

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTRE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE

Université de Mohamed El-Bachir El-Ibrahimi - Bordj Bou Arreridj

Faculté des Sciences et de la technologie

Département d'Electronique

Mémoire

Présenté pour obtenir

LE DIPLOME DE MASTER

Filière : Sciences et de la technologie

Spécialité : Systèmes de Télécommunication

Par

Mr. KHOUDHOUR Sofiane

Mlle. BELKHITER Rima

Thème

Les algorithmes de Débruitage des Séquences
Vidéo (VBM3D, VBM4D, SPTWO).

Soutenu le : 07/09/2019

Devant le jury :

- | | |
|-----------------------|-----------|
| • Dr.BEHIH Mouhammed | Président |
| • Mme.LKHLEF Noura | Examineur |
| • Pr.MESSALI Zoubeida | Encadreur |

Année Universitaire 2018/2019

Résumé

Dans ce mémoire, nous établissons une étude comparative quantitative approfondie du débruitage vidéo à base de patches avec des algorithmes d'estimation du débit optique. A savoir, SPTWO, VBM3D et VBM4D et notre propre algorithme modifié qui a appelé algorithme Hybrid. Le but de cette étude est de comparer ces algorithmes de débruitage vidéo. SPTWO tire parti de l'auto-similarité et de la redondance des trames adjacentes. L'algorithme hybride proposé et les trois algorithmes de débruitage vidéo sont implémentés et testés sur des séquences réelles dégradées par différents niveaux de bruit blanc gaussien. Les résultats obtenus sont comparés en fonction des critères de performance les plus utilisés pour différents cas de test. Les critères de performance calculés dans cette étude sont : RMSE et SSIM en plus du temps de fonctionnement et de la qualité visuelle de la séquence vidéo. Les résultats expérimentaux montrent que l'algorithme proposé et SPTWO offrent la meilleure qualité vidéo et semblent efficaces en termes de préservation de la texture fine et de reconstruction des détails.

Au début, les premières méthodes de débruitage consistaient à utiliser un filtre passe-bas (suppression des hautes fréquences) ce qui avait pour inconvénient d'atténuer les contours de l'image. Pour parer à ce problème, des techniques nouvelles plus performantes [1]. Pour améliorer la qualité de la vidéo (séquence image) obtenu.

Les algorithmes de débruitage moderne fusionne plusieurs techniques pour obtenir des résultats optimale et proche aux donne originale dans ce manuscrit on va étudiée les méthodes de débruitage des séquences vidéo.

D'après l'article [2] l'auteur fait la comparaison entre les trois algorithmes VBM3D, VBM4D, et sont algorithme SPTWO. Le VBM3D débruité la vidéo par la méthode des patches contrairement les deux autres algorithmes base sur le débruitage des pataches avec l'étape de l'estimation du mouvement par le flux optique comme une étape d'amélioration.

Le dernier point de ce mémoire concerne l'évaluation et la comparaison entre les trios algorithmes (SPTWO, VBM3D, VBM4D) avec l'algorithme Hybride proposé. Ce dernier est destiné principalement à traiter les séquences vidéo fortement bruitées. Les meilleurs résultats sont obtenus par l'algorithme Hybride et SPTWO.

Mots-clés : estimation du mouvement, Flot optique, méthode d'estimation, la méthode Variationnelle, Horn et Shunck, Lucas et Kanade, Block matching, VBM3D, VBM4D, ...

Abstract

In this memoir we established an extensive quantitative comparative study of patch based video denoising with optical flow estimation algorithms. Namely, SPTWO, VBM3D and VBM4D and our own modified algorithm which we called Hybrid algorithm. The aim of this study is to compare these video denoising algorithms. SPTWO takes advantage of the self-similarity and redundancy of adjacent frames. The proposed hybrid algorithm and the three video denoising algorithms are implemented and tested on real sequences degraded by various level white Gaussian noise. The obtained results are compared in terms of the most used performance criteria for various test cases. The performance criteria computed in this study are: RMSE and SSIM in addition to the running time and visual quality of the sequence video. Experimental results, illustrate that the proposed algorithm and SPTWO provide the best video quality and appear to be efficient in terms of preserving fine texture and detail reconstruction.

At first, the first methods of denoising consisted in using a low-pass filter (suppression of high frequencies) which had the disadvantage of attenuating the contours of the image. To overcome this problem, new and more efficient techniques [1]. To improve the quality of the video (image sequence) obtained.

Modern denoising algorithms merge several techniques to obtain optimal results and close to the original data in this manuscript we will study the methods of denoising video sequences.

According to the article [2] the author compares the three algorithms VBM3D, VBM4D, and are SPTWO algorithm. The VBM3D disrupts video by the patch method unlike the other two algorithms based on patch denoising with the step of estimating motion by optical flux as an improvement step.

The last point of this memoir concerns the evaluation and comparison between the three algorithms (SPTWO, VBM3D, VBM4D) with the proposed Hybrid algorithm. The latter is mainly intended for processing highly noisy video sequences. The best results are obtained by the Hybrid algorithm and SPTWO algorithm.

Keywords: motion estimation, Optical flow, estimation method, Variational method, Horn and Shunck, Lucas and Kanade, Block matching, VBM3D, VBM4D,...

ملخص

في هذه الأطروحة، نضع دراسة مقارنة متعمقة كمية من تقليل الضوضاء الفيديو القائمة على التصحيح مع خوارزميات تقدير المعدل البصري. وهي SPTWO و VBM3D و VBM4D والخوارزمية المعدلة الخاصة بنا والتي تسمى الخوارزمية المختلطة Hybrid. الغرض من هذه الدراسة هو مقارنة خوارزميات تقليل الضوضاء بالفيديو هذه. SPTWO يستفيد من التشابه الذاتي والتكرار من الأطر المجاورة. يتم تنفيذ الخوارزمية الهجينة المقترحة وخوارزميات تقليل الضوضاء الفيديو الثلاثة واختبارها على تسلسلات حقيقية تدهورت بمستويات مختلفة من الضوضاء البيضاء الجوسية. تتم مقارنة النتائج التي تم الحصول عليها وفقاً لمعايير الأداء الأكثر استخداماً لحالات الاختبار المختلفة. معايير الأداء المحسوبة في هذه الدراسة هي: RMSE و SSIM بالإضافة إلى وقت التشغيل والجودة المرئية لتسلسل الفيديو. تظهر النتائج التجريبية أن الخوارزمية المقترحة و SPTWO توفران أفضل جودة للفيديو ويبدو أنها فعالة من حيث الحفاظ على التفاصيل الدقيقة وإعادة بناء التفاصيل.

في البداية، كانت أول طرق تقليل الضوضاء هي استخدام مرشح تمرير منخفض (حذف الترددات العالية) والذي كان له عيب في التخفيف من معالم الصورة. للتغلب على هذه المشكلة، تقنيات جديدة وأكثر كفاءة [1]. لتحسين جودة الفيديو (تسلسل الصور) التي تم الحصول عليها.

تقوم خوارزميات تقليل الضوضاء الحديثة بدمج عدة تقنيات للحصول على أفضل النتائج وعلى مقربة من البيانات الأصلية في هذه المذكرة، سوف ندرس أساليب تقليل ضوضاء الفيديو.

من خلال المقال [2] يقوم المؤلف بإجراء مقارنة بين الخوارزميات الثلاث VBM3D و VBM4D ، وخوارزميته الخاصة SPTWO. يقوم VBM3D بتعديل مقطع الفيديو عن طريق التصحيحات المقاطع les patches على عكس الخوارزميتين الأخرين اللتان تستندان إلى تقليل التشويش على pataches مع خطوة تقدير الحركة بواسطة التدفق البصري كخطوة للتحسين.

تتعلق النقطة الأخيرة من هذه الرسالة بالتقييم والمقارنة بين الخوارزميات الثلاث (SPTWO، VBM3D، VBM4D) مع الخوارزمية المختلطة Hybrid المقترحة. وهذا الأخير موجه أساساً لمعالجة الفيديو الضوضاء صاخبة. وتم الحصول على أفضل النتائج بواسطة الخوارزمية المختلطة Hybrid و SPTWO.

الكلمات المفتاحية: تقدير الحركة، التدفق البصري، طريقة التقدير، الطريقة المتغيرة، هورن وشانك، لوكاس و كانادي، مطابقة البلوك، VBM3D، VBM4D، ...

Remerciements

*Nous tenons tout d'abord à remercier **ALLAH** le tout puissant de nous avoir donné le courage et la patience pour mener à bien ce modeste travail, qu'il soit béni et glorifié.*

*Nos vifs remerciements accompagnés de toute notre gratitude vont ensuite à notre encadreur Madame **MESSALI Zoubeida**, Professeur à l'université de Bordj Bou Arréridj, qui a dirigé ce mémoire.*

Nous la remercions aussi pour sa confiance, ses encouragements, ses merveilleuses corrections et pour les conseils qu'elle a apportés pour l'achèvement de ce projet.

Nous tenons également à remercier l'ensemble de membres de jury qui nous ont fait l'honneur de juger notre travail.

Nous tenons aussi à exprimer nos remerciements à tous ceux qui nous ont aidés de près ou de loin durant l'élaboration de notre mémoire de fin d'étude.

A la fin nous remercions nos enseignants durant toute notre formation.

Sofiane et Rima

Dédicaces

De plus profond de mon cœur, je dédie ce travail à tous ceux qui me sont chers

A mes chers merveilleux parents **MAHMOUD** et **BAYA**

Aucune dédicace ne serait exprimer mon respect. Vous représentez pour moi le symbole de la bonté par excellence, la source de tendresse et l'exemple du dévouement qui n'a pas cessé de m'encourager de prier pour moi

Aucun mot ne saurait décrire mon immense amour, ma gratitude et ma profonde considération pour tous les sacrifices que vous avez consenti à mon égard, mon instruction et mon bien être pour tous vous encouragement dès le début de mes études.

Je vous remercie pour le soutien et l'amour que vous me portez depuis mon enfance j'espère que votre bénédiction m'accompagnera toujours.

Que ce modeste travail soit l'exaucement de vœux tant formules

Puisse Dieu vous accorder sante bonheur et longue vie.

A ma chère petite sœur **YASMINA** et à mon frère **IMAD**.

A ma deuxième maman ma tante **FARIDA**.

A ma chère amie **KESSAL IKRAM** avec laquelle j'ai partagé des moments inoubliables au niveau de l'université durant cette année.

Je dédié ce travail aussi à mon binôme **KHOUDHOUR SOFIANE** avec le quelle j'ai réalisé ce travail pour leur effort.

A tous ceux qui je n'ai pas cite ici et qui ont une place dans mon cœur.

RIMA

Dédicaces

Je dédie ce mémoire

A mes chers parents

Pour leur soutien et leur
encouragement durant toute
ma vie

A mon frère

A tous mes amis

A mes collègues de travail

A mes professeurs

A " **SELECT * FROM PEOPLE WHERE FRIENDSHIP != '0'** "

Sofiane

Sommaire

Résumé	02
Abstract	03
ملخص	04
Remerciements	05
Dédicaces	06
Sommaire	08
Liste des abréviations /acronymes	13
Liste des principales notations	14
Liste des figures	15
Liste des tableaux	18
Introduction Générale	19
Motivation	21
Guide de lecture	22

Chapitre 1 Concepts de Base des Séquences Vidéo

1.1 Introduction	24
1.2 Définitions de Base Sur l'Image	24
1.2.1 Séquence Vidéo	26
1.3 La Vidéo.....	26
1.3.1 Types de Vidéo	26
-La vidéo Analogique	26
-La Vidéo Numérique.....	27
1.3.2 Hiérarchie des Données Vidéo	27
1.3.2.1 Séquence Vidéo (Video Sequence)	27
1.3.2.2 Groupe D'Images (Group of Pictures)	27
1.3.2.3 Image (Picture)	27
1.3.2.4 Tranche (Slice)	27
1.3.2.5 Macroblocs (Macroblocks)	27

Sommaire

1.3.2.6 Blocs (Block)	27
1.3.3 Classement des Séquences Vidéo	28
-Séquences Rapides	28
- Séquences Lentes	28
- Séquences Stationnaires	28
-Séquences Simples	28
-Séquences Complexes	29
1.4 Bruit dans La Séquence Vidéo.....	29
1.4.1 Bruit	29
1.4.2 Sources de bruit	30
1.4.3 Types de bruit	30
1.4.3.1 Bruit gaussien	30
1.4.3.2 Bruit poivre et sel (Salt & Pepper noise)	30
1.4.3.3 Bruit Uniforme	31
1.4.3.4 Bruit Périodique	31
1.4.4 Modélisation de bruit	31
1.4.4.1 Bruit exponentiel	32
1.4.4.2 Bruit gaussien	32
1.4.4.3 Bruit poivre et sel (Salt & Pepper noise)	32
1.4.5 Bruit dans Les Vidéos	32
1.5 Evaluation Objective.....	33
1.5.1 Méthodes avec Référence Complète	34
1.5.1.1-Erreur Quadratique Moyenne (MSE)	34
1.5.1.2-Rapport Crête Signal Sur Bruit (PSNR)	34
1.5.1.3-Rapport Signal Sur Bruit (SNR)	34
1.5.1.4-Indice de Similarité Structurale (SSIM)	34
1.6 Conclusion	36
Chapitre 2 Etat de L'Art sur Les Méthodes d'Estimation du Mouvement	
2.1 Introduction.....	38

Sommaire

2.2 Notions Fondamentales de L'Estimation du Mouvement	38
2.2.1 Séquence Numérique	38
2.2.2 Mouvement Global, Mouvement Local	38
2.3 Mouvement dans Les Images.....	39
2.3.1 Mouvement réel	39
2.3.2 Mouvement Apparent	40
2.4 Champ Bidimensionnel de Déplacement, Flot optique, Mouvement Apparent 2D.....	40
Flot Optique	41
2.5 Estimation du Mouvement d'une Image (Equation de Contraintes du Mouvement Apparent: ECMA).....	42
2.6 Méthodes d'Estimation du Mouvement.....	45
2.6.1 Méthodes Différentielles	45
Estimation de mouvement par la méthode Variationnelle	45
2.6.1.1 Algorithme de Horn et Schunck	46
2.6.1.2 Algorithme de Lucas et Kanade (LK)	48
Estimation en multi-résolution	50
2.6.2 Méthode de Mise en Correspondance de Blocs	51
2.6.3 Méthodes Fréquentielles	51
2.7 Conclusion.....	52

Chapitre 3 Algorithmes de Débruitage Pour Les Séquences Vidéos

3.1 Introduction	54
3.2 Estimation du Bruit dans les Séquences Vidéo	54
3.2.1 Régularisation des Séquences D'Images	55
3.3 Types de Débruitage des Séquences Vidéo	56
3.3.1 Débruitage par Filtrage Classique (par masques)	56
1) Filtrage Spatial	56
a) Filtre moyenneur	56
b) Filtre Gaussien	57
2) Filtrage Spatial non Linéaire: Médian	58

Sommaire

3) Filtrage fréquentiel: Wiener	59
3.3.2 Algorithmes de Débruitage à base d'approche variationnelle	60
3.3.2.1 Méthodes de Débruitage Locales	60
3.3.2.2 Méthodes de Débruitage Non Locales	60
3.4 Algorithmes de Débruitage des Séquences Vidéo.....	61
1) L'algorithme VBM3D (grayscale-video denoising)	62
2) L'algorithme VBM4D	63
3) L'algorithme SPTWO	64
• PCA	67
• SVD	67
3.5 Algorithme Proposé: L'Algorithme de Débruitage Hybride	67
3.6 Critères d'Evaluation (de performance)	69
3.6.1-Critère Subjectif (visuel)	69
3.6.2-Critère Objectif	69
3.7 Conclusion	69

Chapitre 4 Résultats expérimentaux et interprétations

4.1 Introduction.....	71
4.2 Nature des données	71
4.2.1 Séquences de Test	71
4.2.1.1 Séquences« army »	71
4.2.1.2 Séquence « yosemite »	71
4.2.1.3 Séquence «cooper »	71
4.2.1.4 Séquence « dog »	71
4.2.1.5 Séquence «truck »	71
4.2.1.6 Séquence «basketball »	72
4.2.1.7 Séquence «Traffic»	72
4.2.1.8 Séquence «Bureau »	72
4.3 Réalisation de notre travail	72
4.3.2.1 Matériel utilisé pour le développement	72
4.3.2.2 Langage de programmation	73

Sommaire

4.4 Résultats et Discussions	73
4.5 Etude Comparative	79
4.6 Conclusion	87
Conclusion Générale	88
Annexe	89
Bibliographie	93

Liste des Abréviations /Acronymes

Pixel	PICTure Element
ppp	point par pixel
bpp	bit par pixel
JPEG	Joint Photographic Experts Group
GIF	Graphical Interchange Format
PNG	Portable Network Graphics
TIFF	Tagged Image File Format
BMP	Bitmap
AI	Adobe Illustrator
PS/EPS	Postscript /En capsulated Postscript
SVG	Scalable Vector Graphics
PDF	Portable Document Format
FLA/SWF	Flash
CCD	Charge Couple Device (dispositif a transfert de charge)
LCD	Liquid Cristal Display (l'écran à cristaux liquides)
CMOS	Complementary Metal Oxide Semi-conductor
MSE	Erreur Quadratique Moyenne
SNR	Signal to Noise Ratio
PSNR	Rapport Crête Signal Sur Bruit
SSIM	Indice de Similarité Structurelle
SVH	Système Visuel Humain
ECMA	L'équation de contrainte du mouvement apparent
ROI	Région d'intérêt utilisé pour l'estimation du mouvement

Liste des Principales Notations

t	instant
dt	temps infinitésimal
N	Bruit gaussien
I	Image
σ	Écart type
σ^2	Variance
μ	Moyenne
p	La fonction de densité du bruit gaussien
T	Le nombre de pixels de chaque fenêtre
x_i, y_i	L'intensité d'un pixel
∇I	Le gradient spatial de l'image
\vec{u}	Vecteur vitesse au pixel \vec{x}
\vec{v}	Vecteur vitesse au pixel \vec{y}
\vec{V}	La vitesse du mouvement
d	Déplacement
$\hat{I}(x)$	Image estimé

Liste des Figures

Figure 1.1	Ensemble de pixel	24
Figure 1.2	Image couleur	25
Figure 1.3	Image niveaux de gris et ses nuances de gris (256 niveaux de gris)	25
Figure 1.4	Hiérarchie des données dans le flux vidéo	28
Figure 1.5	Exemple de séquence vidéo avec (a) sa représentation en 2D (b) sa Représentation volumique	29
Figure 1.6	Exemple d'image dégradée par un bruit : (a) gaussien, (b) poivre et sel, (c) Uniforme, (d) Périodique.....	31
Figure 1.7	Distribution de bruit	32
Figure 1.8	Images avec le bruit gaussien	33
Figure 1.9	Diagramme de l'Indice de Similarité Structurale (SSIM)	35
Figure 2.1	Projection mouvement 3D en 2D	39
Figure 2.2	Exemple de flot optique	40
Figure 2.3	Illustration du champ de déplacement bidimensionnel	40
Figure 2.4	Illustration du flot optique (mouvement apparent 2D)	41
Figure 2.5	La séquence de test Yosemite À gauche et au milieu : une paire d'images extraites de la séquence. À droite : le flot optique estimé par la méthode de B.Horn and B.Schunck	41
Figure 2.6	déplacement d'un point dans deux images successives dans le temps I, I+1	43
Figure 2.7	illustration géométrique de l'équation et de la droite de contrainte du mouvement	43
Figure 2.8	Problème de l'ouverture.....	44
Figure 2.9	Principe d'occlusion	44
Figure 2.10	Implémentation pyramidale d'une méthode de calcul du flot optique	50
Figure 2.11	Principe des méthodes de mise en correspondance de blocs	51
Figure 3.1	Estimation du bruit	55

Liste des Figures

Figure 3.2 Application du filtre moyennneur	57
Figure 3.3 Application du filtre gaussien	58
Figure 3.4 Principe du filtre médian	59
Figure 3.5 Application du filtre médian	59
Figure 3.6 Application du filtre de Wiener	60
Figure 3.7 L'algorithme VBM3D	63
Figure 3.8 L'algorithme VBM4D	64
Figure 3.9 Exemples de voisinage spatio-temporel non local. Le pixel central (en rouge) est caractérisé par un vecteur de niveaux de gris défini grâce à un patch de taille 3x3x1	64
Figure 3.10 les étapes de l'estimation de mouvement par l'algorithme SPTWO	65
Figure 3.11 les étapes de l'algorithme TV-L1	66
Figure 3.12 Organigramme de l'algorithme Hybrid	68
Figure 4.1 les séquences de utilise dans notre travail	72
Figure 4.2 Image débruité par différente méthodes ((a) Image originale, (b) Image bruite avec $\sigma=30$, (c) Image filtre par le filtre gaussien, (d) Image débruitée par le filtre médian, (e) Image débruitée par le SPTWO, (f) Image débruitée par le VBM3D, (g) Image débruitée par le VBM4D, (h) Image débruitée par le Hybrid)	73
Figure 4.3 Comparaison des différents résultats de débruitage de quatre algorithmes sur la frame de la séquence de test "army" zoomée sur une région	74
Figure 4.4 Image débruité par différente méthodes ((a) Image originale, (b) Image bruite avec $\sigma=30$, (c) Image filtre par le filtre gaussien, (d) Image débruitée par le filtre médian, (e) Image débruitée par le SPTWO, (f) Image débruitée par le VBM3D, (g) Image débruitée par le VBM4D, (h) Image débruitée par le Hybrid)	75
Figure 4.5 Comparaison des différents résultats de débruitage de quatre algorithmes sur la frame de la séquence de test "army" zoomée sur une région	77
Figure 4.6 les frames de la séquence Bureau qui a une particularité le changement brusque de la Lumière	78
Figure 4.7 une frame de la séquence Bureau débruité par différente méthodes ((a) Image originale, (b) Image bruite avec $\sigma=30$, (c) Image débruitée par le SPTWO, (d) Image débruitée par le VBM3D, (e) Image débruitée par le VBM4D, (f) Image débruitée par le Hybrid)	87
Figure 4.8 SSIM de la séquence "Traffic" débruité par SPTWO, VBM4D, VBM3D, Hybride	84
Figure 4.9 la continuité du flux optique dans le débruitage des vidéos par estimation du mouvement	84

Liste des Figures

Figure 4.10 la moyenne de RMSE des 2 séquences noir et blanc	85
Figure 4.11 la moyenne de SSIM des 2 séquences noir et blanc	85
Figure 4.12 la moyenne de RMSE des séquences en couleur	86
Figure 4.13 la moyenne de SSIM des séquences en couleur	86

Liste des Tableaux

Tableau 4.1 Les valeurs de RMSE en fonction de σ_n	79
Tableau 4.2 La moyenne de RMSE des 2 séquences NB	80
Tableau 4.3 Les valeurs de SSIM en fonction de σ_n	80
Tableau 4.4 La moyenne de SSIM des 2 séquences NB	80
Tableau 4.5 Les valeurs de RMSE en fonction de σ_n	81
Tableau 4.6 La moyenne de RMSE des 7 séquences	82
Tableau 4.7 Les valeurs de SSIM en fonction de σ_n	82
Tableau 4.8 La moyenne de SSIM des 7 séquences	83

Introduction Générale

Le problème de débruitage consiste à restituer le plus fidèlement possible le signal d'origine, à partir d'une version bruitée.

Plusieurs algorithmes ont été élaborés, Le choix d'un algorithme efficace pour le débruitage des séquences vidéo reste encore un défi scientifique. Dans ce qui suit nous allons faire une brève revue sur quelques travaux significatifs dans le domaine de débruitage des séquences d'image. L'analyse du mouvement est un vaste sujet qui englobe un certain nombre de problématiques.

Le but de ce travail est de faire un débruitage optimale des séquences vidéo en temps real en tenant compte de l'estimation du mouvement.

Nous allons donc, dans un premier temps, traiter une séquence vidéo bruitée avec bruit blanc gaussien cette séquence on va les débruité par des algorithmes de débruitage (SPTWO, VBM4D, VBM4D, Hybride) suit les étapes suivant :

Pour estimer un mouvement dans une vidéo (une série d'images séquentielles), la première étape du processus consiste à mesurer le déplacement des objets de la scène observée : une estimation de ce déplacement est fournie par le calcul du flot optique de chaque frame (image).

Le calcul du flot optique est une étape de bas niveau en traitement d'images, permettant d'estimer le déplacement des objets d'une scène [4] Toutefois, les estimations denses et précises du flot optique sont habituellement coûteuses en temps de calcul. L'étude s'attachera donc à analyser les techniques existantes de calcul du flot optique, puis à choisir une méthode parallélisable donnant des résultats répondant aux attentes, par des algorithmes d'estimation de manière à fournir des estimées en temps réel.

Nous décomposons ce travail en quatre chapitres.

Chapitre1 : Dans le premier chapitre, on présente des généralités sur l'image numérique et la vidéo anime (séquence d'image), ainsi le bruit est ces types. Finalement l'évaluation objective de la qualité des images est effectuée.

Chapitre 2 : Le deuxième chapitre est consacré à la description des méthodes usuelles d'estimation du mouvement à partir d'une séquence d'images. Après une brève discussion sur le mouvement et les problèmes généraux de sa mesure, nous présenterons les principales techniques d'estimation proposées dans le domaine de la vision par ordinateur. A la fin de ce Chapitre, nous exposerons la méthode différentielle pour l'estimation du mouvement ainsi qu'une conclusion.

Introduction Générale

Chapitre 3 : le troisième chapitre nous étudierons les algorithmes récents de débruitage des séquences vidéo. Plus précisément, les algorithmes VBM3D, VBM4D, SPTWO seront abondamment détaillés.

Nous présenterons l'algorithme hybride de débruitage que nous avons proposé. Cet algorithme permet de prendre la meilleure frame en terme de niveau du bruit, i.e. le niveau du bruit le plus faible et concaténer toutes les frames pour restituer la séquence vidéo débruitée. Plus précisément, les trois algorithmes sont appliqués sur chaque frame séparément. Un comparateur de la déviation standard du bruit est effectué. La frame qui a le bruit le plus faible est gardée. Cette procédure est appliquée jusqu'à ce que toutes les frames de la séquences sont débruitées. Nous nous sommes basés ainsi sur la robustesse des trois algorithmes.

Chapitre 4 : Le quatrième Chapitre est consacré à la partie pratique, Les expériences de simulation de l'approche proposée ainsi que les différents algorithmes considérés dans notre étude sont réalisés à l'aide des logiciels MATLAB, C++ et le python en utilisant des séquences de teste composée de huit frames. L'approche proposée a pour but de traiter des séquences vidéo fortement bruitées. L'évaluation des algorithmes de débruitage (SPTWO, VBM3D, VBM4D) avec notre nouvel algorithme (méthode hybride) en utilisant est effectuée en utilisant l'Erreur Quadratique Moyenne et l'indice de Similarité Structurale comme critères d'évaluation de performance des algorithmes.

Ces résultats montrent que l'algorithme proposé a de fortes potentialités de surpasser les autres algorithmes (VBM4D, VBM3D, SPTWO) en terme de performance par conséquent une meilleure qualité de débruitage sans un changement important de la complexité du traitement.

Motivation

La réduction du bruit dans les images est un problème délicat et largement traité par les scientifiques. La problématique primordiale, est de réussir à éliminer un maximum de bruit, tout en préservant les structures et les détails de l'image. En général, l'image observée est contaminée par un bruit qui est soit additif ou multiplicatif. Les traitements de restauration sont généralement nécessaires pour améliorer la qualité des images observées.

Dans la littérature du traitement d'image, plusieurs méthodes de débruitage d'image et vidéo ont été proposées et développées. Chaque approche ayant ses propres hypothèses, ses avantages, et ses limites. Toutefois, beaucoup de ces méthodes enlèvent en plus du bruit des détails de l'image. De ce fait, des algorithmes de débruitage continuent à être développés pour remédier à ce problème.

Guide de Lecture

Ce manuscrit est constitué de la présenté un résumé en trois langue avec une introduction générale en suit la motivation et le guide de lecture. On a décomposé la manuscrit on quatre Chapitres illustrant les différents aspects de nos travaux et d'une conclusion à chaque chapitre. Nous avons ajouté, en début de chaque Chapitre, un résumé et une introduction à chaque chapitre encore le contexte et les enjeux de la partie traitée dans le dit Chapitre, ainsi que les travaux effectués.

Le premier Chapitre étudie des généralités sur l'image numérique et la vidéo anime (séquence d'image), ainsi le bruit est ces types. Finalement l'évaluation objective de la qualité des images est effectuée.

Dans la deuxième Chapitre, nous présenterons les méthodes usuelles d'estimation du mouvement à partir d'une séquence d'images. Après une brève discussion sur le mouvement et les problèmes généraux de sa mesure, nous présenterons les principales techniques d'estimation proposées dans le domaine de la vision par ordinateur. A la fin de ce Chapitre, nous exposerons la méthode différentielle pour l'estimation du mouvement ainsi qu'une conclusion.

Le troisième Chapitre nous étudierons les algorithmes récents de débruitage des séquences vidéo. Plus précisément, les algorithmes VBM3D, VBM4D, SPTWO seront abondamment détaillés.

Nous présenterons l'algorithme hybride de débruitage que nous avons proposé,

Cet algorithme est permet de prendre la meilleure frame en terme de niveau du bruit, i.e. le niveau du bruit le plus faible et concaténer toutes les frames pour restituer la séquence vidéo débruitée. Plus précisément, les trois algorithmes sont appliqués sur chaque frame séparément. Un comparateur de la déviation standard du bruit est effectué. La frame qui a le bruit le plus faible est gardée. Cette procédure est appliqué jusqu'à e que toutes les frames de la séquences sont débruitées. Nous nous sommes basés ainsi sur la robustesse des trois algorithmes.

Le quatrième Chapitre et consacre à la partie pratique, Les expériences de simulation de l'approche proposée ainsi que les différents algorithmes considérés dans notre étude sont réalisés à l'aide des logiciels MATLAB, C++ et le python en utilisant des séquences de teste composée de huit frames. L'approche proposée a pour but de traiter des séquences vidéo fortement bruitées. L'évaluation des algorithmes de débruitage (SPTWO, VBM3D, VBM4D) avec notre nouvel algorithme (méthode hybride) en utilisant est effectuée en utilisant l'Erreur Quadratique Moyenne et l'indice de Similarité Structurale comme critères d'évaluation de performance des algorithmes.

L'interprétation de ces résultats annonce une grande réussite de l'algorithme proposé qui a permis d'améliorer les performances de débruitage, tel que la minimisation de l'erreur quadratique inférieure ce qui signifie à une à son tour l'amélioration de la précision, sans pour autant augmenter la complexité du traitement.

Chapitre 1

Concepts de Base des Séquences Vidéo

Résumé

Le traitement d'images est un domaine très vaste qui a connu, et qui connaît encore, un développement important depuis quelques dizaines d'années.

Dans ce chapitre, nous abordons les notions de base de l'image et la vidéo (vidéo analogique et vidéo numérique...). Le bruit est également traité. Finalement l'évaluation objective de la qualité des images est effectuée.

Sommaire

1.1 Introduction	24
1.2 Définitions de Base Sur l'Image	24
1.3 La Vidéo.....	26
1.4 Bruit dans La Séquence Vidéo	29
1.5 Evaluation Objective	33
1.6 Conclusion	36

1.1 Introduction

Le traitement d'images est un domaine très vaste qui a connu, et qui connaît encore, un développement important depuis quelques dizaines d'années.

On désigne par *traitement d'images numériques* l'ensemble des techniques permettant de modifier une image numérique afin d'améliorer ou d'en extraire des informations. De ce fait, le traitement d'images est 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.

Dans ce chapitre, nous abordons les Concepts de base des séquences. Ensuite, nous allons donner un aperçu sur le bruit.

1.2 Définitions de Base Sur l'Image

L'image est une représentation d'une personne ou d'un objet par la peinture, le dessin, la photographie, le film, ...etc. C'est aussi un ensemble structuré d'informations qui, après 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 fonction d'intensité lumineuse et de couleur. Pour voir l'image sur une machine il faut les numérisé.

Une image numérique est une matrice de pixels repérés par leurs coordonnées (x, y) . S'il s'agit d'une image couleur, un pixel est codé par 3 composantes (r, g, b) (chacune comprise au sens large entre 0 et 255), représentant respectivement les doses de rouge, vert et bleu qui caractérisent la couleur du pixel.

S'il s'agit d'une image en niveau de gris, il est codé par une composante comprise au sens large entre 0 et 255, représentant la luminosité du pixel [3]. Le pixel (abréviation de PICTURE Element) représente ainsi le plus petit élément constitutif d'une image numérique. L'ensemble de ces pixels est contenu dans un tableau à deux dimensions constituant l'image numérique [5].

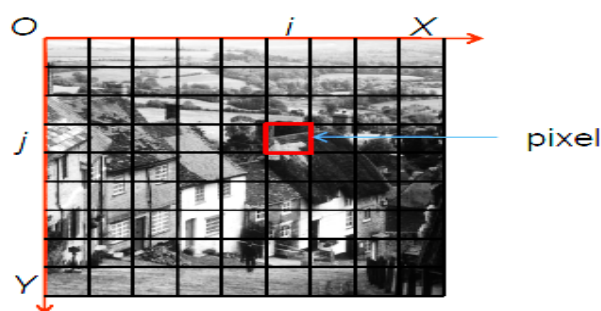


Figure 1.1 Ensemble de pixel [6]

On trouve deux types d'image : Les images Matricielles et Les images Vectorielles

L'image matricielle ou en anglais bitmap : Ce type d'image est adapté à l'affichage sur écran mais peu adapté pour l'impression car bien souvent la résolution est faible (couramment de 72 à 150 ppp pour les images sur Internet) [7].

Dans les images vectorielles les données de l'image sont représentées à l'aide de formules mathématiques.

Cela permet alors d'agrandir l'image indéfiniment sans perte de qualité et d'obtenir un faible encombrement [7].

Lors de son enregistrement une image est stockée suivant un format d'image précis. Ce format doit permettre de stocker l'information de l'image avec un minimum de perte d'informations. Il existe ainsi différents formats qui pourront favoriser soit la conservation de la qualité soit la diminution de la taille du fichier informatique [8]. Pour les formats matriciels, nous citons **JPEG**, **JPEG 2000** (Joint Photographic Experts Group). **GIF** (Graphical Interchange Format) **PNG** (Portable Network Graphics). **TIFF** (Tagged Image File Format) **BMP** (Bitmap).

Pour les formats vectoriels nous citons **AI** (Adobe Illustrator).**PS/EPS** (Postscript /Encapsulated Postscript) **SVG** (Scalable Vector Graphics). **PDF** (Portable Document Format). **FLA/SWF** (Flash).D'après la couleur on constate deux mode d'image : image couleur et image niveaux de gris

-Mode couleurs indexées: permet d'obtenir plus de 256 couleurs fixes, définies à l'avance dans une palette. Si le codage du pixel se fait sur 8 bits (bpp) pour chaque couleur (r,v,b), on aura donc $2^{3 \times 8} = 2^{24}$ possibilités [5]. Un exemple d'une image couleur et une image niveaux de gris est illustré sur les Figures 1.2 et 1.3 respectivement.

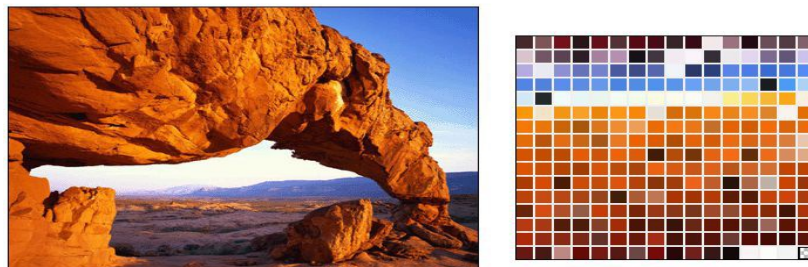


Figure 1.2 Image couleur [5].

- Mode niveau de gris : il permet d'obtenir différentes valeurs de gris, afin d'afficher des images nuancées. Exemple : Codage en 8 bits par pixel (bpp) => $2^8 = 256$ possibilités



Figure 1.3 Image niveaux de gris et ses nuances de gris [5] (256 niveaux de gris).

Après avoir introduit les concepts de base de l'image, nous aborderons dans ce qui suit les séquences vidéo, appelées également images dynamiques.

1.2.1 Séquence Vidéo

Une séquence vidéo brute est une suite d'images fixes, qui peut être caractérisée par trois principaux paramètres :

- ✓ **La résolution en luminance** : Détermine le nombre de nuances ou de couleurs possibles pour un pixel. Celle-ci est généralement de 8 bits pour les niveaux de gris et de 24 bits pour les séquences en couleurs.
- ✓ **La résolution spatiale** : Définit le nombre de lignes et de colonnes de la matrice de pixels, tel qu'il est le cas pour une seule image.
- ✓ **La résolution temporelle** : Est le nombre d'images par seconde.

La valeur de ces trois paramètres détermine l'espace mémoire nécessaire pour stocker chaque image de la séquence. Cet espace mémoire est caractérisé par le débit, qui est le coût de stockage pour une seconde (capacité mémoire nécessaire pour stocker une seconde de vidéo) [9].

1.3 La Vidéo

La vidéo est une succession d'images animées défilant à une certaine cadence afin de créer une illusion de mouvement pour l'œil humain. Elle peut être analogique (signal continu d'intensité de luminance) ou numérique (suite de trames ou images).

On peut distinguer deux types de vidéo :

- ✚ la vidéo entrelacée : où chaque image est formée de deux champs entrelacés. Le premier champ contenant uniquement les lignes impaires de l'image, le second champ, quant à lui, contenant les lignes paires (les signaux de télévision nord-américain et européen sont entrelacés).
- ✚ la vidéo progressive : dans une vidéo progressive, contrairement au mode entrelacé, toutes les lignes d'une image sont lues en une passe (vidéo numérique) [3].

« Images animées » => Vidéo

1.3.1 Types de Vidéo

On distingue deux grandes familles de systèmes vidéo : les systèmes vidéo analogiques et les systèmes vidéo numériques.

✚ La vidéo Analogique

La caméra balaye l'image bidimensionnelle qu'elle a devant elle par un faisceau d'électrons qui se déplace très rapidement de gauche à droite et plus lentement de haut en bas et produit une tension en fonction du temps. Elle enregistre ainsi l'intensité lumineuse, et à la fin du balayage, on a alors une trame. Le faisceau revient à l'origine pour recommencer.

Le récepteur va recevoir cette intensité en fonction du temps, et pour reconstruire l'image, va répéter le processus de balayage [10].

✚ La Vidéo Numérique

Le processus de captation de l'image vidéo en mode numérique est essentiellement le même que pour la vidéo analogique. Un système optique sépare la lumière en trois composants. À la

sortie toutefois, plus d'enregistrement d'un signal électrique mais plutôt l'enregistrement d'une valeur numérique définie pour chacune des trois couleurs de chacun des pixels (RVB).

Le processus de transformation d'un signal électrique et valeurs numériques s'appelle la numérisation ou l'échantillonnage. Deux paramètres servent à échantillonner un signal électrique : la fréquence d'échantillonnage et la quantification de l'échantillonnage ou le nombre de bits utilisés pour le codage [11].

1.3.2 Hiérarchie des Données Vidéo

La hiérarchie des données vidéo est présentée dans la Figure 1.4

1.3.2.1 Séquence Vidéo (Video Sequence)

Qui commence par un en-tête de séquence, contient un ou plusieurs groupes d'images et s'achève par un code de fin de séquence.

1.3.2.2 Groupe D'Images (Group of Pictures)

Qui regroupe un en-tête et une série d'une ou plusieurs images facilitant l'accès direct.

1.3.2.3 Image (Picture)

Qui est l'unité élémentaire pour le codage de la séquence vidéo. Une image est un groupe de trois matrices rectangulaires qui représentent la luminance (Y) et la chrominance (Cb et Cr). Cette représentation YCbCr est équivalente à celle RGB (Red, Green, Blue). Elle lui est préférable, car l'œil étant plus sensible à la luminosité qu'à la chrominance, il n'est pas nécessaire de stocker autant d'informations dans les matrices Cb et Cr que dans la matrice Y, alors qu'en RGB, dimension deux fois plus petite que la matrice Y.

1.3.2.4 Tranche (Slice)

Macroblocs ordonnés de gauche à droite, puis de haut en bas. En cas d'erreur dans la tranche, le décodeur passe à la suivante. Plus il y a de tranches, meilleur est le traitement des erreurs, mais fait perdre de la place.

1.3.2.5 Macroblocs (Macroblocks)

C'est une matrice rectangulaire de dimension 2 et constituée de blocs.

1.3.2.6 Blocs (Block)

C'est un ensemble des valeurs de luminance et chrominance de 8 lignes de 8 pixels [12].

La Figure 1.4 illustre cette hiérarchie des données vidéo.

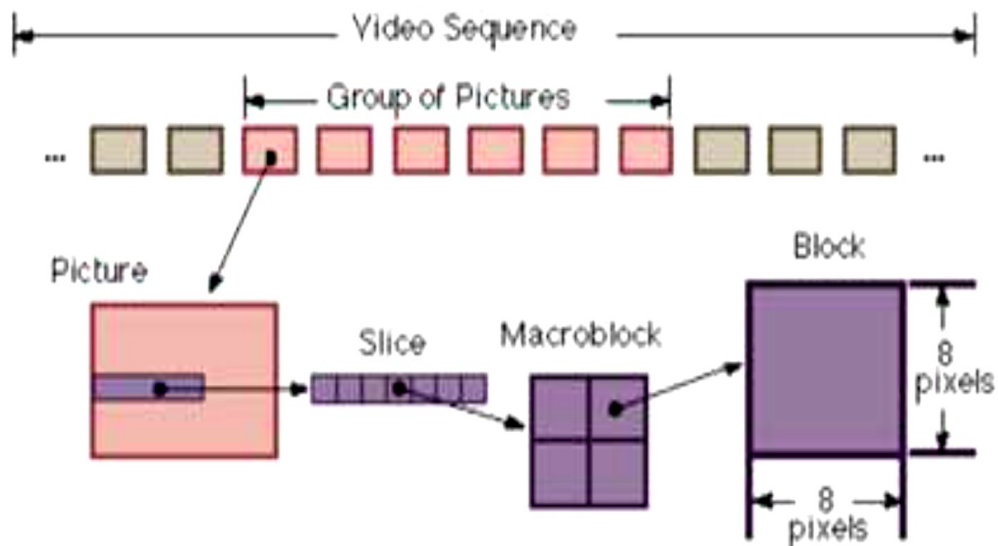


Figure 1.4 Hiérarchie des données dans le flux vidéo [12].

1.3.3 Classement des Séquences Vidéo

▪ Séquences Rapides

Ce sont des séquences où les objets se déplacent d'une image à une autre avec une Distance importante qui peut même dépasser les limites de la zone de recherche prédéfinie, ce qui empêche une compression optimale (les séquences des scènes sportives par exemple). Ce type de séquences nécessite donc l'utilisation de grandes zones de recherche pour atteindre les objets en mouvement, au détriment du temps de calcul bien sûr. Des techniques de recherche permettant de minimiser le nombre de blocs testés sont développées pour surmonter le problème du temps de calcul [13].

▪ Séquences Lentes

Ce sont les séquences où les objets se déplacent lentement ou restent fixes (pour la Plupart du temps) telles que l'exploration d'une scène par une caméra, les séquences de visiophonie ou les séquences des mobiles de la troisième génération. Dans ce cas, une petite zone de recherche suffit pour atteindre un très bon taux de compression en un temps de calcul réduit [12].

▪ Séquences Stationnaires

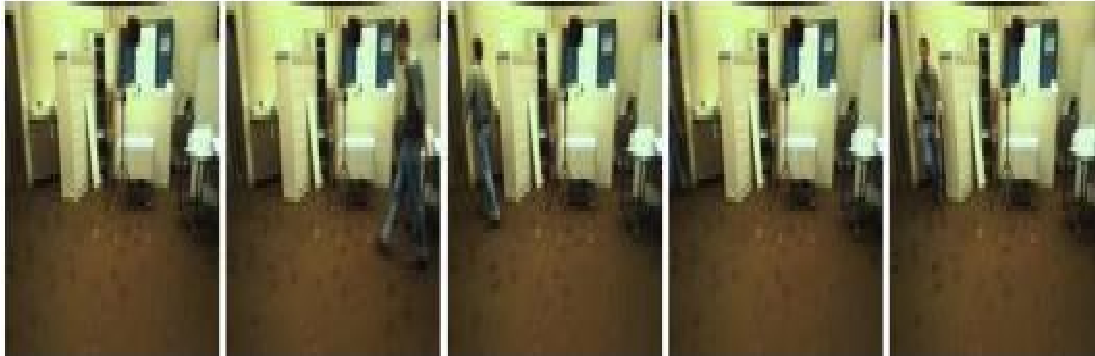
Ce sont les séquences où les blocs gardent la même position d'une image à l'autre tel que les arrières plans. Dans ce cas, le taux de compression est maximal et le temps de recherche est nul si on sait d'avance qu'une portion d'une séquence est stationnaire [13].

▪ Séquences Simples

Dans ce type de séquences, les images ne contiennent pas beaucoup de détails et les objets Connaissent le même mouvement en direction et en vitesse (cas du mouvement de la caméra) :c'est-à-dire que les pixels du même objet se déplacent aussi à la même vitesse et la même direction, ce qui permet de trouver des techniques qui convergent rapidement vers le bloc le plus ressemblant [13].

▪ Séquences Complexes

C'est le contraire du dernier type. Les objets connaissent un mouvement aléatoire (ne peut être prévu) c'est-à-dire qu'un objet peut être jugé meilleur localement mais ne l'est pas globalement (problème des minima locaux) ce qui peut conduire à de faibles résultats en taux de compression et en qualité [13]. La Figure 1.5 montre un exemple d'une séquence vidéo ainsi que sa représentation 2D.



(a) Une séquence de 200 trames vue comme une succession d'images 2D. (Seules les trames 1, 27, 93, 151 et 200 sont représentées.) [13]



(b) Vue compacte de la même séquence obtenue en empilant toutes les images le long de la troisième dimension

Figure 1.5 Exemple de séquence vidéo avec (a) sa représentation en 2D (b) sa représentation volumique [13].

1.4 Bruit dans La Séquence Vidéo

1.4.1 Bruit

Un bruit (parasite) dans une image ou dans une vidéo est considéré comme un phénomène de brusque variation de l'intensité d'un pixel par rapport à ses voisins, il provient de l'éclairage des dispositifs optiques et électroniques du capteur [14]. Ou toute brusque fluctuation d'un pixel par rapport à ses voisins étant considérée comme un bruit provoquant ainsi la dégradation de la qualité d'image que ce soit de luminance ou de contraste [15].

1.4.2 Sources de bruit

Les sources de bruit sont très variées. On peut distinguer plusieurs familles de sources.

- La première source est liée au contexte d'acquisition :
 - ✓ Une sur ou sous illumination réduit l'intervalle de couleurs de la scène mais pas son nombre alors que le nombre de couleurs utilisées pour représenter cette scène sera réduit dans le cas d'un appareil photo numérique.
 - ✓ La perturbation des capteurs, par exemple les perturbations magnétiques pendant une acquisition IRM (Imagerie par Résonance Magnétique), ou des perturbations sur une antenne de télévision lors de la réception sur une carte tuner. Certains capteurs induisent des distorsions géométriques ou d'intensité.
- L'étape d'échantillonnage est source de bruit, notamment le phénomène de moiré lorsque les conditions de Shannon et Nyquist ne sont pas respectées, ou de bruit poivre et sel lorsque des objets de la scène sont projetés dans une image de la taille d'un pixel.
- La quantification apporte un bruit dit de quantification.

Le bruit n'est pas le seul phénomène qui peut dégrader la qualité de l'image ; il y a ce que l'on nomme l'image floue. Cette dégradation peut être causée par plusieurs facteurs entre autres ; le mouvement pendant la prise de la scène [16].

1.4.3 Types de bruit

1.4.3.1 Bruit gaussien

C'est le type de dégradation produite par les composantes électroniques du capteur (relation linéaire avec la température du capteur) et liée à la limite en faible lumière de celui-ci. C'est donc le bruit qui entache majoritairement les images numériques et que nous considérerons plus tard.

Pour créer synthétiquement ce bruit additif gaussien non corrélé (blanc), une variable aléatoire gaussienne a été ajoutée par le système à l'image "idéale". Dans le cas précédent des photos numériques, la source est la précision du capteur CCD ou CMOS mise en évidence par un gain élevé (figure 1.6 (a)) [17].

1.4.3.2 Bruit poivre et sel (Salt & Pepper noise)

Ce bruit peut être créé par la saturation de plusieurs capteurs. Certains pixels sont saturés. Visuellement, cela se traduit par des points blancs et noirs répartis avec une certaine densité dans l'image (figure 1.6 (b)). La source la plus fréquente est un pixel mort dans le capteur dû au processus de fabrication des CCD (que l'on retrouve dans celle des moniteurs LCD). Les appareils photos les "cachent" en les détectant pour chaque appareil avant commercialisation et en utilisant un mécanisme de remplacement par la valeur d'un pixel voisin lors de la prise de photos.

Ce type de bruit est exprimé par le pourcentage du nombre total de pixels dans l'image, dont la valeur est 0 ou 255. Le bruit généré est distribué uniformément dans toute l'image [18].

1.4.3.3 Bruit Uniforme

Une variable aléatoire uniforme (figure 1.6 (c)) a été ajoutée par le système à l'image "idéale". C'est généralement un cas d'école non réaliste mais simple. Visuellement, il est proche du bruit gaussien (seule la distribution diffère) [17].

1.4.3.4 Bruit Périodique

Un bruit de période (figure 1.6 (d)) fixe a été ajouté par le système à l'image "idéale". La source est par exemple le tramage dû à un scannage papier, ou à une compression JPEG (blocs de 8 pixels) ou une interférence haute fréquence sur une image produite par un moniteur utilisant un canon à électron [17].

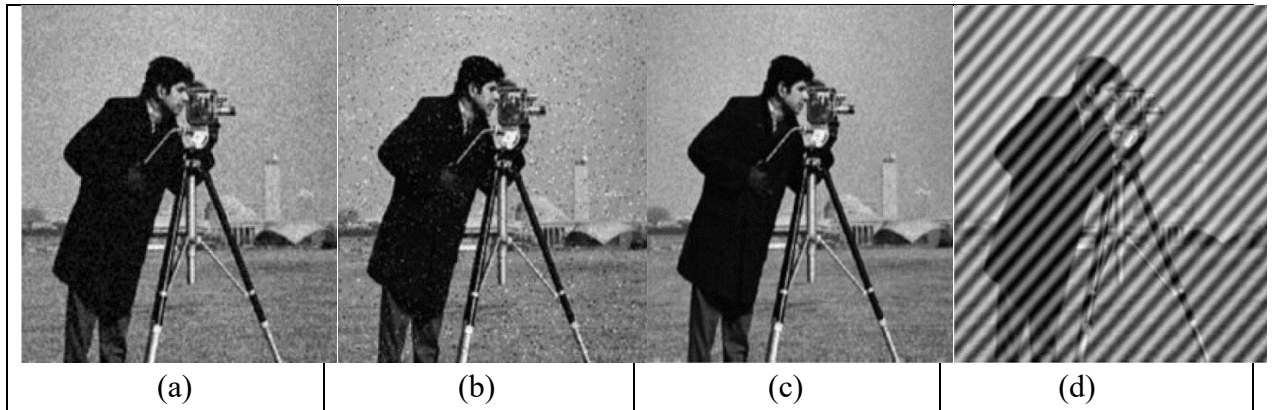


Figure 1.6 Exemple d'image dégradée par un bruit : (a) gaussien, (b) poivre et sel, (c) Uniforme, (d) Périodique [17].

1.4.4 Modélisation de bruit

Le bruit peut être dépendant des données de l'image (par exemple le bruit de quantification) ou indépendant (par exemple les poussières sur l'objectif).

On peut alors modéliser le bruit comme additif ou multiplicatif. Il est bien plus simple de traiter un bruit additif que multiplicatif. Afin d'éliminer le bruit, on peut considérer qu'il concerne des hautes fréquences non présentes dans l'image et qu'il suffit d'effectuer un filtre passe-bas pour améliorer l'image. Malheureusement, cela n'est pas toujours aussi simple.

Le bruit d'image est en toute rigueur considéré comme un champ aléatoire. Du fait de sa haute fréquence, on ne conserve pour le caractériser que le premier ordre (densité de probabilité f sur un pixel de l'image) et parfois le second (corrélacion entre pixels).

On ne conservera ici que le premier ordre, modélisé par :

$$f(a) = C e^{-k|a|^\alpha} \quad (1.1)$$

Avec C et k des constantes de normalisation liées à la variance

1.4.4.1 Bruit exponentiel

Pour $\alpha=1$, il s'agit d'un bruit exponentiel

$$f(a) = C e^{-k|a|} \quad (1.2)$$

1.4.4.2 Bruit gaussien

Pour $\alpha=2$, il s'agit d'un bruit gaussien

$$f(a) = C e^{-k|a|^2} \quad (1.3)$$

Le bruit blanc est un bruit gaussien de moyenne nulle, sa variance est σ^2 .

1.4.4.3 Bruit poivre et sel (Salt & Pepper noise)

Le bruit poivre et sel modélise assez bien les poussières sur un objectif ou scanner, des petits objets sur l'image, on imagine par exemple un objet clair sur un fond foncé et dont la taille dans l'image serait proche du pixel ; il pourrait apparaître ou disparaître lors d'une séquence vidéo créant ainsi du bruit, ainsi que des pertes de données [16]

1.4.5 Bruit dans Les Vidéos

La dégradation de la qualité de l'image (vidéo) ou du son est plus ou moins perceptible plus ou moins progressive et dépend du rapport signal sur bruit S/N ou SNR (Signal to Noise Ratio) [19].

Un des bruits les plus fréquents dans les images ou les vidéos, est le bruit blanc gaussien.

Le bruit gaussien comme c'est illustré dans la Figure 1.7 est une forme idéalisée de bruit blanc, qui est provoqué par des fluctuations aléatoires dans le signal. Nous pouvons observer le bruit blanc en regardant une télévision qui est mal accordée à un canal particulier. Si l'image est représentée comme I , et le bruit gaussien par N , alors nous pouvons modéliser une image bruitée en additionnant simplement les deux [20]. C'est le cas du bruit additif:

$$F = I + N \quad (1.4)$$

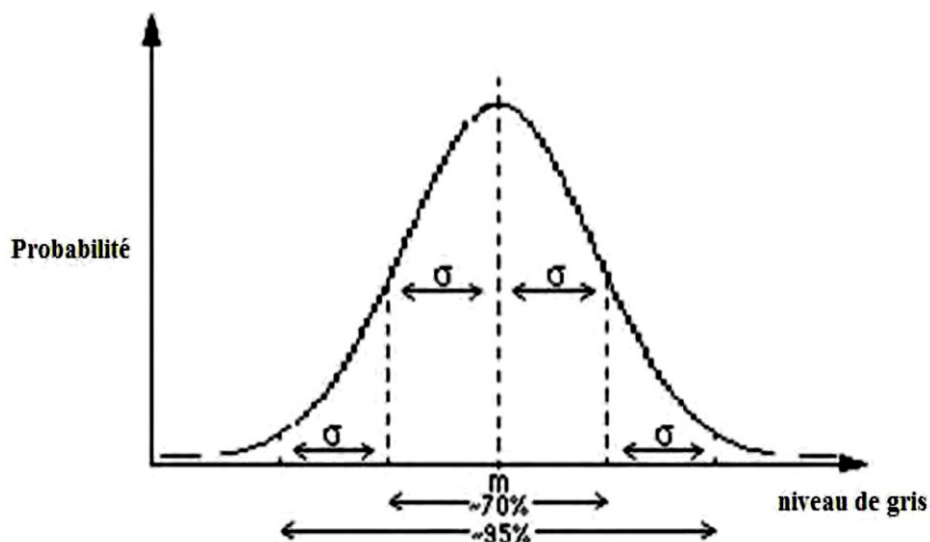


Figure 1.7 Distribution de bruit [20].

Avec les paramètres du bruit (m, σ) qui représentent les statistiques d'ordre un et deux du bruit : la moyenne et l'écart type.

L'ajout d'un bruit blanc gaussien sur une image modifie donc la valeur de chaque pixel (voir figure 1.8).



Figure 1.8 Images avec le bruit gaussien.

La fonction de densité du bruit gaussien, p , avec la moyenne μ et la variance σ^2 est

$$p_q(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} \quad (1.5)$$

1.5 Evaluation Objective

En un mot, la mesure de la qualité objective (par opposition à l'évaluation subjective de qualité par les observateurs humains) cherche à déterminer la qualité des images algorithmiquement. Le but de la recherche de l'évaluation de la qualité objective est de concevoir des algorithmes dont la prévision de la qualité est en accord avec l'évaluation subjective des observateurs humains.

Les méthodes d'évaluation de la qualité d'image (la vidéo) peuvent être classifiées dans trois larges catégories. [20]

- **Les méthodes avec référence complète**, dans lesquelles l'algorithme a un accès à une version parfaite de l'image avec lesquelles il peut comparer la version dégradée. La version parfaite vient généralement d'un dispositif de haute qualité d'acquisition, après elle est dégradée par la compression et les erreurs de transmission.
- **Les méthodes sans référence**, dans lesquelles l'algorithme n'a accès qu'au signal déformé et doit estimer la qualité du signal sans connaissance de la version parfaite. Puisque les méthodes sans référence n'exigent aucune information de référence, elles peuvent être employées dans n'importe quelle application où la mesure de qualité est exigée.
- **Méthodes avec référence réduites**, dans lesquelles l'information partielle concernant la version parfaite est disponible. Un côté canal existe par lequel quelque information concernant la référence peut être rendue disponible à l'algorithme d'évaluation de la qualité. Les algorithmes réduits de référence emploient cette information de référence partielle pour juger la qualité du signal déformé.

Les mesures quantitatives les plus utilisées sont : l'erreur quadratique moyenne (MSE), le rapport crête signal sur bruit (Peak Signal to Noise Ratio, PSNR), le rapport signal sur bruit (Signal to Noise Ratio : SNR) etc...

1.5.1 Méthodes avec Référence Complète

1.5.1.1 Erreur Quadratique Moyenne (MSE)

L'image dégradée \hat{I} est toujours comparée à l'originale I pour déterminer son rapport de ressemblance. Ce critère est le plus utilisé. Il est basé sur la mesure de l'erreur quadratique moyenne (MSE) calculée entre les pixels originaux et dégradés :

$$MSE = \frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N \left(I(m, n) - \hat{I}(m, n) \right)^2 \quad (1.6)$$

Où $(M \times N)$ est la taille de l'image, et I_p et \hat{I}_p sont respectivement les amplitudes des pixels sur les images originale et dégradée. Il est vraisemblable que l'œil tienne beaucoup plus compte des erreurs à grandes amplitudes, ce qui favorise la mesure quadratique.

$$RMSE = \sqrt{MSE} \quad (1.7)$$

1.5.1.2 Rapport Crête Signal Sur Bruit (PSNR)

Au lieu de mesurer la distorsion, cette valeur (Peak Signal to Noise Ratio, PSNR) mesure la fidélité, puisqu'elle est proportionnelle à la qualité. Tout de même, elle est une fonction de MSE ; sa définition et son utilisation proviennent du domaine du traitement de signal :

$$PSNR = 10 \log_{10} \left(\frac{I_{max}^2}{MSE} \right) \quad (1.8)$$

Pour une image à niveau de gris, I_{max} désigne la luminance maximale possible. Une valeur de PSNR infini correspond à une image non dégradée. Cette valeur décroît en fonction de la dégradation. Le *PSNR* relie donc le MSE à l'énergie maximale de l'image.

1.5.1.3 Rapport Signal Sur Bruit (SNR)

On utilise parfois une autre variante du rapport signal sur bruit (Signal to Noise Ratio : SNR), qui relie le MSE à l'énergie moyenne de l'image :

$$SNR = 10 \log_{10} \left(\frac{\frac{1}{N} \sum I^2}{MSE} \right) \quad (1.9)$$

1.5.1.4 Indice de Similarité Structurelle (SSIM)

SSIM [21] est une mesure de similarité entre deux images numériques. Elle a été développée pour mesurer la qualité visuelle d'une image déformée, donc d'une vidéo détériorée, par rapport à l'image originale.

L'idée de SSIM est de mesurer la similarité de structure entre les deux images (x, y) , plutôt qu'une différence pixel à pixel comme le fait par exemple le PSNR. L'hypothèse sous-jacente est que l'œil humain est plus sensible aux changements dans la structure de l'image. La Figure 1.9 la structure qui permet de calculer ce paramètre.

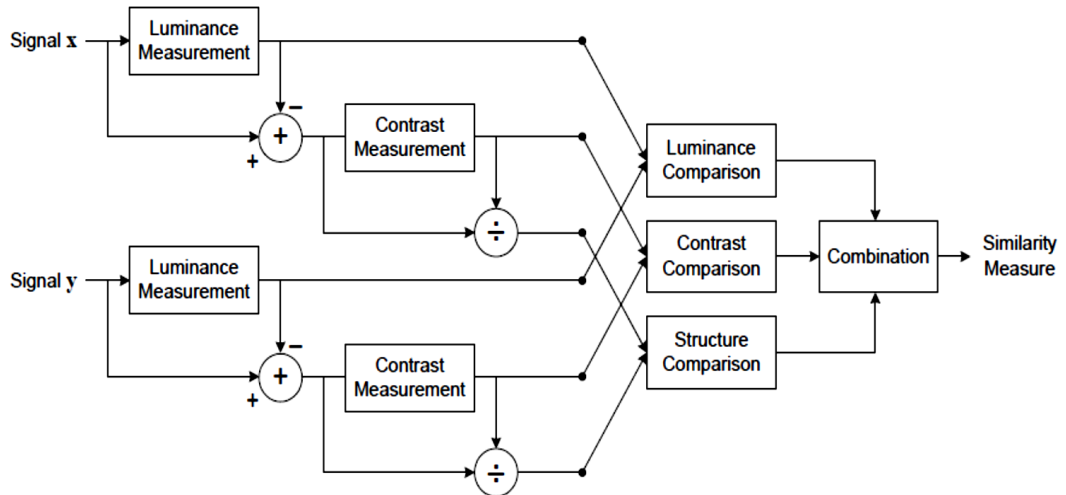


Figure 1.9 Diagramme de l'Indice de Similarité Structurale (SSIM) [22].

La métrique SSIM est calculée sur plusieurs fenêtres d'une image. On dénote x et y l'image originale et l'image déformée respectivement. La similarité compare la luminance, le contraste et la structure entre chaque couple de fenêtres.

La luminance est estimée par la mesure de l'intensité moyenne de chaque fenêtre :

$$\mu_x = \frac{1}{T} \sum_{i=1}^T x_i \quad (1.10)$$

$$\mu_y = \frac{1}{T} \sum_{i=1}^T y_i \quad (1.11)$$

T : le nombre de pixels de chaque fenêtre.

x_i, y_i : L'intensité d'un pixel.

Le contraste pour chaque fenêtre est mesuré par :

$$\left\{ \begin{array}{l} \sigma_x^2 = \frac{1}{T-1} \sum_{i=1}^T (x_i - \bar{x})^2 \\ \sigma_y^2 = \frac{1}{T-1} \sum_{i=1}^T (y_i - \bar{y})^2 \end{array} \right. \quad (1.12)$$

La similarité est déterminée par la perte de corrélation entre les deux fenêtres :

$$\sigma_{xy}^2 = \frac{1}{T-1} \sum_{i=1}^T (x_i - \bar{X})(y_i - \bar{Y}) \quad (1.13)$$

La fonction de comparaison de luminance notée par $l(x, y)$ est une fonction de x et y .

Le contraste entre deux fenêtres noté par $c(x, y)$, comparent les variances de x et y .

La troisième fonction compare les structures des deux fenêtres données par la fonction

$s(x, y)$ En fonction de deux fenêtres normalisées.

Finalement la fonction mesurant la similarité, est une fonction de $l(x, y)$, $c(x, y)$ et de $s(x, y)$. Elle est de la forme :

$$S(x, y) = f(l(x, y), c(x, y), s(x, y)) \quad (1.14)$$

Wang et Al ont proposé dans [16] une fonction de comparaison de la luminance de la forme :

$$l(x, y) = \frac{2\mu_x\mu_y + (K_1L)^2}{\mu_x^2 + \mu_y^2 + (K_1L)^2} \quad (1.15)$$

K_1 Est une constante de très faible valeur. La constante $(K_1L)^2$ évite l'instabilité de la fonction de comparaison lorsque $\mu_x^2 + \mu_y^2$ sont très proches de zéro. Dans le cas des images à niveau de gris, la constante est $L = 255$ (dynamique du capteur).

On remarque que l'équation (1.14) est conforme à la loi de Weber, largement utilisée pour la modélisation de l'adaptation à la lumière dans le SVH (système visuel humain). Selon la loi de weber, la variation de luminance est proportionnelle à la luminance de fond. En d'autres termes, le SVH est sensible aux variations relatives de luminance entre les deux signaux ou encore entre les deux images. La fonction de comparaison de contraste prend une forme similaire :

$$c(x, y) = \frac{2\sigma_x\sigma_y + (K_2L)^2}{\sigma_x^2 + \sigma_y^2 + (K_2L)^2} \quad (1.16)$$

Finalement, la mesure de similarité résultante des trois comparaisons est donnée par une expression simplifiée de l'index de similarité structurelle entre x et y :

$$SSIM(x, y) = l(x, y) \cdot c(x, y) \cdot s(x, y) \quad (1.17)$$

Ou encore sous une forme simplifiée :

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C1)(2\sigma_{xy} + C2)}{(\mu_x^2 + \mu_y^2 + C1)(\sigma_x^2 + \sigma_y^2 + C2)} \quad (1.18)$$

1.6 CONCLUSION

Dans ce chapitre nous avons abordé certaines définitions primaires concernant l'image et la vidéo. Nous avons également présenté le modèle du bruit blanc gaussien qu'on va l'éliminé par les algorithmes de débruitage des séquences vidéo. Finalement, nous avons définis les critères d'évaluation de la qualité objective des images et vidéo (RMSE, PSNR, SSIM). Ces paramètres seront utilisés au cours de nos expériences de simulation, pour évaluer les algorithmes de débruitage considérés.

Chapitre 2

Etat de L'Art sur Les Méthodes d'Estimation du Mouvement

Résumé

Ce deuxième chapitre est consacré à la description des méthodes usuelles d'estimation du mouvement à partir d'une séquence d'images. Après une brève discussion sur le mouvement et les problèmes généraux de sa mesure, nous présenterons les principales techniques d'estimation proposées dans le domaine de la vision par ordinateur. A la fin de ce Chapitre, nous exposerons la méthode différentielle pour l'estimation du mouvement ainsi qu'une conclusion.

Sommaire

2.1 Introduction.....	38
2.2 Notions Fondamentales de L'Estimation du Mouvement	38
2.3 Mouvement dans Les Images.....	39
2.4 Champ Bidimensionnel de Déplacement, Flot optique, Mouvement Apparent 2D	40
2.5 Estimation du Mouvement d'une Image (Equation de Contraintes du Mouvement ECMA) Apparent:	42
2.6 Méthodes d'Estimation du Mouvement	45
2.7 Conclusion	52

2.1 Introduction

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

Par définition, le flux optique est le champ de vitesse (ou de déplacement), produit dans le plan image par des objets en mouvement dans un espace 3D. La plupart des méthodes existantes pour l'estimation du flot optique peuvent être classées en trois catégories : les méthodes différentielles, les méthodes de mise en correspondance et les méthodes fréquentielles basées sur l'énergie et sur la phase [4].

Nous nous intéresserons principalement à l'estimation du mouvement par les méthodes différentielles.

2.2 Notions Fondamentales de L'Estimation du Mouvement

Dans cette section, nous rappellerons les concepts de base de l'estimation du mouvement.

2.2.1 Séquence Numérique

Dans une séquence vidéo numérique, deux trames (images) successives présentent de nombreuses similarités caractérisant le fond et les objets en mouvement. Ces trames sont composées d'un même ensemble d'objets. Une trame est une représentation planaire d'une scène ou d'un objet, généralement situé dans un espace tridimensionnel.

Une image numérique est constituée d'un ensemble de points élémentaires appelés pixels. Le couple de réels (x, y) représente la position spatiale d'un pixel, et la valeur $I(x, y)$ son niveau de gris. Le niveau de gris d'un pixel est une grandeur proportionnelle à l'intensité du signal réfléchi par ce pixel lorsqu'il est radié notamment par une onde électromagnétique. La structure lumineuse d'une image est constituée par un ensemble de motifs brillants [23].

2.2.2 Mouvement Global, Mouvement Local

Un mouvement global est un mouvement tel que tous les pixels de l'image se déplacent d'une manière homogène. Le mouvement local décrit individuellement le mouvement d'un ensemble de points de l'image. En général, ces deux spécifications ne restent qu'une définition, car du point de vue structurel, aucune distinction n'est faite entre mouvement global et local.

Pour mieux comprendre ces définitions, plaçons-nous par exemple dans le contexte de l'analyse d'une image toute entière. Imaginons que nous observons une séquence vidéo en un point arbitraire. Le point est le plus petit élément que nous puissions observer sur un élément d'image. Dans le temps, l'intensité des pixels est susceptible de varier. Trois causes peuvent être la source d'une telle variation

La première est due au mouvement global ou mouvement de la caméra. Même si aucun mouvement n'apparaît dans cette scène, le mouvement de la caméra produit un mouvement global de la scène capturée.

La seconde cause est le mouvement intrinsèque des objets dans la scène. Celle-ci apparaît comme un mouvement local car elle n'affecte pas l'image entière.

Enfin, la troisième raison est la variation de la luminance. Si la condition d'éclaircissement change pendant que la séquence est capturée par la caméra, les intensités du pixel varient.

Nous pouvons donc dire que le mouvement local résume l'aspect de chaque objet de l'image et le mouvement global est affecté à l'image toute entière. Ce qui n'est pas toujours valable car dans un contexte d'analyse d'objet (contenu dans une scène), le mouvement global correspond à la transformation subie par l'objet lui-même (déformations globales de la forme) et le mouvement local correspond à l'information pertinente (déformations résiduelles de la forme) [24]

2.3 Mouvement dans Les Images

Le mouvement est un réel problème en vidéo puisqu'il décrit un contexte en trois dimensions alors que les images sont une projection de scènes 3D dans un plan en 2D.

En général, il est représenté par un vecteur de mouvement qui décrit une transformation d'une image en deux dimensions vers une autre [13]. La Figure 2.1 illustre la projection du mouvement sur un plan 2D.

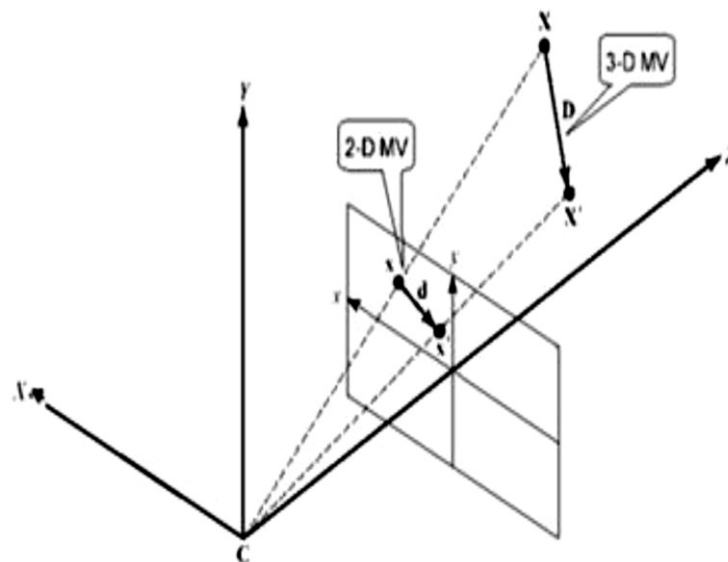


Figure 2.1 Projection mouvement 3D en 2D.

2.3.1 Mouvement réel

Le mouvement 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.

2.3.2 Mouvement Apparent

Le mouvement apparent est un mouvement 2D perçu dans une séquence à travers des variations spatio-temporelles de la luminance. Le champ des vitesses apparentes porte le nom de flot optique.

Le mouvement apparent, appelé flot optique, est une variable cachée et n'est accessible que par l'analyse des variations temporelles de la séquence d'images. Par définition, le flot optique est le champ de vitesse (ou de déplacement) produit dans le plan image par des objets en mouvement dans un espace 3D.

Il est représenté par des vecteurs comme montré dans la Figure 2.2. Chaque vecteur représente le déplacement d'un point de l'image. Le terme de flot optique a été inventé par le psychologue James Jérôme Gibson dans une étude sur la vision humaine [25].

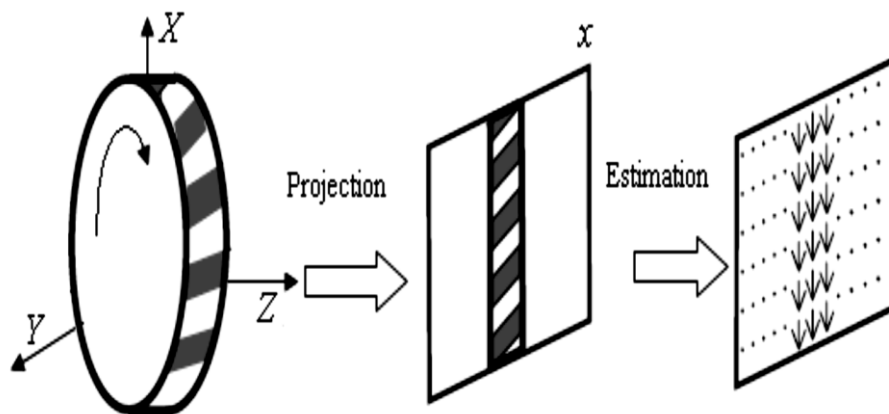


Figure 2.2 Exemple de flot optique [23].

2.4 Champ Bidimensionnel de Déplacement, Flot Optique, Mouvement Apparent 2D

Un mouvement 3D correspond au déplacement entre deux instants d'un objet dans un espace 3D. On appelle champ de déplacement, l'ensemble des vecteurs associés à chaque point de l'objet, décrivant le mouvement des pixels entre deux instants successifs. Comme le montre la Figure 2.3, le champ bidimensionnel de déplacement est obtenu par la projection sur un plan image des déplacements relatifs entre l'objet et l'observateur.

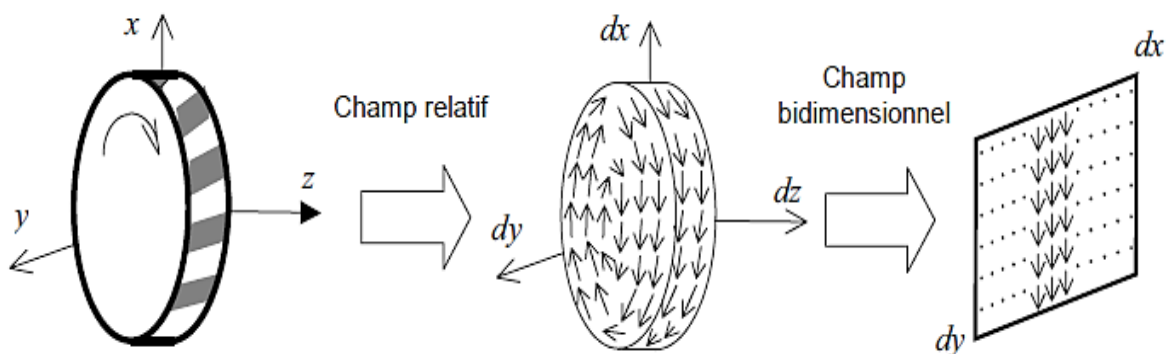


Figure 2.3 Illustration du champ de déplacement bidimensionnel [23].

Dans une séquence vidéo numérique, une scène 3D est ramenée à une image issue d'une projection sur un repère 2D. Dès lors, le champ de déplacement résultant de la variation spatio-temporelle de la luminance d'une image 2D n'est qu'apparent grâce aux effets qualifiés d'illusion d'optique. Reprenons l'exemple d'une roue qui tourne autour de son axe (Figure 2.4). Sur le plan image x - y , son mouvement réel de rotation est perçu comme un mouvement de translation des bandes ornant cette roue. Il s'agit d'un mouvement apparent (perçu) 2D qui s'appelle aussi flot optique.

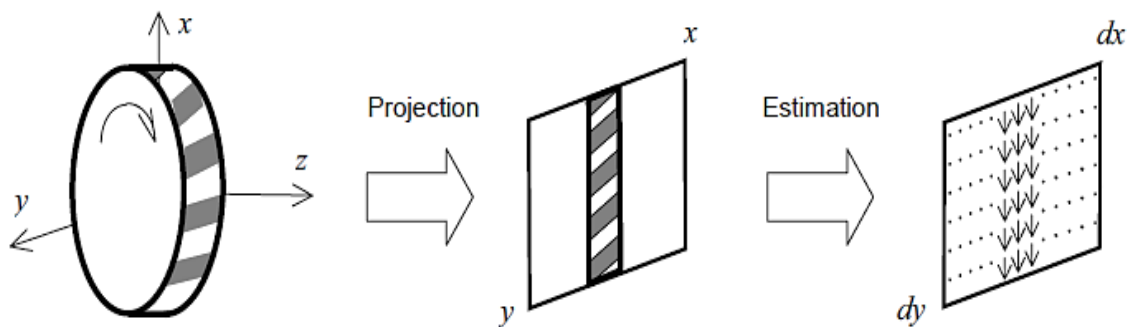


Figure 2.4 Illustration du flot optique (mouvement apparent 2D).

Le flot optique (mouvement apparent 2D) et le champ bidimensionnel de déplacement sont deux aspects différents, mais dans le domaine du traitement numérique des images, ils ne sont pas distingués [23].

Nous signalons ici que notre étude concerne le mouvement apparent 2D et l'expression du mouvement qui devrait être interprété comme flot optique.

Flot Optique [26]

Etant donnée une séquence d'images ordonnée chronologiquement (par exemple issue d'une vidéo), et représentant une même scène, le flot optique désigne le mouvement apparent des objets de la scène dans les images. Dans de nombreux cas, ce mouvement apparent se confond avec la projection du mouvement 3D réel, mais ce n'est pas le cas dans les scènes où l'illumination varie de manière sensible. La Figure 2.5 illustre un exemple de flot optique sur une séquence particulièrement utilisée pour tester les algorithmes par le passé, la séquence Yosemite.



Figure 2.5 La séquence de test Yosemite À gauche et au milieu : une paire d'images extraites de la séquence. À droite : le flot optique estimé par la méthode de B.Horn and B.Schunck [26].

2.5 Estimation du Mouvement d'une Image (Equation de Contraintes du Mouvement Apparent)

L'estimation de mouvement est un procédé qui consiste à étudier le déplacement des objets dans une séquence d'images en cherchant la corrélation entre deux images successives afin de prédire le changement de position du contenu [4].

L'estimation du mouvement par flot optique repose sur une hypothèse de conservation de propriétés photométriques au cours de la séquence d'images. Habituellement, on considère qu'il y a conservation de l'intensité lumineuse, c'est à dire que l'on considère que les pixels sont translatés d'une image à l'autre en conservant leur intensité lumineuse.

$$I(x, y, t) = I(x + dx, y + dy, t + dt) \quad (2.1)$$

Où :

$I(x, y, t)$ est l'image (frame) à l'emplacement (x, y) à l'instant t , dx, dy sont les déplacements spatiaux à l'instant successif $t + dt$. La Figure 2.6 illustre ce déplacement.

Par développement de séries de Taylor :

$$I(x, y, t) = I(x, y, t) + \frac{\partial I}{\partial x} dx + \frac{\partial I}{\partial y} dy + \frac{\partial I}{\partial t} dt \quad (2.2)$$

Nous avons alors :

$$\frac{\partial I}{\partial x} dx + \frac{\partial I}{\partial y} dy + \frac{\partial I}{\partial t} dt = 0 \quad (2.3)$$

Avec un réarrangement de l'équation (2.3) :

$$\frac{\partial I}{\partial x} \frac{dx}{dt} + \frac{\partial I}{\partial y} \frac{dy}{dt} + \frac{\partial I}{\partial t} = 0 \quad (2.4)$$

$$I_x u + I_y v + I_t = 0 \quad (2.5)$$

Avec :

$$I_x = \frac{\partial I}{\partial x}, I_y = \frac{\partial I}{\partial y}, I_t = \frac{\partial I}{\partial t}, u = \frac{dx}{dt}, v = \frac{dy}{dt}$$

Le flot optique vérifie alors l'équation suivante :

$$\nabla I(\vec{x}, t) \cdot \vec{u} + I_t(\vec{x}, t) = 0 \quad (2.6)$$

Avec ∇I le gradient spatial de l'image I , I_t le gradient temporel de l'image I et \vec{u} le vecteur vitesse au pixel \vec{x} [27].

Cette équation est connue sous le nom de l'équation de contrainte du mouvement apparent ECMA. Cette équation seule ne permettant pas de retrouver les deux composantes du champ de vitesse (problème ouvert), nous devons rajouter des contraintes au problème.

Ce problème peut être résolu par des approches Variationnelles où l'on rajoute un terme de régularisation.

Une des approche est la contrainte du champ de vecteurs le plus régulier possible, connue sous l'algorithme de Horn et Schunck [28] ou bien alors une contrainte de régularisation locale du flot optique tel qu'il est cas de l'algorithme de Lucas et Kanade) [27].

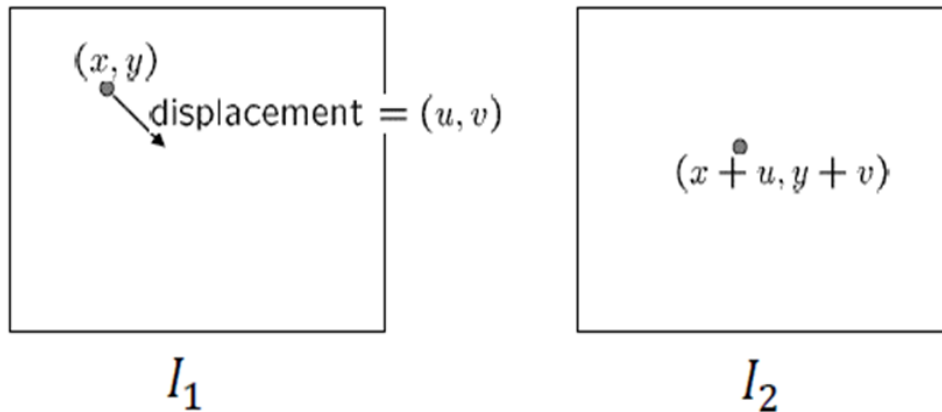


Figure 2.6 Déplacement d'un point dans deux images successives dans le temps $I, I + 1$

En posant l'expression de la vitesse du mouvement $\vec{V} = (u ; v)$ et $\vec{\nabla}I(x, y)$ le gradient de l'image au point $(x; y)$, avec les dérivées partielles de l'image $(I_x; I_y; I_t)$:

$$(I_x u + I_y v) = -I_t$$

Finalement

$$I(x, y) \cdot \vec{V} + I_t = 0 \tag{2.7}$$

L'équation ci-dessus (2.7) est appelée **équation de contrainte du flux optique**.

La Figure 2.7 montre le concept de la conservation d'intensité d'un point sur une trajectoire.

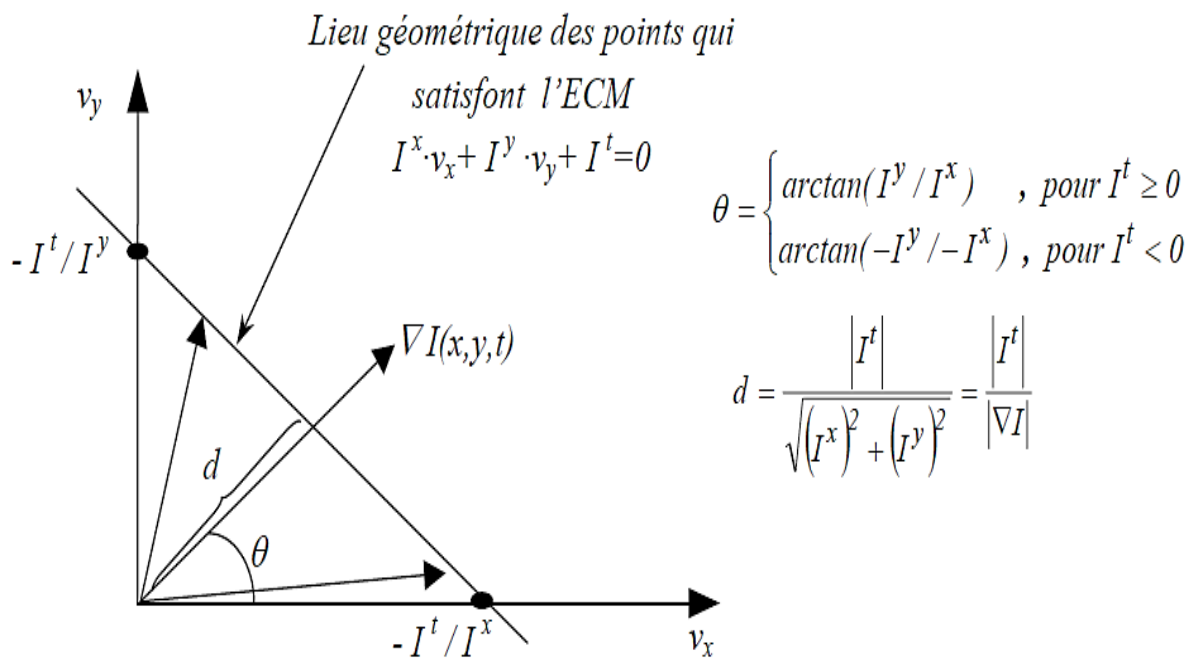


Figure 2.7 Illustration géométrique de l'équation et de la droite de contrainte du mouvement.

En estimation du mouvement en se basant sur l'ECMA, on rencontre deux problèmes majeurs : (a) problème d'ouverture et (b) problème d'occlusion.

- a) **Problème d'ouverture** : le problème d'ouverture est du fait qu'il y a une infinité de valeurs qui satisfont l'équation 2.7, comme c'est illustré dans la Figure 2.8 [29].

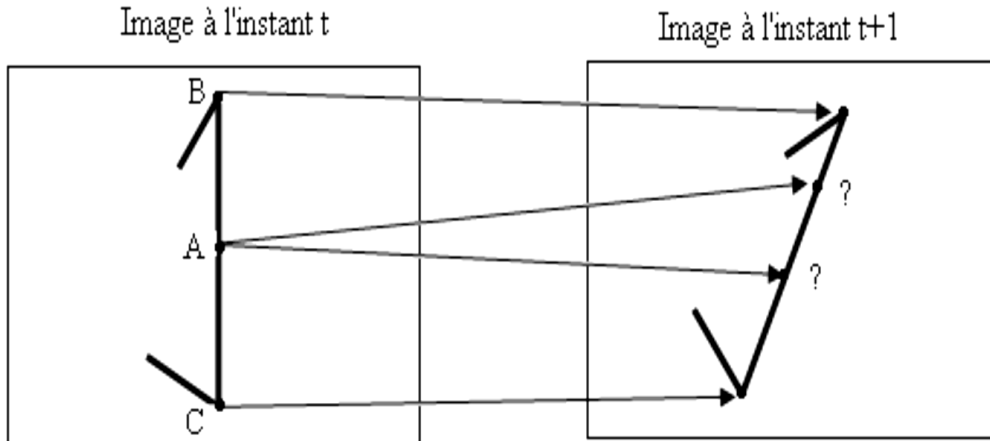


Figure 2.8 Problème de l'ouverture.

b) **Phénomène d'occlusion** : le phénomène d'occlusion dans une séquence vidéo est dû à l'apparition et à la disparition ou aux croisements d'objets dans la vidéo. C'est le résultat du mouvement de la caméra et des objets dans la scène. Cependant, la frontière d'occlusion extérieure d'un objet est en réalité une zone de transition entre cet objet et d'autres objets, plus éloignés de la caméra et partiellement cachés.

Ces objets peuvent effectuer des mouvements différents et il est difficile de savoir à quel objet correspond le mouvement estimé [30]. En effet, les techniques d'estimation de mouvement entre deux images prises à des instants t et $1 + t$ d'une séquence vidéo sont basées sur les ressemblances existantes entre ces deux images. Si des zones apparaissent entre les deux images, les ressemblances diminuent et l'estimation échoue dans ces zones. La Figure 2.9 donne un exemple illustratif du principe d'occlusion.

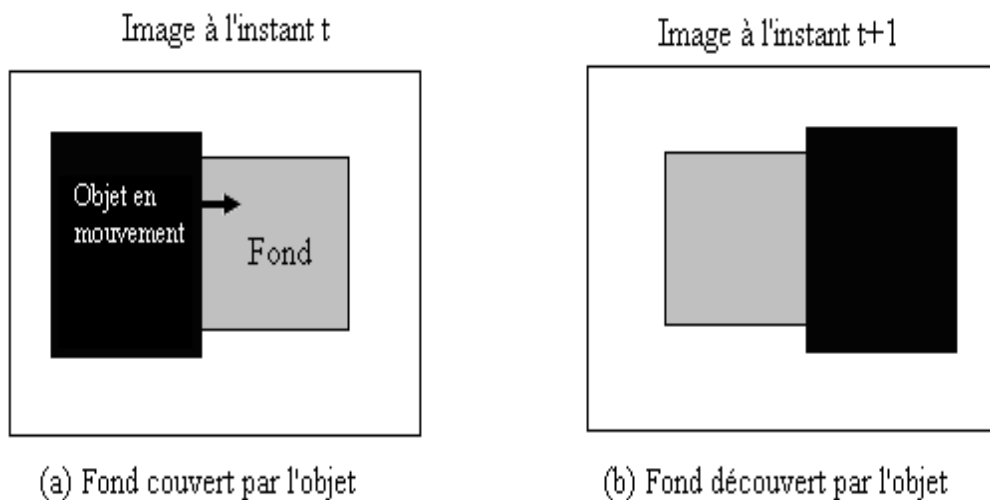


Figure 2.9 Principe d'occlusion.

2.6 Méthodes d'Estimation du Mouvement

Les méthodes d'estimation du mouvement peuvent être regroupées en trois catégories :

- méthodes différentielles ;
- méthodes de mise en correspondance ;
- méthodes fréquentielles.

Dans ce qui suit nous détaillerons ces trois catégories.

2.6.1 Méthodes Différentielles

Les méthodes différentielles calculent la vitesse à partir de dérivées spatiales et temporelles d'intensité lumineuse des pixels. Ce sont des méthodes basées sur une hypothèse forte de conservation de l'intensité des pixels au cours du mouvement dans la séquence d'images.

Parmi les méthodes différentielles existantes, on peut citer les algorithmes de Horn et Schunck [28], de Lucas et Kanade, de Nagel [31], d'Uras et al, le principe de ces algorithmes est le même quelle que soit la technique employée, ils sont basés toujours sur l'équation générale du flux optique (Eq 2.6) pour estimer les mouvements dans les images.

-Estimation de mouvement par la méthode Variationnelle

Les méthodes Variationnelles consistent à résoudre un problème d'optimisation (local ou global) en minimisant une fonctionnelle, généralement basée sur l'équation du flot optique à laquelle on ajoute une contrainte pour particulariser les solutions. Selon le type de contrainte utilisée, on obtient différentes approches Variationnelles :

✓ Méthodes Variationnelles globales

Ce type d'approches consiste à minimiser sur le domaine entier de l'image une fonctionnelle prenant en compte l'équation du flot optique ainsi qu'un terme de lissage, c'est à dire en ajoutant une contrainte de régularisation portant sur le gradient, le Laplacien (ou ordre supérieur) du champ de vitesse. La plus connue de ces méthodes est certainement celle proposée par Horn & Schunck [32] qui impose une contrainte globale de lissage sur le champ des vitesses.

✓ Méthodes Variationnelles locales

Les méthodes locales consistent à prendre en compte des hypothèses supplémentaires sur un domaine de taille réduite pour particulariser le flot optique. On minimise alors un critère sur un petit domaine, et on obtient ainsi le flot optique de ce petit domaine. La méthode locale la plus célèbre est celle de Lucas & Kanade [33].

✓ Méthodes Hybrides

La littérature récente présente des méthodes dites "locales globales" qui ajoutent un filtrage temporel aux méthodes précédentes, donc le lissage et le filtrage temporel peuvent être considérés comme étant effectués sur toute l'image, mais la résolution de l'algorithme reste purement locale [34-35].

Dans notre travail de master, nous nous sommes basé particulièrement sur deux algorithmes : l'algorithme de Horn et Schunck [28] et l'algorithme de Lucas et Kanade [31].

2.6.1.1 Algorithme de Horn et Schunck [28]

L'algorithme de Horn et Schunck fait partie des méthodes différentielles globales qui calculent un champ de vecteurs denses sur le domaine entier de l'image en minimisant une fonctionnelle prenant en compte l'équation (2.7) du flot optique ainsi qu'une régularisation globale, c'est-à-dire en ajoutant une contrainte de régularisation portant sur le gradient du champ de vecteurs.

La régularisation est exprimée comme suit :

$$\iint \left\{ (I_x \mathbf{u} + I_y \mathbf{v} + I_t)^2 + \lambda (u_x^2 + u_y^2 + v_x^2 + v_y^2) \right\} dx dy \quad (2.10)$$

Terme d'intensité constante (Brighness consistancy)

Terme de contrainte (smoothness constraint)

Tel que: $u_x = \frac{u}{dx}$, $u_y = \frac{u}{dy}$ avec: $u = \frac{dx}{dt}$,

$v_x = \frac{v}{dx}$, $v_y = \frac{v}{dy}$ avec: $v = \frac{dy}{dt}$,

Le problème de minimisation sera alors exprimé comme suit:

$$\begin{cases} (I_x \mathbf{u} + I_y \mathbf{v} + I_t) I_x + \lambda (\Delta^2 u) = 0 \\ (I_x \mathbf{u} + I_y \mathbf{v} + I_t) I_y + \lambda (\Delta^2 v) = 0 \end{cases} \quad (2.11)$$

On aura donc à approximer et discrétiser les termes: $\Delta^2 u$ et $\Delta^2 v$, cela est possible en appliquant les masques de Roberts.

Dans un souci de clarification, nous rappelons dans ce qui suit ces différents masques.

Parenthèse sur les masques de Roberts

Nous considérons le cas d'estimation du mouvement entre deux images successives dans le temps: première image et deuxième image. La différentiation (gradient ou laplacien) se fait entre ces deux paires d'images. Nous indiquons les masques pour calculer les deux composantes I_x et I_y ainsi que la composante de différentiation temporelle I_t .

$$\begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix} \text{ première image (first image)}$$

$$\begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix} \text{ deuxième image (second image)}$$

I_x

$$\begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} \text{ first image}$$

$$\begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} \text{ second image}$$

I_y

$$\begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix} \text{ first image}$$

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \text{ second image}$$

I_t

Approximation du Laplacien :

$$\begin{bmatrix} 0 & -\frac{1}{4} & 0 \\ -\frac{1}{4} & 1 & -\frac{1}{4} \\ 0 & -\frac{1}{4} & 0 \end{bmatrix}$$

$I_{xx} + I_{yy}$

$I_{xx} + I = I - I_{av}$ Version discrète du laplacien (discrete version of Laplacian), avec I_{av} la valeur moyenne de I

$$\begin{cases} (I_x u + I_y v + I_t) I_x + \lambda(u - u_{av}) = 0 \\ (I_x u + I_y v + I_t) I_y + \lambda(v - v_{av}) = 0 \end{cases} \quad (2.12)$$

$$\Delta^2 u = \frac{\partial u}{\partial x^2} + \frac{\partial u}{\partial y^2} = u_{xx} + u_{yy} \quad (2.13)$$

$$\Delta^2 v = \frac{\partial v}{\partial x^2} + \frac{\partial v}{\partial y^2} = v_{xx} + v_{yy} \quad (2.14)$$

On va donc résoudre ce système linéaire à deux équations:

$$\begin{cases} u = u_{av} - I_x \frac{P}{D} \\ v = v_{av} - I_y \frac{P}{D} \end{cases} \quad (2.15)$$

Avec :

$$\begin{cases} P = I_x u_{av} + I_y v_{av} + I_t \\ D = \lambda + I_x^2 + I_y^2 \end{cases} \quad (2.16)$$

Ces étapes sont résumées dans l'algorithme de Horn et Schunck comme suit :

Algorithm-1: Horn et Schunck

- $k = 0$,

- Initialisation : $u^k v^k$
- Répéter jusqu'à ce qu'une mesure d'erreur soit satisfaite (converge)

$$\begin{cases} u = u_{av} - I_x \frac{P}{D} \\ v = v_{av} - I_y \frac{P}{D} \end{cases}$$

$$\begin{cases} P = I_x u_{av} + I_y v_{av} + I_t \\ D = \lambda + I_x^2 + I_y^2 \end{cases}$$

2.6.1.2 Algorithme de Lucas et Kanade [31]

Cet algorithme fait parties des méthodes locales. Contrairement aux méthodes différentielles globales les méthodes différentielles locales calculent localement le flot optique et estiment un champ de vecteurs sur un petit voisinage en appliquant une contrainte locale, On minimise alors un critère sur un petit voisinage, et on obtient ainsi le flot optique de ce petit voisinage.

Lucas et Kanade en 1981 ont été les premiers à proposer une telle méthode. Les auteurs ont proposé de contraindre l'équation du flot optique sur un petit voisinage du point considéré. Dans cet algorithme, le problème d'estimation du mouvement revient à résoudre un système d'équations linéaires. Puisque l'algorithme de Lucas et Kanade est local, ils considèrent un voisinage d'une fenêtre de 9 éléments, nous aurons alors le système d'équations suivant :

$$\begin{aligned} I_{x1}u + I_{y1}v &= -I_{t1} \\ &\vdots && \vdots & \vdots \\ I_{x9}u + I_{y9}v &= -I_{t9} \end{aligned}$$

$$\begin{bmatrix} I_{x1} & I_{y1} \\ \vdots & \vdots \\ I_{x9} & I_{y9} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} -I_{t1} \\ \vdots \\ -I_{t9} \end{bmatrix} \quad (2.17)$$

On note $\mathbf{A} \mathbf{u} = \mathbf{I}_t$

Tel que : $\mathbf{A} = \begin{bmatrix} I_{x1} & I_{y1} \\ \vdots & \vdots \\ I_{x9} & I_{y9} \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} u \\ v \end{bmatrix}$ et $\mathbf{I}_t = \begin{bmatrix} -I_{t1} \\ \vdots \\ -I_{t9} \end{bmatrix}$

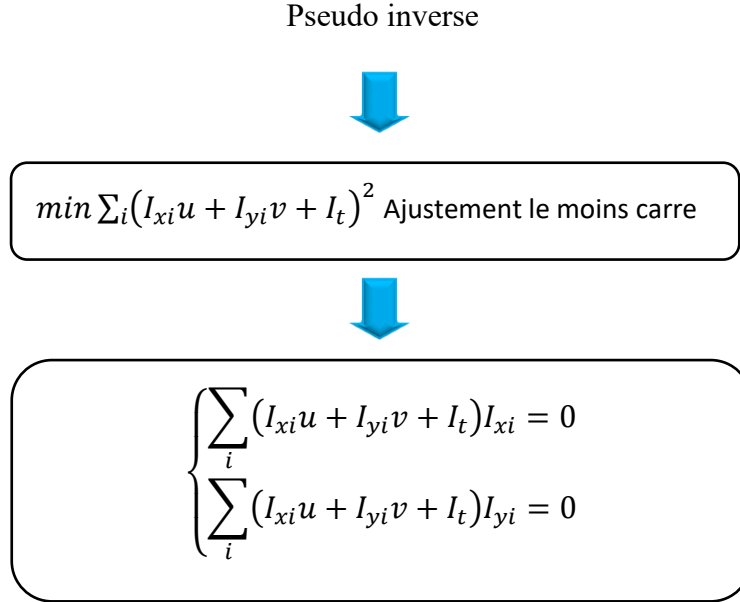
On aura :

$$A^T A u = A^T I_t \quad (2.18)$$

La solution est alors :

$$u = (A^T A)^{-1} A^T I_t \quad (2.19)$$

On résume les étapes de résolution, dans cet organigramme :



Luca and Kanade:

$$\begin{cases} \sum_i I_{xi}^2 u + \sum_i I_{xi}I_{yi}v = -\sum_i I_{xi}I_t \\ \sum_i I_{xi}I_{yi}u + \sum_i I_{yi}^2 v = -\sum_i I_{yi}I_t \end{cases} \quad (2.20)$$

$$\begin{bmatrix} \sum_i I_{xi}^2 & \sum_i I_{xi}I_{yi} \\ \sum_i I_{xi}I_{yi} & \sum_i I_{yi}^2 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} -\sum_i I_{xi}I_t \\ -\sum_i I_{yi}I_t \end{bmatrix} \quad (2.21)$$

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \sum_i I_{xi}^2 & \sum_i I_{xi}I_{yi} \\ \sum_i I_{xi}I_{yi} & \sum_i I_{yi}^2 \end{bmatrix}^{-1} \begin{bmatrix} -\sum_i I_{xi}I_t \\ -\sum_i I_{yi}I_t \end{bmatrix} \quad (2.22)$$

$$\begin{bmatrix} u \\ v \end{bmatrix} = \frac{1}{\sum_i I_{xi}^2 \sum_i I_{yi}^2 - (\sum_i I_{xi}I_{yi})^2} \begin{bmatrix} \sum_i I_{yi}^2 & -\sum_i I_{xi}I_{yi} \\ -\sum_i I_{xi}I_{yi} & \sum_i I_{xi}^2 \end{bmatrix} \begin{bmatrix} -\sum_i I_{xi}I_t \\ -\sum_i I_{yi}I_t \end{bmatrix} \quad (2.23)$$

Parce que :

$$A^{-1} = \frac{1}{\det A} t_{comA}$$

$$u = \frac{1}{\sum_i I_{xi}^2 \sum_i I_{yi}^2 - (\sum_i I_{xi}I_{yi})^2} (-\sum_i I_{yi}^2 \sum_i I_{xi}I_t + \sum_i I_{xi}^2 \sum_i I_{yi}I_t) \quad (2.24)$$

$$v = \frac{1}{\sum_i I_{xi}^2 \sum_i I_{yi}^2 - (\sum_i I_{xi} I_{yi})^2} (\sum_i I_{yi}^2 \sum_i I_{xi} I_t - \sum_i I_{xi}^2 \sum_i I_{yi} I_t) \quad (2.25)$$

Les deux algorithmes détaillés présentent certains inconvénients

- Les méthodes optiques Horn-Schunck et Lucas-Kanade ne fonctionnent que pour les petits mouvements.
- Si l'objet se déplace plus rapidement, la luminosité change rapidement, les masques 2x2 ou 3x3 ne permettent pas d'estimer les dérivées spatio-temporelles
- Les pyramides peuvent être utilisées pour calculer un grand vecteur flux optique (la pyramide réduit le grand mouvement)

C'est pourquoi :

- Le grand mouvement peut être réduit par les pyramides. En d'autres termes, la notion de pyramide est très utile pour estimer le mouvement correct d'un flux optique correct.

Estimation en multi-résolution

Chaque méthode mise en œuvre pour estimer le flot optique peut être améliorée en appliquant deux principes qui expriment la même idée à deux échelles différentes :

- ❖ **Le raffinement itératif** : Cela consiste à raffiner le champ de vitesse estimé en exécutant itérativement l'algorithme après avoir déplacé une des deux images selon le dernier champ de vitesse calculé.
- ❖ **L'implémentation pyramidale** : On définit une hauteur L_m de pyramide (en pratique, $L_m = 2, 3$ ou 4 niveaux de décomposition). A chaque niveau de la pyramide, l'image est sous-échantillonnée d'un facteur 2 pour les deux images successives considérées. Le niveau zéro correspond à l'image initiale. Le niveau L_m correspond au niveau le plus grossier. L'estimation du flot optique commence à l'échelle grossière ensuite ces mesures approximatives sont propagées et raffinées progressivement à des échelles plus fines jusqu'au niveau 0 qui correspond à l'image initiale. On récupère alors le flot optique final (voir Figure 2.10) [36].

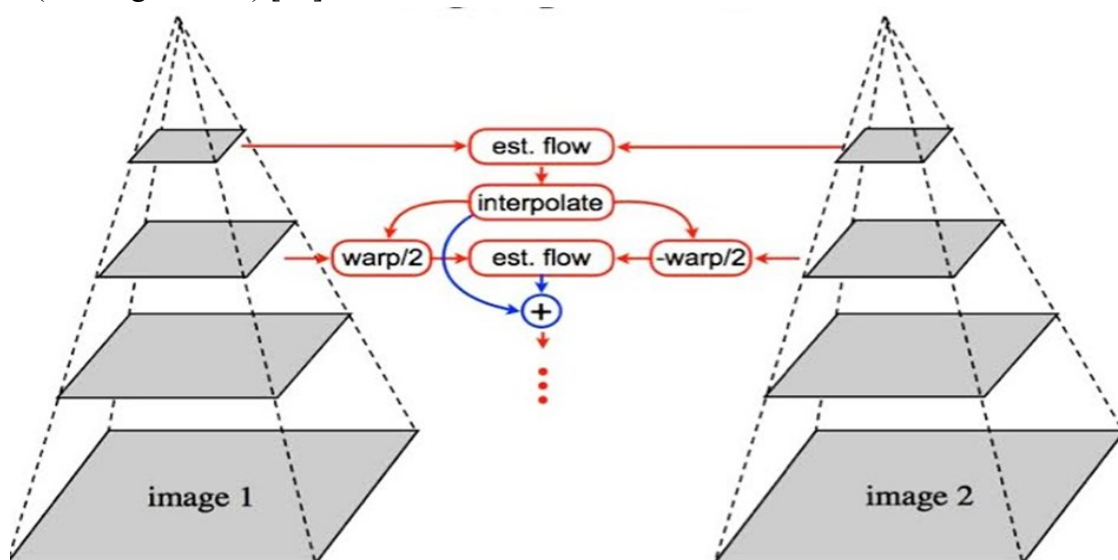


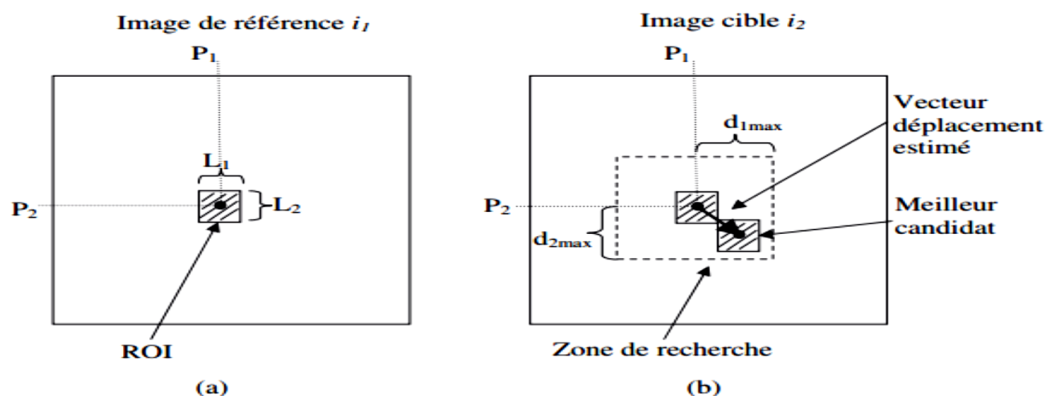
Figure 2.10 Implémentation pyramidale d'une méthode de calcul du flot optique.

2.6.2 Méthode de Mise en Correspondance de Blocs

Lorsque les déplacements considérés dans les images sont trop grands (large motion), typiquement plus que deux pixels par image, ou que le calcul des dérivées spatio-temporelles est sujet à des erreurs numériques, il convient de se tourner vers des méthodes dites de mise en correspondance de blocs ou appariement de blocs ou bloc matching. La Figure 2.11 illustre ce principe.

Ces méthodes consistent à appairer (matching) chaque région R_i (ou bloc) d'une image au temps t à une autre région R'_i au temps $t + 1$ et ce en trouvant le déplacement $d = (dx; dy)$ qui minimise un fonctionnel d'erreur et maximise la similarité entre les deux régions [37].

De nombreux algorithmes d'exploration ont été développés [38-39], le plus simple étant l'exploration de tous les blocs de la fenêtre de recherche ("Full Search").



(a) région d'intérêt (Region of interest ROI) de taille $L_1 \times L_2$ considérée dans l'image de référence.
 (b) zone de recherche considérée dans l'image cible.

Figure 2.11 Principe des méthodes de mise en correspondance de blocs.

La ROI est recherchée dans cette zone de recherche est le meilleur candidat (le plus ressemblant à la ROI en terme d'un critère donné) est retenu. Il est à noter que la taille de la zone de recherche impose le déplacement maximum autorisé pour la ROI courante.

2.6.3 Méthodes Fréquentielles

Les méthodes fréquentielles d'estimation de vitesse dans les séquences d'images sont fondées sur une caractérisation du mouvement dans le domaine des fréquences. Elles ont pour origine des recherches concernant la vision des mammifères, et la mise en évidence de la présence de cellules simples dans l'aire corticale V_1 , qui se comportent comme des filtres passe-bande spatio-temporels. Ces méthodes fournissent des flux optiques de qualité, mais sont réputées gourmandes en calculs.

Dans les approches fréquentielles, les fréquences spatio-temporelles sont mises en relation avec la vitesse du stimulus de mouvement, et le flot optique devient l'identification d'un plan d'énergie dans l'espace de la fréquence spatio-temporelle. En général, les mécanismes sensibles au mouvement fondés sur l'énergie orientée spatio-temporelle dans l'espace de Fourier peuvent estimer le mouvement dans les endroits où les autres approches échouent.

Ainsi, la détection du mouvement dans l'image revient ici à extraire l'orientation spatio-temporelle [40]. On distingue deux sous-approches, Les méthodes fréquentielles basées sur l'énergie et les méthodes exploitant la phase du signal.

2.7 Conclusion

Nous avons rappelé dans ce chapitre la théorie de l'estimation du mouvement. Nous avons notamment présenté les approches basées sur l'estimation du flot optique. Plusieurs techniques d'estimation du flot optique ont donc été présentées. On peut observer que certaines techniques opèrent dans le domaine spatio-temporel, d'autres dans le domaine fréquentiel.

Dans le domaine spatio-temporel, nous avons rappelé l'équation de contrainte spatiotemporelle du flot optique qui relie les composantes du vecteur vitesse aux gradients spatiotemporels.

Les approches qui découlent de cette équation s'appellent méthodes différentielles. Ces dernières ont été développées davantage dans ce deuxième chapitre. Précisément, nous avons détaillé deux algorithmes : Horn and Schrunck et Lucas and Kanade. Nous ne retiendrons que cette technique différentielle par la suite dans notre travail, en raison de sa facilité de mise en œuvre et parce que l'on peut facilement modéliser le déplacement. En outre, comme nous envisageons une application qui nécessite un nombre d'images restreint (deux images), cela justifie également le choix de la technique différentielle car elle peut s'appuyer uniquement sur deux trames consécutives.

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

Chapitre 3

Algorithmes de Débruitage Pour Les Séquences Vidéos

Résumé

Dans ce chapitre nous étudierons les algorithmes récents de débruitage des séquences vidéo. Plus précisément, les algorithmes VBM3D, VBM4D, SPTWO seront abondamment détaillés.

Nous présenterons l'algorithme Hybrid de débruitage que nous avons proposé.

Sommaire

3.1 Introduction	54
3.2 Estimation du Bruit dans les Séquences Vidéo	54
3.3 Types de Débruitage des Séquences Vidéo	56
3.4 Algorithmes de Débruitage des Séquences Vidéo.....	61
3.5 Algorithme Proposé: L'Algorithme de Débruitage Hybrid	67
3.6 Critères d'Evaluation (de performance)	69
3.7 Conclusion	69

3.1 Introduction

Le choix des algorithmes efficaces pour le débruitage d'images numériques et animées reste encore un défi scientifique à l'intersection de l'analyse fonctionnelle, des statistiques et de l'informatique. Il y a eu plusieurs algorithmes [2, 41]. Publiés et chaque approche à ses suppositions, ses avantages, et ses limitations [42]. Le but d'une procédure de débruitage est défini comme le fait de reproduire une image \hat{I} qui pour chaque pixel x donne un estimateur $\hat{I}(x)$ de l'intensité de l'image originale $I(x)$. C'est un problème inverse généralement mal posé.

Le débruitage d'images a pour objectif principal d'atténuer, voire de supprimer dans les cas les plus favorables, les dégradations que subie une image (séquences d'image). Le plus souvent, les dégradations considérées sont celles qui apparaissent au moment de l'acquisition de l'image. On pourrait y ajouter toutes les dégradations liées à la transmission, au changement de format ou au stockage de l'image.

Le débruitage d'images essaye donc de recréer, à partir de l'image dont on dispose, l'image telle que l'on devrait l'avoir en situation idéale. Ce qui est visé, c'est la fidélité de l'image obtenue à la scène qu'elle représente [43].

Le besoin pour la suppression du bruit sans dégrader significativement les contours et autres composants de hautes fréquences de l'image [17], a donc motivé le développement des techniques efficaces qui lissent le bruit et conservent les contours. Dans le chapitre suivant, nous allons décrire les méthodes de débruitage des vidéos à base d'approches variationnelles.

Il convient avant de détailler ces algorithmes, d'effectuer l'estimation du bruit qui contamine les séquences vidéo.

3.2 Estimation du Bruit dans les Séquences Vidéo

De nombreuses applications en traitement d'images comme le débruitage ou la segmentation nécessitent de connaître le niveau de bruit, pourtant son estimation reste un problème délicat.

Lorsque le bruit est additif gaussien, un estimateur populaire de la variance du bruit repose sur l'écart type moyen absolu des coefficients d'ondelettes les plus fins, d'après l'algorithme de Donoho [44].

Plus récemment, Liu et *al.* [45] ont proposé d'estimer la variance du bruit à l'aide d'une analyse en composantes principales (Principal Component Analysis PCA) sur des patches de l'image. Cependant, ces techniques sont efficaces seulement pour un niveau de bruit constant.

Or en pratique, le bruit dépend du signal et son niveau varie spatialement en fonction de l'intensité. Un modèle de bruit dépendant du signal peut être adopté et les paramètres du modèle sont alors estimés avec une méthode dédiée, par exemple dans le domaine des ondelettes pour un modèle Poisson-gaussien [46] ou via une modélisation de patches par des mélanges de Gaussiennes (mixtures of gaussian) pour du bruit additif à variance affine [47]. Les statistiques du bruit peuvent donc être estimées en détectant les régions homogènes puis en évaluant les paramètres du bruit sur ces zones. Cependant, ces détecteurs nécessitent également des hypothèses fortes sur la nature du bruit (additif par exemple), ce qui les rend peu efficaces lorsque la nature du bruit est inconnue [48].

Dans notre travail de simulation, pour tester et confirmer la nature du bruit présent dans les séquences vidéos, (et du moment que nous disposons des séquences vidéo originales, cela nous a permis d'effectuer ce test), nous estimons le type de bruit à travers la différence entre l'image originale et l'image bruitée et nous calculons et traçons la fonction de répartition CDF (Cumulative Density Function) du résultat de différence. Le teste consiste à superposer le résultat des données de simulation avec la CDF Gaussienne théorique. La Figure 3.1 illustre ce teste. Nous pouvons facilement remarqué que la CDF théorique est très proche de celle estimée. Nous avons donc considéré que le bruit présent dans les séquences vidéo de nos expériences est un bruit Gaussien additif. Cette supposition sera respectée tout au long de nos expériences. Rappelons pour mémoire que la CDF est utilisée pour déterminer la probabilité qu'une observation aléatoire extraite de la population soit inférieure ou égale à une certaine valeur. Nous pouvons également utiliser cette fonction pour déterminer la probabilité qu'une observation soit supérieure à une certaine valeur, ou comprise entre deux valeurs [49].

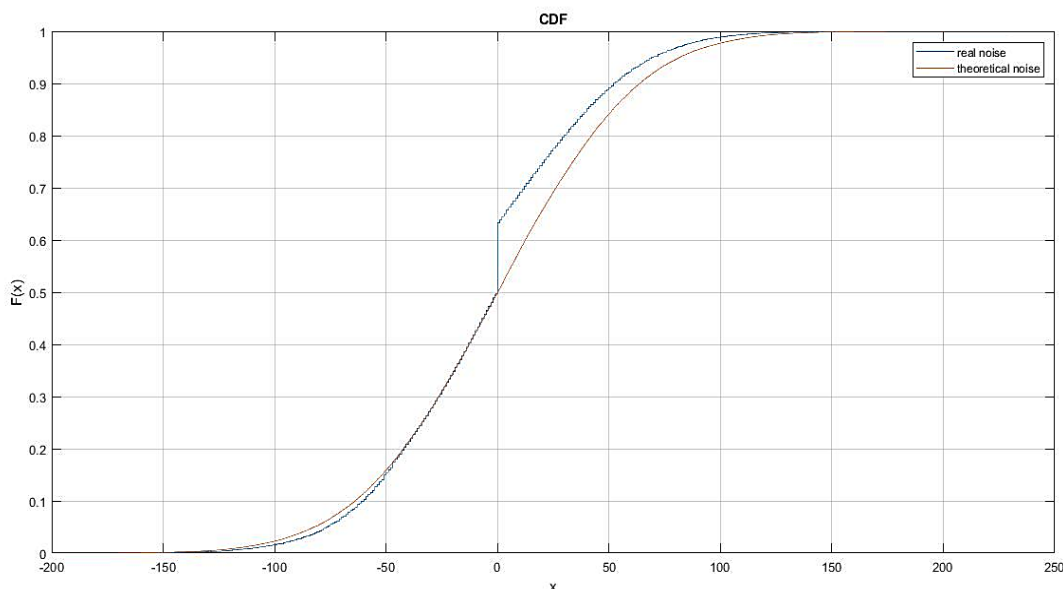


Figure 3.1 Estimation du bruit.

3.2.1 Régularisation des Séquences d'Images

Dans cette section, nous présentons le problème de débruitage des vidéos. La première idée qui viendrait à l'esprit pour débruiter une séquence d'images serait de traiter toutes les images formant la séquence, une par une, en utilisant des méthodes conventionnelles abondamment utilisées pour des images statiques, celles que nous avons présentées dans le Chapitre 2.

Cependant, ces techniques de débruitage (destinées aux images statiques) ne tiennent pas compte des dégradations spécifiques à la vidéo telles que le bougé, la résolution ou encore le zoom. Par ailleurs, elles contraignent le traitement à un espace 2D alors que nous disposons d'une dimension supplémentaire, le temps. Par conséquent, il est plus avantageux de concevoir un traitement $2D + t$ directement et de bénéficier des avantages de la redondance temporelle.

Puisque nous avons déjà cité les dégradations les plus répandues et leurs sources dans le Chapitre 1, nous passons en revue les différentes méthodes de débruitage des vidéos que nous avons considéré dans notre travail de Master.

Dans l'optique, d'améliorer l'algorithme de débruitage proposé dans [2], dans nos expériences, nous avons d'abord appliqué les filtres classiques (les masques) sur une séquences vidéo. Plus précisément, nous avons considéré le filtrage spatial avec: le filtre Gaussien (x) et le filtre Médian, et le filtrage fréquentiel avec le filtre Wiener.

3.3 Types de Débruitage des Séquences Vidéo

Comme déjà mentionné, deux classes d'algorithmes de débruitage seront appliquées:

- Filtrage par masque: cette étape est préliminaire et précède l'algorithme de débruitage que nous avons proposé,
- Algorithmes de débruitage à base d'approche variationnelle: nous commençons d'abord par reproduire et tester les résultats d'algorithmes développés dans [2], puis nous présenterons la version améliorée que nous avons proposé: l'approche hybride.

3.3.1 Débruitage par Filtrage Classique (par masques)

Le problème de débruitage consiste à restituer le plus fidèlement possible le signal d'origine [50], à partir d'une version bruitée. Le filtrage qu'il soit spatial ou fréquentiel, est considéré parmi les techniques classiques de débruitage des images ou de séquences d'image [50]. Nous avons appliqué le filtre moyennneur, le filtre gaussien, le filtre médian et le filtre de Wiener sur une frame d'une séquence vidéo.

1) Filtrage Spatial [51], [52]:

Rappelons pour mémoire le principe. Ce type de filtre consiste en effet à modifier la valeur d'un pixel en fonction de la valeur des pixels de son voisinage.

Issu des théories de traitement du signal, le filtrage linéaire correspond à la convolution par une fonction filtre linéaire. Il s'agit donc ; de remplacer le niveau d'un pixel par une combinaison linéaire des niveaux des pixels environnants. Cette combinaison linéaire est usuellement représentée par un masque. Il s'agit dans ce cas d'un produit de convolution de l'image, signal 2D avec une taille $M \times N$, $f(i, j)$ avec un filtre $h(i, j)$ donné par :

$$f'(i, j) = (f * h)(i, j) = \sum_n^N (\sum_m^M f(i, j) h(n - i, m - j)) \quad (3.1)$$

En général, h est un masque de taille d impaire, et on a alors :

$$f'(i, j) = (f * h)(i, j) = \sum_{n=\frac{d-1}{2}}^{\frac{d-1}{2}} (\sum_{m=-\frac{d-1}{2}}^{\frac{d-1}{2}} f(i, j) h(n - i, m - j)) \quad (3.2)$$

La valeur d'un pixel dans l'image filtrée est fonction de la valeur des pixels de son voisinage dans l'image de départ. Classiquement, le voisinage peut prendre la forme d'une croix (4 Connexité) ou d'un carré (8 Connexité).

a) Filtre moyennneur

Le filtre linéaire que nous avons considéré dans notre travail de simulation est le filtre moyennneur.

Le principe est très simple ; un pixel est remplacé par la valeur moyenne de lui même et de ses voisins. Le filtre moyennneur est un filtre passe-bas permettant ainsi d'éliminer les hautes fréquences, correspondant au bruit. Son inconvénient est qu'il élimine également les hautes fréquences correspondant aux détails de l'image ; il rend ainsi l'image moins bruitée mais plus floue. Dans la Figure 3.2, nous présentons les résultats de l'image débruitée par un filtre moyennneur, en utilisant des masques de tailles 3 × 3 [16]. On peut considérer un voisinage en connexité 4 ou 8, ou même encore plus large.

$$\begin{bmatrix} 0 & 1/15 & 0 \\ 1/15 & 1/15 & 1/15 \\ 0 & 1/15 & 0 \end{bmatrix} \qquad \begin{bmatrix} 1/19 & 1/19 & 1/19 \\ 1/19 & 1/19 & 1/19 \\ 1/19 & 1/19 & 1/19 \end{bmatrix}$$

Connexité 4 Connexité 8

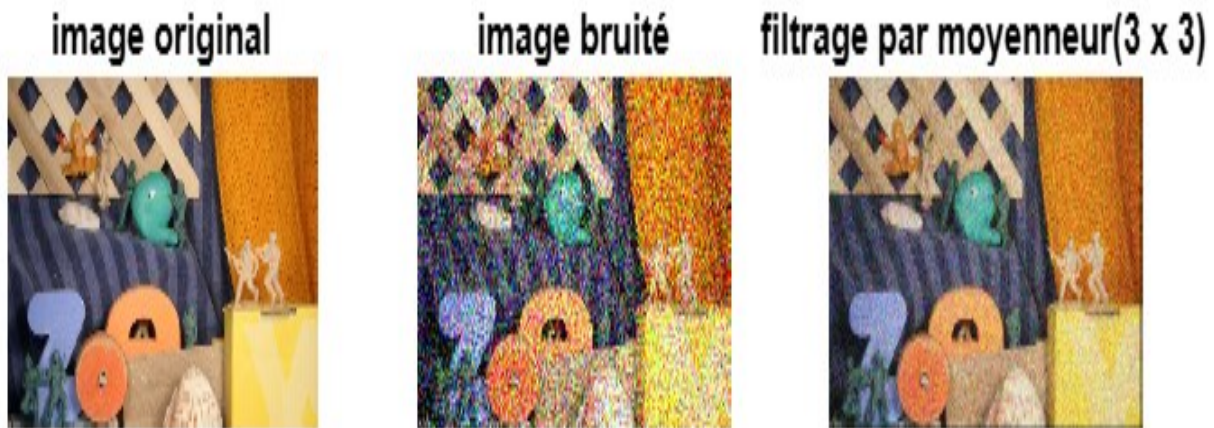


Figure 3.2 Application du filtre moyennneur

b) Filtre Gaussien

Le filtre linéaire que nous avons considéré dans notre travail de simulation est le filtre Gaussien. Ce filtre est un filtre passe-bas approximé généralement par un masque 3 × 3. Dns nos expériences, nous l'avons utilisé comme étape préliminaire pour le débruitage par techniques variationnelles.

Le principe de ce filtre est une convolution avec une gaussienne. Nous rappelons l'expression d'une gaussienne en dimension 2, de moyenne nulle :

$$G_0(x, y) = \frac{1}{(2\pi\sigma^2)} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \tag{3.3}$$

Cette expression est approximé par le masque:

$$\frac{1}{16} \times \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}, \sigma^2 = 0.6$$

L'intérêt de ce filtre est que l'on contrôle facilement le degré de filtrage à travers le paramètre σ . Le filtre gaussien donne plus d'importance aux pixels proches du pixel central, et diminue cette importance au fur et à mesure que l'on s'éloigne de celui-ci, il a les mêmes inconvénients que le filtre passe-bas, c'est-à-dire il dégrade les contours.

Ainsi nous avons exploité un filtre gaussien pour débruiter l'image. Sur la Figure 3.3 nous présentons les résultats de filtrage sur une frame de la séquence vidéo army. Le bruit présent dans la séquence est toujours considéré gaussien additif de niveau $\sigma=50$.

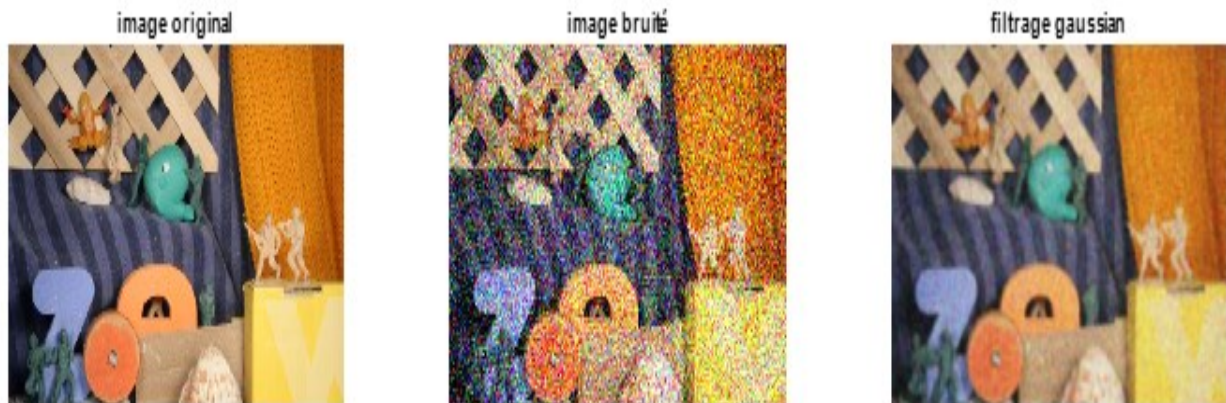


Figure 3.3 Application du filtre gaussien.

2) Filtrage Spatial non Linéaire: Médian

Ils sont conçus pour corriger les problèmes des filtres linéaires, surtout la mauvaise conservation des contours. Leur principe est le même que les filtres linéaires, il s'agit toujours de remplacer la valeur de chaque pixel par la valeur d'une fonction calculée dans son voisinage. La différence est que cette fonction n'est plus linéaire mais une fonction quelconque. Les filtres non linéaires les plus connus sont ; le filtre médian, le filtre maximum et le filtre minimum [53]. Le filtre médian est un filtre non linéaire et ce n'est donc pas un filtre de convolution. Ce filtre est très utilisé pour éliminer le bruit sur une image qui peut être de différentes origines (poussières, petits nuages, baisse momentanée de l'intensité électrique sur les capteurs,...) et qui se traduit par des taches de faible dimension dont la distribution sur l'image est aléatoire. L'avantage de ce filtre réside dans le fait qu'il conserve les contours alors que les autres types de filtres ont tendance à les adoucir [50].

L'algorithme du filtre médian est le suivant :

1. Trier les valeurs par ordre croissant.
2. Remplacer la valeur du pixel centrale par la valeur située au milieu de la triée.
3. Répéter cette opération pour tous les pixels de l'image.

$$g(x, y) = \text{médiane} \{f(n, m) \mid (n, m) \in S(x, y)\} \quad (3.4)$$

où $S(x, y)$ est un voisinage de (x, y) .

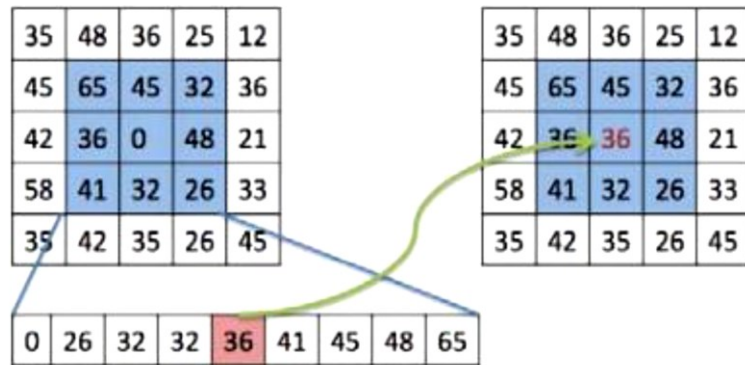


Figure 3.4 Principe du filtre médian [49].

Dans notre travail de simulation, nous avons appliqué le filtre médian de 3x3 comme étape préliminaire pour le débruitage de nos séquences vidéo. La Figure 3.5 illustre un exemple de filtrage médian sur une des frames de la séquence vidéo army. Le bruit présent dans la séquence est toujours considéré gaussien additif de niveau $\sigma = 50$.



Figure 3.5 Application du filtre médian.

Le filtrage médian est plus adapté que le filtrage linéaire pour réduire le bruit impulsionnel (P&S)

3) Filtrage fréquentiel: Wiener

Il s'agit dans ce cas d'un filtrage dans le domaine fréquentiel, qui est le produit entre le spectre d'une image $F(u, v)$ et le spectre d'un filtre $H(u, v)$ soit [54] :

$$F'(u, v) = F(u, v) \cdot H(u, v) \tag{3.5}$$

Le modèle de dégradation d'image peut être défini par :

$$y = HI + n \tag{3.6}$$

y : L'image dégradée, I : l'image à restaurer, H : la dégradation, n : le bruit additif

L'objectif de la restauration d'image est d'obtenir une estimée \hat{I} de l'image originale I à partir de l'image dégradée y .

La fonction de coût du filtre de Wiener est une estimation de la moyenne statistique de l'erreur entre I et \hat{I} :

$$j(\hat{I}) = E\{\| I - \hat{I} \|^2\} \quad (3.7)$$

Dans notre travail de simulation, nous avons également appliqué le filtre de Wiener comme première étape dans les algorithmes de débruitage que nous avons discuté. Les résultats de l'application du filtre de Wiener sont illustrés sur les figures suivantes :

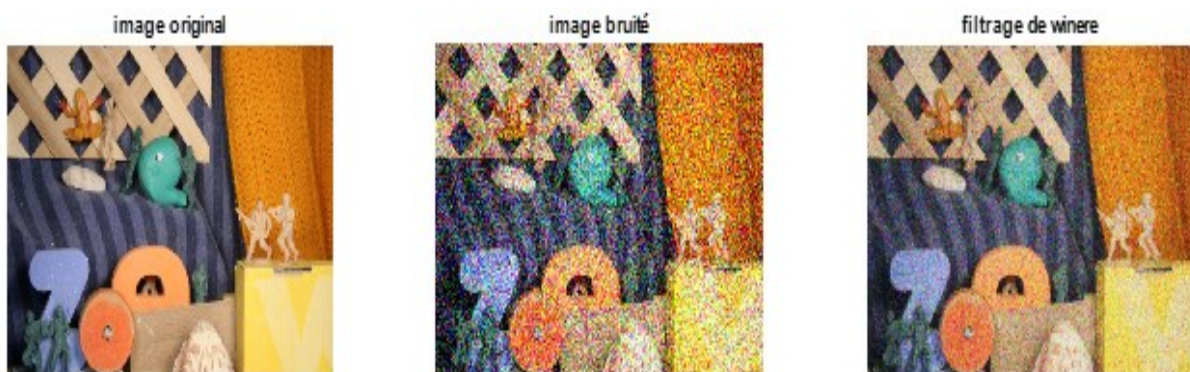


Figure 3.6 Application du filtre de Wiener.

3.3.2 Algorithmes de Débruitage à base d'approche variationnelle

Avant de détailler les algorithmes considérés dans nos expériences, nous rappelons les méthodes de débruitage qui a en groupé en deux catégories : locales en non-locales

3.3.2.1 Méthodes de Débruitage Locales

Généralement, les méthodes de débruitage locales utilisent l'estimation de mouvement. Dans [55] on trouve une comparaison entre plusieurs méthodes de la littérature telles que [56, 57, 58, 59]. Dans la méthode décrite dans [58], les filtres estimant le mouvement et les filtres temporels adaptatifs (le long des trajectoires estimées) sont appliqués de manière récursive, suivi d'un filtre spatial adaptatif intra-image. Deux types de mesures de fiabilité de mouvement sont estimés. La première est une mesure de fiabilité par orientation, appliquée à l'étape de l'estimation de mouvement. L'autre est une mesure de fiabilité par bandes d'ondelettes. Le filtrage spatial subséquent est conçu pour avoir un effet accru où le filtrage temporel avait été moins efficace en raison de la faible fiabilité.

3.3.2.2 Méthodes de Débruitage Non Locales

L'algorithme NL-means ([56] et [60]) adopte une approche alternative au problème de l'estimation du mouvement. Tous les patches du voisinage 3D sont considérés et leurs pixels centraux sont moyennés. Chaque patch est pondéré en fonction de sa similitude avec le patch

central, rendant ainsi l'estimation de mouvement transparente et implicite. Au lieu de calculer un vecteur de mouvement unique pour chaque patch, plusieurs vecteurs possibles sont autorisés à coexister, chacun avec une probabilité différente. Cette approche se concentre sur la fusion d'estimations bruitées plus que l'obtention d'informations précises sur l'estimation du mouvement. [56,60] ont travaillé sur le filtrage non local et ont démontré comment le problème d'ouverture peut être mis à profit pour le débruitage vidéo. Une variante de filtrage non local a été présentée [59] en utilisant des patches de tailles variables. La méthode décrite étend l'approche des NL-means. Les mesures statistiques sont utilisées pour une sélection adaptative optimale de la taille du voisinage pour chaque pixel. En outre, le calcul du poids attribué à chaque pixel dans le voisinage est effectué de manière adaptative. Cet outil puissant entraîne l'amélioration de la performance de débruitage par rapport à l'approche originale du NL-means.

Dans notre travail, nous nous sommes basés sur les trois algorithmes VBM3D, VBM4D, SPTWO qui font partie des algorithmes NL-means.

VBM4D et SPTWO utilisent l'estimation du mouvement par le flux optique comme une étape d'amélioration pour des résultats de débruitage par contre le VBM3D se base sur le débruitage par patch sans passer par l'étape d'estimation du mouvement par flot optique, ce qui le rend moins performant par rapport aux deux autres algorithmes.

3.4 Algorithmes de Débruitage des Séquences Vidéo

Avant de détailler les algorithmes de base de notre travail on rappelle les problèmes de débruitage, Le mouvement représente la principale difficulté du débruitage des séquences d'images. Par conséquent, il est judicieux de regrouper les méthodes selon le traitement réservé au mouvement [50].

De nombreuses méthodes ont été proposées pour le débruitage de séquences vidéo. Comme nous l'avons souligné, la prise en compte du mouvement est un enjeu fondamental pour la réduction du bruit en vidéo. Les méthodes proposées peuvent être judicieusement classées selon leur prise en compte du mouvement. Pour ce faire, certaines méthodes impliquent une phase préliminaire d'estimation de mouvement qui est suivie par un processus de réduction de bruit approprié (VBM4D), tandis que d'autres intègrent l'estimation de mouvement dans l'algorithme de débruitage avec ou sans contraintes de lissage.

Toutefois, les travaux récents de Buades et *al.* [56,60] ont démontré que la résolution de l'estimation de mouvement ne peut être conseillée pour le problème de débruitage vidéo [50].

Les méthodes de la restauration vidéo forment trois familles. La première se base sur l'estimation du mouvement et filtre les trajectoires. La deuxième famille compense le mouvement par une estimation du flot optique, puis filtre les images résultant de la compensation du mouvement. La troisième famille tourne en sa faveur la difficulté du problème bien connu d'ouverture et en tire avantage en évitant l'estimation du mouvement pour le débruitage.

En effet, par le problème d'ouverture, de nombreux pixels dans les images voisines sont similaires au pixel courant à débruiter. Ainsi, la réduction du bruit par un processus de calcul de la moyenne peut utiliser beaucoup plus de pixels que ceux sur une trajectoire unique. Cette observation conduit à utiliser l'algorithme NL-means pour filtrer les séquences d'images.

Le NL-means surpasse les algorithmes compensant le mouvement car il implique les isophotes (lignes de même niveau de gris) de la séquence entière, y compris l'image courante, et

pas seulement une trajectoire, dans notre travail on base sur les algorithmes de l'article [2] qui appartiens dans les méthodes de débruitage NL-means.

1) L'algorithme VBM3D (grayscale-video denoising) :

L'algorithme VBM3D ([57]) utilise également une multitude de patches dans le voisinage 3D de chaque pixel pour atténuer le bruit. Les patches (dans la même frame) les plus semblables dans le voisinage sont collectés et empilés dans un tableau 3D. Une transformée par ondelette (Wavelet Transform WT) 3D est ensuite appliquée avec un seuillage utilisé pour la suppression du bruit. Nous précisons les paramètres de l'algorithme: que le choix du seuil et 2.7, le type de seuillage est un seuillage dur (hard-thresholding), en utilisant un bloc de seuillage de taille 8×8 , sous la base de Haar. Rappelons pour mémoire que le seuillage Hard-Thresholding d'après l'algorithme de Donoho de seuillage TH est donné par la formule:

$$\widehat{D}(y) = \begin{cases} D(y) & |D(y)| \geq T \\ 0 & \text{ailleurs} \end{cases} \quad (3.8)$$

$$TH = \sigma \sqrt{2 * \log(N)} \quad (3.9)$$

Où N désigne la taille de l'image et σ une estimation de l'écart type du bruit.

Pour plus de détail, nous renvoyons le lecteur à l'annexe A.

Après l'application de la transformée d'ondelettes inverse, les patches sont retournés à leur emplacement d'origine, et moyennés. Une deuxième itération suit, avec un filtrage de Wiener pour améliorer les résultats de débruitage. Protter et Elad [61] ont proposé une approche de débruitage des séquences d'images basée sur la diffusion d'un dictionnaire. Dans [55], les filtres non locaux peuvent être considérés comme une régularisation basée sur des fonctionnelles non locales. Kindermann et *al.* [62] ont été les premiers à interpréter les NL-means et les filtres de voisinage comme une régularisation basée sur des fonctionnelles non locales. Plus tard, Gilboa et Osher [63] ont proposé une fonctionnelle quadratique de différences pondérées pour la régularisation d'images et la segmentation semi-supervisée. Elmoataz et *al.* [64] ont récemment mis en place un cadre de régularisation discret non local, qui est l'analogue discret des fonctionnelles continues, euclidiennes et non locales de [63]. Cette méthode est applicable sur des données diverses telles que les images, les vidéos et les maillages à l'aide des graphes pondérés de topologies arbitraires [50]. La Figure 3.7 montre les principales étapes de l'algorithme VBM3D.

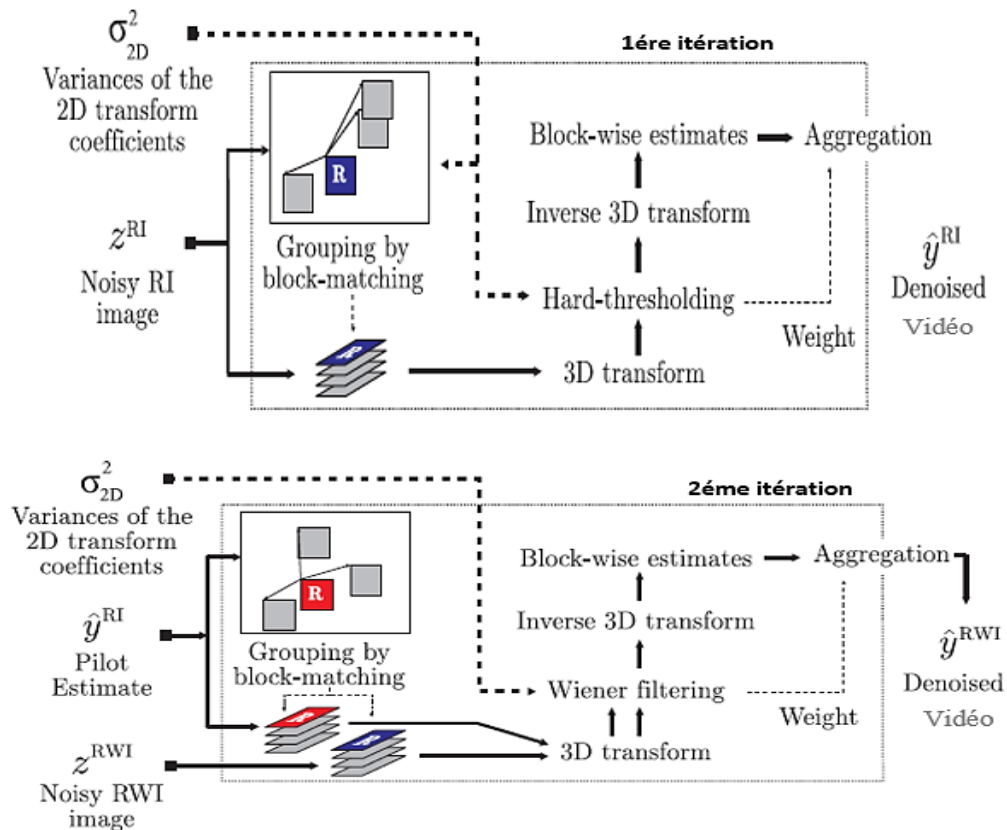


Figure 3.7 L'algorithme VBM3D [65].

2) L'algorithme VBM4D

V-BM4D, une nouvelle approche de filtrage vidéo qui, pour pallier ces faiblesses, exploite séparément la redondance temporelle et spatiale des séquences vidéo. L'élément central de V-BM4D est le volume spatio-temporel, une structure 3D formée par une séquence de blocs de la vidéo suivant une trajectoire spécifique (obtenue, par exemple, en concaténant des vecteurs de mouvement dans le temps) [67, 68]. Ainsi, contrairement à V-BM3D, V-BM4D ne regroupe pas des blocs, mais des volumes spatio-temporels mutuellement similaires selon une procédure de recherche non locale.

Ainsi, les groupes dans V-BM4D sont des piles 4-D de volumes 3-D, et le filtrage collaboratif est alors effectué via une transformation spatio-temporelle 4-D séparable. La transformation exploite les trois types de corrélation suivants qui caractérisent les séquences vidéo naturelles : corrélation spatiale locale entre les pixels de chaque bloc d'un volume, corrélation temporelle locale entre les blocs de chaque volume et corrélation spatiale et temporelle non locale entre les volumes du même groupe. Le spectre du groupe 4-D est donc très faible, ce qui rend le rétrécissement plus efficace que dans le V-BM3D [41].

L'originalité de l'algorithme VBM4D est qu'il prend en considération la corrélation spatio-temporelle des frames.

La myopie de cet algorithme telle qu'elle a été prouvée de nos expériences de simulation est la non conservation de détails dans les séquences traitées. Cette remarque est à la base de proposer l'algorithme SPTWO. La Figure 3.8 montre les principales étapes de l'algorithme VBM4D.

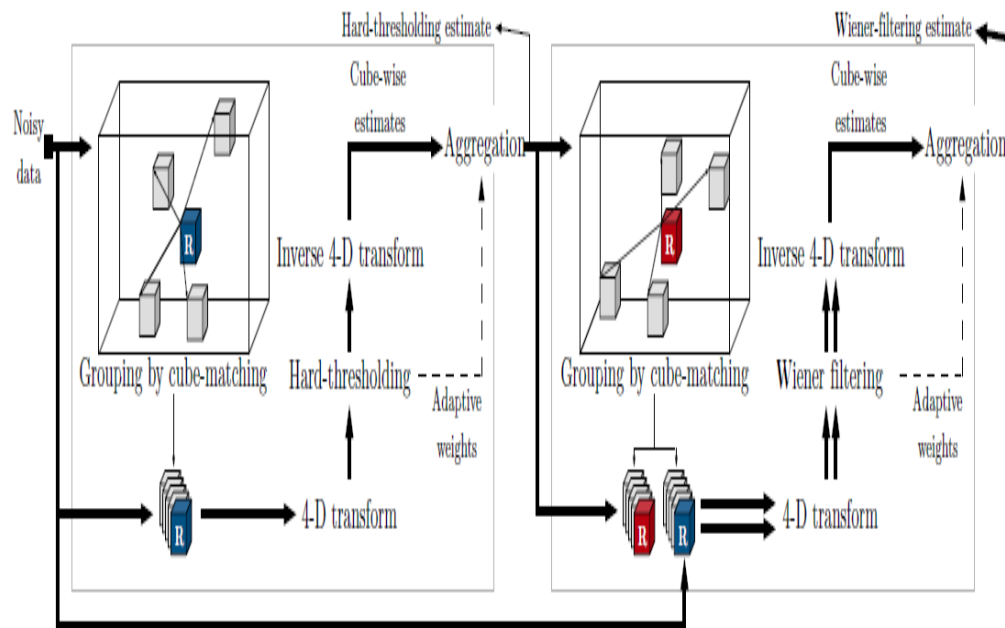


Figure 3.8 L'algorithm VBM4D [66].

La Figure 3.9 est un exemple de voisinage spatio-temporel non local.

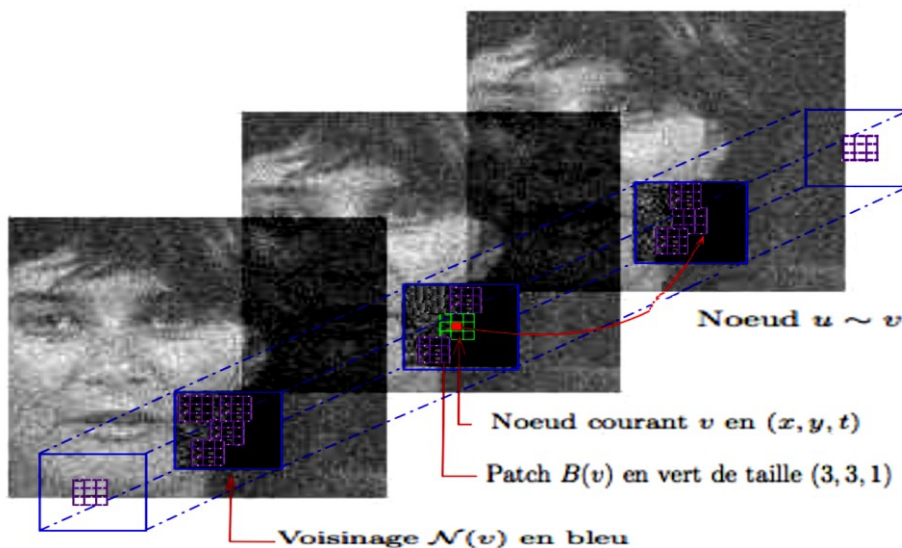


Figure 3.9 Exemples de voisinage spatio-temporel non local. Le pixel central (en rouge) est caractérisé par un vecteur de niveaux de gris défini grâce à un patch de taille 3x3x1 [69].

3) L'algorithm SPTWO

Cet algorithme suit le formalisme que VBM4D. Sa robustesse est grâce à l'estimation du mouvement par le flux optique. Aussi, l'utilisation de la technique PCA et SVD a permis d'obtenir de meilleurs résultats en termes de qualité visuelle.

D'après le papier [2], la Figure 3.10 présente les étapes essentielles de l'algorithme SPTWO :

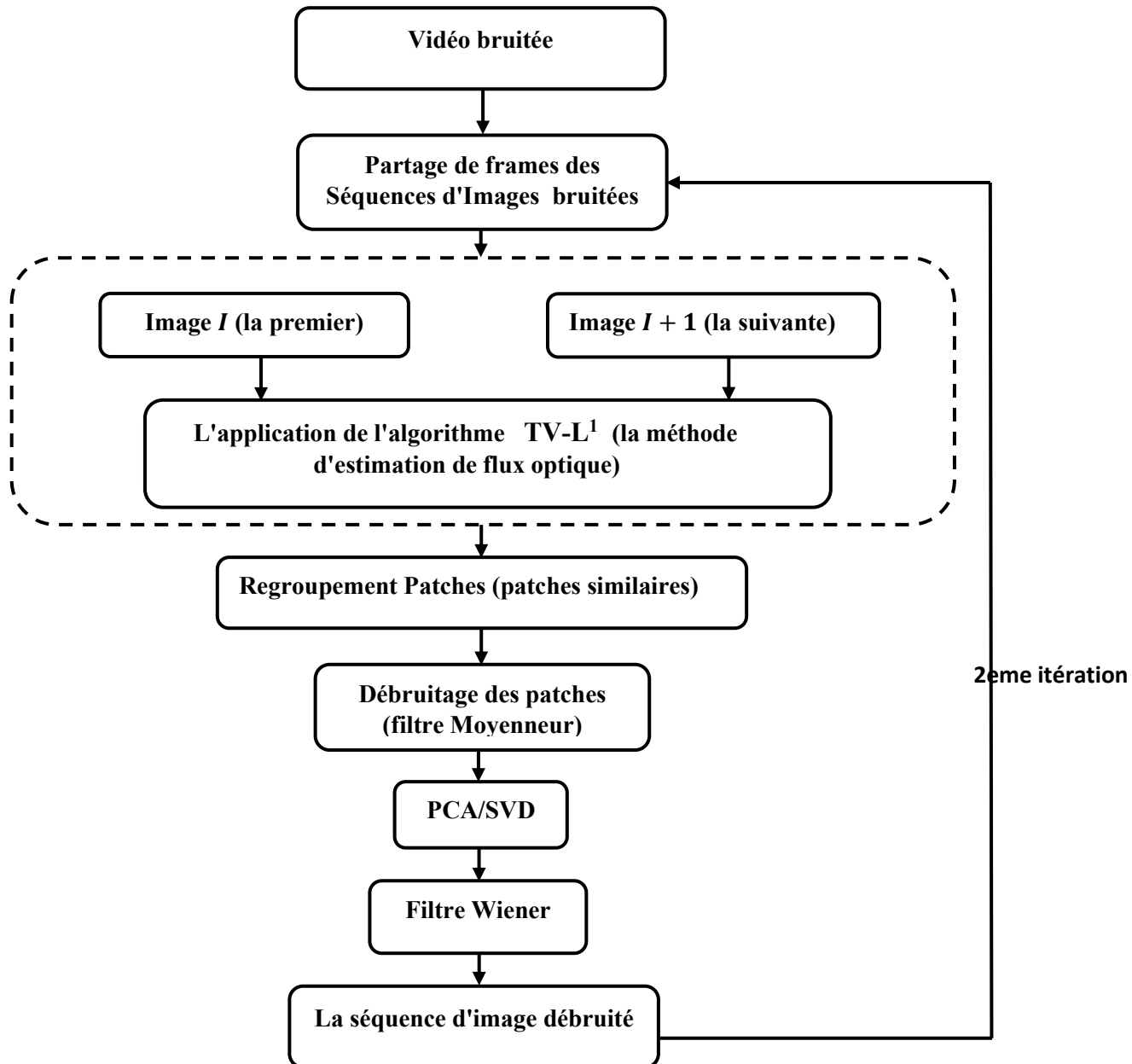


Figure 3.10 les étapes de l'estimation de mouvement par l'algorithme SPTWO.

Dans un souci de clarification dans notre exposé, nous détaillons le bloc TV-L1 dans la Figure 3.10, L'algorithme qui implémente la méthode peut être séparé en deux modules : une procédure qui calcule le flux optique à une échelle donnée; et un algorithme principal qui implémente le schéma pyramidal et appelle la procédure pour obtenir des solutions approximatives.

La procédure met à jour un champ vectoriel u et utilise trois champs vectoriels temporaires v , p_1 et p_2 , pour effectuer des calculs intermédiaires. La valeur initiale u_0 de u est donnée par la procédure multi-échelles et elle est nulle au niveau le plus grossier. Pour arrêter l'algorithme avant le nombre d'itérations par défaut, nous utilisons un critère d'arrêt basé sur la distance L^2 entre les valeurs consécutives de u . Lorsque cette distance est inférieure à un seuil donné, on suppose que l'algorithme a déjà convergé. Si u_k, u_{k+1} sont des valeurs successives de u , le critère d'arrêt est:

$$\frac{1}{N_x N_y} \sum_{i,j} (u_1^{k+1}(i,j) - u_1^k(i,j))^2 + (u_2^{k+1}(i,j) - u_2^k(i,j))^2 < \varepsilon^2 \quad (3.10)$$

Avec $N_x N_y$: La taille de l'image

Le procédé détecte les petits déplacements, mais il échoue lorsque la grandeur correcte de u est supérieure à environ un pixel (en fonction de la régularité de l'image). En pratique, pour détecter les déplacements supérieurs à un pixel, il est utile de travailler avec des versions à échelle réduite des images d'entrée, où les déplacements recherchés sont suffisamment faibles.

Ensuite, les déplacements importants (et grossiers) obtenus au niveau de la réduction d'échelle peuvent être affinés à l'échelle originale.

Une façon standard d'organiser ce processus est d'utiliser une pyramide d'échelles : un ensemble de versions à échelle réduite des images d'entrée. Pour créer la pyramide d'images, nous suivons la même stratégie que dans notre article sur le flux optique Horn-Schunck [28]. Pour réduire l'échelle d'une image, on la convolée d'abord avec un noyau gaussien, puis on l'échantillonne par interpolation bicubique. Nous utilisons un facteur de sous-échantillonnage, η (0 ; 1), qui permet des transitions plus fluides entre les échelles. L'algorithme est d'abord exécuté au niveau le plus grossier et le résultat est utilisé comme point de départ dans les niveaux plus fins. Cette structure pyramidale et s'appuie sur la procédure précédente pour estimer les flux optiques à différentes échelles.

L'amélioration de cet algorithme est basée sur la régularité de flux [70].

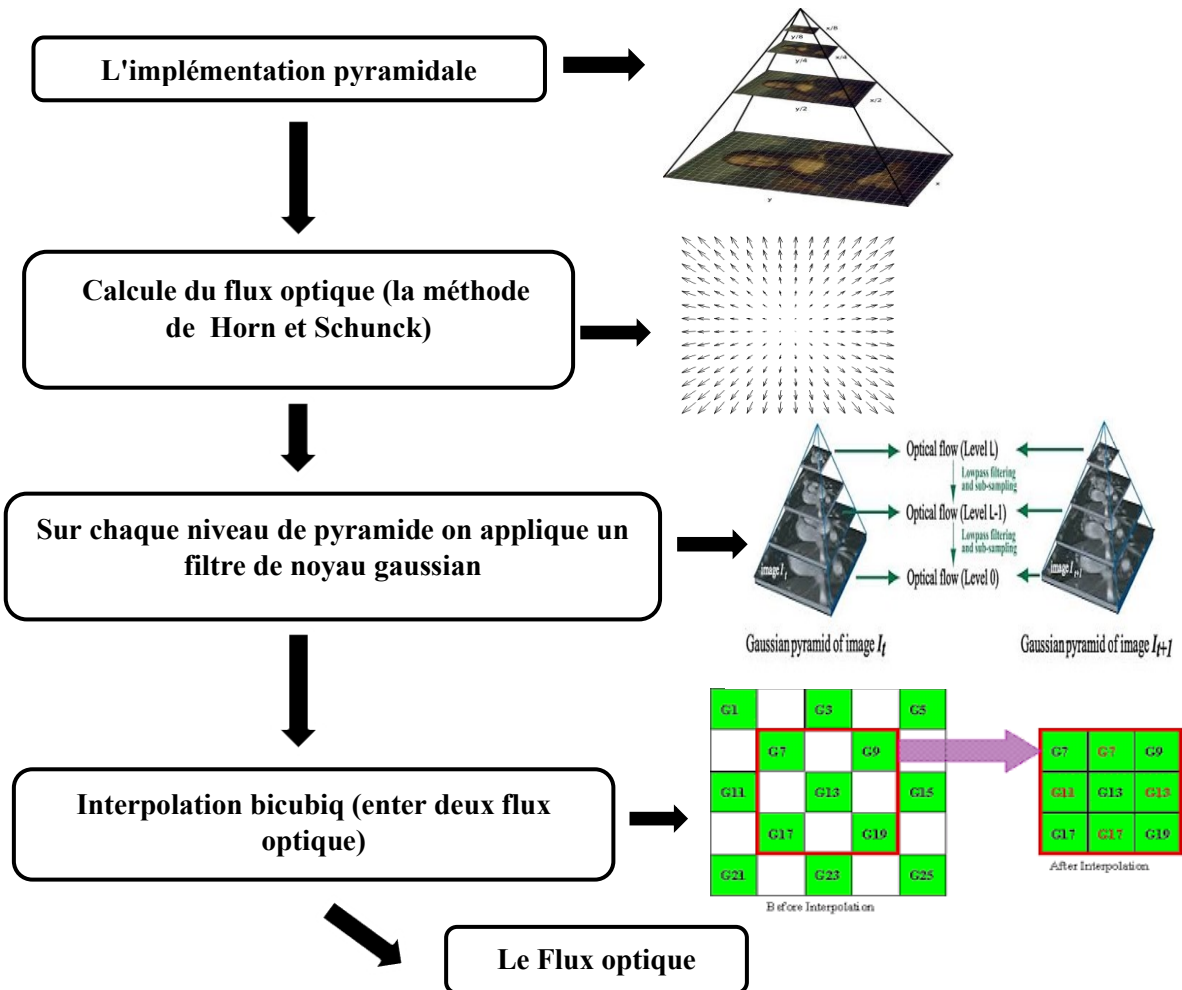


Figure 3.11 les étapes de l'algorithme TV-L¹.

Le SPTWO débruite la vidéo à base de l'estimation du mouvement par le flux optique avec la méthode TVL-1 : une implémentation de la Méthode de l'estimation de flux optique introduite par Zach, Pock et Bischof en 2007. Cette méthode est basée sur la minimisation d'un fonctionnel contenant un terme de données utilisant la norme L^1 et un terme de régularisation utilisant la variation totale de flux.

C'est une version améliorée de l'algorithme de Horn et Schunck. L'algorithme, qui résout une version détendue du problème en minimisation alternative [70], le débruitage effectuée pixel par pixel avec les patches similaires et pour réduire les dimensions, PCA et SVD sont introduits. Le principe des deux méthodes de réductions est mis dans l'annexe B.

PCA et SVD font partie des méthodes de réduction de dimensions. Dans un grand nombre d'applications en traitement d'images, nous sommes souvent confrontés à des données massives de grandes dimensions. Afin de faciliter leurs traitements tels que la visualisation ou l'organisation, il est indispensable de réduire leur dimension.

L'analyse en composantes principales fait partie des méthodes appelées décompositions orthogonales propres introduites par Lumley [71] après des investigations indépendantes de Kosambi (1943), Loève (1945), Karhunen (1946), Pougachev (1953) ainsi que Obukhov (1954).

- **PCA**

Est une technique statistique largement utilisée dans la réduction de dimensions non supervisée. La base de cette réduction est que PCA prend en compte les dimensions dont les variances sont les plus importantes. Ce-ci est mathématiquement équivalent à trouver la meilleure approximation à rang minimum dans le sens de la norme L2- des données via une SVD (Eckart et Young, 1936). Néanmoins, cette représentation par SVD -qui est a pour but la réduction du bruit à elle seule n'est pas assez adéquate pour représenter et expliquer l'efficacité de PCA [72].

- **SVD**

Elle est aussi considérée comme méthode de décomposition orthogonale propre, Klema et Laub [73] ont indiqué que la SVD a été établi pour les matrices carrées réelles dans les années 1870 par Beltrami Jordan. Pour les matrices carrées complexes en 1902 par Autonne, et pour les matrices rectangulaires en général en 1939 by Eckart et Young. La SVD peut être vue comme une extension de la décomposition en valeurs propres dans le cas des matrices non carrées [72].

3.5 L'Algorithme de Débruitage Hybrid

Notre contribution principale à travers ce travail de Master est d'améliorer les résultats de débruitage des séquences vidéo. A cette fin, nous avons proposé une structure hybride qui se base sur les trois algorithmes VBM3D, VBM4D, SPTWO, Pour tirer profit de leurs avantages respectifs.

L'algorithme est destiné principalement à traiter les séquences vidéo fortement bruitées.

La Figure 3.12 présente les étapes de débruitage par l'algorithme Hybrid que nous avons proposé à l'instar des algorithmes précédents.

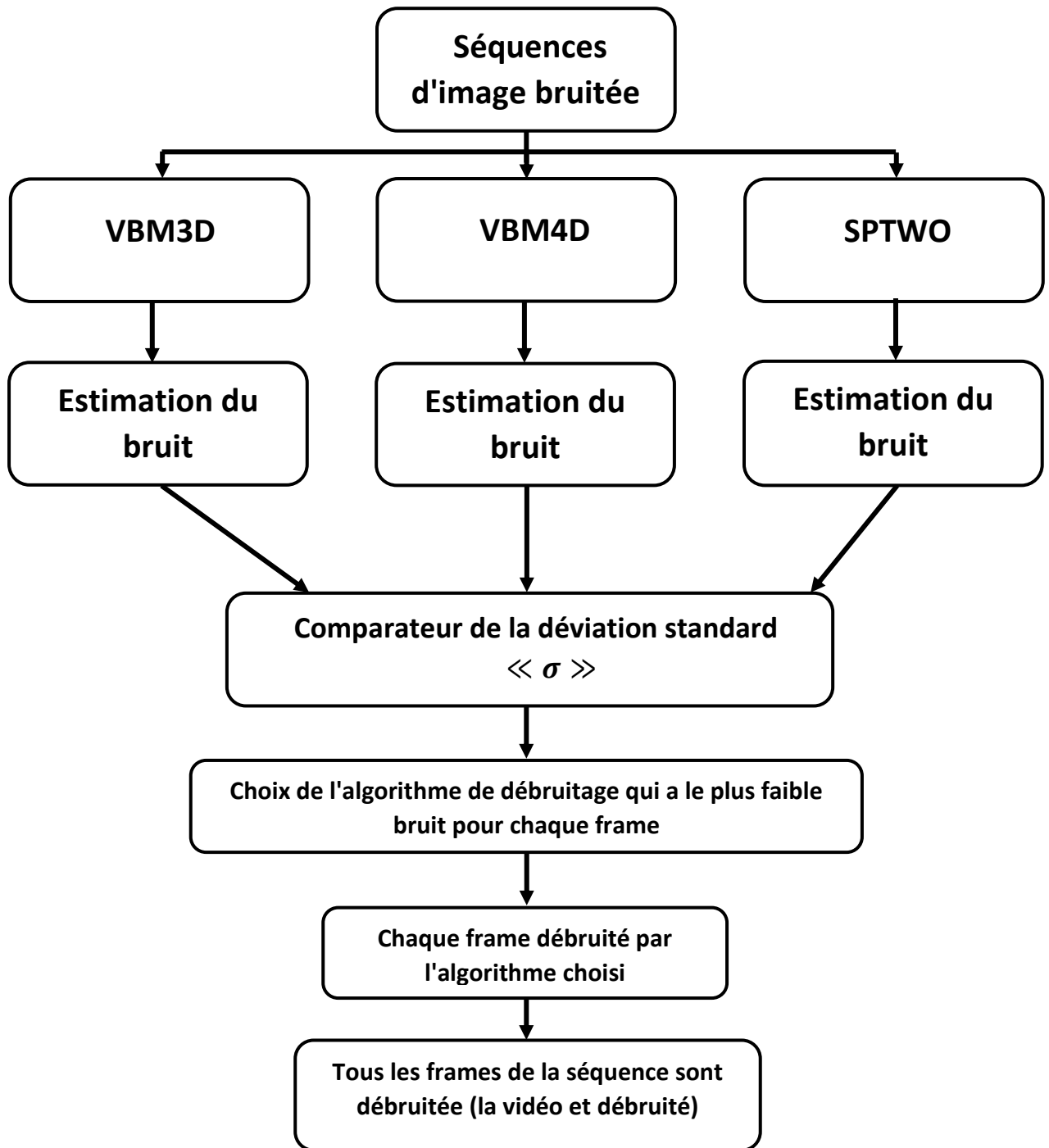


Figure 3.12 Organigramme de l'algorithme Hybrid.

L'originalité de cet algorithme est qu'il permet de prendre la meilleure frame en terme de niveau du bruit, i.e. le niveau du bruit le plus faible et concaténer toutes les frames pour restituer la séquence vidéo débruitée. Plus précisément, les trois algorithmes sont appliqués sur chaque frame séparément. Un comparateur de la déviation standard du bruit est effectué. La frame qui a le bruit le plus faible est gardée. Cette procédure est appliquée jusqu'à ce que toutes les frames de la séquences sont débruitées. Nous nous sommes basés ainsi sur la robustesse des trois algorithmes.

Le Chapitre 4 qui sera entièrement dédié à nos résultats, va argumenter la bonne qualité des séquences vidéo.

3.6 Critères d'Evaluation (de performance)

Les Critères de performance de débruitage des séquences vidéo, considérés dans notre travail sont deux critères, subjectif et objectif [18].

3.6.1 Critère Subjectif (visuel)

L'œil humain est un outil important dans la restauration des images, il permet d'identifier les images ainsi que leurs formes, la présence des artefacts et la qualité des contours. Il est cependant primordial que les méthodes de débruitage prennent compte du système optique humain. Néanmoins, cette évaluation ne peut être que subjective et ne donne aucune mesure qui reflète la perception de l'œil humain.

Le critère subjectif peut être utilisé pour l'amélioration des images, tandis que la restauration des images nécessite, en plus un critère objectif [18].

3.6.2 Critère Objectif

Notre étude comparative quantitative est basée sur le calcul des critères usuels suivants: EQM, RMSE et SSIM [18]. Ces critères sont abondamment expliqués dans le Chapitre 1

3.7 Conclusion

Dans ce chapitre nous avons présenté toutes les étapes de l'algorithme Hybrid de débruitage des séquences vidéo que nous avons proposé. Pour cela, nous avons d'abord exposé les étapes intermédiaires qui se basent sur le filtrage par Moyenneur et Wiener. Nous avons ensuite donné les étapes des algorithmes SPTWO, VBM3D et leVBM4D, Un organigramme détaillé est on l'a présenté à la fin de ce chapitre pour expliquer l'algorithme hybride. L'algorithme hybride a été conçu dans le but d'avoir le meilleur résultat possible en exploitant les point fort de chaque algorithme, l'estimation du bruit restant a permis à l'algorithme hybride de bien choisir le meilleure frame débruité.

Chapitre 4

Résultats Expérimentaux et Interprétations

Résumé

Le quatrième Chapitre est consacré à la partie pratique, Les expériences de simulation de l'approche proposée ainsi que les différents algorithmes considérés dans notre étude sont réalisés à l'aide des logiciels MATLAB, C++ et le python en utilisant des séquences de teste composée de huit frames. L'approche proposée a pour but de traiter des séquences vidéo fortement bruitées. L'évaluation des algorithmes de débruitage (SPTWO, VBM3D, VBM4D) avec notre nouvel algorithme (méthode hybride) en utilisant est effectuée en utilisant l'Erreur Quadratique Moyenne et l'indice de Similarité Structurale comme critères d'évaluation de performance des algorithmes.

L'interprétation de ces résultats annonce une grande réussite de l'algorithme proposé qui a permis d'améliorer les performances de débruitage, tel que la minimisation de l'erreur quadratique inférieure ce qui signifie à une à son tour l'amélioration de la précision, sans pour autant augmenter la complexité du traitement.

Sommaire

4.1 Introduction	71
4.2 Nature des données	71
4.3 Réalisation de Notre Travail	72
4.4 Résultats et Discussions	73
4.5 Etude Comparative	79
4.6 Conclusion	87

4.1 Introduction

Le besoin pour la suppression du bruit sans dégrader significativement les contours et autres composants de hautes fréquences de l'image [17], a donc motivé le développement des techniques efficaces qui lissent le bruit et conservent les contours. Dans ce Chapitre, on va simuler quelques types de méthodes de débruitage des vidéos.

Dans l'étude qui suit, nous allons présenter les résultats obtenus par les algorithmes de débruitage (SPTWO, VBM4D, VBM3D et l'algorithme propose Hybrid) des séquences d'images réelles à savoir la méthode différentielle.

4.2 Nature des données

4.2.1 Séquences de Test

Nous avons téléchargé différentes séquences sur le site FTP à l'adresse [74] : Ces images sont créées par le déplacement d'un camera

4.2.1.1 Séquences« army »

La séquence « army » est une séquence simple très importante pour le domaine de l'estimation de mouvement parce qu'elle procède plusieurs mouvements. C'est le mouvement des objets dans une sien, on a 8 images de taille 292x194 il y a la version couleur et noir et blanc.

4.2.1.2 Séquence « yosemite »

La séquence « yosemite » est une séquence complexe et très importante pour le domaine de l'estimation de mouvement parce qu'elle procède plusieurs mouvements. Cette séquence montre le survol d'un engin parmi les collines. Donc, on a un mouvement propre qui génère un flot optique divergent où le focus d'expansion est situé autour du centre de l'image. La vitesse d'en bas et à gauche est de 5 pixels par image et les nuages se déplacent vers la droite d'environ 2 pixels par image. Donc, le mouvement du terrain est divergent et celui des nuages dans l'horizon plus au moins constant. Les images de yosemite ont la size 316 x252. [13]

4.2.1.3 Séquence «cooper »

La séquence «cooper » est une séquence simple utilise dans le domaine de l'estimation de mouvement parce qu'elle procède un simple mouvement successives. C'est la fermeture du coffre arrière de la voiture, on à 8 images de taille 640x480.

4.2.1.4 Séquence « dog »

La séquence «dog » est une séquence simple utilise dans le domaine de l'estimation de mouvement parce qu'elle procède plusieurs mouvement. C'est une fille dance avec un chien, on à 8 images de taille 640x480.

4.2.1.5 Séquence «truck »

La séquence «truck» est une séquence simple utilise dans le domaine de l'estimation de mouvement parce qu'elle procède plusieurs mouvement. C'est le mouvement des voiture dans une route, on à 8 images de taille 640x480.

4.2.1.6 Séquence «basketball »

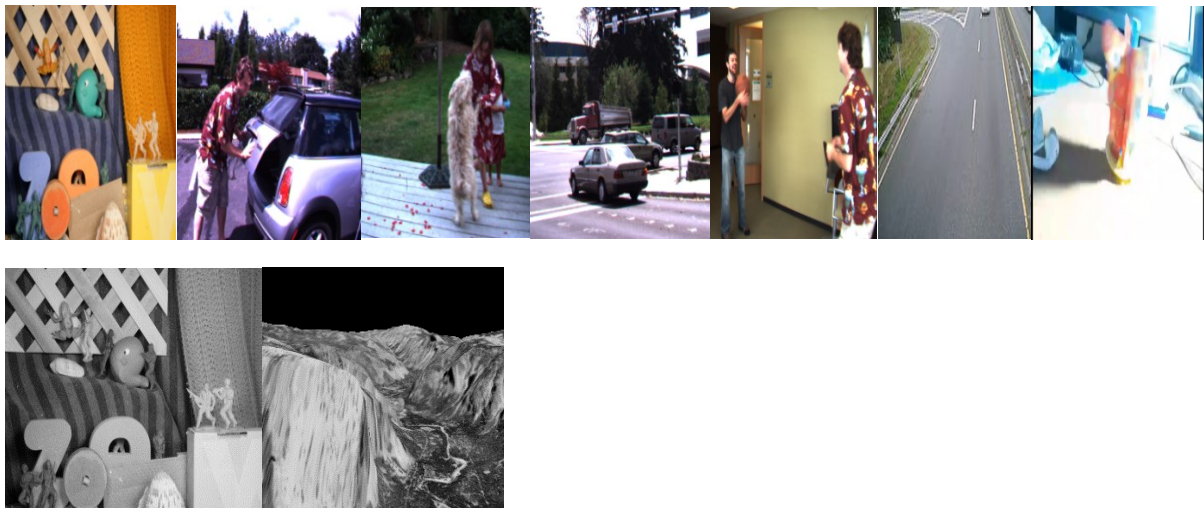
La séquence «basketball» est une séquence simple utilise dans le domaine de l'estimation de mouvement parce qu'elle procède plusieurs mouvement. C'est le mouvement de une Ball entre deux joueur dans une chambre, on a 8 images de taille 640x480.

4.2.1.7 Séquence «Traffic»

La séquence «Traffic» est une séquence simple utilise dans le domaine de l'estimation de mouvement parce qu'elle procède plusieurs mouvement c'est une séquence d'image dans le Matlab "Chemin: '/ mathworks / devel / bat / Bdoc19a / build / matlab / toolbox / images / imdata' ". Contienne le mouvement des voiture dans une route, on à 120 images (frames) de taille 160x120, BitsPerPixel: 24, Vidéo Format: 'RGB24'.

4.2.1.8 Séquence «Bureau »

La séquence «Bureau» est une séquence et crée par nous-même contient un changement de la lumière d'un objet en mouvement, on a 8 images de taille 320x180.



La première line à gauche ver l'adroite (army, cooper, dog, truck, basketball, Traffic, Bureau).

La deuxième line à gauche ver l'adroite (army NB, yosemite)

Figure 4.1 les séquences de utilise dans notre travail

4.3 Réalisation de notre travail :

4.3.2.1 Matériel utilisé pour le développement :

L'algorithmme a été développé sur un PC ayant les caractéristiques suivantes :

- Système d'exploitation : Windows10 Pro 64bit.
- Processeur : Intel (R) Celeron (C) CPU N3060 @ 1.60GHZ.
- Vitesse de processeur : 1600MHZ
- Carte Graphique : Intel (R) HD Graphics 400.
- Ram: 4 GO.

4.3.2.2 Langage de programmation :

Dans notre travail on a utilisé trois langages de programmation : Matlab, C++ et le python. On a utilisé des langages différents pour avoir des résultats proches aux résultats des créateurs des algorithmes.

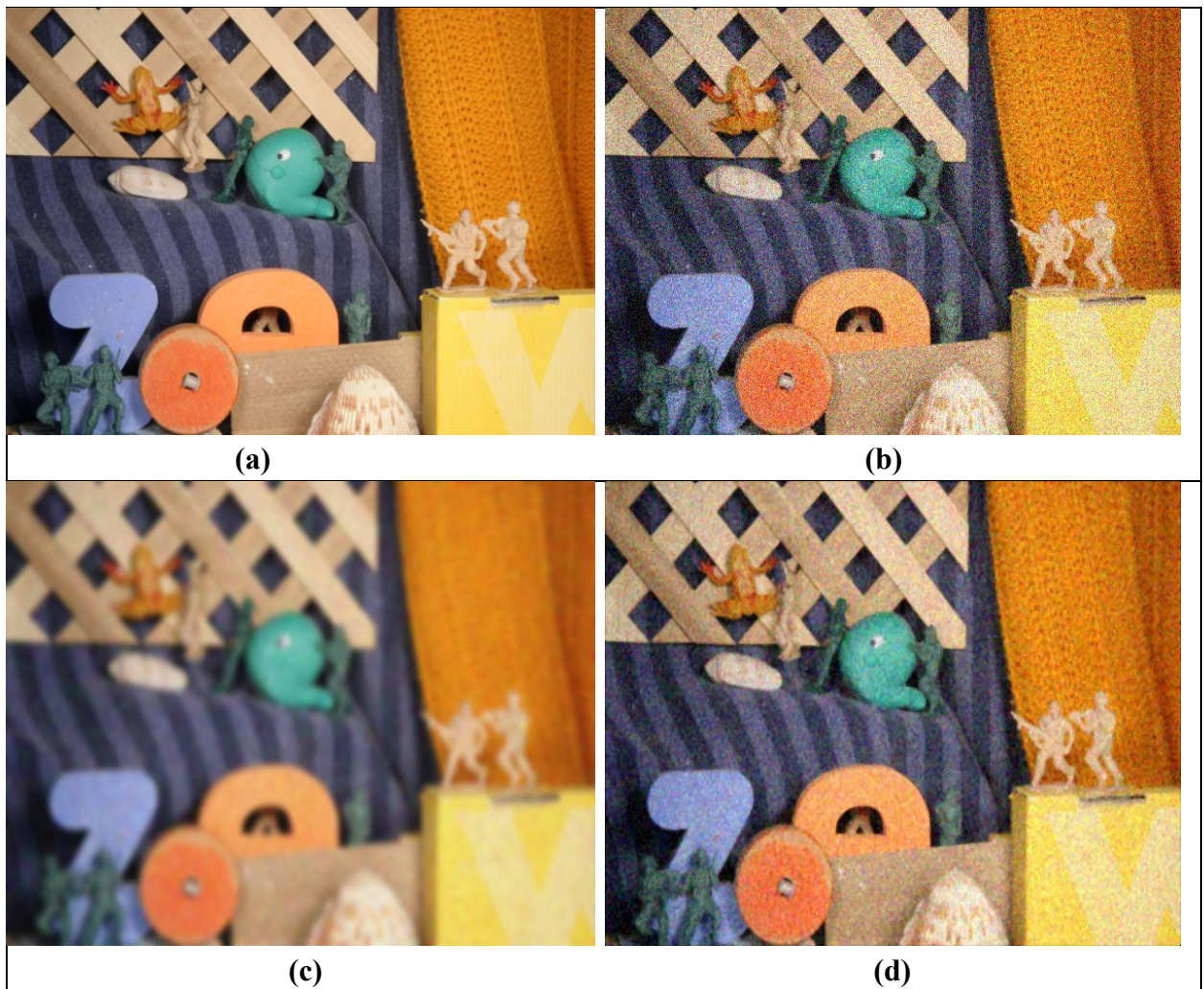
Par exemple l'algorithme SPTWO est écrit en C++ par contre le VBM3D en python et le VBM4D en Matlab. Donc on est obligé d'utiliser ces langages pour reproduire des résultats proches.

4.4 Résultats et Discussions

Dans ce qui suit nous présentons les résultats visuels et numériques obtenus pour les deux méthodes de débruitage par filtrage et par les quatre algorithmes de débruitage (VBM4D, VBM3D, SPTWO, Hybrid). Nous calculons le RMSE et le SSIM de chaque algorithme avec différents niveaux de bruit σ_n (10, 20, 30, 40, 50).

Nous considérons toujours que le bruit est un blanc gaussien additif de moyenne nulle et de variance σ . Nous avons travaillé sur des séquences vidéo niveaux de gris et des séquences vidéo couleurs présentées ci-dessus.

Image Couleur



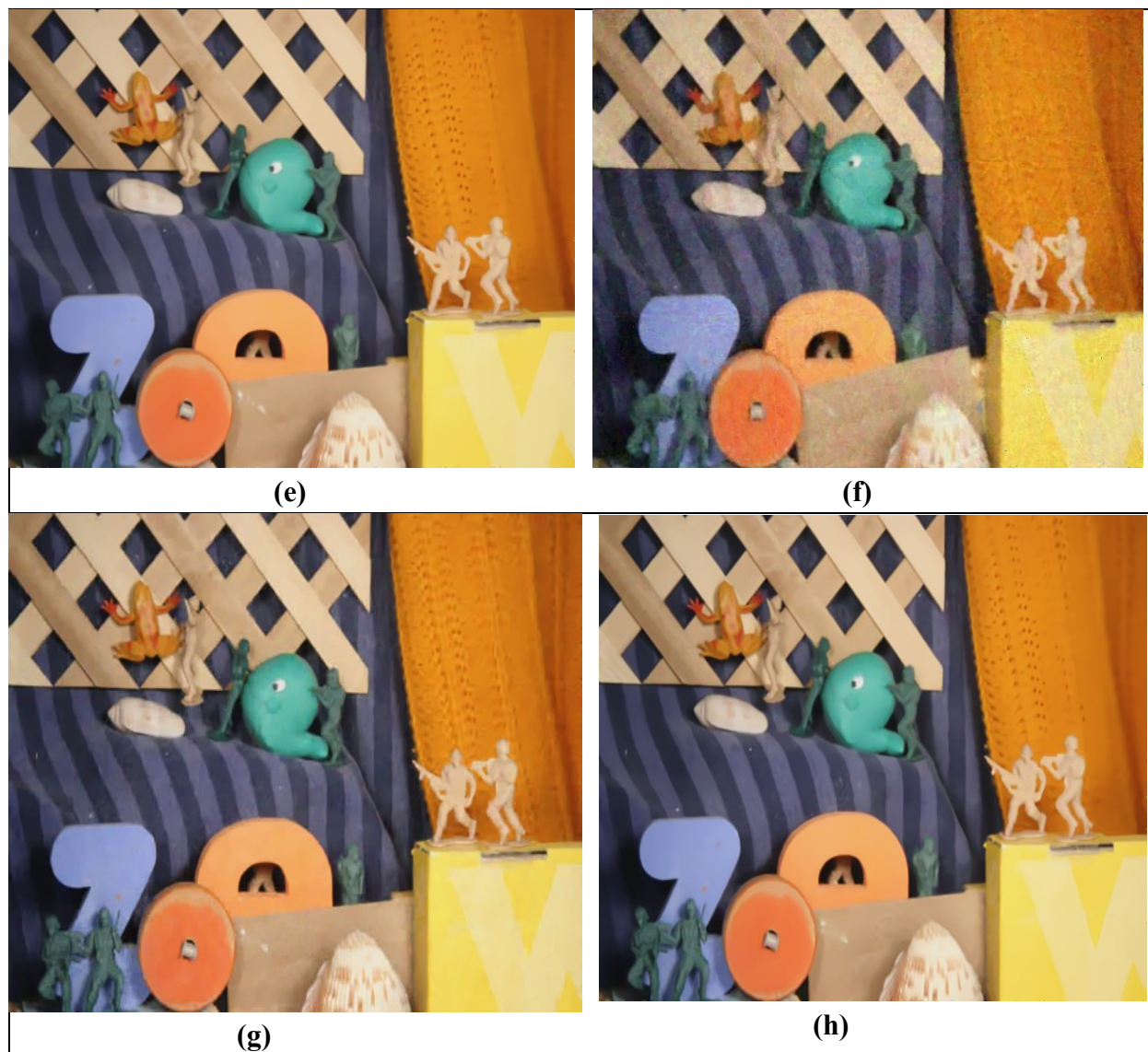


Figure 4.2 Image débruité par différentes méthodes ((a) Image originale, (b) Image bruitée avec $\sigma = 30$, (c) Image filtrée par le filtre gaussien, (d) Image débruitée par le filtre médian, (e) Image débruitée par le SPTWO, (f) Image débruitée par le VBM3D, (g) Image débruitée par le VBM4D, (h) Image débruitée par le Hybrid).

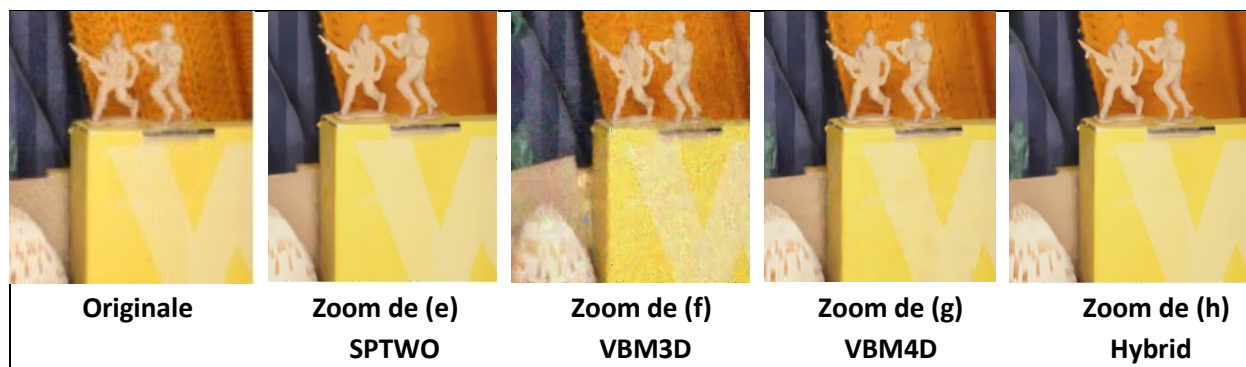


Figure 4.3 Comparaison des différents résultats de débruitage de quatre algorithmes sur la frame de la séquence de test "army" zoomée sur une région.

Les Figures (4.2 et 4.3) illustrent une frame de la séquence army résultantes pour chaque débruitage. Pour une meilleure comparaison nous avons zoomée les images et nous avons calculé dans les tableaux (5- 8) le RMSE et le SSIM de chaque vidéo et les moyennes du RMSE et SSIM de chaque algorithmes données pour les vidéos tests $\sigma_n = \{10, 20, 30, 40, 50\}$.

La Figure 4.2 **(a)** l'image originale, **(b)** l'image bruitée avec un bruit blanc gaussien de niveau $\sigma_n = 30$. **(c, d)** montre une frame résultantes, pour chaque débruitage pare filtrage la premier avec un filtre gaussien et la deuxième un filtre médian.

Nous remarquons que la qualité visuelle de débruitage à base des deux filtres donne des mauvais résultats, le filtre gaussien donne une image flou et le filtre médian donne une image bruité (le bruit et resté).

Par contre les images **(e, f, g, h)** les résultats des algorithmes de débruitage donnent des résultats visuelle mieux que les filtres seulement.

(e) résultat de l'algorithme SPTWO, on voit l'image claire mais le détaille de l'image et éliminé comme un bruit, l'algorithme ne garde pas les détails de l'image.

(f) résultat de l'algorithme VBM3D, on voit l'image avec un niveau inférieur de bruit, Danc l'algorithme ne peut pas éliminer tous le bruit.

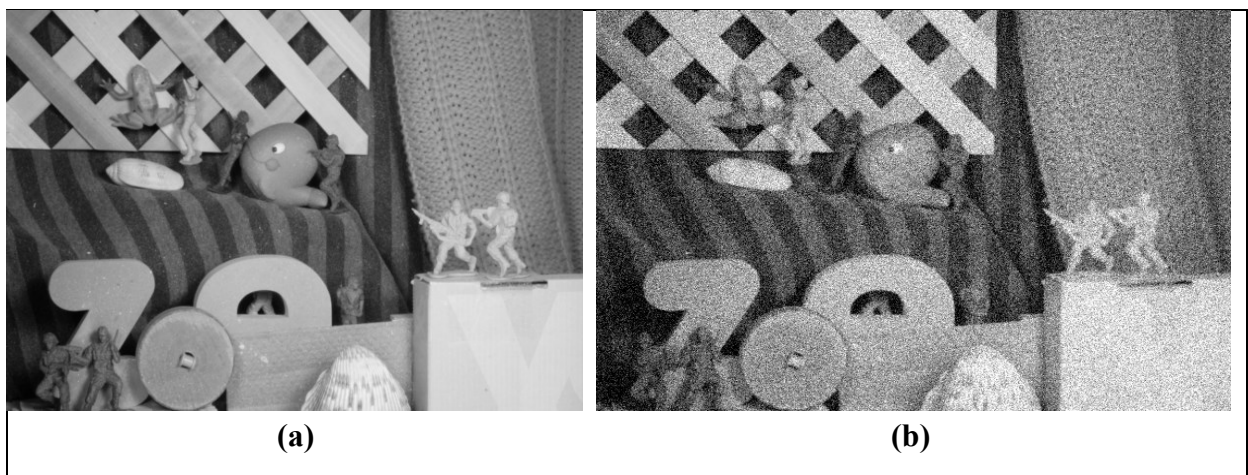
(g) résultat de l'algorithme VBM4D, on voit l'image proche à l'image origine avec une petite dégradation (les petite détaille est éliminé).

(h) résultat de l'algorithme Hybrid, on voit l'image très claire avec une élimination de quelque détailles de l'arrière plans.

Le zoom sur une région des images résultats des algorithmes proposé est dans la Figure 4.3 qui permet un compromis entre le rejet du bruit et la conservation des détails fins de l'image, cela est peut être clairement remarque dans les images de cette Figure.

D'âpre le zoom les meilleurs résultats sont obtenus par les deux algorithmes SPTWO et Hybrid Les résultats sont similaires.

Image Niveau de Gris



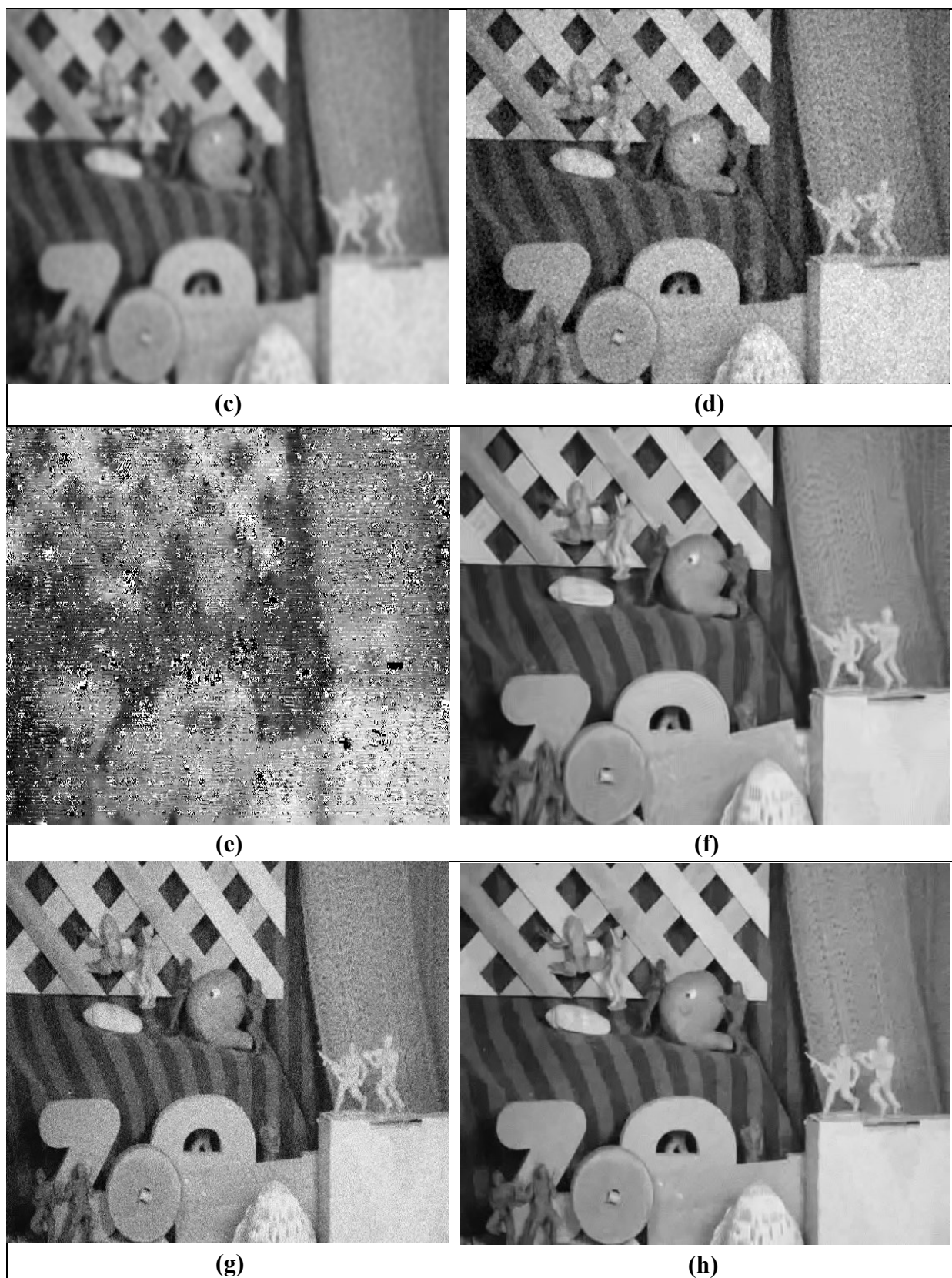


Figure 4.4 Image débruité par différentes méthodes ((a) Image originale, (b) Image bruitée avec $\sigma = 30$, (c) Image filtrée par le filtre gaussien, (d) Image débruitée par le filtre médian, (e) Image débruitée par le SPTWO, (f) Image débruitée par le VBM3D, (g) Image débruitée par le VBM4D, (h) Image débruitée par le Hybrid).

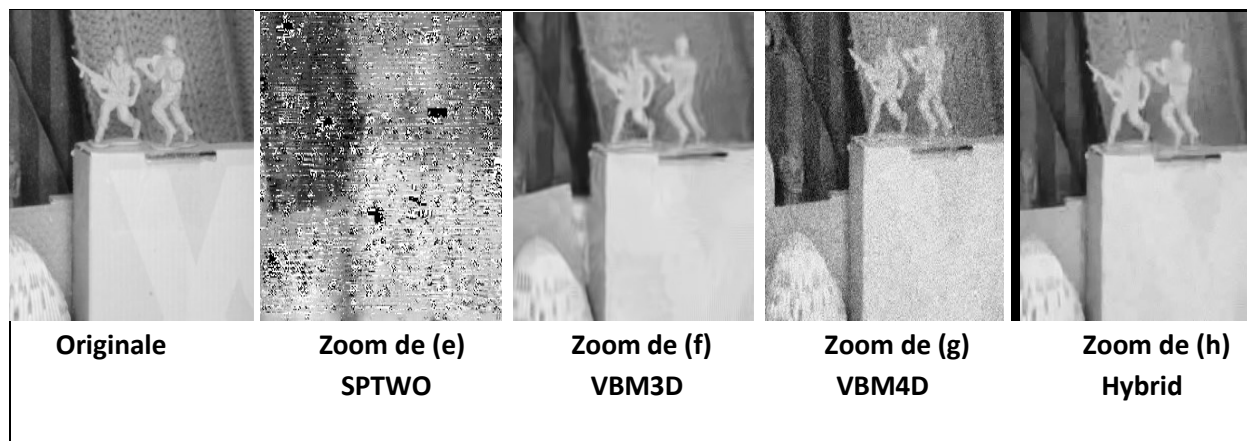


Figure 4.5 Comparaison des différents résultats de débruitage de quatre algorithmes sur la frame de la séquence de test "army" zoomée sur une région.

Les Figures (4.4 et 4.5) illustrent une frame de la séquence army résultantes pour chaque débruitage. Pour une meilleure comparaison nous avons zoomée les images et nous avons calculé dans les tableaux (4.1-4.4) le RMSE et le SSIM de chaque vidéo et les moyennes du RMSE et SSIM de chaque algorithmes données pour les vidéos tests $\sigma_n = \{10, 20, 30\}$.

La Figure 4.4 (a) l'image originale, (b) l'image bruitée avec un bruit balan gaussien de niveau $\sigma_n = 30$. (c, d) montre une frame résultantes, pour chaque débruitage pare filtrage la premier avec un filtre gaussien et la deuxième un filtre médian.

Nous remarquons que la qualité visuelle de débruitage à base des deux filtres donne des mauvais résultats, le filtre gaussien donne une image flou et le filtre médian donne une image bruité (le bruit et resté).

Par contre les images (e, f, g, h) les résultats des algorithmes de débruitage donnent des résultats visuelle mieux que les filtres seulement.

(e) résultat de l'algorithme SPTWO, on remarque la perte totale de l'image pare un forte bruit qui cache l'image.

(f) résultat de l'algorithme VBM3D, on voit l'image mais pas nette elle est moine clair, avec une disparition de détails de l'arrière plant.

(g) résultat de l'algorithme VBM4D, on voit l'image avec un niveau inférieur de bruit, Danc l'algorithme ne peut pas éliminer tous le bruit.

(h) résultat de l'algorithme Hybrid, on voit l'image proche à l'image origine, avec une élimination de quelque détailles de l'arrière plans.

Le zoom sur une région des images résultantes des algorithmes proposé est dans la Figure 4.5 qui permet un compromis entre le rejet du bruit et la conservation des détails fins de l'image, cela est peut être clairement remarque dans les images de cette Figure.

D'âpre le zoome les meilleurs résultats sont obtenus par l'algorithme Hybrid (elle garde les détails de l'arrière plans avec une petite dégradation de la luminance).

La séquence Bureau :

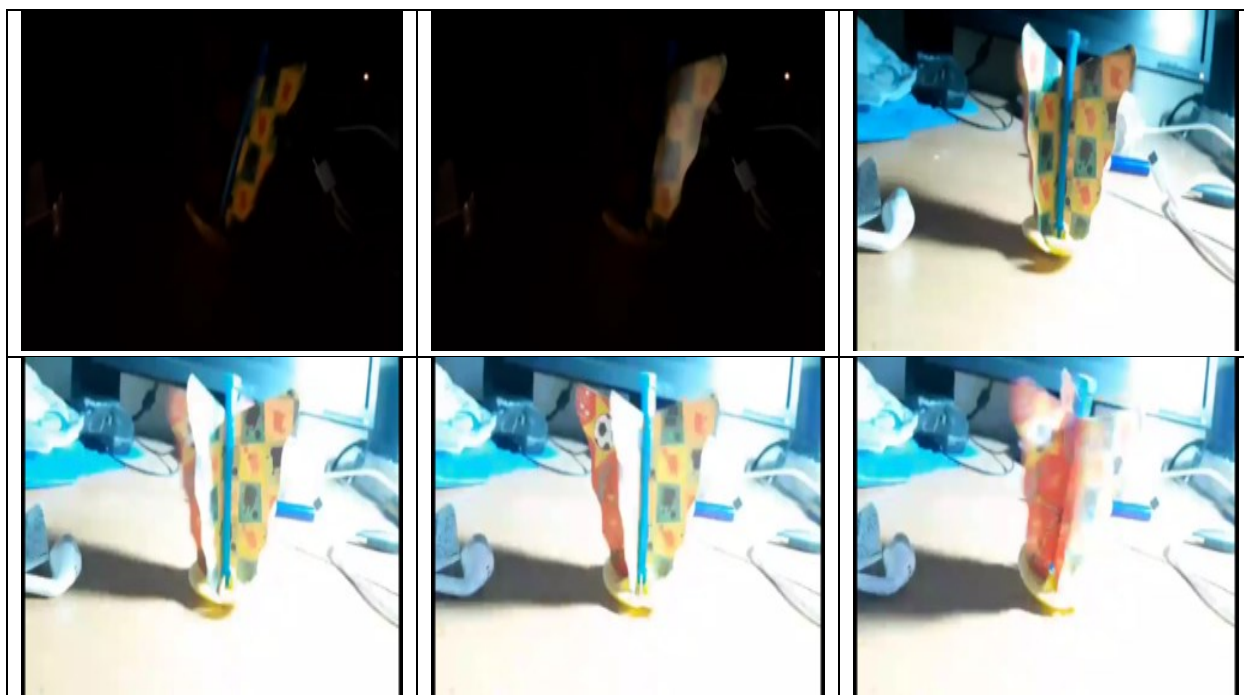
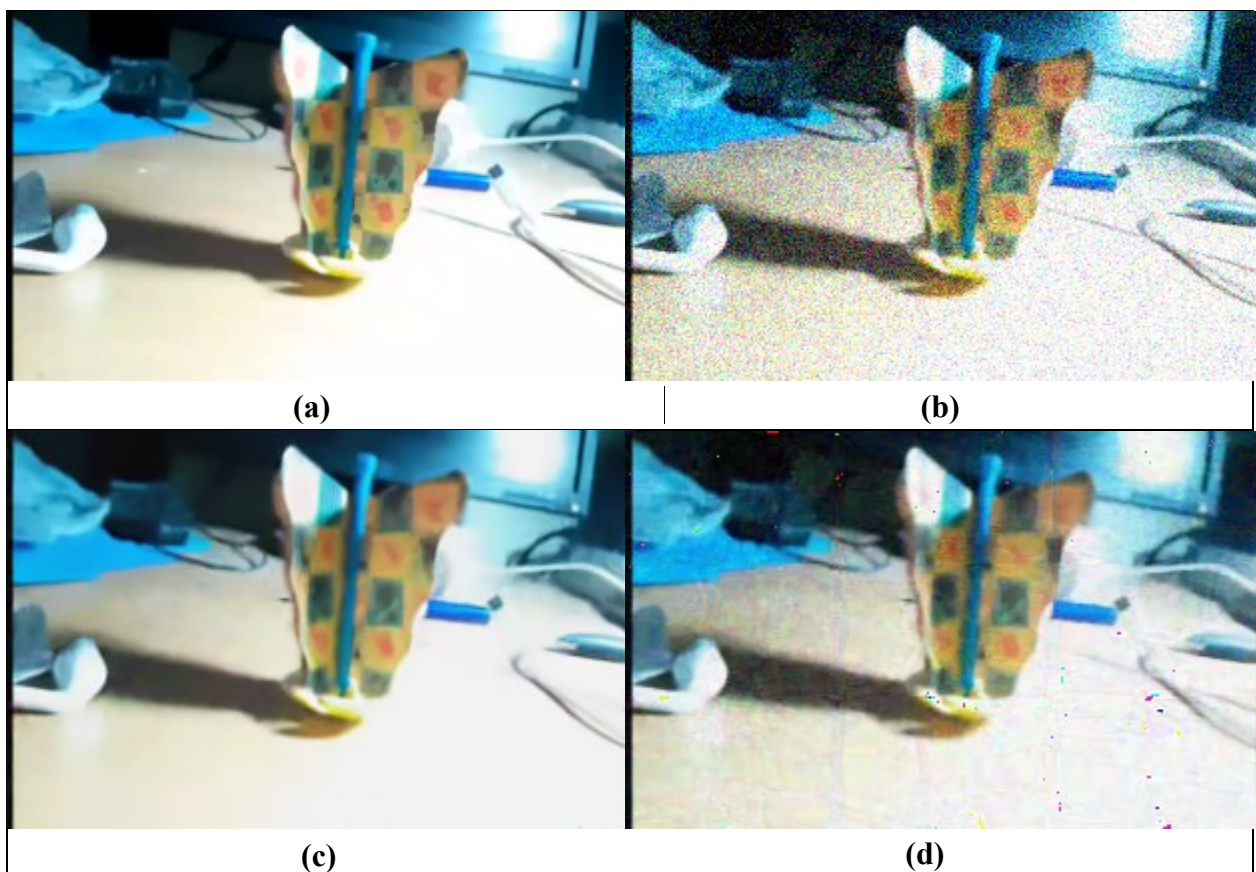


Figure 4.6 les frames de la séquence Bureau qui a une particularité le changement brusque de la Lumière



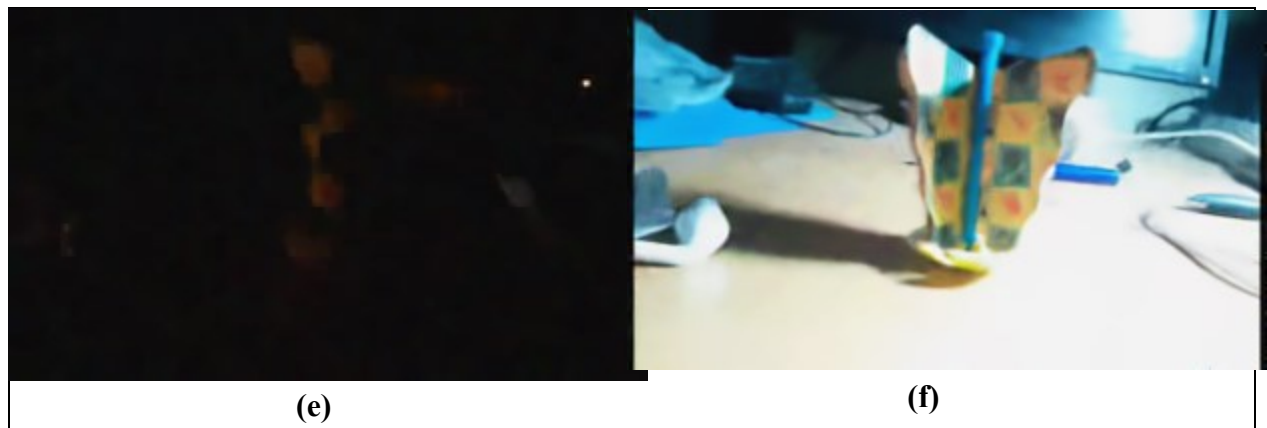


Figure 4.7 une frame de la séquence Bureau débruitée par différentes méthodes. ((a) Image originale, (b) Image bruitée avec $\sigma = 30$, (c) Image débruitée par le SPTWO, (d) Image débruitée par le VBM3D, (e) Image débruitée par le VBM4D, (f) Image débruitée par le Hybrid).

Figure 4.7 montre une frame de la séquence Bureau débruité par les différents algorithmes. Nous constatons que les différents algorithmes ont du mal à débruité la séquence a cause du changement brusque de la lumière qui a affaibli le calcul du flux optique, la puissance de l'algorithme Hybrid est confirmées dans le tableau 8 le choix du meilleure frame lui a permis de garder un SSIM haut et RMSE bas par rapport à ces concurrents.

4.5 Etude Comparative

Pour une meilleure comparaison nous avons également calculé dans les Tableaux (4.1-4.4) le RMSE et le SSIM chaque séquence en suit les moyennes du RMSE et SSIM de chaque algorithme de débruitage, données pour les séquences d'images de niveau de gris tests $\sigma_n = 10, 20$ et 30 .

Image Niveau de Gris

Tableau 4.1. Les valeurs de RMSE en fonction de σ_n

	σ_n	SPTWO	VBM3D	VBM4D	Hybrid
Army 292x194	10	9.98	4.22	6.66	3.50
	20	141.15	6.04	12.16	5.22
	30	47.12	7.39	16.97	6.53
Yosemite 316x252	10	16.56	6.38	6.85	12.13
	20	17.22	9.68	12.47	112.35
	30	117.88	12.38	17.46	20.33

Tableau 4.2 La moyenne de RMSE des 2 séquences NB

		10	20	30
Moyenne	SPTWO	13.27	79.43	82.50
	VBM3D	5.30	7.86	9.88
	VBM4D	6.75	12.31	17.21
	Hybrid	7.81	58.78	13.43

D'après les deux tableaux (4.1-4.2) nous remarquons que les meilleurs résultats de RMSE sont donnés par l'algorithme VBM3D.

Tableau 4.3 les valeurs de SSIM en fonction de σ_n

	σ_n	SPTWO	VBM3D	VBM4D	Hybrid
Army 292x194	10	0.76	0.91	0.78	0.93
	20	0.00	0.85	0.54	0.87
	30	0.25	0.80	0.41	0.82
Yosemite 316x252	10	0.48	0.80	0.74	0.60
	20	0.43	0.67	0.57	0.22
	30	0.02	0.60	0.45	0.53

Tableau 4.4 La moyenne de SSIM des 2 séquences NB

		10	20	30
Moyenne	SPTWO	0.62	0.21	0.13
	VBM3D	0.85	0.76	0.70
	VBM4D	0.76	0.55	0.43
	Hybrid	0.76	0.54	0.67

Le tableau 4.3 donne les valeurs de SSIM de chaque séquence image, le tableau 4.4 donne la moyenne de SSIM de chaque algorithme nous remarquons que les meilleurs résultats aussi donné par l'algorithme VBM3D.

En fin l'algorithme VBM3D donne des meilleurs résultats sur les séquences de niveau de gris.

Les Tableaux (4.5-4.8) le RMSE et le SSIM de chaque séquence en suit les moyennes du RMSE et SSIM de chaque algorithme de débruitage, données pour les séquences d'images en couleur tests $\sigma_n = 10, 20, 30, 40$ et 50 .

*Image couleur*Tableau 5. Les valeurs de RMSE en fonction de σ_n

	σ_n	SPTWO	VBM3D	VBM4D	Hybrid
Army 292x194	10	2.56	7.49	3.27	3.20
	20	4.49	7.72	4.98	4.62
	30	6.11	9.26	6.55	6.18
	40	5.53	12.50	8.11	7.93
	50	9.59	13.70	9.75	9.76
Cooper 640x480	10	4.01	26.02	4.35	4.97
	20	6.72	19.93	7.31	7.66
	30	9.43	17.79	10.06	10.23
	40	12.25	17.69	12.74	12.08
	50	15.24	20.03	15.48	15.33
Dog 640x480	10	3.93	14.75	4.25	4.50
	20	5.90	11.24	6.24	5.92
	30	7.77	11.82	8.05	7.84
	40	9.88	12.94	8.98	9.98
	50	12.22	15.76	12.06	12.35
Truck 640x480	10	3.87	19.73	4.23	4.53
	20	6.34	16.13	6.81	6.95
	30	8.88	16.15	9.26	8.92
	40	11.68	17.73	11.80	11.73
	50	14.73	20.26	14.51	14.75
Basketball 640x480	10	2.89	8.39	3.19	3.23
	20	4.17	7.65	4.69	4.51
	30	5.56	8.75	6.06	5.70
	40	7.06	9.74	7.45	7.06
	50	8.77	11.43	8.95	8.77
Traffic 160x120	10	3.34	6.81	3.16	4.22
	20	4.85	10.52	4.70	5.48
	30	6.32	13.20	4.34	6.72
	40	7.88	14.89	7.49	7.93
	50	9.41	16.53	8.74	9.48
Bureau 320x180	10	3.69	11.19	3.91	3.94
	20	7.18	11.27	7.40	7.25
	30	10.68	13.12	10.90	10.60
	40	14.28	18.88	14.45	14.18
	50	17.96	19.00	18.06	17.96

Tableau 4.6 La moyenne de RMSE des 7 séquences

		10	20	30	40	50
Moyenne	SPTWO	3.56	6.29	8.85	11.66	14.81
	VBM3D	12.52	11.71	12.97	16.56	16.97
	VBM4D	3.82	6.59	9.14	11.93	14.82
	Hybrid	4.0.2	6.55	9.09	11.81	14.85

D'après le tableau 4.5 les valeurs de RMSE de chaque Séquanes avec un niveau de bruit différent.

Le tableau 4.6 donne la moyenne de RMSE de chaque séquence avec un niveau de bruit différent. Nous remarquons que les meilleurs résultats sont donnés par l'algorithme SPTWO en premier rang et l'Hybrid en seconde rang.

Tableau 4.7 Les valeurs de SSIM en fonction de σ_n

		σ_n	SPTWO	VBM3D	VBM4D	Hybride
Army 292x194	10	0.99	0.87	0.98	0.98	
	20	0.93	0.80	0.97	0.97	
	30	0.90	0.74	0.96	0.97	
	40	0.88	0.61	0.95	0.95	
	50	0.88	0.59	0.94	0.94	
Cooper 640x480	10	0.96	0.75	0.97	0.97	
	20	0.95	0.74	0.94	0.94	
	30	0.93	0.71	0.91	0.91	
	40	0.90	0.70	0.87	0.88	
	50	0.88	0.66	0.84	0.84	
Dog 640x480	10	0.93	0.84	0.95	0.95	
	20	0.90	0.78	0.91	0.92	
	30	0.88	0.73	0.88	0.89	
	40	0.85	0.70	0.84	0.85	
	50	0.82	0.66	0.81	0.81	
Truck 640x480	10	0.96	0.80	0.96	0.96	
	20	0.93	0.76	0.92	0.92	
	30	0.91	0.73	0.88	0.89	
	40	0.89	0.70	0.84	0.85	
	50	0.86	0.67	0.80	0.80	

Basketball 640x480	10	0.98	0.95	0.98	0.98
	20	0.97	0.94	0.97	0.97
	30	0.96	0.92	0.96	0.96
	40	0.95	0.91	0.94	0.95
	50	0.94	0.88	0.93	0.94
Traffic 160x120	10	0.96	0.85	0.96	0.96
	20	0.94	0.74	0.92	0.94
	30	0.92	0.64	0.94	0.92
	40	0.89	0.59	0.85	0.90
	50	0.87	0.53	0.81	0.87
Bureau 320x180	10	0.73	0.66	0.73	0.74
	20	0.61	0.56	0.61	0.61
	30	0.56	0.51	0.56	0.56
	40	0.53	0.44	0.53	0.53
	50	0.51	0.46	0.50	0.51

Tableau 4.8 La moyenne de SSIM des 7 séquences

		10	20	30	40	50
Moyenne	SPTWO	0.84	0.77	0.73	0.71	0.69
	VBM3D	0.75	0.67	0.62	0.57	0.56
	VBM4D	0.84	0.77	0.74	0.70	0.67
	Hybrid	0.85	0.77	0.74	0.71	0.68

Le tableau 4.7 donne les valeurs de SSIM de chaque séquence image, le tableau 4.8 donne la moyenne de SSIM de chaque algorithme. Nous constatons que les meilleurs résultats obtenus par l'algorithme Hybrid.

En fin l'algorithme Hybrid donne des meilleurs résultats sur les séquences en couleur quia a moine de détaille en arrière plans.

Les figures 4.8-4.13 facilitent la comparaison des algorithmes

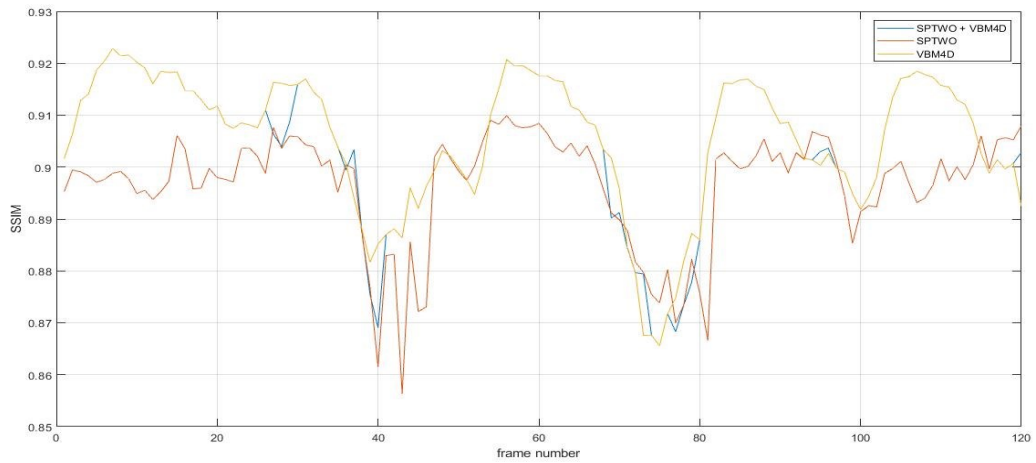


Figure 4.8 SSIM de la séquence "Traffic" débruité par SPTWO, VBM4D, VBM3D, Hybride

D'après la figure 4.8 les quatre algorithmes ont débruité la séquence avec une légère différence de SSIM légère, l'algorithme Hybride emploie l'estimation du bruit ce qui lui a permis d'avoir des résultats similaires ou supérieurs à ceux du VBM4D et SPTWO. Nous remarquons clairement sur le frame 39 l'algorithme hybride dépasse les trois autres algorithmes.

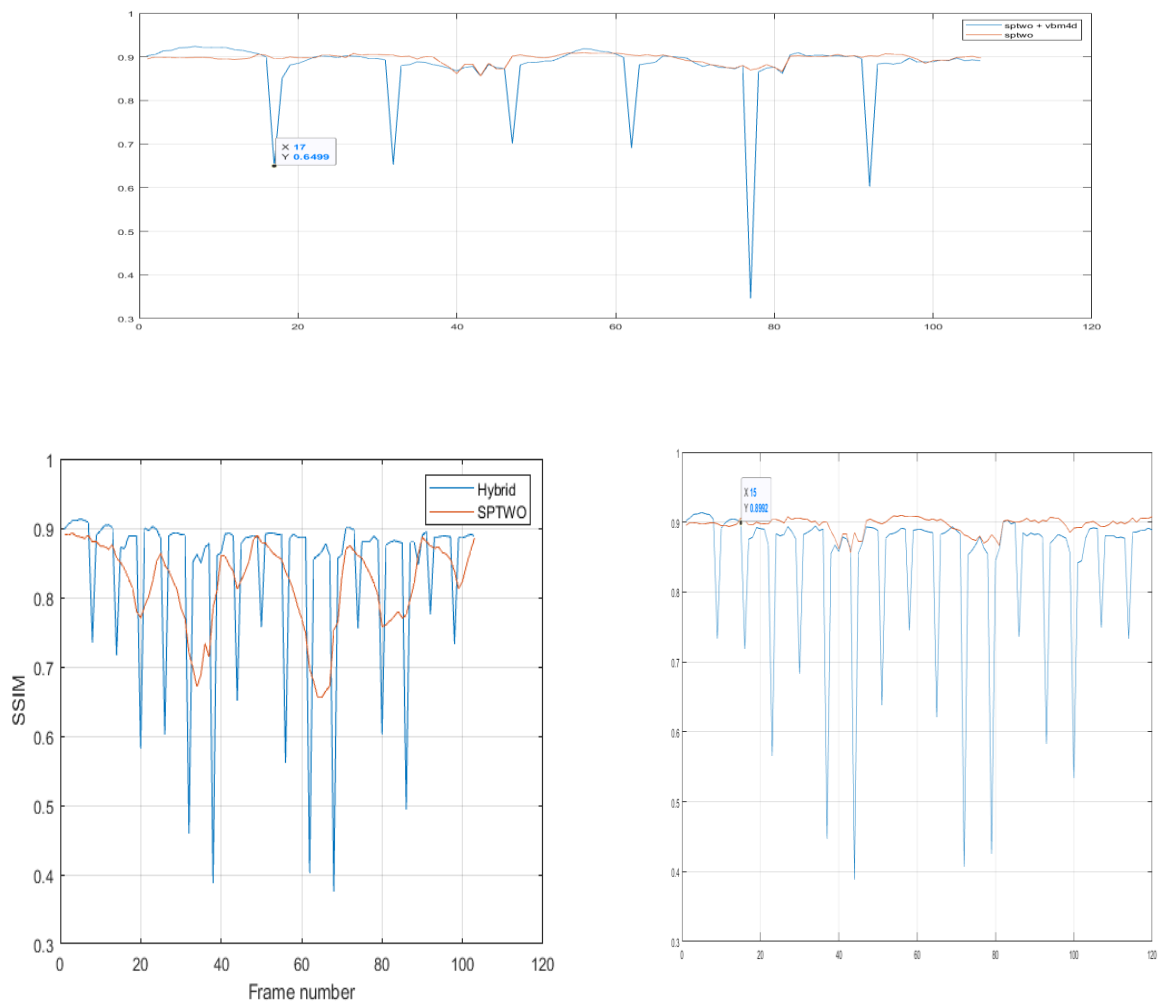


Figure 4.9 continuité du flux optique dans le débruitage des vidéos par estimation du mouvement.

Les trois illustrations de la figure 4.9 montrent l'impact de la discontinuité du calcul du flux optique sur le résultat final du débruitage de l'image, nous remarquons clairement que le SSIM chute lorsque la discontinuité se produit, cela montre l'utilité du regroupement des séquences d'images.

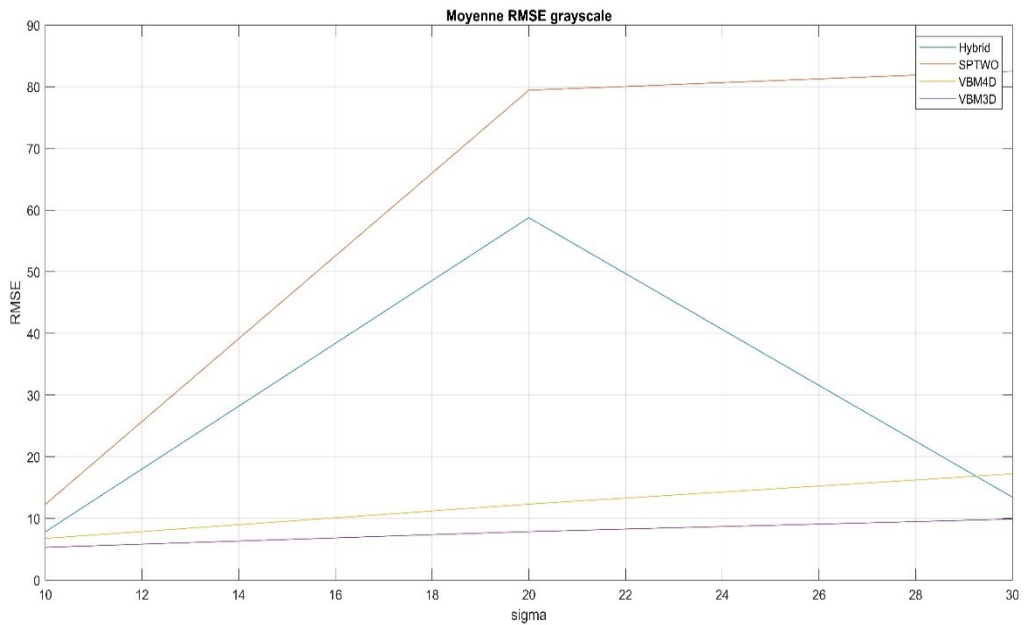


Figure 4.10 la moyenne de RMSE des 2 séquences noir et blanc

D'après la figure 4.10 la courbe en violet (algorithme VBM3D) a les meilleures valeurs de RMSE dans le cas des séquences de niveau de gris

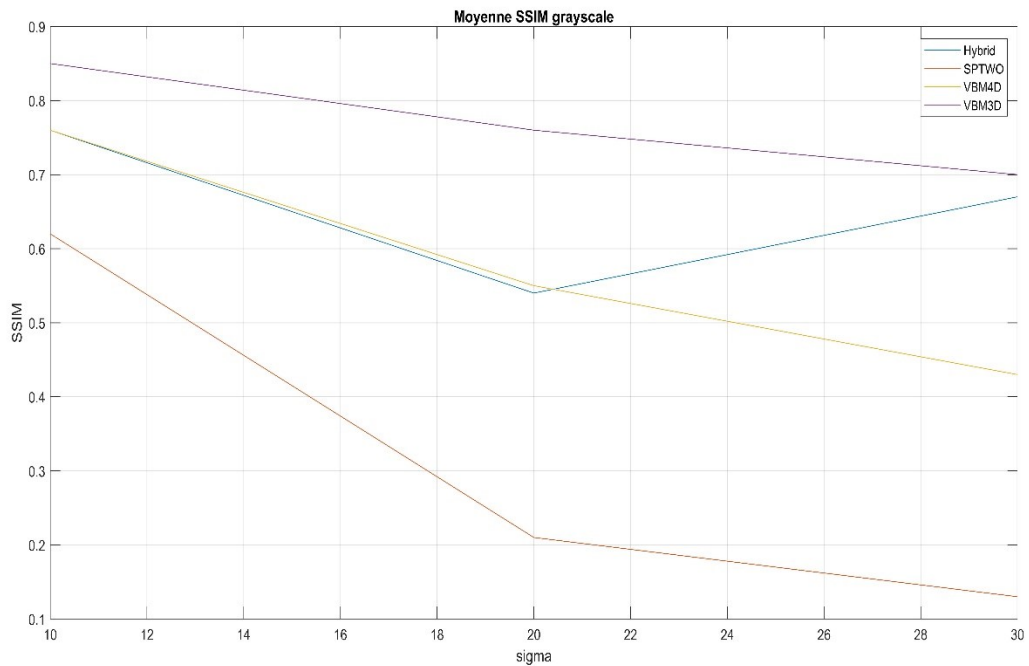


Figure 4.11 la moyenne de SSIM des 2 séquences noir et blanc

D'après la figure 4.11 la courbe en violet (algorithme VBM3D) a les meilleures valeurs de SSIM dans le cas des séquences de niveau de gris

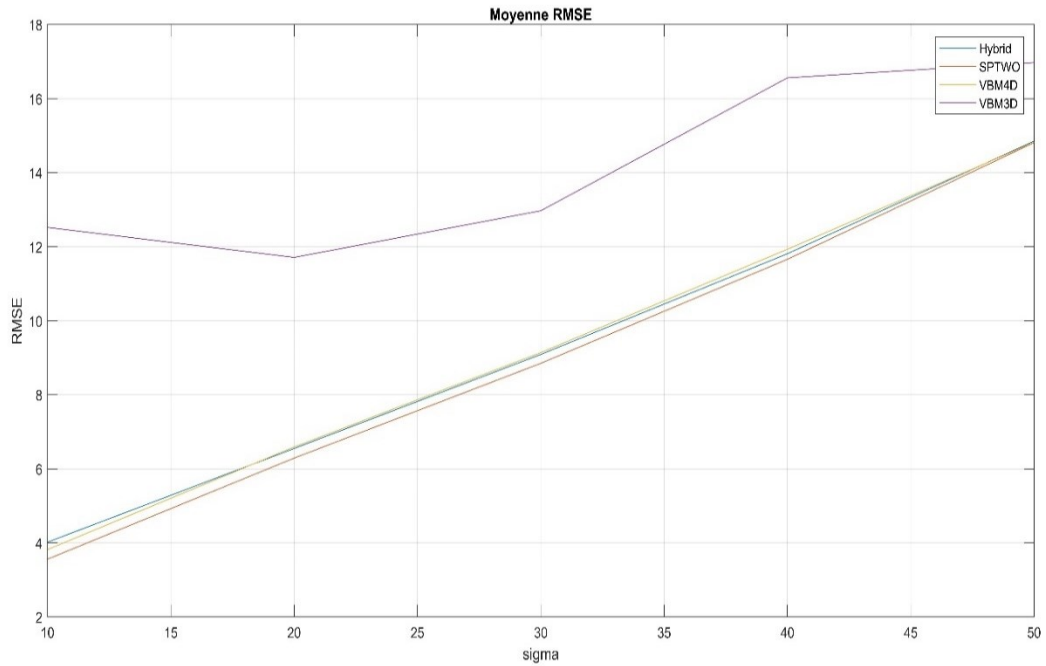


Figure 4.12 la moyenne de RMSE des séquences en couleur

D'après la figure 4.12 les deux courbes (SPTWO et Hybrid) sont presque similaires pour les niveaux de bruit inférieurs à 15 l'algorithme Hybrid donne des meilleurs résultats que le SPTWO, pour les niveaux de bruit entre 45 et 50 les trois algorithmes (SPTWO, Hybrid, VBM4D) sont similaires.

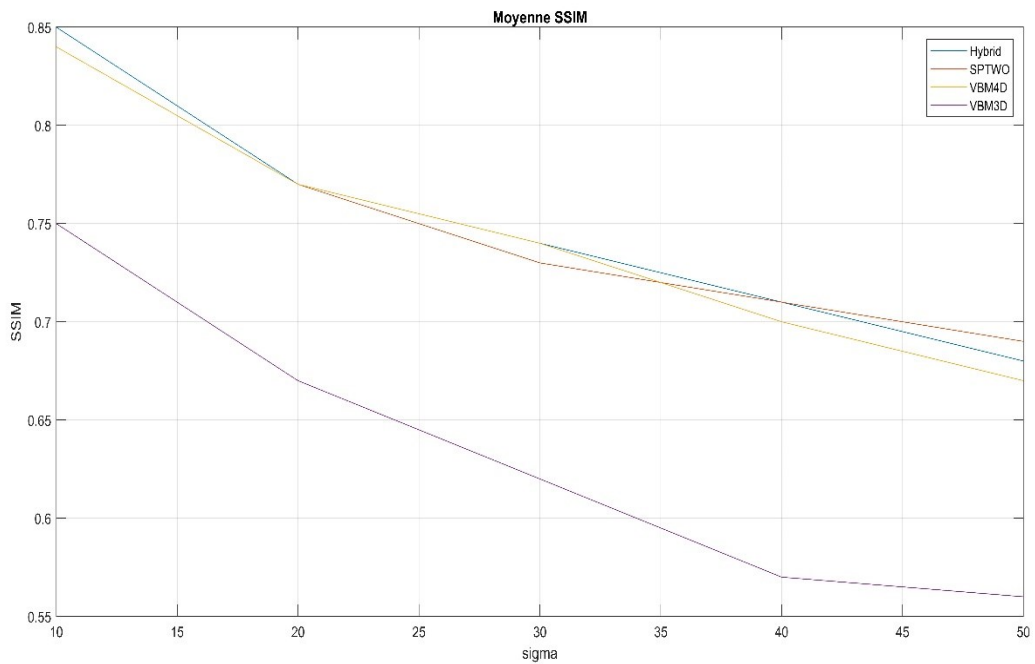


Figure 4.13 la moyenne de SSIM des séquences en couleur

D'après la figure 4.13 la courbe en bleu (algorithme Hybrid) donne des meilleurs résultats dans les niveaux de bruit inférieur à 20, Pour les niveau de bruit entre 20 et 30 les deux courbe (Hybrid et SPTWO) sont similaire, au niveau de bruit entre 30 et 40 l'algorithme Hybrid donne des meilleures résultats que le autre algorithmes, enfin dans les niveaux de bruit entre 40 et 50 le SPTWO donne les meilleurs résultats.

Synthèse

Nous résumons les principales constations de notre étude expérimentale dans ces points :

- Nous remarquons que le filtrage simple ne donne pas des bons résultats.
- L'algorithme VBM3D donne des meilleurs résultats dans le cas des séquences vidéo niveaux de gris.
- L'algorithme Hybrid donne des résultats mieux que SPTWO dans le cas des vidéos du niveau de gris.
- VBM4D diminue l'intensité des images dans les deux cas en couleur niveau de gris.
- Les quatre algorithmes ne conservent pas les détails des images.
- L'algorithme Hybrid reste efficace dans les deux cas en couleur et niveaux de gris parce qu'il tire le profit des avantages des trois algorithmes inspire par les trios algorithmes (SPTWO, VBM3D, VBM4D).

4.6 Conclusion

Dans ce dernier chapitre, un protocole d'évaluation de méthodes de débruitage a été élaboré. Ce protocole est appliqué pour comparer les différentes méthodes décrites dans le troisième Chapitre.

A partir des résultats présentées ci-dessus, le meilleur résultat obtenu est donné par l'algorithme Hybride, ont montré que les performances de nos bruiteurs s'avèrent supérieures à celles des débruiteurs des séquences vidéo, Cette performance reste liée à l'idée de gardés les bonnes performances des trois algorithmes (VBM3D, VBM4D, SPTWO) pour crée un algorithme efficace aux vidéo fortement bruitées.

Conclusion Général

Dans ce document, nous avons étudié le problème du débruitage qui a engendré une importante littérature en Pré-traitement des images (vidéo). D'où l'idée de restaurer une image ou une séquence d'images (vidéo) avec la meilleur qualité possible à partir de sa version dégradée. Nous nous sommes concentrés sur l'estimation du mouvement par le flux optique basé sur la méthode variationnelle dans le domaine spatio-temporel comme solution au problème.

Nous avons d'abord étudié trois algorithmes récents de débruitage appliqués sur des séquences vidéo. Nous avons reproduit les résultats de simulation et établi une comparaison quantitative et qualitative, afin de voir les performances et les limitations de chaque algorithme. L'objectif de cette étude est de proposer une structure hybride qui tire profit des avantages de ces algorithmes. Nous avons donc proposé un algorithme hybride à base de ces trois méthodes.

Les simulations de cette approche sont réalisé à l'aide des logiciels MATLAB, C++ et le python en utilisant des séquences de test composée de huit frames.

L'évaluation exige une phase de comparaison des algorithmes de débruitage (SPTWO, VBM3D, VBM4D) avec notre nouveau algorithme (méthode hybride), en utilisant l'Erreur Quadratique Moyenne et l'indice de Similarité Structurale comme des critères d'évaluations des algorithmes.

L'interprétation de ces résultats montre que le nouveau algorithme a un grand potentielle selon les résultats obtenus. Un SSIM supérieur montre que la méthode Hybrid préserve mieux la qualité visuelle de l'image et donc la séquence vidéo.

En perspective de notre travail, nous proposons une amélioration des performances en terme de temps de calcule dans le but de rendre cette méthode utilisable dans les applications à temps réelle.

Ondelette / Seuillage

1) Transformée en ondelettes

Dès la naissance des ondelettes, elles ont gagné un intérêt considérable pour le traitement des signaux. L'idée est de représenter un signal à plusieurs résolutions, ce qui permet de représenter ses tendances principales dans seulement quelques coefficients. En traitement d'images, les ondelettes ont été employées pour différentes applications telles que le débruitage, la compression et l'estimation de mouvement. Cette section présente les ondelettes dans ses deux versions ; les ondelettes continues unidimensionnelles et les ondelettes discrètes sous les différentes constructions possibles. Puis une extension aux signaux bidimensionnels sera également présentée.

*La transformation en ondelettes

La transformée en ondelettes pour un signal $f(t)$ est défini par l'expression suivante :

$$c_{\Psi}(a, b) = \int_{-\infty}^{+\infty} f(t)\Psi(t)dt \quad (1)$$

Ou
$$\Psi_{a,b}(t) = \frac{1}{\sqrt{a}} \Psi\left(\frac{t-b}{a}\right) \quad (2)$$

"a" et "b" sont respectivement les facteurs de dilatation et de translation.

*Débruitage par ondelettes

La littérature montre que plusieurs techniques de prétraitement sont basées sur l'exploration de la transformation en ondelettes. En effet, les ondelettes offrent une bonne technique pour réduire le bruit dans une image. Les coefficients issus des ondelettes traduisent les discontinuités qui correspondent aux détails de l'image considérée. La sous bande d'approximation représente l'information utile de l'image, tandis que les autres sous bandes représentent les hautes fréquences.

Le bruit est donc concentré dans les sous bandes représentant les détails. Un seuillage des coefficients des sous bandes de détails élimine les éléments les plus fins de l'image, permettant alors la réduction du bruit dans le fond. En fait, il existe plusieurs méthodes de seuillage des coefficients d'ondelettes telles que le seuillage doux (Soft thresholding) et le seuillage dur (Hard thresholding).

2) Différents types de seuillage

Il existe plusieurs types d'estimateurs par seuillage. Nous pouvons d'abord distinguer le seuillage dur et seuillage doux.

- **Seuillage dur ou "hard thresholding"**

C'est le seuillage le plus intuitif. Pour un seuil T préalablement choisi, seuls les coefficients dont l'amplitude (en valeur absolue) est supérieure à ce seuil sont retenus ; les autres sont mis à zéro. Ceci revient à appliquer la fonction suivante :

$$d_m(x) = \rho_T(x) = \begin{cases} x & \text{si } |x| > T \\ 0 & \text{si } |x| \leq T \end{cases} \quad (1)$$

Nous aurons donc le seuillage suivant :

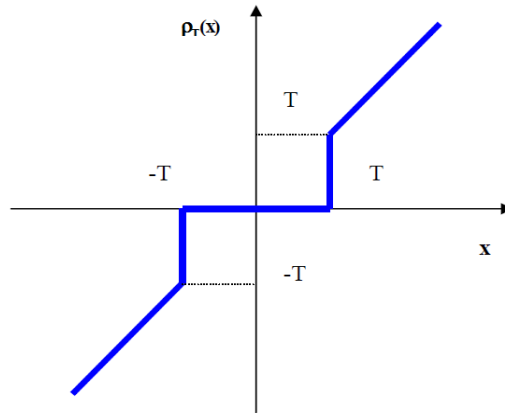


Figure 1 Courbe du seuillage dur

L'opérateur D dans l'équation précédente est alors un projecteur non linéaire dans la base B . Le risque de ce seuillage est :

$$r_t(f) = r(d, f) = \sum_{m=0}^{N-1} E\{|f_B[m] - \rho_T(X_B[m])|^2\} \quad (2)$$

2) Seuillage doux ou "soft thresholding"

Dans le cas du seuillage dur, on met toujours à zéro les coefficients inférieurs à un seuil T . Par contre, pour ceux supérieurs à T , on atténue l'amplitude des coefficients par la valeur du seuil afin de s'assurer d'avoir enlevé l'effet du bruit même pour les forts coefficients

$$d_m(x) = \rho_T(x) = \begin{cases} x - \text{sign}(x)T & \text{si } |x| > T \\ 0 & \text{si } |x| \leq T \end{cases} \quad (3)$$

Dans ce cas, la fonction de seuillage ρ_T est continue :

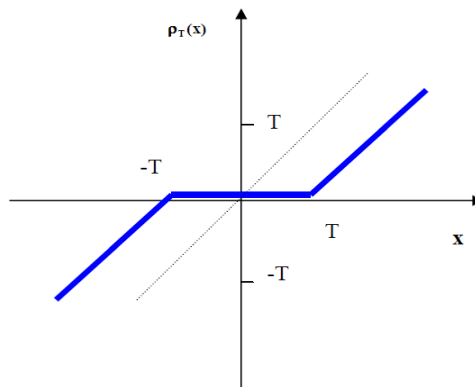


Figure 2 Courbe du seuillage doux

Les deux fonctions de seuillage ci-dessus ont des avantages et des inconvénients ; le seuillage doux n'est pas efficace pour les grands coefficients, le seuillage dur présente des discontinuités.

PCA et SVD

1) Principe (PCA)

L'idée principale de la décomposition orthogonale propre est de trouver un ensemble de vecteurs de base orthonormaux et ordonnés dans un sous espace sans perte conséquente d'information où l'on exprimera de manière optimale les vecteurs aléatoires de données en utilisant les 'l' premiers vecteurs de base obtenus.

C'est une méthode qui souligne les similarités ainsi que les différences des données dans un espace à grandes dimensions où l'on n'a pas le luxe de la représentation graphique en identifiant la dépendance des structures des observations stochastiques multidimensionnelles dans le but d'avoir sa description compacte. Elle est communément considérée comme un outil de visualisation des données, cependant c'est aussi un moyen :

- de décorréler ces données ; dans le nouvel espace, constitué des nouveaux axes.
- de débruiter ces données, en considérant que les axes que l'on oublie sont des axes bruités.

Cette technique peut être vue de deux points de vue : celui de maximiser la variance ou encore celui de minimiser l'erreur quadratique moyenne. Les données représentées dans le nouvel espace sont appelées composantes principales, Elles sont décorrélées et ordonnées dans le sens décroissant de la variance [77].

$$M = \begin{bmatrix} X_{1,1} & \cdots & X_{N,1} \\ \vdots & \ddots & \vdots \\ X_{1,K} & \cdots & X_{N,K} \end{bmatrix} \quad (1)$$

Il est usuel d'appliquer une analyse en composantes principales sur un ensemble de « N » variables aléatoires X_1, \dots, X_N connues de dimension « K ». Ces « N » variables aléatoires peuvent être structurés dans une matrice M à K lignes et N colonnes.

Chaque variable aléatoire $X_n = (X_n, 1, \dots, X_n, K)'$ à une moyenne \bar{X}_n et un écart type σ_{X_n} .

2) Principe de SVD

L'idée essentielle de la SVD est de décomposer la matrice de données en trois matrices simples : deux orthogonales et une diagonale. Du fait qu'elle produise une estimation aux moindres carrés de la matrice de données de même dimension et d'un rang inférieur, elle est équivalente à PCA, et importante autant qu'elle. L'un des avantages de la SVD est son pouvoir de réduction des données après leur blanchissement. En effet, cette technique fournit une description plus compacte des données contenues dans une matrice, exprimée par les premiers modes statistiques. Elle peut être considérée comme une méthode permettant de construire une partition de la variance d'une base de données, c'est à dire qu'elle fournit la base orthogonale qui maxime la variance au sens des moindres carrés. Ceci signifie que le mode « 1 » contient une variance maximale et qu'il contient la structure la mieux corrélée. De manière générale, les premiers modes de la décomposition devraient préférentiellement capter les structures cohérentes et les derniers modes les structures aléatoires du champ, c'est-à-dire sans structure particulière.

La décomposition en valeurs singulières utilise la décomposition en valeur propre d'une matrice semi définie positive obtenue par la multiplication d'une matrice par sa transposé, pour

dériver une décomposition similaire applicable à toutes les matrices rectangulaires composées de nombres réels [77].

2.1 Décomposition en Valeurs Singulières (SVD)

La SVD est une technique mathématique permettant de décomposer une matrice H quelconque en trois matrices dont le produit équivaut à H (Figure 3.1). La première, U est carrée et correspond aux nombres de lignes de H ; la deuxième S est diagonale et a la même taille que H ; La troisième V^T est carrée également et correspond au nombre de colonnes de H . V^T est la transposée de la matrice V . Au sein de S , les valeurs sont classées par ordre décroissant le long de la diagonale, les valeurs les plus grandes correspondant aux signaux les plus intenses, les valeurs les plus faibles correspondant au bruit.

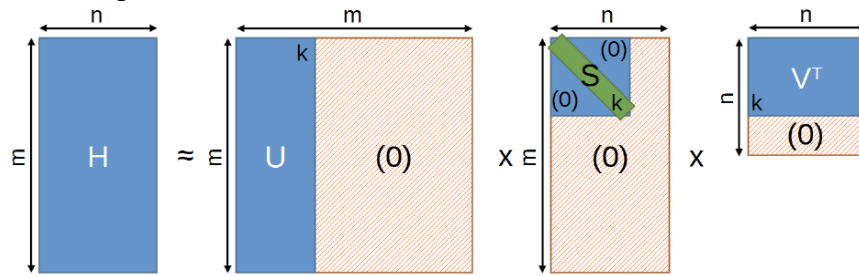


Figure 1 Décomposition en valeurs singulières d'une matrice et approximation de rang réduit

En sélectionnant seulement les k premiers points de la matrice diagonale, on ne conserve que les points correspondant aux signaux d'intérêt, au détriment du bruit. Au contraire, en supprimant les premiers points de la matrice diagonale, il est possible de supprimer un artefact ou un signal de solvant. La matrice d'intérêt est ensuite reconstruite à partir des trois matrices tronquées, ce que l'on nomme l'approximation de rang réduit. Ce procédé est également appelé méthode de Cadzow. Même s'il est beaucoup utilisé dans le domaine de la compression d'image, il reste peu utilisé dans le domaine de la spectroscopie. A noter toutefois que la SVD est une étape clé dans de nombreux traitements statistiques, tels que l'analyse en composante principale (PCA).

2.2 Interprétation Géométrique de la SVD

La SVD peut être interprétée comme une transformation géométrique (une rotation) d'un repère orthogonal dans une nouvelle base ortho normale. Cette rotation géométrique sera précédée généralement, le cas échéant, d'une translation si nous centrons nos données, c'est-à-dire si nous procédons au retrait de la moyenne (suivant le type de données analysées).

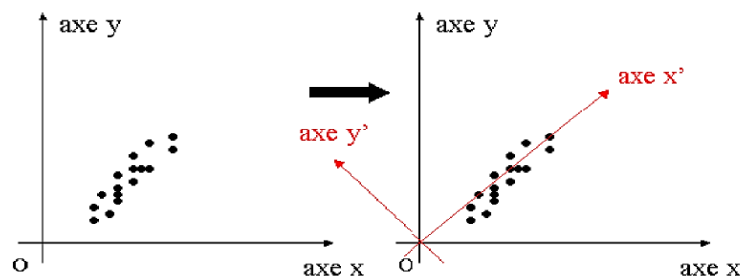


Figure 2 Approche géométrique de la SVD : la SVD est équivalente à un changement de référentiel.

Dans un espace de dimension quelconque (supérieure ou égale à 3), la SVD revient à approximer le nuage de points par un hyperplan.

Bibliographies

Bibliographies

- [1] Humbert Florent, "Techniques de débruitage d'images", 23 janvier 2008
- [2] Antoni Buades, Jose-Luis Lisani Universitat Illes Balears, "Vidéo Denoising with Optical Flow Estimation" Spain toni.buades@uib.es, joseluis.lisani@uib.es, Published in Image Processing On Line, 2018/07/23.
- [3] MEDJAHED Fatiha, "Détection et Suivi d'Objets en Mouvement Dans Une Séquence d'Images", Université des Sciences et de la Technologie d'Oran U. S. T. O., 2012
- [4] DALI Amel, "Estimation de mouvement par les méthodes fréquentielles De flux optique", Université Larbi Ben Mhidi Oum EL Bouaghi, 2013
- [5] Raphaël Isdant , "Traitement numérique de l'image", 2009
- [6] F. Longuetaud (INRA LERFoB) et de E.Ternisien et de B. Kerautret (IUT St-Dié) et de S. Dubuisson (Jussieu), "Cours d'introduction à l'analyse d'images", Sur la base d'un cours de Cours de Master FAGE UE Biophysique des tissus végétaux Mardi 3 mars 2014
- [7] Christophe.alleau@ac-poitiers.fr
- [8] Serge WACKER – C2I niveau 1, <http://www.montpellier.iufm.fr/technoprimaire/>
- [9] Nicolas DUMOULIN, "Compression de séquences vidéo choix des images de références " Université de Rennes1-IFSI, Juin 2003.
- [10] Ahmed Ben Atitallah, "Etude et Implantation d'Algorithmes de Compression d'Images
- [11] <http://gillesboulet.ca/textes/VideoNumerique.pdf>.
- [12] HACHEMI Fouzia, "La détection et suivi des objets en mouvement dans une scène vidéo en utilisant la bibliothèque OpenCV", Université Abou BakrBelkaid– Tlemcen, 2015-2016
- [13] BAKHOUCHE Somaia, BOUALLEG Amine, "Estimation De Mouvement Par La Méthode Fréquentielle", UNIVERSITE LARBI TEBESSI – TEBESSA, 2015 / 2016.
- [14] Chérif TAOUCHE, "Implémentation d'un Environnement Parallèle pour la Compression d'Images à l'aide des Fractales", Université Mentouri Constantine, 2005.
- [15] N. Laouar, M. S. Laraba, Détection d'un mouvement dans une séquence vidéo par filtres morphologiques, Projet de fin d'étude, ENP, Alger, 2009.
- [16] Lingrand, D. "Cours de traitement d'images". Cours non publié, 2004.
- [17] Djenabou Barry, "Débruitage d'image par fusion de filtrage spatio-fréquentielle", Université de Montréal, Mars, 2018.
- [18] Moussa Embarek, R. (2012). Noisy digital image restoration. Mémoire de fin d'étude de master. Université Larbi Ben M'hidi. Oum El Bouaghi, Algérie

Bibliographies

- [19] Gérard Laurent • Daniel Mathiot • Alain Jeanroy, Techniques audiovisuelles et multimédias Tome 2 – Systèmes micro-informatiques et réseaux, diffusion, distribution, réception, 3e édition © Dunod, Paris, 2000, 2007, 2012, ISBN 978-2-10-057600-5
- [20] AHMED SEGHIR Zianou, "Evaluation de la qualité d'image", Université de Mentouri – Constantine, 2012.
- [21] Z. Wang A.C. Bovik, H. R. Sheikh, and E .P. Simocelli, "Image quality assessment: From error measurement to structural similarity," IEEE Trans. Image Processing, vol. 13, no.4, pp.600-612, Apr .2004.
- [22] a. A. C. B. Zhou Wang, "A Universal Image Quality Index," IEEE SIGNAL PROCESSING LETTERS, vol. 9, pp. 81-84, 2002.
- [23] Daorolala Alexandre RANDRIANTSOA, "ESTIMATION DU MOUVEMENT FONDÉE SUR UN SCHÉMA DIRECT ET RÉTROGRADE APPLICATION À LA SEGMENTATION", L'UNIVERSITÉ BORDEAUX I, 2002
- [24] Odobez J. M., Bouthemy P., Fleuet E., "Suivi 2D des pièces métalliques en vue d'un asservissement visuel", RFIA'98, Vol 2, pp. 173-192, Clermont-Ferrand, France, Janvier 1998.
- [25] S. Ambellouis, "Analyse du mouvement dans les séquences d'images par une méthode récursive de filtrage spatio-temporel sélectif ", Thèse de doctorat, Université Pierre et Marie Curie, 2011.
- [26] Frédéric Huguet, "Modélisation et calcul du flot de scène stéréoscopique par une méthode variationnelle", Modélisation et simulation. Université Joseph-Fourier - Grenoble I, 2009. Français. <tel-00421958> <https://tel.archives-ouvertes.fr/tel-00421958>
- [27] Master 1 SIA-AMS Année Universitaire 2017-2018
- [28] B. K. P. Horn, B. G. Shunck, "Determining optical flow", Artificial intelligence, 1981, Vol. 17, p. 185-203.
- [29] Analyse de la vidéo Dernière revu, 13 janvier 2015
- [30] M. R. Fezzani, "Approche parallèle pour l'estimation du flot optique par méthode variationnelle", thèse de doctorat, Université de Lille 1, 2000.
- [31] H. H. Nagel, —Displacement vectors derived from second-order intensity variations in image sequences. CVGIP, 1986. 21(1): p. 85-117.
- [32] B. K. P. Horn, B. G. Schunck, "Determining optical flow", Artificial Intelligence, vol. 17, pp.185–204, 1981.
- [33] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," In Proceedings of the 7th International Joint Conference on Artificial Intelligence, pp. 674–679, 1981.
- [34] A. Bruhn, J. Weickert, "Lucas/Kanade Meets Horn/Schunck: Combining Local and Global Optic Flow Method", IJVC, vol. 61, no. 3, pp. 211-231.2005.

Bibliographies

- [35] Q. Tang, B. Chiang, A. Akinyemi ; A. Zamyatin, B. Shi, S. Nakanishi , “ A combined local and global motion estimation and compensation method for cardiac CT ”, Proc. SPIE 9033, Medical Imaging 2014: Physics of Medical Imaging, 903304 ,March 2014
- [36] M. ALOUACHE Djamel, Analyse multi résolution pour le traitement d'images sphériques, UNIVERSITE MOULOUD MAMMERI, TIZI – OUZOU,2015.
- [37] B. H. Friemel, L. N. Bohs, B. A. McDermott and G. E Trahey, “initial results with the real-time sad vector velocity estimation system: constant velocity calibration”. In Ultrasonics Symposium, 1992. pp. 1179-1982.
- [38] Anderson, M. E., Multi-dimensional velocity estimation with ultrasound using spatial quadrature, IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, volume 45, no. 3, (1998), pp. 852–861.
- [39] Giachetti, A., Matching techniques to compute image motion, Image and Vision Computing, volume 18, (2000), pp. 247–260.
- [40] D.J. Heeger, —Optical Flow using Spatiotemporal Filters, Int. Journal of Computer Vision, 1987, p. 279—302,
- [41] Matteo Maggioni, Giacomo Boracchi, Alessandro Foi, Karen Egiazarian "Video Denoising, Deblocking and Enhancement Through Separable 4-D Nonlocal Spatiotemporal Transforms", 2012, <https://www.researchgate.net/publication/225053819>
- [42] Wiem FOURATI et Mohamed Salim BOUHLEL, "Techniques de Débruitage d'Images" L'Unité Sciences Et Technologies de l'Image et des Télécommunications SETIT, ISBS Sfax-TUNISIE wiem_fourati@yahoo.fr medsalim.bouhlel@enis.rnu.tn
- [43] Stéphane Bres, Jean-Michel Jolion et Frank Lebourgeois. Traitement et analyse des images numériques. Hermes Science Publications, 2003.
- [44] D.L. Donoho. De-noising by soft-thresholding. IEEE Transactions on Information Theory, 41(3) : 613–627, 1995.
- [45] Xi. Liu, M. Tanaka, and M. Okutomi. Single-image noise level estimation for blind denoising. IEEE Transactions on Image Processing, 22(12) : 5226–5237, 2013.
- [46] A. Foi, M. Trimeche, V. Katkovnik, and K. Egiazarian. Practical poissonian-gaussian noise modeling and fitting for single-image raw-data. IEEE Transactions on Image Processing, 17(10) : 1737–1754, 2008.
- [47] L. Azzari and A. Foi. Indirect estimation of signal-dependent noise with nonadaptive heterogeneous samples. IEEE Transactions on Image Processing, 23(8) : 34–59, 2014.
- [48] camille.sutour@u-bordeaux.fr , cdeledal@u-bordeaux.fr , jaujol@u-bordeaux.fr "Estimation du niveau de bruit par la détection non paramétrique de zones homogènes", I Institut mathématique de Bordeaux, CNRS, UMR 525, 2Laboratoire Bordelais de recherche en Informatique Université de Bordeaux, 351 cours de la Libération, 33045 Talence, France
- [49] <https://support.minitab.com/fr-fr/minitab/18/help-and-how-to/probability-distributions-and-random-data/supporting-topics/basics/using-the-cumulative->

Bibliographies

[distributionfunctioncdf/&ved2ahUK](#)EwijirO6jdriAhV5DGMBHTiuCD8QFjAbegQIARAB&usg=AOvVaw0mHiwTiI9vZgpLUHtI6UCN

[50] M. Mahmoud GHONIEM, "Régularisation spatio-temporelle : applications au débruitage, à l'inpainting et à la classification", UNIVERSITÉ de CAEN/BASSE-NORMANDIE U.F.R. Sciences ÉCOLE DOCTORALE SIMEM ,2010

[51] Halary,M. (2009). Etude des symbioses de mytilidés des écosystèmes marins profonds à base chimio synthétique par des techniques de FISH, de microscopie et de traitement d'images. Thèse de Doctorat. Université Pierre et Marie Curie.

[52] Karouche, M. (2013). Développement d'un logiciel didactique pour traitement d'images numériques. Mémoire de fin d'étude de Master. Université Larbi Ben M'hidi.Oum El Bouaghi, Algérie.

[53] Assas, O. (2013). Classification floue des images. Thèse de doctorat. Université de Batna, Algérie.

[54] Lanani, A. (2012). Analyse et Optimisation d'Artefact dans une Image. Mémoire de Magister en Electronique. Université de Batna, Algérie

[55] PROTTER, M. et ELAD, M. (2009). Image sequence denoising via sparse and redundant representations. *Trans. Img. Proc.*, 18(1):27–35.

[56] BUADES, A., COLL, B. et MOREL, J. (2005a). Denoising image sequences does not require motion estimation. In *Advanced Video and Signal Based Surveillance, 2005. AVSS 2005. IEEE Conference on*, pages 70–74.

[57] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3D transform-domain, collaborative filtering," *IEEE Trans. Image Process.* 16, pp. 2080-2095, August 2007.

[58] ZLOKOLICA, V., PIZURICA, A. et PHILIPS, W. (2006). Waveletdomain video denoising based on reliability measures. *CirSysVideo*, 16(8):993–1007.

[59] BOULANGER,J., KERVRANN, C. et BOUTHEMY, P. (2007). Spacetime adaptation for patch-based image sequence restoration. *IEEE Trans. PAMI*, 29: 1096–1102.

[60] BUADES, A., COLL, B. et MOREL, J. M. (2005b). A review of image denoising algorithms, with a new one. *Multiscale Modeling & Simulation*, 4(2):490–530, 2005.

[61] PROTTER, M. et ELAD, M. (2007). Sparse and redundant representations and motion-estimation-free algorithm for video denoising. In *Wavelets XII. Proceedings of the SPIE, Volume 6701*, volume 6701, page 43.

[62] KINDERMANN, S., OSHER, S. et W., J. P. (2005). Deblurring and denoising of images by nonlocal functionals. *Multiscale Modeling & Simulation*, pages 1091–1115.

[63] GILBOA, G. et OSHER, S. (2007). Nonlocal linear image regularization and supervised segmentation. *Multiscale Modeling & Simulation*, 6:595–630.

Bibliographies

- [64] ELMOATAZ, A., LEZORAY, O., BOUGLEUX, S. et TA, V. T. (2008). Unifying local and nonlocal processing with partial difference operators on weighted graphs. In International Workshop on Local and Non-Local Approximation in Image Processing, pages 11–26, Switzerland.
- [65] junsheng Fu, "Real-time Rate-distortion Oriented joint Video Denoising and Compression ", signal processing department, Tampere university of Techenology.
- [66] Matteo Maggioni, Vladimir Katkovnik, Karen Egiazarian, Alessandro Foi, "A Nonlocal Transform-Domain Filter for Volumetric Data Denoising and Reconstruction".
- [67] R. Megret and D. Dementhon, "A survey of spatio-temporal grouping techniques," Tech. Rep., 2002.
- [68] A. Basharat, Y. Zhai, and M. Shah, "Content based video matching using spatiotemporal volumes," Comput. Vis. Image Underst., vol. 110, no. 3, pp. 360–377, 2008.
- [69] Youssef Chahir, "Contributions à la caractérisation structuration d'images et de séquences vidéo", HAL Id: tel-01083958 Submitted on 18 Nov 2014, <https://hal.archives-ouvertes.fr/tel-01083958>
- [70] Javier Sanchez¹, Enric Meinhardt-Llopis², Gabriele Facciolo³, "TV-L1 Optical Flow Estimation", ¹ CTIM, Universidad de Las Palmas de Gran Canaria, Spain (jsanchez@dis.ulpgc.es), ² CMLA, ENS Cachan, France (enric.meinhardt@cmla.ens-cachan.fr), ³ CMLA, ENS Cachan, France (gabriele.facciolo@cmla.ens-cachan.fr)
- [71] J.L.LUMLEY, Stochastic tools in turbulence. New York: Academic Press, 1970.
- [72] Bounneche Meriem Dorsaf, UNIVERSITE MENTOURI CONSTANTINE, 2009, "Reduction de donnés pour le traitement d'image "
- [73] V.C.KLEMA et A.J.LAUB, The singular value decomposition: its computation and some applications, IEEE Transactions on Automatic Control 25, 164-176, 1980.
- [74] <http://vision.middlebury.edu/flow/data>