

Institut National de la Recherche Scientifique
INRS-Télécommunications

**Estimation du mouvement apparent dans les séquences
d'images avec effet d'illumination**

Mémoire présenté comme exigence partielle de la
maîtrise ès sciences (télécommunications).

Paolo Trèves
B.Ing. Honneur (Electrique)
Montréal, Août 1994

REMERCIEMENTS

J'aimerais commencer par remercier le Professeur Janusz Konrad, mon directeur de thèse, qui m'a guidé tout au long de mes travaux et a fait preuve de beaucoup de patience. Sans son support, le développement de cette thèse aurait été beaucoup plus long et laborieux. J'ai beaucoup de reconnaissance envers l'INRS-Télécommunications pour m'avoir donné un support financier de base. Cette aide m'a permis d'entreprendre mes études de deuxième cycle immédiatement. De plus, j'ai pu les poursuivre sans interruption. J'aimerais aussi remercier tous mes collègues d'étude qui, à plusieurs reprises tout au cours des travaux de recherche, m'ont encouragé et ont contribué à une atmosphère professionnelle et scientifique. Enfin, je suis endetté envers ma famille pour le support moral qu'elle m'a donné pendant mes études.

SOMMAIRE

Ce travail présente une nouvelle approche à l'estimation du mouvement apparent dans les séquences d'images avec effet d'illumination. Plusieurs méthodes existent déjà permettant d'estimer le mouvement apparent dans les séquences d'images. Ces méthodes sont basées principalement sur l'hypothèse que l'intensité lumineuse ne change pas le long de la trajectoire de mouvement. Cette hypothèse est souvent combinée avec une contrainte de lissage afin de *régulariser* le problème. Parce que, les changements d'intensité sont uniquement attribués au mouvement, quand une perturbation lumineuse survient, les estimés du mouvement apparent obtenus ne sont pas représentatifs du mouvement apparent. Dans le but d'améliorer la représentation du mouvement apparent sous des conditions d'effets d'illumination, une nouvelle formulation est proposée. Celle-ci est basée sur l'hypothèse que le gradient d'intensité reste constant le long de la trajectoire de mouvement. Une contrainte de lissage est incluse avec cette hypothèse pour compléter la formulation. L'hypothèse que le gradient d'intensité reste constant le long de la trajectoire de mouvement est équivalente à un changement constant d'intensité le long de la trajectoire de mouvement. Avec cette hypothèse, une variation lumineuse est permise, ce que plusieurs formulations couramment utilisées ne permettent pas. Dans la partie expérimentale, les résultats d'estimation du mouvement basés sur ces deux hypothèses sont démontrés. Les images naturelles, ainsi que les images avec trois effets d'illumination synthétique sont utilisées. À partir des résultats expérimentaux, on peut conclure que la nouvelle méthode est beaucoup plus robuste à des effets d'illumination que les méthodes "classiques".

Table des Matières

REMERCIEMENTS	i
SOMMAIRE	ii
<i>TABLE DES MATIERES</i>	iii
<i>LISTE DES FIGURES</i>	v
<i>LISTE DES TABLEAUX</i>	x
1 Introduction	1
2 Fondements théoriques	7
2.1 L'échantillonnage	7
2.2 Mouvement linéaire	9
2.3 Les modèles de base	11
2.3.1 Modèle structurel d'ordre zéro	12
2.3.2 Modèle structurel de premier ordre	14
2.3.3 Modèle de mouvement	14
2.4 Régularisation	15
3 Estimation du mouvement apparent – approches courantes	18
3.1 Le problème de l'ouverture	19
3.2 Méthode de Horn et Schunck	19
3.3 Méthode basée sur le modèle structurel d'ordre zéro avec régularisation	23
3.3.1 Dérivation	23
3.3.2 Images de test	26

3.3.3	Résultats expérimentaux	28
3.4	Méthode fondée sur le modèle structurel de premier ordre	30
3.4.1	Dérivation	30
3.4.2	Travaux antérieurs	31
3.4.3	Résultats expérimentaux	32
4	Estimation du mouvement apparent – nouvelle formulation	39
4.1	Méthode fondée sur le modèle structurel d’ordre un avec régularisation	39
4.1.1	Dérivation	39
4.1.2	Résultats expérimentaux	43
4.2	Estimation du mouvement apparent par multi-résolution	46
5	Estimation à partir des séquences avec effet d’illumination	53
5.1	Modèles d’effet d’illumination	53
5.1.1	Premier effet d’illumination	54
5.1.2	Deuxième effet d’illumination	55
5.1.3	Troisième effet d’illumination	55
5.2	Résultats expérimentaux	57
5.2.1	Séquence avec effet d’illumination	57
5.2.2	Comparaison objective	59
5.2.3	L’impact sur le PSNR	61
6	Conclusion	74
	Appendice A. Les interpolateurs continus	77
	Bibliographie	84

Liste des Figures

1.1	Codeur prédictif avec compensation de mouvement	3
1.2	Reconstruction des trames intermédiaires utilisant l'interpolation compensée par le mouvement.	4
2.1	Mouvement apparent entre deux images.	8
2.2	Echantillonnage 3-D d'une séquence d'images (a) progressive (b) entrelacée.	9
2.3	Mouvement linéaire entre deux trames au temps t_- et t_+ ; le mouvement consiste d'une translation $\mathbf{d}(\mathbf{x}, t)$	10
2.4	Illustration de la paramétrisation s d'une trajectoire de mouvement $\mathbf{c}(\tau; \mathbf{x}, t)$	11
2.5	Les hypothèses structurelles: (a) premier ordre (constance du gradient de l'intensité); (b) ordre zéro (constance de l'intensité).	13
2.6	Voisinage de premier ordre.	15
3.1	Problème de l'ouverture dans le plan de l'image. Composantes parallèles et perpendiculaires à $\nabla f(x, y, t)$	20
3.2	Première séquence test, <i>courtepointe</i> :(a) trame 0; (b) trame 1; (c) trame 2	26
3.3	Deuxième séquence test, <i>femme – et – arbre</i> :(a) trame 4; (b) trame 5; (c) trame 6	27
3.4	Troisième séquence test, <i>pingpong</i> :(a) trame 14; (b) trame 15; (c) trame 16	27

3.5	Champs de déplacement calculés avec le modèle structurel d'ordre zéro avec régularisation (λ égal 20, interpolateur cubique):(a) courtepointe;(b) femme-et-arbre; (c) pingpong.	29
3.6	Champs de déplacement obtenus utilisant le modèle structurel d'ordre un sur la séquence 2 (sans filtrage) avec θ égal à:(a) 0.01; (b) 0.1; (c) 1	33
3.7	Champs de vitesse obtenus avec le modèle structurel d'ordre un avec $\theta = 0.1$ pour la séquence de test 2 avec divers interpolateurs ayant des expansions de Taylor à:(a) 3 points; (b) 5 points	34
3.8	Champs de vitesse obtenus avec le modèle structurel d'ordre un et $\theta = 0.1$ pour la séquence 2 avec divers interpolateurs:(a) cubique; (b) cubique avec compensation de mouvement	34
3.9	Champs de vitesse obtenus sur la séquence 2 avec $\theta = 0.1$ après filtrage avec un filtre gaussien à variance égal à:(a) 0.5; (b) 1; (c) 3.	35
3.10	Méthode itérative pour évaluer les champs de déplacement avant et après filtrage.	36
3.11	Champs de vitesse calculés pour la séquence 2 avec le modèle structurel d'ordre un avec régularisation après 10 itérations:(a) avant filtrage (w_t); (b) après filtrage (v_t).	38
4.1	Champs de déplacement calculés avec la nouvelle formulation (interpolateur cubique), sur les séquences($\lambda = 5$):(a) courtepointe;(b) femme-et-arbre;(c) pingpong.	44
4.2	Champs de déplacement estimés avec la nouvelle formulation sur la séquence 2 avec λ égal à:(a) 0.1; (b) 1.	49
4.3	Champs de déplacement calculés avec la nouvelle formulation sur la séquence 2 avec λ égal à 1 utilisant l'interpolateur:(a) quadratique; (b) cubique	49

4.4	Champs de déplacement calculés avec la nouvelle formulation sur la séquence 2 avec λ égal à 1 utilisant l'interpolateur:(a) cubique de Keys; (b) quartique	50
4.5	Pyramide de multi-résolution pour:(a) la formulation couramment utilisée; (b)la nouvelle formulation; (c) les différents niveaux d'images filtrées.	50
4.6	Différents niveaux de résolution pour la séquence 2:(a) haut de la pyramide; (b) 2eme niveau; (c) bas de la pyramide	51
4.7	Champs de déplacement calculés à différents niveaux de résolution sur la séquence 2 avec la nouvelle formulation avec λ égal à 1:(a) niveau 3; (b) niveau 2; (c) niveau 1.	51
4.8	Champs de déplacement calculés sur différents niveaux utilisant la nouvelle formulation sur la séquence 2 avec λ égal à 5:(a) niveau 1;(b) pyramide	52
4.9	Champs de déplacement calculés sur différents niveaux utilisant la formulation couramment utilisée sur la séquence 2 avec λ égal à 20:(a) niveau 0;(b) pyramide	52
5.1	Illustration du premier effet d'illumination sur une séquence d'images.	54
5.2	Illustration du deuxième effet d'illumination sur une séquence d'images (fonction gaussienne).	56
5.3	Illustration du troisième effet d'illumination sur une séquence d'images (partition aléatoire).	58
5.4	Evolution du PSNR pour les 30 premières trames de la séquence <i>courtepointe</i> utilisant la méthode de constance de DPD avec λ égal à 20.	63
5.5	Evolution du PSNR pour les 30 premières trames de la séquence <i>courtepointe</i> avec une superposition des courbes avec aucune illumination et les modèles 1 et 2, utilisant la méthode de constance de DGD avec λ égal à 20.	63

5.6	Evolution du PSNR pour les 30 premières trames de la séquence <i>femme– et – arbre</i> utilisant la méthode de constance de DPD avec λ égal à 20.	64
5.7	Evolution du PSNR pour les 30 premières trames de la séquence <i>femme– et – arbre</i> avec une superposition des courbes avec aucune illumination et les modèles 1 et 2, utilisant la méthode de constance de DGD avec λ égal à 3.	64
5.8	Evolution du PSNR pour 30 trames à partir de la quatorzième de la séquence <i>pingpong</i> utilisant la méthode de constance de DPD avec λ égal à 20.	65
5.9	Evolution du PSNR pour 30 trames à partir de la quatorzième de la séquence <i>pingpong</i> utilisant la méthode de constance de DGD avec λ égal à 10.	65
5.10	Estimation par constance de DPD sur les séquences <i>courtepointe, femme– et – arbre</i> et <i>pingpong</i> avec λ égal à 20.	66
5.11	Estimation par constance de DPD avec trois effets d’illumination sur la séquence <i>courtepointe</i> avec λ égal à 20.	67
5.12	Estimation par constance de DPD avec trois effets d’illumination sur la séquence <i>femme – et – arbre</i> avec λ égal à 20.	68
5.13	Estimation par constance de DPD avec trois effets d’illumination sur la séquence <i>pingpong</i> avec λ égal à 20.	69
5.14	Estimation par constance de DGD sur les séquences <i>courtepointe, femme – et – arbre</i> et <i>pingpong</i> avec λ égal à 5.	70
5.15	Estimation par constance de DGD avec trois effets d’illumination sur la séquence <i>courtepointe</i> avec λ égal à 5.	71
5.16	Estimation par constance de DGD avec trois effets d’illumination sur la séquence <i>femme – et – arbre</i> avec λ égal à 5.	72
5.17	Estimation par constance de DGD avec trois effets d’illumination sur la séquence <i>pingpong</i> avec λ égal à 5.	73

6.1	Représentation des divers voisinages utilisés par les interpolateurs continus de cette thèse. (a) 3x3, (b) 4x4.	77
6.2	Réponse impulsionnelle des interpolateurs continus.	82
6.3	Première dérivée des réponses impulsionnelles des interpolateurs continus.	82
6.4	Deuxième dérivée des réponses impulsionnelles des interpolateurs continus.	83

Liste des Tableaux

3.1	Changement d'énergie en fonction du nombre d'itérations pour le modèle structurel d'ordre zéro avec régularisation sur la séquence 2.	30
3.2	Changement du MSE en fonction du nombre d'itérations.	37
4.1	Variation d'énergie avec le nombre d'itérations pour le modèle structurel de premier ordre avec régularisation sur la séquence 2.	46
5.1	Changement du PSNR sous les effets d'illumination pour les modèles structurels d'ordre zéro et d'ordre un sur les séquences 1, 2 et 3 avec différentes valeurs de λ	61

Chapitre 1

Introduction

L'estimation du mouvement apparent dans une séquence d'images joue un rôle important dans la compression vidéo, la conversion de standards vidéo, l'estimation de paramètres de la scène et plusieurs autres applications. Toutes ces applications sont sujettes à des effets d'illumination comme les ombrages et les sources lumineuses variables. Un exemple très commun de ce phénomène survient lorsqu'une caméra filme une scène où apparaît un changement d'illumination. La scène devient plus claire ou plus sombre. Chaque séquence d'images contient un mouvement observable représenté par des changements d'intensité d'une trame à l'autre. À partir de ces changements d'intensité il est possible d'estimer un champ de déplacement ou un champ de vitesse. Un champ de déplacement caractérise les déplacements des positions de valeur d'intensité correspondantes d'une image à l'autre. Par contre, un champ de vitesse caractérise le changement de position d'une valeur d'intensité avec le temps le long de la trajectoire de mouvement entre deux images. Les changements d'illumination ont un impact sur les estimés du mouvement apparent utilisant des méthodes courantes. La plupart des méthodes courantes supposent que les changements d'intensité sont dus uniquement au mouvement apparent, ce qui n'est pas le cas avec les effets d'illumination. En ayant un impact sur l'estimation du mouvement apparent, l'aspect visuelle des trames reconstruites utilisant ces estimés se trouve affectée.

Une nouvelle formulation est introduite permettant d'obtenir des estimés du mouvement apparent plus représentatif du mouvement observable dans une séquence sous des conditions d'illumination. En utilisant cette nouvelle méthode plusieurs applications sont possible tenant compte des effets d'illumination linéaire. Voici une courte liste d'applications où l'information contenu dans les champs de déplacement est utilisé.

Lorsqu'une scène est filmée, la caméra projete la scène sur un plan. L'image ainsi obtenue est échantillonnée. Plusieurs standards existent pour définir la fréquence d'échantillonnage horizontal, vertical et temporel. En général, il n'est pas pratique de transmettre chaque échantillon à cause des exigences en terme de mémoire et rapidité de calcul d'ordinateurs ainsi que les contraintes sur le spectre de fréquence. Une compression est donc nécessaire afin d'augmenter la rapidité de transmission. En utilisant l'information obtenu par l'estimation des champs de déplacement, un facteur de compression d'information est réalisable. Un standard de transmission pour des images de télévision utilise une résolution de 720 par 480 pixels dans la recommandation CCIR-601. Pour une représentation temporelle adéquate du mouvement dans une scène, il est nécessaire de transmettre 30 images ou plus par seconde. Le nombre de pixels à transmettre est de l'ordre de $720 \times 480 \times 30$ par seconde (un très grand nombre). Deux approches de compression existent pour diminuer la quantité d'informations transmises tout en conservant l'aspect visuel. La première est basée sur l'élimination de redondances statistiques dans les images. Les redondances statistiques étant attribué avec les propriétés statistiques du signal comme la corrélation et le spectre. Ce type de compression utilise la corrélation de l'information visuelle pour éliminer les redondances sans perte d'information. La deuxième est basée sur l'élimination de redondance perceptuelle dans les images. Les redondances perceptuelles sont les propriétés d'une image qui peuvent être enlevé sans affecter la perception par le système visuel humain. Ce type de compression diminue l'information contenue dans les images de façon à minimiser les effets subjectifs perceptibles par le système visuel humain [3] et [9].

Les estimés du mouvement apparent des images sont souvent utilisés pour le codage prédictif avec compensation de mouvement [5]. Une image est prédite à partir d'images avoisinantes en utilisant les trajectoires de mouvement. Une erreur est établie entre l'image prédite et l'image courante. Cette erreur est ensuite transmise. Ce mode de compression est utilisé dans le codage prédictif avec compensation de mouvement (figure 1.1). Cette approche permet une compression de plusieurs ordres de grandeur.

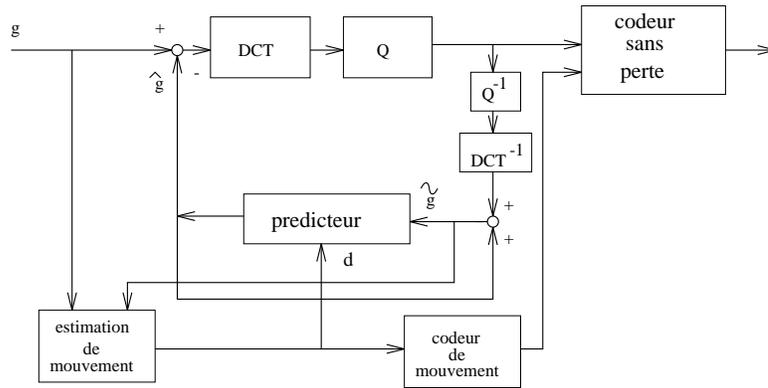


Figure 1.1: Codeur prédictif avec compensation de mouvement

Dans un système conventionnel de transmission d'images, l'émetteur sous-échantillonne souvent un certain nombre de trames. Par exemple, sur 30 images contenues dans une séquence originale, seulement 15 sont transmises dans un système à sous-échantillonnage de rapport 2:1. Au récepteur, les 15 images sont reçues. A partir de ces 15 images, 15 images intermédiaires sont reconstruites en utilisant les estimés du mouvement apparent dans les images transmises. Cette reconstruction des images non-transmises à l'émetteur est fondée sur le principe d'interpolation compensée par le mouvement [5] (figure 1.2). La formulation proposée dans cette étude permet sous des conditions de variations linéaires d'illumination de la scène d'obtenir des estimés du mouvement, représentatif du mouvement observable dans une séquence. Plusieurs standards ont été établis de façon à uniformiser les rapports de compression pour permettre une plus grande accessibilité aux séquences d'images vidéo (télévision,

téléphonie, informatique, etc.). Les standards de compression MPEG-1, MPEG-2, MPEG-4 et d'autres permettent d'atteindre des rapports de compression allant jusqu'à 200:1 dépendant des applications [2].

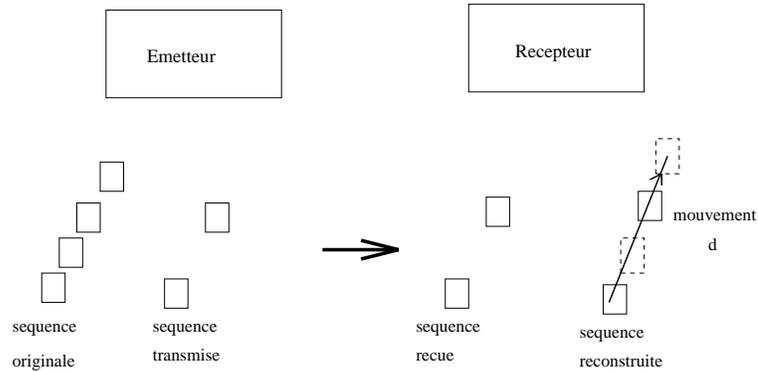


Figure 1.2: Reconstruction des trames intermédiaires utilisant l'interpolation compensée par le mouvement.

La conversion des standards d'échantillonnage constitue une autre application importante de l'estimation du mouvement apparent. Il est fréquent de convertir des séquences d'images d'une fréquence d'échantillonnage à une autre. Par exemple, la conversion du signal NTSC (30 images/seconde) au signal PAL ou SECAM (25 images/seconde), ou la conversion NTSC au rythme de HDTV, et vice versa. Dans le cas de conversion de fréquence de trames, il est nécessaire de reconstruire ou d'éliminer les images intermédiaires à partir de trames originales pour obtenir une nouvelle fréquence de trames compatibles avec le nouveau standard. L'estimation du mouvement apparent joue un rôle important dans ce type d'application [5].

DeMicheli *et al.* [6] ont estimé des paramètres tri-dimensionnels d'une scène à partir d'estimés de mouvement apparent. En analysant la variance, la moyenne et d'autres propriétés du champ de déplacement, il leur a été possible de déterminer certaines caractéristiques de la scène ayant été filmée. Les paramètres estimés se rapportaient à: la vitesse de mouvement d'objets dans la scène, l'orientation et la distance d'objet par rapport à la caméra, parmi plusieurs autres paramètres.

Toutes ces applications, que ce soit la reconstruction de trames par interpolation compensée par le mouvement, le codage prédictif avec la compensation de mouvement, la conversion des standards vidéo ou l'estimation des paramètres d'une scène, nécessitent de précises représentations du mouvement apparent dans les séquences d'images (avec ou sans effet d'illumination). Les méthodes couramment utilisées ne sont pas adéquates pour permettre une représentation satisfaisante du mouvement apparent avec des effets d'illumination; conséquemment plusieurs applications ne sont pas possible dans ces conditions. Une nouvelle formulation est introduite permettant des estimés plus représentatifs du mouvement apparent avec des effets d'illumination. Cette nouvelle formulation est beaucoup plus robuste aux changements d'illumination de la scène qui peuvent survenir lorsque celle-ci est filmée.

Ce mémoire est organisé de la façon suivante: après avoir décrit les motivations et certaines applications de la recherche dans ce chapitre, un bref résumé des fondements théoriques reliés aux méthodes de calcul des champs de déplacement est présenté dans le deuxième chapitre. Les fondements théoriques consistent d'une présentation sur les formats d'échantillonnage d'images, le modèle de mouvement, les modèles structuraux et la régularisation. Au troisième chapitre, une dérivation détaillée de la formulation couramment utilisée est présentée. Celle-ci combine l'hypothèse que l'intensité reste constante le long de la trajectoire de mouvement et l'hypothèse que les champs de vitesse sont lisses. De plus, une simple dérivation mathématique est présentée basée sur l'hypothèse que le gradient d'intensité reste constant le long de la trajectoire de mouvement. Cette hypothèse utilisée par DeMicheli *et al.* [6] est analysée plus en détail dans ce mémoire parce qu'elle est à la base de la nouvelle formulation proposée dans ce travail. Au quatrième chapitre, une nouvelle formulation proposée est introduite. Celle-ci combine l'hypothèse de la constance du gradient et un modèle de mouvement. Cette nouvelle formulation est introduite de façon à obtenir des estimés plus représentatifs du mouvement apparent dans des séquences avec des effets d'illumination. Une analyse détaillée de cette nouvelle formulation est présentée montrant l'impact de celle-ci sur les estimés du mouvement appar-

ent. L'estimation basée sur une approche multi-résolutions est aussi présenté dans ce chapitre. Au cinquième chapitre, trois effets d'illumination synthétique sont proposés. Les algorithmes d'estimation proposés antérieurement (méthode couramment utilisé et la nouvelle formulation) sont utilisés pour estimer le mouvement apparent dans des séquences contenant ces effets. Finalement, au sixième chapitre les conclusions sont présentées.

Chapitre 2

Fondements théoriques

Quelques définitions importantes doivent être présentées de façon à comprendre l'impact que les effets d'illumination ont sur les estimés du mouvement apparent. Pour commencer, une définition du mouvement apparent est nécessaire. Une observation visuelle des séquences d'images révèle un mouvement apparent des objets. Par exemple, dans la figure 2.1, les trames au temps t_- et t_+ montrent le déplacement effectué par une boîte d'un instant de temps à un autre. Pour la boîte en vue, le mouvement consiste en une translation. Dans cet exemple, la translation de la boîte d'une trame à l'autre (le mouvement apparent) peut être représentée par un champ de déplacement ou un vecteur est associé avec chaque pixel de l'image. Donc, les champs de déplacement sont denses. Le but est d'estimer ce champ de déplacement, i.e., d'estimer un vecteur pour chaque pixel dans une image.

2.1 L'échantillonnage

Selon le mode d'échantillonnage des séquences d'images, l'estimation du mouvement apparent ainsi que leur mode de calcul en sont affectés. Une scène est définie en 4 dimensions (x-y-z-t). Une scène projetée sur un plan définit une image dont les coordonnées sont dénotés par $\mathbf{x} = (x, y)$. Une série d'images ou une séquence est définie en 3 dimensions (x-y-t).

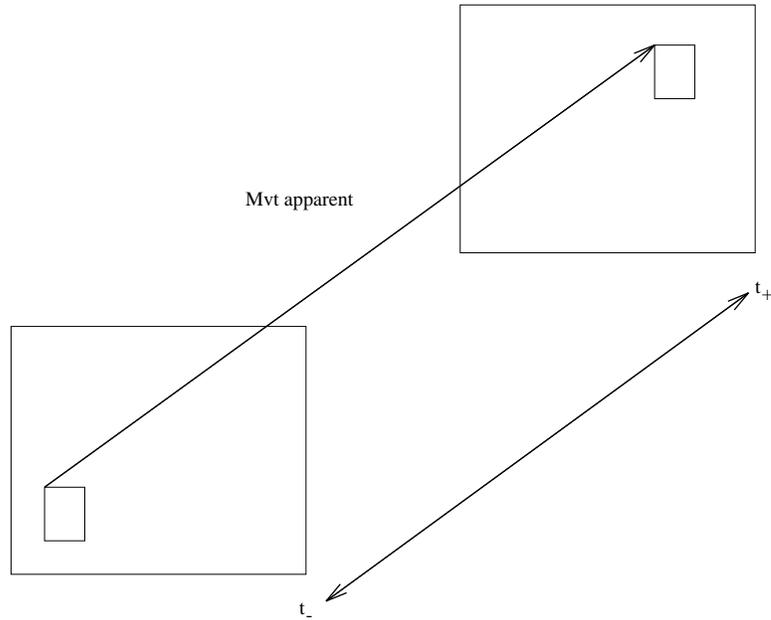


Figure 2.1: Mouvement apparent entre deux images.

Dans la figure 2.2(a), 4 trames sont présentées sur le plan horizontal-vertical (x, y) et sur le plan vertical-temporel (y, t) . Les points représentent des échantillons. Les échantillons d'une trame à l'autre sont tous alignés. Ce format d'échantillonnage représente une séquence progressive. Dans une séquence progressive chaque trame correspond à une image.

Dans la figure 2.2(b), une présentation semblable est faite. Ici, les échantillons sur le plan vertical-temporel (y, t) sont décalés d'une trame à une autre. Dans ce mode d'échantillonnage, deux trames sont nécessaires pour définir une image. Une trame contient les lignes paires et l'autre les lignes impaires. Toutes les simulations contenues dans le mémoire ont été effectuées avec des images échantillonnées de façon entrelacée.

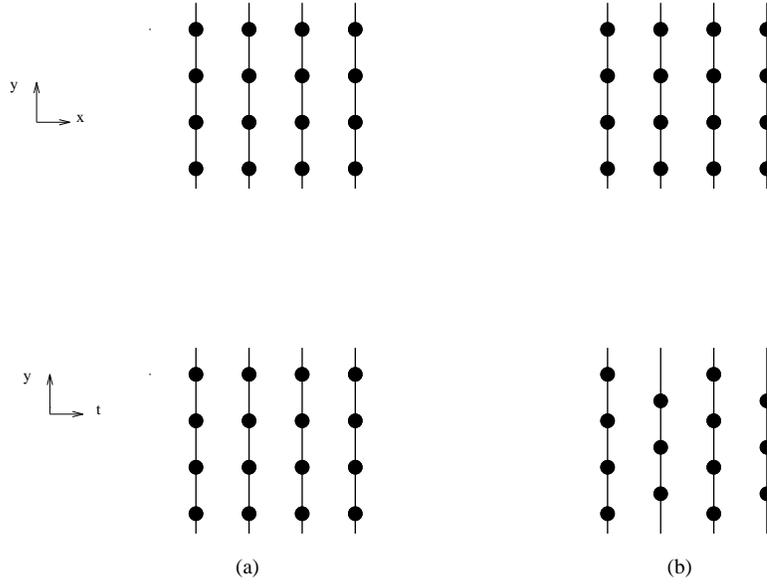


Figure 2.2: Echantillonnage 3-D d'une séquence d'images (a) progressive (b) entrelacée.

2.2 Mouvement linéaire

Sachant que les images sont entrelacées et les champs de déplacement sont denses, un modèle de mouvement entre les trames est défini. Dans les formulations présentées ici, tout calcul de champs de déplacement implique la présence de deux images, une successive t_+ et une précédente t_- à la position temporelle courante t . Entre ces deux images, il est possible de définir une trajectoire de mouvement dénotée $\mathbf{c}(t_+; \mathbf{x}, t)$. Ici, $\mathbf{c}(t_+; \mathbf{x}, t)$ représente la position sur l'image au temps t_+ du point à la position \mathbf{x} au temps t [4].

Dans les travaux effectués dans ce mémoire, les trajectoires de mouvement sont considérées comme étant linéaires. On suppose que la fréquence temporelle des images est assez élevée pour justifier cette hypothèse. La vitesse d'un pixel est définie comme suit:

$$\mathbf{v}(\mathbf{x}, t) = \frac{d}{dt_+} \mathbf{c}(t_+; \mathbf{x}, t)$$

Un déplacement est défini comme suit:

$$\mathbf{d}(t_+; \mathbf{x}, t) = \begin{cases} \mathbf{x} - \mathbf{c}(t_+; \mathbf{x}, t) & \text{si } t_+ < t \\ \mathbf{c}(t_+; \mathbf{x}, t) - \mathbf{x} & \text{si } t_+ > t \end{cases}$$

Pour un mouvement à vitesse constante, le déplacement est simplement égal à $\mathbf{d}(t_+; \mathbf{x}, t) = \mathbf{v}(\mathbf{x}, t) \cdot (t_+ - t)$. Par conséquent:

$$\mathbf{c}(t_+; \mathbf{x}, t) = \mathbf{x} + \mathbf{d}(t_+; \mathbf{x}, t)$$

Pour une image au temps t , \mathbf{d} est l'ensemble des déplacements 2-D qui décrivent le mouvement de pixels entre t et t_+ (figure 2.3). Si t_+ est plus grand que t , $\mathbf{d}(t_+; \mathbf{x}, t)$ est un déplacement vers l'avant. Si t_+ est plus petit que t , $\mathbf{d}(t_+; \mathbf{x}, t)$ est un déplacement vers l'arrière. Pour un mouvement linéaire, l'estimation du mouvement apparent consiste à estimer \mathbf{d} . La dérivation présentée ci-dessus sur le vecteur de déplacement est inspirée de [5].

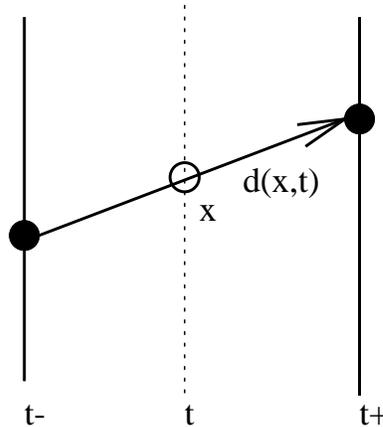


Figure 2.3: Mouvement linéaire entre deux trames au temps t_- et t_+ ; le mouvement consiste d'une translation $\mathbf{d}(\mathbf{x}, t)$.

Parce qu'il est important de modéliser l'effet d'illumination entre les trames le long de la trajectoire de mouvement, un paramètre additionnel est introduit. Une trajectoire de mouvement \mathbf{c} peut être paramétrisée par la variable curviligne s [5].

Une fraction d'intervalle de temps séparant deux trames successives est dénotée par dt . Un intervalle de temps dt est une mesure temporelle entre deux trames successives. Il est possible de définir un gradient d'intensité $\nabla f(x, y, t)$ à chaque position sur le plan de l'image. Le gradient d'intensité est défini par le changement d'intensité dans la direction horizontale et verticale du plan de l'image. Une illustration des définitions entre ces variables est présentée à la figure 2.4. Le rôle de ces variables deviendra clair bientôt.

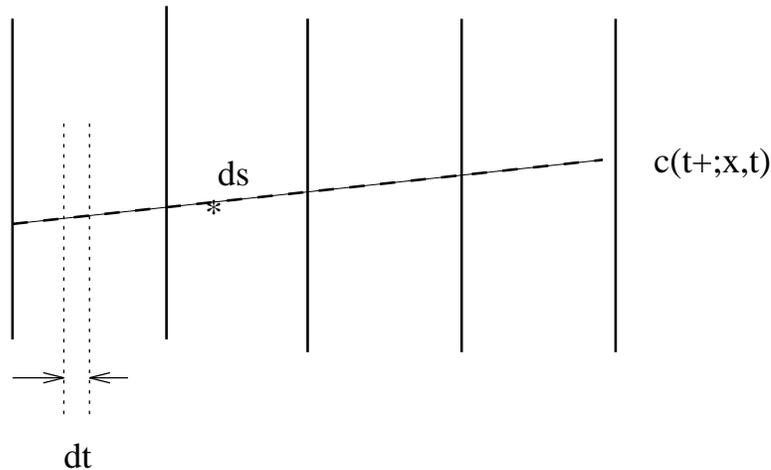


Figure 2.4: Illustration de la paramétrisation s d'une trajectoire de mouvement $\mathbf{c}(\tau; \mathbf{x}, t)$.

2.3 Les modèles de base

L'estimation du mouvement apparent dans une séquence est faite à partir des modèles de base. Deux types de modèle, *le modèle structurel* et *le modèle de mouvement* sont généralement combinés de façon à former une fonction objective. Chaque modèle joue un rôle important dans l'estimation du mouvement apparent. Le modèle structurel se rapporte à l'hypothèse qui est faite sur l'intensité le long de la trajectoire de mouvement. Le modèle structurel définit les propriétés de l'intensité d'une trame à

une autre. Le modèle de mouvement définit le type de trajectoire \mathbf{c} d'une trame à une autre (linéaire, quadratique, etc.).

Deux types de modèles structuraux sont utilisés dans ce mémoire. Le premier repose sur l'hypothèse que l'intensité lumineuse reste constante le long de la trajectoire de mouvement (modèle d'ordre zéro). Le deuxième repose sur l'hypothèse que le gradient de l'intensité lumineuse reste constant le long de la trajectoire de mouvement (modèle de premier ordre). Le modèle structurel le plus utilisé et existant déjà depuis quelque temps est celui reposant sur l'hypothèse que l'intensité ne change pas le long de la trajectoire de mouvement. Conséquemment, les résultats obtenus utilisant ce modèle sont utilisés comme résultats de référence pour la nouvelle formulation proposée. Une illustration des hypothèses structurelles de base est présentée dans la figure 2.5. Dans la figure 2.5 (a) et (c) est illustrée le changement linéaire d'intensité en fonction du temps le long de la trajectoire de mouvement. Ce changement linéaire d'intensité est permis à cause de l'hypothèse de base selon laquelle le gradient d'intensité reste constant le long de la trajectoire de mouvement. Dans la figure 2.5 (b) et (d) est illustrée l'hypothèse structurelle d'ordre zéro. Dans ce cas, il n'y a pas de changement d'intensité en fonction du temps le long de la trajectoire de mouvement. La constance de l'intensité résulte de l'hypothèse structurelle.

2.3.1 Modèle structurel d'ordre zéro

Le premier modèle structurel introduit suppose que l'intensité reste constante le long de la trajectoire de mouvement. La formulation mathématique de cet énoncé est la suivante:

$$\frac{d}{ds}(f(x, y, t)) = 0 \quad (2.1)$$

où s est une variable le long de la trajectoire de mouvement (figure 2.4). La fonction $f(x, y, t)$ représente l'intensité lumineuse sur le plan de l'image idéale (i.e., sans du bruit, recouvrement des spectres, distortions, etc.). Une formulation discrète de

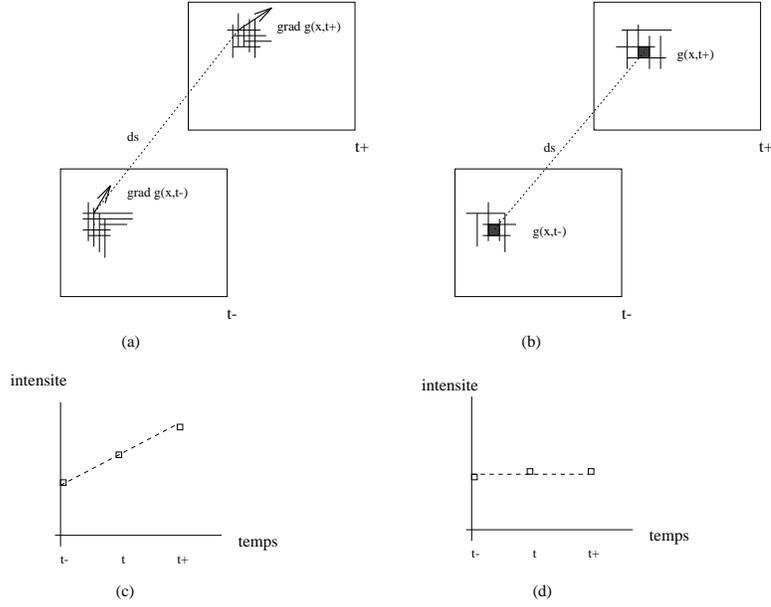


Figure 2.5: Les hypothèses structurelles: (a) premier ordre (constance du gradient de l'intensité); (b) ordre zéro (constance de l'intensité).

l'hypothèse ci-dessus s'appliquant aux images, est la suivante (assumant une trajectoire linéaire):

$$f(\mathbf{x} + (1 - \Delta_t) \cdot \mathbf{d}(\mathbf{x}, t), t_+) - f(\mathbf{x} - \Delta_t \cdot \mathbf{d}(\mathbf{x}, t), t_-) = 0$$

où, $\Delta_t = \frac{(t-t_-)}{(t_+-t_-)}$ représente la position relative dans l'intervalle de temps $[t_-, t_+]$ où l'estimation est faite au temps t relativement à la position de l'image précédente t_- (figure 2.3). Le vecteur $\mathbf{d}(\mathbf{x}, t)$ représente un vecteur de déplacement associé avec un pixel $\mathbf{a}(\mathbf{x}, t)$. Cette équivalence est valide parce que la trajectoire de mouvement entre deux trames est définie comme étant linéaire. Pour des images observées $g(x, y, t)$ (i.e. avec du bruit gaussien, distortions du signal, etc.), la constance de l'intensité devient:

$$g(\mathbf{x} + (1 - \Delta_t) \cdot \mathbf{d}(\mathbf{x}, t), t_+) - g(\mathbf{x} - \Delta_t \cdot \mathbf{d}(\mathbf{x}, t), t_-) = r(\mathbf{d}, \mathbf{x}, t) \quad (2.2)$$

où, $r(\mathbf{d}, \mathbf{x}, t)$ est un scalaire avec des propriétés d'un bruit gaussien [5]. Le modèle structurel d'ordre zéro sera appelé par DPD (Différence de Pixel Déplacé).

2.3.2 Modèle structurel de premier ordre

Le modèle structurel de premier ordre est fondé sur le fait que le gradient d'intensité ne change pas le long de la trajectoire de mouvement. Ce modèle a une formulation mathématique comme suit:

$$\frac{d}{ds}(\nabla(f(x, y, t))) = 0 \quad (2.3)$$

Le gradient d'intensité $\nabla = [\frac{\partial}{\partial x} \frac{\partial}{\partial y}]^T$ reste constant le long de la trajectoire de mouvement. Une formulation discrète équivalente est nécessaire (supposant une trajectoire de mouvement linéaire):

$$\nabla f(\mathbf{x} + (1 - \Delta_t) \cdot \mathbf{d}(\mathbf{x}, t), t_+) - \nabla f(\mathbf{x} - \Delta_t \cdot \mathbf{d}(\mathbf{x}, t), t_-) = 0$$

Idéalement, la différence entre les deux gradients est nulle [23] et [14]. Pour les images observées $g(x, y, t)$, cette différence est non-nulle.:

$$\nabla g(\mathbf{x} + (1 - \Delta_t) \cdot \mathbf{d}(\mathbf{x}, t), t_+) - \nabla g(\mathbf{x} - \Delta_t \cdot \mathbf{d}(\mathbf{x}, t), t_-) = \mathbf{s}(\mathbf{d}, \mathbf{x}, t) \quad (2.4)$$

où $\mathbf{s}(\mathbf{d}, \mathbf{x}, t)$ est un vecteur avec des propriétés du bruit gaussien [5]. Le modèle structurel d'ordre un sera appelé par DGD (Différence de Gradient Déplacé).

2.3.3 Modèle de mouvement

Le modèle de mouvement de vecteurs de déplacement est fondé sur l'observation que les pixels voisins se déplacent à magnitude et direction similaires. Sans modèle de mouvement, chaque vecteur de déplacement est calculé indépendamment des vecteurs avoisinants. La formulation de la fonction objective du modèle de mouvement à la position \mathbf{x}_i est la suivante:

$$\sum_{\mathbf{x}_j \in \eta(\mathbf{x}_i)} \| \mathbf{d}(\mathbf{x}_i, t) - \mathbf{d}(\mathbf{x}_j, t) \|^2$$

Ici, $\mathbf{d}(\mathbf{x}_i, t)$ représente le vecteur de déplacement à la position où l'estimation est faite. $\mathbf{d}(\mathbf{x}_j, t)$ représente le vecteur de déplacement faisant partie du voisinage $\eta(\mathbf{x}_i)$. Une

somme est effectuée sur l'ensemble des vecteurs constituant un voisinage de premier ordre.

En introduisant un modèle de mouvement, il est possible d'établir une certaine correspondance entre vecteur de déplacement avoisinant (une corrélation). À la figure 2.6, une illustration d'un voisinage de premier ordre est présentée. Dans cette figure, \mathbf{d}_g , \mathbf{d}_d , \mathbf{d}_b et \mathbf{d}_h représentent les vecteurs de déplacement à la gauche, droite, vers le bas et le haut de la position courante du vecteur de déplacement \mathbf{d} . De plus, la variable $\bar{\mathbf{d}}$ est défini comme la moyenne des vecteurs de déplacement avoisinant de \mathbf{d} .

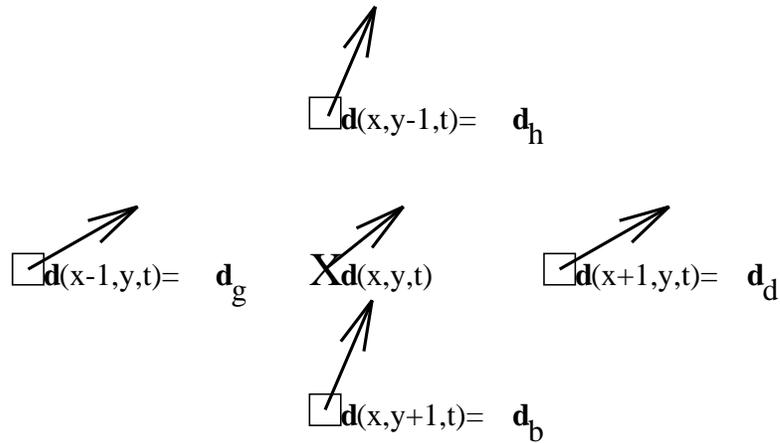


Figure 2.6: Voisinage de premier ordre.

2.4 Régularisation

Le rôle que joue chaque modèle dans la formulation, sous forme d'une fonction objective, est discuté. Les méthodes d'estimation du mouvement apparent courantes utilise un processus de régularisation. Quelques notions importantes doivent être introduites, de façon à justifier cette approche de calcul pour estimer le mouvement apparent dans les séquences d'images (avec ou sans effet d'illumination).

Bertero *et al.* [15] réfèrent à l'estimation du mouvement apparent d'une séquence d'images discrètes bi-dimensionnelles comme étant un problème d'optique inverse. Le mouvement estimé d'une séquence d'images est une approximation du mouvement apparent contenu dans une séquence d'images. Le mouvement des objets de la scène est estimé à partir d'une séquence d'images donc la justification pour le terme d'un problème d'optique inverse. Par analogie, la projection d'une scène (3D+t) sur un plan (2D) représente le problème direct.

L'information contenue dans les images bi-dimensionnelles (2D+t) est insuffisante pour caractériser les propriétés des objets de la scène (3D+t). Simplement, la scène se trouve à être projeté sur un plan. De plus, les images (2D+t) sont quantifiées en un nombre limité de niveaux d'intensité (de 0 à 255) et d'échantillons horizontaux et verticaux (le nombre d'échantillons dépend du standard de représentation). À cause de ces limitations, un problème inverse est mal posé. Un problème mal posé est telle que [15]:

1. la solution n'est pas unique, ou
2. la solution n'existe pas, ou
3. la solution ne dépend pas de l'information contenue dans l'image de façon continue

Par contre, un problème bien posé est tel qu'une solution unique existe qui est fonction continue des données. Le problème d'estimation du mouvement apparent est un problème d'optique inverse et, conséquemment, est mal posé. Le bruit a un impact important sur les problèmes mal posés. Toute image discrète bi-dimensionnelle, contient du bruit gaussien venant de distortions et d'interférences venant du signal, d'erreur de quantification et ainsi de suite. Durant les dernières années, un développement pratique et théorique a été élaboré pour obtenir une solution approximative au problème mal posé. Ce développement porte le nom de *régularisation* de problème mal posé.

La régularisation utilise une propriété du problème d'estimation pour réduire l'ensemble de solution du problème d'optimisation, ce qui permet une solution plus exacte. De plus, la régularisation permet d'obtenir une solution exacte au problème d'estimation du mouvement apparent. En supposant que les champs de déplacement doivent être lisses, une solution unique au problème d'estimation du mouvement apparent est établie.

Une formulation générale d'un problème d'estimation de mouvement apparent utilisant une méthode de minimisation est la suivante:

$$\min_{\{\mathbf{d}\}} \sum_i E_s(\mathbf{x}_i)$$

La fonction de coût E_s est associée à un modèle structurel (équation (2.2) ou (2.4)). Une solution exacte à ce problème d'optimisation n'est pas possible en général parce que l'ensemble des solutions est infini. Dans le but d'obtenir une solution exacte au problème mal posé, une hypothèse additionnelle est introduite, dans notre cas, les champs de déplacement doivent être lisses. En introduisant cette nouvelle hypothèse au problème original, une nouvelle formulation est établie:

$$\min_{\{\mathbf{d}\}} \sum E_s(\mathbf{x}_i) + \lambda E_m(\mathbf{x}_i)$$

Cette nouvelle fonction objective incorpore le modèle de mouvement limitant l'ensemble des solutions. Ce terme régularise la solution au problème d'estimation de mouvement apparent. Le paramètre λ détermine la contribution du modèle de mouvement sur la solution finale. Les deux termes sous sommation sont valides pour une seule position sur le plan de l'image. Une somme est effectuée sur toutes les positions pour obtenir un champ de déplacement représentatif du mouvement apparent sur le plan d'une image. À chaque position sur le plan de l'image où la fonction objective est évaluée, une valeur d'énergie en résulte. Le but des méthodes de descente est de minimiser cette valeur d'énergie. Une fois cette minimisation effectuée, les valeurs des vecteurs de déplacement obtenues à chaque position seront les estimés de mouvement pour cette formulation particulière.

Chapitre 3

Estimation du mouvement apparent – approches courantes

L'hypothèse à la base du modèle structurel d'ordre zéro est que l'intensité lumineuse ne varie pas le long des trajectoires de mouvement. La formulation mathématique de cette hypothèse est donnée par l'équation (2.1). Cette expression peut être réexprimée en terme d'une dérivée directionnelle de $f(x, y, t)$ dans la direction de \vec{z} . Ceci est dénoté mathématiquement comme suit:

$$\frac{d}{ds}f(x, y, t) = \vec{z} \cdot \nabla f(x, y, t)$$

où,

$$\vec{z} = \frac{1}{\sqrt{v_x^2 + v_y^2 + 1}} [v_x v_y 1]$$
$$\nabla f(x, y, t) = \left[\frac{\partial f(x, y, t)}{\partial x} \quad \frac{\partial f(x, y, t)}{\partial y} \quad \frac{\partial f(x, y, t)}{\partial t} \right]^T$$

et v_x et v_y sont les composantes de la vitesse $\mathbf{v} = [v_x v_y]^T$. Donc, l'équation 2.1 peut être réécrite de la façon suivante:

$$\frac{\partial}{\partial x}f(x, y, t) \cdot v_x + \frac{\partial}{\partial y}f(x, y, t) \cdot v_y + \frac{\partial}{\partial t}f(x, y, t) = 0$$

Cette équation est connue dans la littérature comme l'équation de contrainte de mouvement [1]. Elle contient deux inconnues v_x et v_y . Il est clair qu'une autre contrainte doit être introduite pour permettre une solution unique au champ des vitesses. Comme discuté dans le chapitre 2, une deuxième contrainte est nécessaire pour régulariser l'ensemble solution.

3.1 Le problème de l'ouverture

Si on tente d'estimer le mouvement apparent dans une séquence d'images utilisant seulement le modèle structurel d'ordre zéro, il en résulte qu'aucune solution unique n'existe (une équation avec deux inconnues). Le vecteur vitesse \mathbf{v} peut être décomposé en deux composantes, une parallèle \mathbf{v}^{\parallel} et une perpendiculaire \mathbf{v}^{\perp} au gradient d'intensité $\nabla g(x, y, t)$. Donc, $\mathbf{v} = \mathbf{v}^{\parallel} + \mathbf{v}^{\perp}$ et en conséquence:

$$\nabla^T g \cdot \mathbf{v} + \frac{\partial g}{\partial t} = \nabla^T g \cdot \mathbf{v}^{\perp} + \nabla^T g \cdot \mathbf{v}^{\parallel} + \frac{\partial g}{\partial t} = \nabla^T g \cdot \mathbf{v}^{\parallel} + \frac{\partial g}{\partial t} = 0$$

Parce que $\nabla^T g \cdot \mathbf{v}^{\perp} = 0$, une infinité de solutions existent pour la composante perpendiculaire \mathbf{v}^{\perp} au gradient d'intensité. Seulement la composante parallèle \mathbf{v}^{\parallel} est calculée de façon unique. Dans Bertero *et al.* [15], une analogie de ce phénomène est faite. La valeur calculée de v^{\parallel} et v^{\perp} est la même qui serait obtenue en observant l'image à travers un petit orifice où une petite ouverture, et être capable de seulement déterminer v^{\parallel} . Une illustration du problème de l'ouverture est donnée dans la figure 3.1.

Dans le même article, il est démontré que: par le choix d'un modèle structurel et d'un modèle de lissage approprié, il est possible d'éliminer l'effet de l'ouverture, conséquemment il est possible d'obtenir des champs de vitesse plus représentatifs du mouvement apparent de l'image.

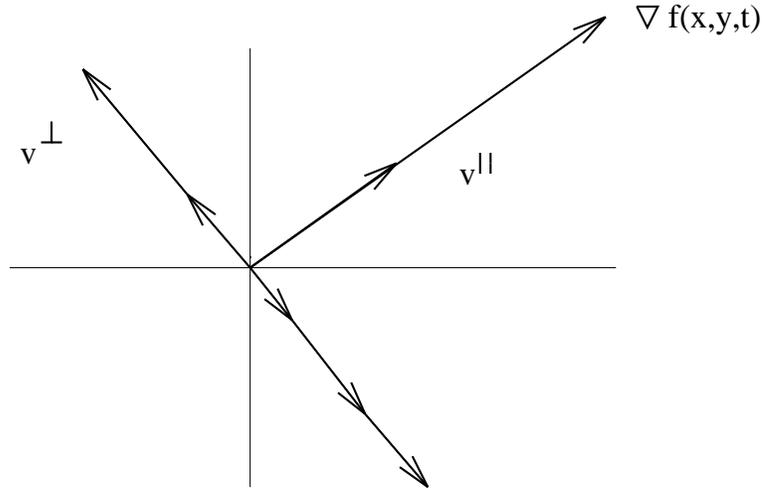


Figure 3.1: Problème de l'ouverture dans le plan de l'image. Composantes parallèles et perpendiculaires à $\nabla f(x, y, t)$.

3.2 Méthode de Horn et Schunck

Pour déterminer $\mathbf{v}(\mathbf{x}, t)$, il est nécessaire d'incorporer une contrainte additionnelle dans le problème d'estimation. Cette contrainte mentionnée dans le deuxième chapitre est le modèle de mouvement qui a pour rôle de régulariser le problème.

Dans les travaux de Horn et Schunck [1], ils ont combiné l'équation de contrainte de mouvement avec un modèle de mouvement ayant une formulation autre que celle présentée dans le deuxième chapitre. Le premier modèle de mouvement qu'ils ont proposé a été le suivant:

$$\varepsilon_c = \left(\left(\frac{\partial v_x}{\partial x} \right)^2 + \left(\frac{\partial v_x}{\partial y} \right)^2 \right) + \left(\left(\frac{\partial v_y}{\partial x} \right)^2 + \left(\frac{\partial v_y}{\partial y} \right)^2 \right)$$

La deuxième formulation du modèle de mouvement proposé utilise le Laplacien des composantes du flot optique. La formulation mathématique de ce modèle de mouvement est la suivante:

$$\nabla^2 v_x + \nabla^2 v_y = \left(\frac{\partial^2 v_x}{\partial x^2} + \frac{\partial^2 v_x}{\partial y^2} \right) + \left(\frac{\partial^2 v_y}{\partial x^2} + \frac{\partial^2 v_y}{\partial y^2} \right)$$

Le modèle structurel qu'ils ont utilisé est le suivant (l'équation de contrainte de mouvement):

$$\varepsilon_b = \frac{\partial}{\partial x}g(x, y, t) \cdot v_x + \frac{\partial}{\partial y}g(x, y, t) \cdot v_y + \frac{\partial}{\partial t}g(x, y, t)$$

Notons que ce terme n'est pas nul parce qu'il est basé sur les observations g (chapitre 2). Ils ont combiné ces deux termes en une fonction objective comme suit:

$$\min_{\{v_x, v_y\}} \int \int (\lambda^2 \varepsilon_c^2 + \varepsilon_b^2) dx dy$$

Donc, ils ont régularisé le problème de départ de minimisation de ε_c par la contrainte additionnel ε_b . Ils ont résolu ce problème par calcul de variations. En calculant les dérivées de la fonction objective, ils ont obtenu deux équations (une pour chaque composante):

$$\begin{cases} g_x^2 v_x + g_x g_y v_y = \lambda^2 \nabla^2 v_x - g_x g_t \\ g_x g_y v_x + g_y^2 v_y = \lambda^2 \nabla^2 v_y - g_y g_t \end{cases} \quad \forall (x, y)$$

où, $g_x = \frac{\partial}{\partial x}g(x, y, t)$, $g_y = \frac{\partial}{\partial y}g(x, y, t)$ et $g_t = \frac{\partial}{\partial t}g(x, y, t)$. Substituant une approximation du Laplacien [1], on obtient:

$$\begin{cases} (\lambda^2 + g_x^2) v_x + g_x g_y v_y = (\lambda^2 \bar{v}_x - g_x g_t) \\ (g_x g_y v_x + (\lambda^2 + g_y^2) v_y = (\lambda^2 \bar{v}_y - g_y g_t) \end{cases} \quad \forall (x, y)$$

où \bar{v}_x et \bar{v}_y sont les moyennes des composantes du flot optique. Avec ce système d'équations, il est possible d'obtenir une solution pour v_x et v_y à chaque position (x, y) en utilisant la méthode de Cramer. On obtient:

$$\begin{cases} (\lambda^2 + g_x^2 + g_y^2) v_x = (\lambda^2 + g_y^2) \bar{v}_x - g_x g_y \bar{v}_y - g_x g_t \\ (\lambda^2 + g_x^2 + g_y^2) v_y = -g_x g_y \bar{v}_x + (\lambda^2 + g_x^2) \bar{v}_y - g_y g_t \end{cases} \quad \forall (x, y)$$

Ces deux équations sont écrites pour chaque point sur le plan de l'image. Cette formulation exige l'évaluation de dérivées spatiales et temporelles g_x , g_y et g_t . Dans leur présentation, Horn et Schunck ont proposé des approximations aux dérivées basées

sur des interpolateurs discrets en trois dimensions (horizontal, vertical et temporel [1]).

Une solution directe pour chaque coordonnée de la grille d'échantillonnage du système d'équation précédent est possible mais pas pratique. Si une matrice de coefficients est construite pour chaque composante du vecteur de vitesse, certains aspects de cette matrice se manifeste. La matrice correspondante sera très grande, deux fois le nombre de pixels et lignes de l'image (chaque vecteur comprend deux composantes) et la matrice contiendra beaucoup de valeurs nulles (chaque vecteur de vitesse est le voisin de quatre autres vecteurs avoisinants). Sur une image de dimension 512x512, la matrice des coefficients sera de dimension (512x512x2)x(512x512x2). De plus, pour chaque rangée, seulement 5 sur (512x512x2) valeurs seront non-nulles (parce que un vecteur contient 4 voisins). A cause de ces aspects, des méthodes récursives sont utilisées. Une solution au problème d'estimation de mouvement est obtenue par plusieurs étapes de relaxation de la méthode de Gauss-Seidel. Cette approche permet d'obtenir une solution initiale en terme de dérivées de l'intensité et d'estimé antérieur du mouvement apparent. A chaque itération un nouvel estimé des composantes du vecteur de vitesse est obtenue. Les équations de récursion sont:

$$v_x^{n+1} = \bar{v}_x^n - \frac{g_x [g_x \bar{v}_x^n + g_y \bar{v}_y^n + g_t]}{(\lambda^2 + g_x^2 + g_y^2)}$$

$$v_y^{n+1} = \bar{v}_y^n - \frac{g_y [g_x \bar{v}_x^n + g_y \bar{v}_y^n + g_t]}{(\lambda^2 + g_x^2 + g_y^2)}$$

Horn et Schunck ont obtenus des résultats satisfaisants avec cette approche sur des images synthétiques. Les séquences qu'ils ont utilisées consistaient de sphères et de cylindres qui effectuaient des translations et des rotations sur leurs axes. Ils ont fait des simulations avec 2, 4, 8, 16 et 32 itérations pour obtenir des champs de déplacement convergant à une solution optimale avec leur méthode de calcul. Ils ont conclu que les vecteurs de vitesse calculés à des positions où le gradient d'intensité est non-nul sont représentatifs du mouvement apparent dans la séquence. Par contre, quand le gradient d'intensité était presque nul, les estimés n'étaient pas bien représentatifs du mouvement apparent dans la séquence. Ce mode de calcul était

satisfaisant pour les images avec beaucoup d'information structurelle (i.e., le gradient d'intensité non-nulle). Des exemples de champs de déplacement obtenus avec cette formulation peuvent être trouvés dans [1].

3.3 Méthode basée sur le modèle structurel d'ordre zéro avec régularisation

3.3.1 Dérivation

Comme discuté dans la section 2.3.1, la constance de l'intensité le long de la trajectoire de mouvement correspond au DPD égale à zéro pour une image idéale. Pour les images observées, Konrad et Dubois [5] ont inclus une contrainte de lissage afin de régulariser le problème d'estimation du mouvement apparent. La fonction objective incorporant cette régularisation est la suivante:

$$\min_{\{\mathbf{d}\}} \sum_{i=1}^N (r^2(\mathbf{d}, \mathbf{x}_i, t) + \lambda \cdot \sum_{\mathbf{x}_j \in \eta(\mathbf{x}_i)} \|\mathbf{d}(\mathbf{x}_i, t) - \mathbf{d}(\mathbf{x}_j, t)\|^2) \quad (3.1)$$

où N est le nombre de pixels de l'image. $r(\mathbf{x}_i, t)$ est défini par l'équation (2.2). Ce terme représente la différence d'intensité entre deux positions compensées par le mouvement sur deux images à t_- et t_+ . Le deuxième terme de la fonction objective est sous forme quadratique. Pour obtenir une forme quadratique pour le premier terme, une approximation linéaire du modèle structurel est faite. Pour accomplir cette tâche, seulement les termes de premier ordre d'un développement de Taylor de la variable $r(\mathbf{d}, \mathbf{x}_i, t)$ sont utilisés. Cette relation peut s'écrire:

$$r(\mathbf{d}, \mathbf{x}_i, t) \approx r(\dot{\mathbf{d}}, \mathbf{x}_i, t) + \nabla_{\mathbf{d}}^T r(\mathbf{d}, \mathbf{x}_i, t)|_{\mathbf{d}=\dot{\mathbf{d}}} \cdot (\mathbf{d} - \dot{\mathbf{d}}) \quad (3.2)$$

Dans ce développement, \mathbf{d} correspond au vecteur de déplacement courant et $\dot{\mathbf{d}}$ est un vecteur représentant les estimés intermédiaires de \mathbf{d} . Initialement, $\dot{\mathbf{d}}$ est nulle. L'approximation linéaire du modèle structurel impose des limitations sur les types de mouvement permit dans une séquence (i.e. les occlusions, les mouvements rapides et

aléatoires). L'approximation (3.2) est vraie pour aussi longtemps que la différence $\mathbf{d} - \dot{\mathbf{d}}$ est petite. Une solution est obtenue quand la valeur de énergie totale de la fonction objective atteint un minimum.

Le gradient $\nabla_{\mathbf{d}}^T r(\mathbf{d}, \mathbf{x}_i, t)|_{\mathbf{d}=\dot{\mathbf{d}}}$, s'exprime de la façon suivante:

$$\nabla_{\mathbf{d}}^T r(\mathbf{d}, \mathbf{x}_i, t)|_{\mathbf{d}=\dot{\mathbf{d}}} = \left[\begin{array}{c} \frac{\partial r(\mathbf{d}, \mathbf{x}_i, t)}{\partial d^x} \\ \frac{\partial r(\mathbf{d}, \mathbf{x}_i, t)}{\partial d^y} \end{array} \right]_{\mathbf{d}=\dot{\mathbf{d}}}$$

où $\mathbf{d} = [d^x \ d^y]^T$. Cette expression peut être réécrite en terme d'intensité lumineuse de l'image:

$$\nabla_{\mathbf{d}}^T r(\mathbf{d}, \mathbf{x}_i, t)|_{\mathbf{d}=\dot{\mathbf{d}}} = \left[\begin{array}{c} \frac{\partial g(\mathbf{x}_i + (1-\Delta_t)\dot{\mathbf{d}}, t_-)}{\partial x} (1 - \Delta_t) + \frac{\partial g(\mathbf{x}_i - \Delta_t \dot{\mathbf{d}}, t_-)}{\partial y} \Delta_t \\ \frac{\partial g(\mathbf{x}_i + (1-\Delta_t)\dot{\mathbf{d}}, t_-)}{\partial y} (1 - \Delta_t) + \frac{\partial g(\mathbf{x}_i - \Delta_t \dot{\mathbf{d}}, t_-)}{\partial x} \Delta_t \end{array} \right] \triangleq \left[\begin{array}{c} r^x \\ r^y \end{array} \right] \quad (3.3)$$

Avec le développement de cette formulation qui comprend une approximation de premier ordre, il est possible de réécrire la fonction objective comme suit:

$$\min_{\{\mathbf{d}\}} \sum_{i=1}^N (r(\dot{\mathbf{d}}, \mathbf{x}_i, t) + \nabla_{\mathbf{d}}^T r(\mathbf{d}, \mathbf{x}_i, t)|_{\mathbf{d}=\dot{\mathbf{d}}} \cdot (\mathbf{d} - \dot{\mathbf{d}}))^2 + \lambda \cdot \sum_{\mathbf{x}_j \in \eta(\mathbf{x}_i)} \| \mathbf{d}(\mathbf{x}_i, t) - \mathbf{d}(\mathbf{x}_j, t) \|^2$$

On peut obtenir une solution en appliquant la condition nécessaire d'optimalité. La combinaison de deux modèles nous donne une fonction objective non-linéaire, qui est ensuite approximée par une fonction objective quadratique dénotée par z . En différentiant par rapport à \mathbf{x}_i un système d'équation peut être obtenue:

$$\nabla_{\mathbf{d}_i} z = 0 \quad (i = 1, \dots, N) \quad (3.4)$$

Ici, z est la fonction objective devant être dérivée par rapport aux composantes du vecteur de déplacement à toutes les positions sur le plan de l'image. Les équations prennent la forme suivante:

$$\begin{cases} \frac{\partial z}{\partial d_i^x} = 2[r_i + r_i^x (d_i^x - \dot{d}_i^x) + r_i^y (d_i^y - \dot{d}_i^y)] r_i^x + 8 \cdot \lambda (d_i^x - \bar{d}_i^x) = 0 \\ \frac{\partial z}{\partial d_i^y} = 2[r_i + r_i^x (d_i^y - \dot{d}_i^y) + r_i^y (d_i^x - \dot{d}_i^x)] r_i^y + 8 \cdot \lambda (d_i^y - \bar{d}_i^y) = 0 \end{cases} \quad (i = 1, \dots, N)$$

où r^x et r^y sont définis précédemment, et la dépendance de $(\mathbf{d}, \mathbf{x}_i, t)$ a été omise pour simplifier la notation. Il est possible de regrouper les termes et de réécrire les équations dérivées ci-dessus comme suit:

$$\begin{cases} Ad_i^x + Bd_i^y = 4 \cdot \lambda \cdot \bar{d}_i^x + r_i^x D \\ Bd_i^x + Cd_i^y = 4 \cdot \lambda \cdot \bar{d}_i^y + r_i^y D \end{cases} \quad (i = 1, \dots, N)$$

avec,

$$\begin{aligned} A &= (r_i^x)^2 + 4\lambda \\ B &= (r_i^x) \cdot (r_i^y) \\ C &= (r_i^y)^2 + 4\lambda \\ D &= r_i^x \dot{d}_i^x + r_i^y \dot{d}_i^y - r_i \end{aligned}$$

où $\bar{\mathbf{d}} = [\bar{d}^x \bar{d}^y]$ représente la moyenne locale de \mathbf{d} . A partir de ce système d'équations, il est possible de déterminer chaque composante du vecteur de déplacement:

$$\begin{cases} d_i^x = \frac{(D \cdot (4 \cdot \lambda \cdot \bar{d}_i^x + r_i^x \cdot D) - B \cdot (4 \cdot \lambda \cdot \bar{d}_i^y + r_i^y \cdot D))}{(A \cdot D - B \cdot C)} \\ d_i^y = \frac{(A \cdot (4 \cdot \lambda \cdot \bar{d}_i^y + r_i^y \cdot D) - C \cdot (4 \cdot \lambda \cdot \bar{d}_i^x + r_i^x \cdot D))}{(A \cdot D - B \cdot C)} \end{cases} \quad (i = 1, \dots, N)$$

Une méthode itérative est essentielle parce que chaque vecteur de déplacement \mathbf{d}^{n+1} calculé dépend de la valeur de la moyenne des vecteurs de déplacement de l'itération précédente $\bar{\mathbf{d}}^n$. Une méthode itérative de calcul du déplacement est:

$$\begin{cases} (d_i^x)^{n+1} = \frac{(D \cdot (4 \cdot \lambda \cdot (\bar{d}_i^x)^n + r_i^x \cdot D) - B \cdot (4 \cdot \lambda \cdot (\bar{d}_i^y)^n + r_i^y \cdot D))}{(A \cdot D - B \cdot C)} \\ (d_i^y)^{n+1} = \frac{(A \cdot (4 \cdot \lambda \cdot (\bar{d}_i^y)^n + r_i^y \cdot D) - C \cdot (4 \cdot \lambda \cdot (\bar{d}_i^x)^n + r_i^x \cdot D))}{(A \cdot D - B \cdot C)} \end{cases} \quad (i = 1, \dots, N)$$

Afin d'utiliser ce système d'équations, le calcul des dérivées de premier ordre est nécessaire. Pour chaque coordonnée \mathbf{x}_i , deux positions compensées par $\mathbf{d}(\mathbf{x}_i, t)$ (figure 2.3) sont calculées sur les images à t_- et t_+ . A ces positions, les dérivées de premier ordre sont évaluées. La moyenne de ces dérivées au temps t_- et t_+ détermine la valeur de $\frac{\partial g}{\partial x}$ et $\frac{\partial g}{\partial y}$. En utilisant ces dérivées et le vecteur de l'itération précédente le nouvel estimé de \mathbf{d} à la position x_i est calculé (figure 2.3). Les dérivées spatiales sur la verticale $\frac{\partial g}{\partial y}$ doivent être divisées par un facteur de deux additionel. Ceci est

pour tenir compte du fait que chaque image est composée de deux trames, l'une, les lignes paires, et l'autre, les lignes impaires (séquence entrelacée).

Le vecteur de déplacement \mathbf{d} calculé à une certaine itération sert à réajuster de nouveau la position à laquelle les dérivées doivent être évaluées. En répétant cette procédure un certain nombre de fois de façon à minimiser la valeur de l'énergie de la fonction objective, il est possible d'obtenir un vecteur de déplacement \mathbf{d} défini par une trajectoire de mouvement qui peut ainsi établir le plus grand niveau de correspondance entre les deux intensités sur les deux trames. Ce type de calcul itératif pour compenser ou modifier la trajectoire de mouvement basée sur les résultats obtenus antérieurement est appelée "interpolation compensé par le mouvement".

3.3.2 Images de test

Trois séquences ont été utilisées pour effectuer les simulations, *courtepointe*, *femme – et – arbre* et *pingpong*. Une fenêtre de dimension 160X80 et positionnée au coin du haut à gauche (89,19) a été choisie sur la trame zéro pour la séquence *courtepointe* (figure 3.2). Cette séquence contient 120 trames et est la première séquence de test.

Une fenêtre de dimension 160X80 et positionnée au coin du haut à gauche (89,19) a été choisie sur la quatrième trame pour la séquence *femme – et – arbre* (figure 3.3). Cette séquence contient 120 trames et est la deuxième séquence de test.

Une fenêtre de dimension 180X152 et positionnée au coin du haut à gauche (180,10) a été choisie sur la quatorzième trame pour la séquence *pingpong*. Cette séquence contient 300 trames avec des scènes représentant un vaste répertoire de mouvements différentes (figure 3.4). Cette séquence est la troisième séquence de test. Seulement la composante de luminance Y a été choisie pour effectuer les calculs.

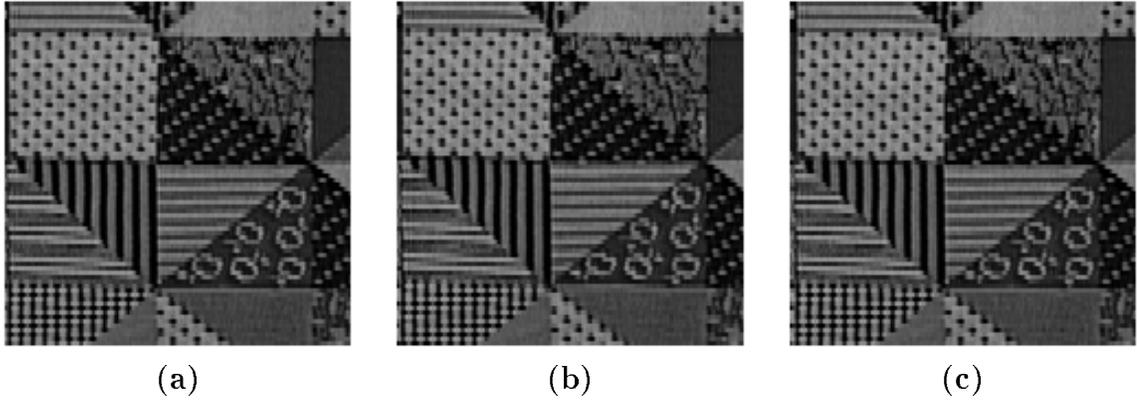


Figure 3.2: Première séquence test, *courtepointe*:(a) frame 0; (b) frame 1; (c) frame 2



Figure 3.3: Deuxième séquence test, *femme – et – arbre*:(a) frame 4; (b) frame 5; (c) frame 6

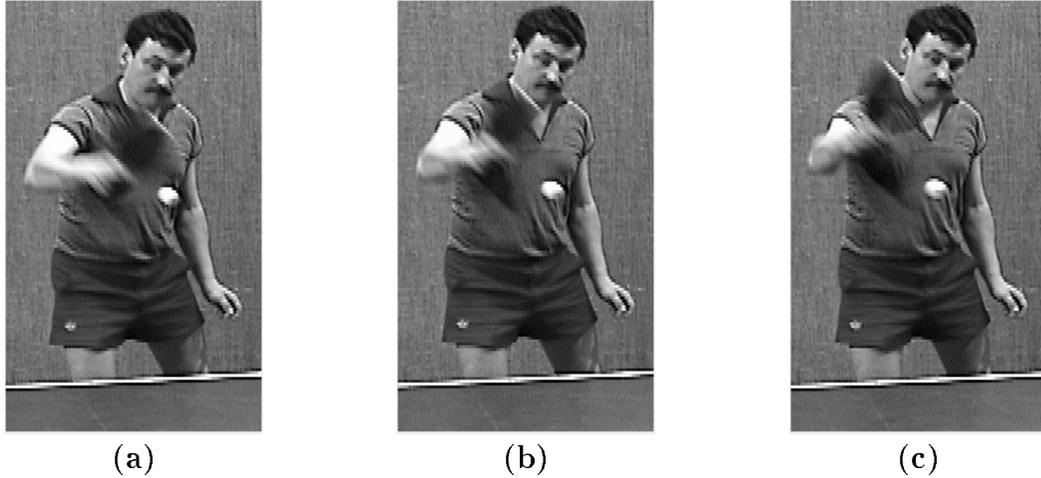


Figure 3.4: Troisième séquence test, *pingpong*:(a) trame 14; (b) trame 15; (c) trame 16

3.3.3 Résultats expérimentaux

Paramètres

Typiquement, lors de l'estimation, un certain nombre d'itérations sont effectuées de façon à minimiser le terme d'énergie pour le paramètre \mathbf{d} . Une fois cette minimisation effectuée, une substitution est faite et $\hat{\mathbf{d}}$ prend la valeur la plus récente de \mathbf{d} . Plusieurs expériences ont été faites afin de déterminer l'impact de plusieurs paramètres sur les champs de déplacement. Par exemple, le modèle structural d'ordre zéro avec régularisation a été utilisé avec un nombre variable d'itérations (i.e., un nombre variable d'étapes de relaxation Gauss-Seidel). En implantant ce modèle, deux boucles ont été définies. Une boucle pour le nombre d'itérations de la variable \mathbf{d} et l'autre, sur le nombre de substitutions de $\hat{\mathbf{d}}$ pour \mathbf{d} . Il a été constaté que, généralement, les termes d'énergie ont tendance à converger en 25 itérations pour \mathbf{d} et 5 itérations pour $\hat{\mathbf{d}}$ (tableau 3.1).

L'interpolateur cubique a été utilisé pour calculer les valeurs d'intensité et les dérivées de l'image. Les champs de déplacement obtenus par l'interpolation cubique ont été jugés les plus représentatifs du mouvement apparent dans les séquences

d'images. Une description des caractéristiques et formulations mathématiques de divers interpolateurs peut être trouvée dans l'appendice A. Des champs de déplacement ont été estimés sur la séquence *femme – et – arbre* ainsi que *pingpong*. Les résultats peuvent être inspectés à la figure 3.5.

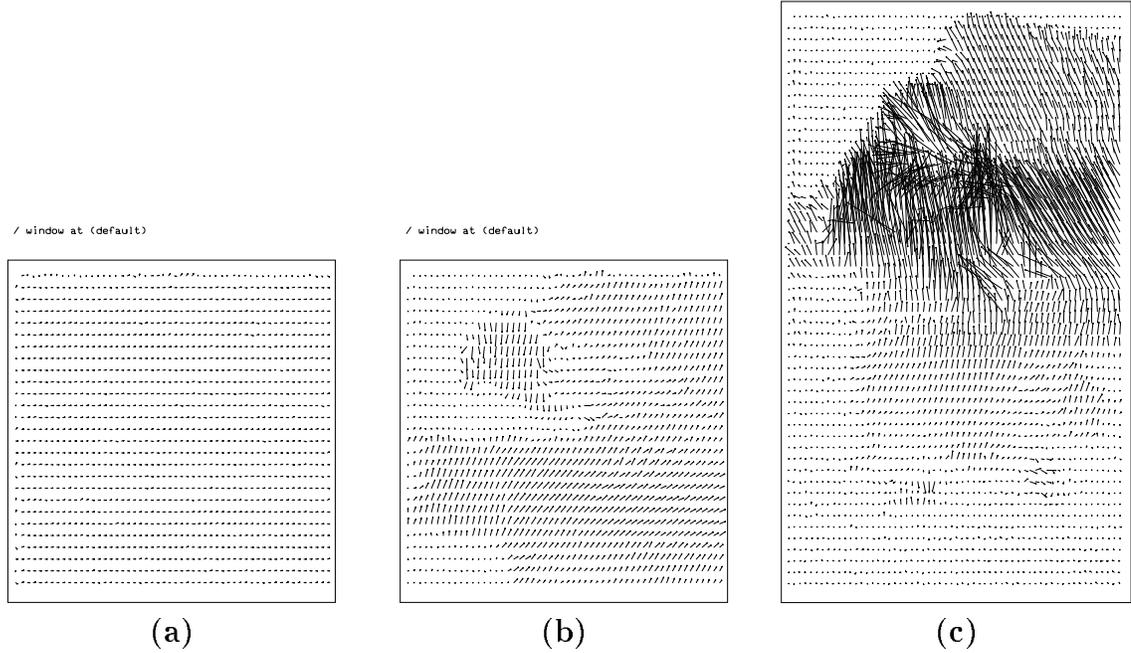


Figure 3.5: Champs de déplacement calculés avec le modèle structurel d'ordre zéro avec régularisation (λ égal 20, interpolateur cubique):(a) courtepoinette;(b) femme-et-arbre; (c) pingpong.

Variation de la valeur d'énergie

Dans le tableau 3.1, il est possible d'inspecter la variation des divers valeurs d'énergie en fonction du nombre d'itérations pour l'équation (3.1). Au début, toute l'énergie se trouve dans le modèle structurel. Ensuite, au fur et à mesure que le nombre d'itérations de la variable \mathbf{d} augmente, l'énergie du modèle structurel diminue, par contre l'énergie du modèle de lissage augmente. Le niveau de correspondance entre intensité sur deux trames augmente avec le nombre d'étapes de relaxation et donc l'erreur entre les deux termes diminue. L'augmentation de l'énergie de lissage est une

itérations		Energie	Energie	Energie
$\dot{\mathbf{d}}$	\mathbf{d}	DPD	lissage	Totale
0	0	24.504627	0.0	24.504627
0	1	7.272609	0.050279	7.775400
0	25	2.658597	0.076828	3.426874
0	75	2.544548	0.078301	3.327555
1	0	2.544548	0.078301	3.327555
1	1	1.948998	0.071648	2.665483
1	25	1.889987	0.069233	2.582318
1	75	1.878089	0.069259	2.570677
5	75	1.815826	0.070176	2.517587

Table 3.1: Changement d'énergie en fonction du nombre d'itérations pour le modèle structurel d'ordre zéro avec régularisation sur la séquence 2.

mesure de la similitude des vecteurs de déplacement avoisinant. Après un certain nombre de substitutions et d'étapes de relaxation, une convergence à une certaine valeur est établie.

3.4 Méthode fondée sur le modèle structurel de premier ordre

3.4.1 Dérivation

Dans le but d'améliorer la représentation du mouvement dans une séquence d'images, Tretiak *et al* [20] ont proposé une nouvelle formulation de calcul des vecteurs de déplacement. Celle-ci est basée sur l'hypothèse que le gradient de l'intensité lumineuse ne doit pas changer le long de la trajectoire du mouvement, exprimer par l'équation (2.3). À chaque position (x, y, t) , cette équation peut être développée de la façon

suivante:

$$\begin{cases} \frac{\partial^2}{\partial x^2} f(x, y, t) \cdot v_x + \frac{\partial^2}{\partial x \partial y} f(x, y, t) \cdot v_y + \frac{\partial^2}{\partial x \partial t} f(x, y, t) = 0 \\ \frac{\partial^2}{\partial x \partial y} f(x, y, t) \cdot v_x + \frac{\partial^2}{\partial y^2} f(x, y, t) \cdot v_y + \frac{\partial^2}{\partial y \partial t} f(x, y, t) = 0 \end{cases} \quad \forall (x, y)$$

Il est facile de résoudre ce système utilisant la méthode de Cramer pour calculer v_x et v_y . On obtient alors l'équation:

$$\mathbf{v} = -H^{-1}(x, y, t) \cdot \begin{bmatrix} \frac{\partial^2}{\partial x \partial t} f(x, y, t) \\ \frac{\partial^2}{\partial y \partial t} f(x, y, t) \end{bmatrix}$$

Dans ce cas-ci, $H(x, y, t)$ représente la matrice des dérivées de deuxième ordre de l'intensité lumineuse. Un terme mathématique souvent utilisé pour représenter cette matrice est la matrice du Hessien ou simplement le Hessien. Son expression exacte est celle-ci:

$$H = \begin{bmatrix} \frac{\partial^2}{\partial x^2} f(x, y, t) & \frac{\partial^2}{\partial x \partial y} f(x, y, t) \\ \frac{\partial^2}{\partial x \partial y} f(x, y, t) & \frac{\partial^2}{\partial y^2} f(x, y, t) \end{bmatrix}$$

Ici, les dérivées de la matrice hessienne $H(x, y, t)$ ne sont pas compensées pour le mouvement. Le mouvement ne joue aucun rôle dans l'évaluation des dérivées, contrairement au modèle couramment utilisé dans l'équation 3.3.

3.4.2 Travaux antérieurs

Il est possible d'analyser certaines propriétés de la matrice du Hessien et d'ainsi acquérir une certaine connaissance sur l'image et son mouvement, avant même d'estimer les champs de déplacement. Les valeurs propres λ_1 et λ_2 , le déterminant d_H ainsi que le numéro de conditionnement $c_H = \frac{\lambda_{min}}{\lambda_{max}}$, sont tous des propriétés qui sont utilisées pour déterminer la stabilité et la précision des champs de déplacement à représenter le mouvement apparent d'une séquence. Le rôle que ces paramètres jouent pour déterminer la précision des estimés du mouvement apparent a été discuté par Nagel dans [18] et [19]. Si $d_H \gg 1$ et $c_H \approx 1$, les estimés obtenus seront représentatif du mouvement apparent comme proposé par Lanczos dans [6] et Kearny dans [11].

D'autres chercheurs comme Verri *et al.* [24] ont analysé $tr(H)$ (la trace) et $det(H)$ (le déterminant) pour déterminer le type de mouvement présent dans une séquence d'image. Ils ont tracé $Tr(H)$ en fonction de $det(H)$, pour distinguer un mouvement de rotation, d'une translation et d'un mouvement général. D'après ces analyses, la matrice du Hessien contient des propriétés pouvant servir à définir plusieurs aspects du mouvement estimé ainsi que certaines propriétés de la scène elle-même.

DeMicheli *et al.* [6] ont calculé des champs de déplacements mettant en application l'équation (2.3). À cause de la sensibilité au bruit du calcul de dérivées de deuxième ordre, les champs étaient pas représentatif du mouvement apparent. Afin d'améliorer la représentation de leurs estimés, ils ont utilisé un filtrage sur le champ de déplacement après estimation. Les filtres utilisés étaient de type Gaussien à deux dimensions avec une variance de 0.5, 1 et 3 et une moyenne de zéro. Ils ont aussi effectué un pré-filtrage spatial et temporel sur les séquences d'images avant l'estimation. Les dérivés ont été estimé avec des interpolateurs discrets utilisant une expansion de Taylor à cinq points (Appendice A). Utilisant leurs estimés du mouvement apparent, ils ont essayé d'établir le type de mouvement présent dans une séquence d'images (translation, rotation, mouvement le long d'un axe). De plus, ils ont estimé certains paramètres de scène (temps de collision, vitesse angulaire, etc).

Pastor et Tretiak [20] ont aussi calculé des estimés du mouvement apparent en utilisant cette formulation. Dans leurs travaux, ils ont déterminé la sensibilité des estimés face à certains paramètres. Ils ont effectué seulement un filtrage spatial de leurs séquences. De plus, ils ont estimé les dérivées de deuxième ordre avec des interpolateurs discrets [20]. Les champs de vitesse qu'ils ont obtenu n'étaient pas très représentatif du mouvement apparent dans une séquence dans certain cas.

3.4.3 Résultats expérimentaux

Seuil et interpolateurs

À partir des travaux effectués avec ce modèle quelques expériences ont été menées

afin d'améliorer les estimés du mouvement apparent obtenus utilisant ce modèle. Un paramètre très significatif et particulier à la méthode proposée par Pastor *et al.* [20] est le seuil du déterminant, θ , qui limite l'utilisation de l'équation (2.3). Si $\det(H) \geq \theta$, les vecteurs de vitesse correspondants sont calculés à partir de l'équation (2.3). Par contre, si $\det(H) < \theta$, les vecteurs correspondants sont mis à zéro.

Dans les résultats qui suivent, il est possible de voir l'effet du seuil sur les champs de vitesse obtenus avec ce type d'algorithme. Le seuil a été fixé à 0.01, 0.1 et 1 (figure 3.6). Dans cette figure, on peut voir l'effet de ce paramètre sur le champ des vitesses. En limitant le seuil à une certaine valeur, on se trouve à limiter $\det(H)$, et donc d'exclure certains estimés du champ de vitesse résultant.

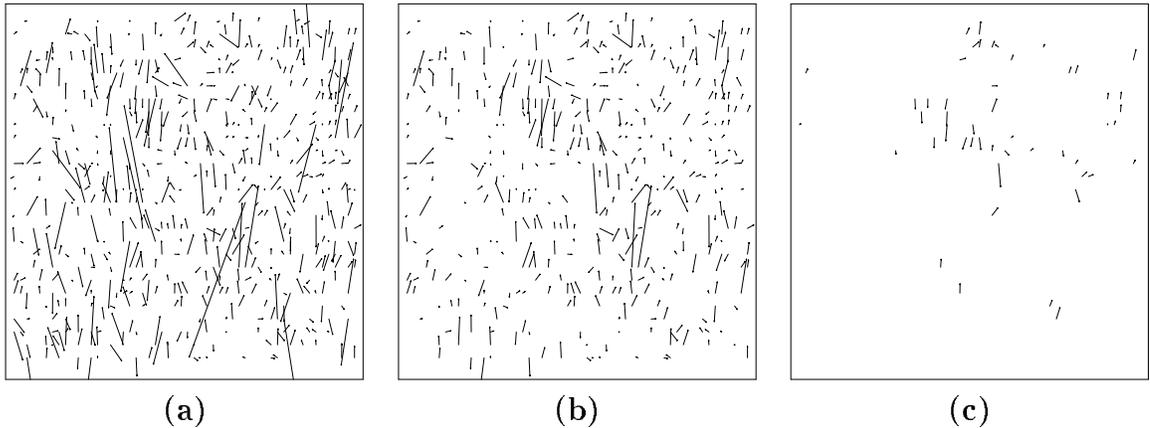


Figure 3.6: Champs de déplacement obtenus utilisant le modèle structurel d'ordre un sur la séquence 2 (sans filtrage) avec θ égal à:(a) 0.01; (b) 0.1; (c) 1

Les estimés du mouvement apparent utilisant des interpolateurs discrets de Taylor à trois et cinq points ont été implémentés [10] (figure 3.7). Ces interpolateurs discrets ne permettaient pas d'utiliser une compensation de mouvement parce que les dérivées pouvaient seulement être évaluées à des positions discrètes sur le plan de l'image. Pour utiliser l'information du mouvement (**d**) dans l'estimation des champs de vitesse, un interpolateur cubique régulier a été utilisé (appendice A). Les champs de vitesse avec cette modification n'étaient pas plus représentatif du mouvement apparent. Le champ des vitesses a été évalués avec cet interpolateur avec et

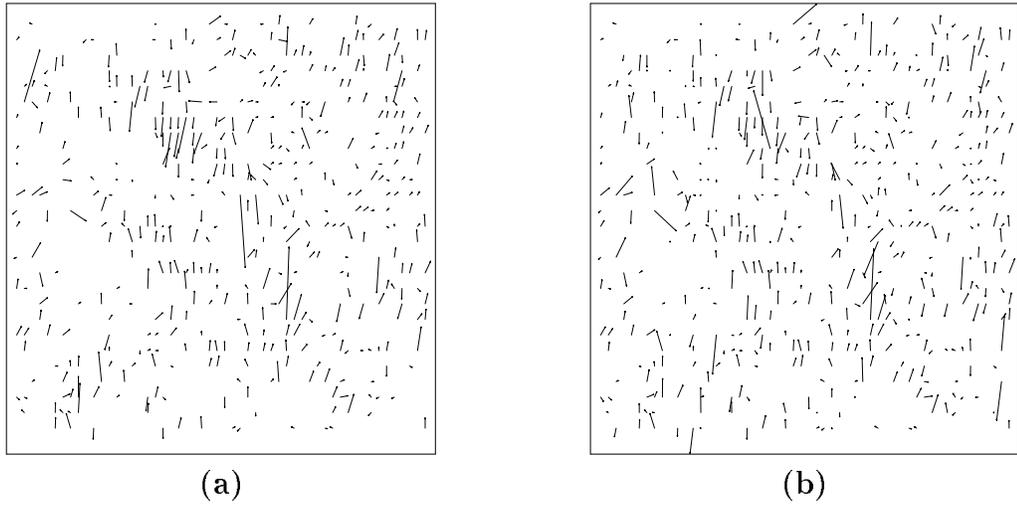


Figure 3.7: Champs de vitesse obtenus avec le modèle structural d'ordre un avec $\theta = 0.1$ pour la séquence de test 2 avec divers interpolateurs ayant des expansions de Taylor à: (a) 3 points; (b) 5 points

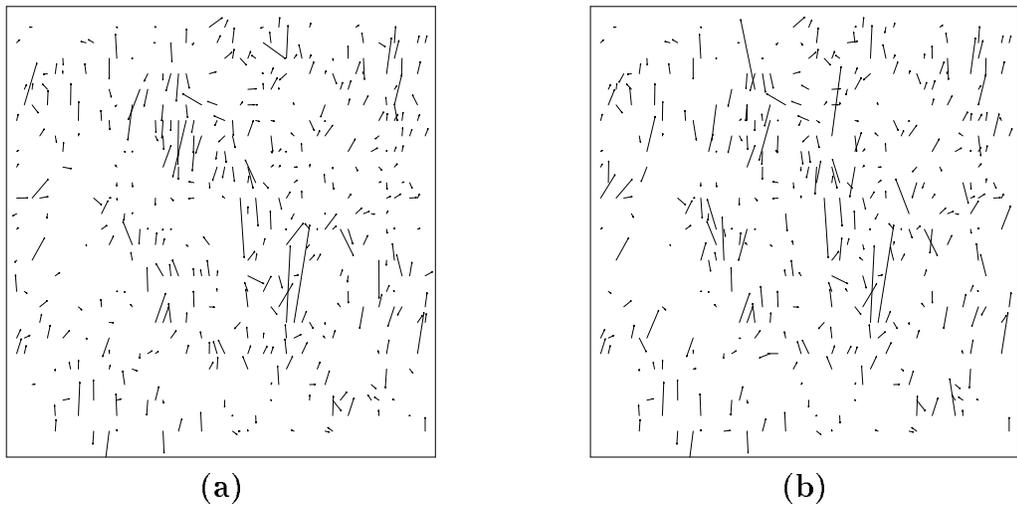


Figure 3.8: Champs de vitesse obtenus avec le modèle structural d'ordre un et $\theta = 0.1$ pour la séquence 2 avec divers interpolateurs: (a) cubique; (b) cubique avec compensation de mouvement

sans compensation de mouvement. Les résultats peuvent être inspectés à la figure 3.8. Le long de la main, des épaules et de la joue, les vecteurs sont les plus représentatifs du mouvement apparent. On peut conclure que la compensation n'a pas permis une amélioration notable. Les vecteurs vitesse pointent dans la direction du mouvement apparent de la séquence *femme – et – arbre* dans les régions de contours d'objets (i.e. la main droite et l'épaule gauche). A l'intérieur des contours d'objets, comme dans la région interne du bras, de la tête et ainsi de suite, les vecteurs n'étaient pas caractéristiques du mouvement observable de la séquence. Comme on peut le constater à partir des résultats obtenus, les vecteurs dans ces régions ont des amplitudes et des orientations très aléatoires, qui n'ont rien à voir avec le mouvement apparent.

Différents filtrages ont été effectués sur les champs de déplacement obtenus avec compensation de mouvement. Des filtres de variance égale à 0.5, 1 et 3 ont été utilisés dans les simulations (figure 3.9). Différentes valeurs de variance ont été utilisées de façon à déterminer l'impact de ce paramètre sur le champ des vitesses estimés. Il est évident que le filtrage a uniformisé les estimés. En augmentant la variance dans les régions de la main droite et de l'épaule gauche, les vecteurs calculés ont été plus représentatif du mouvement apparent. Dans d'autres régions, les estimés étaient uniformisés selon la variance. Ils n'étaient pas très représentatif du mouvement apparent.

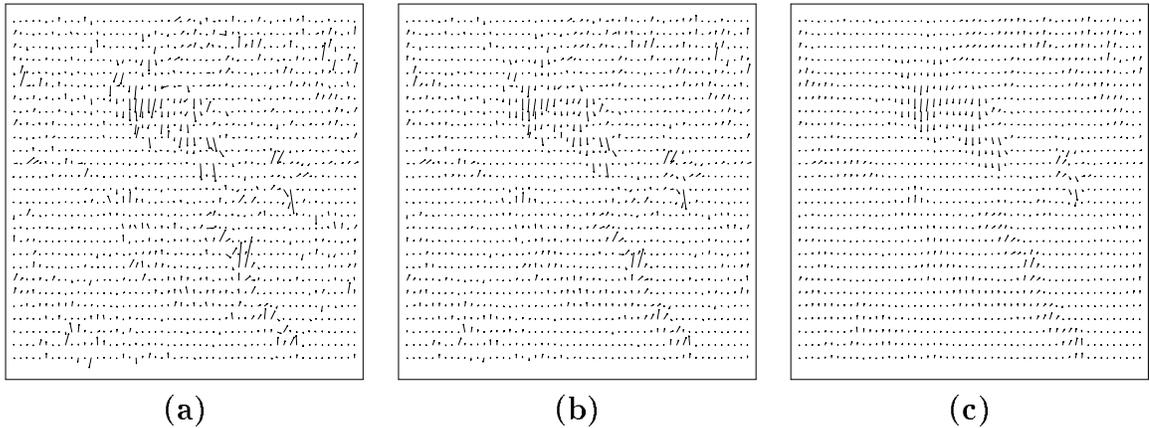


Figure 3.9: Champs de vitesse obtenus sur la séquence 2 avec $\theta = 0.1$ après filtrage avec un filtre gaussien à variance égal à: (a) 0.5; (b) 1; (c) 3.

ECM et calcul itératif

Un critère de performance des champs de déplacement doit être introduit. Pour chaque champ de déplacement, l'erreur carrée moyenne (ECM) est calculée. Dans ce cas, l'erreur carrée moyenne est obtenue en calculant la différence de deux champs de vitesse d'une itération à une autre.

$$ECM = \sum_{i=1}^N \| \mathbf{v}^n(\mathbf{x}_i, t) - \mathbf{v}^{n-1}(\mathbf{x}_i, t) \|^2$$

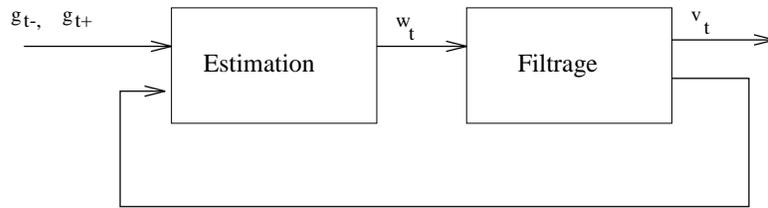


Figure 3.10: Méthode itérative pour évaluer les champs de déplacement avant et après filtrage.

Maintenant qu'un critère d'évaluation des champs de vitesse est établi (ECM), une approche itérative a été introduite pour essayer d'améliorer la représentation du mouvement apparent. Initialement, un champ de vitesse a été estimé puis filtré (filtre gaussien 2D à variance de 1, figure 3.10). Il est important de noter que l'estimation et le filtrage sont faits en deux étapes différentes. Ce résultat a été utilisé pour ajuster les coordonnées auxquelles sont évaluées les dérivés et ensuite recalculer un nouveau champ de vitesse. Après un certain nombre d'itérations, le ECM a diminué de son maximum initial. Pour fin d'expérimentation, 40 itérations ont été choisies. Après ce grand nombre d'itérations, il a été possible de constater que le ECM avait tendance à converger ou même à osciller vers une valeur finale. Une fois cette constatation établie, il était possible de limiter le nombre d'itérations (tableau 3.2). Dans ce tableau, on peut voir le comportement du ECM en fonction du nombre d'itérations dans cette approche. À la figure 3.11, 10 itérations suivi de filtrage ont été effectués.

iterations	MSE (original)	MSE (filtré)
1	0.5123	0.0869
10	0.2246	0.0385
20	0.1107	0.0182
30	0.1159	0.0202
40	0.1242	0.0241

Table 3.2: Changement du MSE en fonction du nombre d'itérations.

Selon ces résultats, avant le filtrage, le champ des vitesses ne représente pas bien le mouvement apparent. Cependant, après le filtrage, le long de la main droite et de l'épaule gauche (comme la figure 4.6 (c)) le champ des vitesses est plus représentatif du mouvement apparent. On peut conclure d'après ces résultats que le filtrage joue un rôle important dans l'estimation du mouvement apparent.

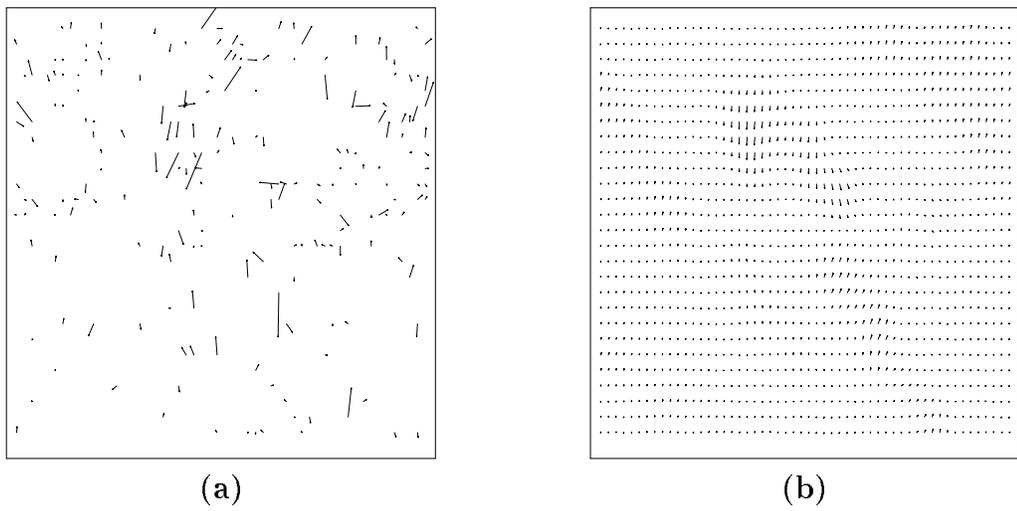


Figure 3.11: Champs de vitesse calculés pour la séquence 2 avec le modèle structurel d'ordre un avec régularisation après 10 itérations:(a) avant filtrage (w_t); (b) après filtrage (v_t).

Chapitre 4

Estimation du mouvement apparent – nouvelle formulation

4.1 Méthode fondée sur le modèle structurel d'ordre un avec régularisation

4.1.1 Dérivation

La contribution de ce travail de recherche est de combiner le modèle structurel basé sur l'hypothèse que le gradient de l'intensité ne change pas le long de la trajectoire de mouvement avec le modèle de mouvement en utilisant la régularisation. Donc, l'estimation et le filtrage de champs de déplacement proposés dans [6] sont combinés en une seule étape. Cette formulation est calculée d'une façon similaire à celle proposée dans la section 3.3.

Le critère de minimisation est le suivant:

$$\min_{\{\mathbf{d}\}} \sum_{i=1}^N (\| \mathbf{s}(\mathbf{d}_i, \mathbf{x}_i, t) \|^2 + \lambda \cdot \sum_{\mathbf{x}_j \in \eta(\mathbf{x}_i)} \| \mathbf{d}(\mathbf{x}_i, t) - \mathbf{d}(\mathbf{x}_j, t) \|^2)$$

où, $\mathbf{s}(\mathbf{d}_i, \mathbf{x}_i, t)$ est défini par l'équation (2.4). La fonction objective que l'on obtient est non-linéaire. Le modèle de mouvement (le deuxième terme de la fonction objective) est sous forme quadratique. Pour obtenir une fonction objective quadratique,

le modèle structurel de premier ordre sera approximé par les termes linéaires de son développement de Taylor. Ceci est fait en supposant que la compensation du mouvement est utilisée seulement pour de petits déplacements. L'approximation prend donc la forme d'une équation vectorielle:

$$\mathbf{s}(\mathbf{d}, \mathbf{x}_i, t) \approx \mathbf{s}(\dot{\mathbf{d}}, \mathbf{x}_i, t) + \nabla_{\mathbf{d}}^T \mathbf{s}(\mathbf{d}, \mathbf{x}_i, t)|_{\mathbf{d}=\dot{\mathbf{d}}} \cdot (\mathbf{d} - \dot{\mathbf{d}})$$

L'approximation présentée ci-dessus correspond à une linéarisation locale. Cette opération est nécessaire pour garantir que la fonction objective est de forme quadratique. Par l'équation (2.4), le gradient de cette approximation s'exprime de la façon suivante:

$$\begin{aligned} \nabla_{\mathbf{d}}^T \mathbf{s}(\mathbf{d}, \mathbf{x}_i, t)|_{\mathbf{d}=\dot{\mathbf{d}}} &= \nabla_{\mathbf{d}}^T (\nabla_{\mathbf{x}} g(\mathbf{x}_i + (1 - \Delta_t) \cdot \mathbf{d}(\mathbf{x}_i, t), t_+))|_{\mathbf{d}=\dot{\mathbf{d}}} - \\ &\quad \nabla_{\mathbf{d}}^T (\nabla_{\mathbf{x}} g(\mathbf{x}_i - \Delta_t \cdot \mathbf{d}(\mathbf{x}_i, t), t_-))|_{\mathbf{d}=\dot{\mathbf{d}}} \end{aligned}$$

Chaque gradient d'intensité de l'image observée est approximé par les termes linéaires de son expansion de Taylor. Développons l'opérateur $\nabla_{\mathbf{d}}^T (\nabla_{\mathbf{x}} g(\mathbf{x}_i + (1 - \Delta_t) \cdot \mathbf{d}(\mathbf{x}_i, t), t_+))$:

$$\begin{aligned} \nabla_{\mathbf{d}}^T (\nabla_{\mathbf{x}} g(\mathbf{x}_i + (1 - \Delta_t) \cdot \mathbf{d}(\mathbf{x}_i, t), t_+)) &= \nabla_{\mathbf{d}}^T \cdot \begin{bmatrix} \frac{\partial g(\mathbf{x}_i + (1 - \Delta_t) \cdot \dot{\mathbf{d}}, t_+)}{\partial x} \\ \frac{\partial g(\mathbf{x}_i + (1 - \Delta_t) \cdot \dot{\mathbf{d}}, t_+)}{\partial y} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial^2 g(\mathbf{x}_i + (1 - \Delta_t) \cdot \dot{\mathbf{d}}, t_+)}{\partial x \partial d^x} & \frac{\partial^2 g(\mathbf{x}_i + (1 - \Delta_t) \cdot \dot{\mathbf{d}}, t_+)}{\partial x \partial d^y} \\ \frac{\partial^2 g(\mathbf{x}_i + (1 - \Delta_t) \cdot \dot{\mathbf{d}}, t_+)}{\partial y \partial d^x} & \frac{\partial^2 g(\mathbf{x}_i + (1 - \Delta_t) \cdot \dot{\mathbf{d}}, t_+)}{\partial y \partial d^y} \end{bmatrix} \end{aligned}$$

en introduisant les dérivées successives, est équivalent à:

$$\nabla_{\mathbf{d}}^T (\nabla_{\mathbf{x}} g(\mathbf{x}_i + (1 - \Delta_t) \cdot \mathbf{d}(\mathbf{x}_i, t), t_+)) = \begin{bmatrix} \frac{\partial^2 g(\mathbf{x}_i + (1 - \Delta_t) \cdot \dot{\mathbf{d}}, t_+)}{\partial x^2} \cdot \frac{\partial x}{\partial d^x} & \frac{\partial^2 g(\mathbf{x}_i + (1 - \Delta_t) \cdot \dot{\mathbf{d}}, t_+)}{\partial x \partial y} \cdot \frac{\partial y}{\partial d^x} \\ \frac{\partial^2 g(\mathbf{x}_i + (1 - \Delta_t) \cdot \dot{\mathbf{d}}, t_+)}{\partial x \partial y} \cdot \frac{\partial x}{\partial d^y} & \frac{\partial^2 g(\mathbf{x}_i + (1 - \Delta_t) \cdot \dot{\mathbf{d}}, t_+)}{\partial y^2} \cdot \frac{\partial y}{\partial d^y} \end{bmatrix}$$

Donc:

$$\nabla_{\mathbf{d}}^T (\nabla_{\mathbf{x}} g(\mathbf{x}_i + (1 - \Delta_t) \cdot \mathbf{d}(\mathbf{x}_i, t), t_+)) = (1 - \Delta_t) \cdot \begin{bmatrix} \frac{\partial^2 g(\mathbf{x}_i + (1 - \Delta_t) \cdot \dot{\mathbf{d}}, t_+)}{\partial x^2} & \frac{\partial^2 g(\mathbf{x}_i + (1 - \Delta_t) \cdot \dot{\mathbf{d}}, t_+)}{\partial x \partial y} \\ \frac{\partial^2 g(\mathbf{x}_i + (1 - \Delta_t) \cdot \dot{\mathbf{d}}, t_+)}{\partial x \partial y} & \frac{\partial^2 g(\mathbf{x}_i + (1 - \Delta_t) \cdot \dot{\mathbf{d}}, t_+)}{\partial y^2} \end{bmatrix}$$

parce que $\frac{\partial x}{\partial d^x} = (1 - \Delta_t)$ et $\frac{\partial y}{\partial d^y} = (1 - \Delta_t)$ et:

$$\nabla_{\mathbf{d}}^T (\nabla_{\mathbf{x}} g(\mathbf{x}_i - \Delta_t \cdot \mathbf{d}(\mathbf{x}_i, t), t_-)) = -\Delta_t \cdot \begin{bmatrix} \frac{\partial^2 g(\mathbf{x}_i - \Delta_t \cdot \dot{\mathbf{d}}, t_-)}{\partial x^2} & \frac{\partial^2 g(\mathbf{x}_i - \Delta_t \cdot \dot{\mathbf{d}}, t_-)}{\partial x \partial y} \\ \frac{\partial^2 g(\mathbf{x}_i - \Delta_t \cdot \dot{\mathbf{d}}, t_-)}{\partial x \partial y} & \frac{\partial^2 g(\mathbf{x}_i - \Delta_t \cdot \dot{\mathbf{d}}, t_-)}{\partial y^2} \end{bmatrix}$$

parce que $\frac{\partial x}{\partial d^x} = -\Delta_t$ et $\frac{\partial y}{\partial d^y} = -\Delta_t$. Finalement:

$$\mathbf{s}(\mathbf{d}, \mathbf{x}_i, t) \approx \mathbf{s}(\dot{\mathbf{d}}, \mathbf{x}_i, t) + \mathbf{G} \cdot (\mathbf{d} - \dot{\mathbf{d}})$$

où,

$$\mathbf{G} \triangleq \begin{bmatrix} G_{xx} & G_{xy} \\ G_{xy} & G_{yy} \end{bmatrix}$$

et,

$$\begin{aligned} G_{xx} &\triangleq \frac{\partial^2 g(\mathbf{x}_i + (1 - \Delta_t) \cdot \dot{\mathbf{d}}(\mathbf{x}_i, t), t_+)}{\partial x^2} \cdot (1 - \Delta_t) + \frac{\partial^2 g(\mathbf{x}_i - \Delta_t \cdot \dot{\mathbf{d}}(\mathbf{x}_i, t), t_-)}{\partial x^2} \cdot \Delta_t \\ G_{xy} &\triangleq \frac{\partial^2 g(\mathbf{x}_i + (1 - \Delta_t) \cdot \dot{\mathbf{d}}(\mathbf{x}_i, t), t_+)}{\partial x \partial y} \cdot (1 - \Delta_t) + \frac{\partial^2 g(\mathbf{x}_i - \Delta_t \cdot \dot{\mathbf{d}}(\mathbf{x}_i, t), t_-)}{\partial x \partial y} \cdot \Delta_t \\ G_{yy} &\triangleq \frac{\partial^2 g(\mathbf{x}_i + (1 - \Delta_t) \cdot \dot{\mathbf{d}}(\mathbf{x}_i, t), t_+)}{\partial y^2} \cdot (1 - \Delta_t) + \frac{\partial^2 g(\mathbf{x}_i - \Delta_t \cdot \dot{\mathbf{d}}(\mathbf{x}_i, t), t_-)}{\partial y^2} \cdot \Delta_t \end{aligned}$$

Les dérivées de deuxième ordre sont compensées par le vecteur de déplacement. Avec cette simplification, il est possible de réexprimer la fonction objective comme suit:

$$\min_{\{\mathbf{d}\}} \sum_{i=1}^N (\| \mathbf{s}(\dot{\mathbf{d}}, \mathbf{x}_i, t) + \mathbf{G} \cdot (\mathbf{d} - \dot{\mathbf{d}}) \|^2 + \lambda \cdot \sum_{\mathbf{x}_j \in \eta(\mathbf{x}_i)} \| \mathbf{d}(\mathbf{x}_i, t) - \mathbf{d}(\mathbf{x}_j, t) \|^2)$$

Pour effectuer cette minimisation, une condition d'optimalité est nécessaire. Posons z_1 , le premier terme de la fonction objective et z_2 le deuxième terme. Pour chaque position \mathbf{x}_i , on doit résoudre le système suivant:

$$\begin{aligned} \frac{\partial z}{\partial d_i^x} &= \frac{\partial z_1}{\partial d_i^x} + \frac{\partial z_2}{\partial d_i^x} \\ \frac{\partial z}{\partial d_i^y} &= \frac{\partial z_1}{\partial d_i^y} + \frac{\partial z_2}{\partial d_i^y} \end{aligned}$$

La dérivée de la fonction objective en terme de la première composante du vecteur de déplacement $\frac{\partial z_1}{\partial d_i^x}$ s'exprime de la façon suivante (le modèle structural seulement):

$$\begin{aligned} \frac{\partial z_1}{\partial d_i^x} &= 2 \cdot G_{xx} \cdot (s_1 + G_{xx} \cdot (d_i^x - \dot{d}_i^x) + G_{xy} \cdot (d_i^y - \dot{d}_i^y))^2 + \\ & 2 \cdot G_{xy} \cdot (s_2 + G_{xy} \cdot (d_i^x - \dot{d}_i^x) + G_{yy} \cdot (d_i^y - \dot{d}_i^y))^2 \end{aligned}$$

où s_1 et s_2 sont évalués à $\dot{\mathbf{d}}$. La dérivée du deuxième terme $\frac{\partial z_2}{\partial d_i^x}$ (le modèle de mouvement) s'exprime de la façon suivante pour chaque coordonné \mathbf{x}_i (figure 2.6):

$$\begin{aligned}\frac{\partial z_2}{\partial d_i^x} &= 2 \cdot \lambda \cdot [(d_i^x - d_{ih}^x) + (d_i^x - d_{ig}^x) - (d_{id}^x - d_i^x) - (d_i^x - d_i^x)] \\ &= 2 \cdot \lambda \cdot [4d_i^x - (d_{ih}^x + d_{ig}^x + d_{id}^x + d_{ib}^x)] = 8 \cdot \lambda \cdot [d_i^x - \bar{d}_i^x] = 0\end{aligned}$$

Ici $\bar{\mathbf{d}}$ est une moyenne locale de \mathbf{d} par rapport au voisinage $\eta(x_i)$. Un développement identique est nécessaire pour dériver la composante $\frac{\partial z}{\partial d_i^y}$. Après quelques simplifications et regroupement de termes, on obtient ce système d'équations:

$$\begin{cases} Ad_i^x + Bd_i^y = D \cdot G_{xx} + E \cdot G_{xy} + 4\lambda \bar{d}_i^x \\ Bd_i^x + Cd_i^y = D \cdot G_{xy} + E \cdot G_{yy} + 4\lambda \bar{d}_i^y \end{cases} \quad (i = 1, \dots, N)$$

où les constantes ont les valeurs suivantes:

$$\begin{aligned}A &= 4\lambda + G_{xx}^2 + G_{xy}^2 \\ B &= G_{xy}(G_{xx} + G_{yy}) \\ C &= 4\lambda + G_{yy}^2 + G_{xy}^2 \\ D &= G_{xx}\dot{d}_i^x + G_{xy}\dot{d}_i^y - s_1 \\ E &= G_{xy}\dot{d}_i^x + G_{yy}\dot{d}_i^y - s_2 \\ s_1 &= \frac{\partial g(\mathbf{x}_i + (1 - \Delta_t) \cdot \dot{\mathbf{d}}(\mathbf{x}_i, t), t_+)}{\partial x} - \frac{\partial g(\mathbf{x}_i - \Delta_t \cdot \dot{\mathbf{d}}(\mathbf{x}_i, t), t_-)}{\partial x} \\ s_2 &= \frac{\partial g(\mathbf{x}_i + (1 - \Delta_t) \cdot \dot{\mathbf{d}}(\mathbf{x}_i, t), t_+)}{\partial y} - \frac{\partial g(\mathbf{x}_i - \Delta_t \cdot \dot{\mathbf{d}}(\mathbf{x}_i, t), t_-)}{\partial y}\end{aligned}$$

Avec ce système d'équations, il est possible de résoudre le système matriciel en utilisant la méthode de Cramer. On obtient alors les deux équations suivantes pour chaque composante du vecteur de déplacement:

$$\begin{cases} d_i^x = \frac{(C \cdot (D \cdot G_{xx} + E \cdot G_{xy} + 4\lambda \bar{d}_i^x) - B \cdot (D \cdot G_{xy} + E \cdot G_{yy} + 4\lambda \bar{d}_i^y))}{(A \cdot C - B \cdot B)} \\ d_i^y = \frac{(C \cdot (D \cdot G_{xy} + E \cdot G_{yy} + 4\lambda \bar{d}_i^y) - B \cdot (D \cdot G_{xx} + E \cdot G_{xy} + 4\lambda \bar{d}_i^x))}{(A \cdot C - B \cdot B)} \end{cases} \quad (i = 1, \dots, N)$$

Ce système peut être résolu de façon récursive comme ceci:

$$\begin{cases} (d_i^x)^{n+1} = \frac{(C \cdot (D \cdot G_{xx} + E \cdot G_{xy} + 4\lambda (\bar{d}_i^x)^n) - B \cdot (D \cdot G_{xy} + E \cdot G_{yy} + 4\lambda (\bar{d}_i^y)^n))}{(A \cdot C - B \cdot B)} \\ (d_i^y)^{n+1} = \frac{(C \cdot (D \cdot G_{xy} + E \cdot G_{yy} + 4\lambda (\bar{d}_i^y)^n) - B \cdot (D \cdot G_{xx} + E \cdot G_{xy} + 4\lambda (\bar{d}_i^x)^n))}{(A \cdot C - B \cdot B)} \end{cases} \quad (i = 1, \dots, N)$$

4.1.2 Résultats expérimentaux

Des expériences similaires à celles de la section 3.3.2 ont été refaites pour le modèle structurel d'ordre zéro avec régularisation. Les champs de déplacement sur les séquences 1, 2 et 3 sont démontrés dans la figure 4.1.

Contrainte de lissage

L'effet de la contrainte de lissage sur les champs de déplacement a été examiné. Un nouvel aspect doit être considéré. La contrainte de lissage pour le modèle structurel de premier ordre avec régularisation a été déterminée expérimentalement à environ 10 fois inférieure à la même contrainte utilisée avec le modèle structurel d'ordre zéro avec régularisation. Des champs de déplacement ont été calculés avec les deux algorithmes. Quand la contrainte de lissage était environ 10 pour l'algorithme de constance de DPD, des champs de déplacement similaires mais non identiques ont été obtenus avec l'algorithme de constance de DGD avec une valeur de la contrainte de lissage égal à environ 1. Ceci démontre clairement que l'hypothèse structurelle de base a un impact important sur l'estimation du mouvement apparent dans une séquence d'images. L'impact de la contrainte de lissage sur les champs de déplacement peut être analysé en comparant les résultats de la figure 4.2 (a) et (b).

Interpolateurs utilisés

En plus, des simulations avec la contrainte de lissage, d'autres simulations ont été entreprises avec des interpolateurs différents. Les interpolateurs sont présentés dans l'appendice A. Les interpolateurs quadratique, cubique, cubique de Keys [22] et quartique ont été implémentés. L'interpolateur cubique donne les champs de déplacement jugés les plus caractéristiques du mouvement apparent dans les séquences d'images (figures 4.3 et 4.4 pour la séquence *femme – et – arbre*). Une inspection de la figure 6.4 démontre que l'interpolateur cubique a une fonction continue avec de petite pente. Donc, les estimés auront tendance à être plus uniformes.

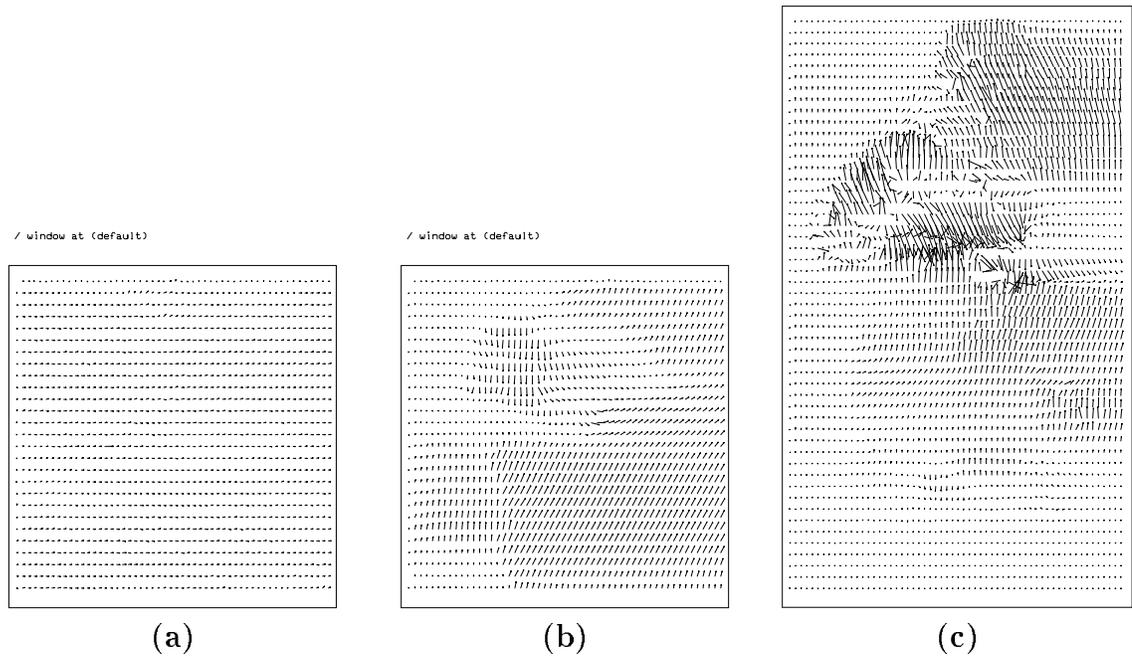


Figure 4.1: Champs de déplacement calculés avec la nouvelle formulation (interpolateur cubique), sur les séquences $(\lambda = 5)$: (a) courtepointe; (b) femme-et-arbre; (c) pingpong.

Les champs de déplacement estimés utilisant les autres interpolateurs ont les caractéristiques suivantes. Une évaluation visuelle fut effectuée pour déterminer l'interpolateur donnant les vecteurs de déplacement les plus représentatifs du mouvement apparent dans la séquence d'image. Pour la séquence *femme – et – arbre*, l'interpolateur quadratique a donné un champ de déplacement avec plusieurs croisements de vecteurs aux contours de la main droite et la manche ainsi que la joue (figure 4.3 (a)). Le champ de déplacement obtenu par calcul de l'interpolateur cubique de Keys avait des vecteurs de déplacement avec des petites amplitudes dans la joue et des vecteurs de déplacement avec une mauvaise orientation dans la région de l'épaule (figure 4.4 (a)). L'interpolateur quartique contenait des croisements de vecteurs de déplacement dans la région de la main droite et la manche. De plus, il y avait un manque de résolution dans la région de la joue (figure 4.4 (b)). L'interpolateur cubique contenait le moins de croisements de vecteurs et la meilleure résolution dans la région de la joue (figure 4.3 (b)).

Pour la séquence *pingpong*, le champ de déplacement obtenu par calcul avec l'interpolateur cubique donne une meilleure résolution du thorax de l'homme avec la raquette que tout autre interpolateur. De plus, l'orientation et l'amplitude des vecteurs de déplacement est représentative du mouvement apparent dans la séquence. Des observations similaires s'appliquent pour la séquence *courtepointe*.

D'après ces résultats, on peut conclure que le mode de calcul des dérivées peut jouer un rôle très important sur les champs de déplacement et la représentation du mouvement apparent dans une séquence.

Variation d'énergie

Le tableau 4.1 représente la variation des différentes énergies impliquées dans le calcul des vecteurs de déplacement. Les mêmes constatations faites pour le modèle d'ordre zéro avec régularisation s'appliquent pour cet algorithme. Après un certain nombre d'itérations, l'énergie a tendance à converger à une certaine valeur.

itérations		DGD	lissage	Totale
$\mathbf{\hat{d}}$	\mathbf{d}	Energie	Energie	Energie
0	0	3.184218	0.0	3.184218
0	1	1.513813	0.211519	1.725332
0	25	1.099988	0.590289	1.690277
0	75	1.119520	0.640120	1.759640
1	0	1.119520	0.640120	1.759640
1	1	0.933616	0.641396	1.575012
1	25	0.705968	0.361020	1.066988
1	75	0.704501	0.359984	1.064485
5	75	0.601718	0.396102	0.997820

Table 4.1: Variation d'énergie avec le nombre d'itérations pour le modèle structurel de premier ordre avec régularisation sur la séquence 2.

4.2 Estimation du mouvement apparent par multi-résolution

Mode d'évaluation

Contrairement aux travaux effectués antérieurement où l'estimation de champs de déplacement était faite sur un seul niveau, l'estimation des champs de déplacement ici, se fait sur plusieurs niveaux: trois pour l'algorithme de constance de DGD et quatre pour l'algorithme de constance de DPD (vu que l'on est intéressé à reconstruire des trames). Cette différence entre niveaux d'estimation vient de la différence d'ordre de dérivée requise pour utiliser chaque algorithme. Pour l'hypothèse de constance de DPD, les dérivées de premier ordre seulement sont requises. Pour l'hypothèse de constance de DGD, les dérivées de premier ainsi que de deuxième ordre sont requises. Vu que le bas de la pyramide (le niveau 0) n'est pas filtré, les estimés de dérivées de

deuxième ordre sont très mauvais (figure 4.6).

Les méthodes de multi-résolution permettent le calcul des champs de déplacement à chaque niveau de la pyramide et permettent ainsi de diminuer d'au moins quatre fois le temps de calcul de chaque niveau par rapport à un autre. Sur le haut de la pyramide, chaque quatre vecteurs de mouvement est calculé (à tous les quatre dans la direction horizontale et verticale, où chaque seizième sur le plan de l'image). Au fur et à mesure que l'on descend de niveau, de plus en plus de vecteurs de déplacement sont calculés. Le temps de calcul augmente d'un facteur de quatre à chaque niveau (2 fois horizontalement et 2 fois verticalement). Les méthodes de multi-résolutions opèrent sur plusieurs séquences d'images qui ont été filtrées à plusieurs reprises de façon à obtenir un résultat plus représentatif du mouvement apparent d'une séquence d'images (figure 4.7). Le filtrage utilisé est basé sur les filtres proposés par Enkelmann [7]. Les méthodes de multi-résolutions permettent une convergence à une solution plus proche du minimum global. Ceci est dû au filtrage qui rend la fonction objective plus lisse et qui minimise l'effet des minimums locaux. Il est important de noter que la contrainte de lissage est doublée pour chaque niveau de résolution pour permettre d'optimiser la correspondance entre vecteurs de mouvement avoisinants au bas de la pyramide.

Résultats

Plusieurs simulations ont été faites dans le but de déterminer la meilleure méthode de calcul, soit un niveau ou hiérarchique, et de déterminer le modèle structurel le plus approprié pour estimer le mouvement apparent dans une séquence.

Dans l'exemple suivant, un interval temporel de quatre trames a été choisi pour effectuer les simulations sur la séquence *femme – et – arbre*. Donc, la sixième trame est reconstruite à partir de la quatrième et de la huitième. D'après nos intuitions expérimentales, il est prévu que les modèles structurels d'ordre zéro sur un niveau soient moins efficaces que le même modèle sur une pyramide. En général, les approches pyramidales sont plus efficaces pour les séquences avec mouvement rapide

parce qu'elles permettent de minimiser les effets de changement rapide d'intensité (à cause du filtrage).

Pour la séquence *femme – et – arbre*, lorsque le bras se déplace vers l'arrière, au fur et à mesure que les trames évoluent, le mouvement accélère. Vers la fin de la séquence, lorsque le mouvement du bras est à son maximum, on peut s'apercevoir que les contours de la manche sont de moins en moins bien définis. Le modèle structurel de premier ordre avec régularisation révèle cette lacune à un degré beaucoup moins prononcé (figure 4.8). Par les figures 4.8 et 4.9, il est facile de conclure que les approches multi-résolutions donnent des estimés du mouvement apparent plus représentatifs d'une observation visuelle des séquences d'images. Les méthodes hiérarchiques donnent des champs de déplacement plus uniformes avec moins de divergence et croisement de vecteurs. Donc, les champs de déplacement sont plus uniformes.

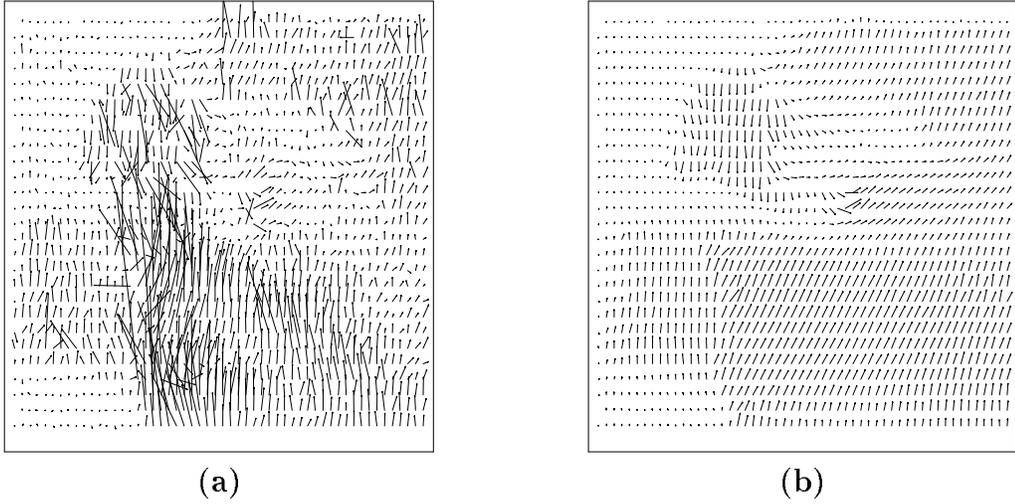


Figure 4.2: Champs de déplacement estimés avec la nouvelle formulation sur la séquence 2 avec λ égal à:(a) 0.1; (b) 1.

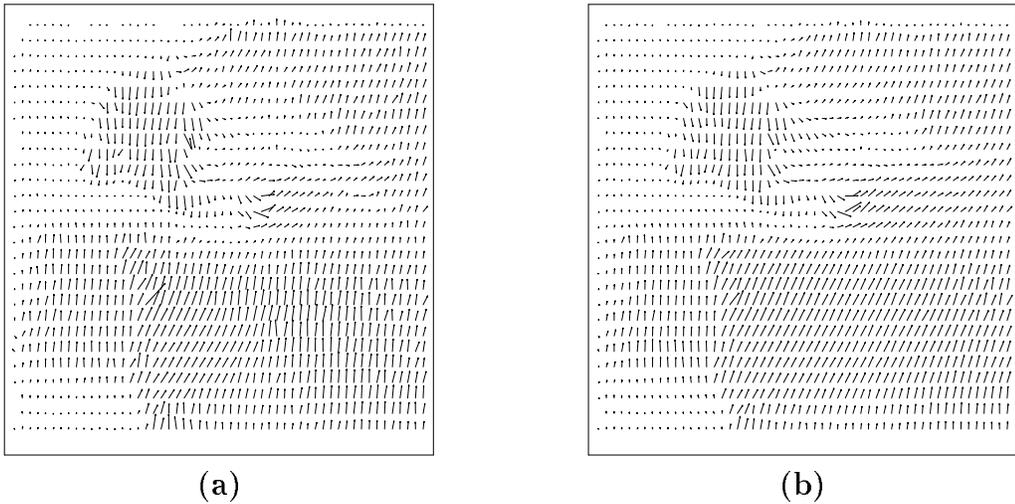


Figure 4.3: Champs de déplacement calculés avec la nouvelle formulation sur la séquence 2 avec λ égal à 1 utilisant l'interpolateur:(a) quadratique; (b) cubique

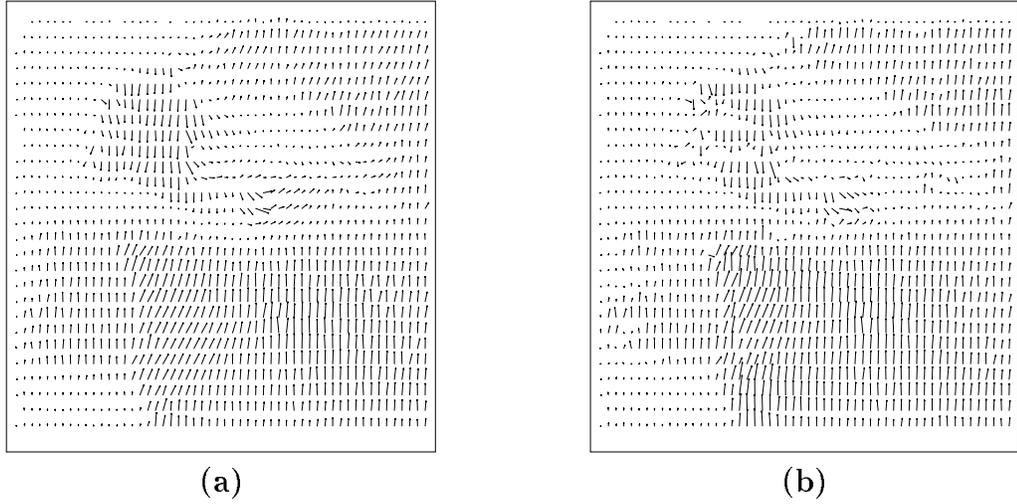


Figure 4.4: Champs de déplacement calculés avec la nouvelle formulation sur la séquence 2 avec λ égal à 1 utilisant l'interpolateur:(a) cubique de Keys; (b) quartique

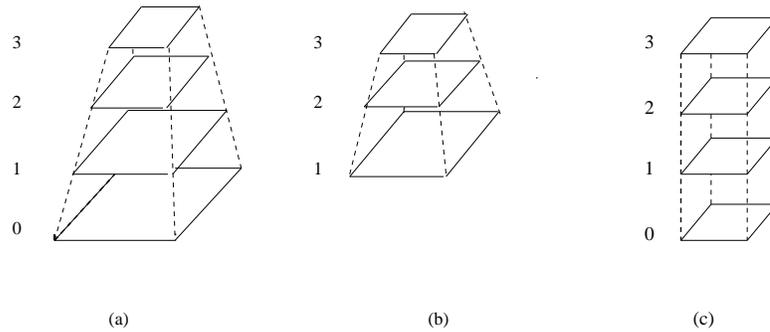


Figure 4.5: Pyramide de multi-résolution pour:(a) la formulation couramment utilisée; (b) la nouvelle formulation; (c) les différents niveaux d'images filtrées.



Figure 4.6: Différents niveaux de résolution pour la séquence 2:(a) haut de la pyramide; (b) 2eme niveau; (c) bas de la pyramide

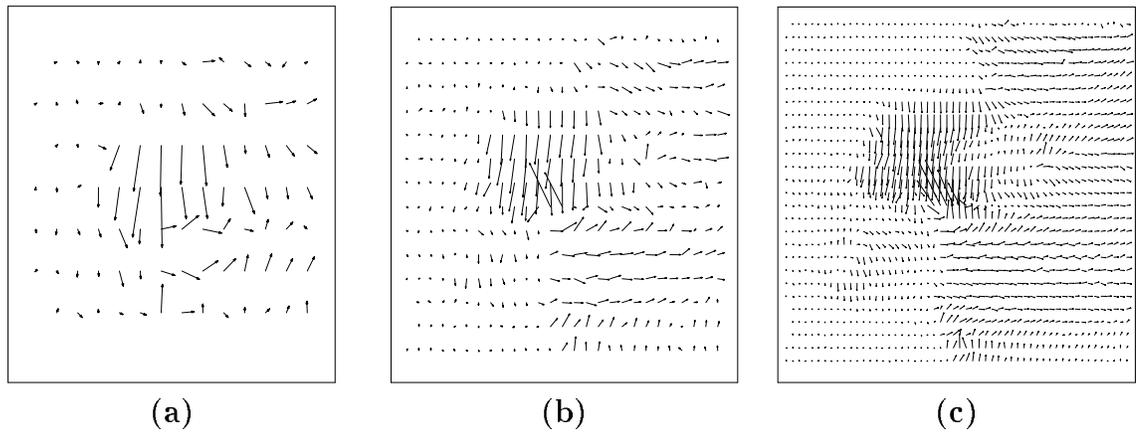


Figure 4.7: Champs de déplacement calculés à différents niveaux de résolution sur la séquence 2 avec la nouvelle formulation avec λ égal à 1:(a) niveau 3; (b) niveau 2; (c) niveau 1.

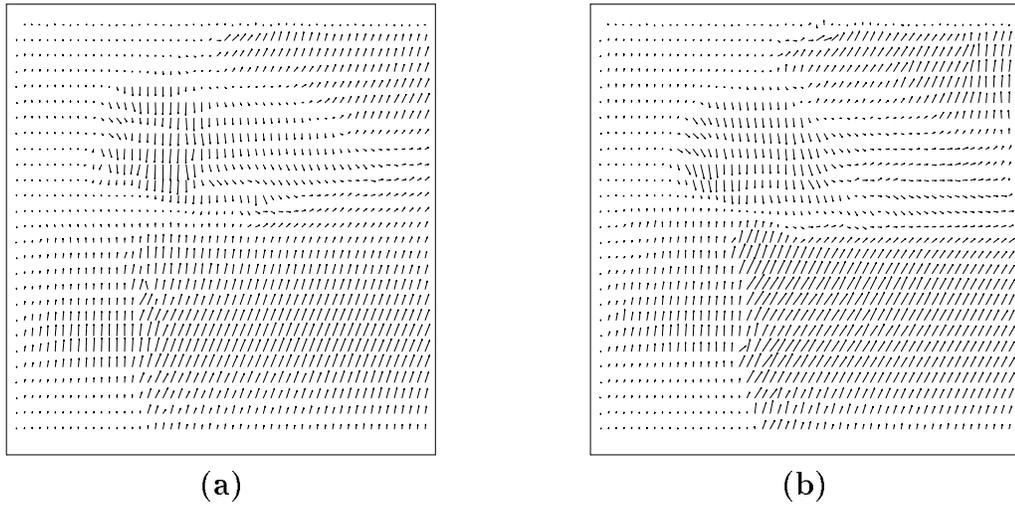


Figure 4.8: Champs de déplacement calculés sur différents niveaux utilisant la nouvelle formulation sur la séquence 2 avec λ égal à 5:(a) niveau 1;(b) pyramide

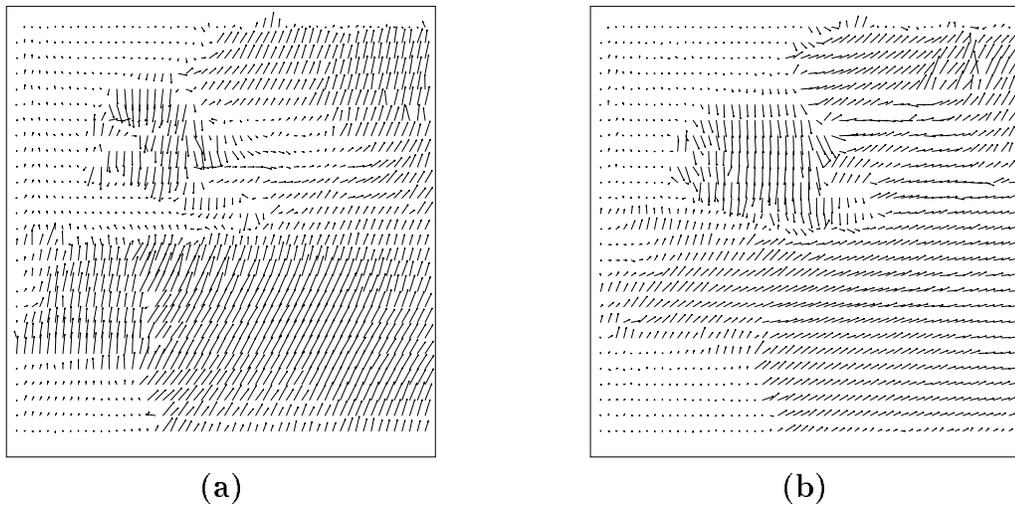


Figure 4.9: Champs de déplacement calculés sur différents niveaux utilisant la formulation couramment utilisée sur la séquence 2 avec λ égal à 20:(a) niveau 0;(b) pyramide

Chapitre 5

Estimation à partir des séquences avec effet d'illumination

Nos algorithmes d'estimation du mouvement apparent établis, il est possible de discuter de leurs performances avec divers modèles d'illumination. Les modèles proposés pour simuler certains effets d'illumination peuvent être divisés en deux classes. La première est constituée de modèles lumineux contenant des variations spatiales d'intensité; la deuxième, de modèles lumineux contenant des variations temporelles du niveau d'intensité d'une trame à une autre. Dans ce travail, trois modèles sont présentés. Le premier consiste d'une intensité spatiale constante combinée avec une variation temporelle linéaire. Le deuxième, consiste d'une distribution gaussienne d'intensité spatiale avec une variation temporelle linéaire. Finalement, une modélisation extrême, avec des variations spatiales aléatoires combinées avec une variation temporelle aléatoires du niveau d'intensité.

5.1 Modèles d'effet d'illumination

Plusieurs travaux ont été fait avec des séquences contenant des effets d'illumination. Des chercheurs comme Moloney [17] ont dérivé des équations qui modélisait certains effets d'illumination. Dans notre cas, la nouvelle formulation introduite permet

d'estimer le mouvement apparent sous n'importe quel type d'effet d'illumination. Dans le but d'établir la performance des algorithmes d'estimation présentés antérieurement, trois modèles d'effet d'illumination ont été proposés.

5.1.1 Premier effet d'illumination

Le premier modèle consiste à établir une variation linéaire temporelle du niveau d'illumination tout en gardant l'intensité spatiale du niveau d'illumination constante (figure 5.1 (a) et (b)). Plus précisément, ceci correspond à additionner un certain niveau d'intensité partout sur une trame. Pour les premières quinze trames, l'intensité est augmentée de façon linéaire d'une trame à une autre. Pour les quinze dernières trames, le niveau d'intensité diminue à nouveau au même taux jusqu'à la dernière trame (figure 5.1 (c)). Ce modèle a été implémenté avec une amplitude spatiale égale à 5 et une pente de 4.

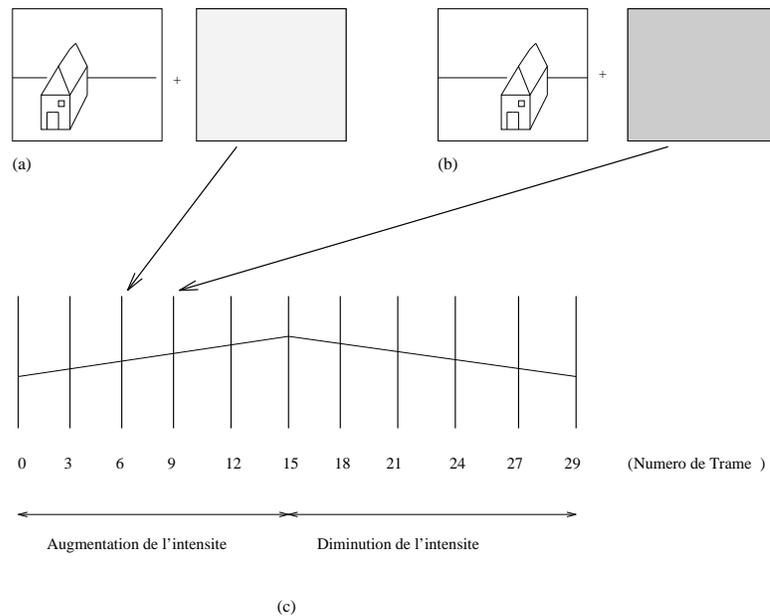


Figure 5.1: Illustration du premier effet d'illumination sur une séquence d'images.

5.1.2 Deuxième effet d'illumination

Le deuxième modèle consiste à établir une variation gaussienne du niveau d'intensité spatiale tout en variant linéairement l'amplitude du niveau d'illumination d'une trame à une autre. Pour une séquence de 30 trames, les quinze premières consistent à additionner sur le plan de l'image, un certain niveau d'intensité distribué selon une fonction gaussienne. L'amplitude maximale de la fonction reste constante sur une même trame mais augmente linéairement pour les quinze premières trames. Pour les quinze dernières trames, l'amplitude maximale de la fonction gaussienne diminue à nouveau à sa valeur originale de façon linéaire (figure 5.2). La forme de la fonction gaussienne est la suivante:

$$f(x, y, t) = A(t) \cdot \exp\left(\frac{(x - m_x)^2 + (y - m_y)^2}{2 \cdot \sigma^2}\right)$$

$A(t)$ représente le changement temporel de l'amplitude maximale de la fonction gaussienne. m_x et m_y représentent les coordonnées auxquelles est centrée la fonction. Finalement, σ^2 représente la variance ou le taux de décroissance spatial de l'effet d'illumination. Ce modèle a été implémenté avec m_x égal à 245, m_y égal à 4, σ^2 égal à 80 et A_0 égal à 30. De plus, $A(t) = A_0 + 7 \cdot \Delta t$.

5.1.3 Troisième effet d'illumination

Le troisième modèle d'illumination consiste à établir une partition aléatoire du plan de l'image et ensuite à appliquer un niveau d'intensité lumineuse choisi selon une distribution gaussienne à chaque région polygonale. A chaque région polygonale définie par la partition, le niveau d'illumination est augmenté de façon linéaire d'une trame à une autre jusqu'à ce qu'un maximum soit atteint. Après quoi le niveau d'intensité d'illumination est réduit de façon linéaire jusqu'à un certain seuil. Une fois le seuil minimum atteint, l'intensité est de nouveau augmentée de façon linéaire (figure 5.3).

Ce modèle est basé sur les travaux de Fries et Modestino présenté dans [8] et [16]. Ils avaient proposé de modéliser l'illumination dans une séquence d'images

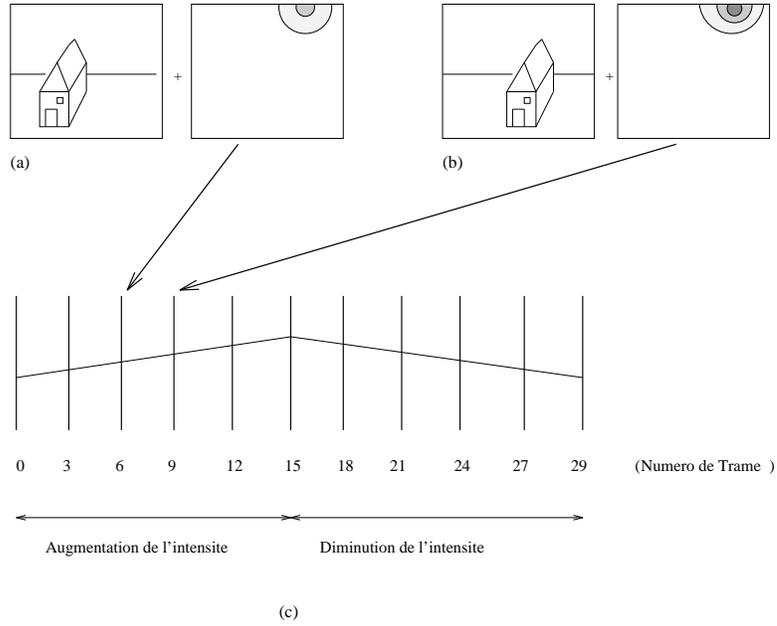


Figure 5.2: Illustration du deuxième effet d'illumination sur une séquence d'images (fonction gaussienne).

en divisant le plan de l'image en régions polygonales et ensuite en assignant une intensité lumineuse à chaque région selon une distribution gaussienne. De plus, le niveau d'intensité de chaque région était corrélé avec le niveau d'intensité des régions avoisinantes par une certaine corrélation ρ . Une certaine quantité de bruit blanc W_i était présente.

Pour convertir le type de distribution d'uniforme à gaussienne, une transformation polaire a été effectuée. Cette approche a été suggérée dans [12] et [21]. La formulation mathématique de cette transformation est la suivante:

$$X_1 = V_1 \cdot \sqrt{\left(\frac{-2 \cdot \ln S}{S}\right)}$$

$$X_2 = V_2 \cdot \sqrt{\left(\frac{-2 \cdot \ln S}{S}\right)}$$

Les variables aléatoires X_1 et X_2 représentent les niveaux d'intensité distribués de façon gaussienne, par contre les variables V_1 et V_2 représentent les niveaux d'intensité distribués de façon uniforme. La variable S vérifie la région de validité de la sélection

aléatoire. En plus de la conversion de standard de distribution, chaque variable X_i est corrélée avec les autres variables de la façon suivante:

$$X_i = \rho X_{i-1} + W_i$$

où, ρ représente le degré de corrélation entre niveaux d'intensité lumineux avoisinants. Comme suggéré dans l'article de Fries et Modestino [16] la corrélation a été fixée à 0. Donc, chaque niveau d'intensité de chacune des régions polygonales est indépendant des régions avoisinantes.

La différence entre le vrai modèle de Fries et Modestino et celui utilisé ici est la suivante. Leur modèle permettait une assignation des niveaux d'intensité lumineuse de façon vraiment aléatoire (comme un kaléidoscope) et non linéaire comme présentée ici. A cause de l'approximation quadratique de la fonction objective à un terme de l'expansion de Taylor, il en résulte que la variation temporelle d'intensité (pour une même région polygonale) doit être petite (petite pente). Ce modèle a été implémenté avec λ égal à 5, ρ égal à 0 et σ égal 2 pour la séquence pingpong et avec λ égal à 5, ρ égal à 0.5 et σ égal 2 pour les deux autres séquences. Dans ce cas, λ représente le paramètre gouvernant la fréquence de partition sur le plan de l'image. Cette partition est gouvernée par une distribution exponentielle. Le paramètre ρ établit la corrélation entre niveaux d'intensité avoisinant. Le paramètre σ est la variance des variables gaussiennes.

5.2 Résultats expérimentaux

5.2.1 Séquence avec effet d'illumination

Une pyramide des séquences avec illumination a été faite pour obtenir les trois différents niveaux de résolution d'intensité lumineuse (niveaux 3-2-1). Toutes les simulations ont été faites sur une fenêtre temporelle de trente trames dans le cas des trois séquences de test. Les algorithmes d'estimation consistaient des modèles structuraux d'ordre zéro et d'ordre un avec régularisation. Ils ont été utilisés pour estimer

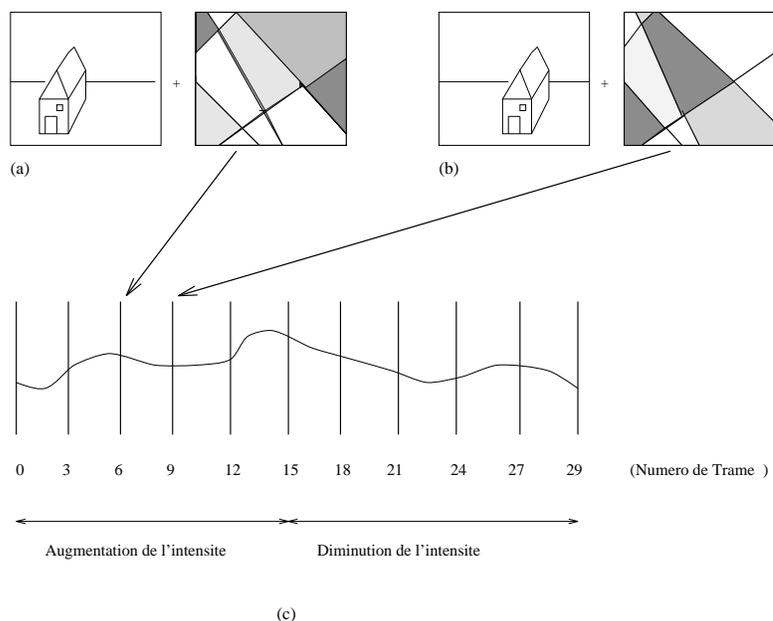


Figure 5.3: Illustration du troisième effet d'illumination sur une séquence d'images (partition aléatoire).

le mouvement apparent dans des séquences avec ces trois effets d'illumination.

Avec ce décalage temporel (30 trames), il a été possible de déterminer la performance des algorithmes d'estimation sous différentes conditions d'intensité et de mouvement. La première séquence de test fut utilisée parce que le type de mouvement consistait en une simple translation sur l'intervalle temporel. Pour la séquence 2 contenant 30 trames, la gamme de rapidité du mouvement des bras et des mains est large. Au début de la séquence, le mouvement est très lent. Au fur et à mesure que la séquence progresse, le mouvement des bras et des mains devient de plus en plus rapide. Dans certaines régions, le niveau d'information structurel avait beaucoup de variabilité, dans d'autres moins. Pour cette séquence, la région du bras ne contient pas beaucoup de variation d'intensité d'un pixel à l'autre. En conséquence, les champs de déplacement sont moins bien définis dans ces régions. Pourtant, la même séquence dans la région de la main droite, le niveau de mouvement ainsi que la variation du niveau d'intensité lumineux laisse beaucoup de manoeuvrabilité. Fi-

nalement, la troisième séquence, est la plus compliquée. Cette séquence contient des occlusions, des mouvements rapides ainsi que des combinaisons de différents types de mouvement comme des translations avec un mouvement de fond et ainsi de suite.

5.2.2 Comparaison objective

Pour comparer subjectivement la performance de chaque algorithme (constance de DPD et constance de DGD) aux trois modèles d'illumination présentés ici, la reconstruction de trame au bas de la pyramide (niveaux 0) a été choisie comme mode d'évaluation. À partir de seize trames originales (paires) transmises et puis reçues, quinze autres trames intermédiaires (impaires) ont été reconstruites. Ce mode d'analyse simulait un facteur de compression 2:1. Le paramètre Δ_t a été fixé à une demie. Avec ces séquences reconstruites qui incluaient les trois modèles d'illumination et combinaient les deux modèles d'estimation (constance de DPD et constance de DGD), une comparaison détaillée des deux algorithmes d'estimation a été faite.

Dans les figures qui suivent cette analyse, les trois modèles d'illumination présentés antérieurement sont incorporés avec les séquences un, deux et trois. Pour un modèle donné, le champ de déplacement ((a), (d),(g)), l'image reconstruite ((b), (e), (h)) et l'image d'erreur ((c), (f), (i)) sont présentés. Tous les résultats présentés ici ont été estimés avec λ égal à 20 pour l'algorithme de constance de DPD et à 5 pour l'algorithme de constance de DGD.

Pour l'image d'erreur, les régions grises ne correspondent à aucune erreur entre l'image reconstruite et l'image originale. Les régions plus claires (blanches) correspondent à des erreurs positives. Les régions sombres (noires) correspondent à des erreurs négatives. Il n'est pas difficile de constater que l'algorithme d'estimation de constance de DGD donne des estimés de mouvement apparent beaucoup plus précis que celui de constance de DPD. Les images reconstruites avec cet algorithme sont mieux définies que celles reconstruites avec l'algorithme de constance de DPD.

Les résultats dans les pages suivantes doivent être inspectés en groupe de la façon suivante. Pour chaque séquence introduite, les champs de déplacement sont calculés

avec l'algorithme de DPD et de DGD dans les figures 5.10 et 5.14. Ces résultats sont introduits pour déterminer l'impact des modèles d'illumination sur ces résultats. Les figures 5.11(a,b,c), 5.12(a,b,c) et 5.13(a,b,c) représentent les résultats obtenus avec l'algorithme de constance de DPD. Les figures 5.15(a,b,c), 5.16(a,b,c) et 5.17(a,b,c) représentent les résultats obtenus avec l'algorithme de constance de DGD. Les champs de déplacement ainsi que les images reconstruites et les images d'erreurs doivent être comparés entre elles. Il n'est pas difficile de constater que les champs de déplacement obtenus avec la nouvelle formulation sont beaucoup plus représentatifs du mouvement apparent dans des séquences avec d'effet d'illumination.

Pour le deuxième effet d'illumination suggéré, les figures 5.11(d,e,f), 5.12(d,e,f) et 5.13(d,e,f) représentent les résultats obtenus avec les algorithmes d'estimation de constances de DPD avec régularisation. Avec cet algorithme, les distortions des champs de déplacement sont plus évidentes. Les images d'erreurs ont plus de distortions dans le coin du haut à droite, indiquant une mauvaise qualité de reconstruction. Les figures 5.15(d,e,f), 5.16(d,e,f) et 5.17(d,e,f) représentent les résultats obtenus avec les algorithmes d'estimation de constance de DGD avec régularisation. Les distortions sont très minimes. Encore une fois la méthode de constance de DGD donne une représentation plus précise.

Pour le modèle d'illumination proposé par Fries et Modestino, les figures 5.11(g,h,i), 5.12(g,h,i) et 5.13(g,h,i) représentent l'algorithme d'estimation combinant l'hypothèse de constance de DPD avec régularisation. Il n'est pas difficile de constater que cet algorithme n'a pas de bon estimé du mouvement apparent sous ces conditions. Les figures 5.15(g,h,i), 5.16(g,h,i) et 5.17(g,h,i) ont été obtenues avec l'algorithme de constance de DGD avec régularisation. Ce modèle, avec des changements extrêmes de variation d'intensité lumineuse ne caractérise aucune situation réelle mais permet d'établir la représentation des algorithmes d'estimation de mouvement apparent dans des conditions très aléatoires d'illumination. Encore une fois, la nouvelle formulation donne des champs de déplacement beaucoup plus robustes aux changements d'intensité que celui de constance de DPD (même dans des régions polygonales différentes, et

modèle	séquence 1			séquence 2			séquence 3		
	DPD	DGD		DPD	DGD		DPD	DGD	
λ	20	5	20	20	3	5	20	5	10
aucun	29.04	28.72	28.87	38.23	38.31	38.59	27.15	27.11	27.57
1	28.59	28.72	28.87	38.08	38.31	38.59	26.95	27.47	27.93
2	27.68	28.73	28.87	36.83	38.29	38.55	26.46	27.70	28.06
3	22.80	25.07	25.32	25.17	26.48	26.51	25.86	26.56	26.76

Table 5.1: Changement du PSNR sous les effets d’illumination pour les modèles structurels d’ordre zéro et d’ordre un sur les séquences 1, 2 et 3 avec différentes valeurs de λ .

même aux frontières de ces régions polygonales).

5.2.3 L’impact sur le PSNR

Dans le tableau 5.1, il est possible de voir l’effet de l’introduction d’un champ d’illumination sur l’estimation du mouvement apparent et sur le critère de performance le PSNR défini comme suit:

$$PSNR = 10 \cdot \log_{10} \frac{255^2}{var(x)}$$

où, $var(x)$ représente la variance cummulative d’erreur de reconstruction sur une séquence d’images. Il n’est pas difficile de constater que le PSNR a tendance à diminuer sous l’effet d’illumination. En introduisant des effets d’illumination dans les séquences, les estimés du mouvement apparent se trouvent à être moins représentatifs du mouvement apparent. Pour effectuer une comparaison plus objective des deux méthodes de calculs, la contrainte de lissage λ est ajustée de façon à obtenir une courbe similaire pour les deux algorithmes sans effet d’illumination. Pour chaque séquence la contrainte est réglée de façon différente.

Le tableau 5.1 démontre que l’algorithme de constance de DGD est beaucoup

plus robuste aux changements d'illumination (irrespectivement de λ). Conséquemment, il conserve le même PSNR pour les modèles d'illumination linéaire 1 et 2. Pour le modèle aléatoire numéro 3, le PSNR a aussi tendance à diminuer mais moins que pour l'algorithme de constance de DPD. Le changement du PSNR en fonction des trames dans une séquence d'images est présenté pour les trois séquences dans les figures 5.5, 5.7 et 5.9 pour l'algorithme de constance de DGD, et les figures 5.4, 5.6 et 5.8 pour l'algorithme de constance de DPD. D'après ces résultats, pour les séquences *courtepointe* et *femme – et – arbre* l'algorithme de constance de DGD est beaucoup plus robuste au changement linéaire d'illumination que ne l'est l'algorithme de constance de DPD. Pour la séquence *pingpong* contenant des phénomènes d'occlusions ainsi que des mouvements rapides pour les 10 premières trames, le comportement du PSNR n'est pas aussi bien défini. Pour cette séquence le PSNR pour les deux modèles d'illumination linéaire a tendance à augmenter légèrement d'environ 1 à 2 dB. Pour cette même séquence le PSNR pour l'algorithme de constance de DPD a tendance à diminuer.

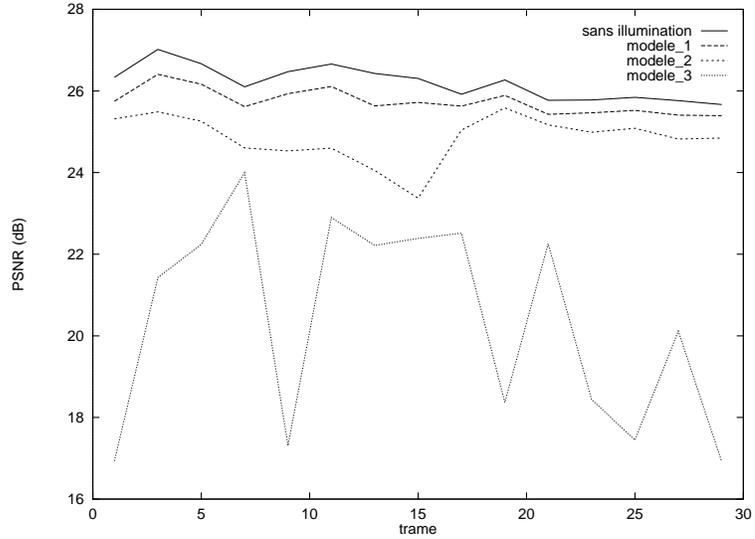


Figure 5.4: Evolution du PSNR pour les 30 premières trames de la séquence *courtepointe* utilisant la méthode de constance de DPD avec λ égal à 20.

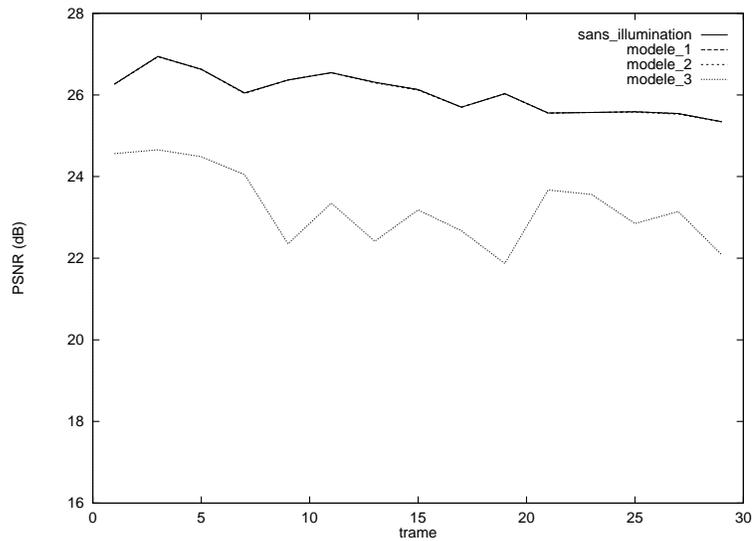


Figure 5.5: Evolution du PSNR pour les 30 premières trames de la séquence *courtepointe* avec une superposition des courbes avec aucune illumination et les modèles 1 et 2, utilisant la méthode de constance de DGD avec λ égal à 20.

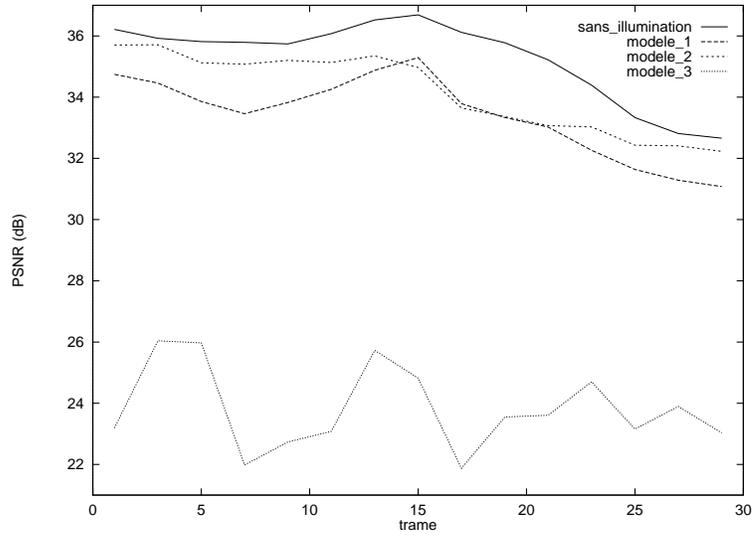


Figure 5.6: Evolution du PSNR pour les 30 premières trames de la séquence *femme – et – arbre* utilisant la méthode de constance de DPD avec λ égal à 20.

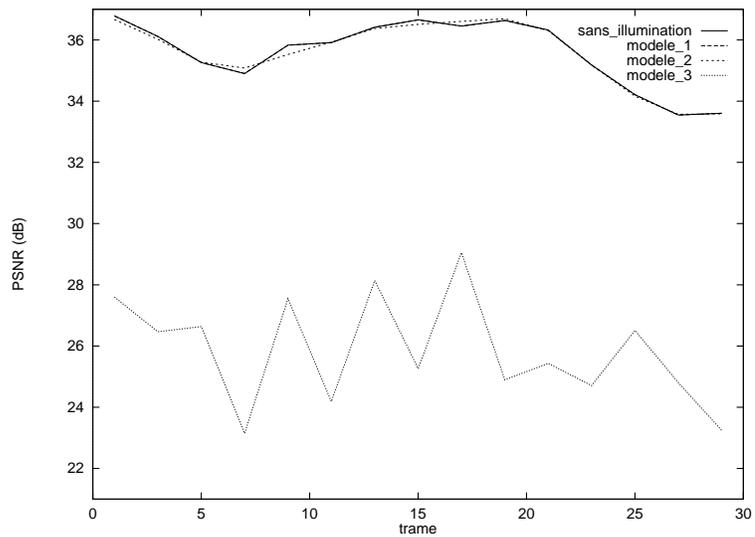


Figure 5.7: Evolution du PSNR pour les 30 premières trames de la séquence *femme – et – arbre* avec une superposition des courbes avec aucune illumination et les modèles 1 et 2, utilisant la méthode de constance de DGD avec λ égal à 3.

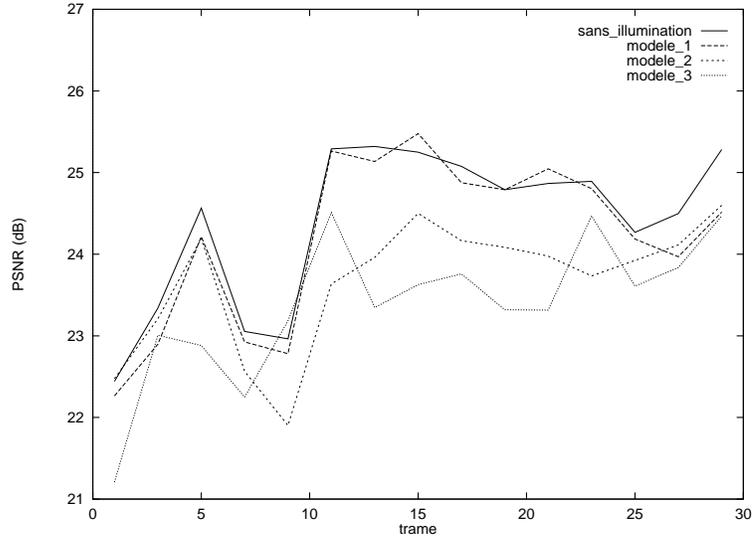


Figure 5.8: Evolution du PSNR pour 30 trames à partir de la quatorzième de la séquence *pingpong* utilisant la méthode de constance de DPD avec λ égal à 20.

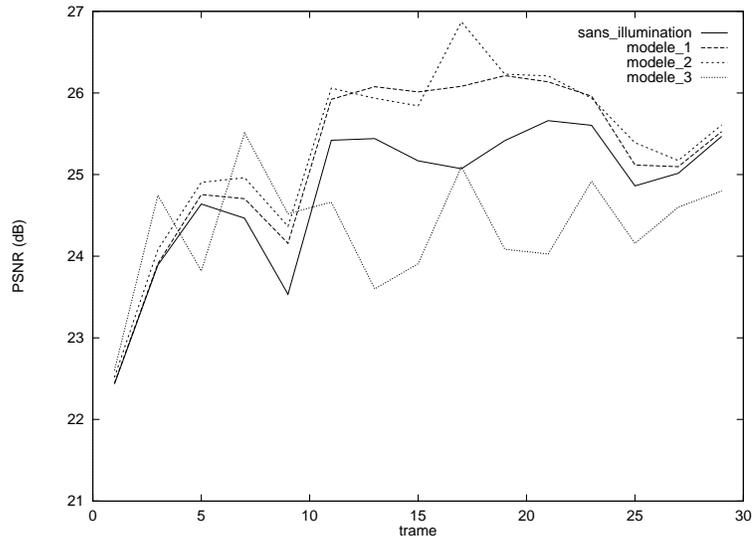
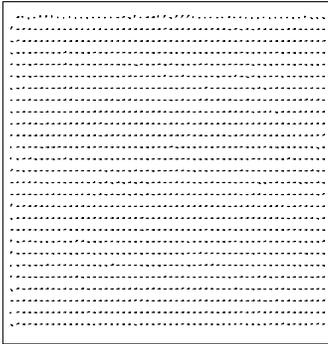
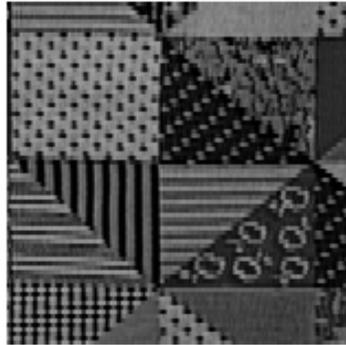


Figure 5.9: Evolution du PSNR pour 30 trames à partir de la quatorzième de la séquence *pingpong* utilisant la méthode de constance de DGD avec λ égal à 10.

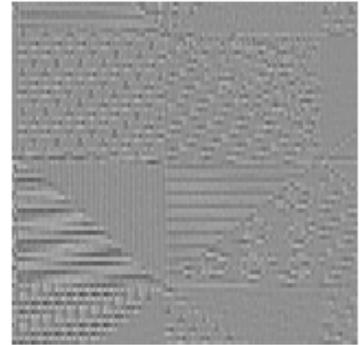
/ window at (default)



(a)

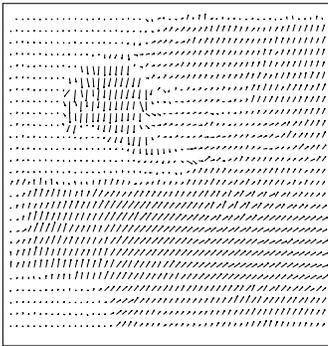


(b)



(c)

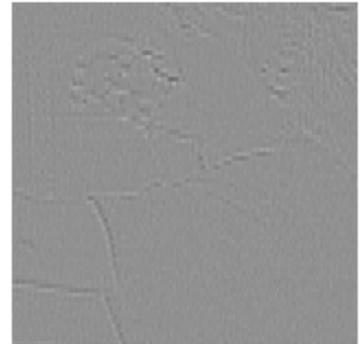
/ window at (default)



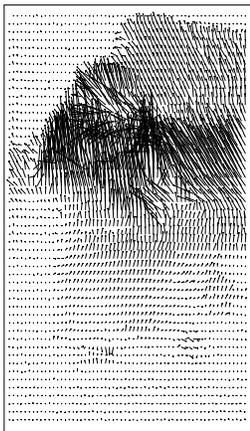
(d)



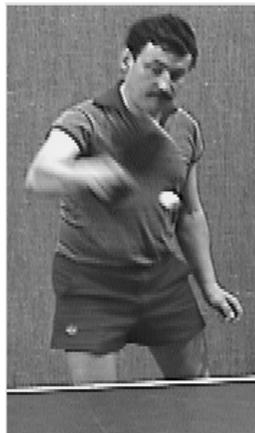
(e)



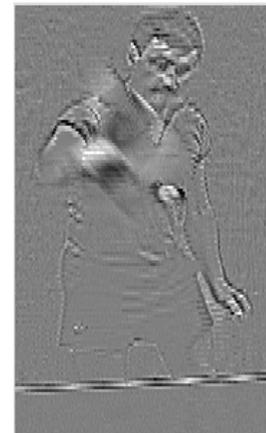
(f)



(g)



(h)



(i)

Figure 5.10: Estimation par constance de DPD sur les séquences *courtepointe*, *femme - et - arbre* et *pingpong* avec λ égal à 20.

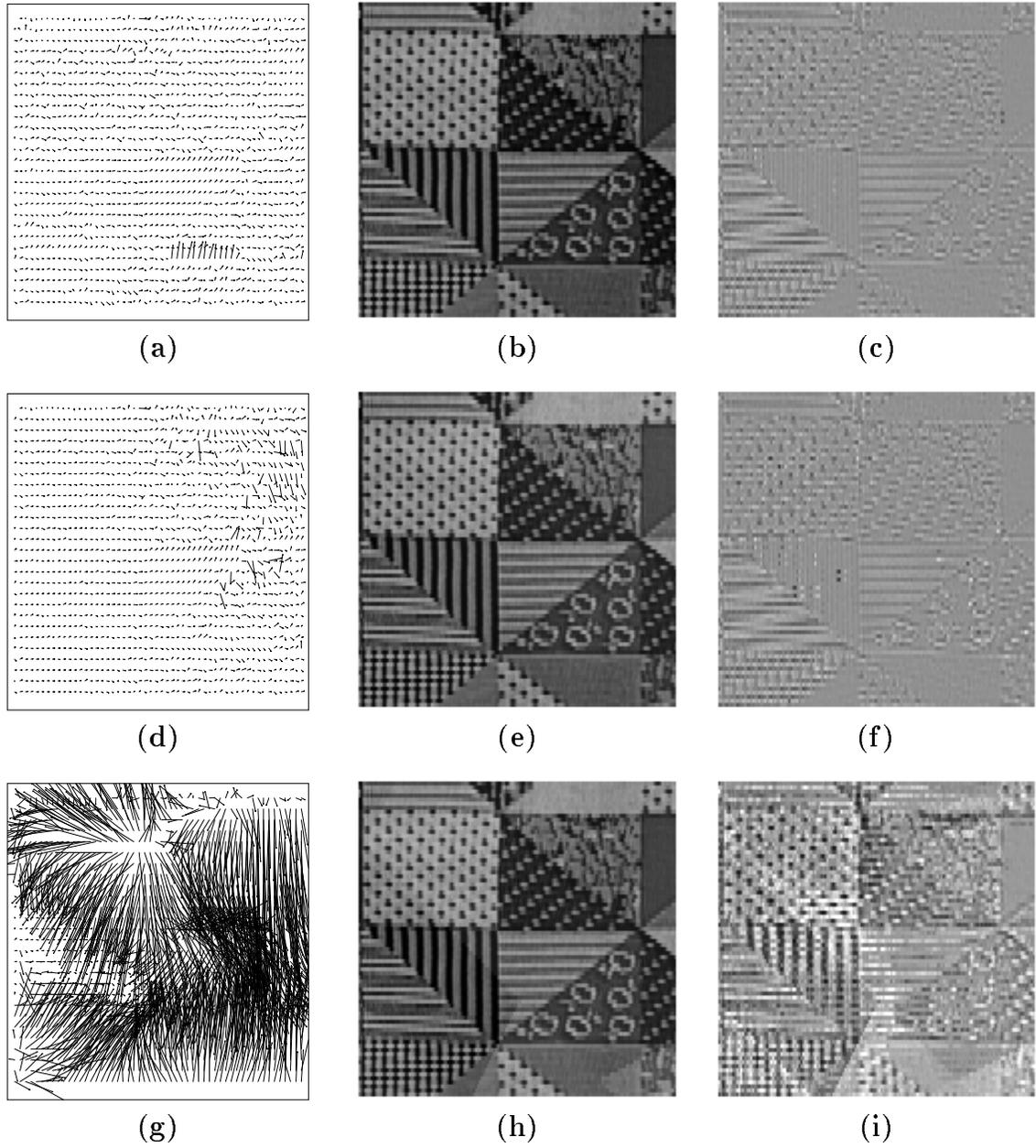


Figure 5.11: Estimation par constance de DPD avec trois effets d'illumination sur la séquence *courtepointe* avec λ égal à 20.

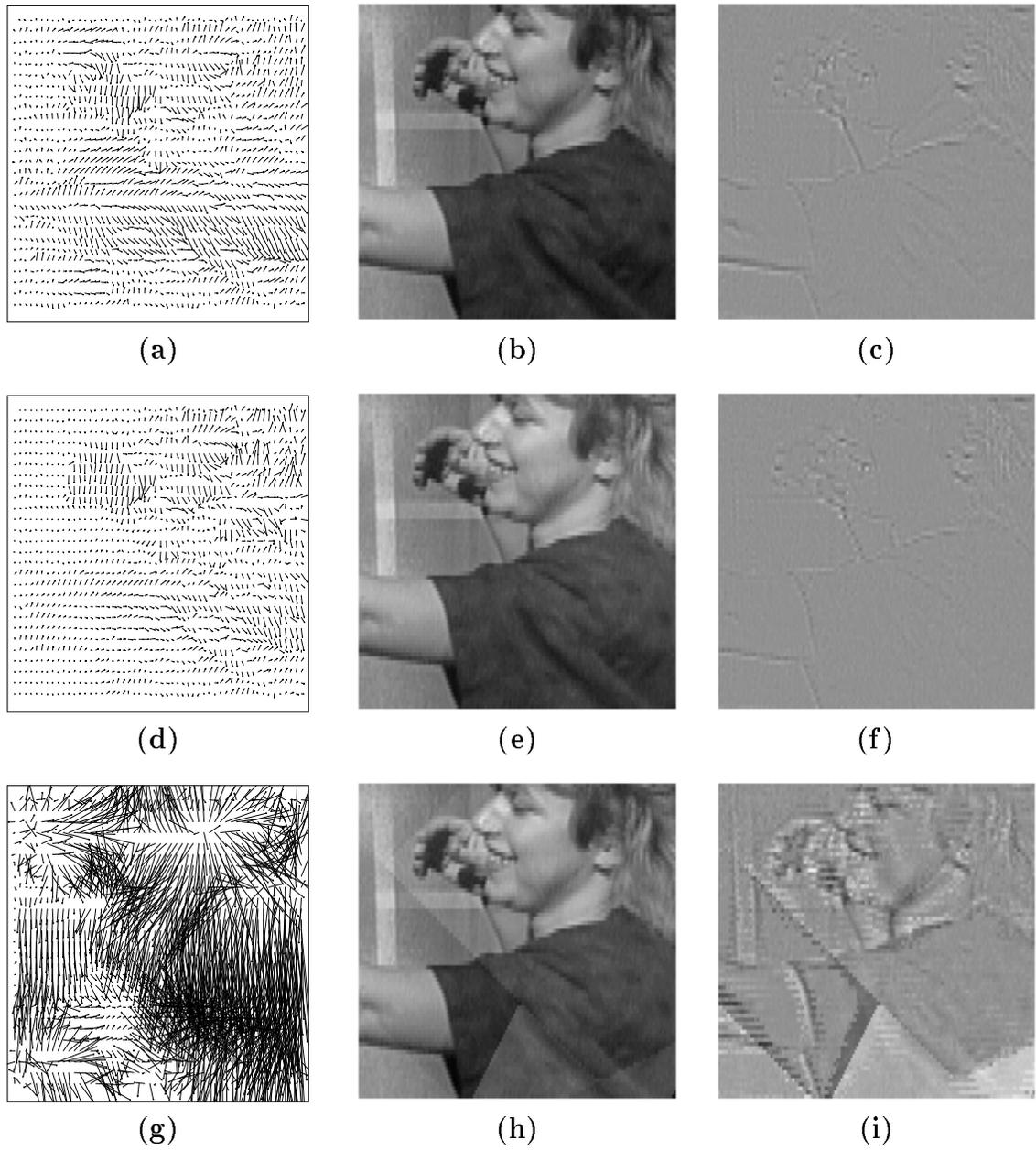


Figure 5.12: Estimation par constance de DPD avec trois effets d'illumination sur la séquence *femme – et – arbre* avec λ égal à 20.

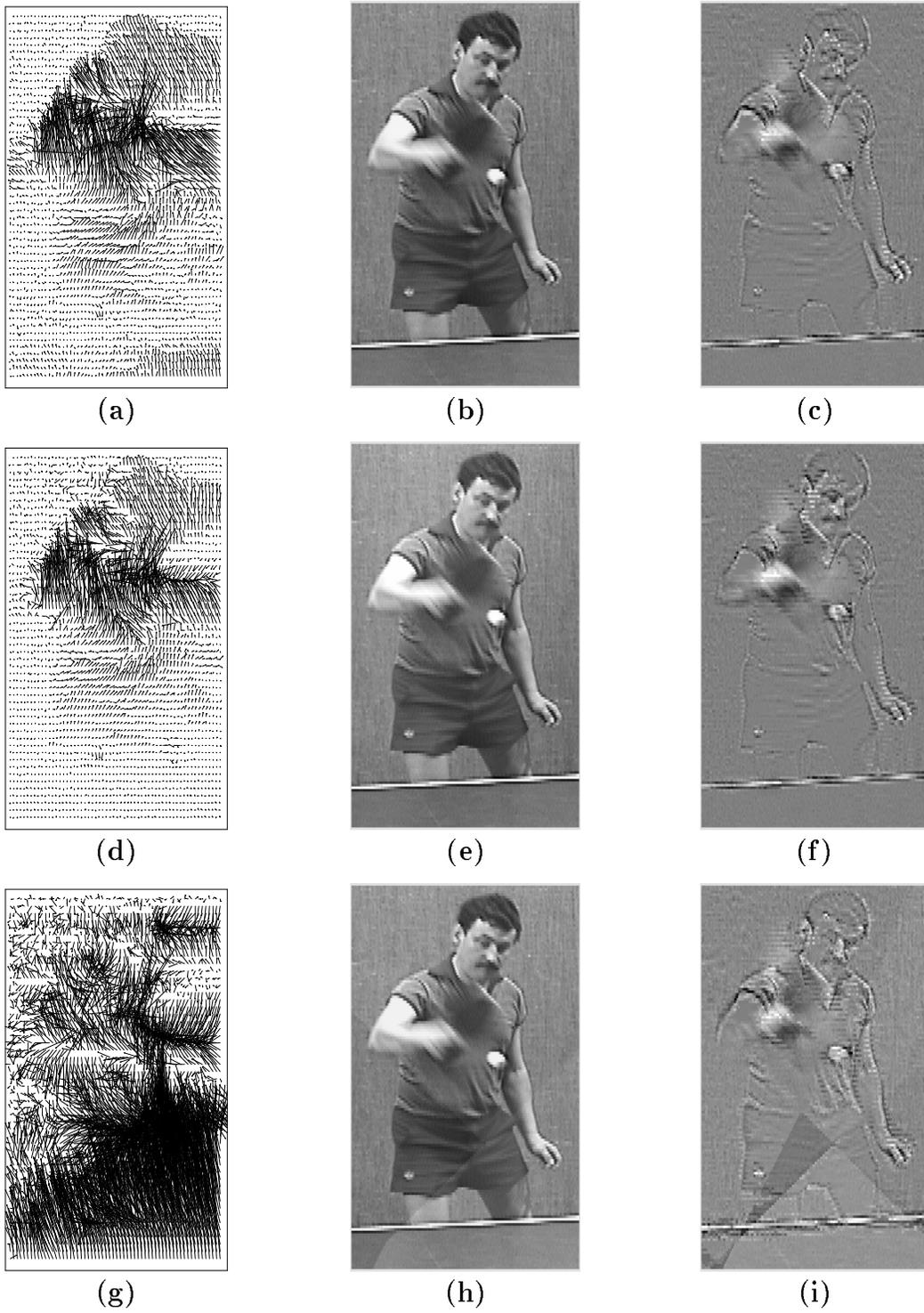
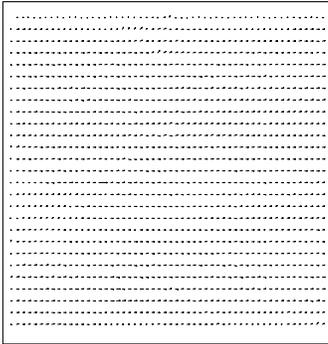
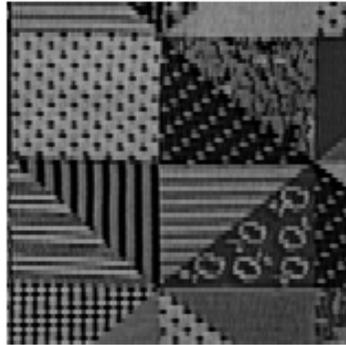


Figure 5.13: Estimation par constance de DPD avec trois effets d'illumination sur la séquence *pingpong* avec λ égal à 20.

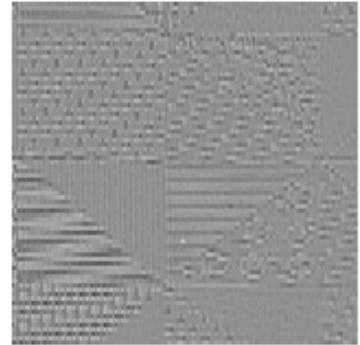
/ window at (default)



(a)

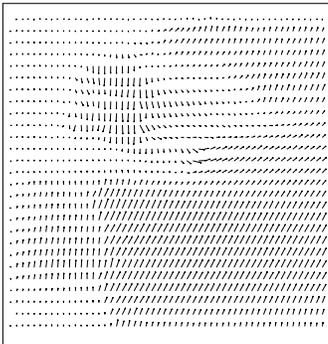


(b)



(c)

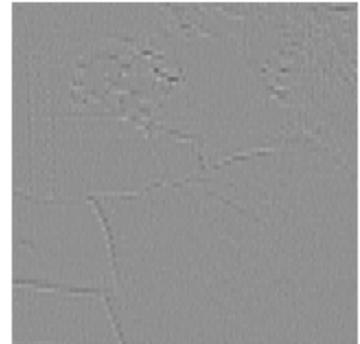
/ window at (default)



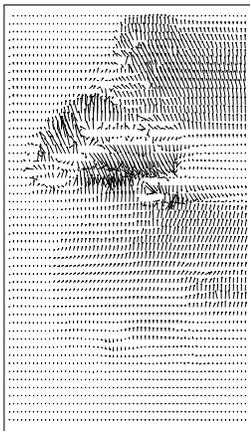
(d)



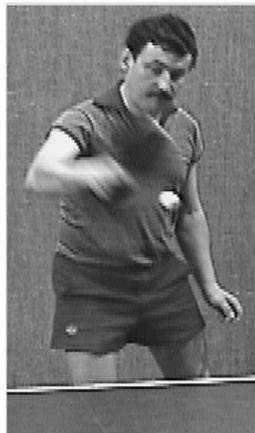
(e)



(f)



(g)



(h)



(i)

Figure 5.14: Estimation par constance de DGD sur les séquences *courtepointe*, *femme – et – arbre* et *pingpong* avec λ égal à 5.

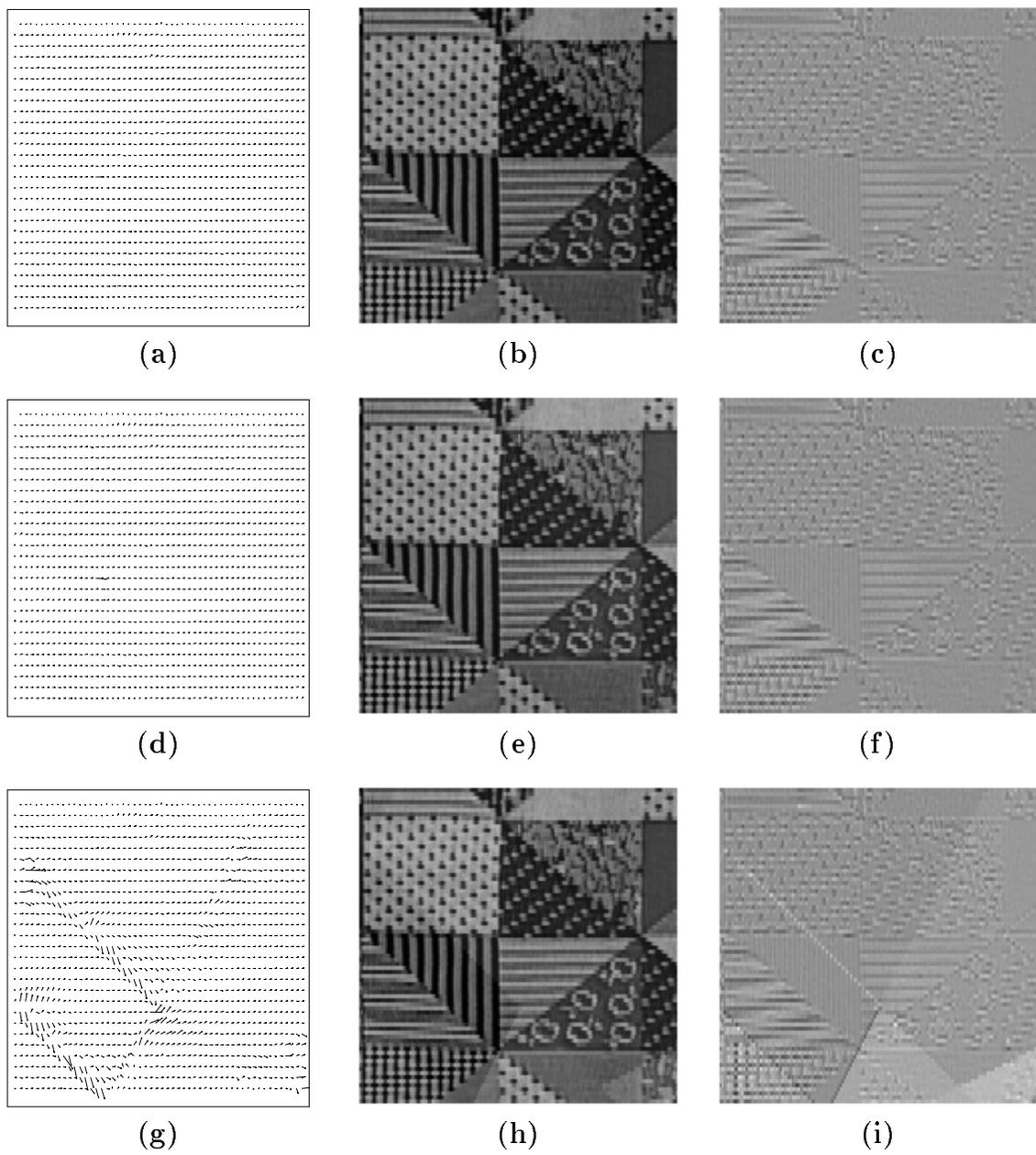


Figure 5.15: Estimation par constance de DGD avec trois effets d'illumination sur la séquence *courtepointe* avec λ égal à 5.

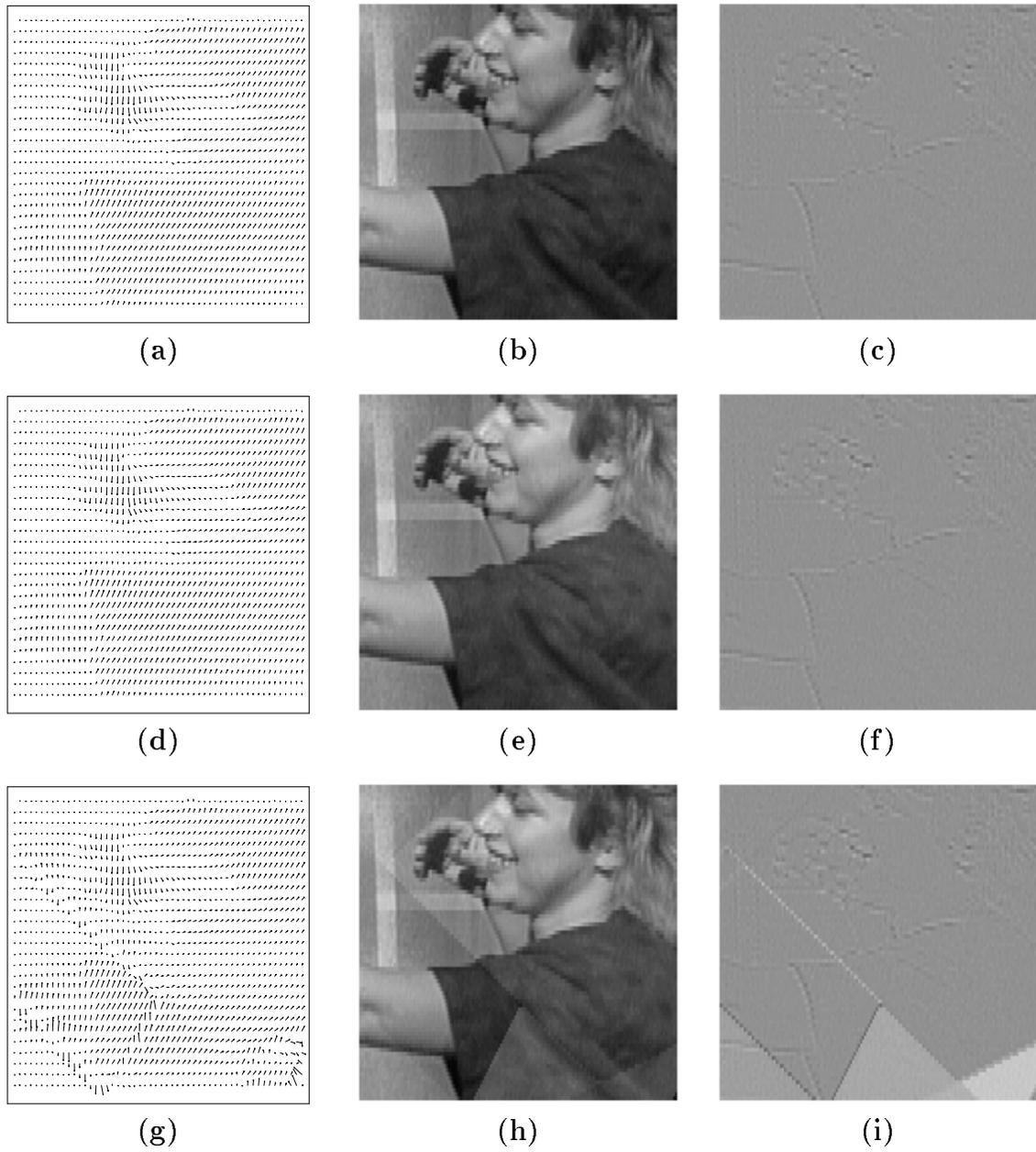


Figure 5.16: Estimation par constance de DGD avec trois effets d'illumination sur la séquence *femme - et - arbre* avec λ égal à 5.

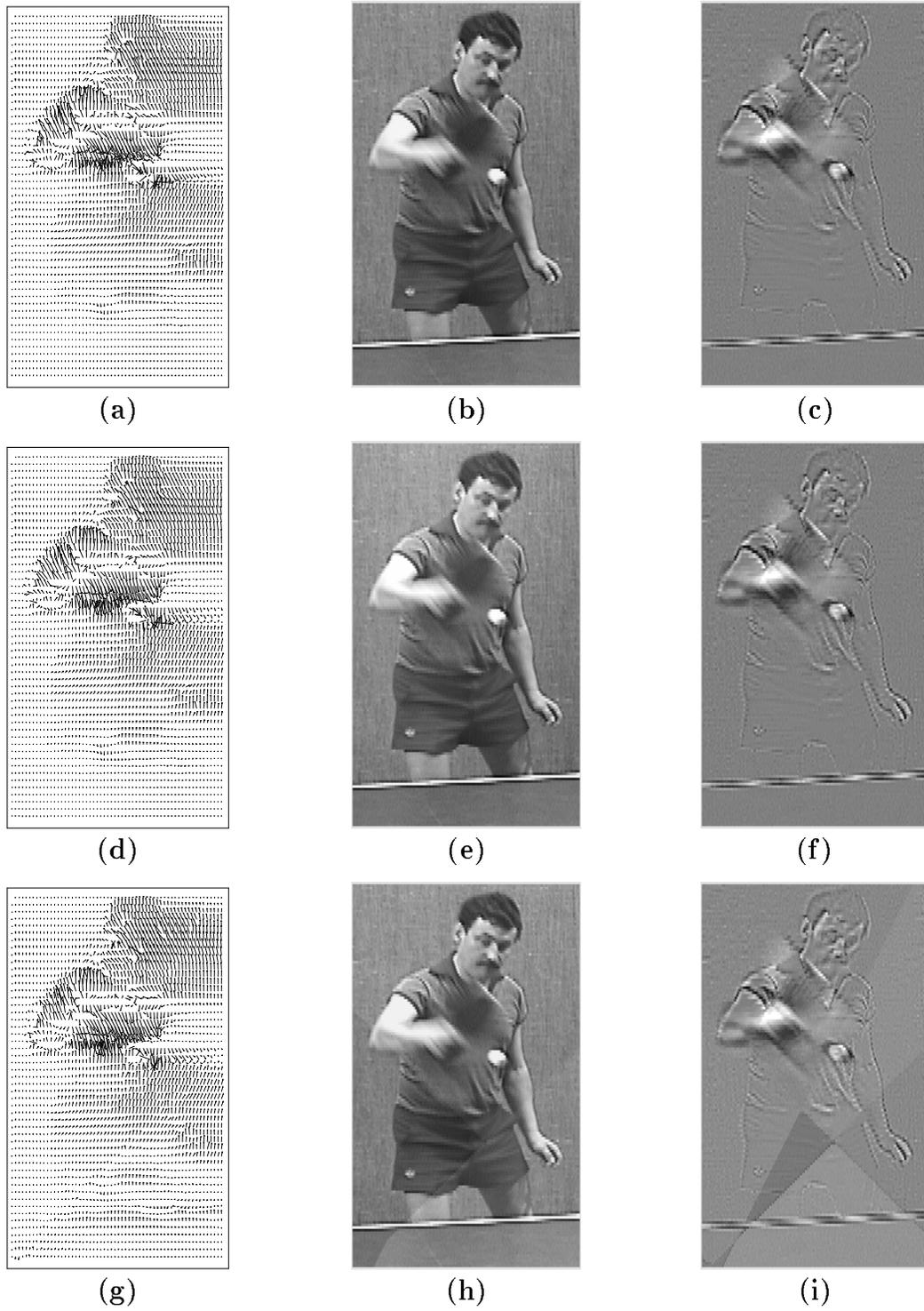


Figure 5.17: Estimation par constance de DGD avec trois effets d'illumination sur la séquence *pingpong* avec λ égal à 5.

Chapitre 6

Conclusion

Plusieurs méthodes pour estimer le mouvement apparent dans les séquences d'images existent. Certaines, telles que l'algorithme de constance de DPD et la méthode proposée par Pastor et Tretyak [20] existent depuis quelques années, mais sont inadéquates pour estimer le mouvement apparent sous des conditions incontrôlables d'illumination de la scène. Ces modèles supposent que tout changement d'intensité est dû au mouvement. Ils ne modélisent pas les effets d'illumination, parce que ces algorithmes ne permettent pas de variations d'intensité d'une trame à l'autre.

Pour surmonter cet obstacle, l'algorithme de constance du gradient a été combiné avec une contrainte de lissage et proposé comme nouvel algorithme d'estimation du mouvement apparent. Au chapitre 4, il a été démontré que cette formulation donne des champs de déplacement caractérisant le mouvement observable dans une séquence d'images sans aucun effet d'illumination.

Trois modèles d'illumination ont été proposés pour vérifier l'invariance de cette nouvelle formulation à des variations linéaires et non-linéaires du niveau d'intensité lumineuse. Afin d'établir une conclusion satisfaisante, les champs de déplacement, images reconstruites ainsi que les images d'erreurs (entre l'image originale et l'image reconstruite) ont été estimés et présentées pour chaque algorithme d'estimation (constance de DPD et constance de DGD) sur trois séquences d'images originales et incorporant ces modèles d'illumination (*courtepointe*, *femme – et – arbre* et *pingpong*).

Il n'est pas difficile de conclure d'après les résultats obtenus au chapitre 5, que la nouvelle formulation est beaucoup plus adéquate pour estimer le mouvement apparent dans des séquences avec des effets d'illumination. Le gradient d'intensité, en étant constant le long de la trajectoire de mouvement, permet une variation constante du niveau d'intensité lumineuse d'une trame à l'autre. Cette propriété est idéale pour modéliser des variations linéaires d'illumination de la scène.

D'après les résultats obtenus avec le PSNR sur les séquences *courtepointe*, *femme – et – arbre* et *pingpong*, il est évident que l'algorithme de constance de DGD a une performance plus robuste aux changements d'illumination que l'algorithme de constance de DPD. En étant invariant au changement linéaire d'illumination de la scène, la nouvelle formulation conserve le même PSNR mais l'algorithme de constance de DPD décroît la valeur correspondante du PSNR. Il est important de noter que même dans le cas de variations non-linéaires, l'algorithme de constance de DGD donne une performance supérieure à celle de l'algorithme de constance de DPD.

Dans le but d'améliorer les estimés du mouvement apparent, des expériences additionnelles pourront être faites avec des interpolateurs d'ordre supérieur à cubique pour essayer d'améliorer la représentation des estimés du mouvement apparent. Une inspection de la réponse impulsionnelle ainsi que de la première et deuxième dérivée de celle-ci révèle que la première dérivée est discontinue. Sachant que l'interpolateur idéal a une réponse impulsionnelle avec un nombre infini de coefficients, il est logique d'expérimenter avec des interpolateurs ayant des réponses impulsionnelles ainsi que leur première et deuxième dérivée continues. Pour ce faire il est nécessaire d'expérimenter avec des polynômes d'ordre supérieur à cubique. De plus, le voisinage d'interpolation pourrait être augmenté de façon à prendre plus de points en considération parce que l'algorithme de DGD utilise des dérivées du deuxième ordre, les estimés du mouvement au niveau 0 sont très mauvais dans les approches pyramidales. En utilisant des dérivées d'ordre supérieur, il serait possible d'estimer les dérivées de deuxième ordre au niveau 0 (une pyramide de DGD sur 4 niveaux et non 3).

Afin d'établir avec certitude l'invariance de la nouvelle formulation à des change-

ments d'illumination, des expériences avec des modèles d'illumination plus élaborées (linéaires et non-linéaires) pourront aussi être réalisées. Des séquences avec modèle d'illumination isolant et combinant chaque type de mouvement (translation, rotation, mouvement de fond, etc.) pourront être créés pour déterminer le rôle qu'ils peuvent avoir sur les estimés du mouvement apparent obtenus avec l'algorithme de constance de DGD.

Appendice A. Les interpolateurs continus

Cette classe d'interpolateurs comprend l'interpolateur quadratique, cubique régulier, cubique de Keys ainsi que quartique parmi d'autres. Cette classe d'interpolateurs permet de calculer des dérivées de premier et de deuxième ordre ainsi que des valeurs d'intensité n'importe où sur le plan de l'image. Grâce à la compensation du mouve-

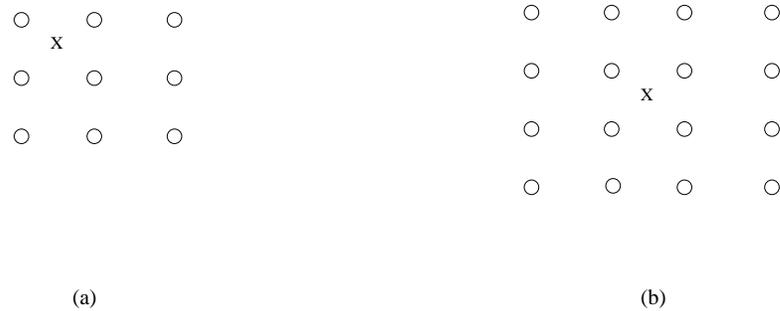


Figure 6.1: Représentation des divers voisinages utilisés par les interpolateurs continus de cette thèse. (a) 3x3, (b) 4x4.

ment, il est possible de devoir évaluer la valeur des dérivées de premier et de deuxième ordre à des positions non-discrètes sur le plan de l'image. Pour ce faire, à partir d'un certain nombre de valeurs avoisinantes, dépendant de l'ordre de l'interpolateur, les dérivées sont calculées. La formulation générale d'un interpolateur est la suivante:

$$w(x) = \sum_{\lambda=1}^N a(\lambda) \cdot w(x - \lambda)$$

Ici, une somme est effectuée sur l'ensemble de tous les points $a(\lambda)$. Ces coefficients caractérisent la fonction d'interpolation. La variable $w(x - \lambda)$ caractérise l'ordre utilisé. Dans les équations à venir, la fonction d'interpolation pour les interpolateurs quadratique, cubique régulier, cubique de Keys et quartique est présentée [13]. Celles-ci sont utilisées dans les travaux d'estimation de dérivées.

Quadratique: cet interpolateur permet d'estimer les valeurs d'intensité ainsi que les valeurs de dérivée de premier et de deuxième ordre partout sur le plan de l'image. Un voisinage de trois points est requis pour permettre l'application de cette

interpolateur. Vu que des dérivées de deuxième ordre sont utilisés dans la deuxième hypothèse, une fonction quadratique est le polynôme du plus bas ordre nécessaire pour obtenir un résultant non-nul pour les dérivées. La réponse impulsionnelle et sa deuxième dérivée sont discontinues.

$$a(x) = \begin{cases} 0 & x < -1.5 \\ 0.5(x^2 + 3x + 2) & -1.5 \leq x < -0.5 \\ -x^2 + 1 & -0.5 \leq x < 0.5 \\ 0.5(x^2 - 3x + 2) & 0.5 \leq x < 1.5 \\ 0 & x > 1.5 \end{cases}$$

$$\frac{d}{dx}a(x) = \begin{cases} 0 & x < -1.5 \\ (x + 1.5) & -1.5 \leq x < -0.5 \\ -2x & -0.5 \leq x < 0.5 \\ (x - 1.5) & 0.5 \leq x < 1.5 \\ 0 & x > 1.5 \end{cases}$$

$$\frac{d^2}{dx^2}a(x) = \begin{cases} 0 & x < -1.5 \\ 1 & -1.5 \leq x < -0.5 \\ 2 & -0.5 \leq x < 0.5 \\ 1 & 0.5 \leq x < 1.5 \\ 0 & x > 1.5 \end{cases}$$

Cubique régulier: cet interpolateur a été jugé comme donnant les meilleurs résultats pour tous les algorithmes d'estimation. Un voisinage de quatre points est requis pour utiliser cet interpolateur. Cet interpolateur a une première dérivée de la réponse impulsionnelle non-continue.

$$a(x) = \begin{cases} 0 & x < -2.0 \\ 0.167(x^3 + 6x^2 + 11x + 6) & -2.0 \leq x < -1.0 \\ -0.5(x^3 + 2x^2 - x - 2) & -1.0 \leq x < 0.0 \\ 0.5(x^3 - 2x^2 - x + 2) & 0.0 \leq x < 1.0 \\ -0.167(x^3 - 6x^2 + 11x - 6) & 1.0 \leq x < 2.0 \\ 0 & 2.0 > x \end{cases}$$

$$\frac{d}{dx}a(x) = \begin{cases} 0 & x < -2.0 \\ 0.167(3x^2 + 12x + 11) & -2.0 \leq x < -1.0 \\ -0.5(3x^2 + 4x - 1) & -1.0 \leq x < 0.0 \\ 0.5(3x^2 - 4x - 1) & 0.0 \leq x < 1.0 \\ -0.167(3x^2 - 12x + 11) & 1.0 \leq x < 2.0 \\ 0 & 2.0 > x \end{cases}$$

$$\frac{d^2}{dx^2}a(x) = \begin{cases} 0 & x < -2.0 \\ 0.167(6x + 12) & -2.0 \leq x < -1.0 \\ -0.5(6x + 4) & -1.0 \leq x < 0.0 \\ 0.5(6x - 4) & 0.0 \leq x < 1.0 \\ -0.167(6x - 12) & 1.0 \leq x < 2.0 \\ 0 & 2.0 > x \end{cases}$$

Cubique de Keys: cet interpolateur a le même ordre que l'interpolateur cubique régulier, donc il a aussi été utilisé avec un voisinage de quatre points. La deuxième dérivée de la réponse impulsionnelle est non-continue.

$$a(x) = \begin{cases} 0 & x < -2.0 \\ 0.5x^3 + 2.5x^2 + 4.0x + 2.0 & -2.0 \leq x < -1.0 \\ -1.5x^3 - 2.5x^2 + 1.0 & -1.0 \leq x < 0.0 \\ 1.5x^3 - 2.5x^2 + 1.0 & 0.0 \leq x < 1.0 \\ -0.5x^3 + 2.5x^2 - 4.0x + 2.0 & 1.0 \leq x < 2.0 \\ 0 & 2.0 > x \end{cases}$$

$$\frac{d}{dx}a(x) = \begin{cases} 0 & x < -2.0 \\ 1.5x^2 + 5x + 4 & -2.0 \leq x < -1.0 \\ -4.5x^2 - 5.0x & -1.0 \leq x < 0.0 \\ 4.5x^2 - 5.0x & 0.0 \leq x < 1.0 \\ -1.5x^2 + 5.0x - 4 & 1.0 \leq x < 2.0 \\ 0 & 2.0 > x \end{cases}$$

$$\frac{d^2}{dx^2}a(x) = \begin{cases} 0 & x < -2.0 \\ 3x + 5 & -2.0 \leq x < -1.0 \\ -9x - 5 & -1.0 \leq x < 0.0 \\ 9x - 5 & 0.0 \leq x < 1.0 \\ -3x + 5 & 1.0 \leq x < 2.0 \\ 0 & 2.0 > x \end{cases}$$

Quartique: cet interpolateur est le plus grand ordre utilisé dans les travaux de cette thèse. Un voisinage de quatre points est choisi pour appliquer cette fonction. La première dérivée de la réponse impulsionnelle est non-continue, sa deuxième dérivée contient des variations brusques de la pente.

$$a(x) = \begin{cases} 0 & x < -1.0 \\ 0.5x^4 + 3.5x^3 + 9x^2 + 10x + 4 & -2.0 \leq x < -1.0 \\ -0.5x^4 - 2.5x^3 - 3.0x^2 + 1 & -1.0 \leq x < 0.0 \\ -0.5x^4 + 2.5x^3 - 3x^2 + 1 & 0.0 \leq x < 1.0 \\ 0.5x^4 - 3.5x^3 + 9x^2 - 10x + 4 & 1.0 \leq x < 2.0 \\ 0 & 2.0 > x \end{cases}$$

$$\frac{d}{dx}a(x) = \begin{cases} 0 & x < -1.0 \\ 2x^3 + 10.5x^2 + 18x + 10 & -2.0 \leq x < -1.0 \\ -2x^3 - 7.5x^2 - 6x & -1.0 \leq x < 0.0 \\ -2x^3 + 7.5x^2 - 6x & 0.0 \leq x < 1.0 \\ 2x^3 - 10.5x^2 + 18x - 10 & 1.0 \leq x < 2.0 \\ 0 & 2.0 > x \end{cases}$$

$$\frac{d^2}{dx^2}a(x) = \begin{cases} 0 & x < -1.0 \\ 6x^2 + 21x + 18 & -2.0 \leq x < -1.0 \\ -6x^2 - 15x - 6 & -1.0 \leq x < 0.0 \\ -6x^2 + 15x - 6 & 0.0 \leq x < 1.0 \\ 6x^2 - 21x + 18 & 1.0 \leq x < 2.0 \\ 0 & 2.0 > x \end{cases}$$

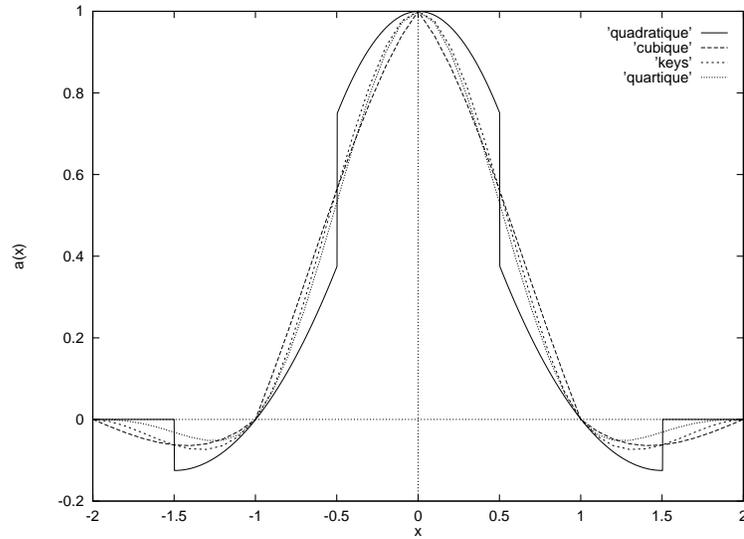


Figure 6.2: Réponse impulsionnelle des interpolateurs continus.

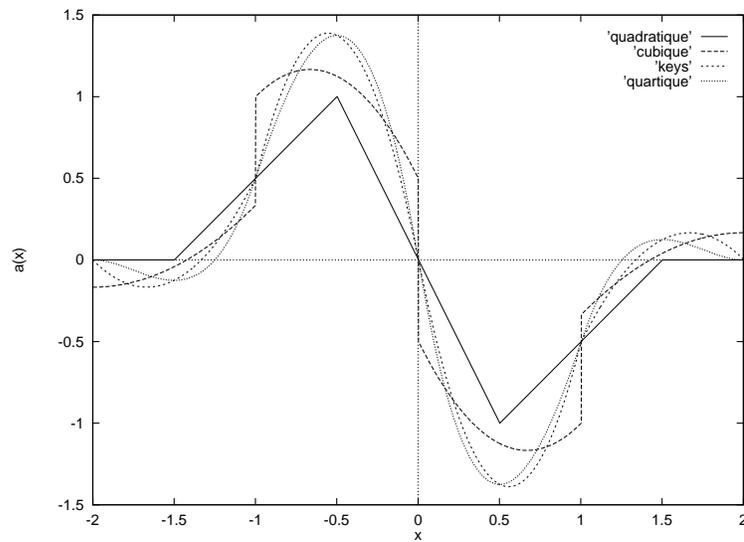


Figure 6.3: Première dérivée des réponses impulsionnelles des interpolateurs continus.

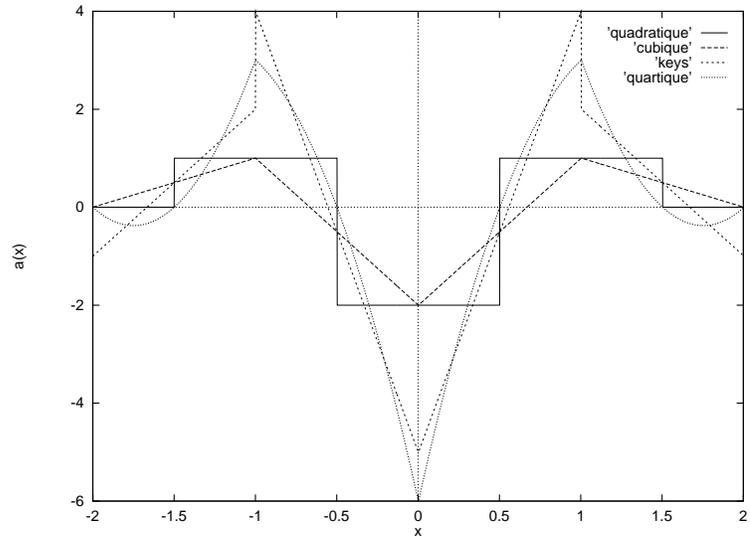


Figure 6.4: Deuxième dérivée des réponses impulsionnelles des interpolateurs continus.

Bibliographie

- [1] B.K.P. Horn and B.G. Schunck, "Determining optical flow," *Artif. Intell.*, vol. 17, pp. 185–203, 1981.
- [2] B. Cole, "The technology framework.," *IEEE Spectrum*, pp. 32–39, March 1993.
- [3] R. Depommier, "Estimation du mouvement considerant les phenomenes d'occlusion pour le codage interpolatif des sequences d'images.," Master's thesis, INRS-Télécommunication, Dec. 1990.
- [4] E. Dubois, "Motion-compensated filtering of time-varying images," *Multidimens. Syst. Signal Process.*, vol. 3, pp. 211–239, 1992.
- [5] E. Dubois and J. Konrad, "Estimation of 2-D motion fields from image sequences with application to motion-compensated processing," in *Motion Analysis and Image Sequence Processing* (M. Sezan and R. Lagendijk, eds.), ch. 3, pp. 53–87, Kluwer Academic Publishers, 1993.
- [6] E. De Micheli and V. Torre and S. Uras, "The accuracy of the computation of optical flow and of the recovery of motion parameters," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 15, pp. 434–437, May 1993.
- [7] W. Enkelmann, "Investigations of multigrid algorithms for the estimation of optical flow fields in image sequences," *Comput. Vis. Graph. Image Process.*, pp. 150–177, March 1988.
- [8] R. Fries and J. Modestino, "Image Enhancement by Stochastic Homomorphic Filtering," *IEEE Trans. Acoust. Speech Signal Process.*, pp. 625–637, May 1979.
- [9] A. Jain, *Fundamentals of digital image processing*, ch. 3, pp. 49–79. Prentice-Hall, 1989.
- [10] M. James, G. Smith, and J. Wolford, *Applied Numerical Methods for digital computation with Fortran and CSMP*, ch. 5, pp. 343–347. Harper and Row Publishers Inc., 2nd ed., 1977.
- [11] J. Kearny, W. Thompson, and D. Boley, "Optical flow estimation: An error analysis of gradient-based methods with local optimization," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-9, no. 2, pp. 229–244, 1987.
- [12] D. Knuth, *Seminumerical Algorithms*, vol. 2, ch. 3, pp. 103–113. Wiley and Sons, 2nd ed., 1981.

- [13] J. Konrad, *Bayesian estimation of motion fields from image sequences*. PhD thesis, McGill University, Dept. Electr. Eng., June 1989.
- [14] J. Konrad and P. Trèves, “Estimation of dense 2-D motion based on the constancy of intensity gradient,” *Signal Process. VII: Theories and Applications (Proc. Seventh European Signal Process. Conf.)*, Sept. (accepté).
- [15] M. Bertero and T. Poggio and V. Torre, “Ill-Posed Problems in Early Vision,” *Proc. IEEE*, vol. 76, pp. 869–889, August 1988.
- [16] J. Modestino and R. Fries, “Construction and Properties of a useful Two-Dimensional Random Field,” *IEEE Trans. Inf. Theory*, pp. 44–50, May 1980.
- [17] C. Moloney, “Estimation of motion fields from image sequences with illumination variation,” Tech. Rep. 90-30, INRS-Télécommunication, Sept. 1990.
- [18] H. Nagel, “Recent advances in image sequences analysis,” *Proc. Premier colloque Image Traitement Synthèse Technologie Applications*, pp. 545–558, 1984.
- [19] H. Nagel, “Extending the orientation smoothness constraint into the temporal domain and the estimation of derivatives of optical flow,” *Proc. First Europ. Conf. Comput. Vision*, 1990.
- [20] O. Tretiak and L. Pastor, “Velocity estimation from image sequences with second order differential operators,” *Proc. IEEE Int. Conf. Pattern Recognition*, vol. 1, pp. 16–19, July 1984.
- [21] W. Press, B. Flannery, S. Teukolsky, and W. Vetterling, *Numerical Recipes in C*, ch. 7, pp. 204–217. Press Syndicate of the University of Cambridge, 1990.
- [22] R. Keys, “Cubic convolution interpolation for digital image processing,” *IEEE Trans. Acoust. Speech Signal Process.*, vol. ASSP-29, pp. 1153–1160, December 1981.
- [23] P. Trèves and J. Konrad, “Motion estimation and compensation under varying illumination,” *Proc. IEEE Int. Conf. Image Processing*, Nov. 1994. (accepté).
- [24] A. Verri, F. Girosi, and V. Torre, “Mathematical properties of the two-dimensional motion field: from singular points to motion parameters,” *J. Opt. Soc. Am.*, vol. 6, pp. 698–712, May 1989.