

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE



UNIVERSITE MOULOUD MAMMERI, TIZI OUZOU

Faculté de Génie électrique et d'informatique

Département : Electronique

MÉMOIRE DE FIN D'ÉTUDES

En vue de l'obtention du diplôme d'ingénieur d'état en électronique

Option : Communication

THÈME :

ESTIMATION DU MOUVEMENT DANS DES SEQUENCES D'IMAGES

Application aux images METEOSAT

Réalisé par :

CHERIFI Hakima

Encadrés par :

M^r ALOUACHE Djamel

M^{me} AMEUR Zohra

2011-2012

Remerciement

Je tiens tout d'abord à remercier Dieu le tout puissant et miséricordieux, qui m'a donné la force et la patience d'accomplir ce modeste travail.

En second lieu, je tiens à remercier mon encadreur Mr : Djamel Alouache, qui s'est toujours montré à l'écoute et très disponible tout au long de la réalisation de ce mémoire, je voudrais lui témoigner ma gratitude pour son orientation, sa confiance, sa patience et son soutien et sans qui ce mémoire n'aurait jamais vu le jour. Qu'il trouve dans ce travail un hommage vivant à sa haute personnalité.

Mes vifs remerciements vont également aux membres du jury pour l'intérêt qu'ils ont porté à mon travail en acceptant de l'examiner et de l'enrichir par leurs propositions.

Mes remerciements s'étendent également à tous les professeurs qui m'ont enseigné et qui par leurs compétences m'ont soutenu dans la poursuite de mes études.

J'adresse aussi mes plus sincères remerciements à tous mes proches et amis, qui m'ont toujours soutenue et encouragée au cours de la réalisation de ce mémoire.

Enfin, je tiens à remercier également toute personne ayant contribué de près ou de loin afin de mener ce projet à terme.

Merci à tous et à toutes.

Dédicace

Ce présent travail est un résultat de grands sacrifices, de recherche, de planification, de fatigue et de patience, et c'est la fin de plusieurs années d'études, de travail et de navettes.

Je dédie ce fruit à ceux qui m'ont aidé et sans leurs soutient ce travail n'aurait jamais vu le jour à :

Allah qui m'a donné le courage et la force et qui m'a toujours orienté vers le bon sens.

Moi-même pour les efforts faits et pour les moments difficiles vécus.

*A la mémoire de mon cher grand père **CHERIFI Si Mohand Saïd**.*

A ceux qui sont la source de mon inspiration, à qui je dois de l'amour : mes chers

*parents **SI AHMED** et **OUARDIA** qui m'ont tout donné pour que je puisse suivre mes*

études dans des bonnes conditions.

*A la mémoire de ma très chère sœur **SAMIA** que dieu ait son âme.*

*A la mémoire de mes très chers oncles **SI SALAH** et **SI AHCEN** que dieu ait leurs âmes.*

*A mes frères **DJAMAL**, **HACEN** et **DJILALI**.*

*A ma sœur **OUIZA**.*

*A mon cousin **MOURAD**.*

*A tous mes amis surtout **LAMIA**, **OUIZA** pour leurs aide et leurs soutient.*

Résumé

Le thème de cette thèse est l'estimation du mouvement dans les séquences d'images, essentiellement des images METEOSAT. L'étape d'estimation consiste à déterminer le vecteur de déplacement entre deux images successives.

Nous avons étudié certains algorithmes d'estimation du mouvement qui s'inscrivent dans le cadre de la méthode du Block-Matching qui repose sur le repère d'un bloc similaire (bloc cible) au bloc source à partir d'une ou plusieurs images de référence qui peuvent apparaître future ($t+1$) ou passée ($t-1$) par rapport l'image courante (t).

Nous avons évalué les performances de ces algorithmes d'où la déduction de la bonne estimation du déplacement de l'algorithme de la recherche exhaustive par rapport aux autres algorithmes.

Son avantage : une bonne précision

Son inconvénient : il est long en temps de calcul et impraticable pour des applications en temps réel.

Mots clés

Prétraitement d'images, acquisition d'images, estimation du mouvement, METEOSAT, block-matching,

Sommaire

Introduction générale	01
 Chapitre I : Notions de base sur le traitement et acquisition d'images	
I.1 Introduction.....	02
I.2 Définition de l'image.....	02
I.3 Technologie de l'image numérique.....	03
I.3.1 Terminologie	03
I.3.2 Caractéristiques d'une image numérique.....	03
➤ Pixel.....	03
➤ Dimension.....	04
➤ Résolution.....	04
➤ Bruit	04
➤ Histogramme	05
➤ Contours.....	06
➤ Texture	06
➤ Luminance.....	06
➤ Contraste.....	07
➤ Images à niveaux de gris.....	07
➤ en couleur Images.....	08
I.3.3 Les capteurs.....	09
I.3.4 Notions d'échantillonnage et de quantification.....	10
I.3.4.1 Echantillonnage.....	10
I.3.4.2 Quantification.....	11
I.3.5 Numérisation de l'image.....	12
I.3.5.1 Le voisinage d'un pixel.....	13
I.3.5.2 Le codage binaire d'une image.....	13

Sommaire

I.3.6 Acquisition d'images.....	13
I.3.6.1 Les deux grands types d'orbites utilisés en télédétection.....	14
I.3.6.1.1 Les satellites à défilement.....	14
I.3.6.1.2 Les satellites géostationnaires.....	14
I.3.6.1.2.1 Les satellites météorologiques METEOSAT....	15
I.3.6.1.2.2 Météosat de première génération.....	16
I.3.5.1.2.3 Météosat de seconde génération.....	18
I.3.6.2 Les images de télédétection.....	19
➤ Le canal visible.....	19
➤ Le canal dans l'infrarouge thermique.....	20
➤ Le canal vapeur d'eau.....	21
➤ Composition colorée.....	22
I.4 Opérateurs de traitement d'images.....	23
I.4.1 Opérateurs image \longrightarrow image.....	23
I.4.2 Opérateurs image \longrightarrow ensemble d'informations.....	24
I.4.3 Opérateurs ensemble d'informations \longrightarrow image.....	24
I.5 Prétraitement.....	24
➤ Modification d'histogramme.....	24
• Expansion de dynamique.....	25
• Egalisation de l'histogramme.....	25
I.5.1 Application à des images de télédétection.....	25
➤ Filtrage.....	27
• Filtre médian.....	28
I.6 Conclusion.....	29
 Chapitre II : Différentes méthodes d'estimation du mouvement	
II.1 Introduction.....	30
II.2 Mouvement réel, mouvement apparent et mouvement estimé.....	30

Sommaire

II.2.1	Champ de mouvement réel et champ de mouvement apparent.....	31
II.2.2	Champ de mouvement apparent et champ de mouvement estimé.....	32
II.3	Méthodes d'estimation du mouvement.....	33
II.3.1	Méthodes différentielles.....	33
II.3.1.1	Equation de contrainte du mouvement.....	33
II.3.1.2	Méthode de Horn et Schunck	35
II.3.1.3	Méthode de Lucas et Kanade.....	36
II.3.1.4	Méthodes différentielles avancées.....	36
II.3.1.4.1	Méthodes différentielles directionnelles.....	37
II.3.1.4.2	Méthodes différentielles hiérarchiques.....	38
II.3.2	Méthodes de mise en correspondance.....	39
II.3.2.1	Méthodes de mise en correspondance dans le plan transformé...39	
II.3.2.2	Méthodes de mise en correspondance dans le plan image.....	39
II.3.3	Méthodes statistiques.....	40
II.3.4	Méthodes fréquentielles.....	40
II.4	Conclusion.....	41

Chapitre III : Algorithme du Block Matching

III.1	Introduction.....	42
III.2	Principe général du Block-Matching.....	42
III.3	Prédiction avant et arrière.....	45
III.4	Critères d'évaluation	46
III.5	Stratégies de recherche du meilleur bloc candidat.....	49
III.5.1	Algorithme de recherche exhaustive (Full Search (FS)).....	50
III.5.2	Recherche en trois pas (Three Step Search (TSS))	50
III.5.3	Recherche en quatre pas.....	51
III.5.4	Recherche selon le gradient (Gradient Search GS).....	52

Sommaire

III.5.5 Recherche en Diamant (Diamond search DS).....	52
➤ Diamant simple.....	53
➤ Predictive diamond search.....	53
III.5.6 Recherche sur une grille hexagonale (Hexagon-Based Search HBS).....	53
III.5.7 Algorithmes 2D-logarithmiques (TDL: Two Dimension Logarithm).....	54
III.6 Les méthodes avancées (Block Matching)	56
III.6.1 Estimation du mouvement subpixelique.....	56
➤ Interpolation subpixelique - principe général.....	56
➤ Méthodes d'interpolation subpixelique.....	57
• Interpolation bilinéaire.....	57
• Interpolation utilisant la norme H26L.....	59
• Interpolation utilisant les courbes de Bézier.....	59
III.6.2 Block- Matching hiérarchique (Hierarchical BM).....	59
III.7 Conclusion.....	60
Chapitre IV : Applications et discussions	
IV.1 : Introduction.....	61
IV.2 : Etude comparative.....	61
IV.3 : Interprétation des résultats.....	63
IV.4 : Application aux images synthétiques.....	63
IV.5 : Application aux images METEOSAT.....	73
IV.6 : Interprétation des résultats.....	77
IV.7 : Conclusion.....	78
Conclusion générale	79
Bibliographie	81
Annexe	82

Introduction générale

Le domaine du traitement d'images a connu ces dernières années un intérêt de plus en plus croissant et suscite par ses différentes applications la convoitise des chercheurs de différents horizons et domaines. L'introduction du traitement d'images dans différents secteurs où l'homme opère s'est avérée des plus nécessaire dans le but de substituer à l'observateur humain la machine, ou au moins l'épauler dans la réalisation de quelques tâches. Étant donné le volume énorme de données qu'un ordinateur peut traiter et en des temps record avec une efficacité déconcertante, il devient nécessaire de faire en sorte que cet outil puisse avoir des données bien formulées et exemptes de toute ambiguïté.

Dans toute la diversité des domaines d'application du traitement d'images, nous allons nous intéresser au domaine d'estimation du mouvement dans des séquences d'images, essentiellement des images METEOSAT.

L'estimation du mouvement dans les séquences d'images, est l'un des problèmes fondamentaux en traitement d'images. Elle consiste principalement à exploiter les redondances temporelles qui existent entre les images consécutives de la séquence, afin de représenter des vecteurs caractérisant les mouvements existant dans cette séquence.

Problématique :

Pour se protéger de certains phénomènes météorologiques ravageurs, il faut avoir une bonne estimation des vecteurs de déplacement des nuages dans les images METEOSAT.

Le problème posé est de trouver la méthode qui nous donne un bon déplacement d'où notre évaluation de certains algorithmes du Block-Matching.

Les critères d'évaluation utilisés sont les PSNR et l'image erreur.

Chapitre I : Notions de base sur le traitement et acquisition d'images

I.1 Introduction

Avec la parole, l'image constitue l'un des moyens les plus importants qu'utilise l'homme pour communiquer avec autrui. Chacun peut analyser l'image à sa manière, pour en dégager une impression et d'en extraire des informations précises.

C'est ainsi que, le traitement d'images se propose comme l'ensemble des méthodes et techniques opérant sur celles-ci, dans le but de rendre cette opération possible, plus simple, plus efficace et plus agréable, d'améliorer l'aspect visuel de l'image et d'en extraire des informations jugées pertinentes.

Le traitement d'images est un domaine très vaste et très utilisé dans divers secteurs, d'où l'intérêt particulier qu'il a connu pendant des décennies et qu'il connaît toujours. Les informations que l'on peut tirer en examinant une simple image, se sont révélées parfois tellement importantes et tellement riches, que beaucoup d'organismes et chercheurs se sont mis à étudier et à développer ce domaine, en vue d'une utilisation pratique et variée.

L'analyse d'une image ou d'une séquence vidéo par le cerveau humain se fait d'une manière naturelle. Le progrès remarquable qu'a connu l'outil informatique en matière de stockage et de vitesse de traitement, a fait que l'on a voulu automatiser ces traitements en imitant un modèle naturel qui n'est autre que le cerveau humain. L'aboutissement à un tel challenge a impliqué la mise en place et l'apparition de différentes méthodes et techniques de traitement, et notamment de nouvelles technologies matérielles adaptées à des besoins relatifs à des secteurs précis (météorologique, médical, militaire, sécurité, robotique...).

Dans ce chapitre nous allons introduire le traitement et acquisition d'images d'une manière générale, en donnant quelques définitions élémentaires.

I.2 Définition de l'image

L'image est un ensemble structuré d'informations qui, après l'affichage sur l'écran, ont une signification pour l'œil humain. Elle peut être décrite sous la forme d'une fonction $I(x, y)$ de brillance analogique continue, définie dans un domaine borné, tel que x et y sont les coordonnées spatiales d'un point de l'image et I est une

Chapitre I : Notions de base sur le traitement et acquisition d'images

fonction d'intensité lumineuse et de couleur. Sous cet aspect, l'image est inexploitable par la machine, ce qui nécessite sa numérisation.

L'image numérique peut être ainsi représentée par une matrice de dimension $N \times M$, N qui est le nombre de lignes et M le nombre de colonnes. Chaque élément de la matrice, appelé pixel, le plus petit élément de l'image, représente l'intensité lumineuse comprise entre 0 et 255, soit 256 niveaux de gris (pour une image noir et blanc). Le noir correspond au niveau de gris 0 et le blanc au niveau de gris 255.

I.3 Technologie de l'image numérique

I.3.1 Terminologie

L'image numérique est obtenue à partir de l'échantillonnage du signal vidéo analogique fourni par des caméras CCD (Charge Coupled Device) , des capteurs CMOS, des scanners ou échantillonneurs spécifiques. La résolution spatiale est le nombre de pixels dans l'image. La densité de résolution est le nombre de pixels par unité de longueur.

I.3.2 Caractéristiques d'une image numérique

L'image est un ensemble structuré d'informations caractérisé par les paramètres suivants :

➤ Pixel

Le pixel est le plus petit point de l'image, c'est une entité calculable qui peut recevoir une structure et une quantification. Si le bit est la plus petite unité d'information que peut traiter un ordinateur, le pixel est le plus petit élément que peuvent manipuler les matériels et logiciels d'affichage ou d'impression.

La quantité d'informations que véhicule chaque pixel donne des nuances entre images monochromes et images couleurs. Dans le cas d'une image monochrome, chaque pixel est codé sur un octet, et la taille mémoire nécessaire pour afficher une telle image est directement liée à la taille de l'image. Dans une image couleur (R, V, B), un pixel peut être représenté sur trois octets : un octet pour chacune des couleurs (R, V, B).

Chapitre I : Notions de base sur le traitement et acquisition d'images

➤ Dimension

C'est la taille de l'image. Cette dernière se présente sous forme de matrice dont les éléments sont des valeurs numériques représentatives des intensités lumineuses (pixels). Le nombre de lignes de cette matrice multiplié par le nombre de colonnes nous donne le nombre total de pixels dans une image.

➤ Résolution

C'est la clarté ou la finesse de détails atteinte par un moniteur ou une imprimante dans la production d'images. Sur les moniteurs d'ordinateurs, la résolution est exprimée en nombre de pixels par unité de mesure (pouce ou centimètre). On utilise aussi le mot résolution pour désigner le nombre total de pixels affichables horizontalement ou verticalement sur un moniteur.

➤ Bruit

Un bruit (parasite) dans une image est considéré comme un phénomène de brusques variations de l'intensité lumineuse d'un pixel par rapport à ses voisins, il provient de l'éclairage, des dispositifs optiques et électroniques du capteur.

Il existe deux types de bruits.

- **Le bruit impulsionnel (poivre et sel)**

Dans ce type de bruit, les valeurs des pixels bruités ne prennent que deux valeurs 0 et M (niveau de gris maximum représentable de l'image).

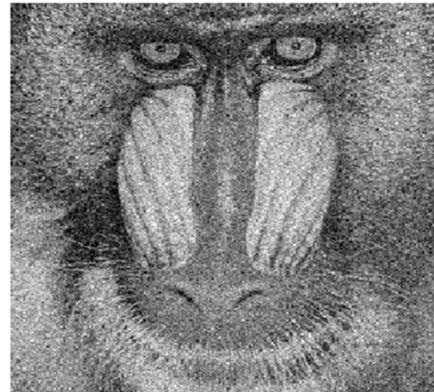
- **Le bruit gaussien**

Le bruit gaussien intervient en ajoutant à chaque pixel de l'image une valeur aléatoire à la luminance. Il est généré par des appareils d'acquisition et sa suppression reste difficile.

Chapitre I : Notions de base sur le traitement et acquisition d'images



(a) : Bruit impulsionnel



(b) : Bruit gaussien

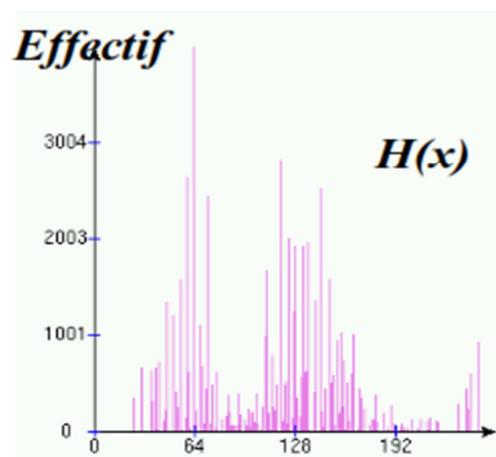
Figure 1 : exemples d'images bruitées

➤ Histogramme

Un histogramme est un graphique statistique permettant de représenter la distribution des intensités des pixels d'une image, c'est-à-dire le nombre de pixels pour chaque intensité lumineuse. Par convention, un histogramme représente le niveau de gris en abscisse en allant du plus foncé (à gauche) au plus clair (à droite). Ainsi, l'histogramme d'une image en 256 niveaux de gris sera représenté par un graphique possédant 256 valeurs en abscisses, et le nombre de pixels de l'image en ordonnées.



(a) : Image originale



(b) : L'histogramme de l'image

Figure 2 : Représentation d'une image par son histogramme

Chapitre I : Notions de base sur le traitement et acquisition d'images

➤ Contours

Les contours représentent la frontière entre les objets de l'image, ou la limite entre deux pixels dont les niveaux de gris représentent une différence significative.

➤ Texture

Dans le domaine du traitement de l'image et de la vision, il n'existe pas de définition satisfaisante de la texture. Les définitions mathématiques construites à partir de propriétés statistiques sont soit trop générales et imprécises soit trop restrictives pour s'adapter à la diversité des cas rencontrés. La définition que nous proposons n'est pas opérationnelle et prête sur ce point à critique. Elle s'appuie sur une constatation expérimentale : une texture est un champ de l'image qui apparaît comme un domaine cohérent et homogène, c'est-à-dire formant un tout pour un observateur. C'est cette propriété de cohérence et de la texture placée dans son contexte d'être perçue comme un tout homogène par l'œil humain qui sera recherchée le plus souvent par le traiteur des images , dans le but d'isoler les textures, soit pour segmenter l'image, soit pour reconnaître des régions.

➤ Luminance

C'est le degré de luminosité des points de l'image. Elle est aussi comme étant le quotient de l'intensité lumineuse d'une surface par l'aire apparente de cette surface, pour un observateur lointain, le mot luminance est substitué au mot brillance, qui correspond à l'éclat d'un objet.

Dans les deux images suivantes, seule la luminance est différente.

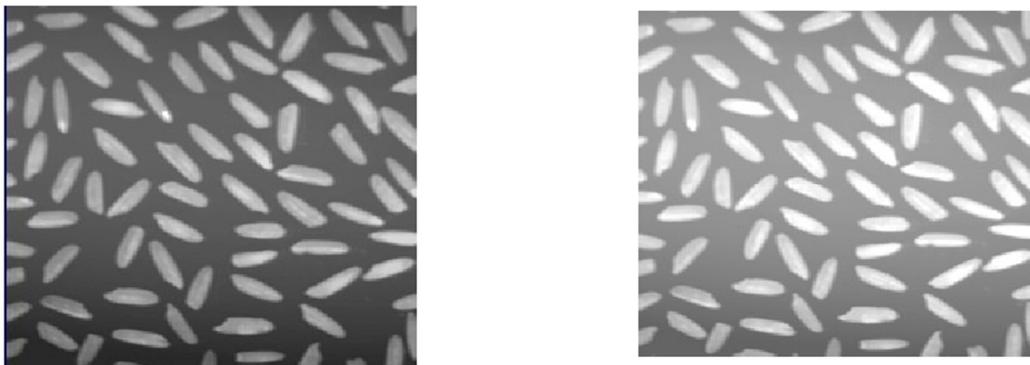


Figure 3 : exemples de luminance

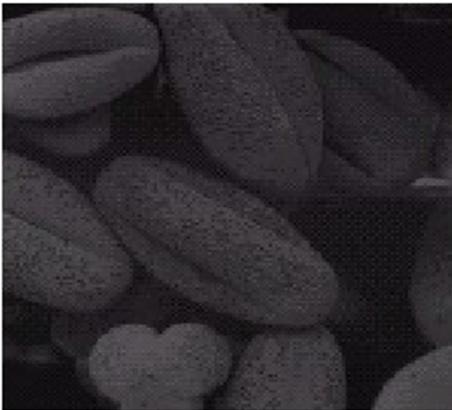
Chapitre I : Notions de base sur le traitement et acquisition d'images

➤ **Contraste**

C'est l'opposition marquée entre deux régions voisines d'une image, plus précisément entre les régions sombres et les régions claires de cette image. Le contraste est défini en fonction des variations de niveaux de gris.

$$c = \frac{L_{max} - L_{min}}{L_{max} + L_{min}}$$

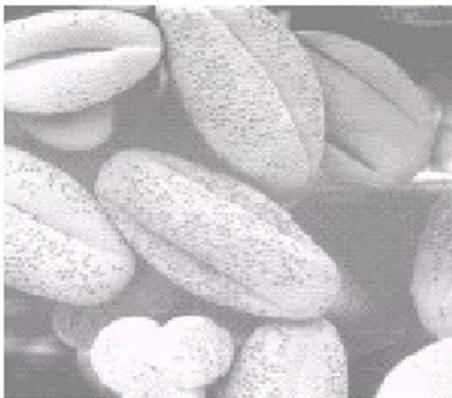
Où L_{max} et L_{min} désignent les degrés d'intensité des deux régions en question



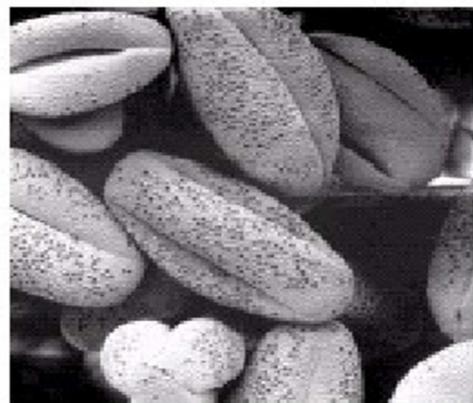
(a) : Dark image



(b) : Low contrast image



(c) : Bright image



(d) : High contrast image

Figure 4 : exemples de contrastes d'une image

➤ **Image en niveaux de gris**

Le niveau de gris est la valeur de luminosité lumineuse en un point. La couleur du pixel peut prendre des valeurs allant du noir au blanc en passant par un nombre fini

Chapitre I : Notions de base sur le traitement et acquisition d'images

de niveaux intermédiaires. Donc pour représenter les images en niveaux de gris, on peut attribuer à chaque pixel de l'image une valeur correspondante à la quantité de lumière renvoyée. Chaque pixel est codé sur huit bits et représente un niveau de gris, on peut définir 256 nuances de gris, allant du noir jusqu'au blanc (0 pour le noir, 255 pour le blanc). Pour cela, il faut que le matériel utilisé pour afficher l'image soit capable de produire les différents niveaux de gris correspondants. Le nombre de niveaux de gris dépend du nombre de bits utilisés pour décrire la « couleur » de chaque pixel de l'image. Plus ce nombre est important, plus les niveaux possibles sont nombreux.

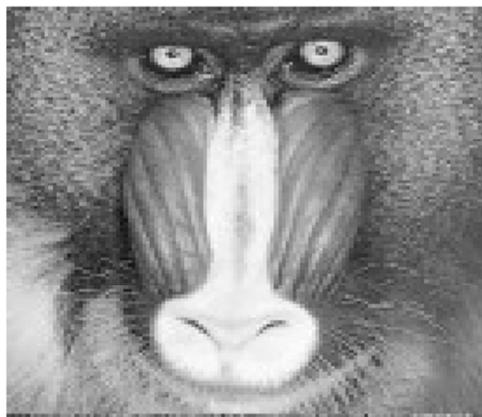


Figure 5 : image en niveau de gris

➤ Images en couleur

Même s'il est parfois utile de pouvoir représenter des images en noir et blanc, les applications multimédias utilisent le plus souvent des images couleurs. La représentation des couleurs s'effectue de la même manière que les images monochromes avec cependant quelques particularités. En effet, il faut tout d'abord choisir un modèle de représentation. On peut représenter les couleurs à l'aide de leurs composantes primaires. Les systèmes émettant de la lumière (écrans d'ordinateurs,...) sont basés sur le principe de la synthèse additive : les couleurs sont composées d'un mélange du rouge, vert et bleu (modèle R.V.B).

Chapitre I : Notions de base sur le traitement et acquisition d'images

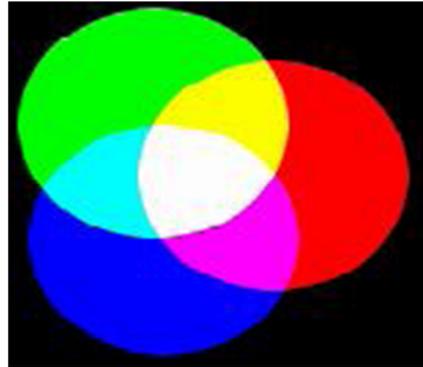


Figure 6 : Superposition des trois couleurs primaires

I.3.3 Les capteurs

Le capteur d'images est un semi-conducteur qui convertit la lumière en tension électrique en fonction du degré de luminosité. Ce capteur est constitué de cellules sensibles appelées photodiodes qui effectuent cette transformation tension-luminosité.

Pour acquérir des images en couleurs, les photodiodes sont associées à des filtres rouges, verts et bleus (RVB), chacune de ces couleurs étant échantillonnée sur 256 niveaux de luminosité (256^3). On distingue :

- Les capteurs chimiques : systèmes biologiques (œil)
- Les films photographiques
- Les capteurs photoélectriques (photodiodes, CCD : Charged Coupled Device ou Dispositif à transfert de charge)

- Les appareils numériques

D'autres capteurs existent tels que :

- IRM en imagerie médicale
- Imagerie sismique

Le signal obtenu est caractérisé par sa dimension et sa nature :

- 1D : image linéique (oscilloscope, barrette CCD,...), une seule ligne

Chapitre I : Notions de base sur le traitement et acquisition d'images

- 2D : est une image souvent plane, de nature :
 - Analogique : continue
 - Numérique : discrète

I.3.4 Notions d'échantillonnage et de quantification

L'échantillonnage et la quantification sont deux opérations permettant de numériser une image. Mais pourquoi cherche-t-on à numériser une image ? Tout simplement pour les visualiser sur un moniteur, les imprimer, les traiter sur un ordinateur, les stocker sur des supports informatiques ou les transmettre sur des réseaux informatiques.

I.3.4.1 Echantillonnage

L'échantillonnage consiste à remplacer une fonction continue dans le temps ou dans l'espace par une suite de valeurs qu'elle prend en des instants ou de zones discret(e)s périodiques. Ces valeurs suffiront pour reconstituer la fonction dans une étape ultérieure. Les valeurs du signal sont prises régulièrement à une période d'échantillonnage T_e .

Dans le cas des images qui sont des signaux 2D, l'échantillonnage concerne la discrétisation de l'espace 2D : c'est le nombre de points que l'on va pouvoir colorier.

L'échantillonnage est limité par la capacité du capteur, donc le nombre de pixels disponibles.

Effets d'échantillonnage

- Contours en marche d'escalier
- Perte de netteté
- Détails moins visibles /moins précis
- Perte de résolution



(a) : Image original



(b) : Image échantillonnée

Figure 7 : image de Léna échantillonnée

I.3.4.2 Quantification

La quantification consiste à accorder pour chaque échantillon une valeur numérique. Dans le cas des images, chaque échantillon représente un pixel, donc la quantification accorde pour chaque pixel un niveau de gris, ou bien, c'est de coder les valeurs réelles de chaque échantillon en valeurs entières de manière optimale.

La quantification concerne la discrétisation de l'espace des couleurs ou niveaux de gris : c'est le nombre de crayons de couleurs différentes que l'on va pouvoir utiliser pour dessiner notre image.

La quantification est limitée par la quantité de tons (de gris) définie dans l'intervalle.

Effets de la quantification

- Apparition de faux contours : des formes fausses qui n'existaient pas dans l'image d'origine
- Bruit de quantification

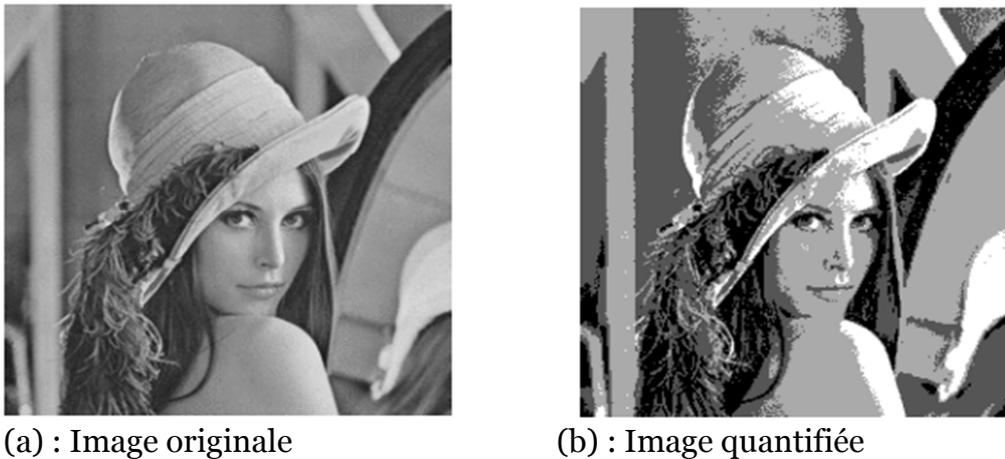


Figure 8 : image de Léna quantifiée

I.3.5 Numérisation (discrétisation) de l'image

La représentation informatique d'une image est nécessairement discrète. Le signal 2-D (deux dimensions) analogique est numérisé par :

- une discrétisation de l'espace : échantillonnage
- une discrétisation de la couleur : quantification

Lors de la numérisation, on cherche à conserver une qualité maximale (un nombre de pixels maximal et un nombre de couleurs maximal) et à obtenir des données les moins volumineuses possibles.

- L'opération d'échantillonnage exige un pavage de l'espace qui est motivé par :
 - l'adéquation au système d'acquisition.
 - des relations géométriques ou topologiques telles que le voisinage.

Quelques exemples de pavage sont représentés sur la figure ci-dessous.

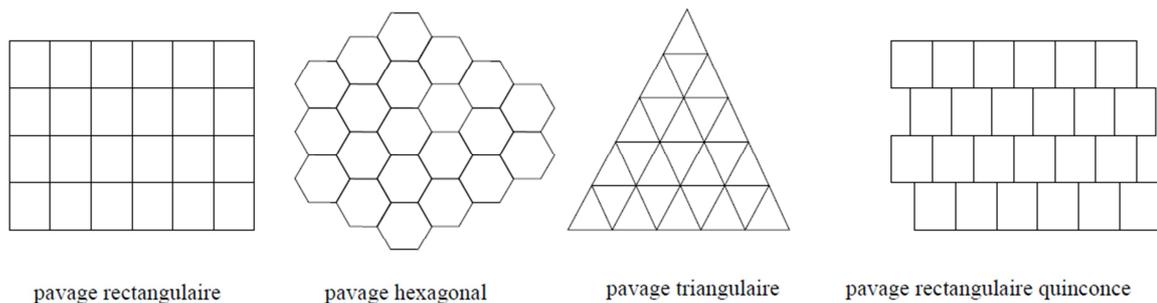


Figure 9 : les différents pavages

Chapitre I : Notions de base sur le traitement et acquisition d'images

I.3.5.1 Le voisinage d'un pixel

Selon le type de pavage, on distingue des différents voisinages qui sont définis par le nombre de pixels voisins à un pixel donné. (Figure 10)

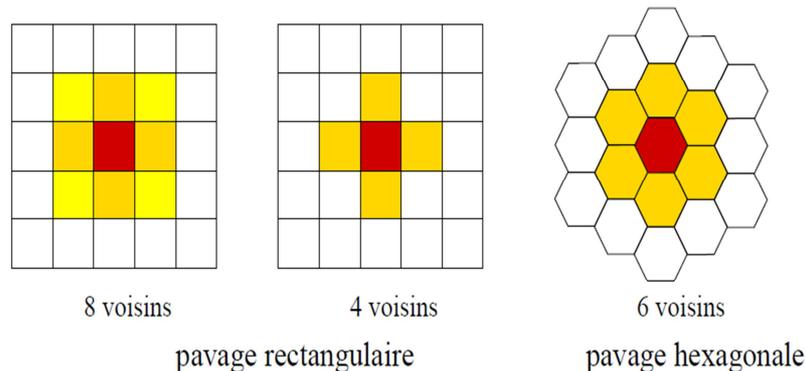


Figure 10 : le voisinage d'un pixel

I.3.5.2 Le codage binaire d'une image

Pour une image en niveau de gris, si l'on code les niveaux sur 8 bits, on pourra décrire $2^8 = 256$ valeurs entières. Pour une image couleur à k composantes, et si chaque composante est codée sur 8 bits, on aura 256^k au total.

I.3.6 Acquisition d'images

L'acquisition d'images constitue un des maillons essentiels de route chaine de conception et de production d'images. Pour pouvoir manipuler une image sur un système informatique, il est avant tout nécessaire de lui faire une transformation qui la rendra lisible et manipulable par ce système. Le passage de cet objet externe (image d'origine) à sa représentation interne (dans l'unité de traitement) se fait grâce à une procédure de numérisation. Ces systèmes de saisie, dénommés optiques, peuvent être classés en deux catégories principales : les caméras numériques et les scanners. A ce niveau, notons que le principe utilisé par le scanner est de plus en plus adapté aux domaines professionnels utilisant le traitement de l'image : la télédétection, les arts graphiques, la médecine, etc.

Chapitre I : Notions de base sur le traitement et acquisition d'images

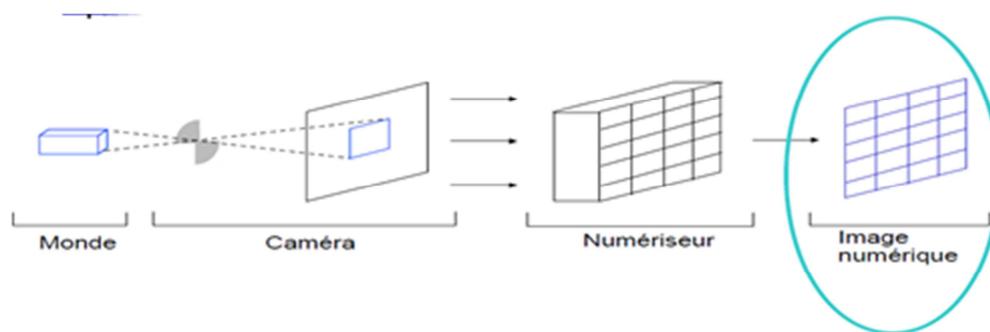


Figure 11 : l'acquisition d'une image

On va s'intéresser au domaine de télédétection, où les satellites ont permis d'atteindre un bon niveau de perception, plus précisément, les satellites météorologiques METEOSAT.

I.3.6.1 Les deux grands types d'orbite utilisés en télédétection.

I.3.6.1.1 Les satellites à défilement

L'orbite d'un satellite « à défilement » est une orbite basse (500 à 1500 km d'altitude), avec une période comprise entre 90 et 120 mn. C'est le cas de la plupart des satellites de télédétection.

I.3.6.1.2 Les satellites géostationnaires

Ces satellites regardent toujours la même région de la surface de la terre (environ un quart), c'est pour cela que nous disons qu'ils ont une orbite géostationnaire. Ils ont une altitude d'environ 36 000 Kilomètres et se déplacent à une vitesse qui correspond à celle de la terre, donnant ainsi l'impression qu'ils sont stationnaires.

Cette configuration orbitale permet au satellite d'observer et d'amasser continuellement de l'information sur une région spécifique.

Les satellites de communication et d'observation des conditions météorologiques sont situés sur de telles orbites.

L'altitude élevée de certains satellites météorologiques leur permet d'observer les nuages et les conditions qui couvrent un hémisphère complet de la terre. Grâce à leur altitude très élevée, ils se trouvent directement éclairés par le soleil, de ce fait ils

Chapitre I : Notions de base sur le traitement et acquisition d'images

n'ont pas besoin de batterie pour le stockage de l'énergie fournie par les cellules solaires (bien qu'ils se trouvent brièvement à l'ombre de la terre, une fois par jour).

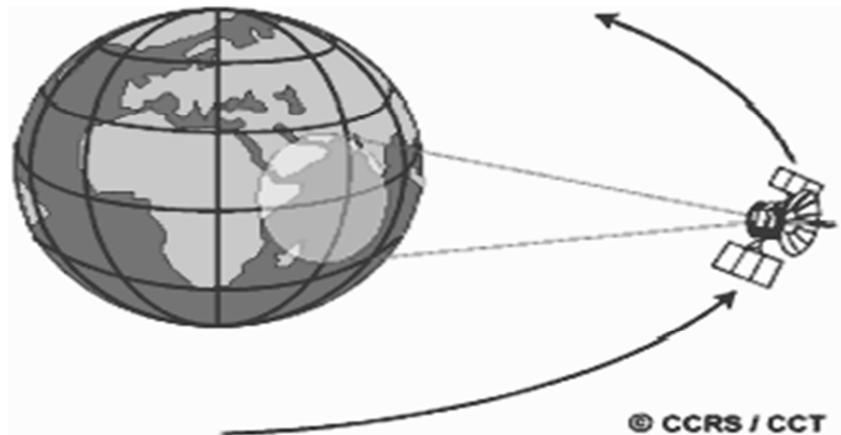


Figure 12 : Orbite et champ de vision d'un satellite géostationnaire

I.3.6.1.2.1 Les satellites météorologiques METEOSAT

METEOSAT est une famille de satellites météorologiques lancés par l'Agence spatiale européenne (ESA). Ce sont des satellites géostationnaires ils permettent donc l'observation en continu d'une zone précise du globe.

Les satellites Météosat envoient régulièrement sur terre de nombreuses images et autres données qui permettent aux météorologues, à partir des observations d'anticyclones, dépressions atmosphériques, masses nuageuses (etc.), d'élaborer les bulletins météorologiques. Ces données sont également essentielles pour anticiper certains phénomènes météorologiques ravageurs tels que les ouragans, tempêtes et pour le suivi de l'évolution du climat de la planète.

Le système opérationnel complet comprend cinq satellites géostationnaires : le satellite européen METEOSAT, les deux satellites américains GOES, un satellite japonais GMS, et un satellite russe GOMS. Tous ces satellites (Fig13) sont dans le plan équatorial, à 36 000 km de la surface de la Terre et disposés de façon à assurer une observation optimale de celle-ci.

Chapitre I : Notions de base sur le traitement et acquisition d'images

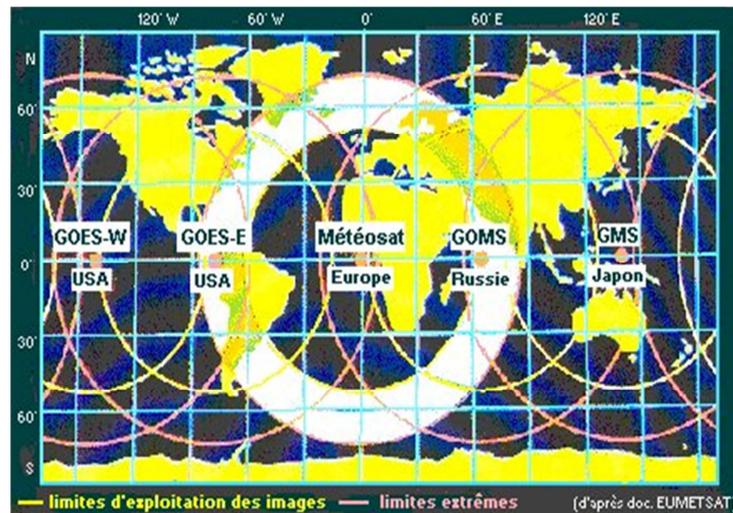


Figure 13 : Couverture des satellites météorologiques géostationnaires opérationnels [EUMETSAT].

I.3.6.1.2.2 Météosat de première génération

Le programme Météosat fait partie du système mondial d'observation de l'atmosphère, mis en place par l'Organisation météorologique mondiale au milieu des années 1970. Le premier satellite, Météosat 1, a été lancé le 23 novembre 1977. En 1995, l'EUMETSAT (organisation européenne pour l'exploitation de satellites météorologiques) a officiellement pris la responsabilité de l'exploitation des satellites Météosat et de la diffusion de leurs données.

Le radiomètre multi spectral de METEOSAT balaye, en rotation dans le sens Est-Ouest, la surface de la terre ligne par ligne par la rotation du satellite à 100 tours/minute. Chaque ligne consiste en une série d'images élémentaires ou pixel. Les lignes successives sont obtenues par décalage du télescope d'une ligne vers le nord à chaque tour du satellite.

Pour chaque pixel, le radiomètre mesure l'énergie radiative dans différentes bandes spectrales. Cette mesure est numérisée, puis transmise à une station au sol où elle est traitée, avant d'être envoyée à la communauté des utilisateurs.

Le radiomètre de METEOSAT possède 3 canaux pour l'imagerie : un canal visible, un canal dans l'infrarouge thermique et un canal vapeur d'eau.

Chaque canal de METEOSAT correspond à une bande spectrale. Le tableau ci-après résume les principales caractéristiques du radiomètre de METEOSAT.

Chapitre I : Notions de base sur le traitement et acquisition d'images

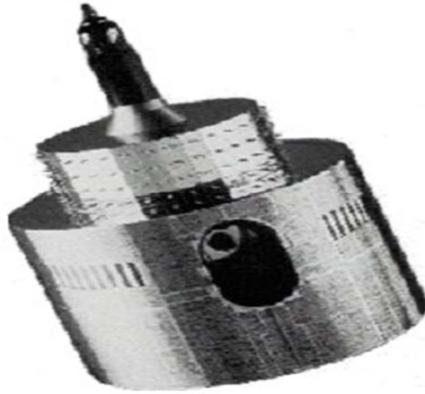


Figure 14 : Météosat première génération

canal	Visible (vis)	Vapeur d'eau (WV)	Infrarouge (IR)
Fenêtre spectrale (mm)	0,45-1,0	5,7-7,1	10,5-12,5
Nombre de détecteurs (+redondants)	2(+2)	1(+1)	1(+1)
Lignes par images	5000	2500	2500
Pixels par lignes	5000	2500	2500
Résolution spatiale au point sub-satellite	2,5 Km	5Km	5Km
Résolution temporelle	30 minutes	48 images par jour	

TAB.1 : Les principales caractéristiques du radiomètre de METEOSAT.

Chapitre I : Notions de base sur le traitement et acquisition d'images

I.3.6.1.2.3 Météosat de seconde génération (MSG)

Le 28 août 2002 a été lancé le premier satellite MSG, Météosat de Seconde Génération. Cette famille de satellites dispose de radiomètres imageurs plus sophistiqués, 12 canaux contre 3 pour la précédente génération, et fournissent des images tous les quarts d'heure, deux fois plus fréquemment.

Depuis 1977, 9 satellites de la famille Météosat ont été mis en orbite autour de la terre, dont 2 satellites MSG. Le dernier lancement en date, celui de MSG 2, a été effectué le 21 décembre 2005



Figure 15 : Météosat seconde génération

Les caractéristiques des différents capteurs du satellite MSG sont illustrées dans le tableau 2.

Canal	Bande spectrale
Visible : canal 0.6	0,56 à 0,71 μ m
canal 0.8	0,74 à 0,88 μ m
Proche infrarouge : canal 1.6	1,5 à 1,78 μ m
Infrarouge moyen : canal 3.9	3,48 à 4,36 μ m
Vapeur d'eau : canal 6.2	5,35 à 7,15 μ m
canal 7.3	6,85 à 7,85 μ m
Infrarouge : canal 8.7	8,3 à 9,1 μ m

Chapitre I : Notions de base sur le traitement et acquisition d'images

canal 9.7	9,38 à 9,94 mm
canal 10.8	9,8 à 11,8 mm
canal 12.0	11 à 13 mm
canal 13.4	12,4 à 14,4 mm
HR VIS (visible)	Haute résolution 0,56 à 0,71 mm

TAB2 : bande spectrale des douze canaux du satellite MSG

I.3.6.2 Les images de télédétection

Une image de télédétection est une représentation sur un plan de la distribution spatiale du rayonnement EM réfléchi (Electromagnétique) ou émis par des objets se trouvant à l'intérieur du champ de vue du capteur dans un intervalle de temps donné. Les variations de l'intensité du REM dans l'espace et dans une bande de longueurs d'ondes (image monochrome) sont traduites sur cette image par des variations de la teinte de gris et perçues par l'homme comme des variations de brillance. Ces variations de brillance créent la sensation visuelle du contraste, de la texture et de la forme des objets. La variation de l'énergie EM dans le spectre produit la couleur (couleur fausse dans la partie invisible du spectre EM). Lorsqu'un capteur peut acquérir plusieurs images monochromes dans différentes parties du spectre de la même scène d'une façon simultanée ou quasi-simultanée, l'ensemble des images monochromes est appelé image multi spectrale (ou multi bande).

L'imageur visible et infrarouge de METEOSAT (MVIRI : « METEOSAT Visible and Infrared Imager ») est un radiomètre haute résolution possédant quatre détecteurs (+ quatre redondants) permettant d'observer dans trois bandes spectrales ou canaux :

Un canal visible, un canal dans l'infrarouge thermique et un canal vapeur d'eau.

➤ Le canal visible

Chapitre I : Notions de base sur le traitement et acquisition d'images

Le canal visible est caractérisé par une bande spectrale de $0.5\mu\text{m}$ à $0.9\mu\text{m}$. Dans celui-ci, on balaye 5000 lignes et chaque ligne est composée de 5000 pixels. Sa résolution est de $2,5\text{km}$.

Ce canal mesure la lumière du soleil réfléchi par les divers éléments ou les divers corps qui se trouvent sous le satellite. Pour ce canal, les images ne sont évidemment utilisables en météorologie que le jour.

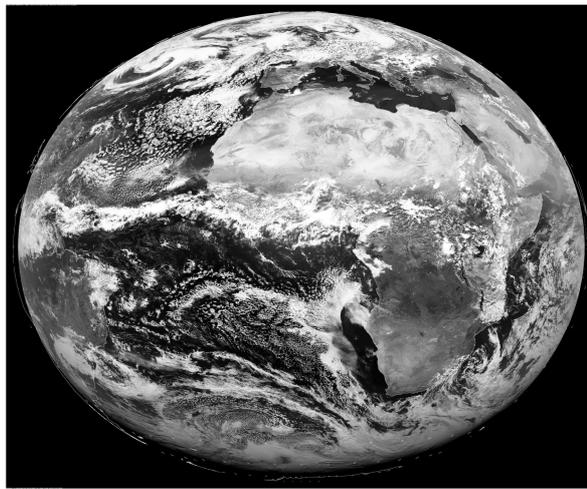


Figure 16 : image du globe dans canal visible.

➤ Le canal dans l'infrarouge thermique

Le canal infrarouge thermique est caractérisé par une bande spectrale de $10.5\mu\text{m}$ à $12.5\mu\text{m}$. Le balayage se fait sur 2500 lignes et chaque ligne est composée de 2500 pixels. Sa résolution est de 5 km .

Il mesure le rayonnement thermique émis par les nuages, par la surface du sol en absence des nuages, et même par les objets qualifiés d'objets froids. L'intérêt de la prise d'image dans ce canal est qu'on peut les obtenir aussi bien la nuit que le jour.

Les objets qui ont des températures très élevées apparaissent en gris sombre et même noir alors que les objets froids apparaissent en claire ou en blanc.

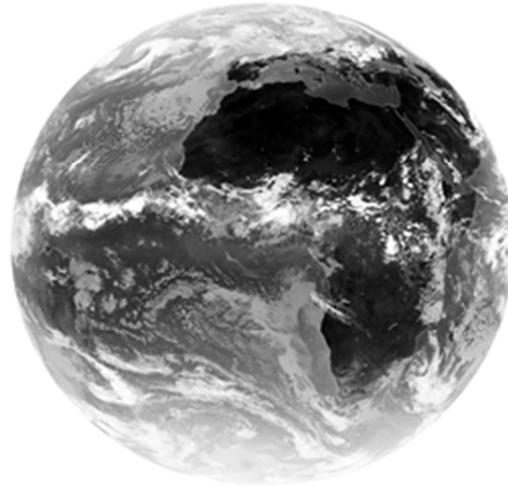


Figure 17 : image du globe dans canal Infrarouge.

➤ Le canal vapeur d'eau

Il s'agit d'une bande d'absorption de rayonnement thermique par la vapeur d'eau. L'énergie mesurée depuis le satellite correspond au rayonnement émis dans cette bande d'absorption. En ciel clair, on ne voit pas la surface terrestre sur les images WV puisque le rayonnement centré sur $6.3 \mu\text{m}$ est complètement absorbé par les basses couches de l'atmosphère, où la concentration en vapeur d'eau est maximum. En présence de nuages épais, les radiances mesurées dans le canal WV sont modulées par la vapeur d'eau présente au-dessus du nuage, et dépendent, comme pour le canal IR thermique, de la température au sommet de nuage.

Dans le spectre visible, la luminance reçue ne dépend que de la réflectivité du nuage. Cette réflectivité dépend de la composition du nuage (gouttelette d'eau, cristaux), et de son épaisseur. Ainsi, des nuages constitués de cristaux de glace épais, apparaîtront blancs. Dans l'infrarouge thermique plus un nuage a un sommet élevé, plus la température du sommet est basse. Le radiomètre mesure alors l'émittance, qui est ensuite convertie en luminance, et éventuellement en température. Ainsi les images infrarouges peuvent permettre la détermination des types (ou genres) des nuages grâce à la séparation en différents niveaux d'altitude et ce, grâce à la température.

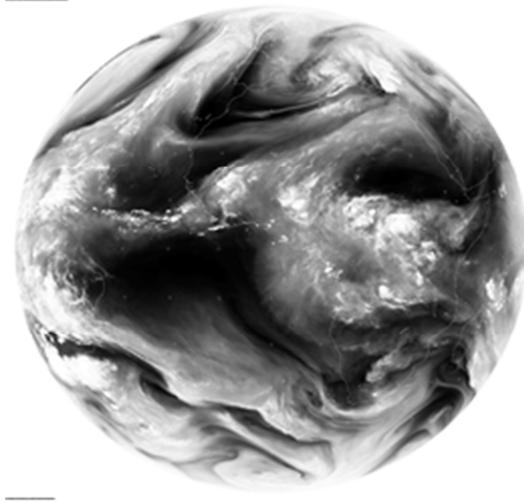


Figure 18 : image du globe dans canal Vapeur d'eau (WV).

➤ Composition colorée

L'emploi conjoint des canaux VIS et IR pour réaliser la composition colorée permet de mettre en valeur diverses formations tant au niveau des sols que des nuages. L'affectation des valeurs numériques des pixels des canaux VIS, IR et du pseudo canal Différence VIS –IR aux 3 plans de couleurs déterminera, selon qu'elles soient basses ou élevées, la couleur prise par chaque pixel dans la composition colorée (Fig.19).

Par exemple, l'eau des océans a un faible albédo : le plan vert sera foncé ; elle absorbe beaucoup les infrarouges, le plan bleu sera foncé ; la différence entre les 2 canaux sera faible, le plan rouge sera aussi foncé : il en résulte que la teinte prise par l'eau dans la composition colorée sera donc foncée. Les variations de couleurs sont telles qu'il devient possible d'identifier plusieurs types de nuages et de dresser une néphanalyse, c'est-à-dire une cartographie de ces nuages.

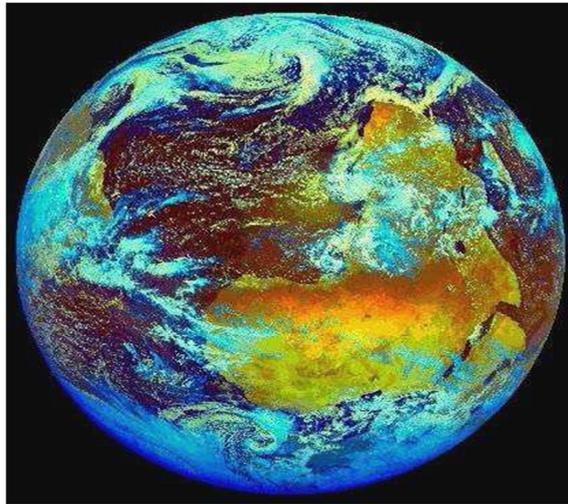


Figure 19 : Image du globe en composition colorée

I.4 Opérateurs de traitement d'images

Par analogie avec les opérateurs mathématiques, les opérateurs de traitement d'image sont des opérateurs plus au moins complexes, prenant en entrée une image en sortie. On classe généralement les opérateurs en différentes familles, en fonction des informations qu'ils acceptent en entrée et qu'ils fournissent en sortie, et en fonction des transformations qu'ils font subir aux données. Ainsi, par exemple, on distingue :

I.4.1 opérateurs image \longrightarrow image

Parmi ces opérateurs, on peut citer :

- Opérateurs de modifications pixel à pixel (opérateur point à point) : qui font le changement de la dynamique de l'image, opérateurs binaires pixel à pixel (et, ou, xor, etc.)

- Opérateurs locaux (traitent les pixels en fonction de leur voisinage) : opérateur de flou, opérateurs morphologique (érosion, dilatation, squelette), opérateurs de détection de contours.

- Opérateurs dans l'espace fréquentiel : opérateurs de réduction de bruit, filtre passe bande (souvent utilisé en première approche pour améliorer l'image, on les appelle alors des opérateurs de prétraitement.

Chapitre I : Notions de base sur le traitement et acquisition d'images

-Opérateurs globaux : calcul des distances.

I.4.2 Opérateurs image \longrightarrow ensemble d'informations

- Opérateurs de segmentation en frontières, en régions ;
- Opérateurs de classification de pixels ;
- Opérateurs de calcul de paramètres ;

I.4.3 Opérateurs ensemble d'informations \longrightarrow image

Ce sont des constructeurs d'image à partir d'une carte de région ou d'une liste de frontière (synthèse d'image)

I.5 Prétraitement

Le prétraitement a pour but de faciliter l'analyse d'une image en renforçant la ressemblance entre pixels appartenant à une même région, ou en accentuant la dissemblance entre pixels appartenant à des régions différentes.

Le prétraitement est nécessaire notamment dans les cas suivants :

- l'éclairage de l'image n'est pas uniforme
- l'image est bruitée
- le contraste n'est pas suffisant

Les méthodes les plus utilisées sont :

- La modification de l'histogramme.
- Le rehaussement de contraste.
- La réduction de bruit.

➤ Modification d'histogramme

On cherche à améliorer l'image en lui appliquant une transformation ponctuelle d'intensité. A tout pixel d'intensité a_s on associe une intensité $a_s' = T(a)$. La transformation T est choisie croissante, de façon à conserver les contrastes relatifs entre régions.

Chapitre I : Notions de base sur le traitement et acquisition d'images

- **Expansion de dynamique :**

Cette méthode consiste à utiliser au mieux l'échelle des niveaux de gris disponibles sur le système d'acquisition d'image.

Soient $A [i, j]$ l'image de départ et $A' [i, j]$ l'image après transformation. Soient $[a_0, a_1]$ l'intervalle des intensités présentes dans l'image et $[a_{\min}, a_{\max}]$ l'intervalle disponible.

L'expansion de dynamique correspond à la transformation linéaire T suivante :

$$A_s' = \alpha + \beta a_s \quad \text{telle que :}$$

$$\forall a \in [a_0, a_1] \quad a \xrightarrow{T} a' \in [a_{\min}, a_{\max}] \quad \text{Avec :}$$

$$\alpha = (a_{\min} \cdot a_1 - a_{\max} \cdot a_0) / (a_1 - a_0) ; \quad \beta = (a_{\max} - a_{\min}) / (a_1 - a_0)$$

- **Égalisation de l'histogramme :**

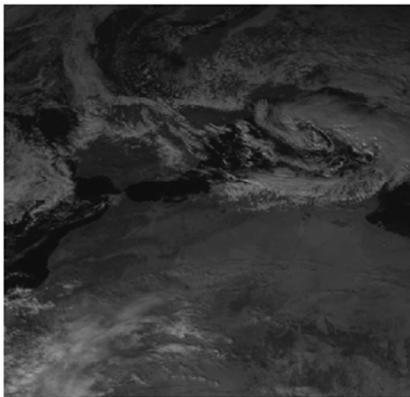
Dans une image, une majorité des pixels ont une valeur inférieure à la lumière moyenne, c'est pourquoi les détails dans les régions sombres sont difficilement perceptibles. Une technique appelée « égalisation de l'histogramme » est utilisée pour pallier à cet inconvénient. Cette transformation consiste à rendre le plus plat possible l'histogramme des niveaux de gris, ce qui donne une meilleure dynamique et un fort contraste à l'image.

I.5.1 Application à des images de télédétection

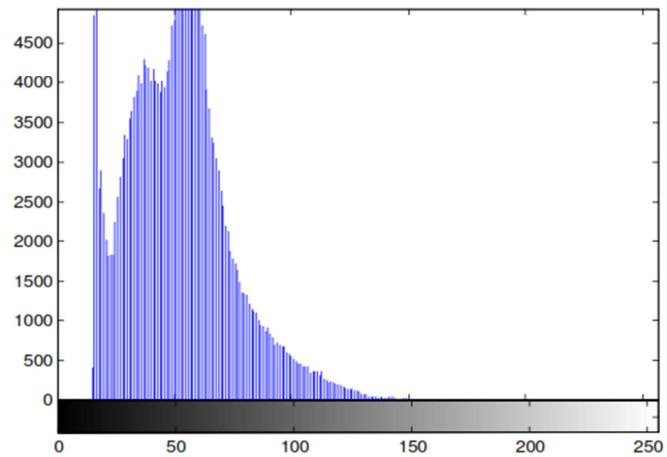
Pour les images Météosat, on peut utiliser indifféremment ces méthodes car, dans le cas des images visibles, les prises de vue se font généralement dans des conditions de faible luminosité et le niveau de gris maximum ne dépasse pas la valeur de 128. Ainsi, les images visible et infrarouge du 16.12.1994 1, ont été rehaussées par égalisation d'histogrammes. Le résultat de ce rehaussement est donné par la figure 20. On trouve que la technique d'égalisation d'histogramme se prête bien au rehaussement des images faiblement contrastées. Mais, lorsque les images sont bien contrastées, on observe une saturation des niveaux de gris dans les régions claires et une dégradation

Chapitre I : Notions de base sur le traitement et acquisition d'images

des contours entre objets clairs. Dans ce cas, les deux autres méthodes de rehaussement sont plus performantes dans la mesure où elles permettent de régler le rehaussement et la finesse des détails de l'image. En outre, grâce à ces réglages, on n'a plus le problème de la saturation des niveaux de gris.



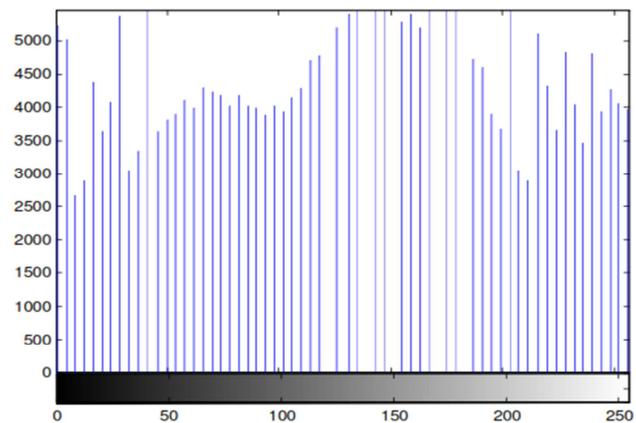
(a): Image du canal visible



Histogramme de (a)

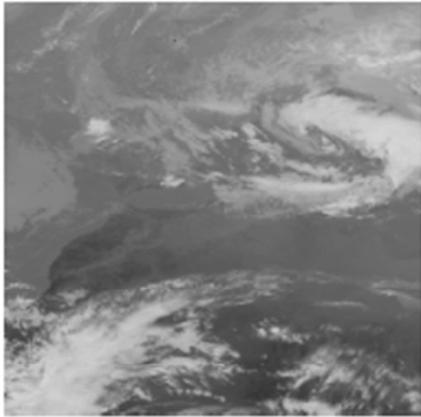


(b) : Image (a) rehaussée

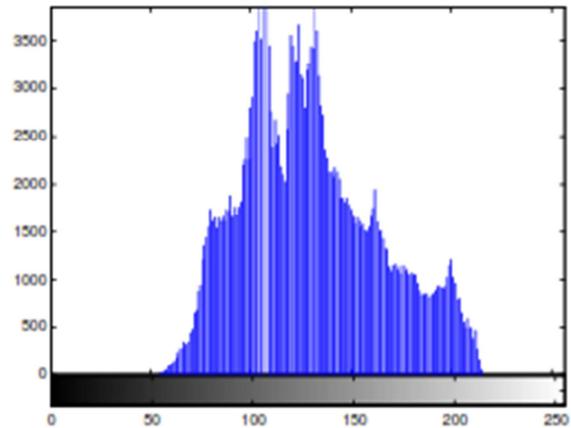


Histogramme de (b)

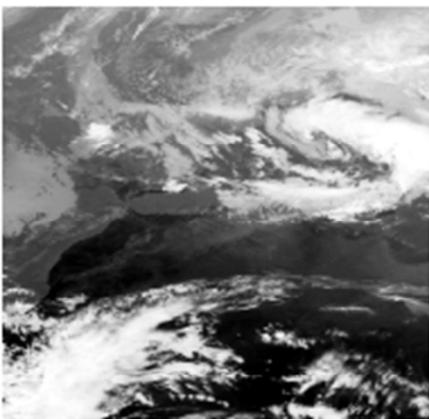
Chapitre I : Notions de base sur le traitement et acquisition d'images



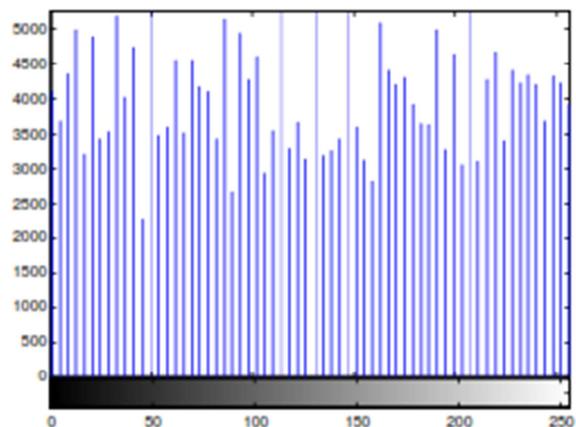
(c) : Image du canal infrarouge



Histogramme de (c)



(d) : Image (c) rehaussée



Histogramme de (d)

Figure20 : exemple de rehaussement d'images par égalisation de l'historgramme.

➤ Filtrage :

L'amélioration de l'image est obtenue essentiellement par une opération de filtrage, qui consiste soit à atténuer les bruits indésirables ou à accentuer les contours. Il existe un grand nombre de filtres parmi lesquelles on peut citer :

- Les filtres linéaires.
- Les filtres morphologiques.
- Les filtres adaptatifs.

Chapitre I : Notions de base sur le traitement et acquisition d'images

L'image satellitaire est souvent accompagnée d'un bruit impulsionnel, dû essentiellement, aux conditions de transmission depuis le satellite. Ce bruit apparaît sous forme de points dispersés aléatoirement sur toute l'image en prenant deux niveaux de gris 0 ou 255. Ce genre de bruit est efficacement éliminé par le filtre médian.

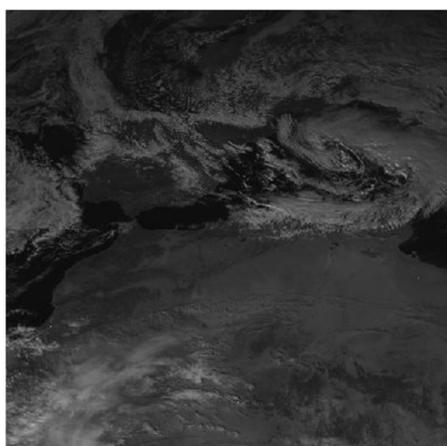
- **Filtre médian :**

C'est un filtre qui prend compte de l'information spatiale de chaque point de l'image.

Les valeurs des niveaux de gris des pixels voisins sont ensuite triées selon un ordre. La figure 21 illustre le résultat obtenu par ce type de filtre.

Exemple :

$$\begin{pmatrix} 1 & 1 & 2 \\ 2 & 3 & 1 \\ 3 & 4 & 2 \end{pmatrix} \rightarrow 1 \ 1 \ 1 \ 2 \ 2 \ 3 \ 3 \ 4 ; \text{ Dans ce cas, la valeur du point Central devra être mise à 2.}$$



(e) : Image du canal visible

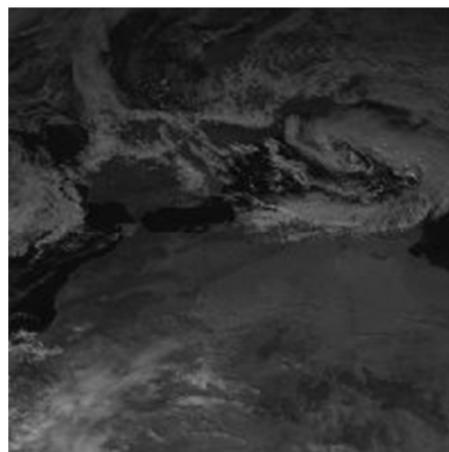
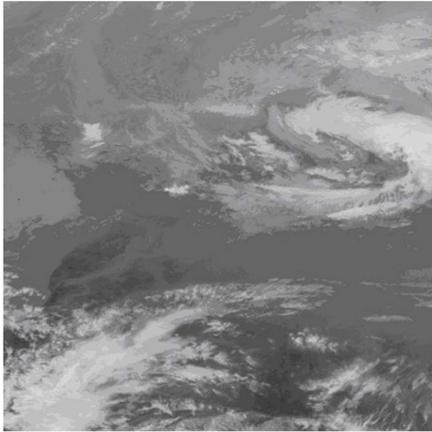


Image (e) filtrée avec le filtre médian.

Chapitre I : Notions de base sur le traitement et acquisition d'images



(f) : Image du canal infrarouge

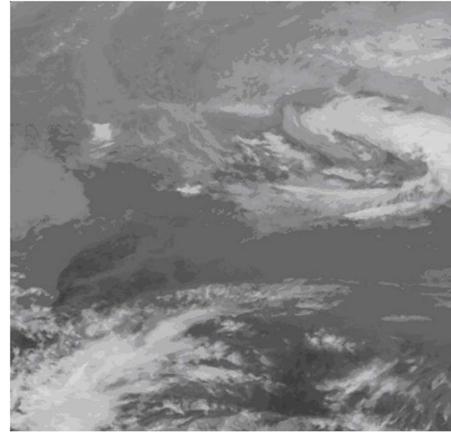


Image (f) filtrée avec le filtre médian.

Figure 21 : Exemple de filtrage avec le filtre médian.

I.7 Conclusion

Dans ce chapitre nous avons présenté les concepts fondamentaux en traitement d'images. Nous avons commencé par donner une définition du traitement d'images.

Par la suite nous avons donné les définitions les plus pertinentes en traitement d'images. Nous avons défini l'image numérique, et donné les principales caractéristiques manipulées pour son traitement (l'histogramme, filtrage...).

Nous avons également présenté les opérations les plus utilisées dans ce domaine du traitement d'images, dont le prétraitement et l'acquisition d'images (essentiellement METEOSAT) en s'appuyant sur des exemples illustratifs.

II.1 Introduction

L'estimation du mouvement dans des séquences temporelles d'images bidimensionnelles et tridimensionnelles est un des problèmes fondamentaux en traitement d'images. Les applications sont nombreuses et parmi les domaines d'application concernés nous pouvons mentionner : la compression d'images utilisant l'information du mouvement, la robotique, la météorologie avec le suivi de masses nuageuses, la médecine avec l'estimation du mouvement d'organes mobiles (comme le cœur ou les poumons), etc.

Les images représentent très souvent la projection de scènes réelles 3D. Pour cette raison, nous pouvons identifier trois types de mouvement : le mouvement réel, le mouvement apparent et le mouvement estimé. Le mouvement apparent, observé à partir des changements de la distribution spatiale d'intensité lumineuse, est aussi appelé le flux optique. Ce mouvement apparent est très souvent différent du mouvement réel et représente en général la projection du mouvement réel 3D dans le plan image.

Afin d'obtenir le mouvement estimé, plusieurs types de méthodes existent. Nous les avons classées dans ce mémoire en quatre catégories : méthodes différentielles, méthodes de mise en correspondance de blocs, méthodes statistiques et méthodes fréquentielles.

II.2 Mouvement réel, mouvement apparent et mouvement estimé

Les images représentent souvent la projection de scènes réelles 3D. C'est pourquoi le mouvement observé (ou mouvement apparent) dans une séquence temporelle d'images représente le plus souvent la projection du mouvement 3D dans le plan image. Une image peut également représenter une coupe d'un objet 3D. Dans ce cas, le champ de mouvement observé à l'intérieur de cette coupe pourra représenter à la fois un déplacement dans le plan image et également la projection sur ce plan d'un mouvement 3D. Le but de l'estimation du mouvement est d'estimer le champ de mouvement 2D ou 3D à partir d'une séquence temporelle d'images 2D ou 3D dont le contenu varie en fonction du temps. On doit différencier:

- le mouvement réel;

- le mouvement apparent (ou observé);
- le mouvement estimé.

II.2.1 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 champ de « mouvement apparent » ou « flux optique » qui, en général, est différent du champ réel de 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.

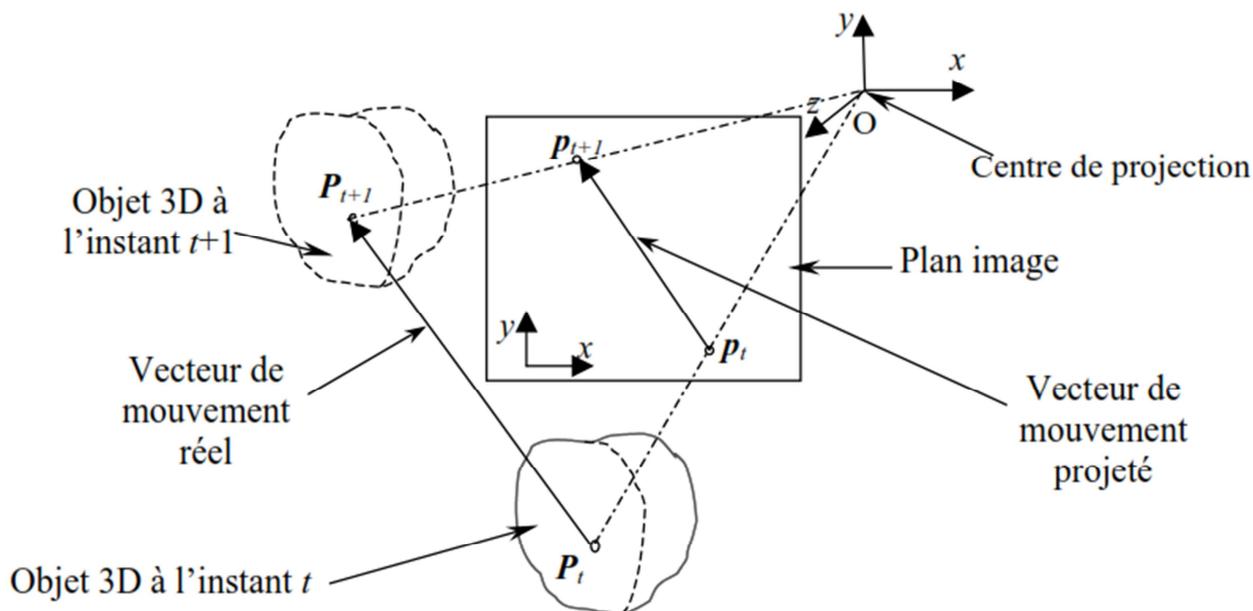


Figure 22 : Illustration des mouvements réel et apparent, dans un système optique de prise de vues

On suppose que le point P_t d'un objet réel 3D, à l'instant t , devient le point P_{t+1} à l'instant $t+1$. La projection perspective des points P_t et P_{t+1} dans le plan image est notée p_t et p_{t+1} respectivement avec $I_t(x, y)$ et $I_{t+1}(x, y)$ leur intensité. Sur la figure 2.1 on note que le déplacement apparent $\overline{p_t p_{t+1}}$ correspond à la projection perspective du mouvement réel $\overline{P_t P_{t+1}}$. Tous les vecteurs déplacements dont les extrémités appartiennent aux droites pointillées correspondant à l'opération de projection ont le même vecteur 2D déplacement apparent (ou projeté).

II.2.2 Champ de mouvement apparent et champ de mouvement estimé

Le vecteur déplacement estimé $d(p) = (d_x(p), d_y(p))$ correspondant au déplacement du Point $p = (x, y)$ du plan image, est déterminé à partir du champ de mouvement apparent, c'est-à-dire des variations locales d'intensité lumineuse $I_t(x, y)$ entre les instants t et $t+\Delta t$, où Δt est la distance en temps inter-images, dans le cas continu. Le vecteur vitesse estimé $v(v_x(p), v_y(p))$ est défini comme la variation temporelle du déplacement par unité de temps, $(v_x, v_y) = (d/dt, d/dt)$. Ceci explique pourquoi une séquence temporelle d'images ne permet que d'estimer le champ de mouvement (déplacement ou vitesse) apparent observable dans la séquence, et non le champ de vitesse réel. On nomme champ de déplacement (respectivement de vitesse), le champ de vecteurs déplacement (respectivement de vitesse) estimé avec une méthode quelconque.

En optique par exemple, la différence entre le champ de vitesse estimé et le mouvement apparent, peut exister si:

-Le gradient spatial d'intensité est trop faible. Pour que le mouvement réel soit observable, il faut que la variation de niveaux de gris (ou des couleurs) soit suffisamment grande dans les régions où il y a mouvement. L'exemple classique qui met en évidence l'importance de ce facteur, est celui d'une sphère ayant une distribution uniforme d'intensité (éclairage homogène), qui effectue un mouvement de rotation autour de son axe propre, dans une scène à éclairage constant. Même si en réalité il y a mouvement, celui-ci ne peut être observé à partir d'une séquence temporelle d'images à niveaux de gris, parce qu'aucune variation apparente de niveaux de gris n'intervient pendant le mouvement.

-L'illumination de la scène varie. Un mouvement observable dans une séquence d'images d'intensité ne correspond pas toujours à un mouvement réel. La même sphère immobile soumise à une illumination variable au cours de la séquence génère un mouvement apparent (c'est-à-dire une variation d'intensité) artificiel.

II.3 Méthodes d'estimation du mouvement

II.3.1 Méthodes différentielles

Les méthodes différentielles sont les méthodes basées sur les gradients spatiaux et temporeux d'intensité lumineuse. Ces gradients sont approchés, dans le cas discret, par des différences finies.

II.3.1.1 Equation de contrainte du mouvement

L'hypothèse d'invariance en temps de l'intensité lumineuse le long de la trajectoire du mouvement est exprimée par l'équation DFD (2.1), qui est :

$$DFD = I(x + d_x, y + d_y, t + \Delta t) - I(x, y, t) = 0 \quad (\text{II.1})$$

En réécrivant cette équation, on obtient :

$$\frac{dI(x,y,t)}{dt} = 0 \quad (\text{II.2})$$

Où x et y varient en temps, le long de la trajectoire du mouvement. Sous l'hypothèse de différentiabilité spatio-temporelle de l'intensité lumineuse et en utilisant les règles de différentiation connues, on obtient:

$$I(x + d_x, y + d_y, t + \Delta t) = I(x, y, t) + \frac{\partial I(x,y,t)}{\partial x} d_x + \frac{\partial I(x,y,t)}{\partial y} d_y + \frac{\partial I(x,y,t)}{\partial t} \Delta t + \text{tos} \quad (\text{II.3})$$

Où tos sont des termes d'ordre supérieur. En remplaçant ce développement dans l'équation de DFD est en négligeant les termes d'ordre supérieur, on obtient :

$$DFD(x, y, t) = \frac{\partial I(x, y, t)}{\partial x} d_x + \frac{\partial I(x, y, t)}{\partial y} d_y + \frac{\partial I(x, y, t)}{\partial t} \Delta t = 0 \quad (\text{II.4})$$

Où : $v_x(x, y, t) = d_x/\Delta t$, $v_y(x, y, t) = d_y/\Delta t$ sont les deux composantes de la vitesse.

CHAPITRE II : Différentes méthodes d'estimation du mouvement

L'équation ECM peut être réécrite, en chaque point (x, y, t) :

$$I^x \cdot v_x + I^y \cdot v_y + I^t = 0 \quad \text{ou: } \langle \nabla I, v \rangle + \frac{dI}{dt} = 0 \quad (\text{II.5})$$

Où $\nabla I = \left(\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right) = (I^x, I^y)$ est le gradient spatial de l'image avec les deux composantes

$I^x = \frac{\partial I}{\partial x}$ et $I^y = \frac{\partial I}{\partial y}$, respectivement $I^t = \frac{\partial I}{\partial t}$ est le gradient temporel de l'intensité, et

$\langle \cdot, \cdot \rangle$ représente le produit scalaire entre deux variables. Dans le cas discret, les gradients spatiaux et temporels de l'image sont approchés par des différences finies.

Comme on peut l'observer, l'équation de contrainte du mouvement n'est pas suffisante pour déterminer d'une manière unique le champ de vitesses, parce que l'ECM représente une équation à deux inconnues (v_x et v_y), en chaque pixel. La composante normale, $v_{\perp}(x, y, t)$ du vecteur vitesse, orientée dans la direction du gradient spatial d'image $\frac{\nabla I}{\|\nabla I\|}$, est accessible. En effet, la composante tangentielle, tangente aux frontières photométriques des objets, disparaît suite au produit scalaire. La résolution du problème d'estimation du mouvement nécessite d'introduire une contrainte supplémentaire.

L'équation de contrainte du mouvement (ECM) décrit une droite dans l'espace de vitesses (v_x, v_y) comme sur la figure II.2

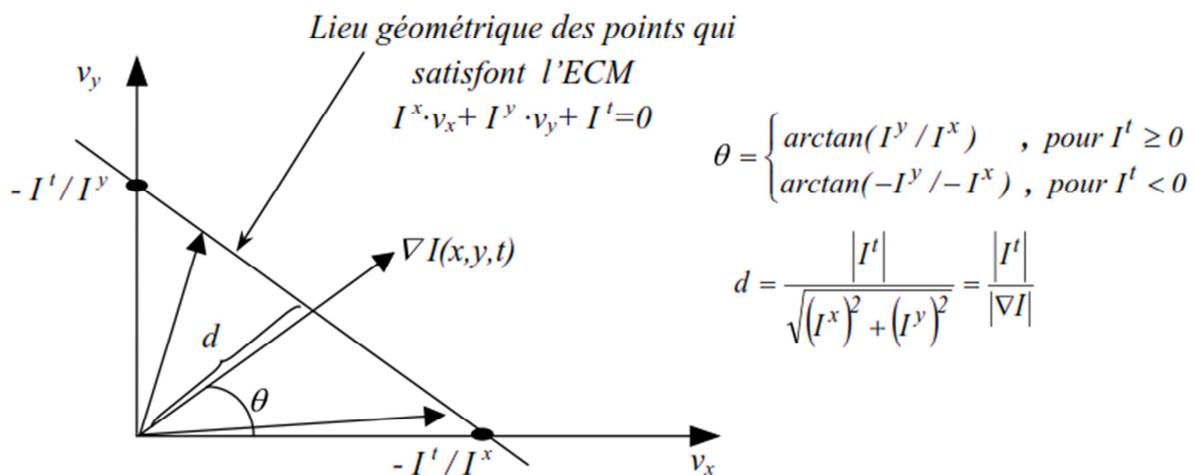


Figure 23 : Illustration géométrique de l'équation et de la droite de contrainte du mouvement.

La distance qui sépare la droite de contrainte du mouvement d'origine dépend du rapport des modules des gradients spatiaux et temporeux, sachant que l'orientation de cette droite est déterminée par l'orientation du gradient spatial au point d'intérêt.

II.3.1.2. Méthode de Horn et Schunck

Cette méthode a été proposée en 1980 par Horn et Schunck en introduisant une deuxième contrainte de lissage du champ de mouvement afin de résoudre l'équation de contrainte du mouvement.

La méthode de Horn et Schunck n'est pas limitée aux translations. C'est une méthode différentielle itérative adaptée à l'estimation des petits déplacements et basée sur le développement en série de Taylor du DFD, des gradients spatiaux et temporeux. Le but est de trouver le champ de vecteurs qui satisfait l'équation de contrainte du mouvement en chaque pixel. Soit:

$$\xi_{flux} = I_x \cdot v_x + I_y \cdot v_y + I^t \quad (II.6)$$

L'erreur par rapport à l'équation de contrainte du mouvement (ECM), en chaque point (x, y, t) . On observe que l'ECM est satisfaite quand l'erreur ξ_{flux} est nulle. Dans la méthode de Horn et Schunck on minimise le carré de l'erreur ξ_{flux} . Une contrainte supplémentaire de lissage qui suppose que tous les points voisins ont un mouvement semblable est utilisée:

$$\begin{aligned} \xi_{uniformité}^2 &= \|\nabla v_x\|^2 + \|\nabla v_y\|^2 = \left(\frac{\partial v_x}{\partial x}\right)^2 + \left(\frac{\partial v_x}{\partial y}\right)^2 + \left(\frac{\partial v_y}{\partial x}\right)^2 + \left(\frac{\partial v_y}{\partial y}\right)^2 = (v_x^x)^2 + (v_x^y)^2 + \\ &(v_y^x)^2 + (v_y^y)^2 \end{aligned} \quad (II.7)$$

On peut montrer que l'erreur $\xi_{uniformité}^2$ est faible si le champ de vecteurs vitesse est lisse. La méthode de Horn et Schunck consiste donc à minimiser une énergie de la forme :

$$\xi^2 = (1 - \gamma) \cdot \xi_{flux}^2 + \gamma \cdot \xi_{uniformité}^2 \quad (II.8)$$

Où : γ est un coefficient de pondération entre les termes d'attaches aux données et de lissage.

Pour minimiser (II.8) on utilise les équations d'Euler-Lagrange qui conduisent aux relations suivantes :

$$\begin{cases} (I^x)^2 \cdot v_x + I^x \cdot I^y \cdot v_y = \gamma \cdot (\bar{v}_x - v_x) - I^x \cdot I^t \\ (I^y)^2 \cdot v_y + I^x \cdot I^y \cdot v_x = \gamma \cdot (\bar{v}_y - v_y) - I^y \cdot I^t \end{cases} \quad (\text{II.9})$$

Où \bar{v}_x (resp. \bar{v}_y) est la valeur motenne locale de v_x (resp. v_y)

A partir de ce système on obtient v_x et v_y :

$$v_x = \frac{\Delta v_x}{\Delta} \text{ et } v_y = \frac{\Delta v_y}{\Delta}$$

Ainsi, en parcourant l'image entière, on obtient pour chaque pixel le couple de valeurs (v_x, v_y) . Si on effectue plusieurs itérations (le nombre des itérations étant le deuxième paramètre de la méthode de Horn et Schunck), ces valeurs vont diminuer, \bar{v}_x et \bar{v}_y étant des fractions de v_x et v_y .

La convergence est obtenue quand (c'est-à-dire la différence entre deux itérations successives de v_x et v_y) est inférieure à un seuil choisis ou pour un nombre (n) d'itérations choisi.

II.3.1.3 Méthode de Lucas et Kanade

La méthode de Lucas et Kanade (1986) est une autre méthode différentielle classique qui repose elle aussi sur l'hypothèse de l'équation de contrainte du mouvement (ECM). Le champ de vitesse est calculé en minimisant au sens des moindres carrés la fonction définie par :

$$E(x, y) = \sum_{(x,y) \in \Omega} W^2(x, y) [\nabla I(x, y, t) \cdot V + I_t(x, y, t)]^2 \quad (\text{II.10})$$

Ω est le voisinage spatial du pixel de coordonnées (x, y) , dans lequel tous les pixels sont soumis à l'hypothèse du mouvement constant. Le terme $W(x, y)$ représente une fenêtre de pondération qui privilégie d'avantage les pixels situés au centre du voisinage.

II.3.1.4 Méthodes différentielles avancées

La méthode de Horn et Schunck impose l'équation de contrainte du mouvement (2.5) et la contrainte de lissage du champ de vitesse au niveau global (II.7), dans l'image entière ou dans la fenêtre d'estimation du mouvement. Cette globalité spécifique aux méthodes non adaptatives a au moins deux effets non souhaités:

- l'effet de "flou" sur les frontières du mouvement estimé, dû à la contrainte de lissage (II.7) dans la direction perpendiculaire à frontière d'un objet en mouvement.

Les frontières du mouvement peuvent être préservées en imposant la contrainte de lissage seulement dans la direction dans laquelle le champ de vitesse varie peu.

-les méthodes non adaptatives imposent l'équation de contrainte du mouvement dans les zones d'occultation, où l'effet de cette équation devrait être annulé. Ce résultat peut être obtenu en modifiant localement le coefficient de pondération γ . Par exemple, dans les zones d'occultation on doit conserver seulement la contrainte de lissage.

En général, les méthodes différentielles adaptatives supposent une contrainte de lissage directionnel, en particulier le long des contours du mouvement et pas dans la direction perpendiculaire.

II.3.1.4.1 Méthodes différentielles directionnelles

La contrainte de lissage directionnel peut être exprimée par la relation

$$\xi_{ud}^2(v) = (\nabla v_x)^T \cdot w \cdot (\nabla v_x) + (\nabla v_y)^T \cdot w \cdot (\nabla v_y) \quad (\text{II.11})$$

Où \mathbf{W} est une matrice de pondération qui pénalise les variations du champ de mouvement en fonction de variations spatiales de niveaux de gris par la séquence d'images. Un exemple de définition de la matrice \mathbf{W} est :

$$w = \frac{F + \delta I}{\text{trace}(F + \delta I)} \quad (\text{II.12})$$

Où I est la matrice identité qui représente le terme de lissage global nécessaire pour assurer une matrice de pondération non nulle dans les zones d'uniformité spatiale, d est un scalaire.

Ainsi, dans le cadre de méthodes d'estimation du mouvement avec un lissage directionnel on minimise la fonction suivante, qui est similaire à la fonction utilisée dans le cas de la méthode de Horn et Schunck (II.8)

$$\xi^2 = (1 - \gamma) \cdot \xi_{flux}^2 + \gamma \cdot \xi_{ud}^2 \quad (\text{II.13})$$

Où γ est un coefficient de lissage.

On observe que la méthode de Horn et Schunck (II.8) représente un cas particulier de cette formulation et correspond à l'équation (II.12), pour: $d=1$ et $\mathbf{F}=0$. Pour minimiser la fonction (II.12) on peut utiliser une minimisation de type Gauss-Seidel,

dans laquelle le terme actualisé à chaque itération peut être calculé par un algorithme linéaire. Les performances de la méthode avec lissage directionnel dépendent de la précision avec laquelle on peut estimer les dérivées partielles d'ordre 1 et surtout celles d'ordre 2 de l'intensité de l'image, qui interviennent dans l'équation (II.12).

II.3.1.4.2 Méthodes différentielles hiérarchiques

Une représentation multirésolution d'une séquence d'images peut être définie de la manière suivante

$$I_\alpha(x, y, t) = \int_S I(x, y, t) \cdot h\left(\frac{m-x}{\alpha}, \frac{n-y}{\alpha}\right) dm dn \quad (\text{II.14})$$

Où α est le paramètre de résolution. S représente le support d'image et h est un noyau de filtrage. Un exemple de filtre passe-bas de paramètre A, B, C et D est donné par :

$$h(x, y) = \begin{cases} A \cdot \exp\left[\frac{-(C^2 \cdot x^2 + D^2 \cdot y^2)}{B^2 - (C^2 \cdot x^2 + D^2 \cdot y^2)}\right], & \text{pour } (C^2 \cdot x^2 + D^2 \cdot y^2) < B^2 \\ 0 & , \text{sinon} \end{cases} \quad (\text{II.15})$$

On peut observer que la résolution spatiale diminue quand a augmente. On peut observer aussi que les dérivées spatiales d'intensité $I(x, y, t)$ peuvent être calculées par la convolution de $I(x, y, t)$ avec les dérivées partielles de $h(x, y)$, qui peuvent être calculées analytiquement. A chaque niveau de résolution on peut utiliser une méthode d'optimisation non-linéaire, comme par exemple la méthode de Quasi-Newton.

L'estimation obtenue à un niveau de résolution est utilisée comme estimation initiale au niveau supérieur de résolution.

Pour augmenter l'efficacité des méthodes différentielles, on peut combiner différentes approches: l'approche hiérarchique avec l'approche directionnelle ou la prise en compte de discontinuités dans le champ de mouvement.

II.3.2 Méthodes de mise en correspondance

Les méthodes de mise en correspondance sont parmi les plus utilisées en pratique grâce à la simplicité de l'implantation physique du modèle de translation utilisé le plus souvent dans ces méthodes. Elles sont utilisées dans les standards de compression vidéo, comme H.261, MPEG-1, MPEG-2 ou MPEG-4.

Parmi les méthodes de mise en correspondance, on distingue deux classes:

- les méthodes de mise en correspondance dans le plan transformé;
- les méthodes de mise en correspondance dans le plan image.

II.3.2.1. Méthodes de mise en correspondance dans le plan transformé

Dans les méthodes de mise en correspondance dans le plan transformé, le champ de déplacement \mathbf{d} est estimé par la mise en correspondance de blocs définis dans un plan transformé. La méthode la plus répandue est la méthode de corrélation de phase obtenue par la transformée de Fourier. Dans cette méthode, on calcule la corrélation de phase entre les transformées de Fourier de deux blocs correspondants dans les deux images successives.

II.3.2.2. Méthodes de mise en correspondance dans le plan image

Parmi les méthodes de mise en correspondance dans le plan image, les méthodes de mise en correspondance de blocs sont les plus utilisées, grâce à leur simplicité d'implantation software et hardware. Ce type de méthode s'est imposé dans les standards de compression d'images, comme H.261 ou MPEG-1, MPEG-2 et MPEG-4

Les méthodes de mise en correspondance par blocs sont basées sur l'hypothèse selon laquelle l'intensité lumineuse des pixels est constante ou faiblement variable le long des trajectoires de mouvement. Pour chaque point de l'image courante à l'instant t on cherche le point d'intensité la plus proche dans l'image cible (instant $t-1$ ou $t+1$). Pour éviter les correspondances incohérentes dans l'image cible, on limite la recherche à une fenêtre centrée sur le point courant (x,y) . Ainsi, le déplacement

CHAPITRE II : Différentes méthodes d'estimation du mouvement

d'un pixel de coordonnées (x,y) , centre du bloc B de dimension $B_x \times B_y$, de l'image courante, est déterminé par la position du bloc de même dimension, le plus proche au sens d'un critère de similarité, dans l'image cible (instants $t-1$ ou $t+1$).

La ressemblance entre deux blocs est calculée sur l'ensemble du bloc. Ceci sous-entend une hypothèse supplémentaire selon laquelle tous les pixels d'un bloc effectuent le même mouvement. On considère en général, un modèle de translation mais on peut considérer aussi des modèles plus complexes.

II.3.3 Méthodes statistiques

Parmi les méthodes statistiques d'estimation du mouvement, les méthodes Markoviennes ou Bayésiennes sont les plus répandues. Le principe de ces méthodes repose sur une formulation probabiliste du champ de déplacement, l'estimation finale étant une réalisation particulière d'un champ aléatoire.

II.3.4 Méthodes fréquentielles

Ces méthodes sont apparues suite à la constatation que certaines propriétés du mouvement sont plus accessibles dans le domaine fréquentiel. En effet les filtres fréquentiels comme le filtre de Gabor agissent comme les filtres passe bande orientés dans la direction du mouvement. Le principe de base est de trouver une équation équivalente à l'équation du mouvement, dans le domaine fréquentiel. Cela suppose que le champ de vitesses soit localement translationnel. Ainsi, après un calcul de transformée de Fourier, nous obtenons l'équation $\omega_x u + \omega_y v + \omega_t = 0$, où ω_x, ω_y et ω_t sont les fréquences selon l'axe des x , l'axe des y et l'axe des t . Cette équation de contrainte du mouvement dans le domaine fréquentiel se traduira par le fait que l'énergie spatiale sera contenue dans un plan passant par l'origine et dont l'orientation est liée à la translation 2D cherchée.

Pour déterminer un tel plan, les filtres sensibles à une direction donnée sont utilisés, dont des filtres de Gabor spatio-temporels pour estimer la vitesse dans l'image.

II.4 Conclusion

Dans ce chapitre, nous avons décrit brièvement les trois types de mouvement représentant l'image, qui n'est rien d'autre qu'une projection d'une scène réelle 3D d'où les trois mouvements, qui sont : réel, apparent et estimé.

Afin d'obtenir ce dernier, plusieurs types de méthodes existent, nous les avons classées en quatre catégories : méthodes différentielles, méthodes du Block Matching, méthodes statistiques et méthodes fréquentielles. La deuxième catégorie fera l'objet de notre étude dans le chapitre suivant.

III.1 Introduction

L'estimation du mouvement dans les séquences d'images, est l'un des problèmes fondamentaux en traitement d'images. Elle consiste principalement à exploiter les redondances temporelles qui existent entre les images consécutives de la séquence, afin de représenter des vecteurs caractérisant les mouvements existant dans cette séquence.

Les méthodes du Block Matching sont parmi les méthodes d'estimation du mouvement les plus utilisées dans la pratique. Elles peuvent être utilisées pour découvrir des redondances temporelles dans une séquence vidéo tel qu'un objet sans (avec) mouvement dans une vidéo, ce qui accroît l'efficacité de compression d'une image. On les retrouve presque dans tous les standards de compression vidéo (H.261, MPEG-1, MPEG-2 ou MPEG-4). Elles s'attaquent aussi aux problèmes d'estimation des mouvements au niveau local. Dans la littérature, l'estimation du mouvement n'est pas utilisée uniquement pour la compression vidéo, mais elle présente aussi un élément principal de plusieurs applications vidéo tels que l'analyse des séquences, le suivi des objets en mouvement (Applications militaires), le calibrage des caméras...etc

III.2 Principe général du Block-Matching

L'objectif principal des méthodes classiques de Block Matching est de repérer un bloc similaire (bloc cible) au bloc source à partir d'une ou plusieurs images de référence qui peuvent apparaître future ($t+1$) ou passée ($t-1$) par rapport l'image courante (t).

CHAPITRE III : Algorithme du Block-Matching

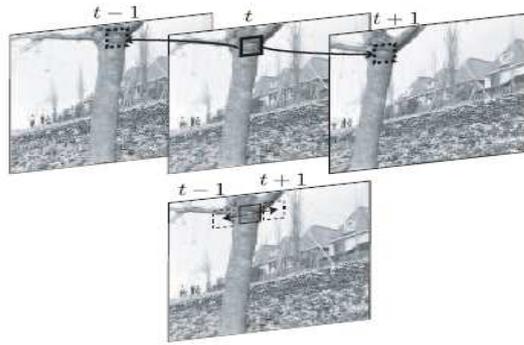


Figure 24 : mise en correspondances des blocs entre l'instant (t), (t-1) et (t+1)

La méthode consiste à diviser l'image courante (t) en blocs non chevauchés de tailles identiques qui sont vus comme des blocs indépendants où les pixels composant chaque bloc ont le même pas de mouvement (généralement 8x8 ou 16x16 pixels).

La deuxième étape consiste à trouver le meilleur bloc cible dans l'image de référence (t-1 ou t+1) par rapport à un bloc de référence choisi dans l'image courante (t). Le meilleur bloc est choisi par un algorithme minimisant / maximisant un critère de comparaison.

Pour réduire le coût des calculs, on définit une fenêtre de recherche qui va limiter le nombre de blocs candidats testés dans l'image de référence. La taille de la fenêtre de recherche dépend du déplacement maximal autorisé pour chaque pixel (choisi par l'utilisateur).

On recommence le procédé avec un autre bloc jusqu'à ce qu'il n'y ait aucun bloc à tester, ou bien, jusqu'à atteindre un critère d'arrêt prédéfini. Une fois la procédure terminée, on obtient pour chaque bloc de l'image courante un vecteur de déplacement qui caractérise son mouvement.

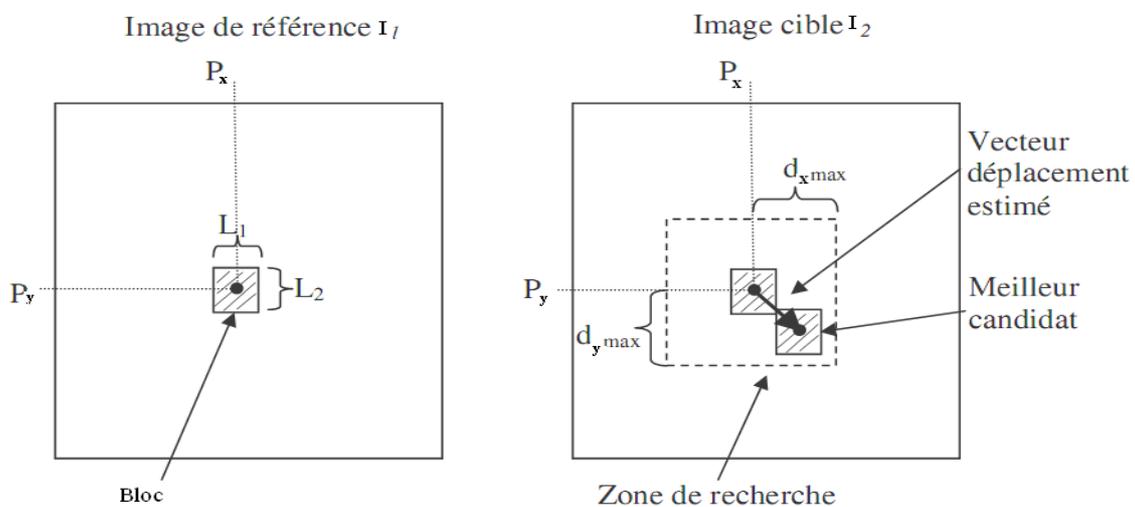


Figure 25 : principe de la méthode de mise en correspondance des blocs

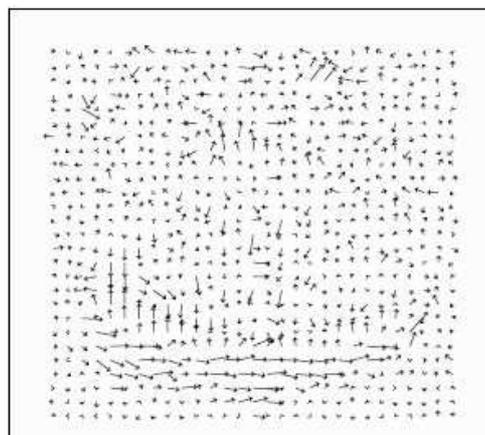
CHAPITRE III : Algorithme du Block-Matching

La figure 25 donne un aperçu global sur le principe de fonctionnement de la méthode du Block Matching (BM). Dans ce qui suit nous allons développer chaque point cité plus haut, afin d'expliquer toutes les étapes de cette approche. On commencera par donner un exemple illustratif de l'estimation d'une manière générale pour mieux assimiler les autres étapes.

• Exemple illustratif de l'estimation



a) Image de la séquence du cube de Rubik



b) Vecteurs des déplacements qui caractérisent le mouvement

Figure 26 : Exemple de l'estimation du mouvement en utilisant le Block Matching

La figure 26 représente une estimation de mouvement avec le Block Matching réalisée avec la séquence du cube de Rubik, les résultats sont représentés sous forme de vecteurs qui caractérisent le déplacement des blocs.

Il existe de nombreuses méthodes de Block Matching cherchant à optimiser l'efficacité et la rapidité de l'algorithme. Pour trouver un bloc cible dans une image de référence, la plus part des algorithmes du Block-Matching se basent sur deux points cruciaux.

- Le parcours d'image bloc par bloc dans un certain ordre. De nombreuses manières existent et chacune possède généralement un ordre de parcours différent.

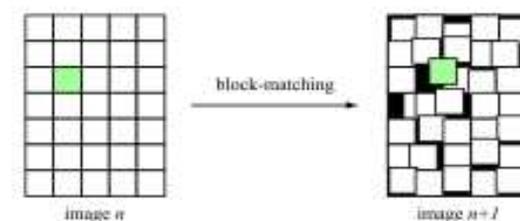
CHAPITRE III : Algorithme du Block-Matching

- Et afin de déterminer le meilleur des blocs similaires, les algorithmes comparent la différence entre le bloc source (image courante) et les autres blocs cibles (des images de référence) à l'aide d'un critère d'évaluation (similarité/disparité) c'est à dire, le bloc ciblé qui minimise/maximise cette règle est retenu. Dans ce dernier point, de nombreuses techniques d'analyse ont été publiées dans la littérature.

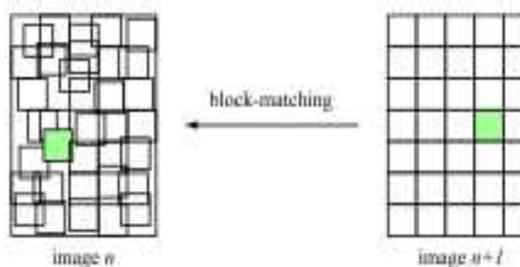
Dans ce qui suit nous allons développer chaque point cité, afin d'expliquer toutes les évolutions de cette approche.

III.3 Prédiction avant et arrière

L'algorithme du Bloc Matching permet d'estimer le mouvement des blocs entre deux images à des temps consécutifs. Le mouvement ainsi calculé va permettre de prédire l'image $n+1$ en combinant l'information contenue dans les pixels présents dans l'image n et les vecteurs mouvements. Nous avons vu précédemment que nous utilisons deux blocs. Cependant, nous avons délibérément omis de spécifier de quelles images sont extraits chacun des blocs. Ainsi, deux types de prédiction sont possibles : la prédiction avant (forward) et la prédiction arrière(backward).



Prédiction forward



Prédiction backward

Figure 27 : Prédiction forward et backward

CHAPITRE III : Algorithme du Block-Matching

La prédiction forward consiste à diviser l'image n en blocs et à trouver leur position dans l'image $n + 1$. Comme le montre la figure 27, le block-matching avec prédiction forward répond pour chaque bloc à la question «Où va le bloc courant ? ».

En revanche, la prédiction backward, illustrée sur la figure 27, divise l'image $n + 1$ en blocs et cherche leur position dans l'image n . Dans ce cas la question est «D'où vient le bloc courant ? ».

Les résultats sont très différents. Il apparaît par construction que l'image prédite par prédiction forward présente des « trous ». En effet, de manière générale les blocs n'ont pas un mouvement unique pour toute l'image. Lors de la prédiction certains blocs se recouvrent et laissent apparaître de zones n'ont prédites et par conséquent noires (voir figure 27). Ce problème est entièrement résolu par l'utilisation de la prédiction backward qui ne laisse entrevoir aucun « trou » dans la prédiction.

III.4 Critères d'évaluation

Comme mentionné précédemment, les blocs sont assignés sur la base d'un critère d'évaluation, ce critère n'est autre qu'une mesure de ressemblance (similarité) ou de dissemblance (disparité) entre les blocs testés.

Dans le cas d'un critère de ressemblance, ce dernier doit être maximisé, sinon, minimisé dans le cas d'une dissemblance.

De manière générale, la plupart des articles traitant de l'estimation du mouvement, avec les algorithmes du Block Matching, considèrent que les blocs sont caractérisés par la luminance (intensité lumineuse). Ceci est dû au fait que l'œil humain est davantage sensible à l'intensité lumineuse qu'à la couleur.

Les critères les plus couramment utilisés sont :

- {Sum of Square Différences (SSD) : somme des différences des pixels de chaque bloc au carré, non adapté à la couleur (critère à rajouter).
- {Som of Absolute différences (SAD) : Somme des différences absolues calcule la somme des valeurs absolues des différences de luminance des pixels de l'image originale et celle de référence.

CHAPITRE III : Algorithme du Block-Matching

- {Zero-mean Normalized Sum of Absolute Differences (ZNSAD) : prend en compte les moyennes des blocs et normalise le résultat (ZNSSD pour SAD)
- {Erreur absolue moyenne EAM: traite les erreurs de manière uniforme.
- {Coefficient d'inter-corrélation (CC) : le critère de ressemblance statistique le plus utilisé.
- {Coefficient d'inter-corrélation normalisé et centré (NCC) : c'est le CC normé et centré à la variance.
- {Erreur quadratique moyenne (EQM) : Ce critère est dérivé de l'erreur moyenne absolue.

Les critères les plus couramment utilisés sont données dans le tableau. Les expressions des critères sont données pour une fenêtre de recherche de taille $L \times L$, centrée sur le pixel de coordonnées $\mathbf{P} = (P_x, P_y)$ (position initiale) dans l'image de référence. Le déplacement d'un bloc candidat à l'intérieur de la fenêtre est caractérisé par un vecteur $\vec{d} = (dx, dy)$ et est borné par la taille de la zone de recherche. La position du meilleur bloc candidat est notée (\hat{d}_x, \hat{d}_y)

Table : Les critères de comparaison

Mesures de disparité : le meilleur candidat correspond au minimum de ces fonctions	
Somme des différences absolues (SAD)	$SAD(p, d) = \sum_{x=-L/2}^{L/2} \sum_{y=-L/2}^{L/2} I1(p_x + x, p_y + y) - I2(p_x + x + dx, p_y + y + dy) $ $(\hat{d}_x, \hat{d}_y) = \arg \min_d (SAD(p, d))$ <p>Ce critère est le même que SSD sauf qu'on somme les différences absolues des valeurs des pixels des deux blocs. Il est moins couteux en calcul vu qu'on ne lève pas la différence au carré.</p>
Somme des carrés des différences (SSD)	$SSD(p, d) = \sum_{x=-L/2}^{L/2} \sum_{y=-L/2}^{L/2} \left(I1(p_x + x, p_y + y) - I2(p_x + x + dx, p_y + y + dy) \right)^2$ $(\hat{d}_x, \hat{d}_y) = \arg \min_d (SSD(p, d))$ <p>Ce critère représente la somme des carrés des différences</p>

CHAPITRE III : Algorithme du Block-Matching

	entre les pixels correspondants aux deux blocs.
<p>Erreur absolue moyenne (EAM)</p>	$EAM(p, d) = \frac{1}{L^2} \sum_{x=-L/2}^{L/2} \sum_{y=-L/2}^{L/2} I1(p_x + x, p_y + y) - I2(p_x + x + dx, p_y + y + dy) $ $(\hat{dx}, \hat{dy}) = \arg \min_d (EAM(p, d))$ <p>Cette mesure est la plus couramment utilisée, car elle est moins coûteuse en calcul que les autres méthodes.</p>
<p>Erreur quadratique moyenne (EQM)</p>	$EQM(p, d) = \frac{1}{L^2} \sum_{x=-L/2}^{L/2} \sum_{y=-L/2}^{L/2} (I1(p_x + x, p_y + y) - I2(p_x + x + dx, p_y + y + dy))^2$ $(\hat{dx}, \hat{dy}) = \arg \min_d (EQM(p, d))$ <p>Ce critère est dérivé de l'erreur moyenne absolue, elle est plus coûteuse en calcul mais donne de meilleurs résultats que l'erreur absolue moyenne</p>
<p>Zero-mean Normalized Sum of Absolute Differences (ZSAD)</p>	$ZSAD(p, d) = \sum_{x=-L/2}^{L/2} \sum_{y=-L/2}^{L/2} I1(p_x + x, p_y + y) - \bar{I1} - I2(p_x + x + dx, p_y + y + dy) + \bar{I2} $ $(\hat{dx}, \hat{dy}) = \arg \min_d (ZSAD(p, d))$ <p>avec $\bar{I1}, \bar{I2}$ représentent les valeurs moyennes des pixels de la zone de recherche et respectivement dans le bloc candidat (même chose pour ZNSSD)</p>
<p>Zero-mean Normalized Sum of Absolute Differences (LSAD)</p>	$LSAD(p, d) = \sum_{x=-L/2}^{L/2} \sum_{y=-L/2}^{L/2} I1(p_x + x, p_y + y) - \frac{\bar{I1}}{\bar{I2}} \cdot I2(p_x + x + dx, p_y + y + dy) $ $(\hat{dx}, \hat{dy}) = \arg \min_d (LSAD(p, d))$ <p>avec $\bar{I1}, \bar{I2}$ représentent les valeurs moyennes des pixels de la zone de recherche et respectivement dans le bloc candidat (même chose pour LSSD)</p>
<p>Mesures de ressemblance : le meilleur candidat correspond au maximum de ces fonctions</p>	

CHAPITRE III : Algorithme du Block-Matching

<p>Coefficient d'inter-corrélation (CC)</p>	$CC(p, d) = \sum_{x=-L/2}^{L/2} \sum_{y=-L/2}^{L/2} I1(p_x + x, p_y + y) \bullet I2(p_x + x + dx, p_y + y + dy)$ $(\hat{dx}, \hat{dy}) = \arg \max_d (CC(p, d))$ <p>Le bloc candidat retenu est celui qui maximise ce critère, il permet d'avoir de bons résultats.</p>
<p>Coefficient d'inter-corrélation normalisé et centré (NCC)</p>	$NCC(p, d) = \frac{\sum_{x=-L/2}^{L/2} \sum_{y=-L/2}^{L/2} (I1(p_x + x, p_y + y) - \bar{I1}) \bullet (I2(p_x + x + dx, p_y + y + dy) - \bar{I2})}{\sqrt{\sum_{x=-L/2}^{L/2} \sum_{y=-L/2}^{L/2} (I1(p_x + x, p_y + y) - \bar{I1})^2 \sum_{x=-L/2}^{L/2} \sum_{y=-L/2}^{L/2} (I2(p_x + x + dx, p_y + y + dy) - \bar{I2})^2}}$ $(\hat{dx}, \hat{dy}) = \arg \max_d (NCC(p, d))$ <p>Ce critère est une version améliorée du coefficient d'inter-corrélation. Il permet d'avoir des résultats plus précis, mais il est plus pénalisant en calculs</p>

III.5 Stratégies de recherche du meilleur bloc candidat

Il existe plusieurs méthodes du Block Matching pour le parcours du voisinage, afin de trouver le bloc candidat qui correspond au mieux au bloc de référence, et qui optimise le critère d'évaluation.

La méthode la plus fondamentale pour trouver le meilleur bloc candidat est de visiter et de tester tous les blocs de la zone de recherche. Cette stratégie, appelée recherche exhaustive, est optimale mais pourrait s'avérer très coûteuse en temps de calcul, surtout pour des amplitudes de déplacement grandes. Cela implique des zones de recherche de taille importante et donc un nombre considérable de blocs candidat à visiter.

D'autres balayages non exhaustifs de la zone de recherche ont été proposés dans la littérature.

A titre d'exemple, nous pouvons citer la recherche en n pas, la recherche 2D logarithmique ou la recherche en Diamant.

Une autre stratégie de recherche classique non exhaustive est la descente de gradient. L'objectif de cette technique est de converger plus rapidement vers le minimum de la fonction de similarité.

III.5.1 Algorithme de recherche exhaustive (Full Search (FS))

Cette méthode consiste à faire la recherche de bloc sur l'ensemble des blocs de l'image référence. Cet algorithme est le plus simple, mais aussi le plus coûteux car il n'optimise pas la recherche: il n'y a pas de seuil pour la fenêtre de recherche. Il sert néanmoins de référence pour les autres algorithmes.

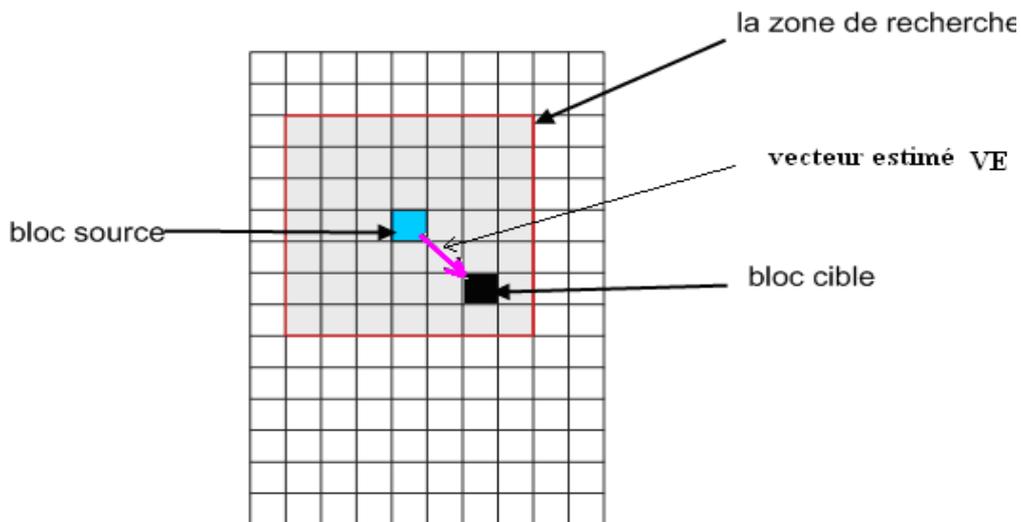


Figure 28 : Full Search Algorithm

III.5.2 Recherche en trois pas (Three Step Search (TSS))

(Proposé par Koga; et al, 1981)

Contrairement à l'algorithme FS, cette méthode (TSS) se limite à trois pas pour trouver le meilleur bloc. Au lieu de parcourir tous les blocs de la zone de recherche, le but de ces algorithmes est de parcourir la fenêtre en se rapprochant pas à pas du bloc candidat.

La Figure 29 illustre les trois étapes nécessaires à l'algorithme pour retourner un résultat.

Dans la première étape le bloc centré se place au vecteur de mouvement nul. Ensuite tester le critère aux alentours du centre (8-connexité) avec un premier pas, et on repère le point qui vérifie le critère. On prend le point trouvé précédemment comme nouveau centre et on recherche aux alentours de ce point avec un pas inférieur au premier pas. On reprend le minimum comme centre de la prochaine étape. Dans cette dernière étape on répète une troisième fois l'opération avec un pas inférieur aux deux

premiers. (Recherche en multi résolution).

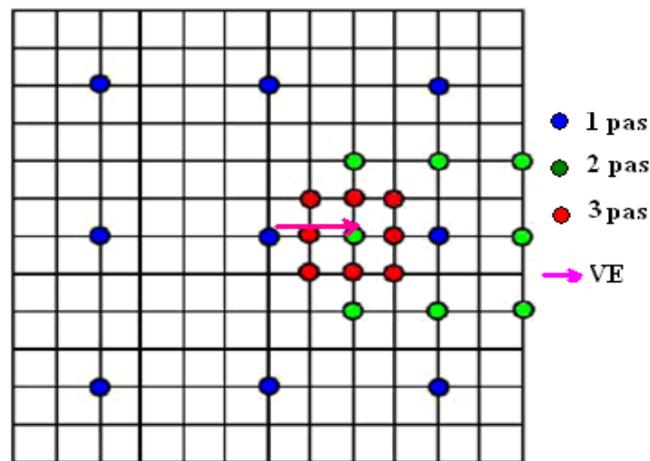


Figure 29 : recherche en 3 pas

Cet algorithme converge. Il est très rapide (3 phases de 8 tests seulement) mais on a une baisse de qualité visuelle.

III.5.3 Recherche en quatre pas

(Proposé par Lai-Man Po et al, 1996)

L'algorithme commence par une comparaison des neuf points et ensuite les autres points sont choisis comme suit :

Début

Pas 1 : L'algorithme commence avec un pas de taille 2, calcul l'altération et prend le point qui a la valeur la plus petite, si le point choisi se trouve au centre alors aller au Pas 4, sinon continuer vers le Pas 2.

Pas 2 : Prendre le nouveau point pour le centre et maintenir la taille du pas à 2, mais le nombre de nouveaux points à explorer dépend de la position précédente du point qui se trouve au centre. Localiser le point avec l'altération minimale. Si c'est au centre, allez au Pas 4 sinon Pas 3.

Pas 3 : La stratégie de recherche est la même cependant il ira finalement au Pas 4.

Pas 4 : La taille du pas est réduite à 1 et tous les huit points autour du centre de la recherche sont examinés.

Fin

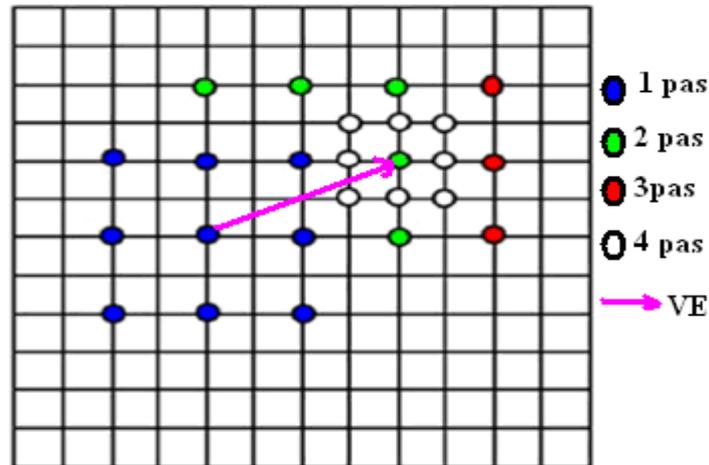


Figure 30 : recherche en 4 pas

III.5.4 Recherche selon le gradient (Gradient Search GS)

(Proposée par Liu et Feig, 1996)

Cet algorithme utilise un pas de recherche fixe et le passage d'un bloc candidat à l'autre se fait en suivant la direction indiquée par la descente du gradient de la fonction de coût utilisée. C'est un algorithme qui fonctionne efficacement dans le cas des séquences stationnaires.

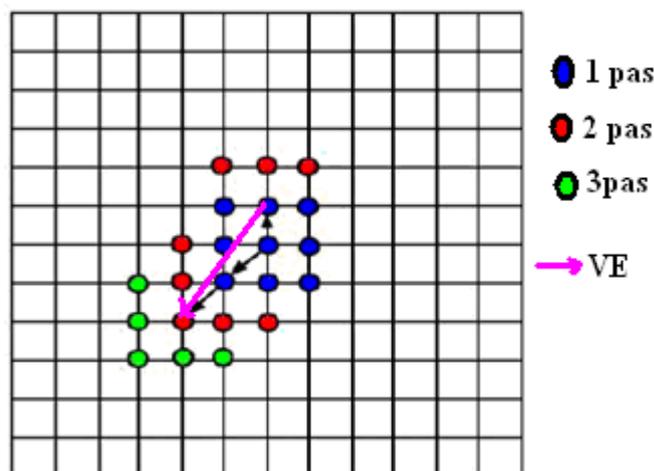


Figure 31 : recherche selon le gradient

III.5.5 Recherche en Diamant (Diamond search DS)

(Proposée par A.M. Tourapis et al, 2000).

Cette méthode vise aussi à réduire le nombre de blocs à visiter.

CHAPITRE III : Algorithme du Block-Matching

➤ **Diamant simple:** La recherche en diamant (Diamond search) commence la recherche au vecteur de mouvement nul. Puis il examine le critère en déplaçant la fenêtre sur 9 points disposés en diamant < large Diamond Search Pattern > (LDSP, grand modèle de recherche en diamant). Une fois le minimum (maximum) des ces neufs points trouvé, on reprend la recherche avec le minimum (maximum) précédent comme centre de la nouvelle recherche. On réitère l'opération tant que le minimum (maximum) ne se trouve pas au centre du diamant. Quand on a effectivement le minimum (maximum) au centre, on effectue une dernière recherche avec un plus petit diamant qui déterminera le minimum (maximum) estimé de l'image. Le second modèle nommé < Small Diamond Search Pattern > (SDSP, petit modèle de recherche en diamant) contient quatre points.

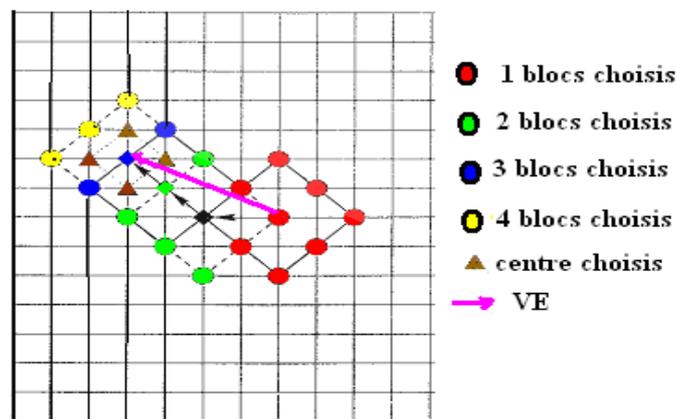


Figure 32 : Algorithme "Diamond search":

➤ **Predictive Diamond Search:** Une amélioration de cet algorithme existe sous le nom de Predictive Diamond Search. Celle-ci utilise l'algorithme diamant, mais utilise les blocs connexes pour positionner sa première recherche. On examine le critère pour les blocs connexes dont les vecteurs de mouvement ont déjà été trouvés: Est, Nord, Nord-ouest, dont on prend les coordonnées médianes. Ces coordonnées correspondront au centre initial de la recherche du bloc. On utilise ici le fait que le vecteur de mouvement d'un bloc a de forte chance d'être corrélé avec ceux des blocs connexes.

III.5.6 Recherche sur une grille hexagonale (Hexagon Based Search HBS)

Le principe de fonctionnement est exactement le même que l'algorithme Diamond search, sauf que les modèles utilisés par le «HEXBS» sont LHSP pour «

CHAPITRE III : Algorithme du Block-Matching

On obtient donc des algorithmes qui consomment environ 10 fois moins de temps que la méthode Full Search, donc plus adaptés à la compression vidéo.

De même que pour les méthodes précédentes, Plusieurs algorithmes rapides d'estimation de mouvement ont été étudiés pour réduire le temps de recherche des blocs, La méthode PHODS (Parallel Hierarchical One Dimension Search), La méthode (Zonal Based Algorithm) sont parmi les méthodes rapides d'estimation de mouvement. Le principe PHODS est de faire en parallèle et indépendamment la recherche du minimum (maximum) sur les deux axes dans la fenêtre de recherche. Et la méthode (Zonal Based Algorithm) propose une solution qui estime le mouvement d'un bloc en effectuant la recherche en fonction de l'ampleur du mouvement.

III.6 Les méthodes avancées (Block Matching)

Une des limitations des méthodes classique du Block Matching est le temps de calcul, qui augmente avec le carré de l'amplitude maximale du mouvement. Pour diminuer le temps de calcul et réduire l'effet de bloc, diverses améliorations ont été proposées à partir des méthodes classiques.

- Le premier a été de remplacer la trame de référence I par une trame interpolée \hat{I} (interpolation subpixélique). Les positions interpolées sont autant de positions supplémentaires possibles pour les vecteurs de mouvement. Il existe différentes méthodes d'interpolation subpixélique :

- Interpolation bilinéaire : moyenne des deux pixels voisins situés à égale distance, imprécise au-delà du demi-pixel.

- Interpolation utilisant la norme H26L : normes de codage uniquement pour le quart et le huitième de pixel.

- Interpolation utilisant les courbes de Bézier : utilisation de B-Splines.

Ensuite, une fois que la méthode est choisie, il suffit d'adapter les méthodes classiques de Block-Matching.

- Il est aussi possible d'effectuer une recherche hiérarchique (Block Matching hiérarchique (hierarchical BM)).

III.6.1 Estimation du mouvement subpixelique

Les techniques proposées jusqu'ici permettent de réaliser une estimation du mouvement de façon plus ou moins efficace, tant au niveau de la qualité qu'au niveau de la rapidité de l'algorithme. Cependant, le déplacement obtenu par ces méthodes fait intervenir des vecteurs aux coordonnées entières ce qui contredit totalement la continuité de l'espace dans lequel nous évoluons. Pour prendre en compte cet aspect, l'interpolation des images et son application au block-matching permettent de se rapprocher du monde réel.

➤ Interpolation subpixelique - principe général

Les champs de vecteurs résultants des différentes méthodes font apparaître naturellement des coordonnées entières car nous considérons que le déplacement se fait sur les pixels. En effet, le pixel est par définition le plus petit élément représentable dans une image et il paraît évident que l'estimation du mouvement entre deux images ne peut donner que des déplacements pixéliques. Cependant, le déplacement réel d'un objet ne se fait pas en fonction de l'échantillonnage du capteur d'acquisition et il en résulte que le mouvement fait généralement intervenir des déplacements non entiers. De ce fait, l'expression du mouvement par des vecteurs ayant des coordonnées non entières devient une nécessité pour une estimation de meilleure qualité. L'interpolation subpixelique trouve ici son intérêt.

Nous disposons d'une image composée de pixels. Chaque pixel est en fait la synthèse des couleurs d'une zone de l'espace. Il n'est en effet pas possible d'imaginer la représentation du monde sous forme continue du fait de la masse d'information à traiter et de la capacité limitée de stockage que nous possédons. Ainsi, une image est une approximation de la projection d'une scène tridimensionnelle sur un plan bidimensionnel.

L'interpolation subpixelique consiste, étant donné les pixels de l'image, à imaginer les pixels qui seraient présents si l'image avait été prise avec une résolution plus importante.

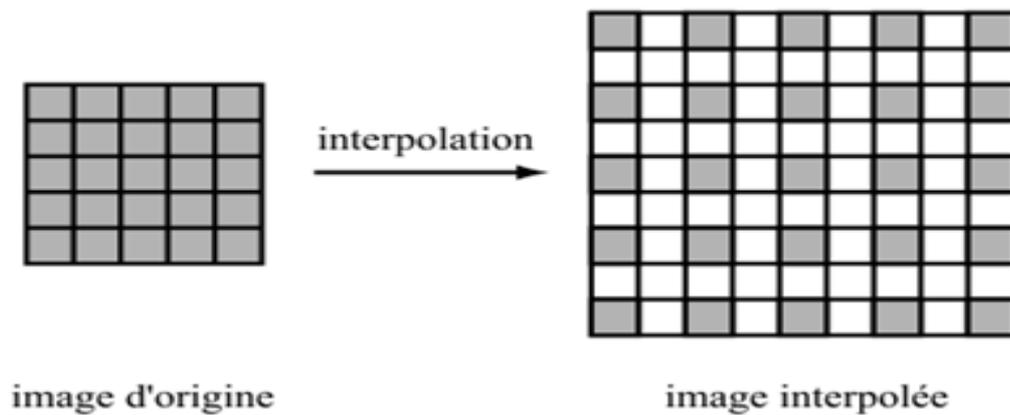


Figure 35 : Interpolation au demi -pixel

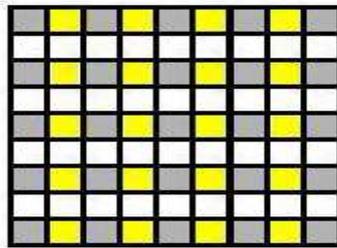


Figure 36 : Interpolation bilinéaire au demi- pixel

Par exemple, l'interpolation avec une précision au demi -pixel revient à insérer un pixel entre deux pixels consécutifs de l'image comme illustrée sur la figure 41. Cette figure illustre bien le travail à réaliser pour interpoler l'image. Bien évidemment la figure présentée ne tient compte que des niveaux de gris. Avec une image en couleur, l'interpolation est à appliquer sur les trois chrominances en mode RGB et sur la luminance et les deux chrominances en mode YUV.

➤ Méthodes d'interpolation subpixelique

Nous présentons ici les différentes méthodes d'interpolation de manière synthétique.

- **Interpolation bilinéaire**

La méthode d'interpolation bilinéaire est l'une des plus simples existant mais elle reste imprécise pour des précisions dépassant le demi- pixel. Elle consiste à faire la moyenne des deux pixels voisins situés à égale distance.

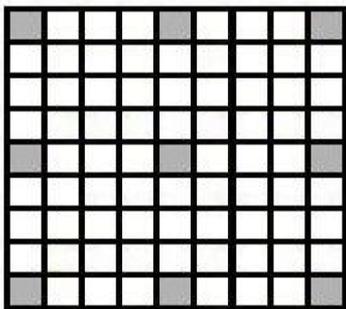
CHAPITRE III : Algorithme du Block-Matching

La figure 36 illustre l'interpolation d'une image au demi-pixel. Cette dernière est réalisée en deux étapes :

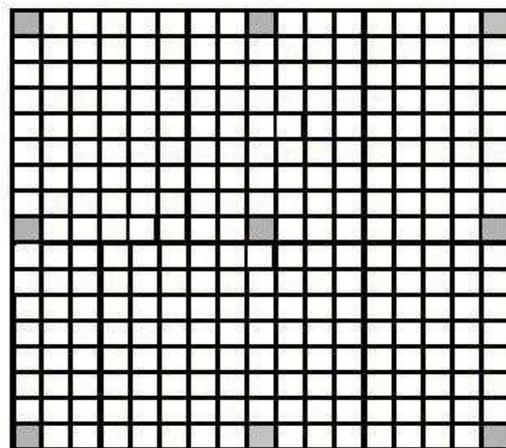
1 ère étape : Tout d'abord, on effectue un balayage des lignes pour calculer les valeurs des nouveaux pixels insérés, qui sont représentés dans la figure en jaune. Ainsi, la valeur d'un pixel jaune est égale à la moyenne des deux pixels en gris voisin de celle-ci.

2ème étape : Cette 2^{ème} étape permet de calculer les valeurs des pixels qui sont représentés en blanc dans la figure précédente. En faisant un parcours vertical des colonnes, de la même manière que la 1ère étape, la valeur d'un pixel blanc est égale à la moyenne des deux pixels voisins.

De la même manière que l'interpolation au demi-pixel, l'interpolation au quart de pixel (respectivement au huitième de pixel) consiste à insérer trois pixels (respectivement sept pixels) entre deux pixels consécutifs de l'image originale (Voir Figure 37).



Interpolation au quart de pixel



Interpolation au huitième de pixel

Figure 37 : Interpolation bilinéaire au quart et au huitième de pixels

La méthode d'interpolation bilinéaire peut être utilisée pour interpoler toutes les précisions qui sont des multiples de deux. Cependant arrivant à une précision d'interpolation au huitième de pixel, les résultats obtenus sont moins précis, ce qui nous pousse à utiliser d'autres méthodes qui donnent de bien meilleurs résultats. Les méthodes les plus pertinentes sont :

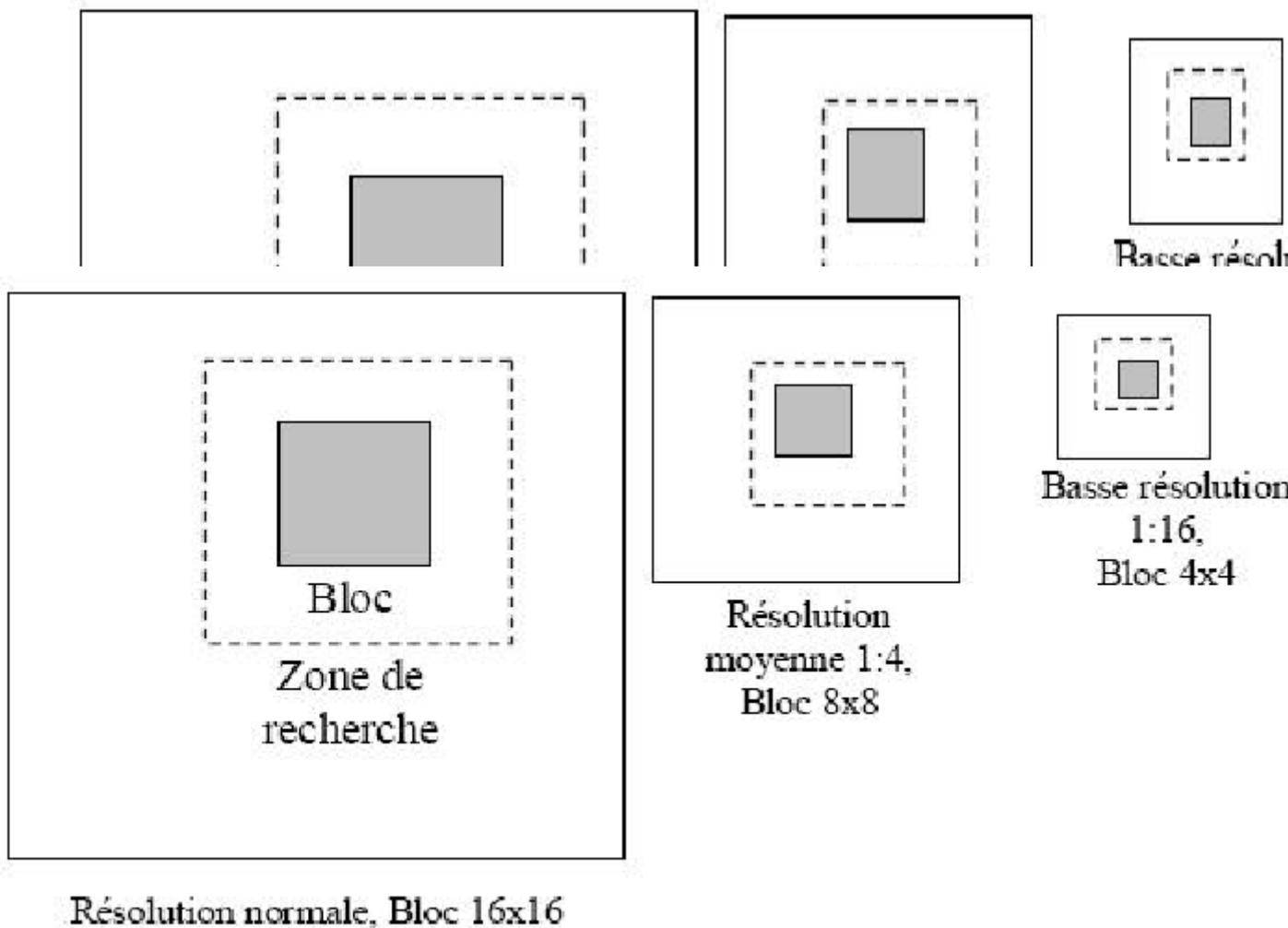
CHAPITRE III : Algorithme du Block-Matching

- **Interpolation utilisant la norme H26L**

C'est une norme de codage qui fournit des méthodes d'interpolation pour les précisions au quart et au huitième de pixel, elle permet d'avoir de bons résultats, mais la complexité des calculs est plus importante que la méthode bilinéaire.

- **Interpolation utilisant les courbes de Bézier**

Les deux méthodes présentées précédemment permettent d'interpoler une image aux précisions du demi, du quart et du huitième de pixel. Cependant, les



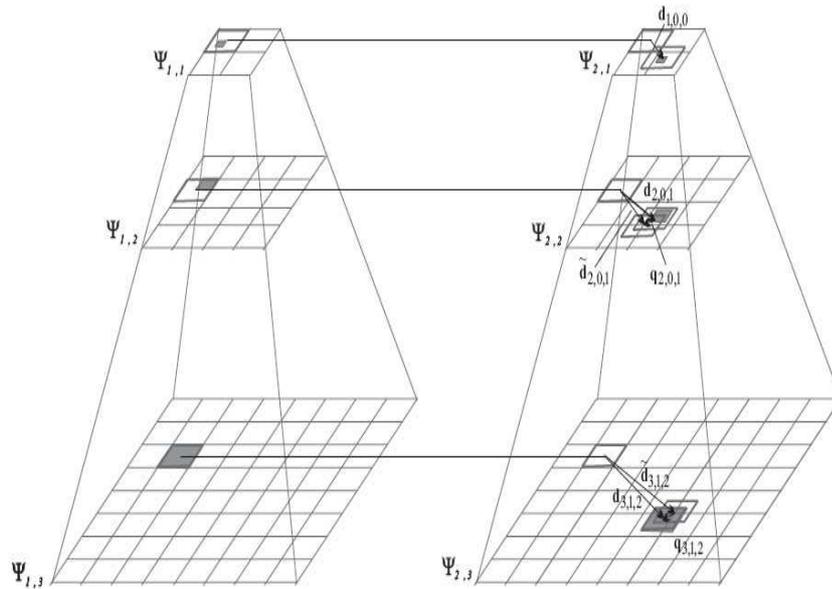


Figure 38 : exemple de représentation multi résolution

III.7 Conclusion

Les méthodes du Block Matching sont parmi les méthodes d'estimation du mouvement les plus utilisées dans la pratique.

Dans ce chapitre, nous avons étudié les principales techniques de la mise en correspondance des blocs (Block Matching) ainsi que le principe général de leur fonctionnement.

IV.1 Introduction

Afin d'estimer le mouvement dans des séquences d'images, on est amené à étudier certaines méthodes dont la méthode du Block Matching.

Pour évaluer les performances de notre méthode(BM), deux critères de performance ont été utilisés pour évaluer nos résultats, à savoir, le PSNR, l'erreur entre l'image prédite et l'image de référence.

1- Le PSNR qui est le rapport signal sur bruit crête, il combine simplicité et bonne performance.

Le PSNR s'exprime dans le cas des images en niveaux de gris par :

$$PSNR=10\log_{10}\cdot\frac{(255)^2}{\frac{1}{N*N}\sum_{j=0}^{N-1}\sum_{i=0}^{N-1}[P(i,j)-\tilde{P}(i,j)]^2}$$

Où: $N * N$ est la taille de l'image, $P(i, j)$ désigne un pixel de l'image courante, $\tilde{P}(i, j)$ désigne un pixel de l'image de référence.

2-Image Erreur

$$IE= (image\ prédite-Image\ de\ référence) = (p(i, j) - \tilde{p}(i, j))$$

IV.2 Etude comparative

Afin d'évaluer les performances des algorithmes du Bloc Matching à savoir : ES, TSS, DS, SS4, on a fait une étude comparative sur les deux séquences rubic et caltrain illustrés ci-dessous :



Vidéo rubic



Vidéo caltrain

CHAPITRE IV : Applications et discussions

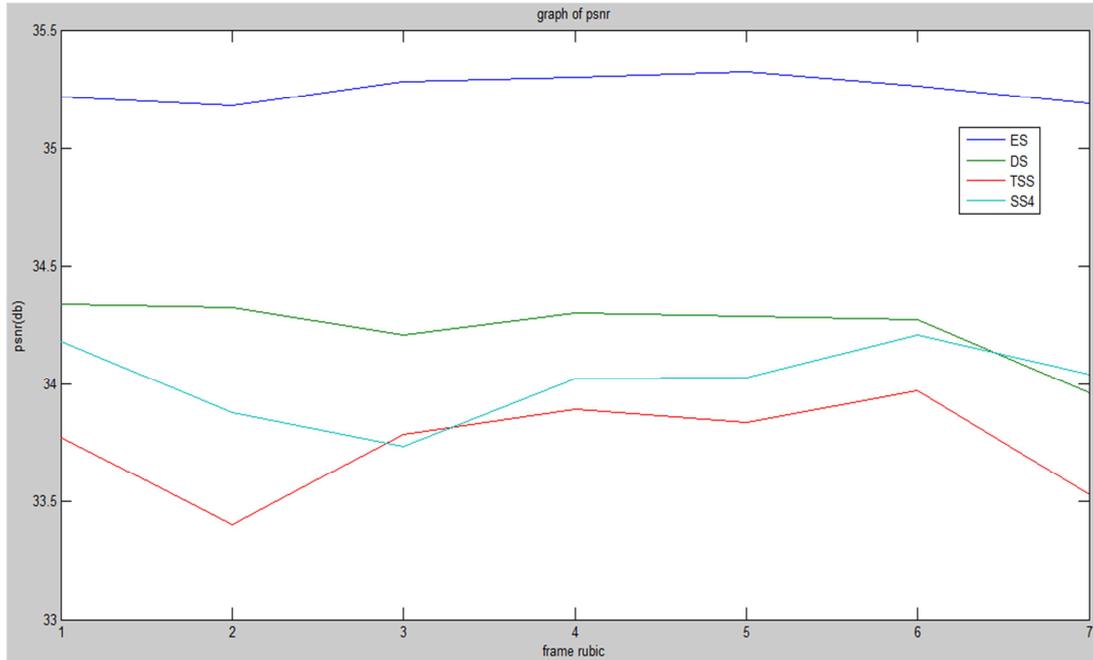


Figure 39 : Résultats obtenus de l'image rubic

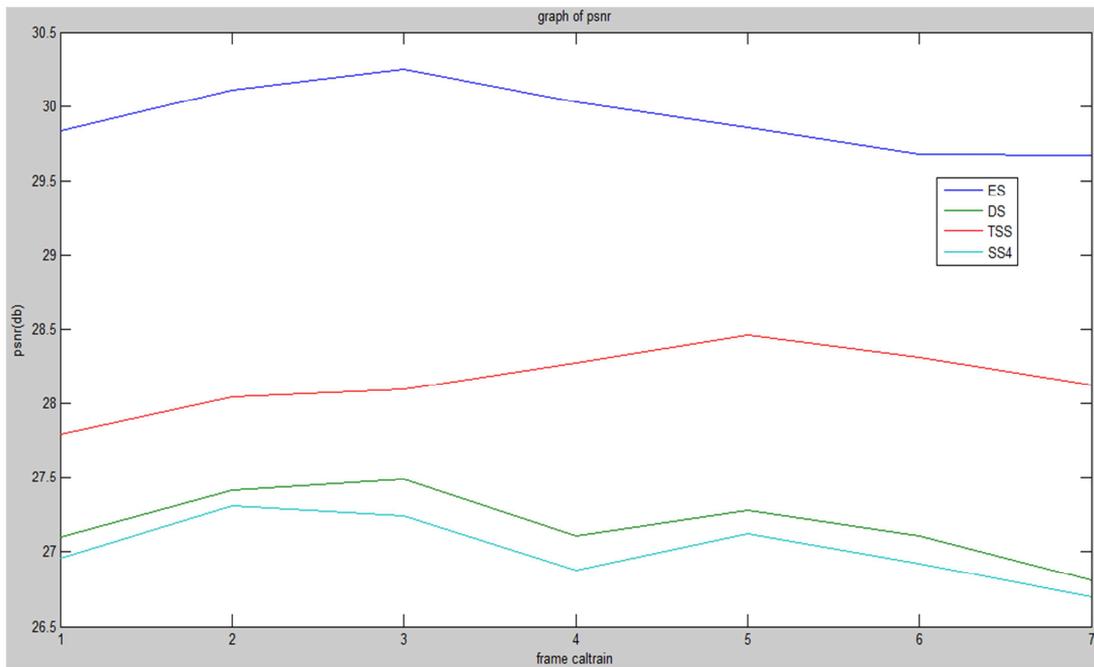


Figure 40 : Résultats obtenus de l'image caltrain

IV.3 Interprétation des résultats

Le critère de performance utilisé pour évaluer nos résultats est le PSNR, qui est défini précédemment.

Les résultats sont obtenus dans les conditions suivantes :

- la taille du Bloc est égale à huit ($mbSize=8$)
- la zone de recherche est égale à 7 ($p=7$)

D'après l'étude comparative qu'on a faite sur les deux séquences d'images rubic ainsi que caltrain, sur la performance des algorithmes du Bloc Matching à savoir : ES, TSS, SS4 et DS, les deux figures 39 et 40 montrent bien la bonne performance de l'algorithme ES par rapport aux autres algorithmes, par les valeurs élevées du PSNR.

Plus le PSNR augmente, meilleure est l'estimation du déplacement.

Vu la performance et la bonne estimation du déplacement de l'algorithme ES, nous allons l'appliquer sur différentes images synthétiques et METEOSAT.

IV.4 Application aux images synthétiques



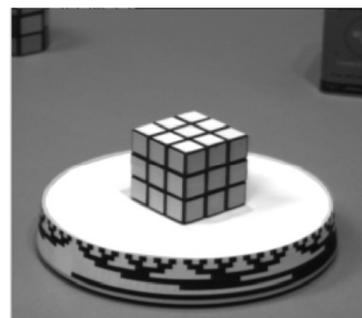
Vidéo 1



Vidéo 2



Vidéo 3

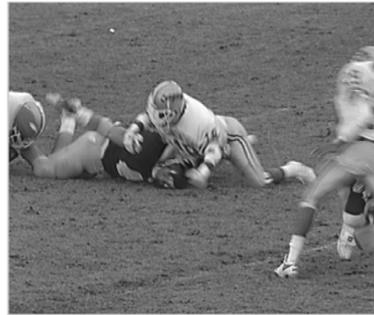


Vidéo 4

CHAPITRE IV : Applications et discussions



Vidéo 5



Vidéo 6

Nous allons présenter les résultats obtenus après la manipulation des images extraites des vidéos, les vecteurs de déplacement estimés ainsi que les images prédites à partir de ces derniers.

Image 1 :



Image1.a



Image1.b



Image prédite

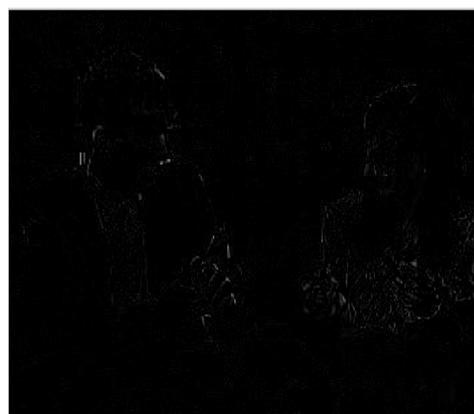
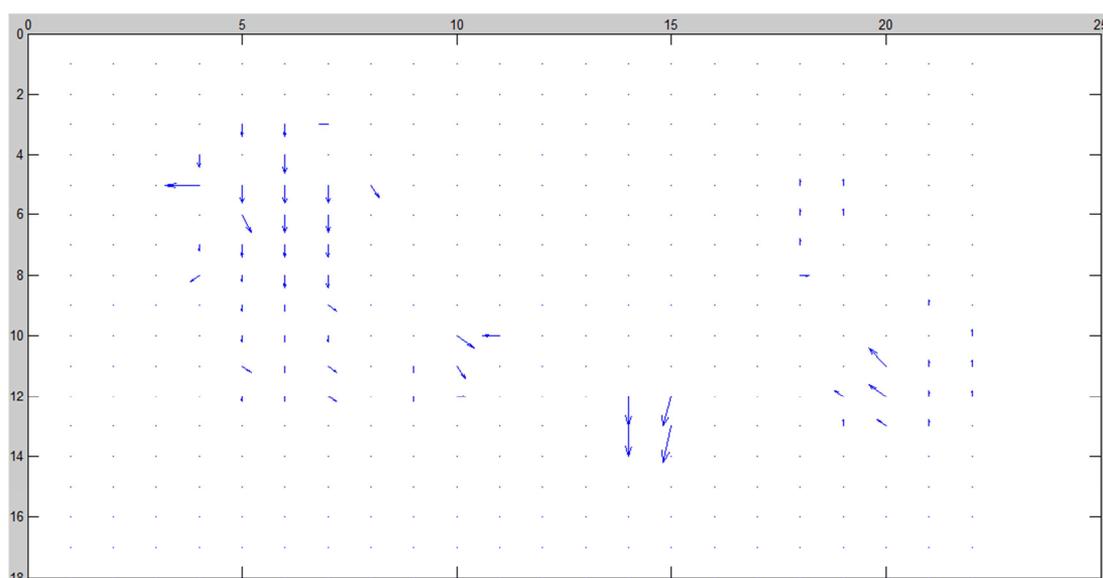


Image erreur

CHAPITRE IV : Applications et discussions



Vecteurs de déplacement



Superposition des vecteurs de déplacement sur l'image

Figure 41 : Résultats de la vidéo 1

Taille du bloc=16

Zone de recherche=7

PSNR=29,8034 db

CHAPITRE IV : Applications et discussions

Image2 :



Image 2.a



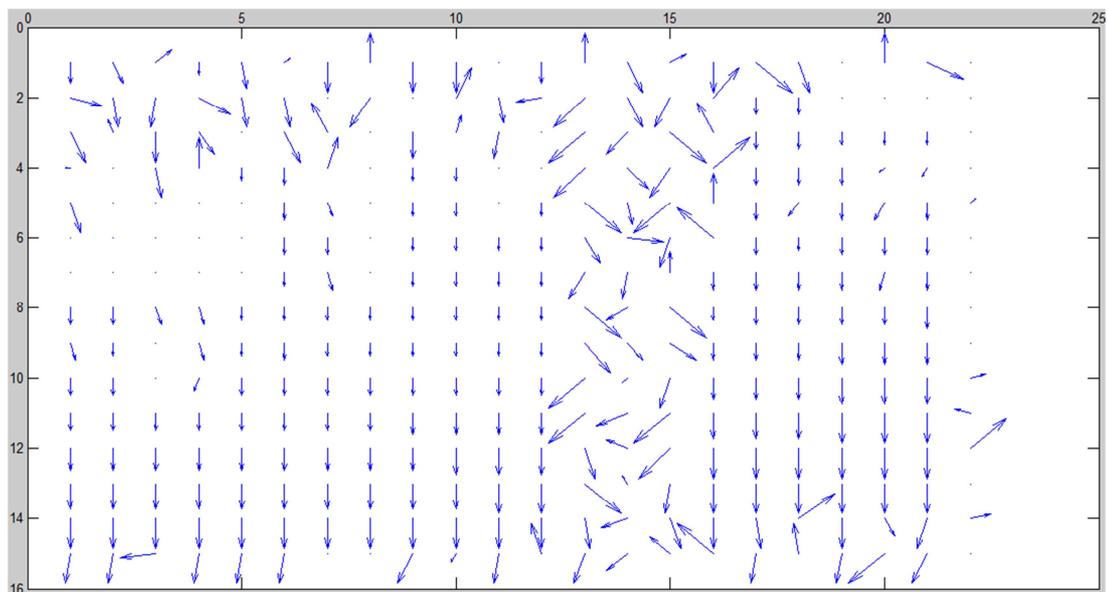
Image 2.b



Image prédite

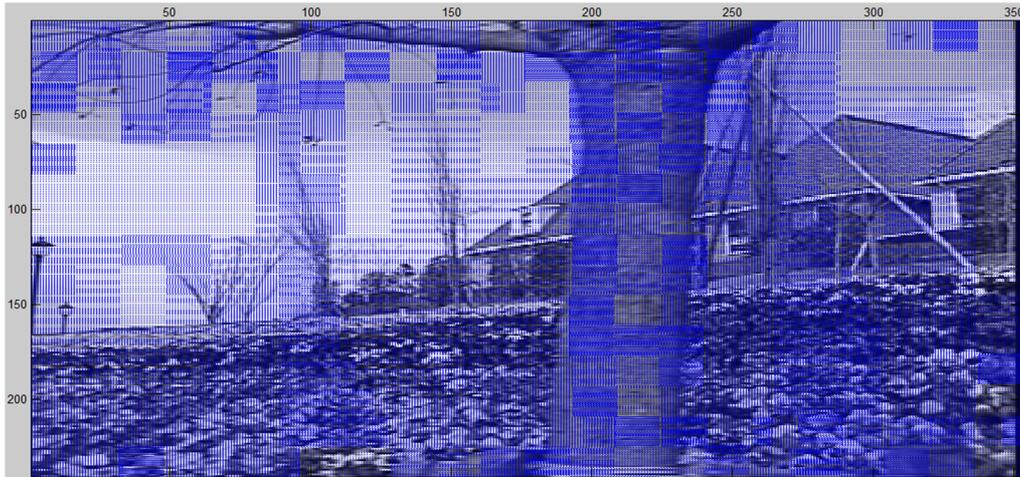


Image erreur



Vecteurs de déplacement

CHAPITRE IV : Applications et discussions



Superposition des vecteurs de déplacement sur l'image

Figure 42 : Résultats de la vidéo 2

Taille du bloc=16

Zone de recherche=7

PSNR=17,3933 db

Image 3 :



Image 3.a



Image 3.b

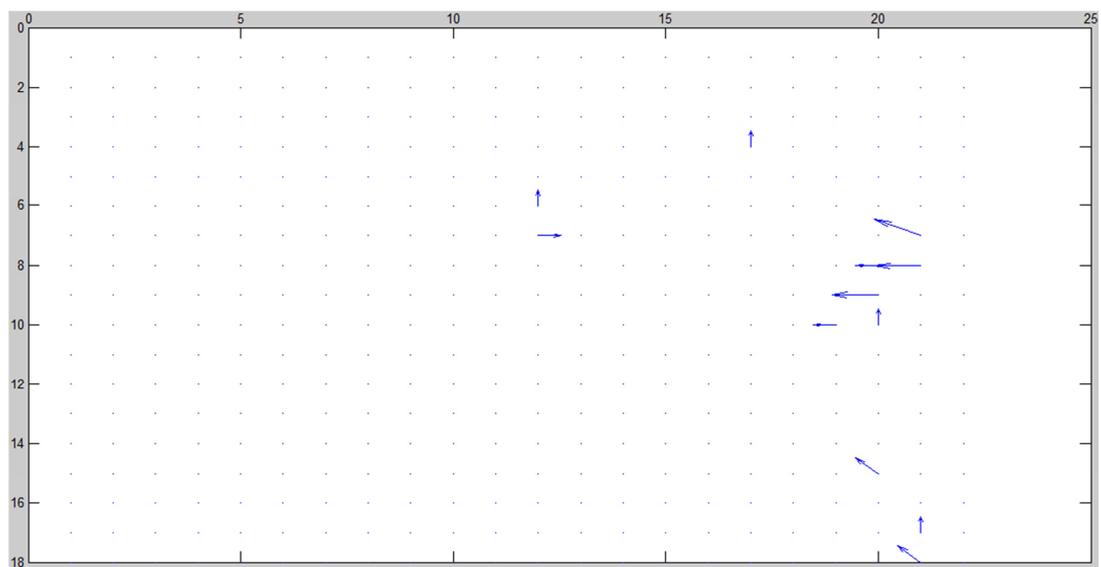


Image prédite



Image erreur

CHAPITRE IV : Applications et discussions



Vecteurs de déplacement



Superposition des vecteurs de déplacement sur l'image

Figure 43 : Résultats de la vidéo 3

Taille du bloc=8

Zone de recherche=7

PSNR=33,4864 db

Image 4 :



Image 4.a



Image 4.b



Image prédite

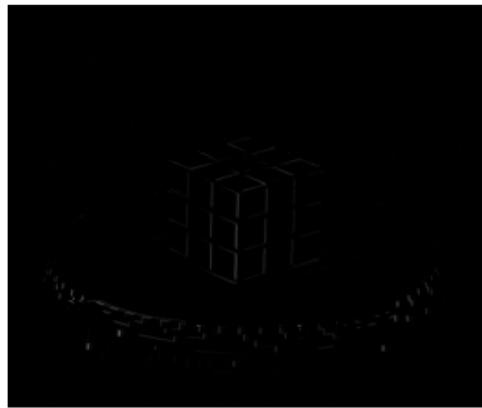
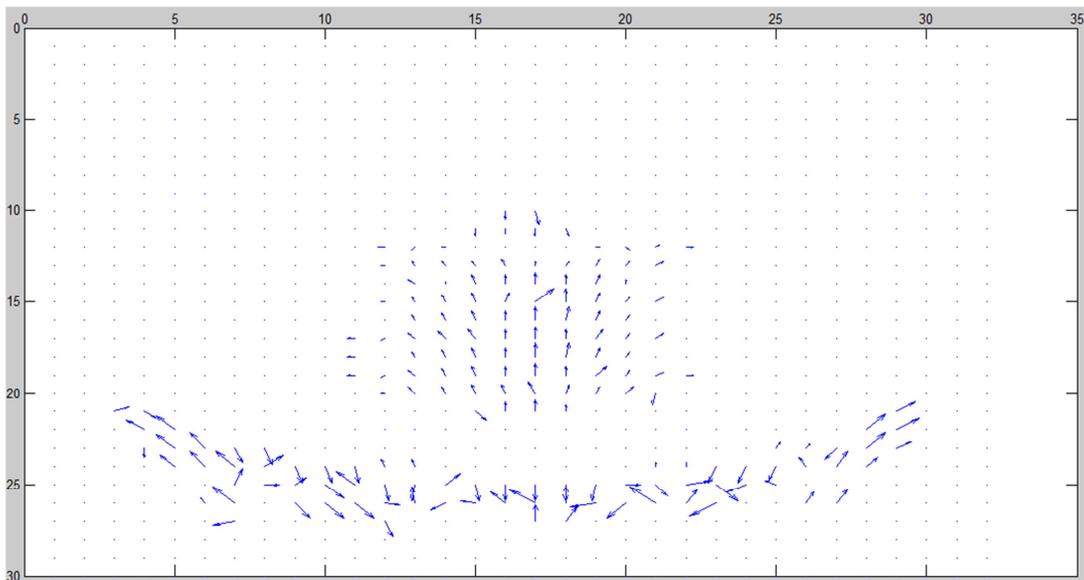
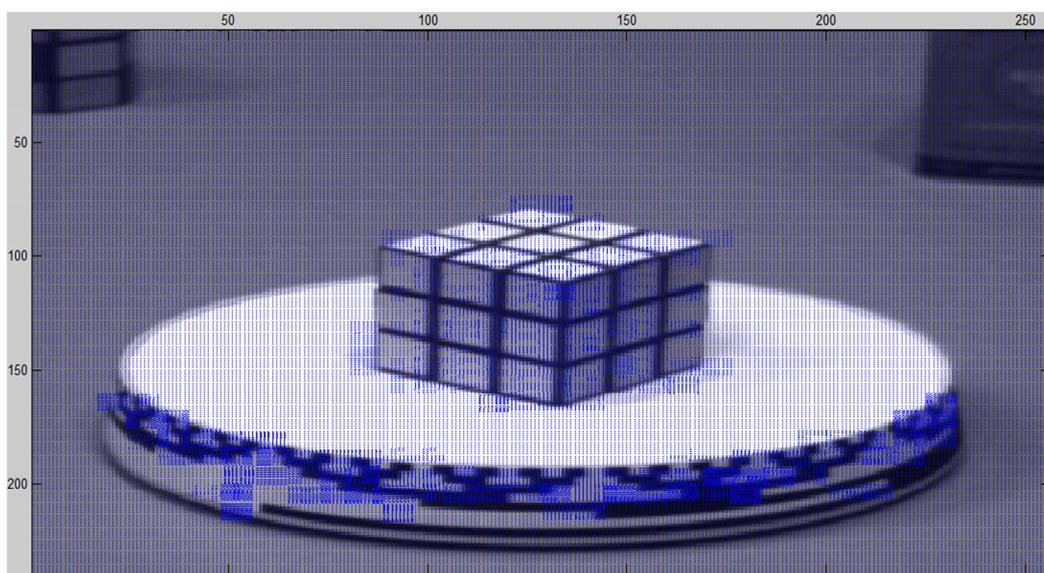


Image erreur



Vecteurs de déplacement

CHAPITRE IV : Applications et discussions



Superposition des vecteurs de déplacement sur l'image

Figure 44 : Résultats de la vidéo 4

Taille du bloc=8

Zone de recherche=7

PSNR=29,0855 db

Image 5 :



Image 5.a



Image 5.b

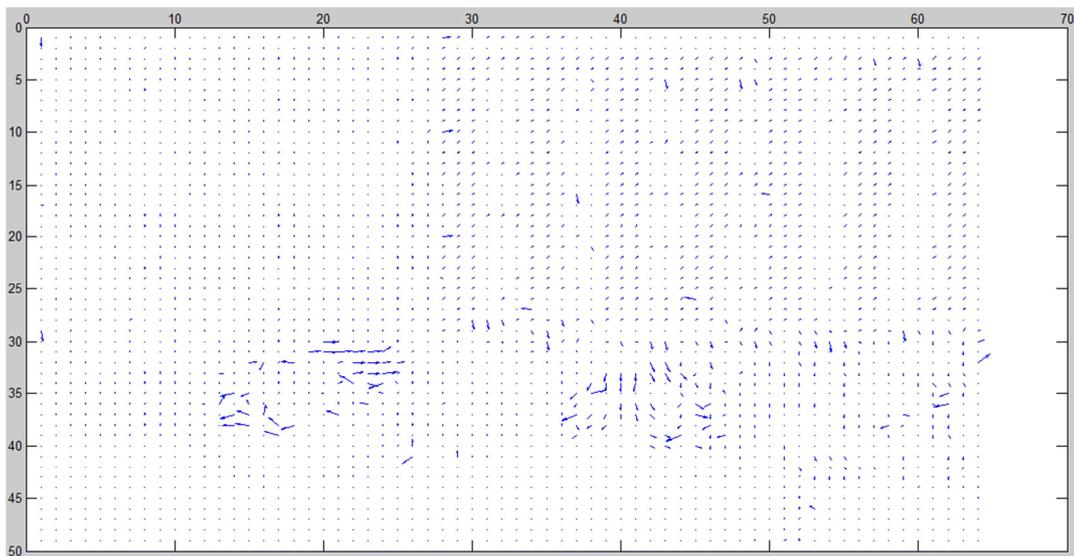
CHAPITRE IV : Applications et discussions



Image prédite



Image erreur



Vecteurs de déplacement

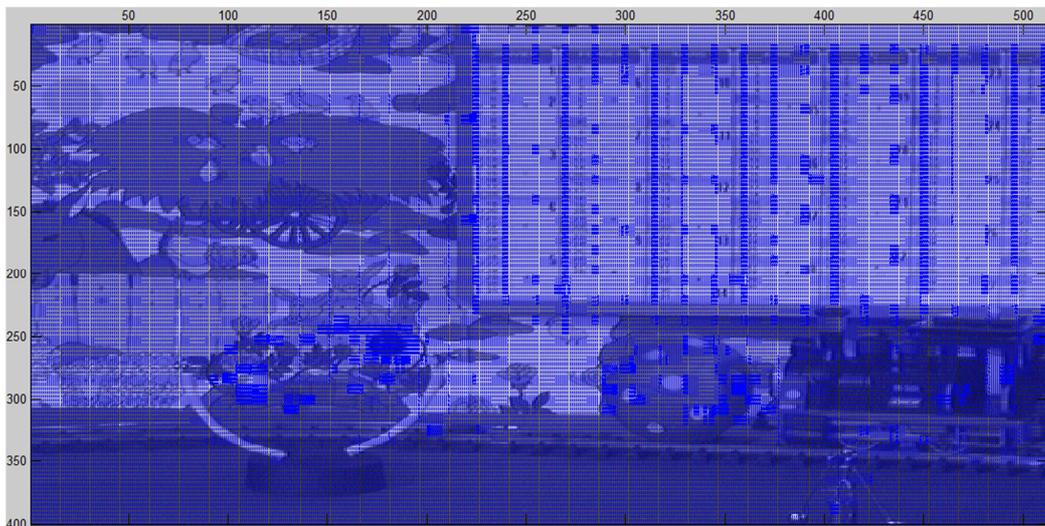


Figure 45 : Résultats de la vidéo 5

Taille du bloc=8, zone de recherche=7 et PSNR=32,7015 db

Image 6 :



Image 6.a



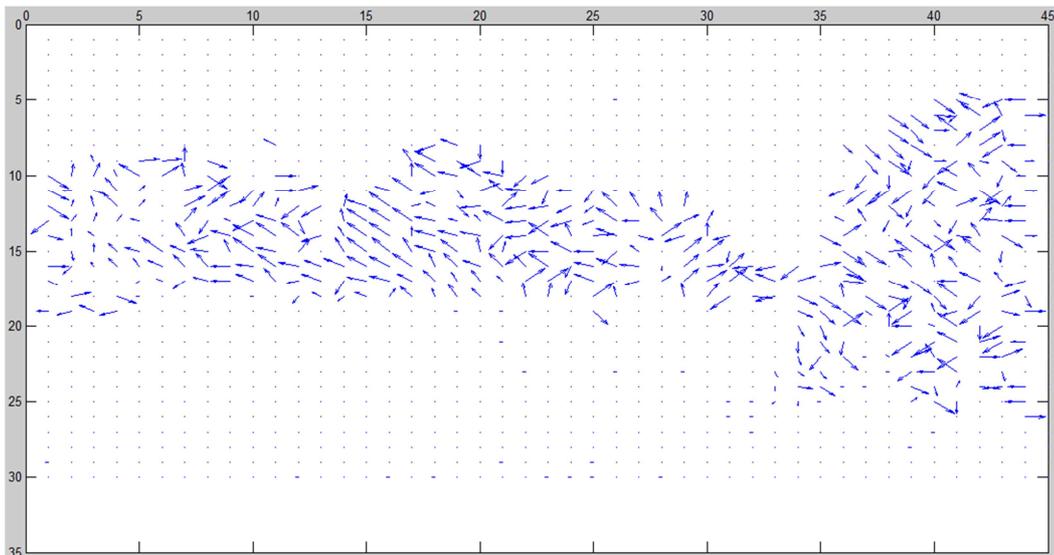
Image 6.b



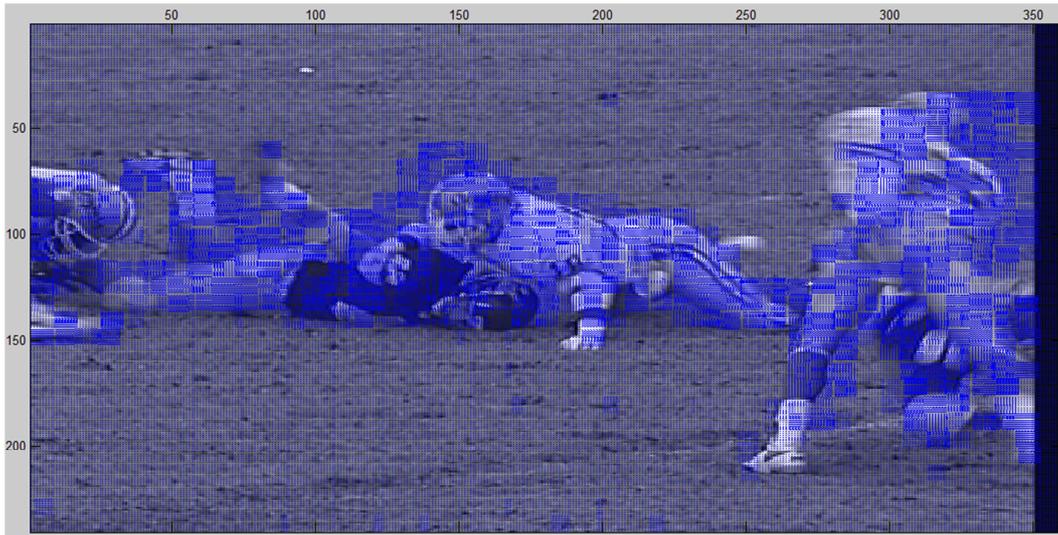
Image prédite



Image erreur



Vecteurs de déplacement



Superposition des vecteurs de déplacement sur l'image

Figure 46 : Résultats de la vidéo 6

Taille du bloc=8

Zone de recherche=7

PSNR=21,3287 db

IV.5 Application aux images METEOSAT

Les images Météosat que nous avons utilisé ont été prises dans le domaine du visible par le satellite météorologique METEOSAT.

Image 1 :

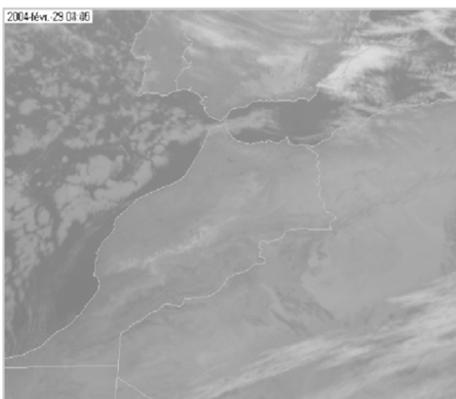


Image 1.a

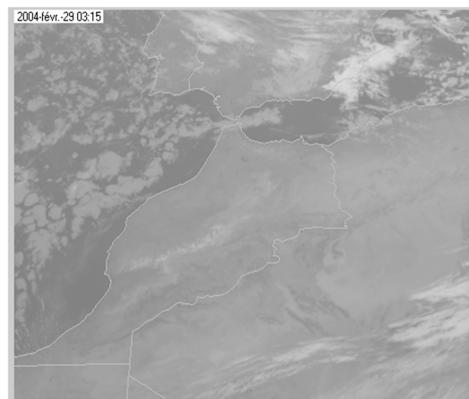


Image 1.b

CHAPITRE IV : Applications et discussions

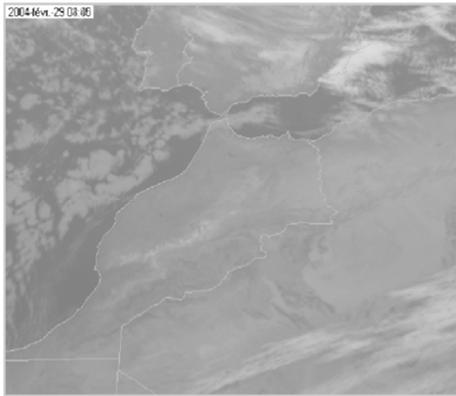
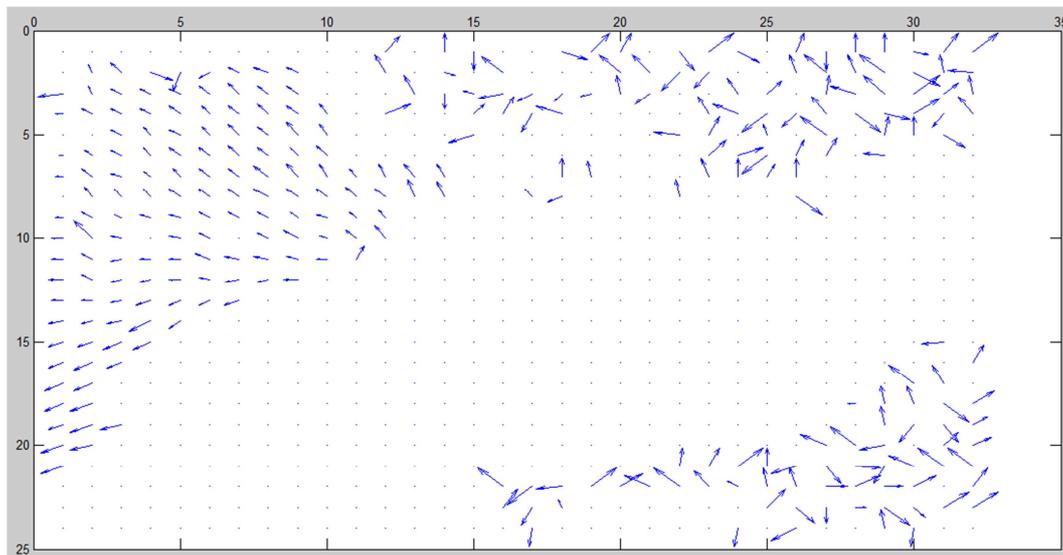


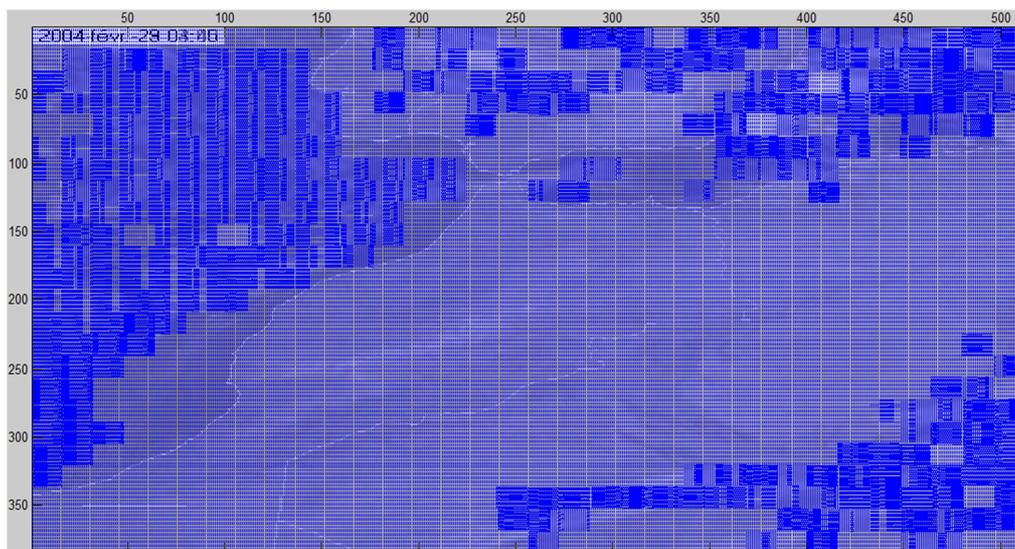
Image prédite



Image erreur



Vecteurs de déplacement



Superposition des vecteurs de déplacement sur l'image

Figure 47 :

CHAPITRE IV : Applications et discussions

Taille du bloc=16

Zone de recherche=7

PSNR=31,4410 db

Image 2 :

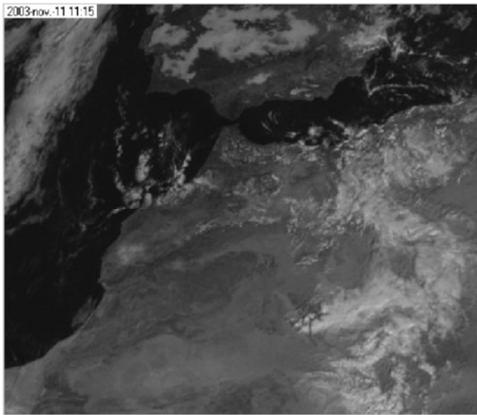


Image 2.a

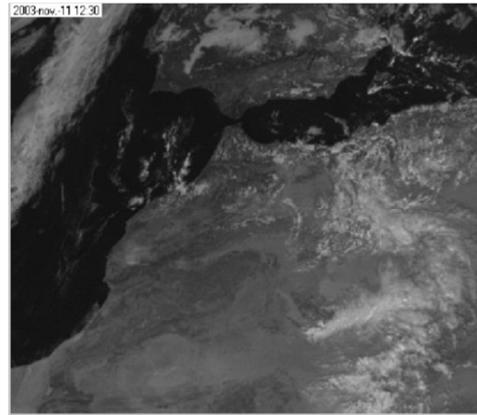


Image 2.b

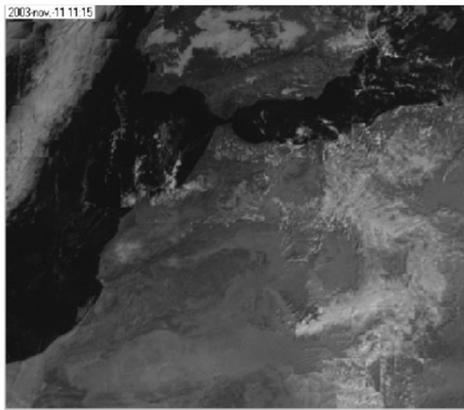


Image prédite

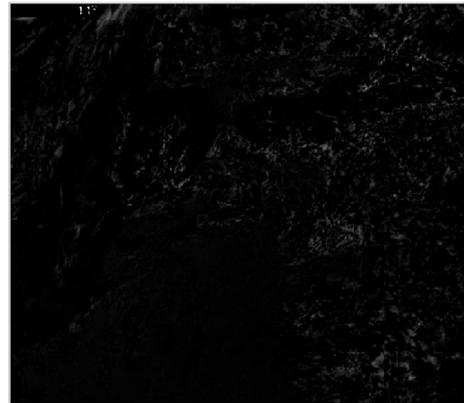
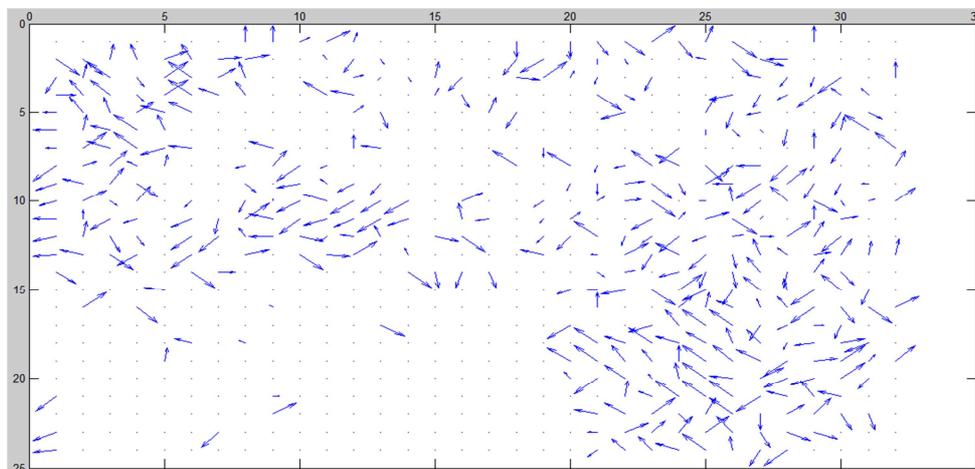
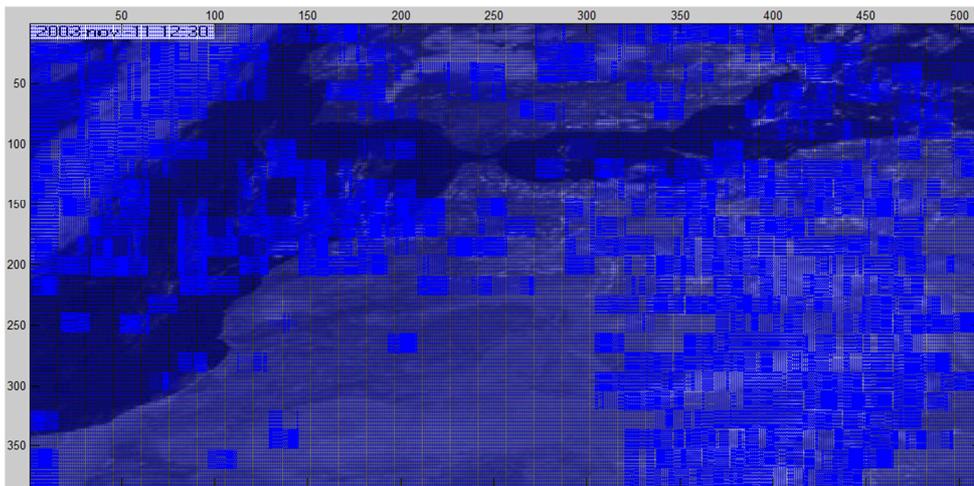


Image erreur



Vecteurs de déplacement

CHAPITRE IV : Applications et discussions



Superposition des vecteurs de déplacement sur l'image

Figure 48 :

Taille du bloc=16, zone de recherche=7

PSNR=24,9808 db

Image 3 :

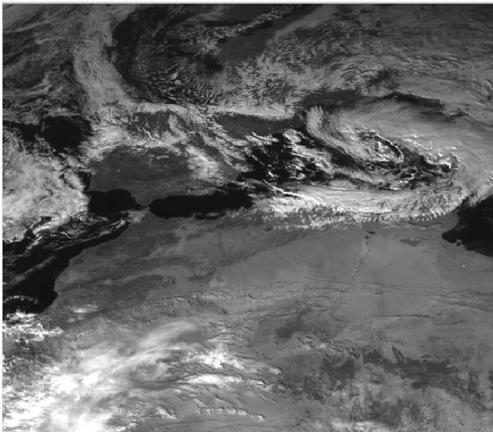


Image 3.a

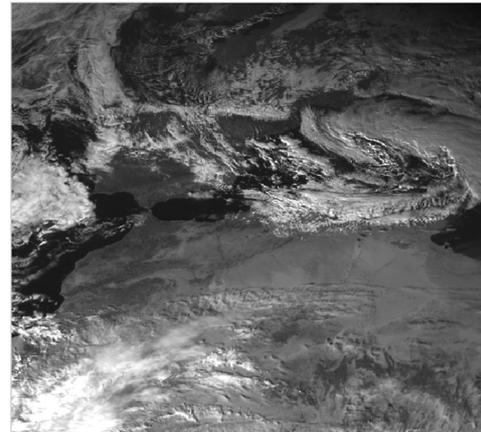


Image 3.b

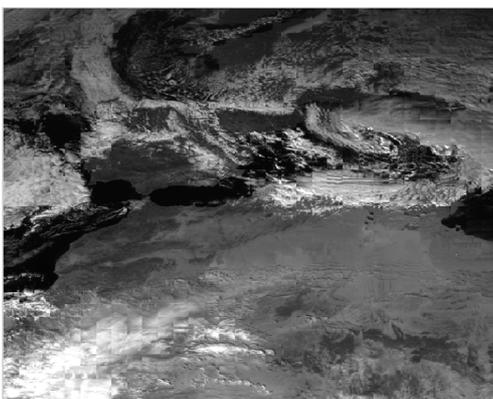
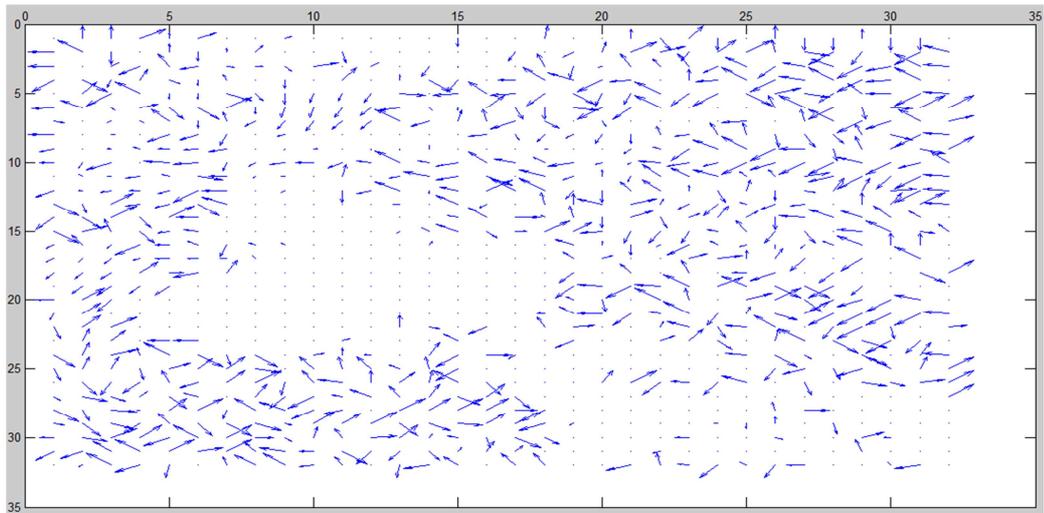


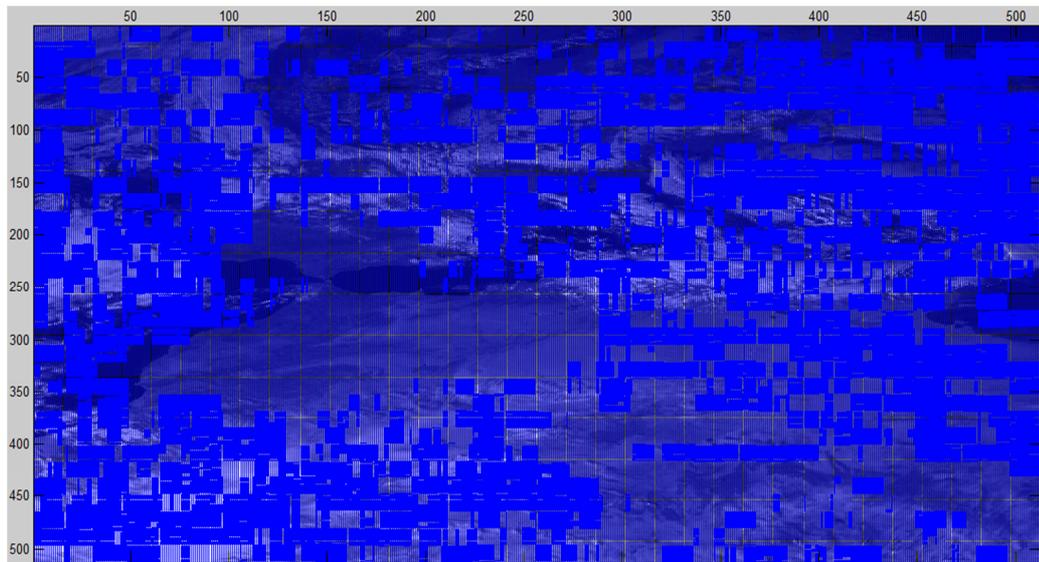
Image prédite



Image erreur



Vecteurs de déplacement



Superposition des vecteurs de déplacement sur l'image

Figure 49 :

Taille du bloc=16,

Zone de recherche=7

PSNR=19,8466 db

IV.6 Interprétation des résultats

Pour chaque image, nous avons estimé le déplacement et les résultats se sont interprétés par les PSNR ainsi que l'image erreur.

CHAPITRE IV : Applications et discussions

Afin de mieux cerner la procédure suivie, voici ses étapes.

Pour chaque image testée, nous avons procédé de la sorte :

L'indice (a) est attribué à chaque image courante et l'indice (b) à l'image de référence, c'est à partir de ces deux images prises à des temps consécutifs qu'on a les vecteurs de déplacement.

Le mouvement ainsi calculé va permettre de prédire l'image $n+1$ (l'image de référence) en combinant l'information contenue dans les pixels présents dans l'image n (l'image courante) et les vecteurs du mouvement.

L'image erreur est la différence entre l'image prédite et l'image de référence, Les zones noires représentent l'absence de la présence du mouvement, contrairement aux zones blanches qui détectent le mouvement.

IV.7 Conclusion

Dans ce travail, nous avons étudié pas mal d'exemples illustratifs sur l'estimation du mouvement sur différentes images synthétiques ainsi que celles de METEOSAT.

D'après l'étude comparative qu'on a faite sur les différents algorithmes du bloc matching d'estimation du mouvement, à savoir : ES, TSS, SS4 et DS, on a constaté que l'algorithme qui nous donne des résultats qualitativement assez appréciables est celui de la recherche exhaustive (ES ou FS) avec des PSNR assez important, car plus la valeur du PSNR est élevée plus le résultat d'estimation du mouvement est satisfaisant.

Conclusion générale

Le sujet abordé dans cette thèse concerne l'estimation du mouvement dans des séquences d'images, synthétiques en générale et météorologiques en particulier. Les informations que peut apporter l'étude des trajectoires et le sens des déplacements des vecteurs de mouvement, peuvent aider à mieux comprendre et à prévoir des phénomènes météorologiques d'envergure tels que les tempêtes tropicales, les ouragans.

Le premier chapitre est consacré à l'étude des différentes notions de base sur le traitement et acquisition d'images, et pour mieux cerner ces dernières, nous avons donné quelques exemples illustratifs.

Vu que notre étude se porte essentiellement sur les images METEOSAT, on s'est étalé sur la présentation des caractéristiques des satellites météorologiques METEOSAT ainsi que le prétraitement et acquisition de ces images.

Au deuxième chapitre, nous avons évoqué les différentes méthodes existantes d'estimation du mouvement dans des séquences d'images que nous avons classées en quatre catégories : Les méthodes différentielles, les méthodes statistiques, les méthodes fréquentielles et les méthodes de mise en correspondance.

Le troisième chapitre, après un tour d'horizon des différentes méthodes d'estimation du mouvement, nous avons présenté une méthode originale de calcul des déplacements qui est le Block Matching.

Au quatrième chapitre, nous avons présenté les différents résultats d'estimation du mouvement fait sur les différentes images synthétiques et METEOSAT avec les différents algorithmes de la méthode du Block Matching.

Nous avons d'abord fait une étude comparative sur deux séquences d'images à savoir la séquence Rubic et la séquence caltrain, les résultats se sont interprétés par le critère d'évaluation PSNR. Cette étude nous ramène à déduire la bonne performance de l'algorithme de la recherche exhaustive (ES) par rapport aux autres algorithmes vu sa bonne précision malgré qu'il est long en temps de calcul et impraticable pour des applications en temps réel.

Conclusion générale

A partir de cette déduction, nous avons appliqué l'algorithme (ES) sur différentes images synthétiques et METEOSAT, et les résultats sont traduits par des images erreurs.

Perspectives

Les méthodes du block Matching sont parmi les méthodes d'estimation du mouvement les plus utilisées dans la pratique. Elles peuvent être utilisées pour :

- La compression vidéo : H.261, H.263, MPEG-1, MPEG-2 ou MPEG-5.
- La détection de mouvement.
- Traking : le suivi des objets en mouvement.
- Matching : images prises dans des conditions différentes.

Ouvrages:

[1]. **ADRIAN BASARAB** « Estimation du mouvement dans des séquences d'images échographiques », Lyon 2008.

[2]. **BAKHOUCHE FARIZA** « Segmentation d'images texturées », Faculté de Génie Electrique et de l'informatique ; Département d'électronique, Mémoire d'ingénieur. UMMTO 2004.

[3]. **KADDOUR BADR, HAMTOUCHE ALI** « Contours actifs paramétriques pour la segmentation d'images et vidéos, I.N.I 2008.

[4]. **NESSRINE ZEGLAM** « Codec de la vidéo appliqué à la visiophonie sur IP », I.N.I, 2008.

[5]. **VINCENT GARCIA** « ESTIMATION DU MOUVEMENT SUBPIXELIQUE PAR BLOCS ADAPTEE A LA COULEUR AVEC MODELE DE MOUVEMENT », CNRS, NICE 2004.

[6]. « Méthodes classiques d'estimation du mouvement », INSA, Lyon.

Documentations :

[7]. ERIC FAVIER. Cours sur l'analyse et le traitement des images, les principes de la vision assistée par ordinateur. ENISE(France).

[8].Caroline Rougier. Cours de traitement d'images. Université de Montréal (Canada).

[9].Antoine Mazanera. Cours de traitement d'images(ENSTRA), France.

Annexe :

L'estimation du mouvement

L'estimation du mouvement est un problème incontournable dans le domaine du traitement de séquences d'images. Le but est d'estimer le mouvement des objets contenus dans une vidéo sous forme d'un champ dense de vecteurs.

Les applications d'une telle estimation sont nombreuses et nous pouvons noter par exemple la segmentation d'objets en mouvement, la compensation du mouvement utilisée en compression vidéo (normes MPEG 2, MPEG 4, H264), etc. L'estimation du mouvement est excessivement coûteuse en temps de calcul.

Cela peut prendre jusqu'à 80% des ressources matérielles de la machine. Une estimation de faible qualité est souvent source de problèmes importants comme par exemple des effets de blocs visuellement gênants. Ainsi, depuis plus de vingt ans, son étude reste un sujet de recherche important.

Estimation du mouvement subpixelique

Méthodes d'interpolation subpixelique

- **Interpolation bilinéaire**

La méthode d'interpolation bilinéaire est l'une des plus simples existant mais elle reste imprécise pour des précisions dépassant le demi-pixel. Elle consiste à faire la moyenne des deux pixels voisins situés à égale distance.

La figure (a) illustre l'interpolation d'une image au demi-pixel. Cette dernière est réalisée en deux étapes :

1- L'interpolation agit premièrement sur les lignes de l'image contenant des pixels réels (marqués en gris et par un « x » sur la figure). Les pixels notés « \diamond » sont obtenus en calculant la moyenne des deux pixels « x » voisins sur la même ligne.

2- Un parcours sur les colonnes est ensuite effectué. Les pixels notés « O » sont obtenus en calculant la moyenne des deux pixels voisins sur la même colonne (les voisins tous deux marqués comme « x » ou comme « \diamond »).

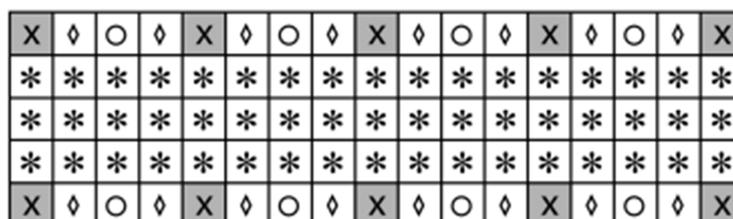


Figure (a) : Interpolation bilinéaire au quart de pixel

X	◇	X	◇	X	◇	X	◇	X
○	○	○	○	○	○	○	○	○
X	◇	X	◇	X	◇	X	◇	X
○	○	○	○	○	○	○	○	○
X	◇	X	◇	X	◇	X	◇	X

Figure (b) : Interpolation H26 au quart de pixel- étape1 : interpolation au demi-pixel

De même, nous pouvons expliquer le fonctionnement de l'interpolation au quart de pixel avec la figure (b).

Comme pour le demi- pixel, l'interpolation au quart de pixel se déroule en deux étapes :

Premièrement l'interpolation en ligne (celles qui contiennent les pixels réels), puis deuxièmement l'interpolation en colonne à partir des lignes complétées. De plus, chaque interpolation nécessite deux opérations qu'on va détailler dans le cas de l'interpolation des lignes :

1-Les pixels notés « O » sont obtenus en calculant la moyenne des deux pixels « x » voisins (à une distance de deux pixels) sur la même ligne.

2-Les pixels notés « ◇ » sont obtenus en calculant la moyenne des deux pixels voisins directs sur la même ligne (« x » et « O »).

L'interpolation en colonne est réalisée de la même manière en procédant en deux étapes

Ce procédé très simple peut naturellement être étendu au huitième de pixel, à toute précision qui est une puissance de deux. Cependant, dès le huitième de pixel, les résultats sont assez peu satisfaisants et d'autres méthodes d'interpolation doivent être employées.

- **Interpolation utilisant la norme H26**

H26 est une norme de codage qui décrit des méthodes d'interpolation pour le quart et le huitième de pixel.

Considérons pour commencer l'interpolation au quart de pixel. Cette interpolation se déroule en deux étapes . Dans un premier temps, l'interpolation va se faire sur les demi-pixels avec l'utilisation d'un filtre numérique.

L'algorithme effectue un parcours sur les lignes d'indices paire (précisions que la première ligne a pour indice 0 et est considérée comme paire). Tous les pixels notés « ◇ » sur la figure (b) sont obtenus en calculant la moyenne pondérée des six voisins

pixéliques (pixels gris et notés par un « × » sur la figure (b)) situés à gauche et à droite sur la même ligne.

Les pondérations sont les suivantes :

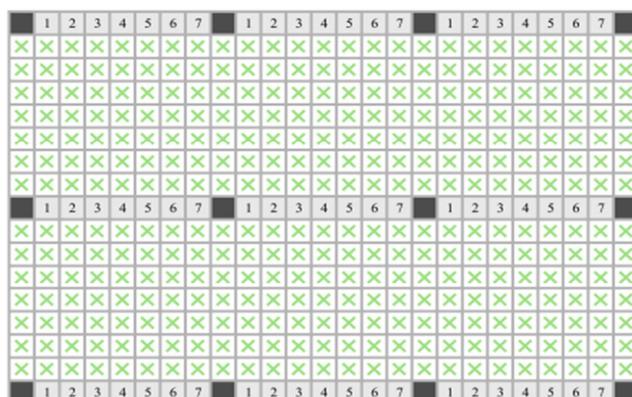


Figure (c) : Interpolation H26L au huitième de pixels

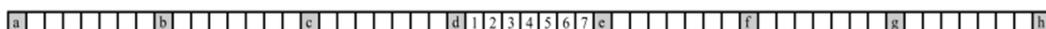


Figure (d) : Représentation d'une portion de ligne

$$\left(\frac{1}{32}, \frac{-5}{32}, \frac{20}{32}, \frac{20}{32}, \frac{-5}{32}, \frac{1}{32} \right)$$

Lorsque tous les demi-pixels des lignes d'indices pairs ont été calculés, nous utilisons le même filtre en colonne pour calculer les lignes d'indice impair.

Nous pouvons à ce stade effectuer deux remarques. Premièrement le résultat est naturellement arrondi à l'entier le plus proche contenu dans l'intervalle [0, 255] généralement utilisé pour décrire la valeur d'un pixel. Deuxièmement, si le filtre déborde de l'image, le pixel situé sur le bord de l'image cumule toute les pondérations des pixels situés à l'extérieur.

L'image est à ce stade interpolée à la précision du demi pixel. La seconde est d'appliquer à cette image interpolée une interpolation bilinéaire au demi pixel. Nous obtenons de cette manière une image interpolée au quart de pixel.

Etudions maintenant l'interpolation au huitième de pixel. Une unique étape est nécessaire pour réaliser cette interpolation. Le principe reste le même que pour toutes les méthodes d'interpolation : l'interpolation complète dans un premier temps les lignes contenant des pixels réels, puis une interpolation sur les colonnes est réalisée à partir des lignes ainsi complétées. De même, comme lors de la première étape de l'interpolation H26L au quart de pixel, un filtre numérique est utilisé.

Annexe

Cependant étant donné que sept pixels sont interpolés entre deux pixels réels (voir figure (c)), le filtre va différer selon la position du pixel entre les pixels réels. La figure (d) permet de fixer les notations employées pour définir, dans le tableau 1, les filtres utilisés par l'algorithme. Sur cette figure les pixels à interpoler sont représentés par des numéros e les pixels réels le sont par des lettres.

Les filtres ainsi définis sont utilisés dans une seconde étape pour réaliser l'interpolation sur les colonnes. Cela permet de calculer les lignes manquantes correspondant aux croix vertes représentées sur la figure (c). A ce niveau, toute l'image a alors été interpolée.

Poids positions	a	b	c	d	e	f	g	h
1	$\frac{-3}{512}$	$\frac{12}{512}$	$\frac{-37}{512}$	$\frac{485}{512}$	$\frac{71}{512}$	$\frac{-21}{512}$	$\frac{6}{512}$	$\frac{-1}{512}$
2	$\frac{-3}{256}$	$\frac{12}{256}$	$\frac{-37}{256}$	$\frac{229}{256}$	$\frac{71}{256}$	$\frac{-21}{256}$	$\frac{6}{256}$	$\frac{-1}{256}$
3	$\frac{-6}{512}$	$\frac{24}{512}$	$\frac{-76}{512}$	$\frac{387}{512}$	$\frac{229}{512}$	$\frac{-60}{512}$	$\frac{18}{512}$	$\frac{-4}{512}$
4	$\frac{-3}{256}$	$\frac{12}{256}$	$\frac{-39}{256}$	$\frac{158}{256}$	$\frac{158}{256}$	$\frac{-39}{256}$	$\frac{12}{256}$	$\frac{-3}{256}$
5	$\frac{-4}{512}$	$\frac{18}{512}$	$\frac{-60}{512}$	$\frac{229}{512}$	$\frac{387}{512}$	$\frac{-76}{512}$	$\frac{24}{512}$	$\frac{-6}{512}$
6	$\frac{-1}{256}$	$\frac{6}{256}$	$\frac{-21}{256}$	$\frac{71}{256}$	$\frac{229}{256}$	$\frac{-37}{256}$	$\frac{12}{256}$	$\frac{-3}{256}$
7	$\frac{-1}{512}$	$\frac{6}{512}$	$\frac{-21}{512}$	$\frac{71}{512}$	$\frac{485}{512}$	$\frac{-37}{512}$	$\frac{12}{512}$	$\frac{-3}{512}$

Tableau 1 : Filtres utilisés pour l'interpolation au huitième de pixel avec la norme H26L