur des traces réelles, des résultats comparables quant à l'allure globale du profil à deux dimensions. Les différences apparaissent simplement au niveau de l'estimation de la valeur des retards. Nous observons une légère différence d'amplitude entre le formalisme simple et le formalisme direct-rétrograde. La valeur de l'estimation obtenue avec l'approche simple est inférieure à celle du formalisme direct-rétrograde. Nous avons également constaté cette différence lors des tests sur des séquences numériques (expérimentations dans la deuxième partie de ce manuscrit).

En outre, comme on peut le voir en comparant les figures 4.20 à 4.22, aux figures 4.17 à 4.19, l'amplitude des retards sur les régions parasites sont atténuées, et par conséquent, le formalisme direct-rétrograde permet d'obtenir des profils plus denses et plus réguliers.

Les présents résultats mettent en valeur l'application du formalisme basé mouvement direct-rétrograde dans l'estimation du retard. L'utilisation d'un tel formalisme permet d'apporter une information plus compacte et plus riche sur l'estimation du retard. Parmi les méthodes que nous avons testées, cette approche est la plus fiable en terme de valeurs quantitatives et également en terme d'aspect qualitatif.

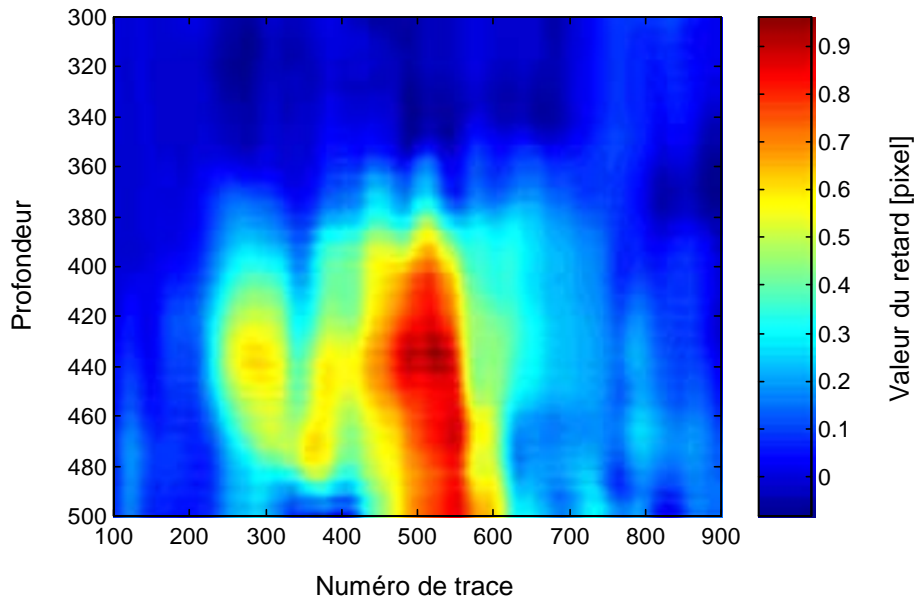


Figure 4.20 : Profil du retard estimé à l'aide de la méthode des moindres carrés, basée sur un mouvement direct-rétrograde des ondes ; profil MO1\_93\_94 ; les données sont filtrées en utilisant le second type de filtrage.

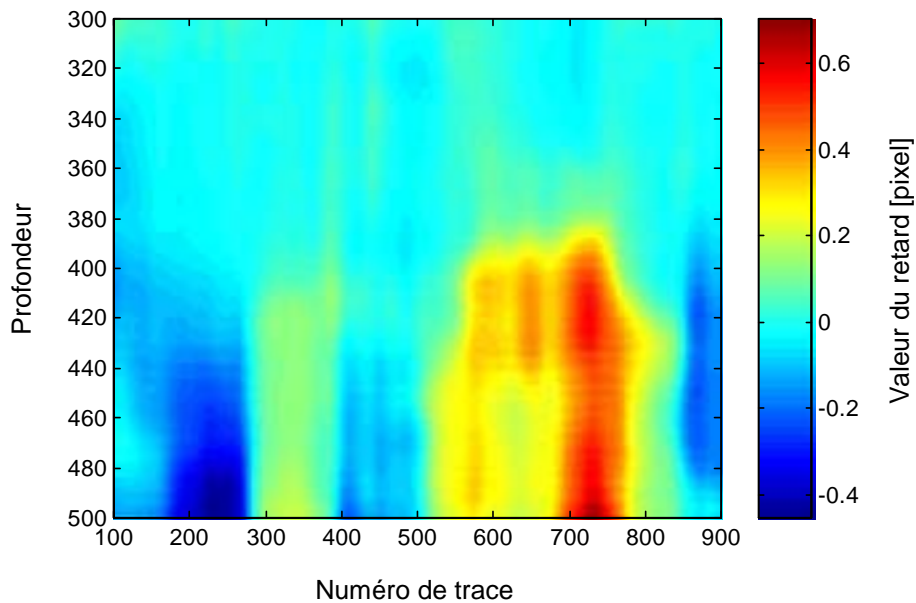


Figure 4.21 : Profil du retard estimé à l'aide de la méthode des moindres carrés, basée sur un mouvement direct-rétrograde des ondes ; profil MO2\_93\_94 ; les données sont filtrées en utilisant le second type de filtrage.

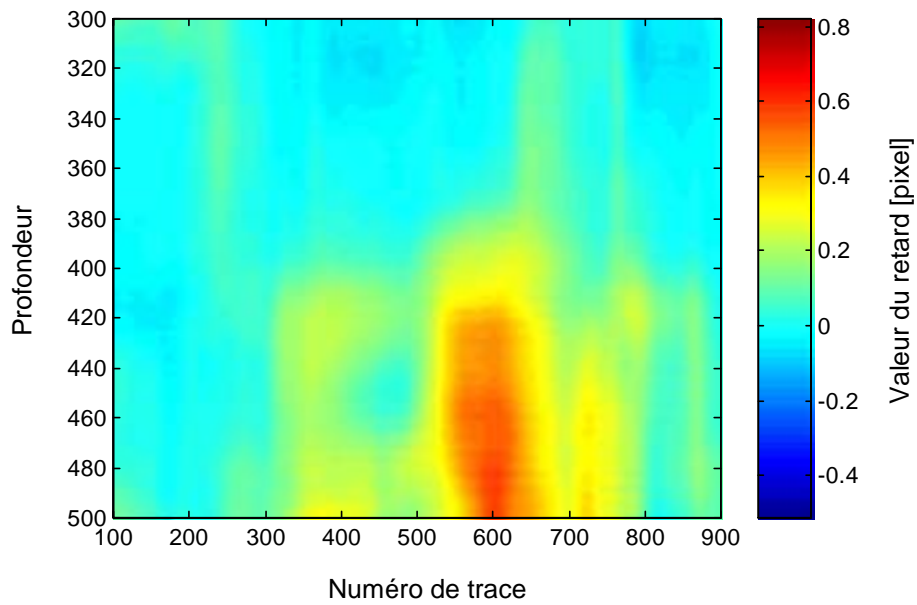


Figure 4.22 : Profil du retard estimé à l'aide de la méthode des moindres carrés, basée sur un mouvement direct-rétrograde des ondes ; profil MO3\_93\_94 ; les données sont filtrées en utilisant le second type de filtrage.



## Chapitre 2

# Segmentation spatio-temporelle en sismique répétitive

### 2.1 Introduction

Une première approche dédiée à l'estimation du retard a été développée en s'appuyant sur le schéma direct-rétrograde. Pour cette approche, le support spatial d'estimation se construit à partir du voisinage direct du pixel courant. Afin de proposer un schéma moins contraint par le caractère local du support d'estimation, nous avons développé une méthode permettant de regrouper les pixels par classe. L'objectif est de permettre une meilleure immunité aux réflexions multiples. En effet, ces réflexions multiples viennent créer des marqueurs qui ne se déplacent pas. Ils viennent donc parasiter la carte de retard.

### 2.2 Principe

Dans le cadre des méthodes différentielles à contrainte locale, la mise en équation du mouvement est effectuée sur un support d'estimation régulier. Pour le cas des données sismiques, de nombreuses perturbations peuvent subsister. C'est le cas des multiples car ils injectent localement un retard nul entre deux profils. La présente technique permet de minimiser l'influence de ce type de perturbation. Pour cela, nous mettons en place une approche axée sur la segmentation dans les domaines spatial et temporel. La technique utilisée s'inspire de la méthode des nuées dynamiques. Nous rappelons que cette technique se déroule en deux étapes :

- La première est caractérisée par une étape de représentation qui induit une métrique (exemple le barycentre d'un nuage) ;
- La seconde a pour objectif l'affectation du point considéré.

Dans notre application, nous considérons comme représentant le vecteur de mouvement. La métrique induite est alors le résidu issu de la contrainte de mouvement. La phase d'affectation d'un pixel à l'une des régions est la comparaison directe de la métrique en fonction du représentant de la classe.

La procédure définit en premier le nombre noté  $k$  de classes. Pour une image considérée  $S$ , nous avons :

$$S = \{S_i, i \in [1, k] / S_i \cap S_j = \emptyset\} \quad (391)$$

Les  $S_i$  représentent le support spatial de chacune des classes. Pour chacune d'elles, nous notons  $\vec{d}_i(dx_i, dy_i)$  le représentant qui est le vecteur de mouvement associé à la classe  $i$ . La métrique utilisée est alors définie par la relation :

$$Res_i(x, y) = dx_i \frac{\partial r(x, y, t)}{\partial x} + dy_i \frac{\partial r(x, y, t)}{\partial y} + Te \frac{\partial r(x, y, t)}{\partial t} \quad (392)$$

Cette relation caractérise le résidu (c'est-à-dire l'écart de modélisation) pour le pixel positionné en  $p$  par rapport au modèle de mouvement de la classe  $i$ .

La méthode des nuées dynamiques est une procédure itérative. Elle part d'une solution initiale pour converger vers un état stable si les données ne présentent pas de mauvais conditionnement. La procédure est décrite par le schéma suivant :

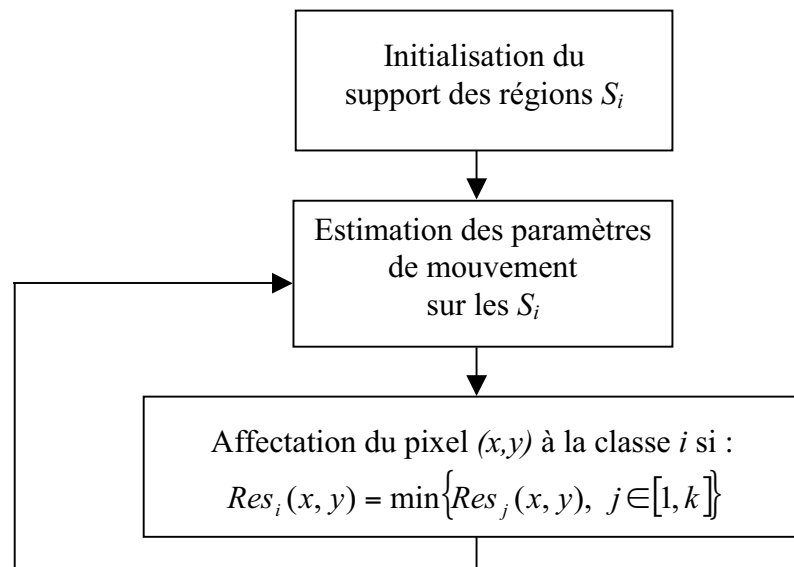


Figure 4.23 : Algorithme de classification en utilisant le principe des nuées dynamiques.

### 2.3 Résultats préliminaires

L'application de ce type de méthode donne des résultats intéressants comme le montrent les figures ci-dessous. La première correspond à une utilisation de la méthode avec 6 classes. La seconde correspond au même paramétrage, mais sur des données filtrées. Nous observons que l'effet du filtrage est de régulariser la répartition spatiale des classes. Les résultats semblent intéressants (figure 4.24), leur exploitation directe met en évidence un problème de stabilité. En effet, une analyse des résultats montre que certaines des classes ont tendance à se vider progressivement et à proposer un support engendrant une estimation instable.

Comme on peut le voir sur la figure 4.25, les poids associés au calcul sont aussi intéressants.

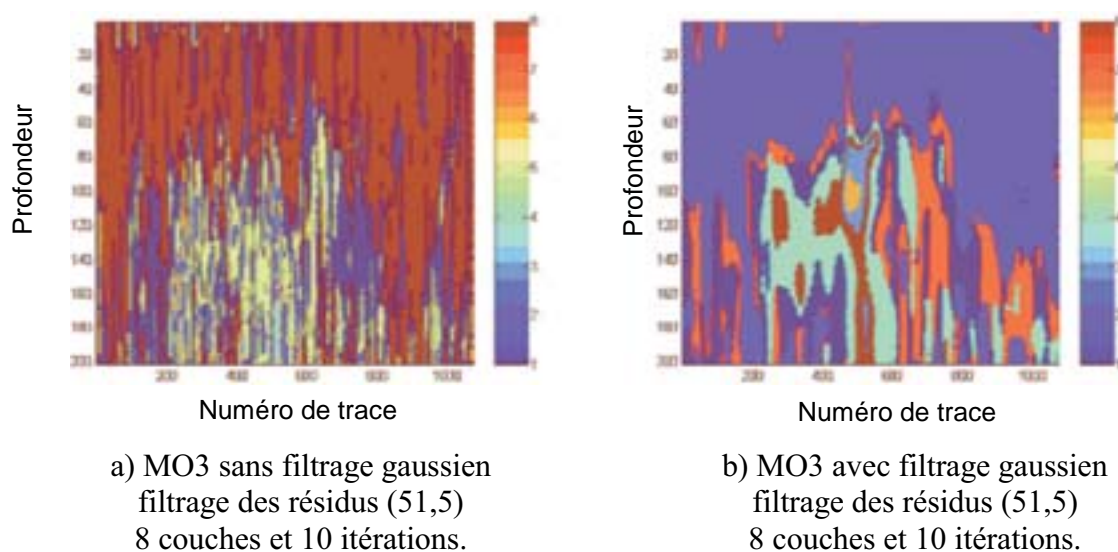


Figure 4.24 : Résultats de la segmentation spatio-temporelle en utilisant le principe des nuées dynamiques.

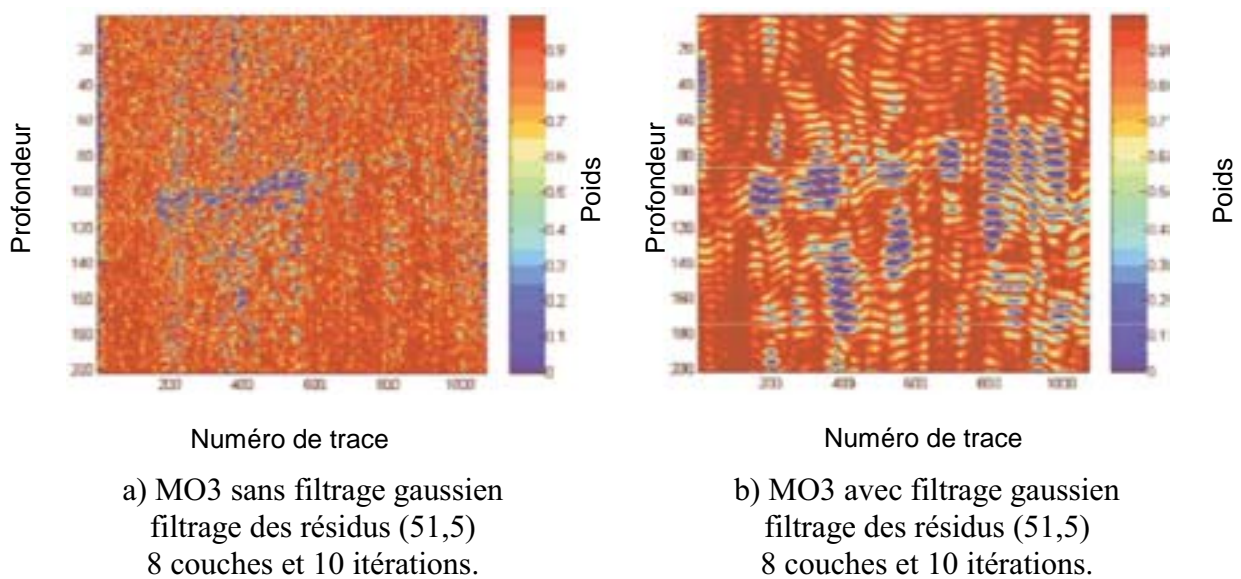


Figure 4.25 : Représentation des matrices de poids associés au calcul des déplacements.

## 2.4 Développement d'une variante

L'extension proposée à partir de la méthode de segmentation initiale suit un processus de blocage des régions à partir d'une estimation. En effet, nous avons remarqué que, si nous ne réalisons pas la phase de filtrage gaussien, les résultats sont très perturbés, mais ne présentent pas d'instabilité dans l'estimation du mouvement. La zone de réservoir est alors peu visible. Par contre, si nous appliquons un filtre gaussien, la zone spatiale est beaucoup plus stable mais avec des zones d'instabilité dans l'estimation du mouvement (Ce phénomène est explicité lors du paragraphe 9.3). Suite à ces observations, nous proposons une variante de la méthode qui permet d'obtenir un compromis stabilité/précision. La stratégie est la suivante:

\**Étape n°1* : Filtrage gaussien global

\**Étape n°2* : Première segmentation à l'aide des nuées dynamiques. Nous introduisons un filtrage important (uniforme ou gaussien) des résidus dans l'étape d'affectation. La caractérisation de ces filtres est définie par le couple (15 traces, 63 points par trace) par défaut. Une première segmentation favorisant la cohérence spatiale du réservoir est obtenue.

\**Étape n°3* : Une seconde segmentation à l'aide des nuées dynamiques. Cette segmentation n'est réalisée que sur les couches, à l'exclusion de la couche du fond, issue de l'étape n°2. Pour plus de cohérence avec l'application, nous prolongeons vers le bas les couches de l'étape n°2. Nous introduisons ensuite un filtrage (uniforme ou gaussien) des résidus dans l'étape d'affectation. La caractérisation de ces filtres est définie par le couple (9 traces, 33 points par trace) par défaut. Une seconde segmentation favorisant l'estimation du retard est obtenue.

Le processus permettant le passage de l'étape n°2 à l'étape n°3 est conduit par une phase de décision sur la stabilité de l'estimation du module de retard pour chacune des couches. Pour chacune, la procédure est la suivante :

- Calcul de deux valeurs moyennes associées à l'estimation du retard  $m_1 = \text{mean}(d_{n,n-4})$  et  $m_2 = \text{mean}(d_{n-2,n-6})$  où  $d$  est le module du retard et  $n$  l'indice associé aux itérations.



- Contrôle d'un seuil de voisinage entre  $m_1$  et  $m_2$  : si la valeur notée  $Tol$  est inférieure à un seuil alors la stabilité est effective sinon nous poursuivons les itérations en restant dans l'étape n°2, avec  $Tol = |m_2 - m_1|$ .

Le passage à l'étape n°3 n'est réalisé que si nous avons une stabilité pour toutes les couches.

Comme indiqué dans le paragraphe précédent, nous avons remarqué que les couches correspondant à une zone présentant un mauvais conditionnement a tendance à se vider. Dans ce contexte, nous avons appliqué un processus d'élimination de la couche qui devient vide. Comme nous pouvons le voir sur les graphes présentés dans la figure ci-dessous, une couche qui se vide va générer une valeur de retard de plus en plus fausse. Dans ce cas, lorsque la couche est vide, nous décrétons le nombre total de couches de 1.

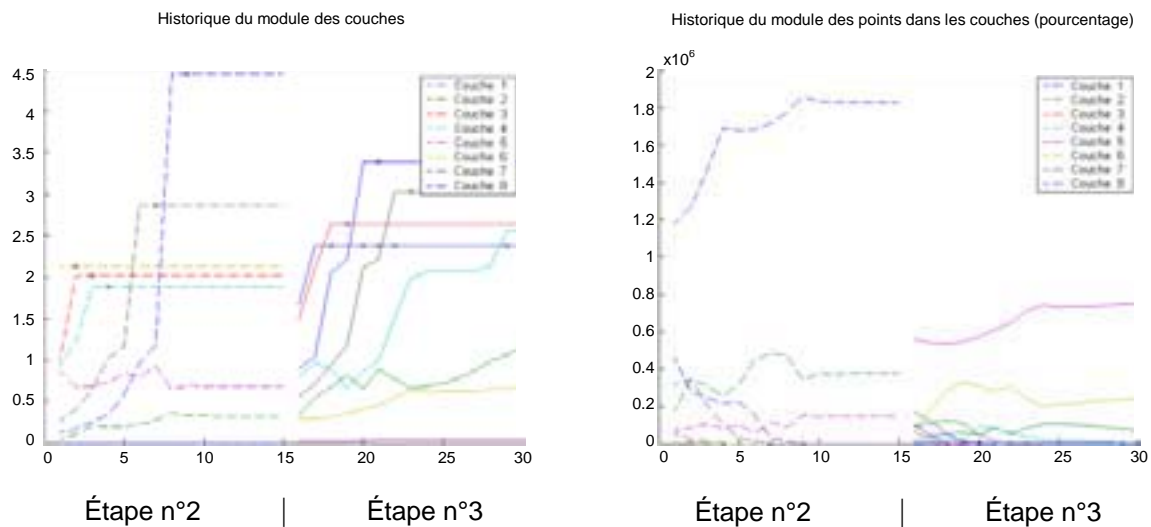


Figure 4.26 : Processus du blocage des couches par décision sur la stabilité de la couverture spatiale de chacune des couches.

Pour le MO2, les résultats que nous obtenons, sont regroupés dans la figure suivante:

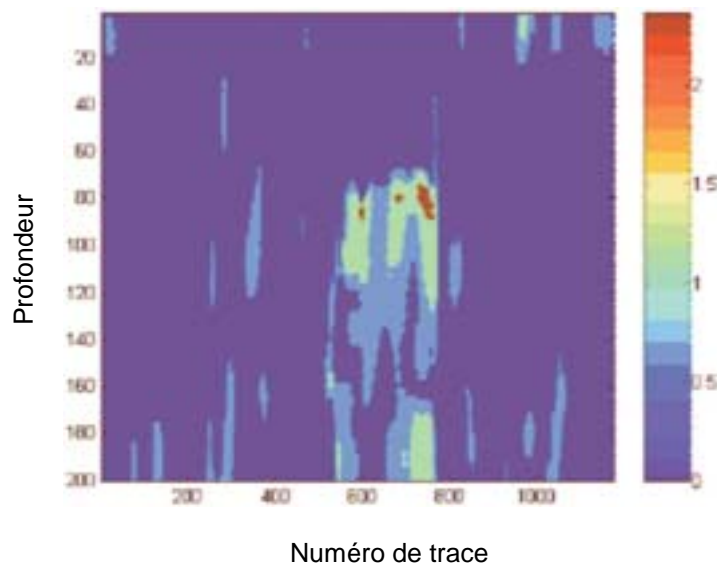


Figure 4.27 : Segmentation spatio-temporelle pour le profil MO2.

## Résumé de la quatrième partie

L'étude décrite dans cette partie consiste en une approche expérimentale, qui vise à améliorer l'estimation du retard associé à la sismique répétitive afin de satisfaire aux besoins de la caractérisation de gaz naturel. Dans les quatre premiers paragraphes, nous avons présenté la nature du problème, l'observation des données fournies et la modélisation du signal sismique. Dans le cinquième paragraphe, un survol des différentes méthodes d'estimation de retard a été présenté. Nous avons examiné entre autres les méthodes corrélatives, méthodes basées sur le bispectre, méthode de l'intercorrélation généralisée, méthode ROTH, méthode PHAT et la méthode de l'estimateur du maximum de vraisemblance.

Le sixième paragraphe décrit la méthode utilisée par la société GDF. Cette méthode utilise le domaine spectral, qui s'appuie sur l'équation de l'interspectre.

Dans les septième et huitième paragraphes, nous avons présenté des méthodes dérivatives qui font l'objet de notre étude. L'analyse a été divisée en deux approches : approche 1D et approche 2D. Dans chacune de ces approches, le formalisme basé sur le mouvement direct-rétrograde des ondes a été pris en compte. Des expérimentations nous ont permis d'évaluer les performances de ces approches.

Les méthodes dérivatives ont apporté une solution aux principaux problèmes liés à l'estimation du temps de retard. Les résultats des expérimentations sont satisfaisants du point de vue de la détection du réservoir et de l'estimation du retard en sismique. Le formalisme basé schéma direct-rétrograde permet d'obtenir des résultats cohérents par rapport à la connaissance a priori du site.

Dans le dernier paragraphe, nous avons présenté l'application de l'approche dérivative à la segmentation spatio-temporelle. La technique utilisée s'inspire de la méthode des nuées dynamiques. On s'est attaché à supprimer toutes les perturbations qui peuvent subsister dans le cadre des données sismiques. Les résultats expérimentaux ont illustré l'efficacité du schéma de segmentation proposé. Nous avons montré également l'intérêt du filtre gaussien modulé dans un schéma de segmentation.



## **Conclusion générale**



## 1 Résumé de l'accomplissement

Ce chapitre conclut la dissertation en résumant les principaux développements et accomplissements de ce travail.

Dans cette dissertation, une alternative à l'estimation du mouvement associée aux objets contenus dans une séquence d'images a été proposée. Les objets sont définis comme les entités qui sont spatialement et temporellement cohérentes dans le temps. L'analyse du mouvement s'appuie sur la transformation des entités entre deux trames successives. La description du mouvement en terme de transformation instantanée infinitésimale, avec l'hypothèse de conservation de la luminance, conduit à une formulation au sens des moindres carrés de l'estimation du mouvement.

Un mouvement d'un objet peut être modélisé par plusieurs types de modèles. Dans ce contexte, pour être concret, nous avons présenté dans le chapitre 1 de la première partie, quelques principes de la description du mouvement. Dans ce même chapitre, un état de l'art sur les différentes méthodes d'estimation du mouvement a été abordé. Notre étude s'est limitée au contexte 2D. Les notions d'estimation et de segmentation sont inséparables. Voilà pourquoi nous avons présenté dans le chapitre 2 un survol des principes de la segmentation.

Après avoir analysé les lacunes des méthodes existantes, nous avons exploré et présenté quelques solutions au problème d'estimation du flot optique. Cela fait l'objet de la deuxième partie. Le principe est de combiner à la fois les informations dans les sens direct et rétrograde. Nous avons baptisé cette approche : "formalisme direct-rétrograde" ou "approche basée sur un schéma direct-rétrograde". Nous avons présenté plusieurs méthodes exploitant ce formalisme :

Pour la première, il s'agit d'une extension de l'équation de contrainte du mouvement : notion de mouvement direct-rétrograde. L'approche développée dans ce contexte est limitée à un modèle de translation. Nous rappelons brièvement que l'équation de contrainte de mouvement est établie à partir de l'hypothèse de conservation de la luminance. Ici, nous exploitons cette hypothèse dans les deux sens temporels. Cela se traduit par une équation de contrainte, exploitant à la fois un déplacement direct et un déplacement rétrograde. La mise en équation d'un ensemble de points conduit à former un système d'équations. Sa résolution est ramenée à un problème des moindres carrés. Il s'agit d'une méthode d'estimation locale dans laquelle un support spatial est nécessaire. Dans ce support, le mouvement de l'ensemble des pixels est supposé uniforme. La résolution fournit donc un vecteur de translation, commun à tous les pixels du support. Les tests ont été effectués sur des séquences de synthèse et des séquences réelles. Contrairement à l'estimation basée sur un schéma simple, les tests de l'approche basée sur un schéma direct-rétrograde sur une séquence synthétique ont montré que ce formalisme était moins sensible au bruit. De même, le schéma proposé présente de meilleures performances pour les séquences réelles. L'avantage du schéma basé mouvement direct-rétrograde est que la prise en compte des gradients spatiaux sur les deux images permet à la méthode d'estimation d'être moins sensible aux fluctuations temporelles aléatoires de mouvement. Les résultats obtenus sont donc plus précis.

Le second scénario consiste simplement à exploiter une méthode voisine de l'approche précédente. Il s'agit de la méthode des moindres carrés totaux basée sur un schéma direct-rétrograde (méthode du tenseur basée sur un schéma direct-rétrograde). Nous rappelons que les méthodes des moindres carrés totaux en général présentent des avantages par rapport à la méthode des moindres carrés classiques. Elles donnent des résultats adéquats, sauf qu'au niveau de la convergence, elle est très critique. Dans notre cas, nous avons associé la méthode des moindres carrés totaux au formalisme basé mouvement direct-rétrograde.

En revanche, jusqu'ici, les approches proposées (moindres carrés classiques et moindres carrés totaux) ne permettent pas de résoudre complètement les carences de l'estimation sur des données erronées (par exemple une image présentant des discontinuités de mouvement).

Voilà pourquoi nous avons étudié l'extension de l'approche sur un schéma robuste. Le troisième scénario s'appuie sur une méthode robuste basée sur le schéma direct-rétrograde. Le but est de rejeter au mieux les points erronés en pondérant les données selon la fonction M-estimateur. L'intérêt du formalisme basé schéma direct-rétrograde est ici de pouvoir pondérer les données des deux sens temporels (ascendant et descendant). En général, avec les méthodes robustes, les pixels appartenant à un mouvement dominant dans une région d'intérêt sont sélectionnés pour obtenir une estimation non biaisée.

Le problème avec les méthodes citées est que l'estimation devient difficile lorsqu'on augmente l'amplitude du déplacement. Pour cela, nous avons introduit la méthode incrémentale. Le quatrième scénario concerne donc l'approche incrémentale basée sur un schéma direct-rétrograde. Nous rappelons que l'avantage de l'approche incrémentale est de pouvoir atteindre une estimation avec une précision accrue. Il s'agit d'une méthode d'optimisation récursive qui permet, à chaque itération, de corriger la valeur de déplacement à l'aide d'un terme correcteur. Le terme correcteur régit un critère à minimiser. Ce critère est différent selon qu'il s'agisse d'une méthode incrémentale classique ou d'une méthode incrémentale basée sur un schéma direct-rétrograde. Dans notre cas, nous nous sommes intéressés à l'exploitation de l'approche incrémentale avec le formalisme "direct-rétrograde". Le point clé de la technique proposée est qu'on peut gérer la correction du déplacement dans les deux sens. De nombreuses expérimentations ont été faites sur l'approche proposée. Les résultats sont satisfaisants.

Jusqu'ici, nos séries de tests se sont limitées à un contexte de mouvement de translation. Il était aussi intéressant de voir comment le formalisme proposé se comporte dans le cas du modèle de mouvement plus complexe. C'est pourquoi nous avons dressé un dernier scénario qui consiste en une extension de l'approche basée sur un schéma direct-rétrograde au modèle paramétrique. C'est une méthode plus générale qui a consisté à estimer les paramètres du mouvement. Le mouvement est modélisé par une transformation affine. Cette modélisation a été effectuée dans les deux sens (direct et rétrograde). Cela aboutit à une équation plus ou moins complexe, dans laquelle sont apparus deux aspects importants : les aspects linéaire et non linéaire. Des études issues de ces deux aspects ont été effectuées. Les tests ont été réalisés sur des séquences synthétiques et réelles. Les résultats obtenus ont permis de déduire la nécessité de linéarisation du modèle et/ou la nécessité de prise en compte du modèle non linéaire. Comparés aux résultats de la méthode paramétrique basée sur un schéma simple, les deux aspects précédents qui sont tous inscrits dans le cadre d'approche paramétrique basée sur un schéma direct-rétrograde, fournissent de meilleurs résultats.

Une extension de l'approche basée sur un schéma direct-rétrograde à une séquence de  $n$  images est proposée. C'est une généralisation du formalisme selon l'axe temporel. Cette proposition a apporté une solution permettant de régulariser le champ de déplacement.

L'intérêt de la prise en compte des informations directe et rétrograde est nettement illustré. Nous avons utilisé ainsi plusieurs séquences, synthétiques ou réelles, pour valider les méthodes développées. Pour chacun de ces estimateurs, une comparaison entre l'approche basée schéma simple et l'approche basée sur un schéma direct-rétrograde a été faite. Pour motiver la nécessité des techniques robustes, nous avons testé la méthode des moindres carrés pondérés sur une image présentant des discontinuités de mouvement, et cela toujours dans le cadre du formalisme direct-rétrograde. Les résultats d'expérimentation ont permis de tester la validité du formalisme proposé s'appuyant sur une double pondération.

Il en est de même pour la méthode du tenseur (moindres carrés totaux), la prise en compte des informations dans les deux sens a permis d'obtenir une meilleure précision. Nous précisons que la méthode du tenseur est sensible au niveau des bruits et voilà pourquoi sa convergence est très critique. En revanche, sa convergence est souvent accompagnée d'une meilleure performance en terme de biais.

Quant au mouvement sous forme paramétrique, la modélisation du formalisme direct-rétrograde dans ce contexte s'avère intéressante. En effet, la mise en équation du formalisme fait apparaître deux aspects différents : aspect linéaire et aspect non linéaire.

Si on se limite à l'aspect linéaire, c'est-à-dire sans la prise en compte du terme non linéaire, la formulation direct-rétrograde est ramenée à une équation dans laquelle les mouvements direct et rétrograde sont égaux mais de signe opposé. En d'autres termes, le mouvement rétrograde est lissé pour être ramené au mouvement direct. Cet aspect est efficace pour l'estimation du mouvement dont les paramètres de translation sont dominants.

Quant à l'aspect non linéaire, c'est dans ce contexte que se voit l'intérêt de l'approche basée sur un schéma direct-rétrograde. La mise en équation du mouvement a permis de traiter mutuellement les deux informations, avec leur modèle approprié. Les deux mouvements ne sont pas uniquement de signe opposé, mais l'un correspond à la transformée inverse de l'autre. Cet aspect est intéressant dans le cadre d'estimation des paramètres de rotation et divergence dominants. Des expérimentations nous ont permis de dresser le domaine d'utilisation de chacune de ces approches. Dans tous les cas, les résultats obtenus ont permis d'avoir une meilleure performance par rapport à ceux de l'approche basée sur un schéma simple. Nous confirmons donc la validité du formalisme basé flot direct-rétrograde dans le cas du modèle paramétrique.

Une conclusion intéressante peut également être faite sur l'approche incrémentale. Cette technique est la plus adaptée à l'estimation du mouvement, notamment pour les déplacements de forte amplitude. L'approche incrémentale basée sur un schéma direct-rétrograde donne une meilleure précision. Cette technique est très conseillée pour une meilleure estimation, mais le fait qu'elle soit itérative pénalise le temps de calcul.

La troisième partie est consacrée à l'application du "formalisme direct-rétrograde" à la segmentation spatio-temporelle. Dans ce contexte, deux familles de méthodes sont présentées. La première dérive de la segmentation markovienne tandis que la seconde est fondée sur la méthode des nuées dynamiques, de type K-moyens. Nous avons étudié séparément l'extension de ces méthodes, en utilisant l'estimation du mouvement. Les résultats obtenus montrent que nous avons une bonne extraction en terme d'objets. En revanche, la définition des objets dépend de l'initialisation des régions. Une meilleure performance est obtenue quand nous faisons intervenir la méthode d'estimation du mouvement dans la génération de l'initialisation des régions. Ces régions sont itérativement fusionnées pour déterminer les objets formant la scène. La prise de décision pour la procédure de division-fusion est aussi guidée par le champ de déplacement. En outre, l'apport du formalisme "direct-rétrograde" dans ces familles de segmentation est nettement démontré. Les résultats de simulation ont montré les performances des méthodes testées. En particulier, les méthodes proposées ont permis une bonne restitution des objets.

Dans la quatrième partie, une application réelle a été présentée. Cette application a pour contexte le traitement de données sismiques en liaison avec la société GDF.

Différentes approches ont été testées afin d'estimer un décalage temporel entre deux observations sismiques. Ce décalage est l'image du gaz dans la cavité naturelle. La première famille de méthodes regroupe les techniques s'appuyant sur l'intercorrélacion simple ou généralisée. La seconde famille se fonde sur les approches développées dans ce manuscrit. Notre étude prend ici tout son intérêt car l'application s'appuie uniquement sur deux relevés. Le premier est une mesure géographique sans présence de gaz et l'autre une mesure géographique après injection de gaz. Du fait du coût élevé d'une campagne de mesures, la méthode d'estimation ne peut utiliser que deux jeux de données. Dans ce contexte, l'approche basée sur le formalisme direct-rétrograde s'adapte donc naturellement au moyen d'observation mis à notre disposition. La technique proposée a permis de mieux localiser la zone de réservoir et offre une estimation plus adaptée à la réalisation sismique.



## **2 Extensions possibles**

Le travail présenté dans ce manuscrit est prolongé par les travaux au sein de l'équipe à travers les thèses de David Izquierdo et Javier Becerra. Ces travaux concernent sur le plan méthodologique :

- Une extension multi-échelle fondée sur une pyramide gaussienne ;
- Une stratégie de segmentation basée sur une approche Expectation-Maximisation (EM) ;
- L'utilisation d'une procédure récursive pour minimiser la fonctionnelle.

En outre, nous travaillons sur des applications utilisant dans la segmentation spatio-temporelle l'approche directe-rétrograde. Ces applications sont les suivantes :

- Vidéosurveillance du trafic routier ;
- Poursuite de voisinage.

# **Annexe**



# Annexe A

## Quelques notions mathématiques utiles en estimation du mouvement

### A.1 Géométrie projective

La figure A.1 illustre la relation entre les coordonnées 2D et 3D.

Soit  $P(X, Y, Z)$  un point de l'objet dont et  $p(x, y)$  sa projection sur le plan focal. Selon le principe de l'optique géométrique, les relations algébriques qui décrivent la géométrie projective sont définies par :

$$\frac{x}{f} = -\frac{X}{Z - f} \quad (\text{A1})$$

$$\frac{y}{f} = -\frac{Y}{Z - f} \quad (\text{A2})$$

Soient :

$$x = -\frac{fX}{f - Z} \quad (\text{A3})$$

$$y = -\frac{fY}{f - Z} \quad (\text{A4})$$

La quantité  $f$  désigne la distance focale (distance entre le centre de projection et le plan focal).

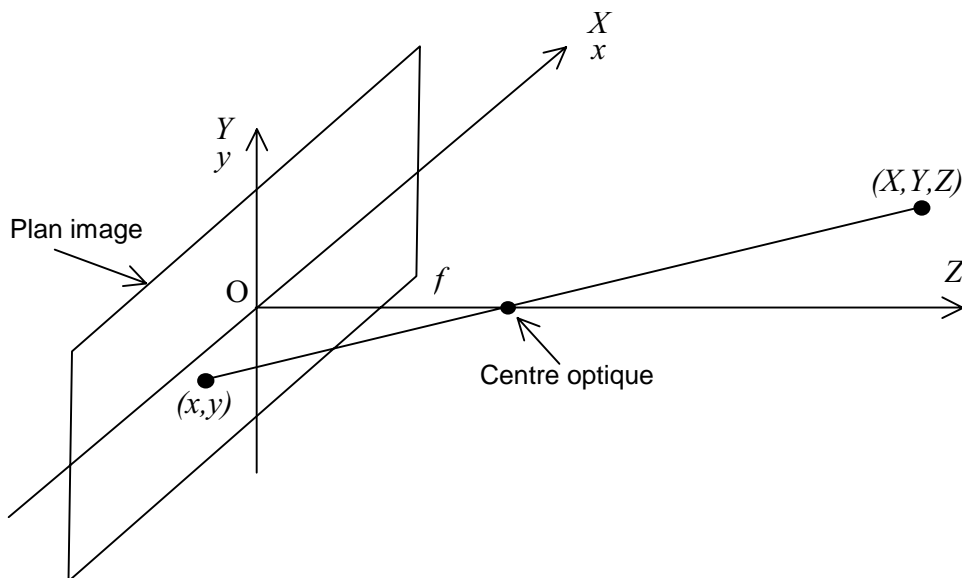


Figure A.1 : Modèle de géométrie projective.

Dans le cas où  $Z \gg f$  avec une inversion de signe, due à l'orientation de l'image étant la même que l'objet, nous avons les relations simplifiées suivantes :

$$x = \frac{fX}{Z} \quad (\text{A5})$$

$$y = \frac{fY}{Z} \quad (\text{A6})$$

Un modèle simplifié dont le centre de projection se confond avec l'origine  $O$  justifie les équations (A5, A6). La figure A.2 représente ce modèle.

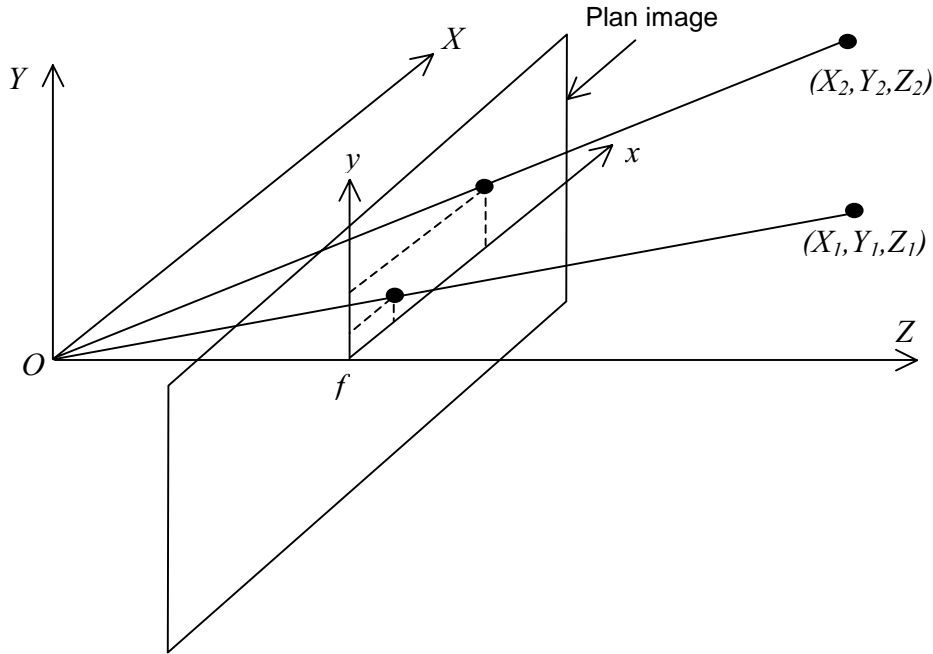


Figure A.2 : Modèle simplifié de géométrie projective.

La géométrie projective est une transformation non linéaire dans le système de coordonnées cartésien parce que l'opération dépend de  $Z$ . En revanche, elle peut être exprimée comme une correspondance linéaire dans le système de coordonnées homogènes, soit :

$$\begin{bmatrix} lx \\ ly \\ l \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} kX \\ kY \\ kZ \\ k \end{bmatrix} \quad (\text{A7})$$

## A.2 Angle d'Euler

Soient  $\theta_x$ ,  $\theta_y$  et  $\theta_z$  les angles de rotation élémentaires individuellement autour de l'axe  $x$ ,  $y$  et  $z$ . Les matrices de rotation correspondantes sont données par :

$$R_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta_x & -\sin \theta_x \\ 0 & \sin \theta_x & \cos \theta_x \end{bmatrix} \quad (\text{A8})$$

$$R_y = \begin{bmatrix} \cos \theta_y & 0 & \sin \theta_y \\ 0 & 1 & 0 \\ -\sin \theta_y & 0 & \cos \theta_y \end{bmatrix} \quad (\text{A9})$$

et :

$$R_z = \begin{bmatrix} \cos \theta_z & -\sin \theta_z & 0 \\ \sin \theta_z & \cos \theta_z & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (\text{A10})$$

Si l'angle de rotation est faible, les approximations suivantes peuvent se faire :

$$\begin{cases} \cos \theta \approx 1 \\ \sin \theta \approx \theta \end{cases} \quad (\text{A11})$$

Les expressions simplifiées des trois matrices précédentes sont ainsi données par :

$$R_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -\theta_x \\ 0 & \theta_x & 1 \end{bmatrix} \quad (\text{A12})$$

$$R_y = \begin{bmatrix} 1 & 0 & \theta_y \\ 0 & 1 & 0 \\ -\theta_y & 0 & 1 \end{bmatrix} \quad (\text{A13})$$

et :

$$R_z = \begin{bmatrix} 1 & -\theta_z & 0 \\ \theta_z & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (\text{A14})$$

La matrice de rotation globale est obtenue par le produit des trois matrices élémentaires, soit :

$$R = \begin{bmatrix} 1 & -\theta_z & 0 \\ \theta_z & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & \theta_y \\ 0 & 1 & 0 \\ -\theta_y & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -\theta_x \\ 0 & \theta_x & 1 \end{bmatrix} = \begin{bmatrix} 1 & -\theta_z & \theta_y \\ \theta_z & 1 & -\theta_x \\ -\theta_y & \theta_x & 1 \end{bmatrix} \quad (\text{A15})$$

La multiplication des matrices n'est pas en général commutative. Ici, l'ordre de multiplication ne change pas le résultat en utilisant l'hypothèse de faible rotation et en négligeant les termes du second ordre.

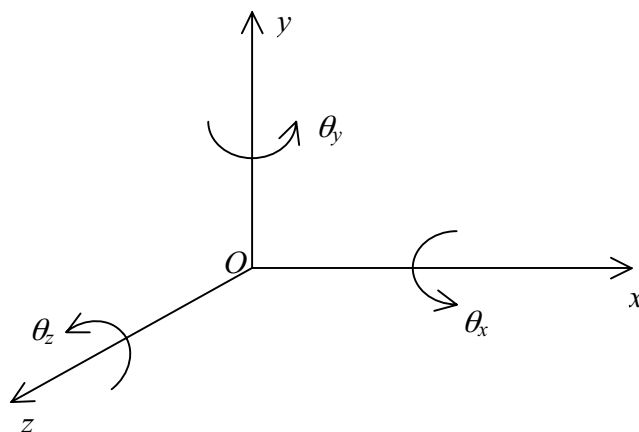


Figure A.3 : Angles d'Euler.

### A.3 Application de la transformée de Fourier en estimation du mouvement

#### a) Théorème de translation

Soit un signal  $g(x_1, x_2, x_3)$  ayant comme transformée  $G_0(f_1, f_2, f_3)$ . Si les variables  $x_1, x_2$  et  $x_3$  subissent respectivement une translation de  $\Delta x_1, \Delta x_2$  et  $\Delta x_3$ , la transformée de Fourier de  $g(x_1 - \Delta x_1, x_2 - \Delta x_2, x_3 - \Delta x_3)$  s'écrit :

$$G(f_1, f_2, f_3) = G_0(f_1, f_2, f_3) e^{-j2\pi(f_1\Delta x_1 + f_2\Delta x_2 + f_3\Delta x_3)} \quad (\text{A16})$$

#### b) Application à l'estimation du mouvement

Considérons, dans un espace spatio-temporel, une séquence d'images définie par  $I(x, y, t)$ , qui est formée à partir d'une translation d'une image 2D  $I_0(x, y)$  avec une vitesse constante  $\vec{V}(u, v)$ . La fonction de luminance suivant l'évolution temporelle peut s'écrire selon la relation suivante :

$$I(x, y, t) = I_0(x - ut, y - vt) \quad (\text{A17})$$

A l'instant  $t = 0$ , l'intensité lumineuse est égale à  $I_0(x, y)$ . La transformée 3D de  $I(x, y, t)$  est donnée par :

$$G(f_x, f_y, f_t) = \int \int \int_{-\infty}^{+\infty} I(x, y, t) e^{-j2\pi(f_x x + f_y y + f_t t)} dx dy dt \quad (\text{A18})$$

En remplaçant la valeur de  $I(x, y, t)$  par  $I_0(x - ut, y - vt)$ , nous avons :

$$G(f_x, f_y, f_t) = \int \left[ \int \int_{-\infty}^{+\infty} I_0(x - ut, y - vt) e^{-j2\pi(f_x x + f_y y)} dx dy \right] e^{-j2\pi f_t t} dt \quad (\text{A19})$$

En appliquant le théorème de translation au terme entre crochets (il s'agit ici d'une translation 2D), l'expression précédente peut se mettre sous la forme :

$$G(f_x, f_y, f_t) = \int_{-\infty}^{+\infty} G_0(f_x, f_y) e^{-j2\pi(f_x u + f_y v + f_t)} e^{-j2\pi f_t t} dt \quad (\text{A20})$$

Soit :

$$G(f_x, f_y, f_t) = G_0(f_x, f_y) \int_{-\infty}^{+\infty} e^{-j2\pi(f_x u + f_y v + f_t)} dt \quad (\text{A21})$$

avec  $G_0(f_x, f_y)$  désigne la transformée de Fourier de  $I_0(x, y)$  qui n'est que la valeur de  $I_0(x - ut, y - vt)$  à l'instant  $t = 0$ .

En introduisant la définition de la fonction de Dirac, nous avons finalement :

$$G(f_x, f_y, f_t) = G_0(f_x, f_y) \delta(f_x u + f_y v + f_t) \quad (\text{A22})$$

### A.4 Analyse des caractéristiques fréquentielles d'une image

#### a) Analyse par la transformée de Fourier

L'analyse de Fourier est une technique mathématique qui permet de séparer une image en ses différentes composantes fréquentielles (il s'agit ici de la fréquence spatiale). Le théorème de Fourier stipule que toute fonction continue  $f(x)$  peut être représentée par la somme d'une série de termes sinusoïdaux de fréquences spatiales variables. Ces termes peuvent être obtenus par la transformation de Fourier de  $f(x)$ , qui s'exprime par :

$$F(f) = \int_{-\infty}^{+\infty} f(x) e^{-2\pi i f x} dx \quad (\text{A23})$$

$f$  est la fréquence spatiale.  $F(f)$  est une fonction dans le domaine de fréquence. La fonction du domaine spatial  $f(x)$  peut être retrouvée à partir de  $F(u)$  par la transformée de Fourier inverse :

$$f(x) = \int_{-\infty}^{+\infty} F(f) e^{2\pi i f x} df \quad (\text{A24})$$

Pour utiliser l'analyse de Fourier en traitement numérique des images, on doit considérer deux extensions des équations (A23) et (A24). Les deux transformées peuvent d'abord être généralisées par un passage à des fonctions à deux dimensions  $f(x, y)$  et  $F(f_x, f_y)$  :

$$F(f_x, f_y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) e^{-2\pi i (f_x x + f_y y)} dx dy \quad (\text{A25})$$

On peut aussi généraliser ces deux transformations par un passage à des fonctions discrètes. La transformée de Fourier discrète bidimensionnelle s'exprime par la relation :

$$F(f_x, f_y) = \frac{1}{MN} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} f(x, y) e^{-2\pi i \left( \frac{f_x x}{N} + \frac{f_y y}{M} \right)} \quad (\text{A26})$$

La fonction  $F(f_x, f_y)$  contient l'information de fréquence spatiale de l'image originale  $f(x, y)$  et elle est appelée spectre de fréquence. Il est à noter que  $F(f_x, f_y)$  est une fonction complexe, car elle contient le terme  $i$  dont la valeur au carré est égale à  $-1$ . On peut écrire toute fonction complexe comme la somme d'une partie réelle et d'une partie imaginaire :

$$F(f_x, f_y) = R(f_x, f_y) + i \text{Im}(f_x, f_y) \quad (\text{A27})$$

En introduisant la notation cartésienne à la relation (A27), nous avons :

$$F(f_x, f_y) = |F(f_x, f_y)| e^{i\Phi(f_x, f_y)} \quad (\text{A28})$$

L'amplitude de la transformée de Fourier est définie par :

$$|F(f_x, f_y)| = \sqrt{F(f_x, f_y)^2 + \text{Im}(f_x, f_y)^2} \quad (\text{A29})$$

Elle peut être affichée comme une image bidirectionnelle. La transformée de Fourier représente l'amplitude et la direction des différentes composantes fréquentielles de l'image  $f(x, y)$ . Trois types de paramètres peuvent être définis à partir du spectre de fréquence. Le premier ensemble est dérivé de la distribution radiale des valeurs de  $|F(f_x, f_y)|$  et est sensible à la grossièreté de la texture de l'image. Une texture grossière aura de grandes valeurs de  $|F(f_x, f_y)|$  concentrées autour de l'origine, alors qu'avec une texture fine ces valeurs sont étalées au-delà de l'origine [WES76]. Les valeurs moyennes de  $|F(f_x, f_y)|$  prises sur des régions en forme d'anneaux concentrés autour de l'origine produisent une mesure de la grossièreté. De même, la distribution angulaire des valeurs de  $|F(f_x, f_y)|$  est sensible à l'orientation de la texture. Par conséquent, les anneaux peuvent être caractérisés par un ensemble de moyennes de  $|F(f_x, f_y)|$  prises sur des régions en forme de cale autour de l'origine. Un deuxième ensemble de caractéristiques, qui sont aussi basées sur le spectre de fréquence, consiste en quatre mesures statistiques suivantes :

- L'amplitude maximale :

$$f_1 = \max \left\{ |F(f_x, f_y)|, (f_x, f_y) \neq (0, 0) \right\} \quad (\text{A30})$$



- L'amplitude moyenne :

$$f_2 = \frac{1}{MN} \sum_{f_x=0}^{N-1} \sum_{f_y=0}^{M-1} |F(f_x, f_y)| \quad (\text{A31})$$

- L'énergie de l'amplitude :

$$f_3 = \left[ \sum_{f_x=0}^{N-1} \sum_{f_y=0}^{M-1} |F(f_x, f_y)|^2 \right]^{\frac{1}{2}} \quad (\text{A32})$$

- La variance de l'amplitude :

$$f_4 = \frac{1}{MN} \sum_{f_x=0}^{N-1} \sum_{f_y=0}^{M-1} \left[ |F(f_x, f_y)| - f_2 \right]^2 \quad (\text{A33})$$

Comme troisième possibilité, les amplitudes d'un ensemble de fréquences choisies peuvent être utilisées comme un ensemble de caractéristiques. La sélection d'un ensemble approprié dépend du problème à résoudre. Dans certains cas, les spectres de fréquence de différents segments peuvent contenir certaines fréquences apparaissant de façon régulière, avec des amplitudes plus grandes que d'autres fréquences. Celles-ci seront désignées comme des fréquences dominantes de l'ensemble de textures. Lorsqu'elles sont présentes, les amplitudes d'un ensemble de fréquences dominantes peuvent produire un ensemble de paramètres appropriés.

## b) Filtres de Gabor

Les statistiques dérivées de l'histogramme du premier ordre donnent principalement une information sur la variation des niveaux de gris, alors que les paramètres de texture basés sur les histogrammes d'ordre supérieur caractérisent les régularités spatiales des pixels à l'intérieur d'un certain voisinage. Une troisième approche pour la sélection des attributs de base est liée à la compacité de la représentation du signal. Dans la plupart des cas, les images sont hautement auto-corrélées. Le degré élevé d'autocorrélation dans les images est dû au fait que les objets à classer tendent à avoir une consistance morphologique. En effet, les valeurs de luminance localement semblables, les continuations des frontières et les textures homogènes induisent respectivement des corrélations du premier, du second et d'ordre supérieur. Pour prendre en compte cette corrélation, la distribution des caractéristiques devrait être telle que très peu de caractéristiques hautement corrélées soient fortement présentes dans la représentation du signal. Ceci motive l'utilisation des filtres ou des fonctions de Gabor (1964), qui peuvent produire une représentation compacte pour les signaux réels. Les filtres de Gabor sont des sinusoides spatiales, localisées pas une fenêtre gaussienne et définies dans un domaine spatial 2D avec la réponse impulsionnelle exprimée par la relation :

$$\psi(x, y, \theta, f) = \exp[i(f_x x + f_y y)] \exp\left[-\frac{(f_x^2 + f_y^2)(x^2 + y^2)}{2\sigma^2}\right] \quad (\text{A34})$$

Dans l'expression (A34), on a  $f_x = f \cos \theta$ ,  $f_y = f \sin \theta$ , où  $x$  et  $y$  sont les coordonnées du pixel. Le nombre  $i$  désigne la partie imaginaire d'un nombre complexe. Le paramètre  $f$  détermine la fréquence centrale du filtre passe bande dans la direction  $\theta$ , telle que  $\theta = \frac{\pi c}{K}$ ,  $c \in \{0, \dots, K-1\}$ . Le paramètre  $\sigma$  détermine la largeur de l'enveloppe gaussienne le long des directions  $x$  et  $y$ . Pour l'extraction des caractéristiques de Gabor, chaque fenêtre d'image  $W$  est convoluée avec des filtres de Gabor de différentes largeurs et orientations, résultant en image de sortie :

$$W_\psi(x, y, \theta, f) = W(x, y, c) \otimes \psi(x, y, \theta, f) \quad (\text{A35})$$

Quatre valeurs de  $\theta$  sont habituellement utilisées [HEI94, LAM92]. Les fréquences  $f$  et la largeur  $\sigma$  sont définies telle que les filtres soient sensibles aux composantes de haute fréquence dans l'image. Ainsi on a fixé :

$$\sigma = \frac{\pi}{2} \quad (\text{A36})$$

et :

$$f \in \left\{ \frac{\pi}{2}, \frac{\pi}{3}, \frac{\pi}{4} \right\} \quad (\text{A37})$$

Les filtres de Gabor complexes capturent le spectre de fréquence entier, en amplitude et en phase. Bien que l'information de phase contienne des informations sur la localisation des frontières et d'autres détails de l'image, elle n'a pas été retenue pour la suite du traitement. Les amplitudes des résultats de convolution, souvent utilisées à chaque fréquence et à chaque niveau d'orientation séparément, ont été calculées. Il y a deux principales raisons pour préférer les amplitudes aux phases des réponses de filtres des Gabor. Premièrement, l'amplitude des sorties des filtres de Gabor change doucement sur un voisinage local de l'image, ce qui est de toute vraisemblance une bonne propriété pour la discrimination des classes naturelles. Deuxièmement, l'amplitude de la transformation de Gabor est connue être robuste et tolérante à des distorsions de formes spatiales, pour plusieurs applications de reconnaissance de formes [LAM92]. Ainsi, les caractéristiques finales de Gabor  $G_w$  pour des traitements successifs sont définies selon l'expression :

$$G_w(\theta, f) = \frac{1}{W_x W_y} \sum_{x=0}^{W_x-1} \sum_{y=0}^{W_y-1} |W(x, y, \theta, f)| \quad (\text{A38})$$

En plus des énergies présentées ci-dessus, la somme des énergies  $G_w^{SUM}$  dans différentes orientations du filtre a été calculée. Elle s'exprime par la relation suivante et elles produit une caractéristique additionnelle pour chaque valeur de  $f$  :

$$G_w^{SUM} = \sum_{\theta} G(\theta, f) \quad (\text{A39})$$

Alors que la transformation en ondelettes utilise une fonction analytique pour filtrer l'image, les filtres de Gabor effectuent un fenêtrage de l'information spatiale pour une bande de fréquences donnée. Ces deux transformations sont des méthodes efficaces en analyse d'images. A l'heure actuelle, les applications des filtres de Gabor en traitement d'images de télédétection sont très peu nombreuses.

## A.5 Minimisation d'une erreur quadratique au sens des moindres carrés

Considérons la fonctionnelle suivante :

$$J(u, v) = \sum_{i=1}^N \left( u \frac{\partial I_i}{\partial x} + v \frac{\partial I_i}{\partial y} + \frac{\partial I_i}{\partial t} \right)^2 \quad (\text{A40})$$

Soit :

$$J = \sum_{i=1}^N \left( u^2 \left| \frac{\partial I}{\partial x} \right|_i^2 + v^2 \left| \frac{\partial I}{\partial y} \right|_i^2 + \left| \frac{\partial I}{\partial t} \right|_i^2 + 2uv \frac{\partial I}{\partial x} \frac{\partial I}{\partial y} \Big|_i + 2v \frac{\partial I}{\partial t} \frac{\partial I}{\partial x} \Big|_i + 2u \frac{\partial I}{\partial t} \frac{\partial I}{\partial y} \Big|_i \right) \quad (\text{A41})$$

La minimisation de cette fonctionnelle peut s'effectuer en résolvant les deux équations suivantes :

$$\begin{cases} \frac{\partial J(u, v)}{\partial u} = 0 \\ \frac{\partial J(u, v)}{\partial v} = 0 \end{cases} \quad (\text{A42})$$

Nous avons :

$$\begin{cases} \frac{\partial J}{\partial u} = \sum_{i=1}^N \left( 2u \frac{\partial I}{\partial x} \Big|_i^2 + 2v \frac{\partial I}{\partial x} \Big|_i \frac{\partial I}{\partial y} \Big|_i + 2 \frac{\partial I}{\partial t} \Big|_i \frac{\partial I}{\partial x} \Big|_i \right) \\ \frac{\partial J}{\partial v} = \sum_{i=1}^N \left( 2v \frac{\partial I}{\partial y} \Big|_i^2 + 2u \frac{\partial I}{\partial x} \Big|_i \frac{\partial I}{\partial y} \Big|_i + 2 \frac{\partial I}{\partial t} \Big|_i \frac{\partial I}{\partial y} \Big|_i \right) \end{cases} \quad (\text{A43})$$

La solution obtenue est donnée par :

$$\begin{cases} u = \frac{\left( \sum_{i=1}^N \frac{\partial I}{\partial t} \Big|_i \frac{\partial I}{\partial y} \Big|_i \right) \left( \sum_{i=1}^N \frac{\partial I}{\partial x} \Big|_i \frac{\partial I}{\partial y} \Big|_i \right) - \left( \sum_{i=1}^N \frac{\partial I}{\partial t} \Big|_i \frac{\partial I}{\partial x} \Big|_i \right) \left( \sum_{i=1}^N \frac{\partial I}{\partial y} \Big|_i^2 \right)}{\left( \sum_{i=1}^N \frac{\partial I}{\partial x} \Big|_i^2 \sum_{i=1}^N \frac{\partial I}{\partial y} \Big|_i^2 \right) - \left( \sum_{i=1}^N \frac{\partial I}{\partial x} \Big|_i \frac{\partial I}{\partial y} \Big|_i \right)^2} \\ v = \frac{\left( \sum_{i=1}^N \frac{\partial I}{\partial t} \Big|_i \frac{\partial I}{\partial x} \Big|_i \right) \left( \sum_{i=1}^N \frac{\partial I}{\partial x} \Big|_i \frac{\partial I}{\partial y} \Big|_i \right) - \left( \sum_{i=1}^N \frac{\partial I}{\partial t} \Big|_i \frac{\partial I}{\partial y} \Big|_i \right) \left( \sum_{i=1}^N \frac{\partial I}{\partial x} \Big|_i^2 \right)}{\left( \sum_{i=1}^N \frac{\partial I}{\partial x} \Big|_i^2 \sum_{i=1}^N \frac{\partial I}{\partial y} \Big|_i^2 \right) - \left( \sum_{i=1}^N \frac{\partial I}{\partial x} \Big|_i \frac{\partial I}{\partial y} \Big|_i \right)^2} \end{cases} \quad (\text{A44})$$

Sous forme matricielle, nous avons :

$$J(d) = \|Ad + B\|_2^2 \quad (\text{A45})$$

$\| \cdot \|_2$  désigne la norme Euclidienne, la quantité  $A$  correspond la matrice de gradients spatiaux et  $B$  le gradient temporel.

La dérivation de  $J$  donne :

$$\frac{\partial J}{\partial d} = 2A^T 2(A\vec{d} + B) \quad (\text{A46})$$

En égalant cette dérivée à zéro, nous obtenons :

$$\vec{d} = \begin{pmatrix} u \\ v \end{pmatrix} = (A^T A)^{-1} AB \quad (\text{A47})$$

## A.6 Différents masques utilisés à l'estimation des gradients

Les différents masques utilisés à l'estimation des gradients spatiaux et gradients temporels sont présentés dans le tableau A.1.

Gradient	Masque	Estimation
Gradients temporels	$\text{SIMPLE}_t = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	$\text{SIMPLE}_t * [I(x, y, t+1) - I(x, y, t)]$
	$\text{GAUSSIEN}_t = \frac{1}{10} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$	$\text{GAUSSIEN}_t * [I(x, y, t+1) - I(x, y, t)]$
	$\text{GAUSSIEN}_t = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$	$\text{GAUSSIEN}_t * [I(x, y, t+1) - I(x, y, t)]$
	$\text{MOYEN}_t = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$	$\text{MOYEN}_t * [I(x, y, t+1) - I(x, y, t)]$
Gradients spatiaux suivant x	$\text{SIMPLE}_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \end{bmatrix}$	$\text{SIMPLE}_x * I(x, y, t)$
	$\text{DOUBLE}_x = \frac{1}{2} \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$	$\text{DOUBLE}_x * I(x, y, t)$
	$\text{PREWITT}_x = \frac{1}{6} \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$	$\text{PREWITT}_x * I(x, y, t)$
	$\text{SOBEL}_x = \frac{1}{8} \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$	$\text{SOBEL}_x * I(x, y, t)$
	$\text{MDIF}_x = \frac{1}{30} \begin{bmatrix} 0 & -1 & -1 & -1 & 0 \\ -1 & -2 & -3 & -2 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 3 & 2 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{bmatrix}$	$\text{MDIF}_x * I(x, y, t)$

Tableau A.1 : Approximation des gradients spatiaux et temporels.

Gradient	Masque	Estimation
Gradients spatiaux suivant y	$\text{SIMPLE}_y = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$	$\text{SIMPLE}_y * I(x, y, t)$
	$\text{DOUBLE}_y = \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$	$\text{DOUBLE}_y * I(x, y, t)$
	$\text{PREWITT}_y = \frac{1}{6} \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 1 & 1 \end{bmatrix}$	$\text{PREWITT}_y * I(x, y, t)$
	$\text{SOBEL}_y = \frac{1}{8} \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$	$\text{SOBEL}_y * I(x, y, t)$
	$\text{MDIF}_y = \frac{1}{30} \begin{bmatrix} 0 & -1 & 0 & 1 & 0 \\ -1 & -2 & 0 & 2 & 1 \\ -1 & -3 & 0 & 3 & 1 \\ -1 & -2 & 0 & 2 & 1 \\ 0 & -1 & 0 & 1 & 0 \end{bmatrix}$	$\text{MDIF}_y * I(x, y, t)$

Tableau A.1 : Approximation des gradients spatiaux et temporels (suite).

## Annexe B

# Rappel des outils mathématiques utiles dans le cadre de la segmentation

### B.1 Classification bayésienne

Soit  $x$  un vecteur d'observation dans un espace à  $p$  dimensions. On note  $x = (x_1, x_2, \dots, x_p)^T$ . Désignons par  $X$  l'ensemble de ces observations. A titre d'illustration, dans le cas d'une image en niveaux gris,  $p$  peut être égal à un pour une image monochrome ;  $x$  correspond à l'intensité lumineuse en un point et  $X$  constitue l'ensemble des niveaux de gris représentatif de l'image. L'objectif de la classification est de répartir le nuage  $X$  en  $k$  sous-nuages  $X_k$ . Cela revient à déterminer si  $x$  appartient à l'une des classes  $\{w_1, \dots, w_k\}$ . La classification repose sur le modèle bayésien. C'est une description statistique du problème de classification ou d'estimation qui assume une connaissance a priori du contexte et ses propriétés. Plus précisément, le modèle bayésien est basé sur deux propriétés :

- Un modèle a priori, qui est donné par une distribution des états possibles  $w$ , nommée  $P(w_i)$ . Cette dernière représente la probabilité qu'un vecteur  $x$  pris au hasard appartienne à la  $i^{\text{ème}}$  structure.

- Un modèle sensoriel, qui décrit les processus stochastiques qui relient l'état  $w$  aux valeurs sensorielles mesurées  $x$ . Ce modèle est donné par  $P(x|w_i)$ , c'est la probabilité a priori de l'observation à la  $i^{\text{ème}}$  structure.

La quantité d'intérêt est le modèle a posteriori,  $P(w_i|x)$  de l'état  $w_i$ , qui est relié par le modèle sensoriel et le modèle a priori par la loi de Bayes ;  $P(x|w_i)$  est la probabilité a posteriori que le vecteur d'observation appartienne à la  $i^{\text{ème}}$  structure. Pour un vecteur  $x$  donné, la classification consiste à l'attribuer à la région qui maximise la probabilité a posteriori (on appelle cette valeur la MAP estimée).

$$P(w_i|x) = \frac{P(x|w_i)P(w_i)}{P(x)} \quad (\text{B1})$$

Selon l'application, une fonction coût  $C$  qui dépend de  $P(x|w_i)$  est désignée. Si un estimateur produit  $w_i$  au lieu d'une réponse réelle  $\hat{w}$ , un coût  $C(w_i, \hat{w})$  est déduit. Le coût prévu compte tenu des mesures est :

$$E(C(w_i, \hat{w})|x) = \int C(w_i, \hat{w})P(w_i|x)dw \quad (\text{B2})$$

L'estimé bayésien est la valeur de  $w_i$  qui minimise  $E(C(w_i, \hat{w})|x)$ . Un cas particulier est fourni par un coût binaire, c'est-à-dire égal à l'unité quand la valeur est fautive et égal à zéro quand la valeur est juste.

$$C(w_i, \hat{w}) = 1 - \delta(w_i - \hat{w}) \quad (\text{B3})$$

Par conséquent, le coût particulier devient :

$$E(C(w_i, \hat{w})|x) = 1 - P(w_i|x) \quad (\text{B4})$$

Le minimum de cette expression correspond au maximum de  $P(w_i|x)$ , c'est-à-dire :

$$P(w_k|x) = \max_i P(w_i|x) \Rightarrow x \in w_k \quad (\text{B5})$$

La structure estimée  $w_k$  s'appelle estimée au sens du maximum a posteriori (MAP) que nous notons par  $w_{MAP}$ . Comme le dénominateur dans l'équation (B1) reste constant, il ne joue pas un rôle dans le procédé de maximisation. Le formalisme MAP devient :

$$w_{MAP} = \arg \max_w [P(x|w)P(w)] \quad (B6)$$

L'erreur de Bayes peut être calculée à partir de l'équation (B6). Elle correspond à la valeur moyenne suivante :

$$\varepsilon = E(1 - \max\{P(w_i|x)\}) \quad (B7)$$

La notation  $E$  indique la valeur moyenne.

## B.2 Champ de Markov

Le champ aléatoire de Markov est la distribution de probabilité définie sur un modèle discret de points (noté  $L$ ), où la probabilité d'une variable particulière dépend uniquement du petit nombre de ses voisinages [GEM84].

$$P(w_i|w) = P(w_i | \{w_j, w_j \in N_i\}) \quad (B8)$$

$N_i$  est le voisinage considéré d'un point  $i \in L$ . Un système de voisinage peut être défini de plusieurs manières. En l'occurrence, un critère de voisinage circulaire est défini selon le principe suivant : soit  $(x_i, y_i)$  le couple de coordonnées spatiales d'un point  $i$  ; le  $n^{em}$  ordre du système de voisinage  $\eta_n$  est défini par :

$$\eta_n(i) = \{j | ((x_j - x_i)^2 + (y_j - y_i)^2) \leq n\} \quad (B9)$$

Un sous-ensemble  $C$  de  $\eta_n(i)$  est à  $n$  cliques si chaque paire de différents sites dans  $C$  sont des voisinages d'ordre  $n$ .

Les probabilités conditionnelles peuvent être utilisées pour générer un modèle a priori  $P(w)$  pour les estimations bayésiennes. Pour cela, une propriété fondamentale du champ de Markov, connue sous le nom de théorème de Hammersley-Clifford, déclare que sa densité de probabilité pourrait être exprimée comme une distribution de Gibbs. La distribution de Gibbs est une distribution exponentielle de type :

$$P(w) = \frac{1}{Z} e^{-E_p(w)} \quad (B10)$$

L'énergie  $E_p(w)$  peut être écrite comme la somme de toutes les énergies locales des cliques :

$$E_p(w) = \sum_{c \in C} E_c(w) \quad (B11)$$

$E_c(w)$  dépend uniquement des valeurs de  $w$  localisé sur les sites de la clique  $c$  ;  $Z$  est un terme de normalisation, appelé fonction de partition :

$$Z = \sum_w e^{-E(w)} \quad (B12)$$

L'approche bayésienne-markovienne essaie de maximiser la probabilité conditionnelle suivant la relation (B1). Le dénominateur  $P(x)$  est une distribution constante par rapport aux étiquettes car la réalisation du champ des observations est connue.  $P(w)$  est la distribution markovienne. Maximiser l'équation (B1) revient à maximiser  $P(x|w)P(w)$ . On a :

$$w_{MAP} = \arg \max_w P(w|x) \quad (B13)$$

Soit après développement :

$$w_{MAP} = \arg \max_w \left\{ \frac{1}{Z} \exp \left[ - \left( - \ln \prod_{s \in L} P(x|w) + \sum_{c \in C} E_c(w) \right) \right] \right\} \quad (B14)$$

L'équation (B14) peut être exprimé de la forme :

$$w_{MAP} = \arg \max_w \left\{ \frac{1}{Z} \exp \left[ - \left( \sum_{s \in L} V(w, x) + \sum_{c \in C} E_c(w) \right) \right] \right\} \quad (B15)$$

avec :

$$V(w, x) = - \ln P(x|w) \quad (B16)$$

Cette maximisation peut s'effectuer en minimisant l'énergie :

$$w_{MAP} = \arg \min_w E(x, w) \quad (B17)$$

avec :

$$E(x, w) = \sum_{s \in L} V(w, x) + \sum_{c \in C} E_c(w) \quad (B18)$$

Pour minimiser cette fonctionnelle, qui est non linéaire, différentes techniques ont été proposées [BES86, KIR83, MET53, DUM99].

### B.3 Classification basée sur l'algorithme K-moyens

Dans ce paragraphe, nous rappelons brièvement la méthode de classification basée sur l'algorithme K-moyens. Le principe de cette technique est le suivant :

- Soit un nuage  $X$  de  $N$  vecteurs de caractéristiques  $x_i$ ,  $X = \{x_i, 1 \leq i \leq N\}$ . On veut répartir ce nuage en  $k$  classes. Un vecteur d'observation est représenté par  $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})^T$  ; la quantité  $p$  désigne le nombre de dimension de l'espace d'observation. Dans le cas d'une image en niveaux de gris,  $N$  correspond au nombre de pixels et  $p$  est égal à 1. En revanche, dans le cas d'une image couleur,  $p$  est égal à 3 (exemple : les composantes Rouge Verte Bleue attribuées à chaque pixel). Supposons que notre objectif consiste à déterminer si  $x$  appartient à l'une des classes  $\{w_1, \dots, w_k\}$ . La classification repose sur l'algorithme K-moyens.

- Soit une distance  $d$  définie sur l'espace des vecteurs. Par exemple, il s'agira de la distance euclidienne canonique classique :

$$\text{si } x = (x_1, x_2, \dots, x_p)^T \text{ et } y = (y_1, y_2, \dots, y_p)^T, \text{ alors } d(x, y) = \sqrt{\sum_{j=1}^p (x_j - y_j)^2} \quad (B19)$$

L'objectif de la classification est de diviser le nuage de points  $X$  en  $k$  sous-nuages  $X_k$ . De manière équivalente, ceci revient à déterminer l'application  $c$  qui à chaque vecteur  $x_i$  fait correspondre le numéro de classe  $c(x_i)$  du sous-nuage auquel appartient  $x$ . La détermination de  $c$  est effectuée par minimisation d'un critère. Les différentes méthodes de classification se différencient en partie par le choix de la distance  $d$ , le choix du critère, et le choix de la méthode de minimisation du critère. Un problème important concerne la détermination automatique du nombre de classes à utiliser.

Dans l'algorithme K-moyens, le nombre de classes  $k$  est une constante. A chacune des classes  $k$ , on associe un vecteur  $q_k$  de représentant cette classe. L'application  $c$  est définie par la formule suivante :

$$c = x_i \mapsto k = c(x_i) / \forall l = 1, \dots, k \text{ on a } d(x_i, q_k) \leq d(x_i, q_l) \quad (B20)$$

Ce qui correspond à effectuer à chaque mesure  $x_i$  le numéro de la classe représentant  $q_k$  le plus proche. L'application  $c$  est donc entièrement caractérisée par la connaissance des représentants  $q_k$ . Le problème posé se ramène donc à la détermination de ces représentants. En utilisant la définition précédente de  $c$ , les sous-nuages de points, qui sont définis comme étant l'ensemble des points qui appartiennent à une classe donnée, sont caractérisés par :

$$X_k = \{x_i / c(x_i) = k\} = \{x_i / \forall l = 1, \dots, k \text{ on a } d(x_i, q_k) \leq d(x_i, q_l)\} \quad (B21)$$



## B.4 Formalisme du maximum de vraisemblance

Dans le cas où aucune distribution a priori n'est connue, on peut considérer la distribution uniforme. Ce principe revient à déterminer un paramètre  $w$  qui maximise  $P(x|w)$ . C'est à dire sélectionner les paramètres pour lesquels la probabilité de l'occurrence des mesures actuelles est maximale. La quantité  $P(x|w)$  considérée comme une fonction de  $w$  est connue sous le nom vraisemblance.

$$w_{ML} = \arg \max_w P(x|w) \quad (\text{B22})$$

La règle de décision du maximum de vraisemblance<sup>23</sup> affecte chaque pixel, de vecteur de mesure  $x$ , à la classe  $w_i$  dont les unités sont susceptibles de produire le vecteur de mesures  $x$ .

La classification par maximum de vraisemblance utilise les statistiques des classes déjà calculées, comprenant le vecteur moyen  $\mu_i$  et la matrice de covariance  $V_i$  pour chaque classe  $i$ . La règle de décision du maximum de vraisemblance affecte le vecteur  $x$  à la classe  $w_i$ , si et seulement si la condition illustrée par l'ensemble des deux équations suivantes est satisfaite :

$$P(x|w_i) = \max_j P(x|w_j) \Rightarrow x \in w_i \quad (\text{B23})$$

$$P(x|w_i) = \frac{1}{(2\pi)^k |V_i|} e^{-\frac{1}{2}(x-\mu_i)^T V_i^{-1}(x-\mu_i)} \quad (\text{B24})$$

$|V_i|$  représente le déterminant de la matrice de covariance  $V_i$ . Par conséquent, pour placer un vecteur de mesures  $x$  d'un pixel inconnu dans une classe, la règle de décision du maximum de vraisemblance calcule la valeur de  $P(x|w)$  pour chaque classe. Puis, elle affecte le pixel à la classe qui présente la plus grande valeur. L'équation (B24) suppose que toutes les classes ont la même probabilité d'exister sur le terrain. Or en pratique, il arrive souvent que certaines classes soient plus fréquentes que d'autres. Il peut donc être utile d'inclure ces informations a priori dans le processus de classification, en pondérant chaque classe  $w_i$  par sa probabilité a priori  $P(w_i)$ . La règle de décision devient alors celle qui est exprimée par la relation suivante:

$$P(x|w_i)P(w_i) = \max_j \{P(x|w_j)P(w_j)\} \Rightarrow x \in w_i \quad (\text{B25})$$

Cette décision bayésienne est identique à la règle de décision du maximum de vraisemblance, excepté le fait qu'elle ne suppose pas que toutes les classes aient des probabilités égales d'être rencontrées dans l'image.

Une décision de classification qui consiste à établir une fonction discriminante peut aussi être utilisée. En général, pour  $k$  classes,  $k$  fonctions discriminantes notées  $g_j(x)$ ,  $\forall j = 1, \dots, k$ , sont établies et la classification est donnée par :

$$\arg \max_{j=1, \dots, k} g_j(x) \quad (\text{B26})$$

A l'exception de certains cas, les fonctions de densité de probabilités normales ne sont jamais estimées.  $P(x|w_i)$  peut être estimé en adaptant une fonction de densité de probabilités normales à la moyenne  $\mu_i$  et la covariance  $V_i$  de  $x$  et seulement l'ordre de rang de la fonction de décision nécessite d'être conservé. Des fonctions discriminantes plus simples, exprimées par les équations (B27) et (B28), sont ainsi utilisées.

$$g_j(x) = \ln P(w_j) - \frac{1}{2} \ln |V_j| - \frac{1}{2} (x - \mu_j)^T V_j^{-1} (x - \mu_j) \quad (\text{B27})$$

<sup>23</sup> Maximum likelihood

Seule la forme quadratique en  $w$  est calculée, et il existe des méthodes efficaces pour le faire. En outre, si on suppose que les classes ont la même probabilité a priori, on obtient encore une forme plus simple des fonctions discriminantes :

$$g_j(x) = \frac{1}{2} \ln(|V_j|) + \frac{1}{2} (x - \mu_j)^T V_j^{-1} (x - \mu_j) \quad (\text{B28})$$

Dans ce cas, un pixel de vecteur de mesures  $x$  sera affecté à la classe  $i$  si et seulement si la condition exprimée par l'équation suivante est satisfaite :

$$g_i(x) = \min_j g(x|w_j), \forall j = 1, \dots, k \quad (\text{B29})$$

## B.5 Formalisme MDL pour la détermination de la complexité du modèle

La difficulté majeure dans l'estimation des paramètres modélisant un processus aléatoire est la détermination de l'ordre du modèle adéquat. La tendance étant que l'on estime un ordre trop élevé pour le modèle de représentation. Ceci est dû au fait que les mesures utilisées pour qualifier les modèles de représentation sont en parties basées sur une mesure de l'erreur de prédiction, et que cette dernière diminue avec l'ordre du modèle utilisé. Le problème avec le formalisme du maximum de vraisemblance est qu'il n'y a pas de limite dans la complexité du modèle mixte, c'est-à-dire le nombre de populations. Généralement, plus le nombre de modèles est élevé, meilleure sera la description obtenue. Dans ce cadre, Rissanen propose une technique intitulée MDL<sup>24</sup> basée sur la minimisation du coût de description global afin d'obtenir un ordre adéquat du modèle de représentation [RIS83]. Le but est de déterminer les paramètres du modèle qui minimise le coût total de description. Le formalisme MDL conduit à une fonction objective sans seuil arbitraire. Avec le formalisme MDL, le modèle peut être codé à faible coût tandis que la description de l'illustration est meilleure. L'avantage de l'approche MDL est aussi la manière uniforme dans laquelle on peut combiner des modèles purement stochastiques (tels que le bruit blanc) avec des modèles déterministes (tels que les modèles polynomiaux).

Avant de traiter du formalisme MDL proprement dit, nous allons d'abord introduire la notion du coût de description<sup>25</sup> [RIS83, FOS98]. Le coût total de description d'une séquence de  $n$  variables aléatoires  $x_1, x_2, \dots, x_n$ , basé sur un vecteur des paramètres  $\Phi$  est défini par :

$$L(x, \Phi) = \log_2^* [C(k) \|\Phi\|^k] - \log_2 V(x|\Phi) \quad (\text{B30})$$

$V(x|\Phi)$  est la fonction de vraisemblance d'une variable  $x$  pour un modèle de paramètres  $\Phi$ ,  $C(k)$  est le volume de la boule unité de dimension  $k$ . La norme du paramètre est défini par :

$$\|\Phi\| = \sqrt{\Phi' H \Phi} \quad (\text{B31})$$

avec :

$$H = \frac{-\partial^2 \log_2 V(x|\Phi)}{\partial \Phi^2} \quad (\text{B32})$$

La fonction  $\log_2^*$  désigne le logarithme itératif à base 2.

$$\log_2^* \cdot = \log_2 \cdot + \log_2 \log_2 \cdot + \dots \quad (\text{B33})$$

où la somme s'étend uniquement aux termes positifs. Le coût de description consiste en l'approximation du nombre de bits utile pour encoder les paramètres du modèle.

D'après l'équation (B30), le coût de description total contient deux parties : une partie pour le modèle et une autre pour les données utilisant ce modèle. L'équation (B30) peut se généraliser par :

<sup>24</sup> Minimum Description Length

<sup>25</sup> Description Length

$$L(x, \Phi) = L(\Phi) - \log_2 V(x|\Phi) \quad (\text{B34})$$

$L(\Phi)$  représente le coût de description en terme de bits nécessaire pour décrire les paramètres  $\Phi$ . Le terme  $-\log_2 V(x|\Phi)$  correspond au coût de codage idéal de l'observation  $x$  connaissant le modèle de représentation  $\Phi$ . Afin de choisir le meilleur modèle paramétrique parmi  $k$  possibilités, Rissanen propose le modèle permettant d'obtenir le coût de description minimal.

$$MDL(x, \Phi) = \min_{\Phi} \{L(\Phi) - \log_2 V(x|\Phi)\} \quad (\text{B35})$$

Comme critère de sélection d'un modèle, MDL présente une compensation explicite de la complexité du modèle et la bonne adéquation aux données. Pour caractériser la distribution utilisée pour la description des  $n$  points [RIS86], si l'ordre du modèle est  $k$ , le coût de description  $L(\Phi)$  peut s'exprimer par :

$$L(\Phi) = \frac{k}{2} \log n \quad (\text{B36})$$

Concrètement, le formalisme MDL consiste à minimiser le coût de la description des données (de longueur  $n$  et de paramètres  $\Phi$ ) en minimisant le critère (B35). Soit :

$$MDL(x, \Phi) = \min_{\Phi} \left\{ \frac{k}{2} \log n - \log_2 V(x|\Phi) \right\} \quad (\text{B37})$$

Dans le cadre de la segmentation du mouvement, pour un pixel  $p_j$ , le logarithme à base 2 de la fonction de vraisemblance des paramètres  $\Phi$  peut s'écrire [AYE95b] :

$$\log_2 V(I(p_j)|\Phi) = \log_2 \prod_{j=1}^n P(I(p_j)|\Phi) \quad (\text{B38})$$

Dans la formulation du modèle mixte, le vecteur  $\Phi$  contient tous les paramètres inconnus.

$$\Phi = (\Pi, \Sigma, \Theta)^T \quad (\text{B39})$$

avec  $\Pi = (\pi_1, \dots, \pi_m)^T$ ,  $\Sigma = (\sigma_1, \dots, \sigma_m)^T$  et  $\Theta = (\theta_1, \dots, \theta_m)^T$ . Dans cette notation,  $\theta_i$  désigne le paramètre du modèle de mouvement,  $\sigma_i^2$  la variance et  $\pi_i$  la proportion de l' $i^{\text{em}}$  modèle dans la mixture telle que :

$$\sum_{i=1}^m \pi_i = 1, \pi_i > 0 \quad (\text{B40})$$

En introduisant les expressions des paramètres inconnus, la relation (B38) peut s'écrire :

$$\log_2 V(I(p_j)|\theta) = \sum_{j=1}^n \log_2 \left[ \sum_{i=1}^g \pi_i P(I(p_j)|\sigma_i) \right] \quad (\text{B41})$$

Sous l'hypothèse de la distribution normale de résidus, et si les résidus sont quantifiés au voisinage d'une précision  $\varepsilon$ , nous avons :

$$P(I(p_j)|\theta) \approx \frac{\varepsilon}{\sqrt{2\pi}\sigma_i} \exp\left[\frac{-r_i^2(p_j)}{2\sigma_i^2}\right], \text{ quand } \varepsilon < \sigma_i \quad (\text{B42})$$

Le coût de description total est donné par :

$$MDL(I(p_j), \theta) = \min_{\theta} \left\{ \frac{k}{2} \log n - \sum_{j=1}^n \log_2 \left[ \sum_{i=1}^g \pi_i \frac{\varepsilon}{\sqrt{2\pi}\sigma_i} \exp\left(\frac{-r_i^2(p_j)}{2\sigma_i^2}\right) \right] \right\} \quad (\text{B43})$$

Ce formalisme peut être utilisé pour optimiser le modèle de représentation d'un modèle de chaque région dans la phase de fusion.

## Annexe C

# Analyse des erreurs d'estimation associées à la méthode différentielle

### C.1 Analyse

Considérons l'équation de contrainte du mouvement, appliquée sur deux pixels voisins  $p_i$  et  $p_j$  :

$$\begin{cases} uI_x^{(i)} + vI_y^{(i)} + I_t^{(i)} = 0 \\ uI_x^{(j)} + vI_y^{(j)} + I_t^{(j)} = 0 \end{cases} \quad (C1)$$

Ce système d'équation est entaché d'une erreur dont l'origine peut venir des trois facteurs suivants :

- Comme nous avons vu dans le paragraphe 1.5.1.b de la première partie, les dérivées partielles sont estimées à partir des données discrètes. Les erreurs sur les gradients spatio-temporels dépendent des approximations utilisées.
- L'erreur qui vient de l'hypothèse de cohérence, supposant que le déplacement des deux pixels  $p_i$  et  $p_j$  est le même.
- L'erreur de propagation caractéristique du système linéaire.

Un système est bien-conditionné s'il est insensible aux petites perturbations. Nous allons ici reprendre l'analyse décrite dans [KEA87].

#### a) Erreur due à l'estimation des gradients

Les estimateurs différentiels introduisent des biais à cause des bruits dans la structure lumineuse de l'image et de l'erreur due à l'échantillonnage spatial et temporel [KEA87, KEA83]. Ce dernier peut s'exprimer en fonction du terme d'ordre supérieur<sup>26</sup>, de l'intensité lumineuse :

$$I(x + \Delta x, y, t) = I(x, y, t) + \hat{I}_x \Delta x + \hat{I}_{xx} \Delta x^2 + hod \quad (C2)$$

La quantité  $\Delta x$  désigne le pas d'échantillonnage,  $\hat{I}_x$  et  $\hat{I}_{xx}$  sont respectivement les dérivées première et seconde théoriques de l'intensité lumineuse. La valeur discrète d'une dérivée est calculée par l'estimateur suivant :

$$I_x = \frac{I(x + \Delta x, y, t) - I(x, y, t)}{\Delta x} \quad (C3)$$

soit :

$$I_x = \hat{I}_x + \hat{I}_{xx} \Delta x + hod \quad (C4)$$

L'erreur  $\varepsilon_{I_x}$  est définie comme la différence  $I_x - \hat{I}_x$ . De l'équation (C2), nous obtenons la relation approchée :

$$\varepsilon_{I_x} = \hat{I}_{xx} \Delta x \quad (C5)$$

Par analogie, les erreurs d'échantillonnage dans les estimés  $I_y$  et  $I_t$  sont respectivement évaluées selon les équations (C6) et (C7).

$$\varepsilon_{I_y} = \hat{I}_{yy} \Delta y \quad (C6)$$

---

<sup>26</sup> Hod : Higher-order derivatives

et :

$$\varepsilon_{I_t} = \hat{I}_u \Delta t \quad (C7)$$

L'erreur sur l'estimation des gradients spatiaux dépend de la résolution spatiale de la caméra et les dérivées spatiales secondes de l'intensité lumineuse.

Quant au gradient temporel, l'erreur d'échantillonnage est influencée par la fréquence temporelle et les dérivées d'ordre supérieur de l'intensité par rapport au temps. Le contexte suivant consiste à analyser cette fréquence temporelle de façon plus détaillée. Pour ce faire, nous faisons appel aux équations de contrainte du second ordre. Le vecteur vitesse  $\vec{V}(u, v)$  est implicitement considéré localement constant. Reprenons le système d'équations des contraintes définie dans le paragraphe consacré aux méthodes différentielles du second ordre (première partie, chapitre 1, paragraphe 1.5.7.c). Dans ce système d'équation, si on substitue la valeur de  $I_{xt}$  de la première équation et de  $I_{yt}$  de la deuxième équation dans la troisième équation, on obtient :

$$\begin{bmatrix} u & v \end{bmatrix} H \begin{bmatrix} u & v \end{bmatrix}^T = \hat{I}_u \quad (C8)$$

$H$  est la matrice du Hessian constituée par les dérivées secondes théoriques.

Dans [KEA87], sans perte de généralité, on fait subir au système des coordonnées une rotation de telle sorte que le déplacement en un point s'étende sur l'axe des  $x$ . Dans ce nouveau système des coordonnées, nous pouvons écrire :

$$u^2 \hat{I}_{xx} = \hat{I}_u \quad (C9)$$

L'erreur d'échantillonnage temporel devient :

$$\varepsilon_{I_t} = u^2 \hat{I}_{xx} \Delta t \quad (C10)$$

L'erreur donnée par cette équation est proportionnelle à la dérivée seconde de la luminance suivant la direction du mouvement, au carré de la valeur du déplacement ainsi qu'à la fréquence temporelle. L'erreur sur l'estimation du gradient temporel est plus sensible quand on se trouve dans une région où le déplacement est important et la fonction de luminance est non linéaire.

## b) Erreur due à l'hypothèse de cohérence

Le mouvement dans un voisinage local n'est pas forcément constant. Le système d'équations de contrainte pour les deux pixels  $i$  et  $j$  s'écrit ainsi [KEA87] :

$$\begin{cases} uI_x^{(i)} + vI_y^{(i)} + I_t^{(i)} = 0 \\ (u + \Delta u)I_x^{(j)} + (v + \Delta v)I_y^{(j)} + I_t^{(j)} = 0 \end{cases} \quad (C11)$$

$\Delta u$  ( ou  $\Delta v$  ) représente le décalage local entre la valeur du déplacement d'un pixel  $i$  et celle du pixel voisin  $j$ . Nous ne distinguons pas ici la notation entre les gradients théoriques et les gradients estimés car le but n'est pas d'analyser l'erreur sur l'estimation des gradients.

En arrangeant la deuxième équation, ce système s'exprime sous la forme :

$$\begin{cases} uI_x^{(i)} + vI_y^{(i)} + I_t^{(i)} = 0 \\ uI_x^{(j)} + vI_y^{(j)} + I_t^{(j)} + \varepsilon_{\Delta d} = 0 \end{cases} \quad (C12)$$

avec :

$$\varepsilon_{\Delta d} = \begin{bmatrix} I_x^{(j)} & I_y^{(j)} \end{bmatrix} \begin{bmatrix} \Delta u & \Delta v \end{bmatrix}^T \quad (C13)$$

L'erreur due à l'hypothèse d'uniformité du mouvement pourrait être considérée comme un bruit additionnel sur le gradient temporel. Exprimons la grandeur de cette erreur relative au gradient temporel . Soit :

$$\left| \frac{\varepsilon_{\Delta d}}{I_t} \right| = \frac{\left| \begin{bmatrix} I_x^{(j)} & I_y^{(j)} \end{bmatrix} \begin{bmatrix} \Delta u & \Delta v \end{bmatrix}^T \right|}{\left| \begin{bmatrix} I_x^{(j)} & I_y^{(j)} \end{bmatrix} \begin{bmatrix} u + \Delta u & v + \Delta v \end{bmatrix}^T \right|} \quad (\text{C14})$$

Le numérateur ou le dénominateur du second membre n'est qu'un produit scalaire de deux vecteurs (figure C.1).

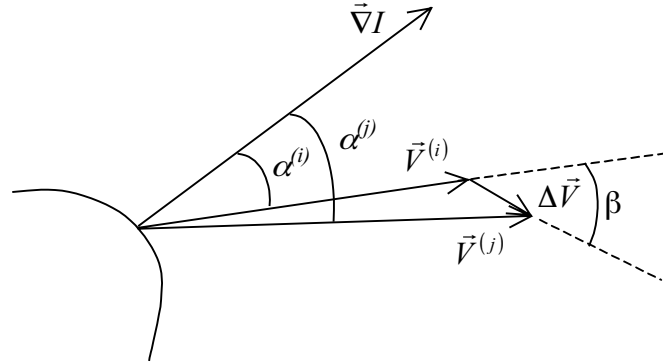


Figure C.1 : Représentation graphique de l'écart entre le gradient spatial et le déplacement.

- Si  $\beta \neq 0$ , il y a un décalage angulaire entre le vecteur de déplacement du pixel  $i$  et celui du pixel  $j$ . On n'a ni cohérence spatiale en orientation, ni cohérence spatiale en amplitude. L'équation (C14) peut s'exprimer sous la forme

$$\left| \frac{\varepsilon_{\Delta d}}{I_t} \right| = \frac{\|\vec{v}\| \cos(\alpha^{(i)} + \beta)}{\|\vec{v} + \Delta\vec{v}\| \cos \alpha^{(j)}} \quad (\text{C15})$$

L'erreur relative en  $I_t$  dépend de l'amplitude relative de  $\vec{v}$  et  $\Delta\vec{v}$ , la valeur du cosinus de l'angle formé par le gradient spatial et le vecteur du déplacement, et finalement de l'angle formé par le gradient spatial et le vecteur du décalage local.

- Si  $\beta = 0$ , c'est à dire nous avons une cohérence spatiale en orientation mais pas en amplitude. Pour tous les pixels du voisinage, l'angle formé par le vecteur gradient spatial et le déplacement est le même. C'est le cas plus général car l'orientation du gradient spatial  $\vec{\nabla}I$  est indépendante de l'amplitude du déplacement  $\vec{v}$  et le décalage local  $\Delta\vec{v}$ . L'équation (C15) peut s'écrire :

$$\left| \frac{\varepsilon_{\Delta d}}{I_t} \right| = \frac{\|\vec{v}\|}{\|\vec{v} + \Delta\vec{v}\|} \quad (\text{C16})$$

L'erreur relative en  $I_t$  dépend fortement de l'amplitude du déplacement et de celle du décalage spatial. Dans la majorité des scènes, le flot changera lentement sur la majeure partie de l'image. Sur les contours, nous pouvons trouver fréquemment des discontinuités dans le champ de déplacement. Dans ce cas là, la variation du flot contribuera par une erreur résiduelle et l'estimé est généralement biaisé. En revanche, une grande partie de l'image est constituée par des surfaces variables en terme de lissage. Quand les points voisins se trouvent sur la même surface homogène, le flot sera généralement uniforme et par conséquent, l'erreur engendrée par des variations du flot sera petite.

### c) Erreur de propagation caractéristique du système linéaire : moindres carrés

Il y a plusieurs manières d'analyser les erreurs générées par la résolution au sens des moindres carrés. Ce paragraphe se contente de passer brièvement en revue quelques tendances prises par l'erreur en fonction du type de bruits présents dans l'image [FER99]. Reprenons l'équation de contrainte du mouvement sous la forme :

$$\hat{A}_i u + \hat{B}_i v = \hat{C}_i \quad (C17)$$

L'indice  $i$  montre un pixel dans un espace spatio-temporel ou dans un espace fréquentiel. La notation « $\hat{\phantom{x}}$ » signifie qu'il s'agit d'une valeur théorique. Exprimons cette expression en introduisant les termes d'erreurs, qui sont les différences entre les données réelles et les données estimées,  $(\Delta A_i = A_i - \hat{A}_i)$  :

$$(A_i - \Delta A_i)u + (B_i - \Delta B_i)v = C_i - \Delta C_i \quad (C18)$$

L'équation (C18) peut aussi s'exprimer sous la forme :

$$(E_i - \Delta E_i)\vec{V} = C_i - \Delta C_i \quad (C19)$$

$E$  et  $\Delta E$  sont des matrices  $n \times 2$  qui contiennent respectivement les données  $[A \ B]$  et  $[\Delta A \ \Delta B]$ . Le vecteur  $\vec{V}(u, v)$  représente le flot optique. Par définition, la solution au sens des moindres carrés est donnée par :

$$\vec{V} = (E' E)^{-1} E' C \quad (C20)$$

S'il n'y a aucune erreur dans l'estimation des coefficients  $A_i$  et  $B_i$ , alors sous l'hypothèse usuelle que les différences  $\Delta C_i$  soient non corrélées et ont la même variance, la méthode des moindres carrés donne un estimé non biaisé et il est aussi simple de donner des limites de confiance pour la solution. Une estimation de la variance de l'erreur  $\Delta C_i$  est donné par la relation suivante [LEH83, FUL87] :

$$\rho = \frac{\|C - E\vec{V}\|_2^2}{n - 2} \quad (C21)$$

L'effet usuel des termes  $\Delta A_i, \Delta B_i$  est de produire une surestimation sur la grandeur  $\vec{V}$ . Le biais influe aussi sur la direction de l'estimé [FUL87, GLE81, STO90]. Nous allons considérer les deux cas suivants pour l'estimation du biais :

- Dans le premier cas on suppose que le bruit est symétrique autour des valeurs réelles.
- Dans le deuxième, le bruit est symétrique autour des valeurs estimées.

#### • Pour du bruit symétrique autour des valeurs réelles

Étant donné les gradients théoriques  $\hat{E}$  et  $\hat{C}$ , les distributions de bruits  $\Delta E = E - \hat{E}$  et  $\Delta C = C - \hat{C}$  sont supposées symétriques, mais pas nécessairement gaussienne. Dans ce cas, pour une série de mesures importante, il y aura une décroissance du biais. L'estimé peut s'écrire sous la forme :

$$\vec{V} = \left[ (\hat{E}' + (\Delta E)') (\hat{E} + \Delta E) \right]^{-1} (\hat{E} + \Delta E)' (\hat{C} + \Delta C) \quad (C22)$$

S'il y a corrélation entre le bruit temporel  $\Delta C$  et le bruit spatial  $\Delta E$ , cette corrélation peut affecter la valeur de  $\vec{V}$ . En revanche, si nous avons les conditions suivantes :

- ces deux termes sont indépendants et considérés d'être faibles, leur produit est négligeable,
- $\Delta C$  est indépendant de  $\hat{E}$ ,
- $\Delta E$  est indépendant de  $\hat{C}$ ,

alors la solution au sens des moindres carrés peut être estimée par l'équation (C23) :

$$\vec{V} = \left[ \left( \hat{E}' + (\Delta E)' \right) \left( \hat{E} + \Delta E \right) \right]^{-1} \hat{E}' \hat{C} \quad (C23)$$

L'expression (C23) peut être écrite sous la forme :

$$\vec{V} = \left[ \hat{E}' \hat{E} + (\Delta E)' \Delta E + \hat{E}' \Delta E + (\Delta E)' \hat{E} \right]^{-1} \hat{E}' \hat{C} \quad (C24)$$

Dans le cas où il y a suffisamment d'équations, l'expression (C24) peut être approchée par :

$$\vec{V} = \left[ \hat{E}' \hat{E} + (\Delta E)' \Delta E \right]^{-1} \hat{E}' \hat{C} \quad (C25)$$

Notons au passage que pour les matrices définies positives  $A$ ,  $B$  et  $C$ , si  $A > B$ , nous avons  $A^{-1} < B^{-1}$  et  $AC > BC$ . A cet effet, nous pouvons écrire :

$$\left[ \hat{E}' \hat{E} + (\Delta E)' \Delta E \right]^{-1} < \left( \hat{E}' \hat{E} \right)^{-1} \quad (C26)$$

Et par la suite, nous avons :

$$\left[ \hat{E}' \hat{E} + (\Delta E)' \Delta E \right]^{-1} \hat{E}' \hat{C} < \left( \hat{E}' \hat{E} \right)^{-1} \hat{E}' \hat{C} \quad (C27)$$

L'estimé est donc inférieur à la valeur théorique.

### • Le bruit est symétrique autour des valeurs estimées des gradients

Les différentes erreurs  $\Delta A_i$ ,  $\Delta B_i$  et  $\Delta C_i$  sont supposées indépendantes et distribuées identiquement. La distribution de probabilité du bruit est gaussienne.

- Pour le cas monodimensionnel, c'est à dire  $B_i = 0$ , nous avons :

$$\hat{A}_i \hat{u} = \hat{C}_i \quad (C28)$$

En supposant que les  $\Delta C_i$  sont nuls, c'est à dire que  $C = \hat{C}$ . Nous avons ainsi :

$$\hat{A}_i - A_i = \frac{C_i - A_i \hat{u}}{\hat{u}} \quad (C29)$$

Pour deux vecteurs  $A$  et  $B$  donnés, nous cherchons la probabilité pour que le flot réel dans la direction  $x$  ait une certaine valeur  $\hat{u}$ . Il serait commode d'argumenter en terme d'énergie mais pas en terme de probabilité. Si  $P$  est la probabilité d'un événement, par définition, l'énergie  $T$  est donnée par la relation :

$$P = k_1 e^{-k_2 T} \quad (C30)$$

Les termes  $k_1$  et  $k_2$  sont des constantes. Nous avons introduit la notion d'énergie dans le paragraphe intitulé estimation au sens des moindres carrés. Dans le cas présent, l'énergie ou la fonctionnelle peut être exprimée par la relation (C31).

$$\sum_i \left( \frac{C_i - \hat{A} \hat{u}}{\hat{u}} \right)^2 = \sum_i \frac{(C_i - A_i u)^2 + (A_i u - A_i \hat{u})^2 + 2 A_i (u - \hat{u})(C_i - A_i u)^2}{(\hat{u})^2} \quad (C31)$$

Par définition de la solution au sens des moindres carrés, nous avons :

$$\sum_i A_i C_i = \sum_i A_i A_i u \quad (C32)$$

La fonction d'énergie est une fonction monotone croissante de  $\frac{T + \tau^2 (u - \hat{u})^2}{(\hat{u})^2}$  [FER99] où

$T = \frac{1}{n} \sum_i (C_i - A_i u)^2$  et  $\tau = A_i^2$  sont des termes qui ne dépendent pas de  $\hat{u}$ . Si les coordonnées sont choisies telles que  $u > 0$ , alors pour  $s > 0$ , la fonction d'énergie associée à  $\hat{u} = u + s$  est inférieure à celle associée  $\hat{u} = u - s$ .

- Considérons maintenant le cas bidimensionnel. Choisissons les coordonnées telles que  $v = 0$ . Cette hypothèse nous permet d'analyser le terme d'énergie dans la direction suivant  $x$ . Admettons également que  $\Delta C = \Delta B = 0$ .



Nous avons :

$$\hat{A}_i = \frac{C_i - B_i \hat{v}}{\hat{u}}. \quad (\text{C33})$$

La fonction d'énergie est donnée par :

$$\sum_i \left( \frac{C_i - A_i \hat{u} - B_i \hat{v}}{\hat{u}} \right)^2 = \sum_i \frac{(C_i - A_i u - B_i \hat{v})^2 + (A_i u - A_i \hat{u})^2 + 2A_i (u - \hat{u})(C_i - A_i u - B_i \hat{v})^2}{(\hat{u})^2} \quad (\text{C34})$$

La fonction d'énergie peut être écrite sous la forme  $\frac{T + \tau^2 s^2 + Q \hat{v} s}{(\hat{u})^2}$  où  $T$ ,  $\tau$  et  $Q$  ne sont des paramètres qui ne dépendent ni de  $\hat{u}$  ni de  $\hat{v}$ . A cet effet, nous faisons l'analyse paramétrique en fixant les valeurs de  $|\hat{v}|$  et  $|s|$ . Comme précédemment, choisissons les coordonnées telles que  $u > 0$ .

La fonction d'énergie est une fraction. Selon le signe de  $s$ , nous avons :

- Deux possibilités pour le numérateur :  $q_1 = T + \tau^2 s^2 + |Q \hat{v} s|$  et  $q_2 = T + \tau^2 s^2 - |Q \hat{v} s|$  ;

- Deux possibilités pour le dénominateur :  $r_1 > r_2 > 0$ .

Considérons le cas de  $s < 0$ , l'énergie peut être une fraction  $q_1/r_1$  ou  $q_2/r_1$ . Si on remplace le dénominateur  $r_1$  par  $r_2$ , on obtient les possibilités qui se présentent quand  $s > 0$ , mais si on remplace le dénominateur  $r_1$  par  $r_2$ , on fait augmenter l'énergie. Cependant, il est moins probable que  $s$  soit positive que négative. Cet argument assure que les valeurs de  $|\hat{v}|$  et  $|s|$  doivent être fixées, mais aussi, il n'implique pas qu'elles soient égales. Par conséquent,  $s$  est plus probablement négatif que le contraire et nous avons ainsi une sous-estimation.

Nous pouvons généraliser le cas avec  $\Delta B_i \neq 0$  et  $\Delta C_i \neq 0$ . Selon [FER99], nous pouvons avoir une fonction d'énergie de la forme  $R + \frac{1}{\hat{u}^2} \left[ T + \tau^2 s^2 + Q \hat{v} s + \left( \sum_i A_i \delta C_i \right) s \right]$  où  $R$ ,  $T$ ,  $\tau$  et

$Q$  ne dépendent pas de  $\hat{u}$  et  $\hat{v}$ . Une analyse paramétrique qui consiste à fixer temporairement les valeurs de certaines quantités peut être utilisée. C'est à dire, nous avons besoin de fixer les valeurs de  $|\hat{v}|$  et  $|s|$ , et de trouver des vecteurs de bruit  $K_3$ ,  $K_4$ , et de supposer que  $|\Delta B| = K_3$  (c'est à dire  $\Delta B = \pm K_3$ ) et  $|\Delta C| = K_4$ . Selon le signe de  $s$ ,  $\Delta B$  et  $\Delta C$ , nous pouvons faire la même analyse que précédemment et nous pouvons montrer qu'une surestimation est plus probable qu'une sous-estimation.

#### d) Erreur angulaire

Les performances peuvent être quantifiées par l'écart angulaire entre l'estimé et le flot réel [FLE92]. Parce que la vitesse  $\vec{V}(u, v)$  peut être représentée comme une direction spatio-temporelle  $\vec{V}(u, v, 1)$ , nous pouvons mesurer l'erreur angulaire en considérant la vitesse comme un vecteur dans l'espace spatio-temporel. Dans ce cas, l'estimé  $V$  est représenté par

$$\vec{V} \equiv \frac{1}{\sqrt{u^2 + v^2 + 1}}(u, v, 1) \quad (\text{C35})$$

La vitesse réelle peut aussi être représentée par :

$$\hat{V} \equiv \frac{1}{\sqrt{\hat{u}^2 + \hat{v}^2 + 1}}(\hat{u}, \hat{v}, 1) \quad (\text{C36})$$

L'erreur angulaire est donnée par :

$$\psi \equiv \arccos(\hat{v}, v) \quad (\text{C37})$$

Cette mesure permet de s'affranchir des problèmes causés par les déplacements de très forte amplitude qui feraient exploser les valeurs si on utilisait l'erreur quadratique moyenne<sup>27</sup> ou l'erreur absolue moyenne<sup>28</sup>.

Nous n'allons pas analyser en détail tous les paramètres qui influent sur cette erreur angulaire. Dans [FLE92], les auteurs précisent que la valeur de l'erreur est moins biaisée dans le cas d'une vitesse extrêmement faible ou une vitesse extrêmement forte.

## C.2 Conséquence des erreurs dans la mise en équation : le besoin de contraintes additionnelles

Comme expliqué antérieurement, l'estimation des gradients est nécessaire en méthodes différentielles, sur des images discrètes. Différents types d'approximation peuvent être utilisés. Plus l'approximation est simple, plus le coût de calcul exigé sera faible, d'où son importance pour les applications en temps réel. Les erreurs dues à l'approximation des gradients font que les droites de contraintes ne présentent pas un point d'intersection unique. La figure C.2 montre la solution réelle du système de  $n$  équations de contraintes ( $n$  pixels). Il est évident qu'on n'a pas de solution unique si plusieurs intersections apparaissent. La solution du système sera ainsi biaisée et nous nous retrouvons dans cas de la figure C.3.

$$\begin{cases} (ECM_1): uI_x^{(1)} + vI_y^{(1)} + I_t^{(1)} = 0 \\ (ECM_2): uI_x^{(2)} + vI_y^{(2)} + I_t^{(2)} = 0 \\ \vdots \\ (ECM_n): uI_x^{(n)} + vI_y^{(n)} + I_t^{(n)} = 0 \end{cases} \quad (\text{C38})$$

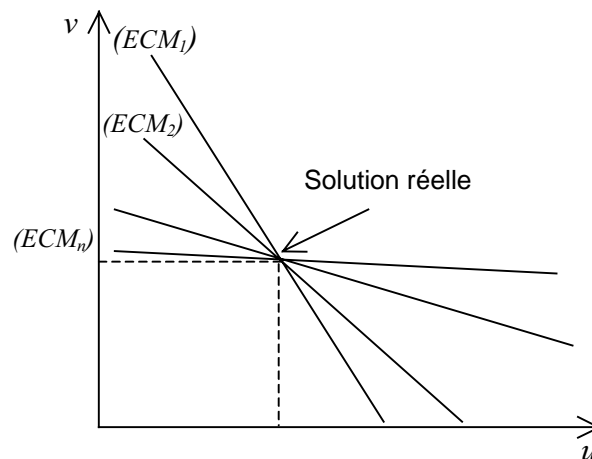


Figure C.2 : Représentation graphique des droites de contraintes et le point d'intersection associé à la solution réelle du système.

<sup>27</sup> Mean square error

<sup>28</sup> Mean absolute error

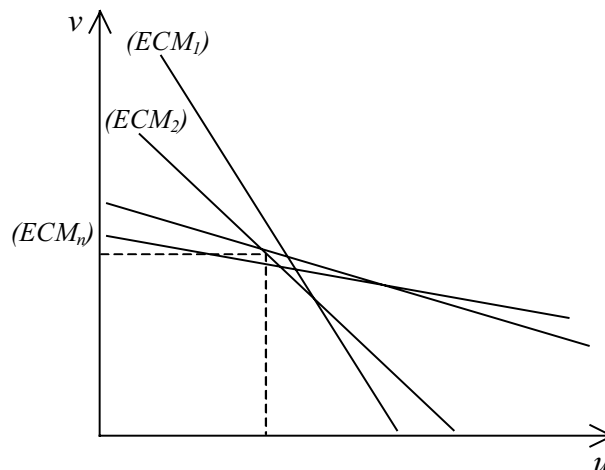


Figure C.3 : Plusieurs points d'intersection, effet des différentes sources d'erreurs lors de la résolution du système d'équations de contrainte du mouvement.

### C.3 Critère de choix du support de calcul

Dans les méthodes à contraintes locales, ce problème explique comment définir un bon support de calcul. Prenons un exemple tel que le support ne contienne que des points à luminance constante. Localement, il ne pourrait pas être défini sur l'autre image étant donné qu'il est homogène le long de toutes les directions. Si le support a suffisamment de gradients le long d'une direction, il est possible de localiser sa position sur l'autre trame, uniquement le long de cette direction. Ces deux aspects illustrent encore le problème d'ouverture. Alors que, dans les deux directions, le support de calcul choisi doit présenter suffisamment de gradients. Prenons un cas plus proche du cas théorique où la luminance est continue dans l'échelle spatio-temporelle (cas de la spirale). L'approximation des gradients discrets peuvent donner des valeurs proches des valeurs théoriques et, par conséquent, un faible nombre de points sont instables dans le système d'équation. Dans ce cas, seulement à partir d'un faible support de calcul, l'estimation est capable de produire des résultats stables. Nous avons déjà mentionnée cette remarque dans le test précédent. En effet, dans la zone où la taille du support est très faible, il n'y a pas suffisamment des points pour contraindre le système d'équation et les résultats ne sont pas stables. De même, si la taille est trop grande (si par exemple, elle dépasse la surface de l'objet), tous les points de la zone d'intérêt ne sont plus prépondérants par rapport à ceux de la surface externe. Nous risquons d'avoir une mauvaise estimation. Entre ces deux zones, les estimés sont stables.

En pratique, la fonction de luminance n'est pas forcément continue et par conséquent, il y a toujours quelques points qui sont instables au conditionnement de l'équation linéaire. Donc, le problème ne vient pas réellement du nombre de points de la matrice mais il vient de la relation entre les points de la matrice. Différentes valeurs aboutissent à des estimations différentes. Nous pouvons donc en déduire que la détérioration vient non pas du changement de taille mais du changement de points. L'équation suivante permet de savoir si le support de calcul est bien défini [FEN98].

$$\sigma_{\min} \left( \sum_{\Omega} \begin{bmatrix} I_x & I_y \\ I_x & I_y \end{bmatrix} \right) > \tau \quad (\text{C39})$$

Le terme entre parenthèses est une matrice carrée correspondant aux carrés des gradients

---

spatiaux. Ce terme s'appelle somme des différences carrées<sup>29</sup>. Le symbole  $\sigma_{\min}$  note la plus petite valeur singulière,  $\tau$  est un seuil et  $\mathcal{Q}$  un voisinage donné.

#### **C.4 Critère de choix des techniques d'estimation du gradient**

Ce sujet sera analysé plus en détail dans la partie simulation. Dans le cadre d'une image fortement bruitée, le choix du gradient peut jouer un rôle assez important sur les résultats d'estimation. En général, nous remarquons que le choix d'une technique d'approximation du gradient est moins important que le choix de l'algorithme d'estimation du mouvement utilisée. A titre d'exemple, les résultats obtenus avec le masque de Prewitt et celui de Sobel pourront être différents pour une résolution au sens des moindres carrés classiques. A partir du moment où on ajoute des termes de régularisation à l'équation de contrainte, cette différence ne sera plus remarquable.

---

<sup>29</sup> Sum of squared difference



# **Bibliographie**



- [ADE85] E. H. Adelson and J. R. Bergen, "Spatiotemporal energy models for the perception of motion", *Journal of the Optical Society of America A*, Vol. 2, N°1, pp. 284-299, January 1985.
- [ADE86] E. H. Adelson and J. R. Bergen, "The extraction of spatio-temporal energy in human and machine vision", in *Proceedings of IEEE Workshop on Motion : Representation and Analysis*, pp. 151-156, Charleston, S Carolina, 1986.
- [ADE90] E. H. Adelson and P. Anandan, "Ordinal characteristics of transparency", *AAAI-90 Workshop on Qualitative Vision*, July 20 1990, Boston, MA.
- [ADI85] G. Adiv, "Determining three-dimensional motion and structure from optical flow generated by several moving objects", *IEEE Trans. Pattern Anal. And Machine Intell.*, Vol. PAMI-7, N°4, pp. 384-401, July 1985.
- [ADI89] G. Adiv, "Inherent ambiguities in recovering 3D motion and structure from a noisy flow field", *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. PAMI-11, N°5, pp. 477-489, May 1989.
- [AGG81] J. K. Aggarwal and W. N. Martin, "Analyzing dynamic scenes containing multiple moving objects", in T. S. Huang, editor, *Image Sequence Analysis*, pp. 255-380, Springer-Verlag, New-York, 1981.
- [AGG88] J. K. Aggarwal and N. Nandhakumar, "On the computation of motion from sequences of images-a review", *Proceedings of the IEEE*, Vol.76, N°8, pp. 917-925, August 1988.
- [ANA93] P. Anadan, J. R. Bergen, K. J. Hanna and R. Hingorani, "Hierarchical model-based motion estimation", in M. I. Sezan and R. L. Lagendijk, editors, *Motions Analysis and Image Sequence Processing*, pp. 1-22, Kluwer academic Publishers, 1993.
- [AYE93a] S. Ayer and P. Schroeter, "Hierarchical robust motion estimation for segmentation of moving objects", in *IEEE Workshop on Image and Multidimensional Signal Processing*, pp. 122-123, Cannes, France, September 1993.
- [AYE93b] S. Ayer, P. Schroeter and J Bigün, "tracking based on hierarchical multiple motion estimation and robust regression", in *Time Varying Image Processing and Moving Objects Recognitions*, Florence, Italy, June 1993.
- [AYE95a] S. Ayer and H. S. Sawhney, "Layered representation of motion video using robust maximum-likelihood estimation of mixture models and MDL encoding", in *Fifth International Conference on Computer Vision*, pp. 777-784, Boston, MA, USA, June 1995.
- [AYE95b] S. Ayer, "Sequential and competitive methods for motion of multiple motions", Thèse présentée à la section de systèmes de communication, Ecole Polytechnique de Lausanne, 1995.
- [BAL82] D. H. Ballard and C. M. Brown, *Computer Vision*, Englewood Cliffs, NJ : Prentice-Hall, 1982



- [BAR92] J. L. Barron, D. J. Fleet, S. S. Beauchemin and T. A. Burkitt, "Performance of optical flow techniques", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 236-242, Champaign, Illinois, June 15-18 1992.
- [BAR92] J.L. Barron, D.J. Fleet and S. S. Beauchemin, "Performance of optical flow techniques", Technical Report 299, Queens University, Outario, Canada, 1992.
- [BAR94] J. L. Barron, S. S. Beauchemin, D. J. Fleet, "On Optical Flow", *AIICSR*, Bratislava, Slovakia, p. 3-14, September 1994.
- [BAR94] J.L. Barron, D.J. Fleet, and S.S Beauchemin, "Performance of optical flow techniques", *The International Journal of Computer Vision*, Vol. 12 N°1, pp. 43-77, 1994.
- [BER00] Yannick Berthoumieu, Benoît Escrig, "Estimation de retard en sismique répétitive", Rapport de Contrat GDF/IXL/ADERA, Laboratoire Étude de l'Intégration des Composants et Systèmes Electroniques IXL, Décembre 2000.
- [BER90] J.R. Bergen, P. J. Burt, R. Hingorani and S. Peleg, "Computing two motions from three frames", in *Proceedings of the Third International Conference on Computer Vision, ICCV'90*, Osaka, Japan, pp. 27-32, 1990 .
- [BES86] J. Besag, "On statistical analysis of dirty pictures", *Journal of the Royal Statistical Society B*, Vol. 48, N° 3, pp. 259-302, 1986.
- [BIG91] J. Bigün, G. Granlud and J. Wibelund, "Multidimensional orientation estimation with applications to texture analysis and optical flow", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 13, pp. 775-790, August 1991
- [BLA92] M. J. Black, "Robust incremental optical flow", PhD thesis, Yale University, Computer Science Department, September 1992.
- [BLA93] M. J. Black and P. Anadan, "A framework for the robust estimation of optical flow", in *ICCV'94*, pp. 231-236, Berlin, Germany, May 1993.
- [BOL95] Ph. Bolon, J. M. Chassery, J. P. Cocquerez, D. Demigny, C. Graffigne, A. Montanvert, S. Philipp, R. Zéboudj, J. Zérubia, "Analyse d'images : filtrage et segmentation", Enseignement de la physique, Traitement du signal, Edition Masson, Paris, 1995.
- [BOU87] P. Bouthemy and J. Santillana Rivero, "A hierarchical likelihood approach for region segmentation according to motion-based criteria", in *ICCV'87*, pp. 463-467, London, UK, 1987.
- [BOU93] P. Bouthemy and E. Francois, "Motion segmentation and qualitative dynamic scene analysis from an image sequence", *International Journal of Computer*, Vol. 10, N°2, pp. 157-182, 1993.
- [BUR84] P. J. Burt, "The pyramid as a structure for efficient computation", in A. Rosenfeld, editor, *Multiresolution Image Processing and Analysis*, pp. 6-35, Springer-Verlag, 1984.

- [BUR89] P. J. Burt, J. R. Bergen, R. Hingorani, R. Kolczynski, W. A. Lee, A. Leung, J. Lubin and H. Shvaytser, "Object tracking with a moving camera, an application of dynamic motion analysis", in *IEEE Proc. Workshop on Visual Motion*, pp. 2-12, Irvine, CA, March 1989.
- [BUR91] P. J. Burt, J. R. Bergen, R. Hingorani, R. Kolczynski, W. A. Lee, A. Leung, J. Lubin and H. Shvaytser, "Object tracking with a moving camera, an application of dynamic motion analysis", in *IEEE Proc. Workshop on Visual Motion*, pp. 2-12, Irvine, CA, March 1989.
- [CAF83] C. Cafforio and F. Rocca, "The differential method for image motion estimation", *Image Sequence Processing and Dynamic Scene Analysis*, pp. 104-124, Berlin, April 1983.
- [CAM90] M. Campani and A. Verri, "Computing optical flow from an overconstrained system of linear algebraic equations", in *Proceedings of the Third International Conference on Computer Vision, ICCV'90*, pp.22-26, Osaka, Japan, 1990.
- [CAR87] G. C. Carter, "Coherence and time delay estimation", *Proc. IEEE*, Vol. 75, pp. 236-255, Feb. 1987.
- [COH93] J. Cohen, "Nonlinear variational method for optical flow computation", in *Proceedings 8<sup>th</sup> SCIA*, Vol. 1, pp. 523-530, 1993.
- [DEP92] R. Depommier and E. Dubois, "Motion estimation with detection of occlusion areas", in *Proceedings IEEE Int. Conf. Acoustics Speech Signal Processing*, Vol. 3, pp. 269-272, Mars 1992.
- [DER87] R. Deriche, "Using Canny's Criteria to Derive a Recursively Implemented Optimal Edge Detector", *International Journal of Computer Vision*, Vol. 1, N°2, pp. 167-187, 1987.
- [DIE91] N. Diehl, "Object-oriented motion estimation and segmentation in image sequence", *Signal Process., Image Commun.*, Vol. 3, pp. 23-56, February 1991.
- [DRI91] J. N. Driessen and J. Biemond, "Motion field estimation for complex scenes", in *Proc. SPIE Visual Communications and Image Process.*, pp. 511-521, 1991.
- [DUC94] B. Duc, "Motion estimation using invariance under group transformations", in *12<sup>th</sup> International Conference on Pattern Recognition*, pp. 159-163, Jerusalem, October 1994.
- [DUC95] B. Duc, P. Schroeter and J. Bigün, "Spatio-temporal robust motion estimation and segmentation", in *6<sup>th</sup> International Conference on Computer Analysis of Images and patterns*, pp. 238-245, Prague, September 5-8 1995.
- [DUF95] F. Dufaux, F. Moscheni and A. Lippman, "Spatio-temporal segmentation based on motion and static segmentation", in *Proc. ICIP'95*, Vol. 1, pp. 306-309, Washington, DC, October 1995.
- [DUM99] C. Dumontier, F. Luthon and J. P. Charras, "Real Time DSP Implementation for MRF-Based Video Motion Detection", *IEEE Transactions On Image Processing*, Vol. 8, N° 10, October 1999.

- [FEN98] Xiaolin Feng and Pietro Perona, "Real Time Motion Detection System and Scene Segmentation", CDS Technical Report : CIT-CDS-98-004, California Institute of Technology, Pasadena, CA 91125, March 11 , 1998.
- [FER99] C. Fermüller, D. Shulman and R. Pless, "The Statistics of Optical Flow", Center for Automation Research, University of Maryland, College Park, MD 20742-3275, November 1999.
- [FLE90] D. J. Fleet and A. D. Jepson, "Computation of component image velocity from local phase information", *International Journal of Computer Vision*, Vol. 5, N°1, pp. 77-104, 1990.
- [FLE92] D. J. Fleet, "Measurement of Image Velocity", Luwer Academic Publishers, Norwell, 1992.
- [FLE93] D. J. Fleet and A. D. Jepson, "Stability of phase information", *IEEE PAMI (in press)*, 1993.
- [FOS98] Dean P. Foster and Robert A. Stine, "Local Asymptotics and the Minimum Description Length", The Wharton School of the University of Pennsylvania, Philadelphia, PA 19104-6302, March 27, 1998.
- [FUL87] W. Fuller, *Measurement Error Models*, Wiley, 1987.
- [GEL96] M. Gelgon, P. Bouthemy, "A region-level graph labeling approach to motion-based segmentation", Publication interne N°1070, IRISA, Rennes, France, Décembre 1996.
- [GEM84] G. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the bayesian restoration of images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 6, N° 5, pp. 721-741, November 1984.
- [GIO85] A. A. Giordano and F. M. Hsu, "Least Square Estimation with Applications to Digital Signal Processing", A Wiley-Interscience Publication, John Wiley & Sons, Inc. New York, 1985.
- [GLE81] L. J. Gleser, "Estimation in a multivariate « errors in variables » regression model : Large sample results", *Annals of Statistics I*, 9 : 24-44, 1981.
- [GON90] S. Gong and M. Brady, "Parallel computation of optical flow", In *Proc. 1<sup>st</sup> European Conference en Computer Vision*, pp. 124-133, Antibes, France, 1990.
- [GU96a] C. Gu, "Multivalued morphology and segmentation-based coding", PhD thesis, Swiss Federal Institute of technology, Lausanne, 1996.
- [GU96b] H. Gu, Y. Shirai and M. Asada, "MDL-based segmentation and motion modeling in a long image sequence of scene with multiple independently moving objects", *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. PAMI-18, N°1, pp. 58-64, January 1996.
- [HAR83] R. M. Heralick and J. S. Lee, "The facet approach to optic flow, In L. Baumann, editor", in *Proceedings Image Understanding Workshop*, pp. 84-93, Science Applications, Arlington, VA, 1983.

- [HEE87] D. J. Heeger, "Model for the extraction of image flow", *Jouranal of the Opical Society of America A.*, Vol. 4, N° 8, pp. 1455-1471, 1987.
- [HEE88] D. Heeger, "Optical flow using spatiotemporal filters", *International Journal of Computer Vision*, vol. 1, pp. 279-302, 1988.
- [HEI90] F. Heitz and P. Bouthemy, "Multimodal motion estimation and segmentation using Markov random fields", in *Proc. IEEE Int. Conf. Pattern Recognition*, pp. 378-383, june 1990.
- [HEI93] F. Heitz and P. Bouthemy, "Multimodal estimation of discontinuous optical flow using Markov random fields", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, N°12, pp. 1217-1232, December 1993.
- [HEI94] J. Heikkonen, Subsymbolic Representations, "Self-Organizing Maps and Object Motion Learning", PhD Technical Thesis, Lappeenranta University of Technology, Finland, 1994.
- [HEE96] D. J. Heeger, "Notes on Motion Estimation", Psych 267, CS 348D, EE 365, October 30, 1996.
- [HIL84] E. C. Hildreth, "Computaion underlying the measurement of visual motion", *Artificial Intelligence*, N°23, pp. 309-354, February 1984.
- [HOL77] P.W. Holland and R.E. Welsch, "Robust regression using iteratively reweighted least-squares", *Commun. Stat. Theor. Meth.*, A6, pp. 813-828, 1977.
- [HOR81] B. K. P. Horn and B. G. Schunck, "Determining optical flow", *Artificial Intell.*, Vol. 17, pp. 185-204, 1981.
- [HOR86] B. K. P. Horn, *Robot Vision*, pp. 278-294, The MIT Press, 1986.
- [IRA92] M. Irani, B. Rousso and S. Peleg, "Detecting and tracking multiple moving objects using temporal integration", in Sandini G., editor, *Second European Conference on Computer Vision*, pp. 282-287. Springer-Verlag, S. Margherita, Italy 1992.
- [IRA94] M. Irani, B. Rousso and S. Peleg, "Computing Occluding and Transparent Motions", Institute of Computer Science, The Hebrew University of Jerusalem, 91904 Jerusalem, ISRAEL.
- [IZQ00] D. Izquierdo, Y. Berthoumieu, "Estimation-segmentation en poursuite spatio-temporelle", rapport de DEA, Université Bordeaux I / ENSERB, Laboratoire d'Automatique et de Productique, 1999-2000
- [JAH93] B. Jähne, "Spatio-Temporal Image Processing", N° 751 in lecture Notes in Computer Science, Springer-Verlag, 1993.
- [JAH97] B. Jähne, "Digital Image Processing", Springer-Verlag, Berlin Heidelberg, 1997.

- [JEP93] A. D. Jepson and M. J. Black, "Mixture models for optical flow computation", in *IEEE Proc. Conf. on Computer Vision and Pattern Recognition I*, pp. 760-761, New-York, USA, June 1993.
- [KAR97] M. Karczewicz, J. Nieweglowski, P. Haavisto, "Video coding using motion compensation with polynomial motion vector fields", *Signal Process., Image Commun.*, Vol. 10, pp. 63-91, July 1997.
- [KAU90] L. Kaufman and P. J. Rousseeuw, "Finding groups in data", John Wiley & Sons, Inc., New York, 1990.
- [KEA83] J. K. Kearney, "Gradient-based estimation of optical flow", PhD dissertation, Dep. Comput. Sci., Univ. Minnesota, Aug. 1983.
- [KEA87] J. K. Kearney, W. B. Thompson and D. L. Boley, "Optical Flow Estimation : An Error Analysis of Gradient-Based Methods with Local Optimization", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 9, N°2, 1987.
- [KER95] C. Kervann and F. Heitz, "A Markov random field model-based approach to unsupervised texture segmentation using local and global spatial statistics", *IEEE Trans. On Image Processing*, Vol. 4, N°6, pp. 856-862, June 1995.
- [KIR83] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing", *Science*, Vol. 220, 1983.
- [LAC97] J. L. Lacoume, P. O. Amblard, P. Comon, "Statistiques d'ordre supérieur pour le traitement du signal", Masson, Paris, 1997.
- [LAM92] Lampinen, "Neural pattern recognition : Distorsion tolerance by self-organizing maps", PhD Tech. Thesis, Lappeenranta University of Technology, Lappeenranta, Finland
- [LEC89] Yvan G. Leclerc, "Constructing Simple Stable Descriptions for Image Partitioning", *International Journal of Computer Vision*, Vol. 3, pp. 73-102, 1989.
- [LEH83] E. H. Lehmann, *Theory of Point Estimation*, Wiley, 1983.
- [LUC81] B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision", *Proc. DARPA Image Understanding Workshop*, pp. 121-130, 1984.
- [MAR82] D. Marr, *Vision*, W. H. Freeman New York, 1982.
- [MAR87] J. Marroquin, S. Mitter and T. Poggio, "Probabilistic solution of ill-posed problems in computational vision", *Journal of the American Statistical Association*, Vol. 82, pp. 76-89, March 1987.
- [MAR95] F. Marqués, M. Pardàs and P. Salembier, "Coding-oriented segmentation of video sequences", in L. Torres and M. Kunt, editors, *Video Coding : The Second Generation Approach*, pp. 79-123, Kluwer Academic Publishers, 1995.

- [MAR97] J. L. Mari, F. Glangeaud, F. Coppens, "Traitement du signal pour géologues et géophysiciens", Editions Technip, Paris, 1997.
- [MAX87] J. Max, "Méthodes et Techniques de Traitement du signal et application aux mesures physiques", Tome I et II, Masson 4<sup>ème</sup> Edition, 1987.
- [MEE91] P. Meer, D. Mintz and A. Rosenfeld, "Robust regression methods for computer vision : a review", *International Journal of Computer Vision*, 6 (1), pp. 59-70, 1991.
- [MET53] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equations of state calculations by fast computing machines", *J. Chem. Phys.*, Vol. 21, pp. 1087-1094, 1953.
- [MOL91] C. R. Moloney and E. Dubois, "Estimation of motion fields from image sequences with illumination variation", in *Proc. ICASSP'91*, Vol. 4, pp. 2425-2428, Toronto, Canada, May 1991.
- [MUS85] H. G. Musmann, P. Pirsch and H. J. Grallert, "Advances in picture coding", *Proceedings IEEE*, Vol. 73, N° 4, pp. 523-548, April 1985.
- [NAG83] H. H. Nagel, "Displacement vectors derived from second-order intensity variations in image sequences", *Computer Vision, Graphics, and Image Processing*, Vol. 21, pp. 85-117, 1983.
- [NAG86] H. H. Nagel and W. Enkelmann, "An investigation of smoothness constraint for the estimation of displacement vector field from images sequences", *IEEE transactions on Pattern Analysis and Machine Intelligence*, Vol. 8, pp. 565-593, 1986.
- [NAG87] H. H. Nagel, "On the estimation of optical flow : relations between different approaches and some new results", *Artificial Intelligence*, Vol. 33, pp. 299-324, 1987.
- [NAG89] H. H. Nagel, "On constraint equation for the estimation of displacement rates in image sequences", *IEEE Trans. Patt. Anal. Mach. Intell.*, Vol. 11, pp.13-30, 1989
- [NAG92] H. H. Nagel, "Direct estimation of optical flow and its derivatives", in G. A. Orban and H. H. Nagel, editors, *Artificial and Biological Vision Systems*, Basic Research Series, pp. 191-224, Springer-Verlag, 1992.
- [NAM96] S. H. Nam and M. K. Lee, "Flexible VLSI Architecture of Motion Estimator for Video Image Compression", *IEEE transactions on Circuits and Systems -II : Analog and Digital Signal Processing*, Vol. 43, N°6, June 1996.
- [NEG93] S. Negahdaripour and C. H. Yu, "A generalized brightness change model for computing optical flow", in *IEEE Proc. Int. Conf. on Computer Vision I*, pp. 2-11, Berlin, Germany, May 1993.
- [NET79] A. Netravali and J. D. Robbins, "Motion compensated television coding part I", *Bell Syst. Tech. Journal*, Vol. 58, N°3, pp. 629-668, 1979.

- [NIC92] H. Nicolas, "Hiérarchie de modèles de mouvement et méthodes d'estimation associées, Application au codage de séquences d'images", Thèse présentée devant l'Université de Rennes I, September 1992.
- [NIK87] C. L. Nikias, M. R. Raghuvver, "Bispectrum Estimation, A digital Signal Processing Framework," *Proc. IEEE*, Vol. 75, N°7, pp. 869-891, July 1987.
- [NIK88] C. L. Nikias, R. Pan, "Time Delay Estimation in Unknown Gaussian Spatially Correlated Noise", *IEEE Trans. Acoust., Speech, Signal Processing*, Vol. 36, N°11, pp. 1704-1714, November 1988.
- [ODO94] J. M.M Odobez and P. Bouthemy, "Robust multiresolution estimation of parametric models in complex image sequences", in *7<sup>th</sup> European Conference on Signal Processing, EUSIPCO '94*, Edinburgh, Scotland, September 1994.
- [ODO98] Odobez J. M., Bouthemy P., Fleuet E., "Suivi 2D des pièces métalliques en vue d'un asservissement visuel", *RFLA '98*, Vol 2, pp. 173-192, Clermont-Ferrand, France, Janvier 1998.
- [ORK96] M. Orkisz and P. Clarysee, "Estimation du flot optique en présence de discontinuités", *Traitement du Signal*, Vol. 13, N°5, pp.489-513, 1996.
- [OTT94] M. Otte and H. H. Nagel, "Optical flow estimation : Advances and comparaisons", in Jan-Olof Eklundh, editor, *Proceedings of the 3<sup>rd</sup> European Conference on Computer Vision*, Vol. 800 of Lecture Notes in Computer Science, pp. 51-70, Springer-Verlag, 1994.
- [PAR94] M. Pardàs and P. Salembier, "3D morphological segmentation and motion estimation for image sequences", *Signal Processing*, Vol. 38, N° 2, pp. 31-43, September 1994.
- [PAU99] F. Pauget, "Utilisation du Flux Optique pour l'Aide à l'Interprétation d'Images Sismiques", Thèse, Spécialité Sciences Cognitives, Université Paris VI, 1999.
- [PEL90] S. Peleg and H. Rom, "Motion based segmentation", in *IEEE Int. Conference on Pattern Recognition*, pp. 109-113, 1990.
- [PEL96] D. Pellerin, A. Spinéi et A. Guérin-Dugué, "Calcul du flot optique par filtrage de Gabor combinés", *Traitement du Signal*, Vol 13, N°1, pp. 13-23, 1996.
- [PER91] P. Perez and F. Heitz, "Une approche multiéchelle à l'analyse d'images par champs markoviens", Technical report, INRA RENNES, Juillet 1991.
- [POG87] T. Poggio, V. Torre and C. Koch, "Computational vision and regularisation theory", *Nature*, Vol. 317, N° 26, pp. 314-319, 1985.
- [PRO94] M. Proesmans, L. Van Gool, E. Pauwels, and A. Oosterlinck, "Determination of Optical Flow and its Discontinuities using Non-Linear Diffusion", in *Proceedings of the 3<sup>rd</sup> ECCV*, Vol. II, N° 801 in Lecture Notes in Computer Science, pp. 295-304, Springer-Verlag, 1994.

- [PUR87] A. Puri, H. M. Hang, and D. L. Schiling, "An efficient block-matching algorithm for motion compensated coding", in *Proceedings IEEE ICASSP' 87*, pp.25.4.1-25.4.4, 1987.
- [RAN00a] A. Randriantsoa, Y. Berthoumieu, "Optical flow estimation using forward-backward constraint equation", *IEEE International Conference on Image Processing, ICIP'2000*, Vol 2, pp. 578-582, Vancouver, Canada, Sept. 2000.
- [RAN00b] A. Randriantsoa, Y. Berthoumieu, P. Marchegay, "Robust total Least-Squares Method Using Forward and Backward Errors for Motion Estimation", *International Symposium in Image and Video Communications over Fixed and Mobile Networks, ISIVC'2000*, Vol. 2, pp. 212-219, Rabat, Morocco, Apr. 2000.
- [RAN00c] A. Randriantsoa, D. Izquierdo, Y. Berthoumieu, "Mouvement direct et rétrograde, applications en estimation et en segmentation vidéo", *Compression et Représentation des Signaux Audiovisuels, CORESA'2000*, pp. 179-187, France Télécom et Laboratoire IRCOM-SIC de l'Université de Poitiers, France, Oct. 2000.
- [RIS78] J. Rissanen, "Modeling by shortest description length", *Autommatica*, Vol. 14, pp. 465-471, 1978.
- [RIS83] J. Rissanen, "A Universal Prior for Integers and Estimation by Minimum Description Length", *The Annals of Statistics*, Vol. 11, N°2, pp. 416-4312, 1983.
- [RIS86] J. Rissanen, "Minimum description length", in *Encyclopedia of Statistics*, Kotz and Johnson, Ed, pp. Wiley, New York.
- [ROB67] E. A. Robinson, "Predictive decomposition of time series with application to seismic exploration", *Geophysics*, Vol. 32, pp. 418-484, 1967.
- [ROB83] J. D. Robbins and A. N. Netravali, "Recursive motion compensation : A review", in *Image sequence Processing and Dynamic Scene Analysis*, T. S. Huang, ed., pp. 76-103, Berlin, Germany : Springer-Verlag, 1983.
- [ROU87] P. J. Rousseeuw and A. M. Leroy, "Robust regression and outlier detection", John Wiley & Sons, Inc., New York, 1987.
- [SAL94] P. Salembier and M. Pardàs, "Hierarchical morphological segmentation for image sequence coding", *IEEE Trans. Image Process*, Vol. 3, N° 5, pp. 639-651, September 1994.
- [SAL95] P. Salembier, L. Torres, F. Meyer and C. Gu, "Region-based video coding using mathematical morphology", *Proc. IEEE*, Vol. 83, N° 6, pp. 843-857, June 1995.
- [SAM80] H. Samet, "Quadtree from boundary codes", *Communication of the ACM*, pp. 163-170, Mars 1980.
- [SAN94] H. Sanson, "Region-based motion analysis for video coding at low bit rates", MPEG-4, Paris, Mars 1994.
- [SAT77] T. Sato, K. Sasaki, "Bispectral holography", *J. Acoust. Soc. Amer.*, Vol. 62, pp. 404-408, 1977.



- [SCH94] P. Schroeter and S. Ayer, "Multi-frame based segmentation of moving objects by combining luminance and motion", in *Proc. EUSIPCO 94*, Edinburgh, U. K., September 1994.
- [SIM91] E.P. Simoncelli and E. H. Adelson, "Computing optical flow distributions using spatio-temporal filters", Technical Report 65, MIT Media Laboratory Vision and Modeling, March 1991.
- [SIM93] E. P. Simoncelli, "Distributed Representation and Analysis of Visual Motion", Vision and Modeling Group Technical Report #209, MIT Media Laboratory, E15-385, Cambridge, 1993.
- [STI92] C. Stiller, R. Suntrup, "Parametric object-motion estimation", *Proc. Int. Symp. On Information Theory and its Applications*, pp. 633-637, November 1992.
- [STI94] C. Stiller, "Object-oriented video coding employing dense motion fields", *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, Vol. 5, pp. 273-276, April 1994.
- [STI99] C. Stiller and J. Konrad, "Estimating Motion in Image Sequences", *IEEE Signal Processing*, Vol. 16, N°4, pp. 70-91, July 1999.
- [STO90] G. R. Stoner, T. D. Albright and V. S. Ramachandran, "Transparency and coherence in human motion perception", *Nature*, 344 : 153-155, 1990.
- [TEK95] A. Murat Tekalp, "Digital Video Processing", *Prentice Hall Signal Processing Series, University of Rochester*, February 1995.
- [TEK95] A. Murat Tekalp, "Digital video processing", *Prentice hall signal processing series*, Upper Saddle River, NJ 07458, 1995.
- [THO80] W. B. Thompson, "Combining motion and contrast for segmentation", *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. PAMI-2, N°6, pp. 543-449, November 1980.
- [TIS94] M. Tistarelli, "Multiple constraints for optical flow", in *Proc. 3<sup>rd</sup> European Conf. on Computer Vision*, Stockholm, pp.61-70, 1994.
- [TOM98] Carlo Tomasi, "Mathematical Methods for Robotics and Vision", CS205, Stanford University, Fall 1998.
- [TRE84] O. Tretiak and L. Pastor, "Velocity estimation from image sequences with second order differential operators", in *Proceedings International Conference on Pattern Recognition*, Montréal, pp. 16-19, 1984.
- [TRE94] P. trevs and J. Korad, "Motion estimation and compensation under varying illumination", in *Proc. ICIP'94*, Vol. 1, pp. 373-377, Austin, TX, November 1994.
- [VER89] A. Verri and T. Poggio, "Motion field and optical flow : qualitative properties", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, N°5, pp. 490-498, 1989.

---

[VER90] A. Verri, F. Girosi, and V. Torre, "Differential techniques for optical flow", *Journal of the Optical Society of America A*, Vol. 7, pp. 912-922, 1990.

[WAL84] D. R. Walker and K. R. Rao, "Improved pel-recursive motion compensation", *IEEE Trans. Comm.*, Vol. 32, pp. 1128-1134, October 1984.

[WAL96] F. Walter, "Détection de failles sismiques", Rapport de DEA Signal Image Parole Option Signaux Naturels et Télédétection, Grenoble, France, 1996.

[WAN94a] J. Y. A. Wang and E. H. Adelson, "Spatio-temporal segmentation of video data", in *SPIE Proc. Image and Video Processing II*, Vol. 2182, San Jose, CA, February 1994.

[WAN94b] J. Y. A. Wang and E.H. Adelson, "Representing moving images with layers", *IEEE Trans. Image Process.*, Vol. 3, N° 5, pp. 625-638, September 1994.

[YIL87] O. Yilmaz, "Seismic data processing", Society of Exploration Geophysicists, 1987.

## **Estimation du mouvement fondée sur un schéma direct et rétrograde - Application à la segmentation**

---

**Résumé** - Cette thèse traite de la détection et de l'estimation du mouvement dans le contexte de l'analyse de séquence d'images. Plus précisément, le but de ce manuscrit est le développement d'algorithmes d'estimation et de segmentation du flot optique, qui tendent à déterminer les changements relatifs de position des objets visibles dans une scène.

Un problème majeur de l'estimation du mouvement est l'insuffisance des contraintes pour rendre l'estimation robuste. Dans notre cas, nous proposons un formalisme qui permet de renforcer les contraintes dans la mise en équation du mouvement. Le formalisme est construit à partir de l'information spatio-temporelle causale et anti-causale (mouvements direct et rétrograde) de deux trames consécutives. Dans un premier temps, le formalisme direct-rétrograde est testé sur un mouvement de translation. Le second volet dédié à l'estimation consiste en l'extension du formalisme proposé sur un modèle plus général.

Pour la segmentation spatio-temporelle, deux catégories de méthode sont évoquées. La première est l'application du schéma direct-rétrograde à la segmentation markovienne. La deuxième catégorie est basée sur la classification utilisant les nuées dynamiques. Pour cela, nous développons une méthode de segmentation qui consiste à associer la méthode K-moyens et le formalisme direct-rétrograde.

Le dernier point concerne l'application sur des données réelles. Cette application est réalisée dans le contexte de l'analyse en imagerie sismique fournie par la société GDF.

**Mots clef** - Séquence d'images, estimation du mouvement, détection, segmentation, flot optique, direct, rétrograde, champ de Markov, K-moyens, imagerie sismique.

**Abstract** - This thesis deals with the motion detection and estimation in the context of the image sequence analysis. More precisely, the goal of this manuscript is the development of optical flow estimation and segmentation algorithms, which rely on the changes of image brightness.

The main problem of the motion estimation is the insufficient of constraint for provide a robust estimate. In our case, we tray to develop a formalism which allows to add some constraints in the motion equation. The formalism is carried out in starting from the causal and anti-causal spatio-temporal information (forward an backward motions) of two consecutive frames. Initially, the forward-backward formalism is tested on a translation motion. The second shutter consists of the extension of the suggested formalism on a general model. It is about the parametric model.

For the spatio-temporal segmentation, two categories of methods are evoked. The first is the application of the forward-backward formalism to the markovian segmentation. The second category is based on classification using the dynamic clouds. For that, we developed a method of segmentation in which, we associate the K-means algorithm and the forward-backward formalism.

Finally, we deal with the application on real data. This application was carried out in the context of the analyse in seismic imagery, provided by the GDF company.

**Keywords** - Image sequence, motion estimation, detection, segmentation, optical flow, forward, backward, Markov random field, K-means, seismic imagery.