

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

**Université Mouloud MAMMERI de TIZI-OUZOU
Faculté de Génie Electrique et d'Informatique
Département d'Electronique**



**Mémoire de Magister en Electronique
Option : Télédétection**

Présenté par :
AYOUNI Boudjemma

Thème

**Estimation du mouvement dans une séquence
d'images par approche probabiliste**

Mémoire soutenu le: / /

devant le jury composé de

Président : Ziani Rezki

Maitre de conférences UMMTO

Rapporteur : Ameer Soltane

Professeur UMMTO

Examineur : Ameer Zohra

Maitre de conférences UMMTO

Examineur : Lahdir Mourad

Maitre de conférences UMMTO

Examineur : Bennamane Kamal

Maitre de conférences UMMTO

Remerciements

Ce travail a été réalisé au Laboratoire d'Analyse et de Modélisation des Phénomènes Aléatoires LAMPA de l'UMMTO sous la direction de Mr Ameur Soltane directeur du laboratoire, je le remercie pour son accueil.

Je tiens à exprimer ma reconnaissance pour M. Ameur Soltane professeur à l'UMMTO et directeur du laboratoire LAMPA d'avoir accepté d'être rapporteur de mon mémoire.

Je remercie également Mme Saida Abdat étudiante en doctorat à l'UMMTO pour m'avoir accompagné tout au long de mon mémoire.

Je remercie M. Mostafaoui Ghiles Enseignant en embarqué et traitement d'images à l'université de Cergy-Pontoise pour l'aide précieuse qu'il m'a apporté durant l'élaboration de ce travail avec ses différentes corrections et avis.

Un grand merci à M Arezki Ziani maitre de conférences à l'UMMTO de m'avoir fait l'honneur de présider le jury de mon mémoire.

Je suis également très reconnaissant à Mme Zohra Ameur maitre de conférences à l'UMMTO de m'avoir fait l'honneur d'être membre du jury de mon mémoire.

Je remercie M Lahdir Mourad maitre de conférences à l'UMMTO de me faire l'honneur d'être membre du jury de mon mémoire.

Je remercie M Bennamane Kamal maitre de conférences à l'UMMTO de me faire l'honneur d'être membre du jury de mon mémoire.

Résumé

Dans cette étude, nous nous sommes intéressés à l'estimation du mouvement par une approche probabiliste, après un état de l'art des différentes méthodes d'estimation du mouvement existantes et le développement de deux méthodes connues à savoir Horn et Shunck HS et Lucas Kanade LK [1], [2], nous avons développé l'approche probabiliste dans laquelle l'estimation du mouvement 2-D est formulée comme un problème bayésien ; par ailleurs, une contrainte de lissage est introduite par la modélisation d'un champ de mouvement 2-D par la distribution de Gibbs [7].

Le modèle proposé dans cette thèse est un modèle globalement lisse basé sur les champs de Markov aléatoires et le critère d'estimation est celui de la probabilité du maximum à posteriori (MAP) [5], dans lequel cette probabilité est maximisée compte tenu des données d'entrée. L'estimation du MAP est réalisée par le biais de l'algorithme du recuit simulé [8], dans lequel l'échantillonnage de l'espace des solutions se fait au moyen de l'échantillonneur de Gibbs. Les mauvaises données sont éliminées en utilisant la méthode de rejet des valeurs aberrantes [6].

Mots-clés : Estimation du mouvement, flot optique, Horn et Shunck, Lucas Kanade, approche probabiliste, théorème de Bayes, Réseaux de Markov, distribution de Gibbs, Maximum à Posteriori, modèle de vraisemblance, modèle à priori, recuit simulé.

Table des matières

Introduction2

Chapitre I : état de l'art

I.1 Préambule	5
I.2 Estimation du mouvement (Motion estimation)	5
I.3 Le mouvement	6
I.4 Champ de mouvement réel et apparent	6
I.5 Problèmes de l'estimation de mouvement	7
I.5.1 Problème d'occlusion	7
I.5.2 Problème d'ouverture	8
I.6 Représentation du tracé du champ de mouvement	9
I.7 Les différentes méthodes d'estimation du mouvement	10
I.7.1 Equation de contrainte du flot optique	10
I.7.2 Méthodes différentielles	11
I.7.2.1 Horn et Shunck	11
I.7.2.2 Lucas et Kanade	12
I.7.3 Approche de la mise en correspondance de blocks	12
I.7.3.1 Block Matching	12
I.7.4 Approche Probabiliste	13
I.8 Discussion	15

Chapitre II : calcul du flot optique

II.1 Préambule.	17
II.2 Flot optique.	17
II.3 Le problème restreint de domaine	17
II.4. Méthode de Horn et Shunck	18
II.4.1 Contraintes.	18
II.4.2 La contrainte de lissage	20
II.4.3 Estimation des dérivées partielles	21
II.4.4 Estimation du laplacien de la vitesse d'écoulement	22
II.4.5 Minimisation	23
II.4.6 Contrainte de minimisation	24
II.4.7 Solution itérative	25
II.5 Résultats pratiques	25
II.6 Méthode de Lucas et Kanade	28
II.6.1 Formulation mathématique	28
II.6.2 Améliorations apportées à cette méthode	29
II.6.2.1 Implémentation pyramidale	29
II.6.2.1 Raffinement itératif	30

II.6.3 Description de l'implémentation	30
II.6.4 Résultats pratiques	31
II.7 Discussion	33

Chapitre III : méthode probabiliste

III.1 Préambule	35
III.2 Formulation du problème du MAP	35
III.3 Formulation mathématique	38
III.4 Le modèle de vraisemblance	40
III.5 Le modèle à priori	41
III.6 Méthodologie à suivre	43
III.7 Organigrammes détaillés de la méthode	49
III.8 Résultats pratiques	53
III.9 Discussion	57

Conclusion

Annexes

Annexe A : formulation de l'équation du flot optique	61
Annexe B : introduction au champ de Markov	62
Annexe C : l'échantillonneur de Gibbs	67

Bibliographie

Introduction

Durant ces dernières années, nous assistons à la convergence rapide des ordinateurs et de la communication, conduisant à une profonde révolution de l'information, ce qui changera finalement non seulement la façon dont les gens font des affaires, mais aussi l'apprentissage et le divertissement.

Le traitement d'image, inventé au début des années soixante dix, est de nos jours utilisé dans divers domaines avec des applications de plus en plus nombreuses.

Le concept du traitement d'images est assez simple, il a en effet très souvent pour but la reconnaissance d'objets ou de formes que l'œil humain perçoit rapidement et avec efficacité. En pratique et malgré le foisonnement important de travaux que suscite ce domaine de recherche, il reste difficile d'extraire du flux de données que contient une image, des informations claires, précises et utiles pour l'application souhaitée. Les problèmes rencontrés sont d'autant plus frustrants qu'ils peuvent souvent être résolus par un observateur humain qui reconnaîtra aisément une personne ou un objet sur une image un peu floue. Néanmoins, l'analyse d'image est aujourd'hui dotée d'outils mathématiques et informatiques assez puissants qui permettent de repousser ces limites et d'étendre les domaines d'applications notamment à l'analyse des séquences d'images.

L'estimation du mouvement est l'un des domaines du traitement d'images qui joue un rôle considérable dans l'analyse des séquences d'images de part le nombre considérable de ses applications.

La mesure du champ de mouvement est une étape de traitement de l'image dite de bas niveau. On lui trouve de nombreuses applications comme l'analyse de mouvements de fluides en physique expérimentale, la compression de séquences d'images vidéo par compensation de mouvement, ou encore son utilisation pour des phases de traitement des images de plus haut niveau, comme la reconstruction de scènes tridimensionnelles.

En 1980, Horn et Schunck ont proposé une méthode d'estimation du mouvement basée sur la régularisation. Ce premier travail a été suivi d'un grand nombre de contributions qui ont proposé différentes méthodes alternatives.

Parmi ces méthodes, nous citerons les méthodes différentielles, les méthodes de mises de correspondances de blocks, les méthodes probabilistes ...etc.

Structure du mémoire

Dans le premier chapitre du mémoire, nous avons présenté une revue du travail existant dans le domaine de l'estimation du mouvement, ainsi, les méthodes différentielles, la méthode de mise

en correspondance de blocks et l'approche probabiliste sont présentées tout en évoquant les principaux problèmes rencontrés dans l'estimation du mouvement.

Dans le second chapitre, nous avons retenu les méthodes différentielles pour présenter les fondements mathématiques régissant les algorithmes nous permettant d'estimer le mouvement, tout en proposant des améliorations telles que l'implémentation pyramidale et le raffinement itératif, rendant ainsi les algorithmes plus précis que ceux proposés au départ (début années 80)

Dans le troisième chapitre, le problème de l'estimation du mouvement 2D est formulé comme un problème bayésien. Le modèle champ de mouvement proposé est un modèle basé sur les champs de Markov aléatoire et le critère d'estimation est celui de la probabilité du maximum à posteriori (MAP), cette dernière est maximisée par le biais de l'algorithme du recuit simulé, dans lequel l'échantillonnage de l'espace des solutions se fait au moyen de l'échantillonneur de Gibbs. Les mauvaises données sont éliminées en utilisant la méthode « local outlier rejection » rejet des valeurs aberrantes.

Les résultats expérimentaux seront présentés après le développement de chacune des méthodes citées.

Chapitre I

Etat de l'art

I.1 Préambule

Ce chapitre introductif est consacré à la description des différentes méthodes d'estimation du mouvement.

Nous avons dans un premier temps donné les différentes définitions relatives au domaine et par la suite quelques méthodes de calcul du flot optique ainsi qu'à la façon choisie pour représenter les champs des vecteurs mouvements en mentionnant également les principaux problèmes relatif à l'estimation du mouvement.

I.2 Estimation du mouvement (Motion estimation)

L'estimation du mouvement est un procédé qui consiste à étudier le déplacement des objets dans une séquence vidéo, en cherchant la corrélation entre deux images successives afin de prédire le changement de position du contenu.

Le mouvement est un réel problème en vidéo puisqu'il décrit un contexte en trois dimensions alors que les images sont une projection de scènes 3D dans un plan en 2D. En général, il est représenté par un vecteur de mouvement qui décrit une transformation d'une image en deux dimensions vers une autre.

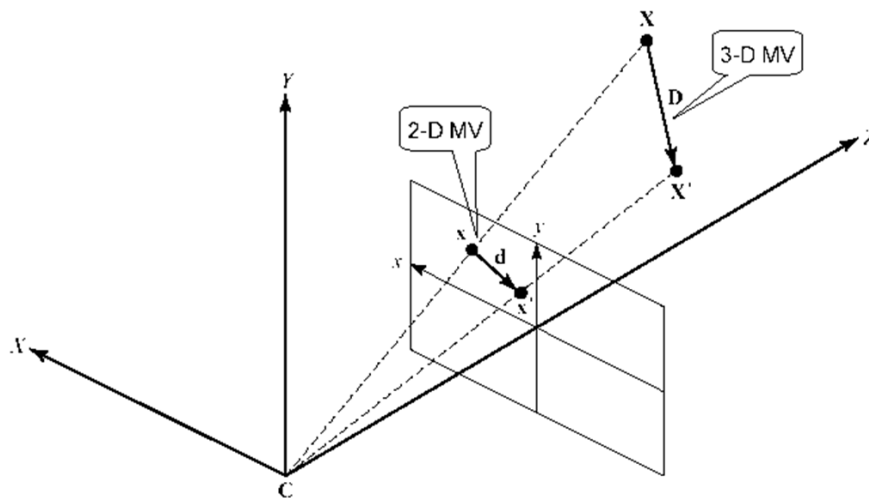


Figure I.1 Projection mouvement 3D en 2D

Les vecteurs de mouvement lient deux blocs par projection des blocs sur une même image, c'est-à-dire que les coordonnées du vecteur ne sont définies que spatialement sans contraintes temporelles. Les blocs en question sont différents selon l'algorithme; toute l'image peut être liée au vecteur comme c'est le cas pour l'estimation de mouvement global, ou juste des parties spécifiques de l'image, tels que des blocs rectangulaires, des formes arbitraires ou même

par pixel. Les vecteurs peuvent être représentés par un modèle traditionnel ou par de nombreux autres modèles qui peuvent se rapprocher du mouvement d'une caméra vidéo réelle, tel que la rotation et la translation dans les trois dimensions sans oublier le zoom. Cette technique de prédiction est surtout utilisée en compression vidéo, en robotique et en traitement d'image.

I.3 Le mouvement

Le mouvement dans une séquence d'images bidimensionnelles (2D), est perceptible grâce aux changements de la distribution spatiale des intensités lumineuses. Le mouvement ainsi perçu est appelé « apparent », car il ne correspond pas nécessairement à la projection dans le plan de l'image, du mouvement ayant lieu dans l'espace tridimensionnel (3D). Ainsi, par exemple, les vitesses apparentes des points situés sur une sphère uniforme en rotation sont nulles, donc différentes des projections des vraies vitesses de ces points. Le champ des vitesses apparentes porte le nom de « flot optique ». Nous utiliserons indifféremment les termes « flot optique » et « champ de vitesses » pour désigner le champ des vitesses apparentes.

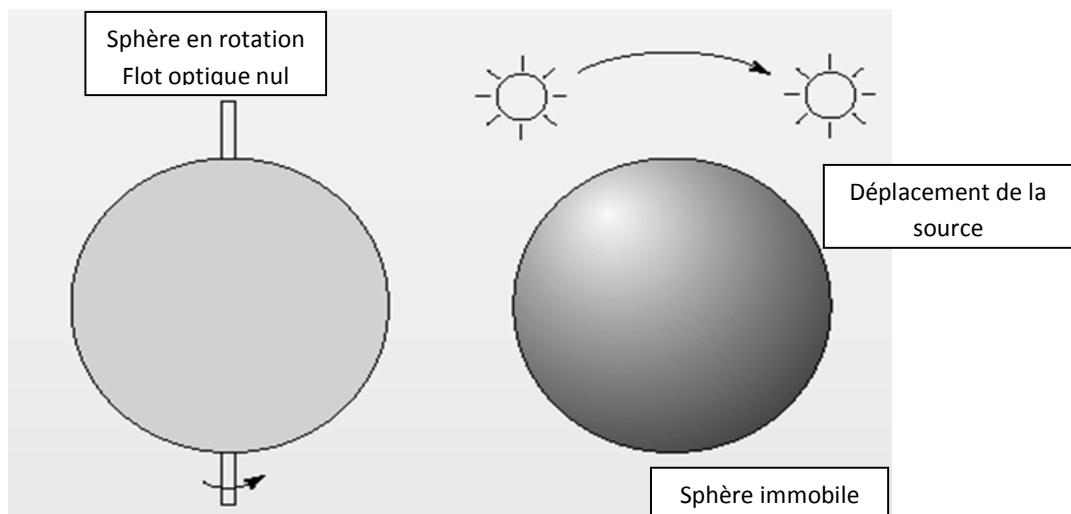


Figure I.2 Illustration du phénomène de la sphère immobile et en rotation

I.4 Champ de mouvement réel et champ de mouvement apparent

Le mouvement (ou déplacement) réel anime la scène réelle, dans l'espace réel 3D. Ce mouvement réel est observé soit par l'œil humain, soit par un système de prise de vue. Le second fournit un enregistrement du mouvement observé ou apparent sous la forme d'une séquence temporelle d'images.

Souvent, le champ de mouvement réel n'est pas directement mesurable dans une séquence d'images 2D ou 3D si l'on s'appuie seulement sur l'intensité des pixels, dans ce cas, on observe en fait des changements de la distribution spatiale d'intensité lumineuse. Le mouvement ainsi perçu est appelé « mouvement apparent » ou « flot optique » qui, en général, est différent du champ réel du mouvement.

Le champ de mouvement ou de déplacement apparent représente, en général, la projection (orthographique ou projective) du mouvement réel 3D dans le plan image c'est pourquoi, le champ de mouvement apparent s'appelle aussi « mouvement projeté », il représente une approximation du mouvement réel.

I.5 Problèmes de l'estimation de mouvement

L'estimation de mouvement est très sensible au bruit présent dans les images qui peut être interprété comme étant le résultat d'un mouvement dans la scène réelle. Deux autres problèmes sont souvent rencontrés et posés lors de l'estimation de mouvement :

Problème d'occlusion

Problème d'ouverture

I.5.1 Problème d'occlusion

Le problème des occlusions est un problème non résolu dans le traitement de séquences vidéo. Une séquence vidéo donne une représentation en deux dimensions d'une scène dynamique de trois dimensions. La scène est dynamique non seulement parce que les éléments qui la composent ne sont pas fixes et peuvent bouger et/ou interagir entre eux, mais également parce que la caméra qui requiert cette séquence n'est pas fixe et peut bouger par rapport au référentiel de la scène.

Le phénomène d'occlusion dans une séquence vidéo est du à l'apparition et à la disparition ou aux croisements d'objets dans la vidéo. C'est le résultat du mouvement de la camera et des objets dans la scène.

L'apparition et/ou la disparition de texture sont dues au mouvement des objets dans la scène mais peuvent être également dues aux déformations des objets. Par exemple, un ballon immobile qui se dégonfle fait apparaître de la texture du fond qui n'était pas visible lorsque le ballon était gonflé.

Les occlusions rendent difficiles l'estimation de mouvement apparent dans la séquence vidéo. En effet, les techniques d'estimation de mouvement entre deux images prises à deux instants t et $t+1$ d'une séquence vidéo sont basées sur les ressemblances existantes entre ces deux

images. Si des zones apparaissent ou disparaissent entre les deux images, les ressemblances diminuent et l'estimation échoue dans ces zones.

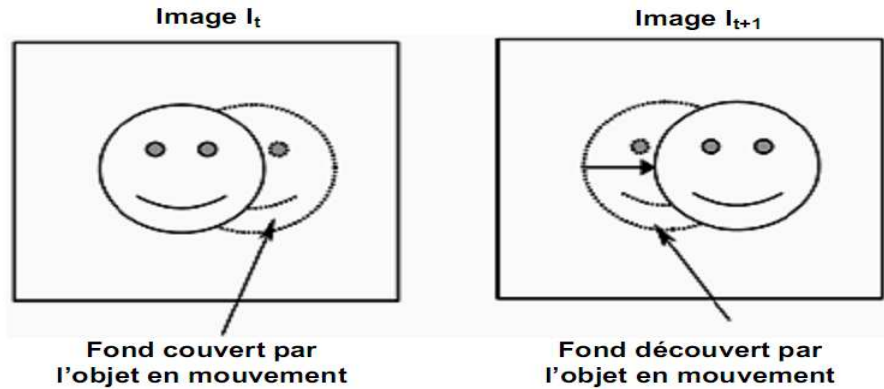


Figure I.3 : Illustration du principe d'occlusion

I.5.2 Problème d'ouverture

Le problème d'ouverture représente une formulation du fait que la solution du problème d'estimation de mouvement n'est pas unique. Si on suppose les vecteurs de mouvement en chacun des pixels comme des variables indépendantes, le nombre d'inconnues sera (dans le cas 2D) deux fois plus grand que le nombre d'équations disponibles. Ceci est dû au fait que le nombre d'équations est égal au nombre de pixels de l'image et le vecteur de mouvement en chaque pixel a deux composantes [1].

En absence de contraintes supplémentaires, on ne peut déterminer en chaque point que la composante normale du déplacement qui est orientée dans la direction du gradient spatial de l'intensité au point considéré.

Le problème d'ouverture est illustré sur la Figure I.4. On suppose que l'un des coins d'un objet est en mouvement dans la direction verticale haute. Pour estimer le mouvement, on utilise un bloc noté B1, il n'est pas possible de déterminer si l'objet est en mouvement dans la direction verticale supérieure (comme en réalité) ou dans la direction normale au bord de l'objet.

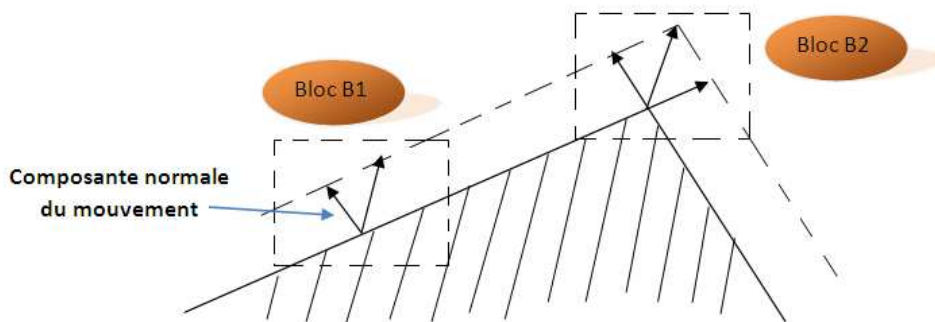


Figure I.4 : Illustration du problème d'ouverture

Si pour estimer le mouvement, on utilise le bloc B2, on peut déterminer correctement le mouvement, parce que le gradient d'intensité dans ce bloc a deux composantes normales entre elles. Par conséquent, pour éviter le problème d'ouverture, on doit utiliser le bloc qui contient suffisamment d'informations traduites par des variations locales des niveaux de gris.

I.6 Représentation du tracé du champ de mouvement

Pour être à même de comparer les différentes estimations obtenues, il faut utiliser la même représentation visuelle du champ de vitesses. De nombreuses solutions sont possibles pour représenter visuellement le flot optique calculé. La plus couramment utilisée est le tracé du champ (en prenant 1 point sur 10 par exemple) des vecteurs vitesses superposé à l'image initiale, l'inconvénient étant que l'on ne visualise pas le champ dense, et donc que les points aberrants risquent d'être masqués ou au contraire mis en valeur, conduisant à une interprétation biaisée.

On peut également représenter en niveaux de gris respectivement u et v . La solution la plus intéressante est l'utilisation d'une carte de couleurs, qui permet de représenter la direction du flot ainsi que son intensité de manière dense. La carte des couleurs que nous utiliserons est la suivante (Figure I.5). On présente également le tracé du champ de vitesse comme on le trouve dans la littérature, ce qui permet de bien comprendre la carte.

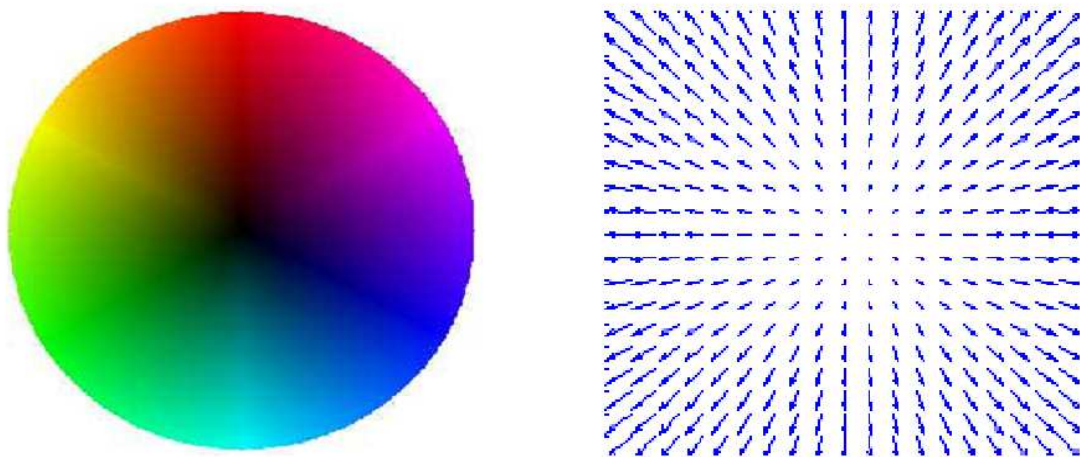


Figure I.5. Colormap et affichage du même champ de vitesse sous forme de vecteurs

Les vecteurs vitesses sont représentés par les couleurs contenues dans ce cercle. Chaque vecteur vitesse est codé par la couleur qu'il indique en plaçant son origine au centre du cercle. En fonction de l'angle le point prend une couleur différente, et son intensité varie du noir à la couleur complète selon la norme du vecteur. Il faut également donner une vitesse maximale au delà de laquelle l'affichage est alors un point blanc. Sinon, Cette cartographie revient finalement

à construire une application qui associe à l'angle et à la norme du vecteur vitesse un point dans l'espace (R,G,B).

Ce type de représentation permet une représentation dense (une couleur pour chaque pixel), et permet de visualiser très rapidement la cohérence des résultats, la sensibilité au bruit de la méthode testée et l'aspect (lissage, discontinuités ...) du flot optique.

I.7 Les différentes méthodes d'estimation du mouvement

Dans cette partie, on s'attachera à décrire les approches de calcul du flot optique existantes :

I.7.1 Equation de contrainte du flot optique

Considérons un patch (block) d'une certaine luminosité qui se déplace sur une distance δx dans la direction des x et δy dans la direction y en δt dans le temps [1], et soit $f(x, y, t)$ l'intensité de l'image à l'instant t . L'hypothèse fondamentale utilisée dans le domaine de l'estimation du mouvement est que l'intensité lumineuse de ce block se conserve entre deux images successives.

Cela s'écrit sous la forme générale $f(x, y, t+1) - f(x, y, t) = 0$ I.1

L'hypothèse de conservation peut également s'écrire sous la forme différentielle suivante:

$$\frac{df(x, y, t)}{dt} = \frac{dE(x, y, t)}{dt} = 0 \quad \text{Ce qui conduit à} \quad \frac{dE(x, y, t)}{dt} = \frac{\partial E}{\partial x} \frac{dx}{dt} + \frac{\partial E}{\partial y} \frac{dy}{dt} + \frac{\partial E}{\partial t} = 0 \quad \text{I.2}$$

$$\text{Puis} \quad \left[\frac{\partial E}{\partial x} \quad \frac{\partial E}{\partial y} \right] \begin{bmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{bmatrix} + \frac{\partial E}{\partial t} = 0 \quad \text{et donc enfin} \quad (\nabla E)^T w + E_t = 0 \quad \text{I.3}$$

Est appelée équation de contrainte du mouvement apparent. Avec $w = \begin{pmatrix} u \\ v \end{pmatrix}$

Remarque: L'hypothèse de conservation de la luminance est forte car elle suppose que les variations temporelles de luminance sont uniquement dues au mouvement. On a par ailleurs gardé simplement les termes du premier ordre.

Cette contrainte est insuffisante pour déterminer le flot complet $\vec{w}$ car le problème est mal posé; on dispose d'une équation linéaire pour deux inconnues. Cela signifie que l'on est simplement en mesure d'évaluer, la vitesse normale $w_{\perp}$ dirigée selon le gradient local.

Concrètement, on est en présence de plusieurs solutions possibles pour évaluer la solution réelle, il faut donc particulariser la solution, ajouter une contrainte de régularisation notamment.

On peut évaluer la vitesse normale de la manière suivante:

Suite à (1), on a $\vec{w} = \frac{E_t}{(\nabla E)}$ On définit par ailleurs le vecteur normal $\vec{n} = \frac{\nabla E}{|\nabla E|}$

La projection de la vitesse sur la normale est alors $s = \vec{w}^T \cdot \vec{n}$

La vitesse normale vaut $\vec{w}_{\perp} = (\vec{w}_n \cdot \vec{n}) \cdot \vec{n} = \frac{-E_t}{(\nabla E)} \cdot \frac{\nabla E}{|\nabla E|} \cdot \frac{\nabla E}{|\nabla E|} = \frac{-E_t \nabla E}{|\nabla E|^2}$ I.4

Le flot optique est exactement équivalent au mouvement image si les hypothèses suivantes sont respectées :

- éclairage uniforme
- surface à réflectivité Lambertienne
- mouvement de translation pure, parallèle au plan de l'image.

Ces conditions ne sont bien sûr jamais rigoureusement respectées dans le cas d'images réelles. On peut toutefois les supposer vérifiées localement.

I.7.2 Méthodes différentielles

I.7.2.1 Horn et Shunck

C'est la première méthode de calcul du flot optique, développée en 1980 par Horn & Schunck, ce type d'approche consiste à minimiser sur le domaine entier de l'image la fonctionnelle I.5 prenant en compte l'équation I.1 du flot optique ainsi qu'un terme de lissage, c'est à dire en ajoutant une contrainte de régularisation portant sur le gradient, le laplacien (ou ordre supérieur) du champ de vitesse [1].

Afin d'évaluer le déplacement de chaque pixel, Horn et Schunck ont utilisé la méthode du multiplicateur de Lagrange pour minimiser une somme pondérée de l'erreur dans le flot optique et la variation d'un pixel à un autre de la vitesse:

$$u(x, y, t), v(x, y, t) = \arg \min_{u, v} \int (e_b^2(u, v) + \alpha e_c^2(u, v)) dx dy \quad I.5$$

α : Un paramètre qui contrôle la contrainte de lissage, est faible pour un mouvement fluide et grand pour un mouvement moins lisse.

e_b : Erreur dans l'équation du flot optique.

e_c : Variation pixel à pixel de la vitesse.

Cette méthode a l'avantage de fournir un flot dense (un résultat pour chaque pixel).

I.7.2.2 Lucas et Kanade

Leur méthode consiste à prendre en compte des hypothèses supplémentaires sur un domaine de taille réduite pour particulariser le flot optique. On minimise alors un critère sur un petit domaine, et on obtient ainsi le flot optique de ce petit domaine [3].

La vitesse locale est supposée constante sur un voisinage spatial Ω , on minimise alors la fonctionnelle E_{LK} sur ce domaine:

$$E_{LK} = \sum_{\Omega} \left[\frac{\partial E(x, y, t)}{\partial x} u(t) + \frac{\partial E(x, y, t)}{\partial y} v(t) + \frac{\partial E(x, y, t)}{\partial t} \right]^2 \quad I.6$$

$$\text{Avec } \begin{cases} u(x, y, t) = u(t) \\ v(x, y, t) = v(t) \end{cases} \quad \text{pour tout } (x, y) \in \Omega$$

Cette méthode est très intéressante car hautement parallélisable (chaque calcul sur une petite fenêtre est indépendant des autres).

I.7.3 Approche basée sur la mise en correspondance de blocks

I.7.3.1 Block Matching

A partir d'un bloc de taille donnée de l'image, on cherche le déplacement de ce bloc (approximation de la vitesse) entre les deux images. On autorise pour cela un déplacement maximal et on cherche (différentes techniques d'exploration) le bloc qui correspond au mieux au bloc initial, de manière à minimiser un critère d'erreur [2].

Les critères de corrélation les plus souvent employés sont les suivants :

$$\int_{\Omega} f_1(x + v_x, y + v_y) f_2(x, y) dx dy \quad \text{Corrélation simple} \quad \text{I.7}$$

$$SSD = \sum \sum (E(x, y) - J(x + u, y + v))^2 \quad \text{Somme des différences au carré} \quad \text{I.8}$$

$$SAD = \sum \sum |E(x, y) - J(x + u, y + v)| \quad \text{Somme des différences absolues} \quad \text{I.9}$$

De nombreux algorithmes d'exploration ont été développés, le plus simple étant l'exploration de tous les blocs de la fenêtre de recherche.

L'inconvénient majeur des méthodes de block matching est qu'elles postulent un déplacement maximal sur une certaine fenêtre, on exclue ainsi les grands déplacements et le résultat est biaisé. Une solution consiste donc à utiliser une pyramide comprenant des niveaux de raffinement différents.

I.7.4 Approche Probabiliste

La méthode probabiliste se base sur la nature stochastique du contenu de l'image, ainsi un modèle de différence d'images consécutives à travers une distribution exponentielle.

Afin d'optimiser la fonction d'estimation du mouvement en utilisant cette approche probabiliste, Très souvent deux modèles sont nécessaires. L'estimation de la probabilité du maximum à posteriori, à savoir le modèle de vraisemblance qui permet de donner le mouvement terrain utilisant la fonction de densité de probabilité conditionnelle de l'intensité de l'image observée, et la fonction de densité de probabilité a priori des vecteurs de mouvement, appelée le modèle de champ de mouvement ou le modèle à priori.

La formule de base de l'estimation de la probabilité du maximum à priori est formulée ainsi ; ayant deux images consécutives $f(x, y, t_1)$ et $f(x, y, t_2)$, trouver les champs de mouvement $u(x, y, t_1)$ et $v(x, y, t_1)$ tels que :

$$(u(x, y, t_1), v(x, y, t_2)) = \arg \max_{u, v} p(u(x, y, t_1), v(x, y, t_2) | f(x, y, t_1), f(x, y, t_2)) \quad \text{I.10}$$

Le théorème de Bayes nous donne :

$$(u(x, y, t_1), v(x, y, t_2)) = \frac{p(f(x, y, t_2) | u(x, y, t_1), v(x, y, t_2), f(x, y, t_1)) \cdot p(u(x, y, t_1), v(x, y, t_2) | f(x, y, t_1))}{p(f(x, y, t_2) | f(x, y, t_1))} \quad \text{I.11}$$

Le premier facteur du numérateur est la fonction du modèle de vraisemblance qui représente la probabilité d'avoir l'image courante, connaissant l'image précédente et en appliquant les champs de mouvement $u(x, y, t_1)$ et $v(x, y, t_1)$ à cette dernière.

Notons que si $f(x, y, t)$ est l'image réelle et $f(x, y, t)$ l'image observée, alors :

$f(x, y, t) = f(x, y, t) + n(x, y, t)$ avec $n(x, y, t)$ le bruit observé (blanc, gaussien, ayant pour variance σ^2 et une moyenne $m = 0$).

Cela nous conduit à formuler cette fonction de vraisemblance de la façon suivante :

$$p(f(x, y, t_2) | u(x, y, t_1), v(x, y, t_1), f(x, y, t_1)) = (2\pi\sigma^2)^{-\frac{1}{N}} \exp \left(- \sum_{(x, y) \in I} \frac{[f(x, y, t_2) - f(x - u(x, y, t_1), y - v(x, y, t_1), t_1)]^2}{2\sigma^2} \right)$$

N: nombre total de pixel de l'image I

I.12

Le second facteur de l'équation I.11, représente le modèle à priori, la probabilité d'avoir $u(x, y, t_1)$ et $v(x, y, t_1)$ comme champs de mouvement horizontal et vertical connaissant l'image $f(x, y, t_1)$.

Cette fonction du modèle à priori est assimilé à un champ de markov (annexeB) ayant une distribution Gibbsienne donnant ainsi :

$$p(u(x, y, t_1), v(x, y, t_1) | f(x, y, t_1)) = \frac{1}{Z} \exp \left[-U_p(u(x, y, t_1), v(x, y, t_1) | f(x, y, t_1)) \right]$$

I.13

Z : fonction de partition et $U_p(u(x, y, t_1), v(x, y, t_1) | f(x, y, t_1)) = \lambda \sum_{c \in C} V_c^p(u(x, y, t_1), v(x, y, t_1) | f(x, y, t_1))$

C : représente l'ensemble des cliques du système de voisinage

V_c^p : Le potentiel de la clique $c \in C$, et λ une constant positive.

En remplaçant les équations I.12 et I.13 dans I.11, on obtient la formule basic de l'estimation de la probabilité du Maximum A Postérieur (MAP)

Nous modélisons q un champ de vecteur aléatoire avec une distribution de Gibbs comme suit:

$$p(q) = \frac{1}{Z} \exp \left(\frac{U(q)}{T} \right)$$

I. 14

Où T est le paramètre de température de l'algorithme du recuit simulé et Z est la constante de normalisation telle que $\sum_{q \in Q} p(q) = 1$.

Ensuite un algorithme d'optimisation sera appliqué afin d'obtenir la meilleure estimation des vecteurs mouvements en combinant les potentiels énergies des pixels de l'image et les probabilités conditionnelles locales, et ceci jusqu'à convergence.

I.8 Discussion

Nous avons présenté les techniques d'estimation de mouvement les plus utilisées à savoir les méthodes basées sur la mise en correspondance de blocs ainsi que les méthodes différentielles et l'approche probabiliste qui feront l'objet de notre étude dans les chapitres à venir.

Nous avons défini les différents types de mouvement ainsi que les principaux problèmes rencontrés dans l'estimation des vecteurs mouvements.

Chapitre II

Calcul du flot optique

II.1 Préambule

Dans cette partie, il s'agit de tester les approches retenues de manière à décrire le processus de l'estimation du flot optique.

Nous présenterons les deux méthodes les plus connues dans ce domaine, à savoir la méthode différentielle de Horn et Schunck et celle de Lucas et Kanade auxquelles nous attribuerons des améliorations telles l'implémentation pyramidale et le raffinement itératif afin d'optimiser les algorithmes initiaux.

II.2 Flot optique

Le flot optique est la distribution apparente du mouvement de l'intensité lumineuse des pixels de l'image. Il peut être du au mouvement des objets de la scène captée ou à celui du point de prise de vue.

Le flot optique nous renseigne à la fois sur l'arrangement spatial des objets de la scène et sur le taux de changement dans l'objet en question, les discontinuités dans le flot optique peuvent nous aider lors de la segmentation de l'image nous permettons d'avoir les régions correspondantes aux différents objets. Dans quelques cas, des informations sur la forme de ces objets peuvent être retrouvées.

Le flot optique ne peut être calculé à un point de l'image indépendamment de ses voisins sans l'introduction de contraintes additionnelles, parce qu'il résulte à chaque point de l'image en deux composantes alors que le changement dans l'intensité lumineuse de ce point due au mouvement résulte uniquement en une seule composante.

II.3 Le problème restreint de domaine

Pour éviter des variations de luminosité dues aux effets d'ombrages nous allons d'abord supposé que la surface imagée est plate. Nous supposons en outre que l'éclairage incident est uniforme sur toute la surface

La luminosité d'un point dans l'image est alors proportionnelle à la réflectance de la surface au point correspondant sur l'objet. Aussi, nous supposons d'abord que la réflectance varie régulièrement et n'a pas de discontinuités spatiales. Cette dernière condition nous assure que la luminosité de l'image est dérivable. Nous excluons les situations où des objets occlurent sur d'autres, en partie parce que des discontinuités de la réflectance sont présentes à la périphérie de l'objet.

Dans la situation décrite, le mouvement de la luminosité dans les modèles pris est déterminé directement par le déplacement des points correspondants sur la surface de l'objet. Le calcul de la vitesse des points sur l'objet est une question de géométrie simple une fois que le flot optique est connu.

II.4 Méthode de Horn et Schunck

Nous allons dans un premier temps illustrer le fonctionnement de la méthode globale HS, en suivant les différentes étapes décrites dans la figure qui suit.

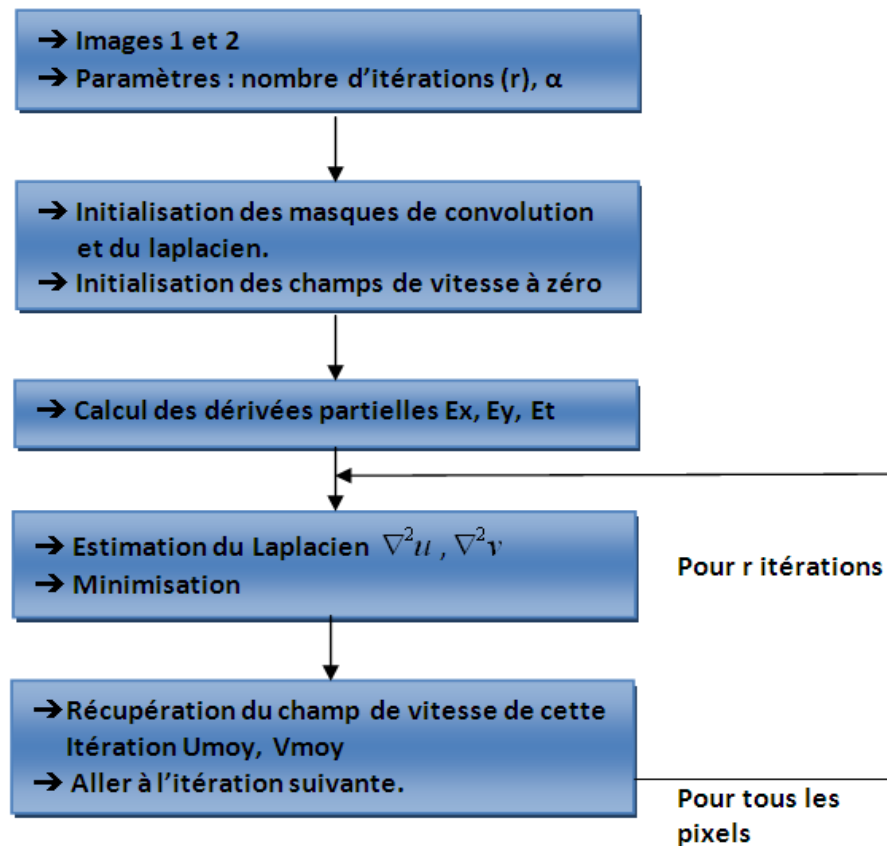


Figure II.1 Les étapes à suivre pour l'estimation du flot optique avec la méthode globale de Horn et Schunck plus solution itérative.

II.4.1 Contraintes

Nous allons tirer une équation qui relie la variation de luminosité de l'image en un point au mouvement du motif contenant la même distribution de luminosité. Notons la distribution de luminosité d'un point (x, y) dans le plan de l'image au temps t , $E(x, y, t)$. Considérons maintenant ce qui se passe lorsque cette structure se déplace.

La luminosité d'un point particulier pris dans le motif est constante, de sorte que :

$$\frac{dE}{dt} = 0 \quad \text{II.1}$$

L'utilisation de la règle de la chaîne de différenciation, nous donne :

$$\frac{\partial E}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial E}{\partial y} \frac{\partial y}{\partial t} + \frac{\partial E}{\partial t} = 0 \quad \text{II.2}$$

En posant :

$$u = \frac{dx}{dt}, \quad v = \frac{dy}{dt} \quad \text{II.3}$$

Il est alors facile de voir que nous avons une seule équation linéaire à deux inconnues **u** et **v**,

$$E_x u + E_y v + E_t = 0 \quad \text{II.4}$$

Où nous avons également introduit les abréviations supplémentaires, E_x , et E_y , pour les dérivées partielles de luminosité de l'image par rapport à x , y et t , respectivement.

La contrainte sur le flot local exprimée par cette équation est illustrée sur la Figure II.2

Écriture de l'équation d'une autre façon,

$$(E_x, E_y) \cdot (u, v) = -E_t \quad \text{II.5}$$

Ainsi, la composante du mouvement dans le sens des gradients (E_x, E_y) est égal à

$$- \frac{E_t}{\sqrt{E_x^2 + E_y^2}} \quad \text{II.6}$$

Nous ne pouvons pas cependant déterminer la composante du mouvement formant un angle droit avec la composante dans le sens des gradients. En conséquence, la vitesse d'écoulement (u, v) ne peut pas être calculée localement sans l'introduction de contraintes supplémentaires.

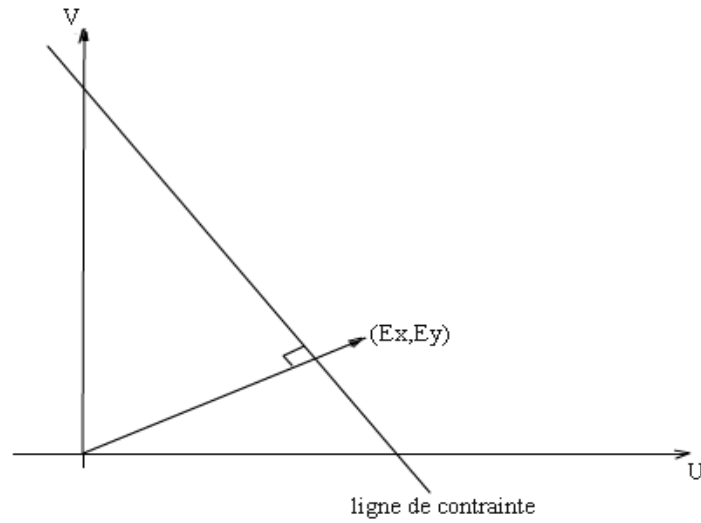


Figure II.2 : La vitesse (u, v) se situe le long d'une ligne perpendiculaire au vecteur gradient de (Ex,Ey). La distance de cette ligne de l'origine est égal à Et, divisé par l'amplitude de (Ex, Ey).

II.4.2 La contrainte de lissage

Si tous les points du diagramme de luminosité peuvent se déplacer de façon autonome, il y a peu d'espoir de récupérer les vitesses. Le plus souvent nous allons voir des objets opaques de taille finie subissant un mouvement rigide ou de déformation. Dans ce cas les points voisins des objets ont des vitesses similaires et le champ de vitesse de luminosité dans l'image varie de façon régulière un peu partout. Des discontinuités dans le flot optique peuvent avoir lieu où un objet obstrue un autre. Un algorithme basé sur une contrainte de lissage est susceptible d'avoir des difficultés surtout pour le calcul du flot sur les bords. Une façon d'exprimer la contrainte supplémentaire est de minimiser le carré de l'intensité du gradient du flux optique:

$$\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2, \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 \quad \text{II.7}$$

Une autre mesure du lissage du flux optique est la somme des carrés des laplaciens des composantes x et y du flot.

Le Laplacien de u et v sont définis comme suit :

$$\nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \text{ et } \nabla^2 v = \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} \quad \text{II.8}$$

Dans des situations simples, les deux laplaciens sont nuls. Si l'observateur se positionne parallèlement à un objet plat, tourne autour d'une ligne perpendiculaire à la surface ou voyage

orthogonalement par rapport à la surface, alors les dérivées secondes partielles de u et v disparaîtront.

Nous utiliserons ici le carré des amplitudes du gradient comme mesure de lissage. Notez que notre approche est en contraste avec celle de Fennema et Thompson, qui propose un algorithme intégrant des hypothèses supplémentaires telles la constance du flot optique dans des régions discrètes de l'image. Leur méthode, basée sur une analyse typologique, ne peut pas traiter avec des objets en rotation, étant donné que c'est donner lieu à un continuum de vitesses d'écoulement.

II.4.3 Estimation des dérivées partielles

Nous devons estimer les dérivés de la luminosité de la série de mesures discrètes disponible de l'image. Il est important que les estimations de E_x, E_y, E_t soient compatibles. Autrement dit, elles devraient se référer toutes au même point de l'image en même temps. Bien qu'il existe de nombreuses formules pour approximer les différenciations, nous allons utiliser un ensemble qui nous donne une estimation de E_x, E_y, E_t , à un point situé au centre d'un cube formé par huit mesures. La relation dans l'espace temps entre ces mesures est représentée sur la Figure II.3. Chacune des estimations est la moyenne des quatre premières différences mesures prises dans le cube parmi les mesures adjacentes.

$$\begin{aligned} E_x &\approx \frac{1}{4} \{ E_{i,j+1,k} - E_{i,j,k} + E_{i+1,j+1,k} - E_{i+1,j,k} \\ &\quad + E_{i,j+1,k+1} - E_{i,j,k+1} + E_{i+1,j+1,k+1} - E_{i+1,j,k+1} \} \\ E_y &\approx \frac{1}{4} \{ E_{i+1,j,k} - E_{i,j,k} + E_{i+1,j+1,k} - E_{i,j+1,k} \\ &\quad + E_{i+1,j,k+1} - E_{i,j,k+1} + E_{i+1,j+1,k+1} - E_{i,j+1,k+1} \} \\ E_t &\approx \frac{1}{4} \{ E_{i,j,k+1} - E_{i,j,k} + E_{i+1,j,k+1} - E_{i+1,j,k} \\ &\quad + E_{i,j+1,k+1} - E_{i,j+1,k} + E_{i+1,j+1,k+1} - E_{i+1,j+1,k} \} \end{aligned}$$

II.9

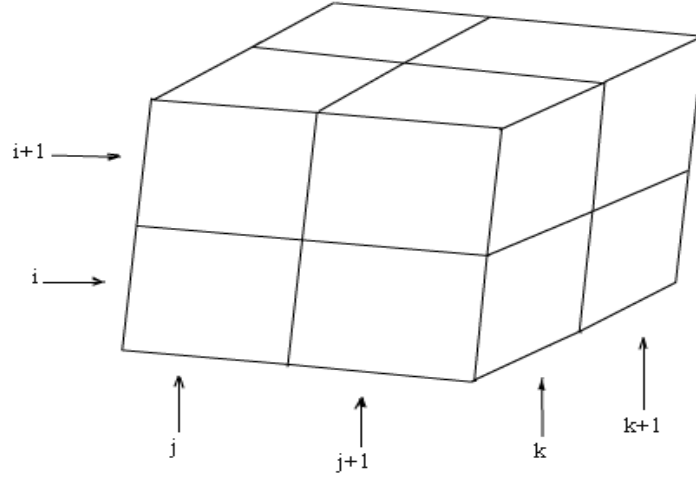


Figure II.3 : illustration des trois dérivées partielles de la luminosité des images à l'aide d'un cube, ici, les j indice de colonne correspondent à la direction x de l'image, l'indice de ligne i à la direction y. alors k est dans le sens du temps.

L'unité de longueur est l'espacement des intervalles dans chaque trame d'images et l'unité de temps est la période d'échantillonnage de l'image. Nous évitons les formules d'estimation avec un large support, car il s'agit généralement de petites formules pour soutenir la recherche appliquée à des images lissées.

II.4.4 Estimation du laplacien de la vitesse d'écoulement

Nous avons également besoin d'estimer les laplaciens de u et de v. Une méthode d'approcher ces valeurs prend la forme suivante :

$$\nabla^2 u \approx K (\bar{u}_{i,j,k} - u_{i,j,k}) \quad \text{et} \quad \nabla^2 v \approx K (\bar{v}_{i,j,k} - v_{i,j,k}) \quad \text{II.10}$$

Où les moyennes locales $\bar{u}$ et $\bar{v}$ sont définies comme suit :

$$\begin{aligned} \bar{u}_{i,j,k} &= \frac{1}{6} \{ u_{i-1,j,k} + u_{i,j+1,k} + u_{i+1,j,k} + u_{i,j-1,k} \} \\ &\quad + \frac{1}{12} \{ u_{i-1,j-1,k} + u_{i-1,j+1,k} + u_{i+1,j+1,k} + u_{i+1,j-1,k} \} \\ \bar{v}_{i,j,k} &= \frac{1}{6} \{ v_{i-1,j,k} + v_{i,j+1,k} + v_{i+1,j,k} + v_{i,j-1,k} \} \\ &\quad + \frac{1}{12} \{ v_{i-1,j-1,k} + v_{i-1,j+1,k} + v_{i+1,j+1,k} + v_{i+1,j-1,k} \} \end{aligned} \quad \text{II.11}$$

Le facteur de proportionnalité K est égal à 3 si la moyenne est calculée comme indiqué et nous supposons de nouveau que l'unité de longueur est égal à l'espacement de la grille d'intervalle.

La figure suivante illustre l'affectation des poids aux points voisins.

1/12	1/6	1/12
1/6	-1	1/6
1/12	1/6	1/12

Figure II.4 poids des points voisins.

II.4.5 Minimisation

Le problème alors, est de réduire au minimum la somme des erreurs dans l'équation du taux de variation de la luminosité de l'image

$$e_b = E_x u + E_y v + E_t \quad \text{II.12}$$

Et la mesure de l'erreur dans la vitesse d'écoulement pixel à pixel.

$$e_c^2 = \left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 + \left(\frac{\partial v}{\partial x} \right)^2 + \left(\frac{\partial v}{\partial y} \right)^2 \quad \text{II.13}$$

Quel devrait être le poids relatif de ces deux facteurs? Dans la pratique, les mesures de luminosité de l'image sont quantifiées de sorte que nous ne pouvons pas attendre que e_b soit nulle. Cette quantité a tendance à avoir une grandeur d'erreur qui est proportionnelle au bruit dans la mesure. Ce qui nous pousse à choisir un facteur de pondération approprié, notée α^2 , comme on le verra plus loin.

Soit l'erreur totale à minimiser :

$$e^2 = \iint (\alpha^2 e_c^2 + e_b^2) dx dy \quad \text{II.14}$$

La minimisation consiste à trouver des valeurs convenables pour la vitesse de flot optique (u, v).

En utilisant le calcul des variations, nous obtenons :

$$\begin{aligned} E_x^2 u + E_x E_y v &= \alpha^2 \nabla^2 u - E_x E_t \\ E_x E_y u + E_y^2 v &= \alpha^2 \nabla^2 v - E_y E_t \end{aligned} \quad \text{II.15}$$

En utilisant l'approximation du laplacien introduite dans la section précédente :

$$\begin{aligned} (\alpha^2 + E_x^2) u + E_x E_y v &= (\alpha^2 \bar{u} - E_x E_t) \\ E_x E_y u + (\alpha^2 + E_y^2) v &= (\alpha^2 \bar{v} - E_y E_t) \end{aligned} \quad \text{II.16}$$

Le déterminant du coefficient de la matrice est égal à $\alpha^2 (\alpha^2 + E_x^2 + E_y^2)$. La résolution pour u et v nous donne :

$$\begin{aligned} (\alpha^2 + E_x^2 + E_y^2) u &= (\alpha^2 + E_y^2) \bar{u} - E_x E_y \bar{v} - E_x E_t \\ (\alpha^2 + E_x^2 + E_y^2) v &= -E_x E_y \bar{u} + (\alpha^2 + E_x^2) \bar{v} - E_y E_t \end{aligned} \quad \text{II.17}$$

II.4.6 Contrainte de minimisation

Lorsque nous faisons tendre α^2 vers zéro, nous obtenons la solution de la contrainte de minimisation de notre problème. En appliquant la méthode des multiplicateurs de Lagrange au problème de minimisation de e_c^2 tout en maintenant $e_b = 0$ nous obtenons:

$$E_y \nabla^2 u = E_x \nabla^2 v \quad E_x u + E_y v + E_t = 0 \quad \text{II.18}$$

L'approximation du laplacien par la différence de la vitesse en un point et la moyenne de ses voisins nous donne alors :

$$\begin{aligned} (E_x^2 + E_y^2)(u - \bar{u}) &= -E_x [E_x \bar{u} + E_y \bar{v} + E_t] \\ (E_x^2 + E_y^2)(v - \bar{v}) &= -E_y [E_x \bar{u} + E_y \bar{v} + E_t] \end{aligned} \quad \text{II.19}$$

Se référant à nouveau à la Figure II.5, nous constatons que le point calculé ici est à l'intersection de la ligne de contrainte et de la ligne à angle droit par le point $(\bar{u}, \bar{v})$. Nous n'utiliserons pas ces équations, puisque nous nous attendons à avoir des erreurs dans l'estimation des dérivées partielles.

II.4.7 Solution itérative

Nous avons maintenant une paire d'équations pour chaque point de l'image. Il serait très coûteux en termes de ressources de résoudre ces équations en même temps par l'une des méthodes standard, comme celle de Gauss-Jordan. La matrice correspondante est très grande, puisque le nombre de lignes et de colonnes est égal à deux fois le nombre des cellules de l'image. Les méthodes itératives, telles que celles de Gauss-Seidel s'imposent d'elles même. On peut calculer un nouvel ensemble de champ de vitesses (u^{n+1}, v^{n+1}) estimé à partir des dérivés et la moyenne des estimations des champs précédents (u^n, v^n) de la manière suivante :

$$\begin{aligned} u^{n+1} &= \bar{u}^n - E_x \left[E_x \bar{u}^n + E_y \bar{v}^n + E_t \right] / (\alpha^2 + E_x^2 + E_y^2) \\ v^{n+1} &= \bar{v}^n - E_y \left[E_x \bar{u}^n + E_y \bar{v}^n + E_t \right] / (\alpha^2 + E_x^2 + E_y^2) \end{aligned} \quad \text{II.20}$$

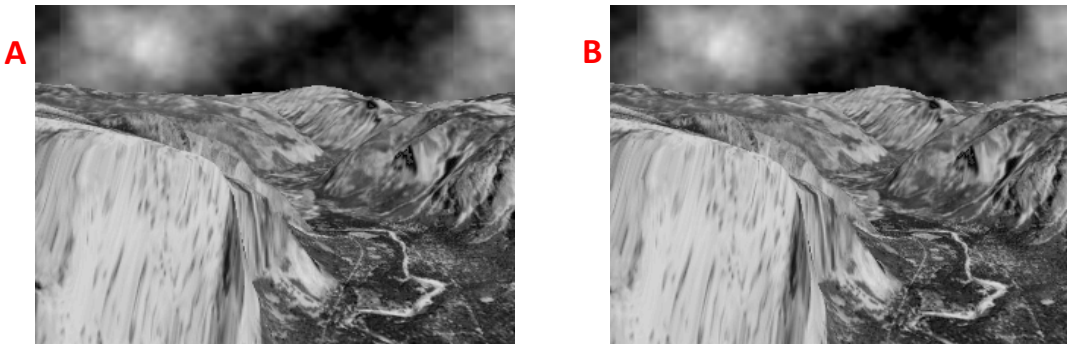
Remarque:

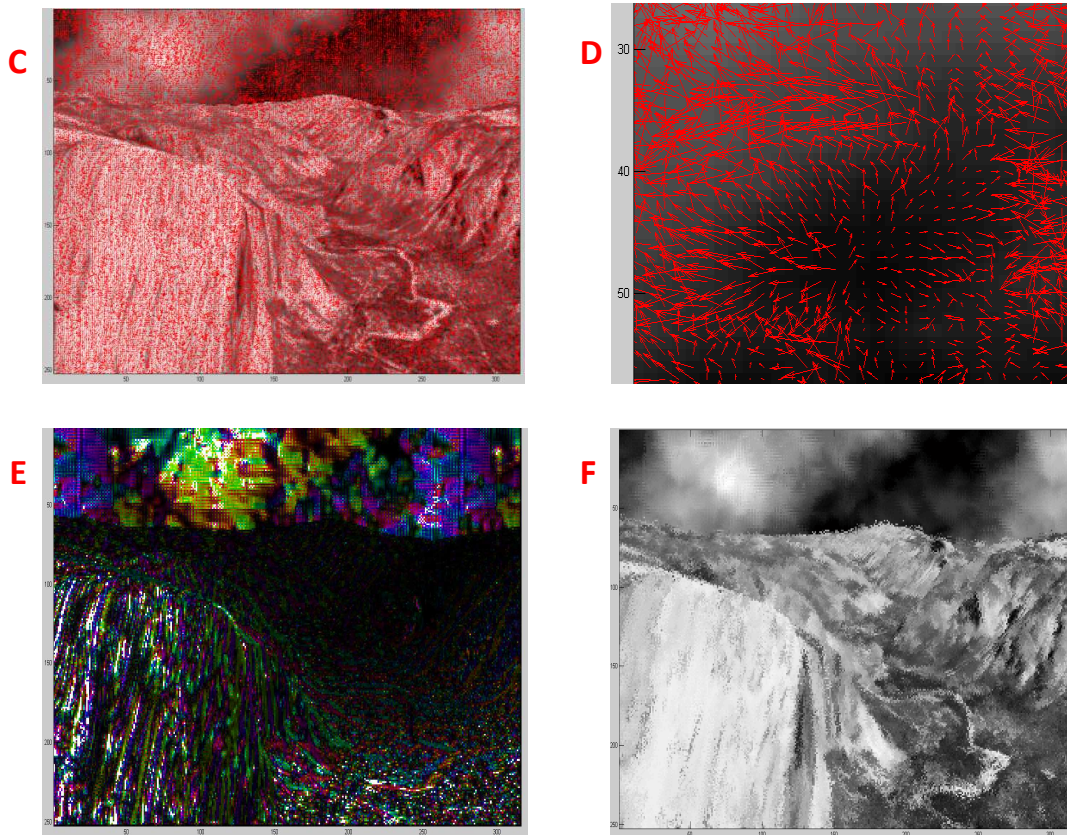
Il est intéressant de noter que les nouvelles valeurs estimées du flot à un point donné ne dépendent pas directement des estimations précédentes au même point. Les conditions aux limites naturelles pour le problème variationnel se révèlent être une dérivé de zéro normal. Au bord de l'image, quelques-uns des points nécessaires pour calculer la moyenne locale de la vitesse se trouvent en dehors de l'image. Ici nous avons simplement copié les vitesses des points adjacents.

α est le coefficient de régularisation qu'il faut régler, de même que le nombre d'itérations de la méthode de résolution.

II.5 Résultats pratiques

Les meilleurs résultats sont obtenus pour (alpha=0.1, N° d'itérations= 11)



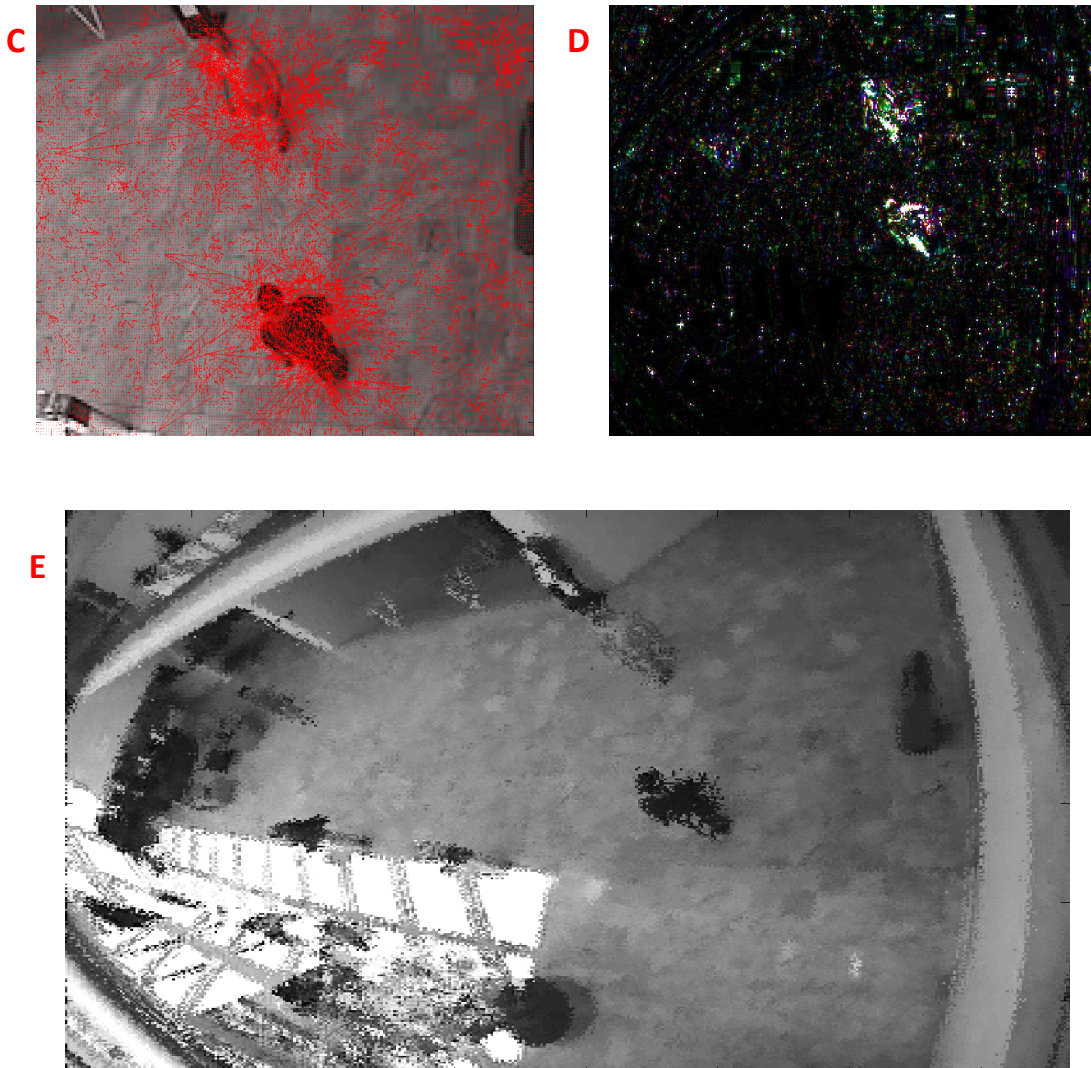


Figures II.5 : **A**-image test YOS9, **B**-image test YOS10, **C-D** tracé des vecteurs mouvements en utilisant la méthode de HS **E**-représentation des vecteurs vitesses par la carte des couleurs. **F**-reconstruction de la seconde image test grâce aux vecteurs mouvements

Nous remarquons que nous obtenons une bonne estimation du mouvement dans l'ensemble, avec un léger flou dans les zones de discontinuités, et un flot optique dense.

Le même test pour une autre séquence:





Figures II.6 : **A-B**-images tests, **C**-tracé des vecteurs mouvements en utilisant la méthode de HS
D-représentation des vecteurs vitesses par la carte des couleurs. **E**-reconstruction de la seconde image test grâce aux vecteurs mouvements

Notons que pour ce deuxième test, on voit que la méthode de Horn et Shunck est très sensible au bruit, le caractère globale de la méthode lui confère un caractère très lisse, du coup les résultats sont imprécis, néanmoins l'estimation demeure acceptable.

La reconstruction de l'image est un moyen qui nous permet de savoir visuellement, si l'estimation est bonne ou pas, mais techniquement, il n'y a que les vecteurs vitesses et la carte des couleurs qui sont les indicateurs d'une bonne ou mauvaise estimation du mouvement

II.6 Méthode de Lucas et Kanade

Comme pour la méthode de Horn et Shunck, On fait toujours l'hypothèse de conservation de la luminance entre deux images successives, à laquelle on ajoute une hypothèse (forte mais locale) de constance du flot sur un voisinage Ω centré sur le pixel dont on veut calculer le déplacement.

II.6.1 Formulation mathématique

La fonctionnelle à minimiser sur tout le voisinage Ω du pixel prend la forme suivante :

$$E_{LK} = \sum_{\Omega} \left[\frac{\partial E(x, y, t)}{\partial x} u(t) + \frac{\partial E(x, y, t)}{\partial y} v(t) + \frac{\partial E(x, y, t)}{\partial t} \right]^2 \quad \text{II.21}$$

La solution la plus simple (maximum de vraisemblance) consiste à réaliser une estimation au sens des moindres carrés, c'est à dire à résoudre de la façon suivante sur le voisinage Ω :

$$\text{Soient } A = \begin{bmatrix} E_{x1} & E_{y1} \\ E_{x2} & E_{y2} \\ \dots & \dots \\ E_{xn} & E_{yn} \end{bmatrix}, \quad b = \begin{bmatrix} E_{t1} \\ E_{t2} \\ \dots \\ E_{tn} \end{bmatrix} \quad \text{II.22}$$

Où $x_1, x_2, \dots, x_n$ sont les pixels dans la fenêtre, Ω et $E_{x1}, E_{x2}, \dots, E_{xn}$ sont les dérivées partielles de l'image I .

$$\text{Alors } \begin{bmatrix} u \\ v \end{bmatrix} = (A^T A)^{-1} A^T b \quad \text{II.23}$$

Où A^T est la transposée de la matrice A .

$$\text{Ce qui est également équivalent à : } G = \begin{bmatrix} \sum_{\Omega} E_x^2 & \sum_{\Omega} E_x E_y \\ \sum_{\Omega} E_x E_y & \sum_{\Omega} E_y^2 \end{bmatrix}, \quad \eta = \begin{bmatrix} \sum_{\Omega} E_t E_x \\ \sum_{\Omega} E_t E_y \end{bmatrix} \quad \text{II.24}$$

$$\text{Alors } \begin{bmatrix} u \\ v \end{bmatrix} = G^{-1} \cdot \eta \quad \text{II.25}$$

Il peut arriver que la matrice $A^T A$ (ou G , de manière équivalente) soit mal conditionnée (déterminant quasi nul). Cela provient du fait que le pixel considéré se trouve dans une zone où l'intensité lumineuse est constante sur le domaine, d'où un gradient nul (problème de l'ouverture). Le calcul par les moindres carrés risque alors de fournir une estimation aberrante.

Par ailleurs, l'estimateur des moindres carrés a la particularité d'être très sensible au bruit et aux erreurs de mesure.

On peut remédier à ces problèmes en utilisant une technique de régularisation, comme la régularisation quadratique. La nouvelle fonctionnelle s'écrit :

$$E_{LK2} = \sum_{\Omega} \left[\nabla E \cdot \vec{w} + E_t \right]^2 + \alpha |w|^2 \quad \text{II.26}$$

Cet estimateur est toujours linéaire, et s'exprime tout calcul fait de la manière suivante:

$$\begin{bmatrix} u \\ v \end{bmatrix} = (A^T A + \alpha E)^{-1} A^T b \quad \text{II.27}$$

Avec α réglable, représentant la régularité de la solution.

II.6.2 Améliorations apportées à cette méthode

II.6.2.1 Implémentation pyramidale

On définit une hauteur L_m de pyramide (en pratique, $L_m = 2, 3$ ou 4). A chaque niveau de la pyramide, on sous-échantillonne l'image d'un facteur 2 pour les deux images successives considérées, le niveau 0 correspondant à l'image initiale, le niveau L_m correspondant au niveau le plus grossier [4].

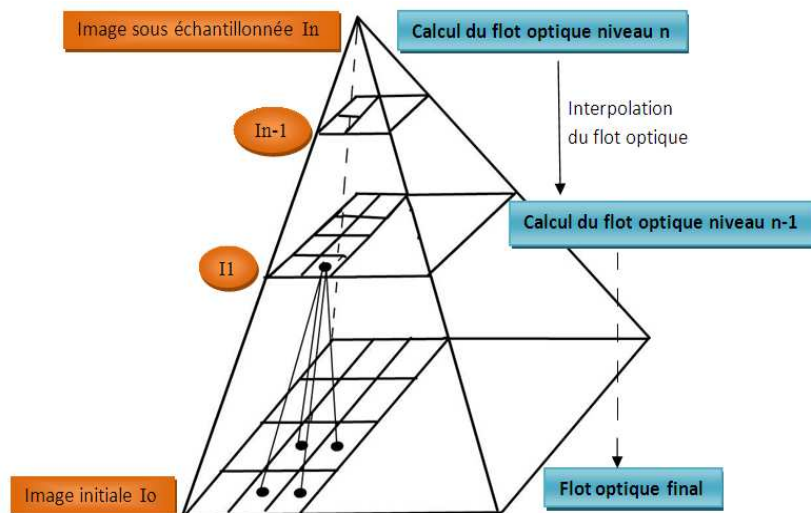


Figure II.7 Implémentation pyramidale d'une méthode de calcul du flot optique.

Au niveau L_m , on calcule le flot optique, puis on le propage au niveau inférieur en translatant l'image f_1 avec l'a priori calculé au niveau supérieur, avant d'exécuter l'algorithme à nouveau et

ainsi de suite jusqu'au niveau 0 qui correspond à l'image initiale. On récupère alors le flot optique final.

Pyramide gaussienne: le niveau zéro est l'image originale. Pour passer au niveau supérieur, on effectue une moyenne pondérée (masque gaussien) des pixels dans une fenêtre 3x3 (ou 5x5) qui devient la valeur du pixel correspondant à la fenêtre au niveau supérieur (filtrage passe-bas), suivi d'un sous-échantillonnage de facteur 2 et ainsi de suite.

$$g_0 = I, g_k = \text{reduire}(g_{k-1}) \text{ pour } k \geq 1$$

II.6.2.1 Raffinement itératif

Cela consiste à minimiser l'écart entre les deux images successives en exécutant à nouveau l'algorithme après avoir déplacé une des deux images selon le dernier champ de vitesse calculé. Si l'algorithme est stable, la méthode converge.

II.6.3 Description de l'implémentation

Les étapes à suivre pour résoudre cet algorithme sont décrites dans la figure suivante.

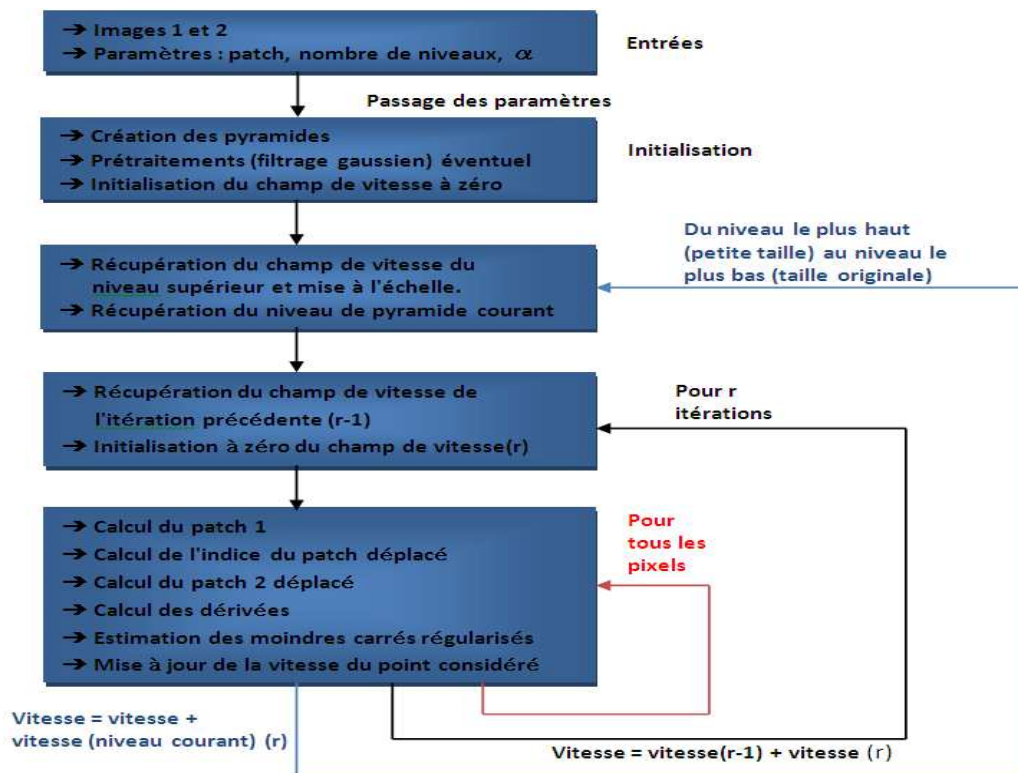


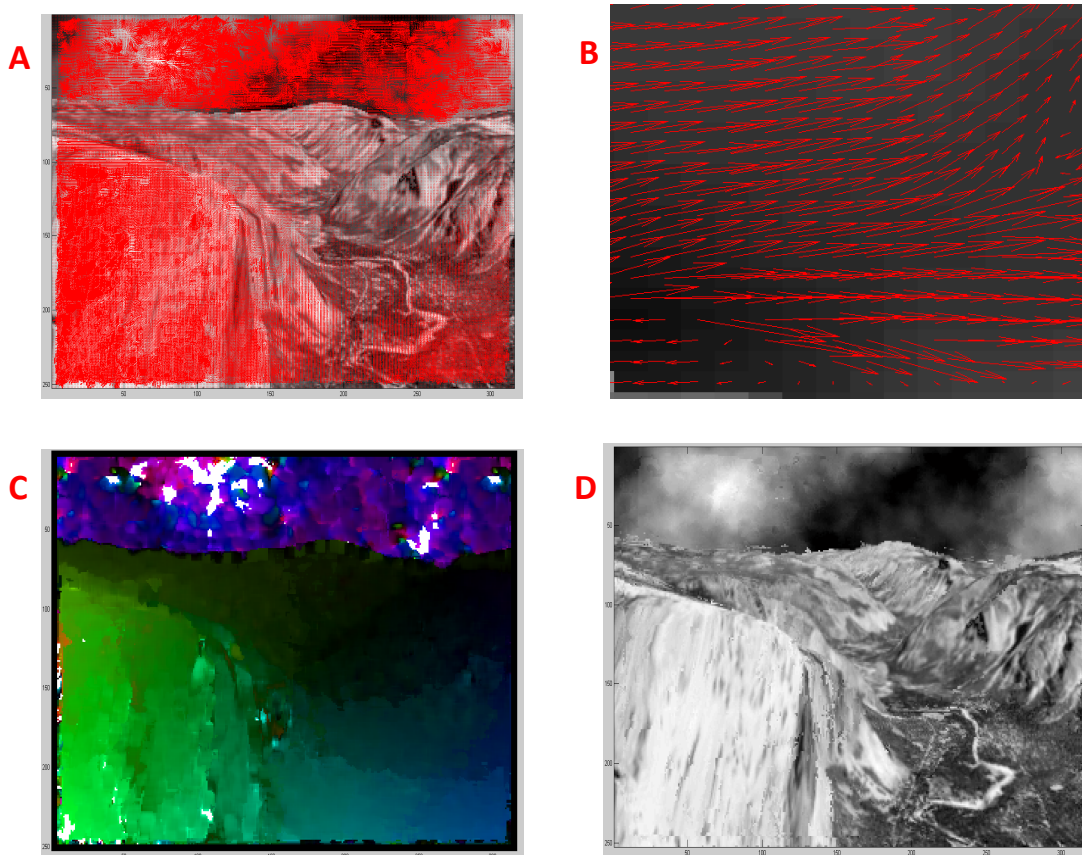
Figure II.8 Les étapes suivies pour l'estimation du flot optique avec l'implémentation pyramidale à la méthode locale de Lucas Kanade avec raffinement itératif.

A partir de deux images successives, il s'agit tout d'abord de créer les pyramides gaussiennes (sous échantillonnage) pour chaque image. L'algorithme de calcul en lui même commence alors avec l'interpolation du champ de vitesse du niveau supérieur, puis le calcul pour chaque pixel.

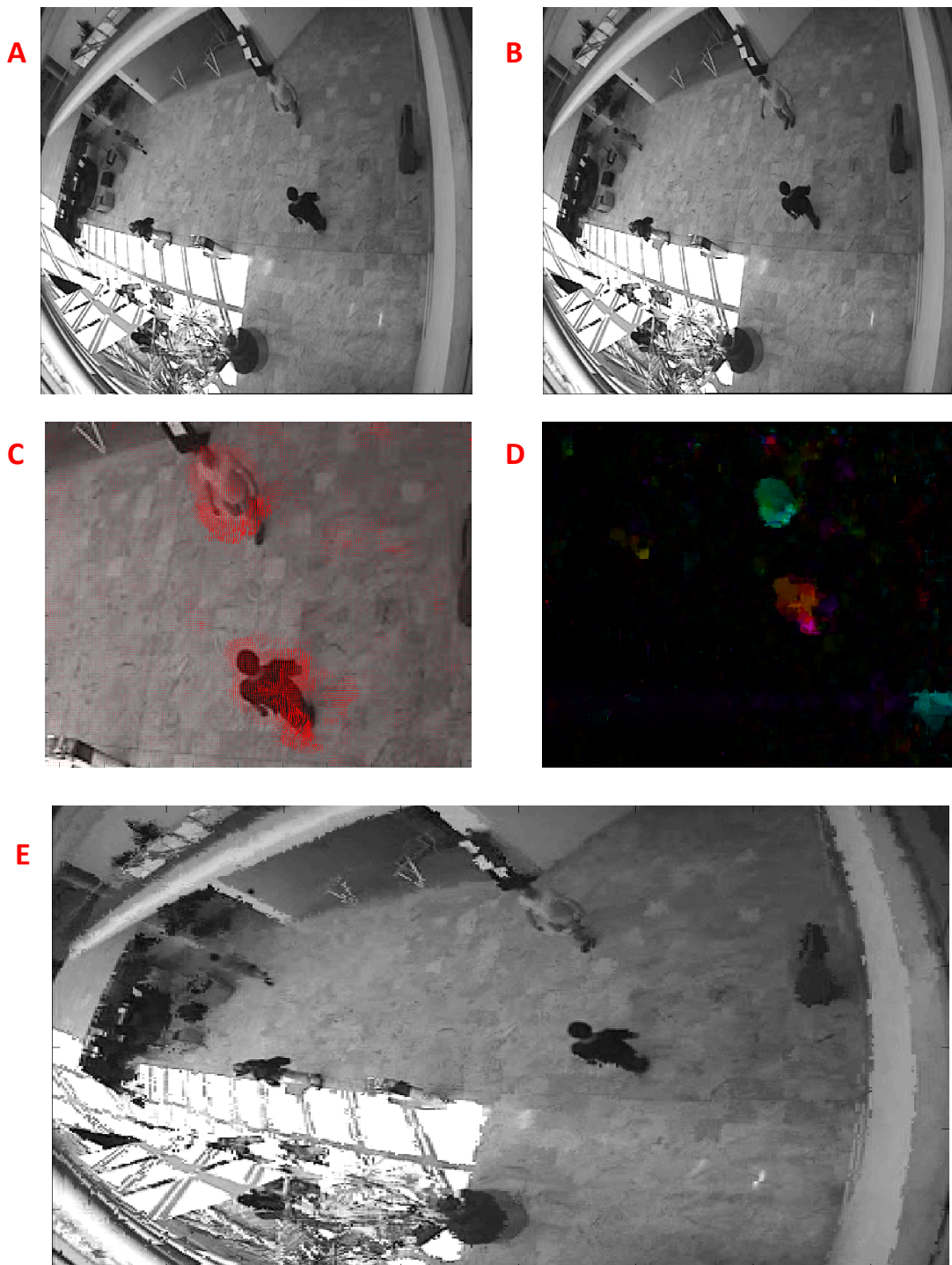
II.6.4 Résultats pratiques

Nous effectuons les mêmes séries de test sur les séquences d'images précédentes, afin de pouvoir juger la différence dans l'estimation du mouvement entre les deux méthodes.

Les meilleurs résultats sont obtenus pour ($\alpha=0.1$, N° d'itérations= 10, niveau pyramide=3)



Figures I.4 : **A-B** tracé des vecteurs mouvements en utilisant la méthode de LK, **C**-représentation des vecteurs vitesses par la carte des couleurs. **D**-reconstruction de la seconde image test grâce aux vecteurs mouvements



Figures II.12 : **A-B** images tests **C**- tracé des vecteurs mouvements en utilisant la méthode de LK+implémentation pyramidale, **D**- représentation des vecteurs mouvement grace à la carte des couleurs **E**-la reconstruction de la seconde image test grâce aux vecteurs mouvements

On remarque que les résultats obtenus avec la méthode LK sont très satisfaisants, les contours des objets sont très bien respectés, avec une bonne segmentation et enfin, le flot est dense et le plus exact possible.

II.7 Discussion

De même que pour la méthode d'Horn & Schunck, il est possible de réaliser une implémentation pyramidale et d'utiliser le raffinement itératif à chaque niveau. Les paramètres à régler sont alors la taille du patch, le nombre de niveaux de pyramide et le nombre d'itérations

Quelle que soit la méthode employée, les bords ainsi que le bas de l'image (« essuie-glace ») comporteront toujours des valeurs aberrantes car il s'agit de points de l'image qui disparaissent au cours du mouvement, il est donc impossible de deviner leur déplacement.

Chapitre III

Méthode Probabiliste

III.1 Préambule

Toutes les méthodes présentées dans le deuxième chapitre, sont des approches déterministes d'estimation du mouvement 2D, elles minimisent l'erreur dans l'équation du flot optique. **L'approche probabiliste** que nous allons détailler dans ce chapitre est une **approche bayésienne** qui s'appuie sur la nature stochastique du contenu de l'image.

Afin d'estimer le mouvement dans la séquence d'images, la méthode bayésienne se base sur deux modèles permettant la résolution de la probabilité du MAP (Maximum A Posteriori), le modèle de vraisemblance, et le modèle à priori.

L'optimisation de la fonction cout de l'estimation du mouvement par cette approche probabiliste est atteinte en faisant appel à différentes techniques, parmi ces techniques l'algorithme du recuit simulé, qui est un algorithme de relaxation stochastique qui effectue une partie aléatoire ou à chaque itération une solution est exposée, puis une perturbation aléatoire est appliquée donnant naissance à une nouvelle solution candidate, qui est acceptée avec une certaine probabilité.

III.2 Formulation du problème du MAP

La formulation du problème de la probabilité du **Maximum A Posteriori** [5] est défini comme suit : ayant pour images $f(x, y, t_1)$ et $f(x, y, t_2)$, trouvant $u(x, y, t_1)$ et $v(x, y, t_1)$ les vecteurs mouvements, tels que :

$$(u(x, y, t_1), v(x, y, t_2)) = \arg \max_{u, v} p(u(x, y, t_1), v(x, y, t_2) | f(x, y, t_1), f(x, y, t_2)) \quad \text{III.1}$$

Ou : $p(u(x, y, t_1), v(x, y, t_2) | f(x, y, t_1), f(x, y, t_2))$ représente la fonction de densité de probabilité du champ de mouvement donnant les deux images.

En appliquant le **théorème de Bayes** plusieurs fois, nous obtenons

$$\begin{aligned} p(u(x, y, t_1), v(x, y, t_2) | f(x, y, t_1), f(x, y, t_2)) &= \frac{p(u(x, y, t_1), v(x, y, t_2), f(x, y, t_1), f(x, y, t_2))}{p(f(x, y, t_2) | f(x, y, t_1))} \\ &= \frac{p(u(x, y, t_1), v(x, y, t_2), f(x, y, t_2) | f(x, y, t_1)) \cdot p(f(x, y, t_1))}{p(f(x, y, t_2) | f(x, y, t_1)) \cdot p(f(x, y, t_1))} \\ &= \frac{p(f(x, y, t_2) | u(x, y, t_1), v(x, y, t_2), f(x, y, t_1)) \cdot p(u(x, y, t_1), v(x, y, t_2) | f(x, y, t_1))}{p(f(x, y, t_2) | f(x, y, t_1))} \end{aligned} \quad \text{III.2}$$

Le dénominateur n'ayant aucun rapport avec $u(x, y, t_1)$ et $v(x, y, t_1)$, il n'intervient donc pas dans l'estimation du mouvement.

Le premier facteur du numérateur est la fonction du modèle de vraisemblance qui représente la probabilité d'avoir l'image courante, connaissant l'image précédente et en appliquant les champs de mouvement $u(x, y, t_1)$ et $v(x, y, t_1)$ à cette dernière.

Notons que si $f(x, y, t)$ est l'image réelle et $f(x, y, t)$ l'image observée qui est la résultante de l'image $f(x, y, t)$ et d'un bruit de distorsion, alors :

$f(x, y, t) = f(x, y, t) + n(x, y, t)$ avec $n(x, y, t)$ le bruit observé (blanc, gaussien, ayant pour variance σ^2 et une moyenne $m = 0$).

En ignorant les régions d'occlusion de l'image ainsi que les changements au niveau de la luminance de la scène, nous avons :

$$f(x, y, t_2) = f(x - u(x, y, t_1), y - v(x, y, t_1), t_1) \quad \text{III. 3}$$

Cependant, du au bruit observé, l'équation précédente devient erronée, et ceci nous conduit à formuler la fonction de vraisemblance postulant que la fonction de densité $p(f(x, y, t_2) | u(x, y, t_1), v(x, y, t_2), f(x, y, t_1))$ peut s'exprimer de la façon suivante en tenant compte du bruit observé.

$$p(f(x, y, t_2) | u(x, y, t_1), v(x, y, t_2), f(x, y, t_1)) = (2\pi\sigma^2)^{\frac{1}{N}} \exp \left(- \sum_{(x,y) \in I} \frac{[f(x, y, t_2) - f(x - u(x, y, t_1), y - v(x, y, t_1), t_1)]^2}{2\sigma^2} \right)$$

N: nombre total de pixel de l'image I III. 4

Le second facteur de l'équation III.2, représente le modèle à priori, la probabilité d'avoir $u(x, y, t_1)$ et $v(x, y, t_1)$ comme champ de mouvement horizontal et vertical connaissant l'image $f(x, y, t_1)$.

Cette fonction du modèle à priori est assimilé à un **champ de markov** (Annexe.B) ayant une **distribution Gibbsienne** donnant ainsi :

$$p(u(x, y, t_1), v(x, y, t_2) | f(x, y, t_1)) = \frac{1}{Z} \exp \left[-U_p(u(x, y, t_1), v(x, y, t_1) | f(x, y, t_1)) \right] \quad \text{III. 5}$$

$$Z : \text{fonction de partition et } U_p(u(x,y,t_1), v(x,y,t_1) | f(x,y,t_1)) = \lambda \sum_{c \in C} V_c^p(u(x,y,t_1), v(x,y,t_1) | f(x,y,t_1))$$

III. 6

C : représente l'ensemble des cliques du système de voisinage

V_c^p : Le potentiel de la clique $c \in C$, et λ une constant positive.

En remplaçant III. 4, III. 5 dans III. 2, on obtient la formule basique de l'estimation de la probabilité du Maximum A Posteriori (MAP).

$$p(u(x,y,t_1), v(x,y,t_2) | f(x,y,t_1), f(x,y,t_2)) = \frac{(2\pi\sigma^2)^{\frac{1}{N}} \exp \left(- \sum_{(x,y) \in I} \frac{[f(x,y,t_2) - f(x-u(x,y,t_1), y-v(x,y,t_1), t_1)]^2}{2\sigma^2} \right) \frac{1}{Z} \exp[-U_p(u(x,y,t_1), v(x,y,t_1) | f(x,y,t_1))]}{p(f(x,y,t_2) | f(x,y,t_1))}$$

III. 7

III.3 Formulation mathématique.

La technique retenue dans cette étude pour l'estimation du flot optique dans une séquence d'images vidéo, est basée sur l'approche de base de l'estimation de la probabilité du Maximum A Posteriori, elle utilise une méthode dite « **local outlier rejection** » rejet de valeurs aberrantes proposée par Lu [6] permettant de résoudre le problème des discontinuités du mouvement aux frontières des objets. Tout d'abord, nous discuterons du cadre dans lequel les modèles de base pour l'estimation de la probabilité du Maximum A Posteriori sont développés et ensuite nous présenterons les deux modèles et nous étendrons l'application à l'estimation de mouvement, basée sur les différences dans les séquences d'images vidéo.

L'amélioration essentielle suggérée par cette méthode sur celle proposée par Lu [6] est d'employer une élimination des valeurs aberrantes, en classant par ordre croissant les valeurs afin de calculer la contrainte de lissage, tandis que Lu a utilisé un seuil à partir duquel les données seront rejetées.

La contribution originale de cette étude comprend une technique d'estimation du mouvement basée sur les différences d'images vidéo successives au lieu des images vidéo elles mêmes. Nous avons choisi d'utiliser la différence entre les images successives, car cela réduit la gamme dynamique du niveau de l'intensité des pixels des images et aussi parce que cela élimine les redondances spatiales et temporelles.

Tout d'abord, nous discutons du cadre dans lequel les modèles d'estimation de la probabilité du maximum a posteriori sont développés, ensuite nous présentons les deux modèles et essayons d'étendre leur application à l'estimation de mouvement basée sur les différences entre les images successives d'une vidéo.

Soit $I = \{(x, y) | 0 \leq x < H, 0 \leq y < W\}$ un site de l'image 2D avec un système de voisinage (exemple système à 8 voisins V_8 Figure III.1), alors, un champ aléatoire $q = \{q(x, y) | (x, y) \in I\}$ est dit champ aléatoire de Markov si et seulement si :

1. Toutes les réalisations ont des probabilités non nulles tels que : $p(q) > 0 \forall q$
2. La probabilité conditionnelle est Markovienne

$$p(q(x, y) | q(x', y'), (x', y') \in I \setminus \{(x, y)\}) = p(q(x, y) | q(x', y'), (x', y') \in V_8 \setminus \{(x, y)\})$$

III. 8

La seconde condition stipule que la fonction de densité des probabilités conditionnelles locales en un site particulier (x, y) ne dépende que de la valeur du champ aléatoire de ses voisins V_8 .

Le **théorème de Hammersley-Clifford**, publié par Besag [8], stipule que si q est un Champ de Markov, alors, sa distribution conjointe est Gibbsienne, c'est à dire,

$$p(q) = \frac{1}{Z} \exp(-U(q)) \quad \text{III. 9}$$

Où U est appelée la fonction de l'énergie et Z facteur de normalisation appelé la fonction partition et est donné par: $Z = \sum_q \exp(-U(q))$ III. 10

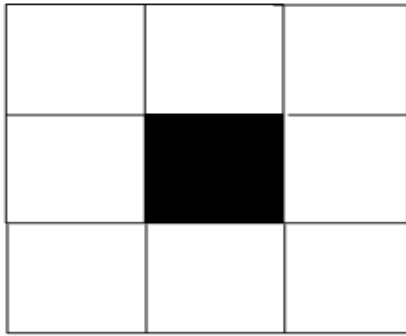


Figure III.1 : Le système à huit voisins V_8

Pour être en mesure de déterminer la fonction d'énergie de toute l'image, nous devons définir une clique. Une clique c , définie sur l'image I pour le voisinage du système V_8 est une partie de I , de telle sorte que c soit constituée d'un ou plusieurs sites tous voisins les uns aux autres dans le système de voisinage représenté dans la figure III.1. Ainsi la fonction d'énergie (potentiel de Gibbs) peut s'exprimer de la façon suivante: $U(q) = \sum_{c \in C} V^c(q(x, y) | (x, y) \in c)$ III. 11

Où C est l'ensemble des cliques de l'image et V^c est appelé le potentiel clique et ne dépend que des valeurs du champ aléatoire tel que $(x, y) \in c$. A noter que la distribution de Gibbs est une distribution exponentielle, qui inclue un cas dit Gaussien.

Un vecteur de mouvement pour un champ aléatoire Gibbsien pour une clique de deux pixels, est défini comme suit :

$$V^c(u(x, y), v(x, y), u(x', y'), v(x', y')) = [u(x, y) - u(x', y')]^2 + [v(x, y) - v(x', y')]^2 \quad \text{III. 12}$$

Ou $u(x,y)$ et $v(x,y)$ sont les déplacements horizontal et vertical du pixel $(x,y) \in I$ et $u(x',y')$, $v(x',y')$ sont les déplacements du pixel au site $(x',y') \in I$ et $(x,y), (x',y')$ sont voisins dans V_g . Evidemment, plus la distance entre deux vecteurs de mouvement voisins est grande, plus la probabilité d'avoir cette configuration est petite. Comme la somme des potentiels cliques passe par toutes les cliques de l'image, les configurations dans lesquelles les vecteurs de mouvement varient considérablement d'un pixel à un autre, se verront affectés des probabilités petites.

Un autre choix approprié pour les potentiels cliques est défini comme suit:

1. Pour une clique singleton : $V^c(x,y) = \alpha$
2. Pour une clique double : $V^c((x,y), (x',y')) = \begin{cases} -\beta & \text{les deux pixels ont la mm intensité} \\ \beta & \text{sinon} \end{cases}$
3. Pour les cliques à trois ou quatre pixels, le même principe pour les cliques doubles leur est appliqué avec $\pm\gamma$ et $\pm\delta$. Ou $\alpha, \beta, \gamma, \delta$ choisis expérimentalement.

Comme le nombre de chiffres avec virgule croît avec le carré du nombre de cliques, nous avons décidé de ne considérer que les cliques double, comme dans l'équation III. 12.

Cependant, si la vitesse de calcul est négligée, le deuxième choix des potentiels cliques peut fournir une contrainte de lissage plus robuste car elle emploie plus de cliques, aussi, susceptible de prédire une bonne estimation des vecteurs de mouvement parce qu'ils prennent en compte un plus grand nombre de pixels afin de prendre la décision.

III.4 Le modèle de vraisemblance

Soit f une image observée à travers un appareil photo idéal, et soit f l'image observée acquise à partir d'un appareil normal correspondant à f mis à part quelques transformations dans le processus d'acquisition telles que le filtrage, la correction gamma, la quantification et la distorsion due au bruit. L'image observée f est échantillonnée dans un réseau rectangulaire

$$T.Q : I = \{(x,y) | 0 \leq x < H, 0 \leq y < W\}$$

Avec l'échantillonnage vertical, horizontal et temporel (T_v, T_h, T_t) , le nombre total d'éléments de l'image (pixels) est donné par : $N = H * W$ III. 13

Soit $f(x,y,t_2)$ la fonction d'intensité lumineuse de l'image courante, et $f(x,y,t_1)$ celle de l'image précédente, ou $t_2 \in \mathfrak{R}$ et $t_1 = t_2 - T_t$.

Sans perdre de vue les généralités, nous nous sommes intéressés au problème d'estimation des vecteurs mouvements à partir des images $f(x, y, t_1)$ et $f(x, y, t_2)$. Si $u(x, y, t_1)$ et $v(x, y, t_1)$ indiquent le vecteur déplacement estimé entre les deux images au niveau du pixel $(x, y) \in I$ puis le long de la vraie trajectoire du mouvement d'un pixel, on peut écrire (en ignorant les variations d'intensité): $f(x, y, t_2) = f(x - u(x, y, t_1), y - v(x, y, t_1), t_1)$ III. 14

Selon la fonction III. 4, la probabilité d'obtenir $f(x, y, t_2)$ à partir de $f(x, y, t_1)$ étant donné les déplacements $u(x, y, t_1)$ et $v(x, y, t_1)$ dépend de la différence entre les images déplacées tel que:

$$p(f(x, y, t_2) | u(x, y, t_1), v(x, y, t_1), f(x, y, t_1)) = (2\pi\sigma^2)^{\frac{1}{2N}} \exp \left\{ - \sum_{(x,y) \in I} \frac{[f(x, y, t_2) - f(x - u(x, y, t_1), y - v(x, y, t_1), t_1)]^2}{2\sigma^2} \right\}$$

III. 15

Ou, σ^2 représente la variance du bruit Gaussien observé, en exprimant, cette fonction de densité de probabilité telle une fonction de potentiel on peut l'écrire sous la forme suivante :

$$U(f(x, y, t_2) | u(x, y, t_1), v(x, y, t_1), f(x, y, t_1)) = \frac{\log(2\pi\sigma^2)}{2N} + \frac{1}{2\sigma^2} \sum_{(x,y) \in I} [f(x, y, t_2) - f(x - u(x, y, t_1), y - v(x, y, t_1), t_1)]^2$$

Ou, U est le potentiel de Gibbs.

III. 16

III.5 Le modèle à priori

L'articulation de la fonction de densité de **probabilité a priori** du champ de vecteur mouvement est directement déduite de l'équation III. 5, comme la fonction de densité de probabilité d'une distribution de Gibbs [5]. Etant donné le champ des vecteurs mouvements est un champ markovien:

$$p(u(x, y, t_1), v(x, y, t_1) | f(x, y, t_1)) = \frac{1}{Z} \exp \left\{ -U_p(u(x, y, t_1), v(x, y, t_1) | f(x, y, t_1)) \right\}$$

III. 17

Ou l'énergie u_p est donnée par :

$$U_p(u(x, y, t_1), v(x, y, t_1) | f(x, y, t_1)) = \lambda \sum_{c \in C} V_p^c(u(x, y, t_1), v(x, y, t_1) | f(x, y, t_1))$$

III. 18

Est composée des mêmes potentiels cliques V_p^c comme dans III. 12

De l'équation III. 2, notre but est de maximiser l'expression suivante :

$$p(f(x, y, t_2) | u(x, y, t_1), v(x, y, t_1), f(x, y, t_1)) \cdot p(u(x, y, t_1), v(x, y, t_1) | f(x, y, t_1))$$

III. 19

Ou une fonction monotone croissante de III. 19 telle que :

$$\log(p(f(x, y, t_2) | u(x, y, t_1), v(x, y, t_1), f(x, y, t_1))) + \log(p(u(x, y, t_1), v(x, y, t_1) | f(x, y, t_1)))$$

III. 20

De III. 16, III. 18 et III. 20, nous déduisons que la maximisation de la **probabilité à posteriori** III. 2, est équivalente à la minimisation de l'expression suivante:

$$\min_{u, v} \left\{ \sum_{(x, y) \in I} [f(x, y, t_2) - f(x - u(x, y, t_1), y - v(x, y, t_1), t_1)]^2 + \lambda \sum_{c \in C} V_p^c(u(x, y, t_1), v(x, y, t_1) | f(x, y, t_1)) \right\}$$

III. 21

Ou en tenant compte des potentiels cliques définis en III. 12, pour les cliques doubles :

$$\min_{u, v} \left\{ \sum_{(x, y) \in I} [f(x, y, t_2) - f(x - u(x, y, t_1), y - v(x, y, t_1), t_1)]^2 + \frac{\lambda}{2} \sum_{(x, y) \in I} [(u(x, y, t_1) - u(x', y', t_1))^2 + (v(x, y, t_1) - v(x', y', t_1))^2] \right\}$$

III. 22

Où les deux sommes sont faites sur tous les pixels de l'image, à noter que dans la seconde somme (x, y) et (x', y') sont voisins et chaque clique de deux pixels apparait deux fois, d'où la nécessité de diviser par deux.

Comme la seule contrainte imposée aux fonctions $f(x, y, t_1)$ et $f(x, y, t_2)$, est l'hypothèse qu'il s'agisse de réalisations particulières d'un champ Markovien, sans aucune perte, nous pouvons donc étendre cette formule aux fonctions représentants la différence entre les images successives d'une vidéo. Ainsi, dans les étapes suivantes, nous allons utiliser $f(x, y, t_1)$ pour dénoter la différence entre l'image courante et l'image précédente et $f(x, y, t_2)$ la différence entre l'image suivante et l'image actuelle.

III.6 Méthodologie à suivre.

Afin d'obtenir un champ de vecteur de mouvement unique à partir d'images données, il est nécessaire de spécifier un modèle structurel reliant les vecteurs de mouvement et la valeur de l'intensité de l'image, ainsi que quelques hypothèses afin d'avoir le mouvement réel.

Il est courant de supposer, comme décrit dans les chapitres précédents, que l'intensité de l'image le long des trajectoires de mouvement ne change pas et les vecteurs de mouvement subissent des déplacements petits pour un voisinage restreint.

A partir de ces deux hypothèses, on peut estimer le champ des vecteurs mouvements en minimisant la fonction d'énergie, et en utilisant un modèle de mouvement défini dans III. 22.

$$\min_{u,v} \left\{ \sum_{(x,y) \in I} \left[f(x,y,t_2) - f(x-u(x,y,t_1), y-v(x,y,t_1), t_1) \right]^2 + \frac{\lambda}{2} \sum_{(x,y) \in I} \left[(u(x,y,t_1) - u(x',y',t_1))^2 + (v(x,y,t_1) - v(x',y',t_1))^2 \right] \right\}$$

III. 23

Le paramètre λ représente le rapport de poids de l'importance entre l'information de mouvement à partir de l'intensité de l'image et l'hypothèse de lissage.

Cette approche, conduit à un inconvénient fondamental, tel que décrit au chapitre précédent: l'effet de propagation du à l'hypothèse de lissage tend à brouiller le champ des vecteurs mouvements à travers les frontières de ces derniers; Le problème vient du fait qu'aux frontières du mouvement il ya deux tendances différentes du mouvement, un pour chaque côté de la frontière. LU [6] a appelé ces mauvaises données dans l'estimation du processus des «outliers» autrement dit des **valeurs aberrantes** et les a éliminées localement en utilisant une approche dite de seuil.

En utilisant cette méthode, on peut écrire III. 23, comme suit :

$$\min_{u,v} \left\{ \sum_{(x,y) \in I} \left[f(x,y,t_2) - f(x-u(x,y,t_1), y-v(x,y,t_1), t_1) \right]^2 + \lambda' \sum_{(x,y) \in V_8(x,y)} \left[(u(x,y,t_1) - u(x',y',t_1))^2 + (v(x,y,t_1) - v(x',y',t_1))^2 \right] \right\}$$

III. 24

Avec $\lambda' = \lambda/2$, La seconde somme de l'équation s'étend sur tous les voisins V_8 de (x, y) , y compris ceux que nous souhaitons éliminer.

Par conséquent, nous allons remplacer la somme à l'intérieur par certains estimateurs robustes qui permettront d'éliminer tous les potentiels cliques avec rang plus grand qu'un certain seuil prédéfini. Ainsi, nous pouvons exprimer III. 24, comme suit :

$$\min_{u,v} \left\{ \sum_{(x,y) \in I} \left[f(x, y, t_2) - f(x - u(x, y, t_1), y - v(x, y, t_1), t_1) \right]^2 + \lambda'' \sum_{(x', y') \in V_8(x, y)} \delta(x', y') \left[(u(x, y, t_1) - u(x', y', t_1))^2 + (v(x, y, t_1) - v(x', y', t_1))^2 \right] \right\}$$

Avec $\lambda'' = \lambda' \times |V_8| / N_t$ et $N_t = \sum_{(x', y') \in V_8(x, y)} \delta(x', y')$ Le cardinal de l'ensemble $|V_8| = 8$ III. 25

Et la fonction d'indication $\delta(x', y')$ est égale à 1 pour le plus petit potentiel de la clique et est égale à zéro pour les autres, comme suit :

$$\delta(x', y') = \begin{cases} 1 & \text{si le rang du potentiel clique (ordre croissant) est inférieur à } N_t \\ 0 & \text{sinon} \end{cases} \quad \text{III. 26}$$

La valeur par défaut de N_t est 3 (c'est à dire, nous utilisons uniquement les 3 premiers potentiels cliques afin de déterminer l'énergie). De toute évidence, le rejet des valeurs aberrantes n'est pas sensible à l'amplitude des vecteurs de mouvement, car le rejet s'effectue localement.

Dans le but de minimiser la fonction d'énergie dans l'équation III. 23, afin de trouver le champ des vecteurs de mouvement, nous utilisons la **méthode du recuit simulé** avec un espace d'état discret pour les valeurs possibles du vecteur de mouvement à la résolution en pixels, semblable à ceux décrits dans Konrad et Dubois [11] et Lu [6]. Le recuit simulé a donné des résultats assez impressionnant depuis sa première application proposée par Kirkpatrick, Gelatt et Vechhi [9] pour résoudre les problèmes d'optimisation combinatoire à grande échelle.

Il a été utilisé dans de nombreux problèmes d'optimisation, y compris l'estimation du mouvement et la restauration de l'image. Par ailleurs, Geman et Geman [7] ont prouvé que, sous certaines conditions, l'algorithme du recuit simulé pourrait fournir la solution globale optimale pour n'importe quelle valeur initiale.

Nous employons une méthode semblable à celle suggérée par Geman et Geman [7], à savoir l'échantillonneur de Gibbs [Annexe C], car il fournit un allègement du temps de calcul comparé à l'algorithme de Metropolis. Une caractéristique très importante des méthodes de relaxation stochastique est qu'elles produisent des états en fonction des probabilités de leur occurrence, à savoir, les états peu probables sont également générés, mais moins fréquemment, bien sûr. Lorsqu'ils sont incorporés dans le recuit simulé, cette propriété permet aux minima locaux d'être éliminés, contrairement au cas des méthodes de gradients.

L'échantillonneur de Gibbs pour le processus de déplacement discrets $u(x, y)$ et $v(x, y)$ est une procédure de remplacement de pixels, dans lequel chaque site de pixel (x, y) est visité dans l'ordre naturel et le nouveau vecteur mouvement est généré pour chaque site selon une certaine probabilité conditionnelle comme suit : Soit M l'ensemble des valeurs possibles des vecteurs de mouvement avec une précision Δ (pixel, demi pixel, quart de pixel ...etc.) et le total des niveaux $(2 \times MAX_U + 1) \times (2 \times MAX_V + 1)$ c'est-à-dire :

$$M = \{(u, v) \mid u = 0, \pm\Delta, \pm2\Delta, \dots, \pm MAX_U \times \Delta, \text{ et } v = 0, \pm\Delta, \pm2\Delta, \dots, \pm MAX_V \times \Delta\}$$

Comme le montre la figure suivante.

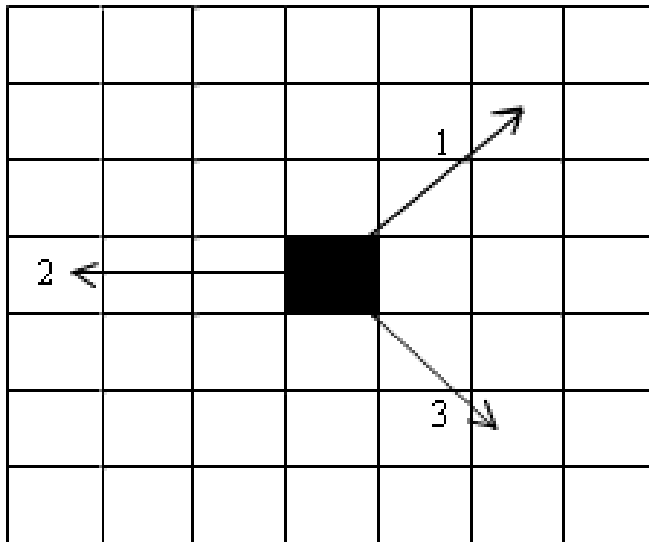


Figure III.2 : Les 49 déplacements possibles M pour $MAX_U = MAX_V = 3$

exemple

- 1) $u = -2$, $V = 2$
- 2) $u = 0$, $V = -3$
- 3) $u = 2$, $V = 2$

Supposons que nous voulons estimer un vecteur de mouvement inconnu avec la fonction d'énergie $U(q)$, où $q \in Q$ est une réalisation du champ des vecteurs de mouvement, telle que

$q = \{ (u(x, y), v(x, y)) \mid (x, y) \in I \}$, et Q représente l'ensemble des réalisations possibles.

Nous modélisons q comme un champ de vecteur aléatoire avec une distribution de Gibbs comme suit:

$$p(q) = \frac{1}{Z} \exp\left(-\frac{U(q)}{T}\right) \quad \text{III. 27}$$

Où T est le paramètre de température de l'algorithme du recuit simulé et Z est la constante de normalisation telle que $\sum_{q \in Q} p(q) = 1$.

Soit $(x_0, y_0) \in I$ le pixel étudié. Si q' est une réalisation du champ de vecteur mouvement Q telle que $q' = q$ pour tous les sites sauf (x_0, y_0) , puis nous calculons la probabilité conditionnelle de $q(x_0, y_0) = (u_0, v_0) \in M$ donnant q' , comme suit :

$$\begin{aligned} p(q(x_0, y_0) = (u_0, v_0) \mid q(x, y) = q'(x, y), \forall (x, y) \in I \setminus \{(x_0, y_0)\}) &= \\ &= \frac{p(q(x, y) = q'(x, y), \forall (x, y) \in I \setminus \{(x_0, y_0)\} \cup q(x_0, y_0) = (u_0, v_0))}{\sum_{(u, v) \in M} p(q(x, y) = q'(x, y), \forall (x, y) \in I \setminus \{(x_0, y_0)\} \cup q(x_0, y_0) = (u, v))} \end{aligned}$$

III. 28

$U(q)$ étant défini comme dans III. 23, on peut donc écrire :

$$\begin{aligned} U(q(x, y) = q'(x, y), \forall (x, y) \in I \setminus \{(x_0, y_0)\} \cup q(x_0, y_0) = (u_0, v_0)) &= \\ &= \sum_{(x, y) \in I \setminus \{(x_0, y_0)\}} \left[f(x, y, t_2) - f(x - u(x, y, t_1), y - v(x, y, t_1), t_1) \right]^2 + \\ &\quad + \frac{\lambda}{2} \sum_{(x', y') \in V_8(x, y)} \left[\left(u(x, y, t_1) - u(x', y', t_1) \right)^2 + \left(v(x, y, t_1) - v(x', y', t_1) \right)^2 \right] + \\ &\quad + \left[f(x_0, y_0, t_2) - f(x_0 - u_0, y_0 - v_0, t_1) \right]^2 + \\ &\quad + \frac{\lambda}{2} \sum_{(x', y') \in V_8(x, y)} \left[\left(u_0 - u(x', y', t_1) \right)^2 + \left(v_0 - v(x', y', t_1) \right)^2 \right] \end{aligned}$$

III. 29

Et d'une façon similaire $U(q(x, y) = q'(x, y), \forall (x, y) \in I \setminus \{(x_0, y_0) \cup q(x_0, y_0) = (u, v)\})$. Notez que la somme dans III. 23, est découpée en une sommation sur $(x, y) \in I \setminus \{(x_0, y_0)\}$ et le terme correspondant pour $(x, y) = (x_0, y_0)$. En soustrayant le potentiel clique en (x_0, y_0) à partir de la première somme, cela devient complètement indépendant de (x_0, y_0) . Par conséquent, nous pouvons réorganiser III. 26, comme suit:

$$\begin{aligned}
 U(q(x, y) = q'(x, y), \forall (x, y) \in I \setminus \{(x_0, y_0) \cup q(x_0, y_0) = (u_0, v_0)\}) = \\
 = T_1 + \left[f(x_0, y_0, t_2) - f(x_0 - u_0, y_0 - v_0, t_1) \right]^2 + \\
 + \frac{\lambda}{2} \sum_{(x', y') \in V_8(x, y)} \left[(u_0 - u(x', y', t_1))^2 + (v_0 - v(x', y', t_1))^2 \right]
 \end{aligned} \tag{III. 30}$$

Où T_1 est indépendant de (x_0, y_0) . Comme (x_0, y_0) est choisi arbitrairement parmi tous les éléments de M , il s'ensuit que III. 30, soit vérifiée pour toute valeur possible de l'ensemble M . Ce qui nous donne.

$$\begin{aligned}
 p(q(x, y) = (u_0, v_0) | q(x, y) = q'(x, y), \forall (x, y) \in I \setminus \{(x_0, y_0)\}) = \\
 = \frac{\exp\left(-\frac{U(q(x, y) = q'(x, y), \forall (x, y) \in I \setminus \{(x_0, y_0)\}) \cup q(x_0, y_0) = (u_0, v_0)}{T}\right)}{\sum_{(u, v) \in M} \exp\left(-\frac{U(q(x, y) = q'(x, y), \forall (x, y) \in I \setminus \{(x_0, y_0)\}) \cup q(x_0, y_0) = (u, v)}{T}\right)}
 \end{aligned} \tag{III. 31}$$

Notons que dans III. 31 le terme T_1 est supprimé, alors que les deux derniers termes de III. 30, consistent uniquement en des calculs locaux impliquant le pixel (x_0, y_0) et ses voisins. Ainsi, nous pouvons calculer les nouveaux vecteurs de mouvement de tous les pixels simultanément, ce qui signifie que nous pouvons résoudre le problème de minimisation d'une manière très parallèle.

Soit T_0 la température initiale, MAX_ITER le nombre maximum d'itérations, et soit q_0 une certaine réalisation du champ de vecteur mouvement. L'algorithme du recuit simulé permettant de retrouver les vecteurs mouvements en maximisant une probabilité a posteriori (MAP) peut être résumé comme suit :

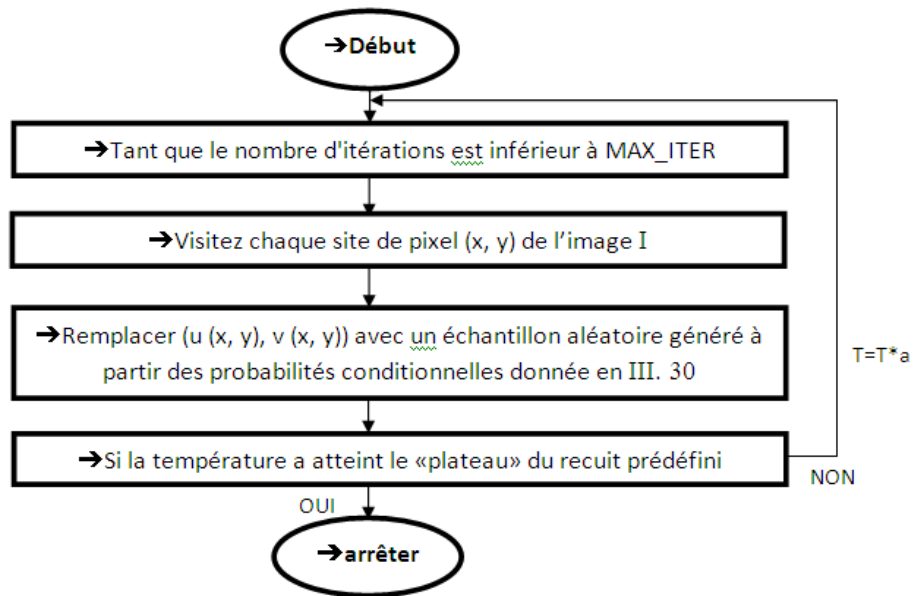


Figure III. 3 fonctionnement du recuit simulé

Pour comprendre le remplacement de la valeur actuelle de u et v au pixel (x, y) , on indique qu'à chaque endroit une probabilité de distribution bi-variée doit être calculée (exp : la probabilité pour tous les déplacements possibles des vecteurs). Alors, la distribution 2D est accumulée le long de la direction horizontale afin d'obtenir une distribution marginale cumulative sur une seule dimension 1D. Cette distribution est échantillonnée avec un nombre aléatoire uniformément réparti, générant ainsi la valeur du déplacement vertical u . Le déplacement horizontal v est généré par l'échantillonnage de la distribution 1D distribution correspondant au déplacement vertical u .

Le recuit simulé doit commencer par une température très élevée (par exemple $A = 500$) et se termine à une très basse température (par exemple, inférieure à 0,1), la température à la K -ième itération est donnée par $T_k = T_0 * a^{k-1}$ avec a = constante, légèrement inférieur à 1.

III.7 Organigrammes détaillés de la méthode d'optimisation.

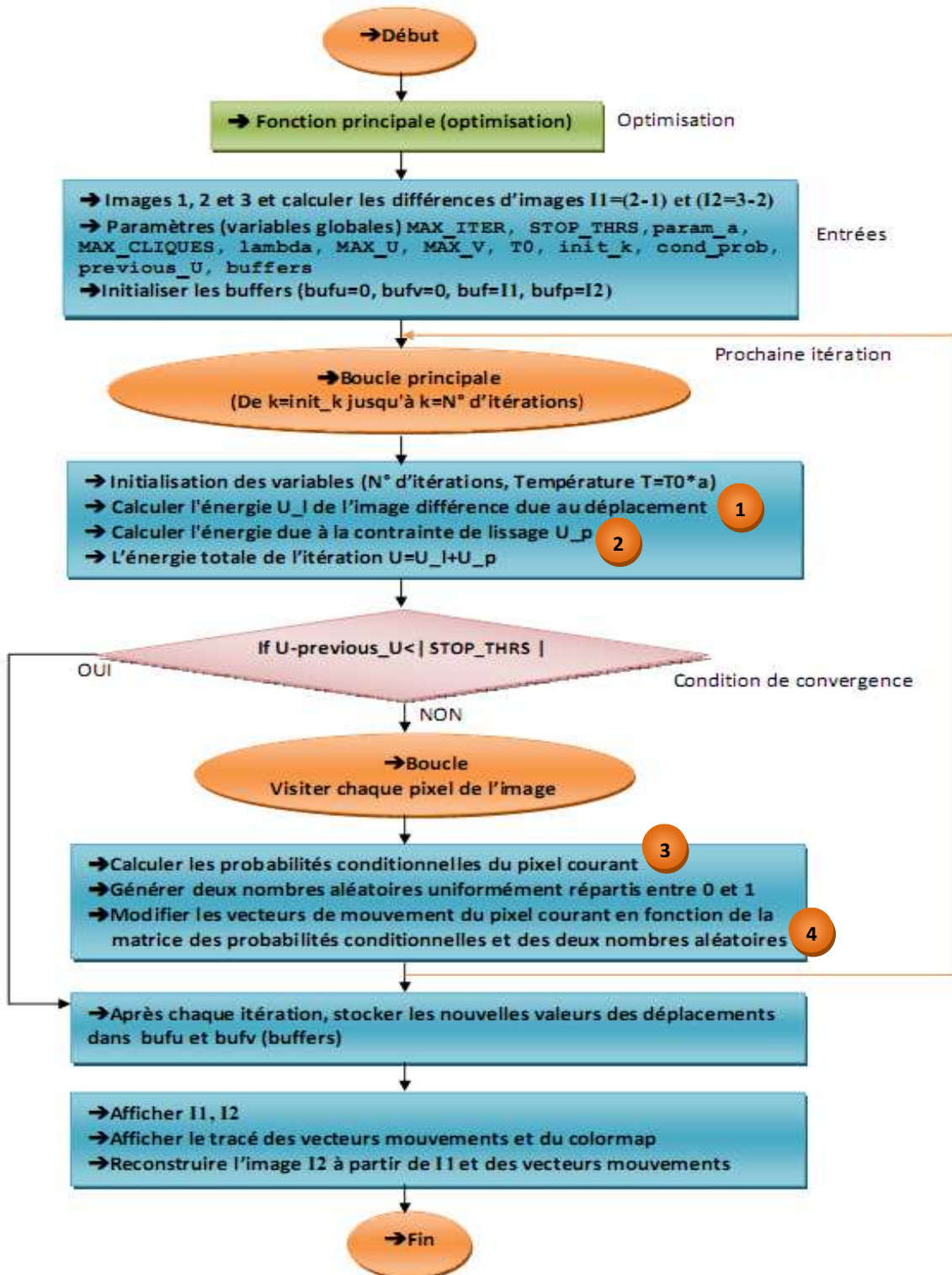


Figure III.4 : Organigramme montrant la fonction principale du programme d'optimisation

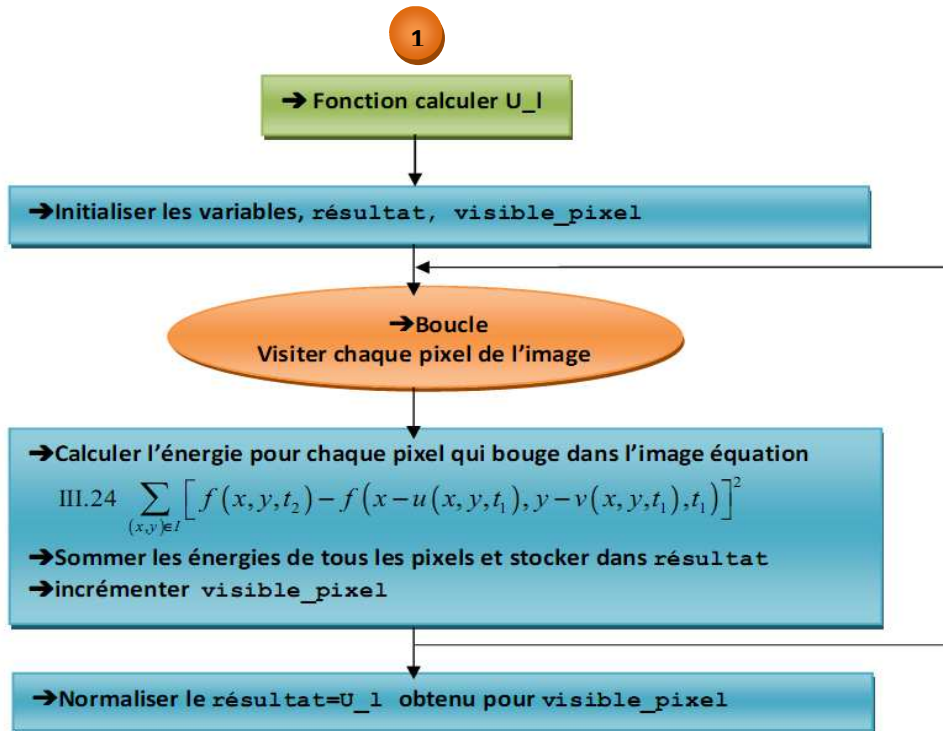


Figure III.5 : Calcul de l'énergie du aux déplacements existants entre les deux images

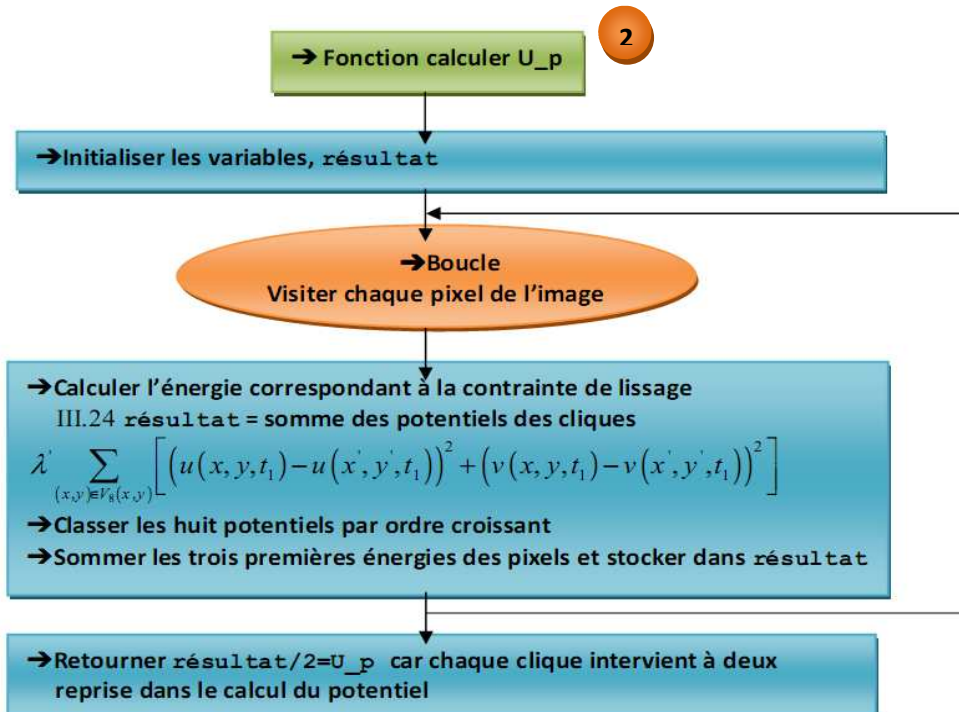


Figure III.6 : Calcul du potentiel cliques des huit voisins de chaque pixel

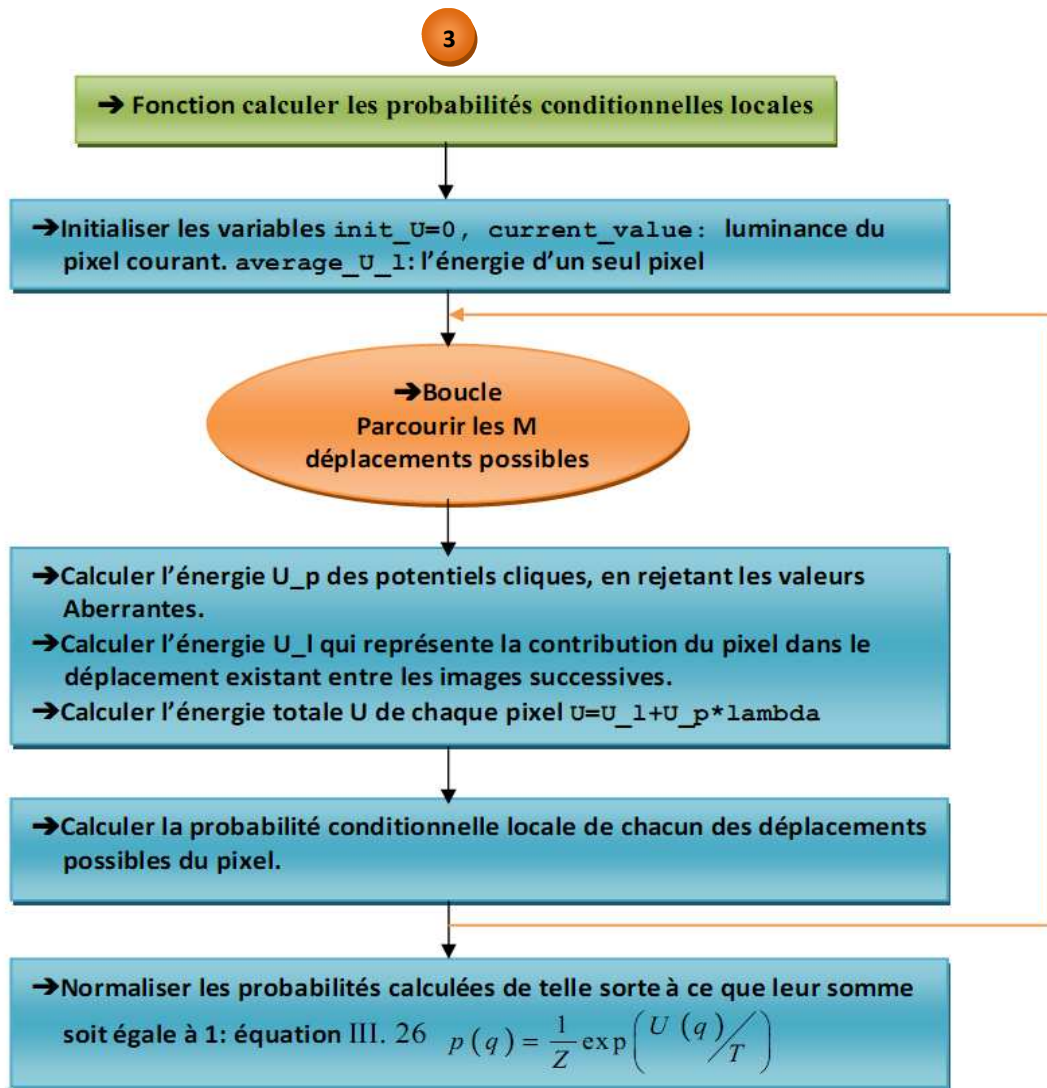


Figure III.7 : Calcul des probabilités conditionnelles locales de chaque pixel

A noter que dans notre application, nous avons choisi d'utiliser la méthode dite « local outlier rejection » rejet des valeurs aberrantes, cette technique nous aide à avoir une meilleure estimation de la probabilité conditionnelle, en choisissant uniquement les potentiels des pixels les plus importants.

Sachant que les pixels des bords supérieurs et inférieurs de l'image ne contiennent que 3 voisins, nous avons opté pour un choix de 3 potentiels sur les huit, une boucle nous permet d'ordonner les huit potentiels par ordre croissant et d'en choisir la somme des trois premiers potentiels,

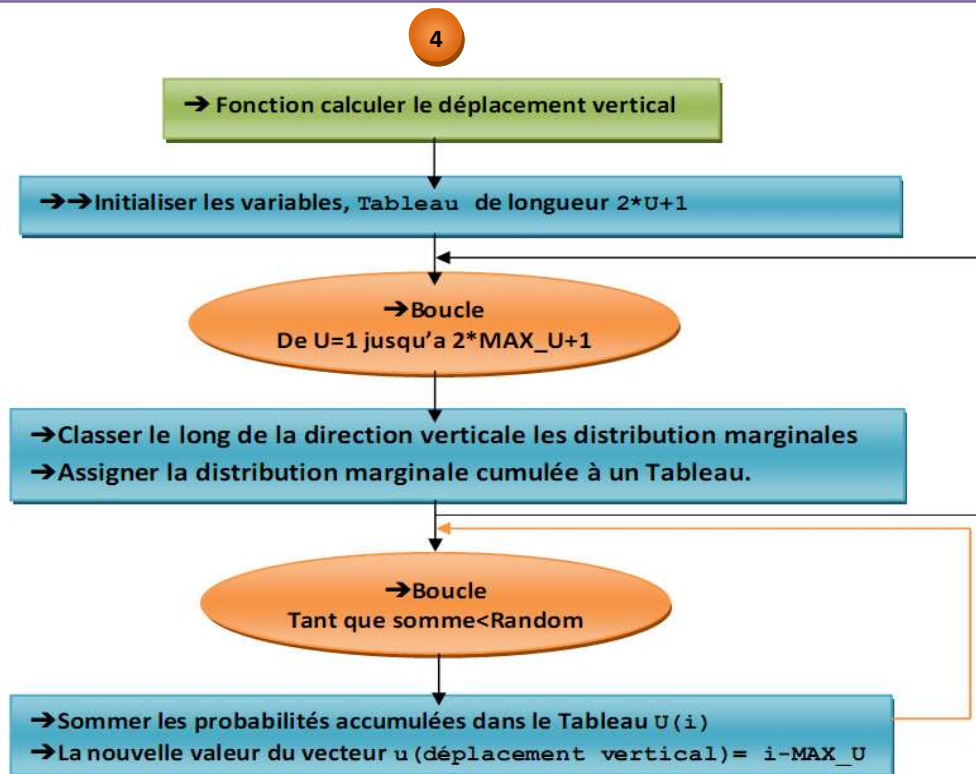


Figure III.8 : Calcul du déplacement vertical.

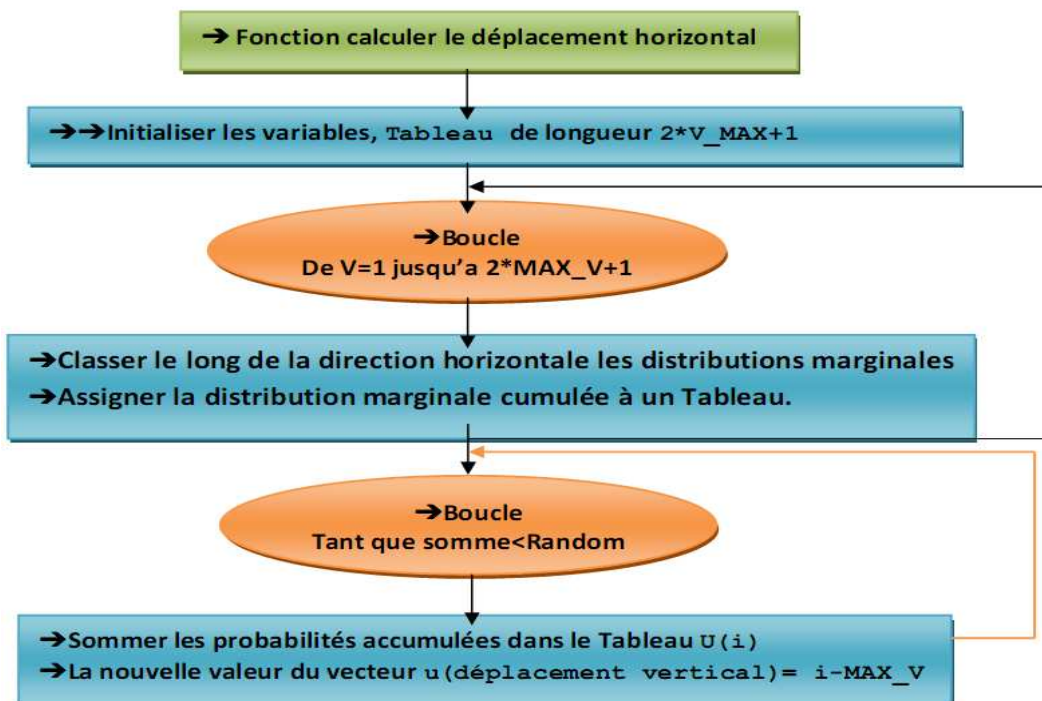
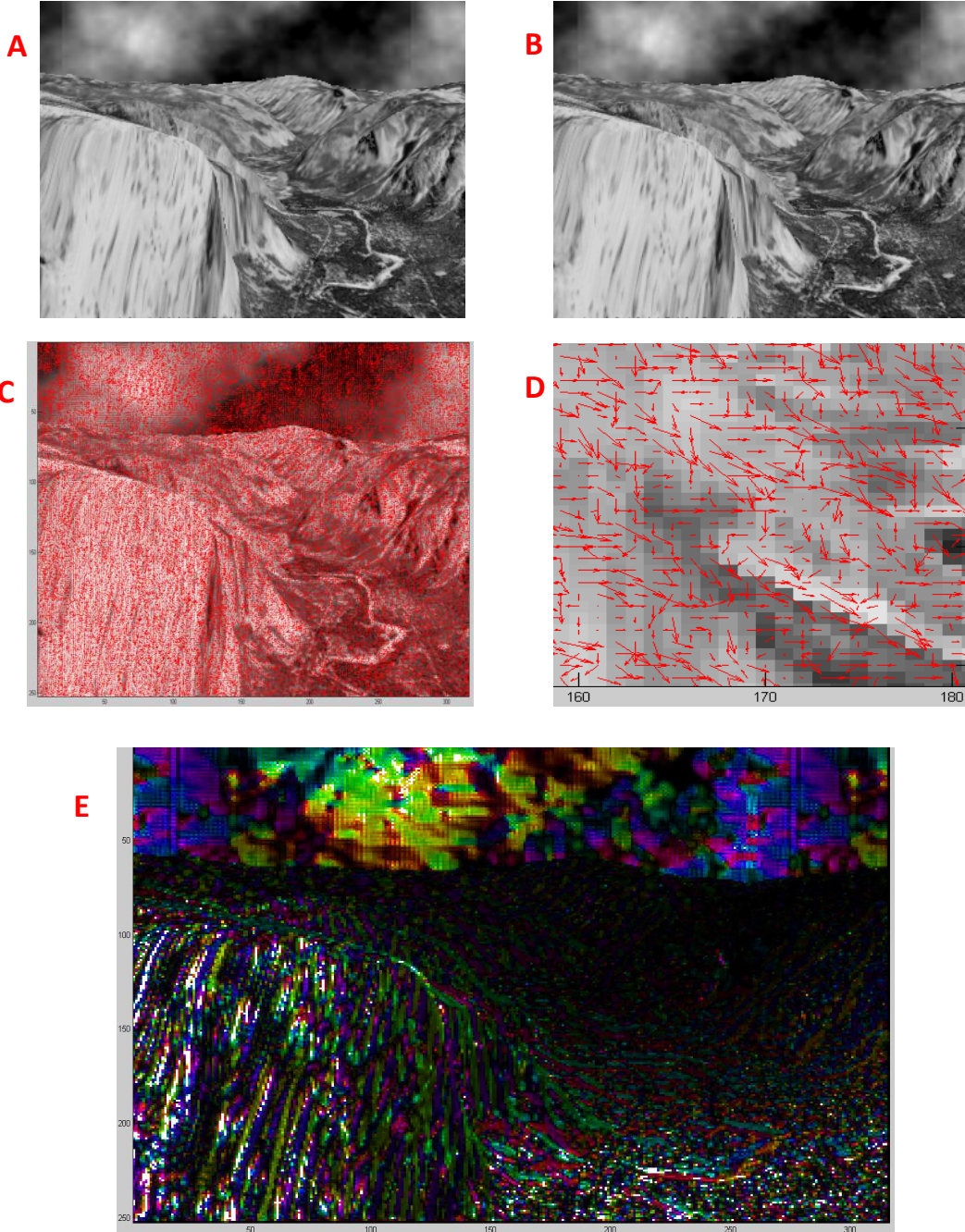


Figure III.9 : Calcul du déplacement horizontal.

III.8 Résultats pratiques





Figures III.10 : **A**-image test YOS9, **B**-image test YOS10, **C-D** tracé des vecteurs mouvements en utilisant la méthode probabiliste, **E**-représentation des vecteurs vitesses par la carte des couleurs.

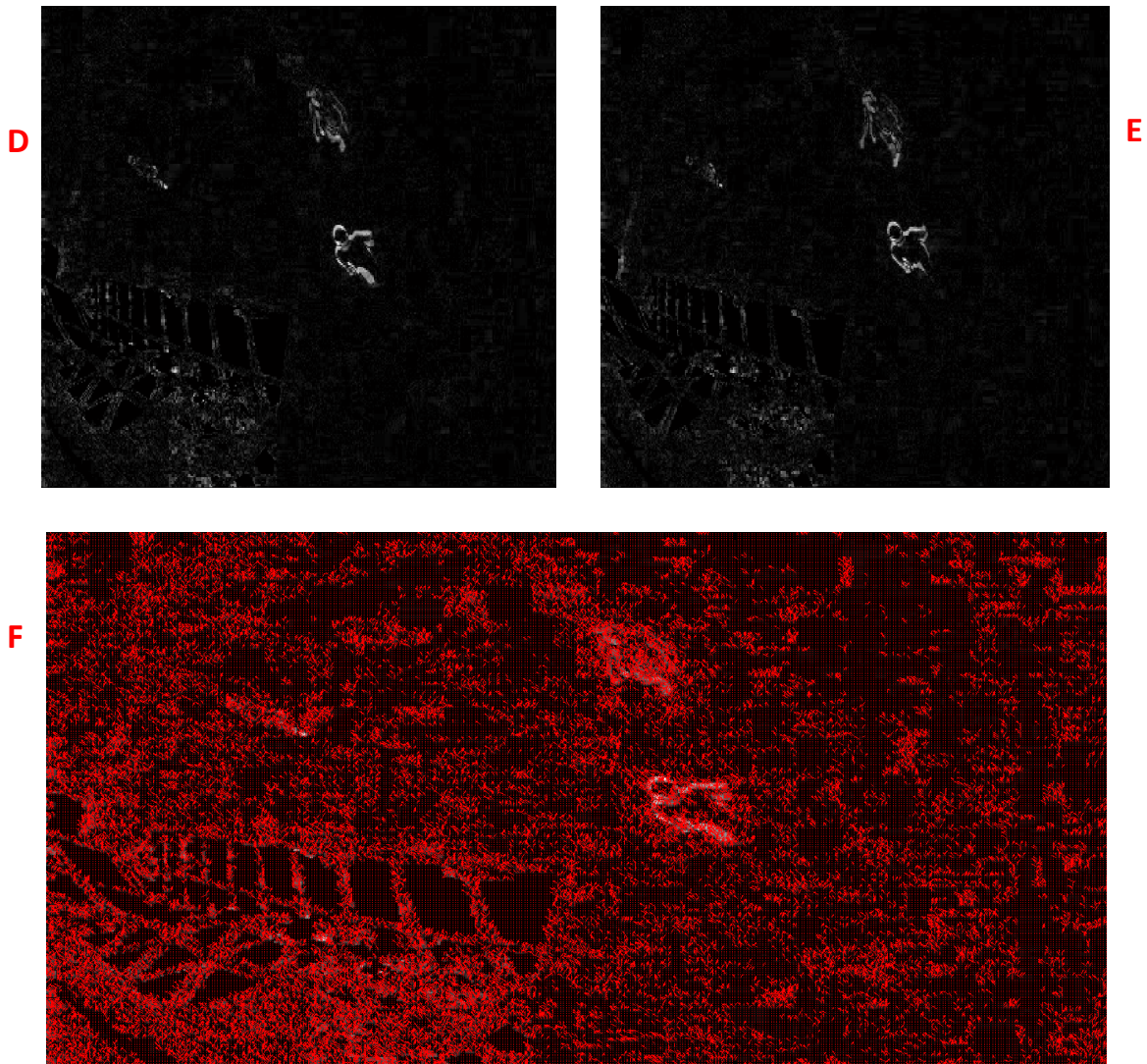
F-reconstruction de la seconde image test grâce aux vecteurs mouvements

Nous avons obtenu, des résultats satisfaisants pour cette séquence d'image, qui sont analogues à ceux obtenus par HS et LK, avec comme inconvénient principal, des temps de calcul très lent.

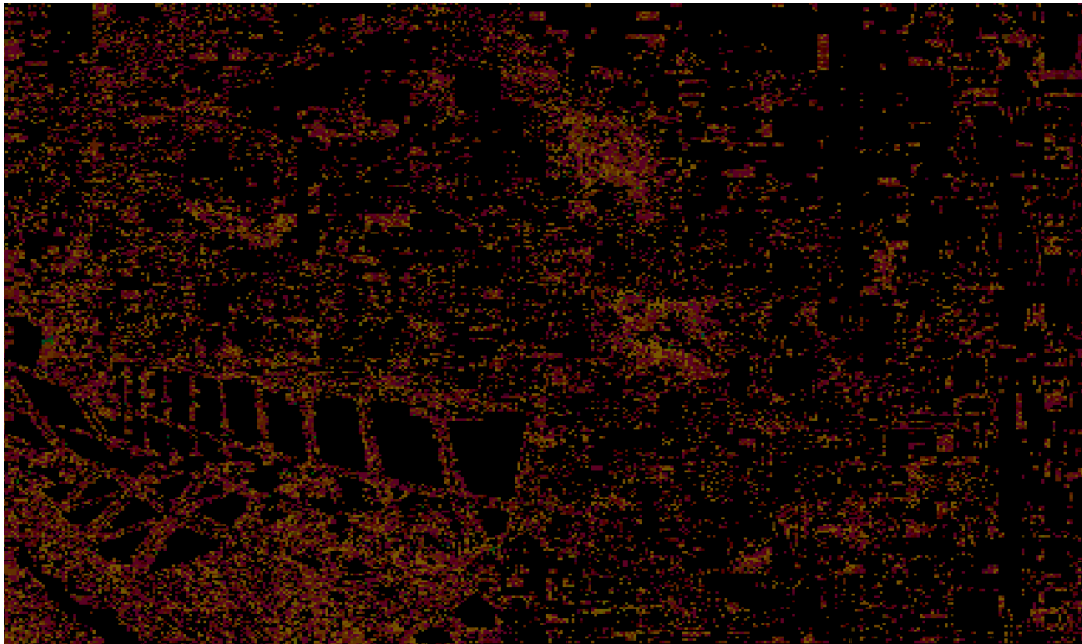
Dans l'exemple suivant, contrairement au premier, nous n'utiliserons pas les images elle mêmes pour l'estimation du mouvement, mais les différences d'images, pour cela, nous aurons besoin d'une séquence de trois images nous permettant de récupérer deux images différences auxquelles sera appliqué notre algorithme d'estimation du mouvement.



En faisant la différence (B)-(A) et (C)-(B), nous obtenons les images différences suivantes.



G



Nous constatons que nous avons plutôt une assez bonne estimation du mouvement comme illustré par la carte des couleurs (**G**) et les vecteurs mouvement (**F**).

H



Figures III.11 : **A-B-C** images tests **D-E** images différences tests. **F**-tracé des vecteurs mouvements, **G**-représentation des vecteurs vitesses par la carte des couleurs. **H**-reconstruction de la seconde image test grâce aux vecteurs mouvements

Nous obtenons une assez bonne segmentation avec cette méthode probabiliste, toutefois, elle est moins bonne que celle obtenue avec la méthode de Lucas et Kanade.

On constate aussi lorsque l'on effectue la reconstruction de l'image (déplacement des pixels par la vitesse calculée point à point) que l'image déplacée est de bonne qualité, ce qui signifie que l'algorithme a bien retrouvé les pixels correspondants.

III.9 Discussion

Nous venons de développer l'approche probabiliste permettant d'estimer le mouvement dans une séquence d'images (différences d'images) et nous remarquons que globalement, la méthode probabiliste permet d'avoir des résultats analogues ou très proche de ceux des méthodes de Horn et Shunck & Lucas et Kanade, néanmoins, un inconvénient majeur est à souligner ; en termes de temps de calcul, cette méthode probabiliste est extrêmement coûteuse, ce qui limite beaucoup son utilisation, surtout pour des applications nécessitant une estimation du mouvement en temps réel.

Conclusion

Conclusion

Dans cette étude, nous nous sommes intéressés à l'estimation du mouvement dans une séquence d'images vidéo par approche probabiliste.

Une étude bibliographique des différentes techniques d'estimation du mouvement a été d'abord abordé dans le premier chapitre, en introduisant dans un premier temps, la contrainte fondamentale du flot optique, avant d'énoncer les principales techniques existantes et surtout les plus connues.

Deux tests ont été effectués dans le second chapitre sur les méthodes différentielles de Horn et Shunck et celle de Lucas et Kanade, tout en optimisant nos résultats en jouant sur les paramètres suivants :

- Nombre d'itérations
- Paramètre de lissage alpha
- Choix de la fenêtre de recherche...etc.

Deux améliorations ont été apportées à la méthode différentielle de Lucas et Kanade toujours dans le but d'optimiser les résultats, à savoir l'implémentation pyramidale et le raffinement itératif.

Dans le troisième chapitre, nous avons étudié l'approche probabiliste qui consiste en une technique bayésienne pour l'estimation du mouvement 2D dans une séquence d'images vidéo. Nous avons introduit une contrainte de lissage stochastique pour modéliser le domaine 2D des vecteurs de mouvement basé sur une distribution de Gibbs.

L'optimisation est appliquée à des différences successives d'images vidéo au lieu des images elles-mêmes, et ceci afin de réduire efficacement la redondance temporelle et spatiale et la gamme dynamique des niveaux d'intensités des pixels.

Nous avons reformuler la technique de l'échantillonneur de Gibbs pour échantillonner la solution d'un seul pixel à la fois suivant les probabilités conditionnelles calculées pour chaque valeur possible des composantes horizontales et verticales du vecteur mouvement.

Le processus est itératif et les probabilités impliquées dans le calcul dépendent d'un paramètre de contrôle appelé température. Du point de vue informatique, l'algorithme a coûté énormément en temps de calcul comparé aux autres méthodes développées dans le deuxième chapitre, pour cause, l'algorithme du recuit simulé converge relativement très lentement.

Néanmoins les résultats obtenus sont analogues à ceux des méthodes les plus connues, et des perspectives d'améliorations peuvent être envisagées telles un meilleur modèle mathématique pour caractériser la différence d'image vidéo, un meilleur algorithme d'optimisation, car l'algorithme utilisé est certes très prometteur, mais il converge lentement.

Annexes

Annexe A

Formulation du flot optique

Considérons un patch d'une certaine luminosité qui se déplace sur une distance δx le long de la direction des X et de δy le long de la direction des y. La luminosité du patch est supposée rester constante ; de sorte que.

$$E(x, y, t) = E(x + \delta x, y + \delta y, t + \delta t)$$

En développant le coté droit, nous obtenons,

$$E(x, y, t) = E(x, y, t) + \delta x \frac{\partial E}{\partial x} + \delta y \frac{\partial E}{\partial y} + \delta t \frac{\partial E}{\partial t} + \varepsilon$$

ou ε contient les termes de second ordre fonction de $\delta x, \delta y$ et δt . Après soustraction de $E(x, y, t)$ de part et d'autre de l'équation et en divisant par δt , on obtient :

$$\frac{\delta x}{\delta t} \frac{\partial E}{\partial x} + \frac{\delta y}{\delta t} \frac{\partial E}{\partial y} + \frac{\partial E}{\partial t} + o(\delta t) = 0$$

Ou $o(\delta t)$, est un terme fonction de δt et que δx et δy varient de la même façon que δt .

Pour la limite δt tend vers zéro ($\delta t \rightarrow 0$) on a :

$$\frac{\partial E}{\partial x} \frac{dx}{dt} + \frac{\partial E}{\partial y} \frac{dy}{dt} + \frac{\partial E}{\partial t} = 0$$

Annexe B

Définition et introduction au champ de Markov

Dans cette annexe, nous introduisons tout d'abord de façon intuitive la notion d'énergie locale avant de définir plus formellement un champ de Markov et d'énoncer le théorème d'équivalence entre champs de Markov et champs de Gibbs.

Description de l'image

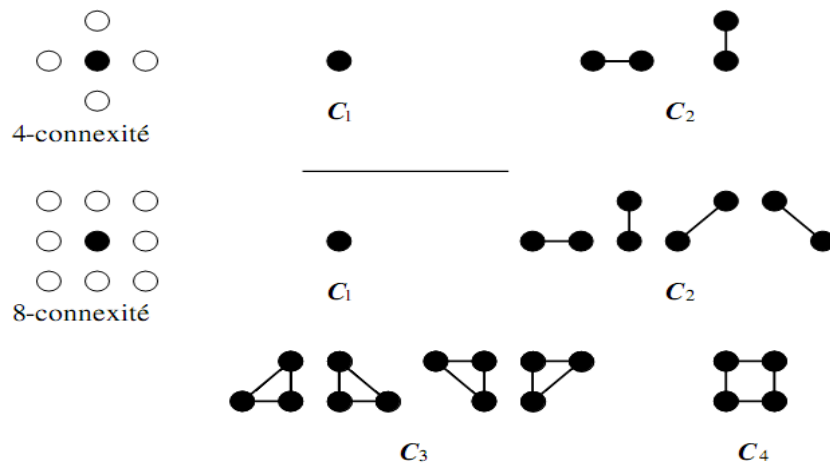
L'image est formée d'un ensemble fini S de sites s_i correspondant aux pixels. S est donc essentiellement un réseau discret fini, partie de $\mathbb{Z}^d$, si on note d la dimension de l'espace ("2" le plus classiquement, "3" pour les volumes, etc.). A chaque site est associé un descripteur représentant l'état du site et qui peut être son niveau de gris, une étiquette, ou une information plus complexe, et prenant ses valeurs dans E .

La notion d'interactions locales nécessite de structurer les relations spatiales entre les différents sites du réseau.

Pour ce faire, on munit S d'un système de voisinage V défini de la façon suivante :

$$V_s = \{t\} \text{ tels que } \begin{cases} s \notin V_s \\ t \in V_s \Rightarrow s \in V_t \end{cases}$$

A partir d'un système de voisinage, un système de cliques peut être déduit : une clique est soit un singleton de S , soit un ensemble de sites tous voisins les uns des autres. En fonction du système de voisinage utilisé, le système de cliques sera différent et fera intervenir plus ou moins de sites comme illustre sur la figure suivante. On notera C l'ensemble des cliques relatif à V , et C_k l'ensemble des cliques de cardinal k .



Les interactions locales entre niveaux de gris (ou descripteurs) de sites voisins peuvent alors s'exprimer comme un potentiel de clique. Soit c une clique, on lui associe le potentiel U_c dont la valeur dépend des niveaux de gris (ou descripteurs) des pixels constituant la clique. En poursuivant ce raisonnement, on peut définir l'énergie globale de l'image comme la somme des potentiels de toutes les cliques : $U = \sum_{c \in C} U_c$

Et l'énergie locale en un site comme la somme des potentiels de toutes les cliques auxquelles il appartient : $U_s = \sum_{c \in C / s \in c} U_c$

Nous avons jusqu'ici considéré le cas d'une image pour illustrer les notions de voisinage, de clique et de potentiel, mais le formalisme markovien se définit très généralement sur tout graphe. Soit un ensemble de sites S dénombrables (sommets du graphe), et une relation de voisinage, les cliques sont alors définies comme les sous-graphes complets du graphe. C'est l'utilisation de graphes plus généraux que ceux définis sur la grille de l'image qui permet des traitements de plus haut niveau.

Modélisation probabiliste de l'image

La définition des champs de Markov qui sera donnée dans la section suivante nécessite une modélisation probabiliste de l'image. Ainsi, l'image dont nous disposons va être considérée comme une réalisation d'un champ aléatoire. Soit s un site de l'image, on peut en effet lui associer une variable aléatoire (v.a) X_s prenant ses valeurs

Définition et simulation d'un champ de Markov

Dans E . Le niveau de gris X_s en s n'est ainsi qu'une réalisation de la v.a X_s . On définit alors le champ aléatoire $X = (X_s, X_t, \dots)$ prenant ses valeurs dans $\Omega = E^{|S|}$. On trouvera aussi le terme de processus aléatoire pour X ; en toute rigueur, "processus" devrait être réservé au cas d'un ensemble d'indexation continu, et champ au cas discret. Dans ce cadre probabiliste, l'image considérée est simplement une réalisation x du champ. La probabilité globale de x , $P(x = x)$, permet d'accéder en quelque sorte à la vraisemblance de l'image, et les probabilités conditionnelles locales d'une valeur en un site permettent de mesurer le lien statistique entre un niveau de gris et le reste de l'image. L'hypothèse markovienne permet d'évaluer ces quantités. Notons que nous nous plaçons dans le cas où E , l'espace de valeurs des descripteurs, est quantité, ce qui nous permet de manipuler des probabilités. Dans le cas où cet espace est continu, il faut remplacer P par une densité de probabilité, mais dans ce cas le théorème que nous allons voir ci-dessous n'est plus valable.

Champs de Markov - Champs de Gibbs

Définition d'un champ de Markov : Considérons x_s la valeur du descripteur prise au site s et $x^s = (x_t)_{t \neq s}$ la configuration de l'image excepte le site s . La définition d'un champ de Markov est alors la suivante :

X est un champ de Markov si et seulement si la probabilité conditionnelle locale en un site n'est fonction que de la configuration du voisinage du site considéré

Ce qui s'exprime de façon formelle par : $P(X_s = x_s / x^s) = P(X_s = x_s / x_t, t \in V_s)$

Ainsi, le niveau de gris en un site ne dépend que des niveaux de gris des pixels voisins de ce site. Cette hypothèse markovienne se justifie bien dans le cas de la plupart des images naturelles constituées de zones homogènes ou texturées ainsi que pour une large gamme d'images de synthèse. Plus généralement, une connaissance locale de l'image suffit souvent à réaliser son interprétation partielle et donc cette hypothèse markovienne sera souvent justifiée sur des graphes plus globaux que le graphe des pixels. Notons qu'en l'absence de contrainte sur le système de voisinage, tous les champs aléatoires peuvent être considérés comme markoviens à condition de prendre un voisinage suffisamment grand. L'intérêt de cette modélisation réside bien sur dans le cas où la propriété markovienne est vérifiée pour des voisinages restreints permettant des calculs rapides.

Equivalence entre champs de Markov et champs de Gibbs :

La modélisation markovienne prend toute sa puissance grâce au théorème que nous allons voir maintenant. En effet, celui-ci permettra d'accéder aux expressions des probabilités conditionnelles locales. Il nous faut au préalable définir un certain nombre de notions relatives aux mesures et champs de Gibbs.

Définition d'une mesure de Gibbs: La mesure de Gibbs de fonction d'énergie (ou d'hamiltonien) $U : \Omega \rightarrow \mathbb{R}$ est la probabilité P définie sur Ω par:

$$P(X = x) = \frac{1}{Z} \exp(-U(x))$$

$$U(x) = \sum_{c \in C} U_c(x)$$

Où C est le système de cliques associé au système de voisinage V de U^2

$Z = \sum_{x \in \Omega} \exp(-U(x))$ Est une constante de normalisation appelée fonction de partition de Gibbs. En pratique, il est quasi impossible de calculer cette constante à cause du très grand

nombre de configurations possibles. Ne serait-ce que dans le cas d'une image binaire ($\text{card}(E) = 2$) et de taille = 512×512 , on a $\text{card}(\Omega) = 2^{512 \times 512} = 2^{262144}$ configurations possibles.

La notation couramment utilisée pour $U(x)$ est abusive car $U_c(x)$ ne dépend pas de l'ensemble de la configuration x mais seulement de x restreinte à la clique c ($U_c(x) = U_c(x_t, t \in c)$). Nous pouvons maintenant définir le champ de Gibbs de potentiel associé au système de voisinage V : c'est le champ aléatoire X dont la probabilité est une mesure de Gibbs associée au système de voisinage V , ce qui implique :

$$P(X = x) = \frac{1}{Z} \exp(-U(x)) = \frac{1}{Z} \exp\left(-\sum_{c \in C} U_c(x)\right)$$

L'énergie globale d'un champ de Gibbs possède donc la propriété de se décomposer sous forme d'une somme d'énergies locales, qui comme on le verra par la suite permettront d'accéder aux probabilités conditionnelles locales. Notons ici que plus une configuration d'un champ de Gibbs a une énergie faible, plus elle est probable.

Le théorème de Hammersley-Clifford [Besag, 1974] établit alors le résultat fondamental suivant sous les hypothèses :

- S fini ou dénombrable,
- le système de voisinage V est borné,
- l'espace des états E est discret.

X est un champ de Markov relativement à V et $P(X = x) > 0 \forall x \in \Omega$ si et seulement si X est un champ de Gibbs de potentiel associé à V .

Par exemple, si nous considérons un champ de Markov de voisinage 4-connexe, nous pouvons écrire l'énergie de la configuration x sous la forme :

$$U(x) = \sum_{c=(s) \in C_1} U_c(x_s) + \sum_{c=(s,t) \in C_2} U_c(x_s, x_t)$$

Notons que rien n'interdit la non-stationnarité du champ, c'est à dire la variation des potentiels U_c en fonction de la localisation de la clique c dans l'image. D'autre part, rien n'impose la symétrie des potentiels et on peut avoir $U_c =_{(r,s)}(0,1) \neq U_c(0,1)$

Le théorème de Hammersley-Clifford, et la forme bien spécifique de probabilité de X qui en résulte, va permettre de lier les probabilités globales et locales comme nous allons le voir maintenant. En effet si nous cherchons à écrire la probabilité conditionnelle locale $P(x_s / X^s = x^s)$, nous avons grâce au résultat précédent:

$$P(X_s = x_s / X^s = x^s) = \frac{P(X = x)}{P(X^s = x^s)} = \frac{\exp(-U(x_s, x^s))}{\sum_{\xi \in E} \exp(-U(\xi, x^s))}$$

Définissons l'énergie U_s locale par :

$$U_s(x_s / x_t, t \in V_s) = \sum_{c \in C / s \in c} U_c(x_s, x_t, t \in V_s) = \sum_{c \in C / s \in c} U_c(x_s, V_s)$$

En notant $V_s = (x_t, t \in V_s)$. Cette énergie locale ne fait donc intervenir que les voisins de s . On peut alors écrire l'énergie globale $U(x)$ sous la forme :

$$U(x) = \sum_{c \in C / s \notin c} U_c(x) + \sum_{c \in C / s \in c} U_c(x) = \sum_{c \in C / s \notin c} U_c(x_s) + U_s(x_s / V_s)$$

En simplifiant l'expression de la probabilité conditionnelle locale en supprimant les termes communs qui font intervenir les cliques ne contenant pas le site s au numérateur et au dénominateur, on a :

$$\begin{aligned} P(X_s = x_s / X^s = x^s) &= \frac{\exp\left(-\sum_{c \in C / s \notin c} U_c(x) - U_s(x_s, V_s)\right)}{\sum_{x_s \in E} \exp\left(-\sum_{c \in C / s \notin c} U_c(x) - U_s(x_s, V_s)\right)} \\ &= \frac{\exp(-U(x_s / V_s))}{\sum_{x_s \in E} \exp(-U_s(x_s / V_s))} \end{aligned}$$

L'expression obtenue, qui ne fait intervenir que les potentiels des cliques contenant le site s (ce qui nous permet de retrouver au passage l'hypothèse markovienne), est très importante. En effet, autant il n'est pas possible partant d'une configuration x d'accéder à sa probabilité à cause de la constante de normalisation, autant il est possible de calculer en chaque site la probabilité conditionnelle locale. Cette expression sera à la base de tous les algorithmes de simulation de champs markoviens.

Annexe C

L'échantillonneur de Gibbs

Echantillonnage du MRF

Si nous résumons les résultats précédents, la définition d'un champ de Markov passe par la définition de sa fonction d'énergie U . Celle-ci nécessite la définition d'un système de voisinage, qui définit alors le système de cliques, et de fonctions de potentiel associées aux cliques. Ces fonctions de potentiel permettent d'accéder à la probabilité globale d'une configuration, et aux probabilités conditionnelles locales. Le problème qui se pose alors est, étant défini un champ de Markov, comment pouvons-nous réaliser le tirage d'une configuration (une image ici) en suivant la loi de probabilité de Gibbs caractéristique de ce champ ? Afin de synthétiser des réalisations d'un champ de Markov, nous allons décrire l'algorithme de l'échantillonneur de Gibbs.

L'échantillonneur de Gibbs

L'échantillonnage d'un champ de Markov (échantillonneur de Gibbs) est présenté dans cette annexe, ainsi que les modèles markoviens les plus courants. L'utilisation des champs markoviens en traitement d'images dans un cadre bayésien montre la nécessité de pouvoir accéder aux configurations les plus probables d'un champ markovien et nous amène à la présentation du recuit simulé.

Cet algorithme, proposé par Geman et Geman [7], repose sur la construction itérative d'une suite d'images. A la convergence, i.e. après un nombre d'itérations suffisant, les images construites sont des réalisations tirées selon la loi de Gibbs globale.

La méthode de construction de l'image à l'itération n , partant de l'image à l'itération $(n-1)$ se fait par mises à jour successives des sites de l'image. A l'étape n :

- choix d'un site s
- au site s , selon la configuration des voisins V_s pour l'image $x^{(n-1)}$, calcul de la probabilité conditionnelle locale :

$$P(X_s = x_s | V_s) = \frac{\exp(-U_s(x_s | V_s))}{\sum_{\xi \in E} \exp(-U_s(\xi | V_s))}$$

- mise à jour du site s par tirage aléatoire selon la loi $P(X_s = x_s | V_s)$.

On considère que l'algorithme a convergé après un grand nombre d'itérations ou lorsque le nombre de changements est faible. Le choix du site s considéré à l'étape n peut se faire de n'importe quelle façon à condition de balayer tous les sites un très grand nombre de fois

(théoriquement un nombre infini de fois). Les méthodes usuelles consistent à tirer un site selon une loi uniforme, ou effectuer un balayage classique, ligne par ligne, de l'image.

Cet algorithme construit en réalité une suite d'images $x(n)$ qui sont les observations d'une suite $x(n)$ de champs aléatoires constituant une chaîne de Markov pour un certain noyau de transition. On peut montrer le théorème suivant, lorsque la séquence balaye chaque site une infinité de fois: $\forall x^{(o)} \forall x \in \Omega \quad \lim_{n \rightarrow \infty} P(X^{(n)} = x \mid X^{(o)} = x^{(o)}) = P(x)$

Où P est la mesure de Gibbs associée au champ de Markov considéré. Ainsi, après un grand nombre d'itérations, les images $x(n)$ générées sont des réalisations de la loi globale $P(x)$, et cela indépendamment de la configuration initiale $x(o)$.

La preuve de ce théorème, et donc de la convergence de l'algorithme, est donnée dans [Winkler] [14]. On parle de l'échantillonneur de Gibbs comme d'un algorithme de relaxation, car il procède par mises à jour successives des sites, et probabiliste car celle-ci est fondée sur un tirage aléatoire.

Bibliographie

- [1]** Horn, B. K. P., and Schunck, B. G., 1981, Determining optical flow. *Artificial Intelligence*, v. 17, pp. 185-203.
- [2]** Lucas, B. D., and Kanade, T. 1981, An iterative image registration technique with an application to stereo vision. *Proceedings DARPA Image Understanding Workshop*, pp. 121-130
- [3]** S. Baker, I. Matthews, Lucas-Kanade. 2002, 20 Years On: A Unifying Framework, tech. report CMU-RI-TR-02-16, Robotics Institute, Carnegie Mellon University.
- [4]** Julien Marzat - INRIA - Estimation Temps Réel du Flot Optique, ISA 2008
- [5]** Florence TUPIN et Marc SIGELLE, Définition et simulation d'un champ de Markov
- [6]** Lu, S.-L., 1993, Robust estimation of motion vector fields with discontinuity and occlusion using local outliers rejection, *SPIE*, v. 2094, pp. 588-599.
- [7]** Geman, S., and Geman, D., 1984, Stochastic relaxation, Gibbs distribution, and Bayesian restoration of images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 6, pp. 721-741
- [8]** Besag, J., 1974, Spatial interaction and the statistical analysis of lattice systems, *Journal of Royal Statistical Society B*, v. 36, no. 2, pp. 192-236.
- [9]** Kirkpatrick, S., Gellat, C. D., and Vecchi, M. P., 1983, Optimization by Simulated Annealing, *Science*, v. 220, pp. 671-680.
- [10]** Konrad, J., and Dubois, E., 1991, Comparison of stochastic and deterministic solution methods in Bayesian estimation of 2D motion. *Image and Vision Computing*, v. 9, pp. 215-228.
- [11]** Konrad, J., and Dubois, E., 1992, Bayesian estimation of motion vector fields, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 14, pp. 910-927
- [12]** G. Mustaphaoui, 2005, Détection de mouvement et suivi de personnes dans les séquences d'images couleur, thèse de doctorat à l'université Pierre et Marie curie (France)
- [13]** Stephen.G et Laval.G 1996, Échantillonnage de Gibbs et autres applications économétriques des chaînes markoviennes Université de Montréal ; *L'Actualité économique*, vol. 72, n° 1, 1996, p. 27-49.
- [14]** G. Winkler (Springer, 2002), *Image Analysis, random fields and dynamic MC Methods*: Springer Verlag. Berlin

- [15]** Snyder, M. A., 1991, On the mathematical foundations of smoothness constraints for the determination of optical flow and for surface reconstruction, IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 13, pp. 1105-1.
- [16]** E. P. Simoncelli, "Bayesian multi-scale differential optical flow," in Handbook of Computer Vision and Applications, pp. 397–422, Academic Press, 1999.
- [17]** E. Thiebaut and J. M. Conan, Strict a priori constraints for maximum-likelihood blind deconvolution, J. Opt. Soc. Am. 12, 485-492 (1995).
- [18]** M. Black and P. Anandan. The robust estimation of multiple motions: parametric and piecewise smooth flow fields. CVIU, 63(1):75–104, January 1996
- [19]** Nagel, H. H., 1987, On the estimation of optical flow: Relations between different approaches and some new results. Artificial Intelligence, v. 33, pp. 299-324,
- [20]** Nakaya, Y., and Harashima, H., 1994, Motion compensation based on spatial transformations, IEEE Transactions on Circuits and Systems for Video Technology, v. 4, pp. 339-356.
- [21]** Orland, H., 1985, Mean-field theory for optimization problems, J. Physique Lett., v. 46, no. 17, pp. 163-110