

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITE MOULOUD MAMMARI DE TIZI-OUZOU



FACULTE DU GENIE ELECTRIQUE ET D'INFORMATIQUE
DEPARTEMENT D'AUTOMATIQUE

Mémoire de Fin d'Etudes de MASTER PROFESSIONNEL

Domaine : Sciences et Technologies

Filière : Génie électrique

Spécialité : **Automatique et informatique
industrielles**

Présenté par

Samira FLISSI

Lynda FETTOUHI

Thème

Fusion d'images de la Kinect par la transformée en ondelettes

Mémoire soutenu publiquement le 25 / 09 / 2016 devant le jury composé de :

M^r Ahmed DIRAMI

M.C.B, à l'UMMTO, Président

M^r Kamal HAMMOUCHE

Professeur, à l'UMMTO, Encadreur

M^{me} Ouiza ADJEMOUT

M.A.A, à l'UMMTO, Co-Encadreur

M^{lle} Ouardia CHILALI

M.A.A, à l'UMMTO, Examineur

M^{me} Zohra HALICHE

M.A.A, à l'UMMTO, Examineur

Remerciements

*Nous remercions notre DIEU qui nous a aidé à faire
ce modeste travail.*

*Nous tenons à remercier vivement notre dévoué promoteur
M^r K.HAMMOUCHE pour son encadrement, pour sa contribution à
travers sa disponibilité, son temps, son expérience et ses conseils
qu'il nous a prodigué qui nous ont beaucoup éclairé et à travers
lesquels nous avons apprécié l'utilité de notre mémoire*

*Nous remercions également notre co-promotrice M^{me}
O.ADJEMOUT pour tous les efforts qu'elle a bien voulu nous
consacrer durant toute la phase de préparation de ce modeste travail,
pour sa disponibilité, sa serviabilité.*

*Nous remercions vivement les membres de jury qui nous feront
l'honneur d'évaluer notre travail.*

*Nos remerciements les plus chaleureux s'adressent à tous (tes) nos
amis (es) et tous ceux qui ont contribué de près et de loin à la
réalisation de ce travail.*

*Nous remercions tous ceux qui croient à la science et au
Savoir.*

Dédicaces

Je dédie ce modeste travail à :

- ❖ *Mes très chers parents pour leur aide et leur soutien tout au long de mes études, et qui ont fait de moi ce que je suis aujourd'hui et j'espère qu'un jour je serai capable de leur donner au moins le minimum car quoi qu'on face on arrivera jamais à leurs rendre tout.*
- ❖ *Mes grands parents et à la mémoire de mon grand-père paternel*
- ❖ *Mes chers frères MOUMOUH, AHMED, SAID, MAKHLOUF et sa femme FAIZA et leurs filles AYA et ASSIL*
- ❖ *Mes très chères sœurs : FAZIA, LILA, DJAMILA ainsi que leurs époux FARID, HAMID, RABAH et leurs enfants FERJEL, ANIA, IMAD et la petite NILYA*
- ❖ *Mon futur mari SAID qui m'a soutenu et encouragé tout au long de ce travail*
- ❖ *Ma belle famille*
- ❖ *Mes oncles et tantes et leurs familles et à la mémoire de mon grand oncle MOHAMED*
- ❖ *Mes chères amies SIHAM, CHABHA, SARA et mon binôme LYNDIA ainsi que sa famille*
- ❖ *Toute la promotion 2015/2016*

SAMIRA

Je dédie ce modeste travail à :

- ❖ *Mes très chers parents pour leur aide et leur soutien tout au long de mes études, et qui ont fait de moi ce que je suis aujourd'hui et j'espère qu'un jour je serai capable de leur donner au moins le minimum car quoiqu'on face on arrivera jamais à leurs rendre tout.*
- ❖ *Mes grand parents et à la mémoire de mon grand-père paternel*
- ❖ *Mes chers frères : DJAHID, SAID, RACHID, MOUKRAN, OMAR, ABDENNOUR.*
- ❖ *Ma très chère sœur : KAHINA*
- ❖ *Mes oncles et tantes et leurs familles*
- ❖ *Mes chères amies SABRINA, HABIBA et tous mes amis(es) de près et de loin et mon binôme SAMIRA ainsi que sa famille*
- ❖ *Toute la promotion 2015/2016*

LYNDA

Table des matières

Table des matières

| | |
|----------------------------|---|
| Tables des figures..... | 2 |
| Introduction générale..... | 5 |

Chapitre 1 : Présentation de la Kinect

| | |
|--|----|
| 1.1 Introduction..... | 8 |
| 1.2 Présentation de la Kinect..... | 8 |
| 1.3 Fonctionnement de la Kinect..... | 10 |
| 1.3.1 Caméra couleur | 10 |
| 1.3.2 Caméra de profondeur..... | 12 |
| 1.3.2.1 Technique de triangulation ou de lumière structurée..... | 13 |
| 1.3.2.2 Capteur infrarouge..... | 15 |
| 1.4 Outils logiciel pour utiliser la Kinect..... | 16 |
| 1.4 Avantages et limites de la Kinect..... | 17 |
| 1.4.1 Avantages..... | 17 |
| 1.4.2 Limitations..... | 18 |
| 1.5 Applications de la Kinect..... | 21 |
| 1.6 Conclusion..... | 21 |

Chapitre 2 : Fusions d'images

| | |
|--|----|
| 2.1 Introduction..... | 23 |
| 2.2 Fusion d'images..... | 23 |
| 2.3 Niveaux de la fusion d'images..... | 24 |
| 2.4 Les méthode de la fusion d'image..... | 25 |
| 2.5 Fusion d'images basée sur la transformée en ondelettes | 25 |
| 2.5.1 Transformée en ondelettes..... | 25 |
| 2.5.1.1 L a transformé en ondelettes est monodimensionnelle..... | 25 |
| 2.5.1.2 Transformée en ondelettes discrète..... | 27 |
| 2.5.1.3 Transformée en ondelettes bidimensionnelle..... | 28 |
| 2.5.2 Types d'ondelettes..... | 31 |
| 2.6 Domaine d'application..... | 31 |
| 2.6.1 Imagerie aérienne et satellitaire..... | 31 |
| 2.6.2 Imagerie médicale..... | 32 |
| 2.6.3 Imagerie industrielle..... | 33 |

Table des matières

| | |
|---|----|
| 2.7 Conclusion..... | 34 |
| | |
| Chapitre 3 : Fusion d'images de la Kinect à base de la transformée en ondelettes | |
| 3.1 Introduction..... | 36 |
| 3.2 Fusion d'images de la Kinect par la transformée en ondelettes..... | 36 |
| 3.3 Résultats de la Fusion d'images de la Kinect par la transformée en ondelette..... | 40 |
| 3.3.1 Critères d'évaluation des performances de la fusion..... | 40 |
| 3.3.2 Attributs de texture..... | 42 |
| 3.4 Tests et Résultats..... | 44 |
| 3.5 Conclusion..... | 50 |
| Conclusion générale..... | 52 |
| Bibliographie..... | 55 |

Table des figures

Liste des figures et tableaux

| | |
|---|----|
| Figure (1.1) : Composants de la Kinect..... | 8 |
| Figure (1.2) : Champ de vision de la Kinect..... | 9 |
| Figure (1.3): Portée du capteur Kinect..... | 10 |
| Figure (1.4) : Illustration du processus du système RVB..... | 10 |
| Figure (1.5): Capteur CMOS recouvert d'une grille photosensible de Bayer..... | 11 |
| Figure (1.6) : Surface d'un capteur CMOS. | 12 |
| Figure (1.7) : Schéma de la déformation de motif..... | 13 |
| Figure (1.8) : Principe de calcul de la disparité..... | 14 |
| Figure (1.9) : Capteur CMOS infrarouge..... | 16 |
| Figure (1.10) : Exemple d'images acquises par la Kinect..... | 17 |
| Figure (1.11) : Distance minimale de détection..... | 18 |
| Figure (1.12) : Bruit sur les contours des objets..... | 19 |
| Figure (1.13) : Ombrages et occlusions..... | 19 |
| Figure (1.14) : Réflections sur les surfaces lisses..... | 20 |
| Figure (1.15) : Désalignement entre l'image couleur et l'image de profondeur..... | 20 |
| Figure (1.16) : Décalage entre la caméra RGB et la caméra de profondeur..... | 20 |
| | |
| Figure (2.1) : Différents niveaux de fusion d'images..... | 24 |
| Figure (2.2) : Exemple d'une ondelette mère contractée et dilatée..... | 26 |
| Figure (2.3) : Exemple de transformée en ondelettes continu..... | 27 |
| Figure (2.4) : Algorithme de décomposition de Mallat..... | 30 |
| Figure (2.5) : Décomposition pyramidale d'une image..... | 30 |
| Figure (2.6) : Exemple de fusion d'images satellitaire et radar..... | 32 |
| Figure (2.7) : Fusion d'une image PET et une image CT en médecine..... | 33 |
| | |
| Figure (3.1) : Fusion des coefficients d'image de la Kinect par la T.O..... | 36 |
| Figure (3.2) : Fusion des coefficients d'image de la Kinect par la T.O..... | 37 |
| Figure (3.3) : Fusion de la moyenne des coefficients d'image de la Kinect par la T.O..... | 37 |
| Figure (3.4) : Fusion du minimum des coefficients d'image de la Kinect par la T.O..... | 38 |
| Figure (3.5) : Fusion du maximum des coefficients d'image de la Kinect par la T.O..... | 38 |
| Figure (3.6) : Fusion du max-min des coefficients d'image de la Kinect par la T.O..... | 39 |
| Figure (3.7) : Fusion du min-max des coefficients d'image de la Kinect par la T.O..... | 39 |
| Figure (3.8): Plus proches voisins du pixel 'x' selon 4 directions..... | 43 |

Liste des figures et tableaux

| | |
|--|----|
| Figure (3.9) : Images tests..... | 44 |
| Figure (3.10): Décomposition de l'image couleur en niveau de gris..... | 45 |
| Figure (3.11): Décomposition de l'image de profondeur..... | 45 |
| Figure (3.12) : Images fusionnées de la 1 ^{er} image test..... | 46 |
| Figure (3.13) : Images fusionnées de la 2 ^{ème} image test..... | 47 |

Liste des tableaux :

| | |
|---|----|
| Tableau (3.1) : Critères d'évaluation obtenus par les différentes méthodes de fusion sur les images de l'exemple 1..... | 48 |
| Tableau (3.2) : Critères d'évaluation obtenus par les différentes méthodes de fusion sur les images de l'exemple 2..... | 48 |
| Tableau (3.3) : Attributs de textures obtenues sur les images de l'exemple 1..... | 49 |
| Tableau (3.4) : Attributs de textures obtenues sur les images de l'exemple 2..... | 49 |

Résumé : L'objectif principal de la fusion d'image basé sur la transformée en ondelettes est la combinaison entre deux image prises par la caméra Kinect (image couleur convertie en niveau de gris et l'image de profondeur). Dans ce contexte, le but de ce travail est l'utilisation des techniques de la fusion d'image de différentes sources pour obtenir une nouvelle image plus riche en information que l'image initiale, dont on a fait une comparaison sur les sept différentes stratégies de fusion, pour connaître la meilleure méthode qui nous donne des images fusionnées proches aux images originales.

Mots clés: Kinect, fusion d'images, la transformée en ondelettes, attributs de textures, les critères de performances

Introduction générale

Avec l'évolution des technologies, les appareils d'imagerie se diversifient et deviennent de plus en plus abordable. De ce fait, on assiste à un intérêt croissant pour des systèmes composés de plusieurs type de caméras dont les plus communément utilisées sont les caméras couleur, les caméras thermiques et les caméras de profondeur.

On assiste aujourd'hui à l'émergence d'un dispositif permettant de fournir en même temps deux images de modalités différentes grâce aux deux caméras qu'elle dispose (camera couleur et caméra infra-rouge). Ce système appelé Kinect a été initialement destiné pour le divertissement pour être ensuite intensivement exploitée dans divers applications (vidéo surveillance, reconnaissance de visage, imagerie médicale, télédétection, robotique).

Cependant, les images fournies par divers caméras contiennent des informations complémentaires ou redondantes qu'il convient de combiner ou de fusionner dans le but d'améliorer ou de faciliter l'interprétation de la scène. C'est dans cette optique que s'est développé la fusion des images.

Le but de notre travail est de fusionner les images issues de la Kinect (image en niveau de gris) et image de profondeur en une seule image par l'intermédiaire de la transformée en ondelettes.

Parmi les nombreuses méthodes de fusion, nous avons choisi d'utiliser la transformée en ondelettes. La transformée en ondelettes est une technique de représentation temps échelle très utilisée dans le domaine du traitement d'images. Elle permet de décomposer une image en sous images dites "approximation" contenant les basses fréquences et sous images qui contiennent les hautes fréquences et qui indiquent les détails horizontal, vertical et diagonal de l'image.

Le mémoire est principalement scindé en trois chapitres.

Le premier chapitre est dédié à la présentation de la caméra Kinect. Sa description, ses caractéristiques, son principe de fonctionnement ainsi que son utilité y sont principalement exposés.

Dans le deuxième chapitre, nous présentons un bref aperçu sur la fusion d'images. Nous définissons la fusion des images, son intérêt ainsi que les différentes approches. Nous décrivons la technique de fusion d'images basée sur la transformée en ondelettes.

Introduction générale

Le troisième chapitre est consacré à la description de la méthode de fusion d'images de la Kinect à savoir l'image en niveaux de gris (couleur) et l'image de profondeur en se basant sur la transformée en ondelettes. Différentes stratégies qui consistent à combiner l'information basses fréquences et les informations hautes fréquences des deux images sont présentées. Les résultats de ces méthodes de fusion sont évalués d'une manière objective par l'intermédiaire des critères de fusion tirés de la littérature. L'impact de chaque stratégie de fusion sur l'information texturale est également évalué par l'intermédiaire des attributs de texture extraits à partir des matrices de cooccurrence.

Chapitre 1

1.1 Introduction

La Kinect est un dispositif d'acquisition optique développé par Microsoft en 2010. Elle est popularisée au sein du grand public grâce à son usage dans le divertissement. La version pour PC avec son kit de développement, sortie l'année suivante l'a rendue également populaire dans la communauté scientifique ainsi que chez les développeurs indépendants.

Dans ce premier chapitre, nous tacherons de décrire les caractéristiques et le fonctionnement de la Kinect.

1.2 Présentation de la Kinect

La Kinect est une caméra utilisant des techniques d'interaction développée par la société Israélienne Prime Sense. Elle a longtemps été nommée par son nom de code « Project Natal » avant d'être officialisée et baptisée sous le nom Kinect. Elle est connectée via un périphérique d'entrée directement branché sur la console de Microsoft, qui porte le nom de Xbox 360. Ce qui lui permet d'interagir par commande vocale, reconnaissance de mouvement et d'image. D'ailleurs, le nom "Kinect" est issu des mots anglais "*kinetic*" (qu'on peut traduire par "cinétique") et "*connect*" (qu'on peut traduire par "connecter").

Plus précisément, la Kinect est un dispositif composée de plusieurs capteurs audio, d'une caméra couleur, d'un projecteur infrarouge, d'une caméra infrarouge et d'un pied motorisé (Fig.1.1).

Grâce à tous ces composants, la Kinect fournit trois types de données qui sont les données audio, des images couleur et des images de profondeur.



Figure (1.1) : Composants de la Kinect

Chapitre 2

2.1 Introduction

La fusion de données a été appliquée depuis déjà quelques années dans le domaine du traitement du signal, de l'automatisme et de la prise de décision. Très tôt, avec le développement de nouvelles et nombreuses méthodes d'imagerie, l'idée de combiner les différentes images d'un même objet ou d'une même scène est apparue intéressante et utile, dans de nombreux domaines d'application.

Nous consacrons ce chapitre à la présentation des notions liées à la fusion d'images. Sa définition, son intérêt ainsi que les différentes approches sont brièvement rappelés.

2.2 Fusion d'images

En traitement des images, la fusion d'informations se préoccupe de combiner au mieux des images d'origines différentes pour mieux connaître l'objet d'observation.

Le terme « fusion d'images » est souvent utilisé comme un synonyme de combinaison ou de superposition des images.

D'une manière générale, la fusion d'images consiste à combiner des images de différentes sources afin de synthétiser de nouvelles images plus riches en information. Elle permet de produire une nouvelle image qui conserve une partie de l'information contenue dans chacune des images originales. L'objectif est ainsi de créer une synergie, c'est-à-dire d'obtenir une image fusionnée géométriquement et/ou sémantiquement plus riche qu'une image initiale.

Les informations à fusionner peuvent provenir de plusieurs capteurs d'une même scène (multimodale) ou (multi sensor), d'une même scène prise à des temps différents généralement par un même capteur (multi temporal) et/ou d'une même scène prise par un capteur au même instant mais à des angles différents ou sous différentes conditions (multi-view).

De nombreuses méthodes sont capables de réaliser une fusion d'images. Elles diffèrent par la manière selon laquelle elles favorisent telle ou telle caractéristique des images originales.

2.3 Niveaux de la fusion d'images

La fusion d'images peut être effectuée de trois manières selon l'information ou le niveau considérée dans l'image (Fig.2.1). Soit au niveau soit du pixel (fusion bas niveau), soit au niveau des caractéristiques (fusion de moyen niveau), soit au niveau de décision (ou fusion haut niveau).

La fusion de bas niveau fonctionne dans le domaine spatial, le contenu de l'information liée à chaque pixel est déterminé à partir d'un ensemble de pixel dans l'image source. Dans la fusion moyenne niveau, les informations à fusionner sont les caractéristiques extraites sur les images sources qui correspondent à des attributs de texture ou les contours. Quand au niveau haut, les informations à fusionner sont des images d'entrée qui seront traitée individuellement pour l'extraction de l'information.

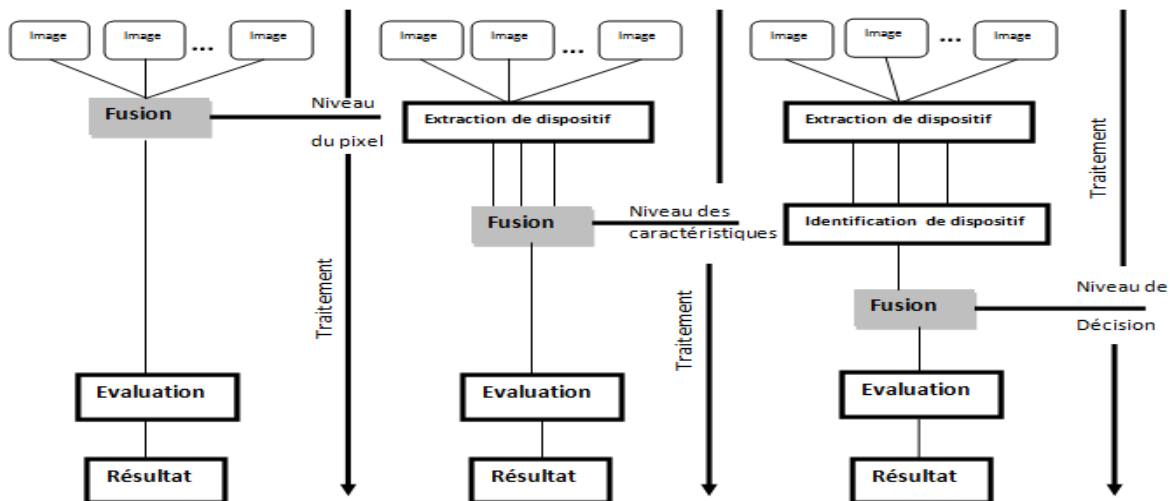


Figure (2.1) : Différents niveaux de fusion d'images.

Notons qu'avant d'effectuer la fusion, il est nécessaire de procéder à certaines opérations comme le redimensionnement des images, dans le cas où les images à fusionner ont des tailles différentes. Cette opération doit se faire sans perte d'informations et sans introduction de distorsions. Un autre problème qu'on rencontre souvent est celui du recalage. En effet dans certains cas, les images sources à fusionner sont acquises avec des capteurs ayant des positions et des optiques nécessaires de transformer ces images afin de les situer sur un même plan de référence. Ceci permet de représenter les informations issues des images dans un même référentiel afin de pouvoir les comparer ou les fusionner [6]

2.4 Méthodes de la fusion d'images

Plusieurs méthodes de fusion d'images ont été proposées [7]. Une manière simple de fusionner des images consiste à calculer la valeur moyenne des pixels de ces images [7]. Une autre technique est basée sur l'analyse des composantes principales également dite transformée de Karhunen-Loeve [7]. Toutes ces méthodes utilisent les valeurs des pixels (domaine spatial) lors de la fusion. Une autre catégorie de méthodes utilise les caractéristiques fréquentielles des pixels lors de la fusion. Parmi ces méthodes, on trouve celles basées sur la transformée en ondelettes [7].

2.5 Fusion d'images basée sur la transformée en ondelettes

La transformée en ondelettes offre un outil de choix pour la fusion d'image à cause de sa capacité à conserver toute l'information importante des images à fusionner et à présenter aucun objet façonné ou contradictoire avec le contenu des images à fusionner. De plus, elle permet de supprimer les caractéristiques indésirables comme le bruit et d'autres détails non pertinents [8], [9].

2.5.1 Transformée en ondelettes

2.5.1.1 La transformé en ondelette monodimensionnelle

La transformée en ondelettes est une technique de représentation temps échelle performante. Elle permet de décrire l'évolution temporelle des caractéristiques du signal relativement à une échelle d'observation donnée.

L'analyse est réalisé au moyen d'une fonction Ψ appelée ondelette de base (ou ondelette mère) qui permet de spécifier les caractéristiques du signal que l'on souhaite détecter. L'analyse en ondelettes consiste alors de positionner, dans le domaine temporel, l'ondelette mère en regard de la partie du signal à traiter, on parlera alors de la translation par l'utilisation du facteur b . L'ondelette mère est ensuite dilatée ou contractée, par l'utilisation du facteur d'échelle notée a , permettant de concentrer l'analyse sur une gamme donnée d'oscillations. Quand l'ondelette est dilatée, l'analyse explore les oscillations rapides comme celle contenues dans une discontinuité du signal. Par ce changement d'échelle (contraction et dilatation), la transformée en ondelettes $s(t)$ d'un signal est formulée par :

$$w(a, b) = \int_{-\infty}^{+\infty} s(t) \cdot \bar{\Psi}_{a,b}(t) dt \quad (2.1)$$

Ou $\bar{\Psi}_{a,b}(t) = \frac{1}{\sqrt{a}} \frac{\Psi(t-b)}{a}$ représente une famille de fonctions, $\Psi(a,b)$ est appelée ondelette analysante (fille) qui découle d'une ondelette,

Ψ est appelée l'ondelette mère, b est le paramètre de translation et a le paramètre d'échelle.

$\bar{\Psi}_{a,b}$ est le complexe conjugué de $\Psi(a,b)$

Le coefficient d'ondelette $w(a,b)$ d'un signal $s(t)$ dépend de la forme de celui-ci au voisinage du temps b . Quand $s(t)$ varie peu dans le temps, son produit par l'ondelette Ψ engendre une petite aire, autrement dit $w(a,b)$ est petit. Quand, au contraire, la séquence du signal est irrégulière et que ses variations de fréquence sont comparables à celles de l'ondelette, l'aire du produit entre le signal et l'ondelette est importante

La valeur des coefficients est d'autant plus grande que la fréquence du signal analysé coïncide avec celle de l'ondelette. La variation du paramètre 'a' permet de déterminer les modes du signal qui correspond à la fréquence $f_s = \frac{f_c}{a}$

Avec f_c : La fréquence centrale.

Une autre caractéristique de l'ondelette dilatée est d'être plus diffuse dans le temps est donc d'avoir un spectre plus concentré autour de sa fréquence centrale. L'inverse est constaté pour l'ondelette contractée.

La figure (2.2) présente les propriétés temporelles (image de gauche) et fréquentielle (image de droite) de 3 échelles différentes d'une ondelette :

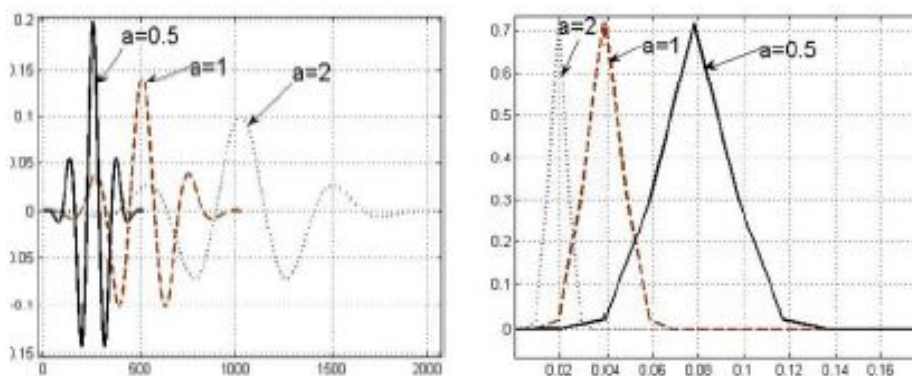


Figure (2.2) : Exemple d'une ondelette mère contractée et dilatée : $a=0.5$ correspond à une ondelette contractée, $a=1$ correspond à l'ondelette de base ou l'ondelette mère et $a=2$ correspond à une ondelette dilatée

La figure (2.3) montre le résultat de la transformée en ondelettes continu d'un signal.

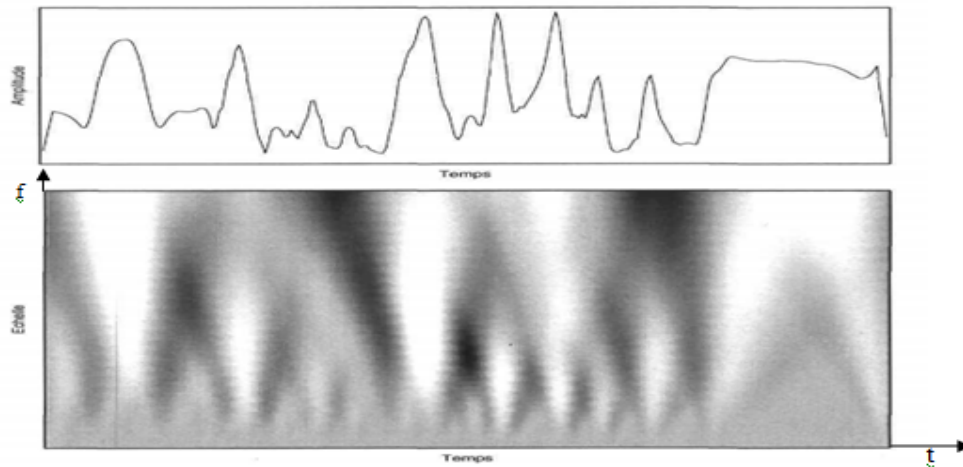


Figure (2.3) : Exemple de transformée en ondelettes continu.

Le signal $s(t)$ peut être reconstruit à partir des coefficients $w(a, b)$ et les fonctions d'ondelettes utilisées lors de la décomposition. Cette reconstruction s'effectue par le biais de la transformée en ondelettes inverse formulée comme suit :

$$s(t) = \frac{1}{C_\Psi} \iint_{-\infty}^{+\infty} \frac{w(a,b)}{a^2} \Psi_{a,b}(t) da db \quad (2.2)$$

C_Ψ est une constante telle que :

$$C_\Psi = 2\pi \int_{-\infty}^{+\infty} \frac{|\Psi(w)|^2}{|w|} dw \quad (2.3)$$

2.5.1.2 Transformée en ondelettes discrète

La T.O est dite continue quand les paramètres de dilatation et translation (a, b) varient continument.

La T.O discrète est une implémentation utilisant un ensemble discret d'échelles (a) et de translations (b) d'ondelettes obéissant à certaines règles. En d'autres termes, cette transformée décompose le signal en un ensemble d'ondelettes mutuellement orthogonales, ce qui constitue la différence principale avec la transformée en ondelettes continue.

La T.O discrète d'un signal $s(t)$ utilise une fonction ondelette $\Psi_{a,b}(t)$ telle que :

$$a = a_0^j \quad ; \quad b = k b_0 a_0^j \quad \text{avec : } b_0 \in \mathbb{R}_+^* \quad \text{et } j, k \in \mathbb{Z}$$

Dans ce cas, la T.O discrète prend la forme suivante :

$$C_s(j, k) = a_0^{-j/2} \int_{-\infty}^{+\infty} s(t) \overline{\Psi_{j,k}}(t) dt \quad (2.4)$$

En pratique, $a_0 = 2$ et $b_0 = 1$, pour aboutir à

$$C_s(j, k) = 2^{-j/2} \int_{-\infty}^{+\infty} s(t) \cdot \Psi(2^{-j}t, k) dt \quad (2.5)$$

2.5.1.3 Transformée en ondelette bidimensionnelle

La T.O 2D découle directement de la T.O 1D. Ainsi, la méthode la plus simple pour obtenir les fonctions d'ondelettes et d'échelle en deux dimensions est de multiplier deux fonctions à une dimension.

La reconstruction du signal à partir de ses coefficients est obtenue par :

$$S(t) = \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} C(j, k) \Psi_{j,k}(t) \quad (2.6)$$

On obtient la fonction d'échelle par la multiplication de deux fonctions d'échelles. L'obtention des fonctions d'ondelettes est réalisée par multiplication de deux fonctions d'ondelettes 1D ou d'une fonction d'ondelettes et d'une fonction d'échelle.

$$\Phi(x, y) = \Phi(x) \cdot \Phi(y) \quad (2.7)$$

Avec $\Phi(x)$ = Fonction d'échelle a une variable x

$\Phi(y)$ = Fonction d'échelle a une variable y

Ainsi une T.O 2D possède trois fonctions d'ondelettes

$$\Psi^H(x, y) = \Psi(x) \cdot \Phi(y) \quad (2.8)$$

$$\Psi^V(x, y) = \Phi(x) \cdot \Psi(y) \quad (2.9)$$

$$\Psi^D(x, y) = \Psi(x) \cdot \Psi(y) \quad (2.10)$$

Avec $\Psi(x)$ = Fonction d'ondelettes a une variable x

$\Psi(y)$ = Fonction d'ondelettes a une variable y

Le but de l'algorithme multi-résolution est de calculer les coefficients de détails $d_{j,k}$, avec $k, j \in \mathbb{Z}$ ou d'ondelettes à partir des différentes approximations. Les coefficients c_j^k sont exprimés en fonction des coefficients c_k^{j-1}

$$c_k^j = \sum_{n=-\infty}^{+\infty} \bar{h}(2k-1) c_k^{j-1} \quad (2.11)$$

Avec $\bar{h} = h(-n)$

Où h est un filtre qui est donné par l'équation suivante :

$$h(n) = \int_{-\infty}^{+\infty} \Phi(x) \Phi_{-1,n}(x) dx$$

Les coefficients d'ondelettes représentant la différence d'information entre ces deux approximations Mallat (1989) sont calculés par :

$$d_k^j = \sum_{n=-\infty}^{+\infty} \bar{g}(2k-n) c_k^{j-1} \quad (2.12)$$

avec : $\bar{g}(n) = g(-n)$.

Où g est un filtre qui est donné par l'équation suivante :

$$g(n) = \int_{-\infty}^{+\infty} \Phi(x) \Psi_{-1,n}(x) dx = (-1)^n h(1-n) \quad (2.13)$$

Les coefficients d_k^j représentent la différence d'information entre les approximations d'une fonction avec les coefficients c_k^j et c_k^{j-1} .

Connaissant d_k^j et c_k^j , les coefficients c_k^{j-1} peuvent être calculés de façon exacte par :

$$c_k^{j-1} = \sum_{n=-\infty}^{+\infty} h(k-2n) c_n^j + \sum_{n=-\infty}^{+\infty} g(k-2n) d_n^j \quad (2.14)$$

D'un point de vue pratique, la décompositions en ondelettes est effectuée selon l'algorithme de Mallat [9]. Il consiste à filtrer l'image suivant les lignes puis les colonnes par deux filtres passe-haut (\bar{g}) et passe bas (\bar{h}), afin d'obtenir une qui contient les basses fréquences dites "approximation" dans l'image (l'approximation) et 3 images qui contiennent les hautes fréquences dites "les détails horizontal, vertical et diagonal".

Les équations de décomposition s'écrivent :

$$c_{k,l}^j = \sum_n \sum_m \bar{h}(2k-n) \bar{h}(2l-m) c_{n,m}^{j-1} \quad (2.15)$$

$$d_{k,l}^{H,j} = \sum_n \sum_m \bar{h}(2k-n) \bar{g}(2l-m) c_{n,m}^{j-1} \quad (2.16)$$

$$d_{k,l}^{V,j} = \sum_n \sum_m \bar{g}(2k - n) \bar{h}(2l - m) c_{n,m}^{j-1} \tag{2.17}$$

$$d_{k,l}^{D,j} = \sum_n \sum_m g(2k - n) \bar{g}(2l - m) c_{n,m}^{j-1} \tag{2.18}$$

Ou :

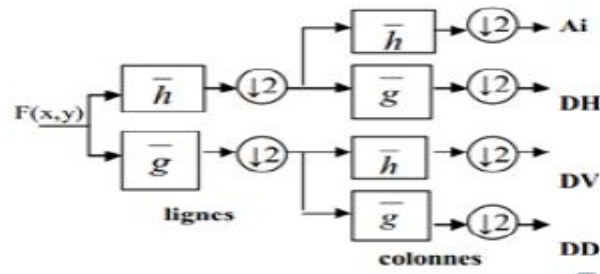
$c_{k,l}^j$ est l'approximation à la résolution j,

$d_{k,l}^{H,j}$ est le coefficient des détails horizontaux

$d_{k,l}^{V,j}$ est le coefficient des détails verticaux

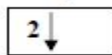
$d_{k,l}^{D,j}$ est le coefficient des détails diagonaux

Les figures (2.4) et (2.5) illustre le principe de cet algorithme

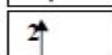


h Filtre passe bas

g Filtre passe haut



: Décimation d'une ligne (colonne) sur deux



: Interpolation des lignes (colonnes), c'est-à-dire mettre une ligne (colonne) de zéros entre deux

Figure (2.4) : Algorithme de décomposition de Mallat.

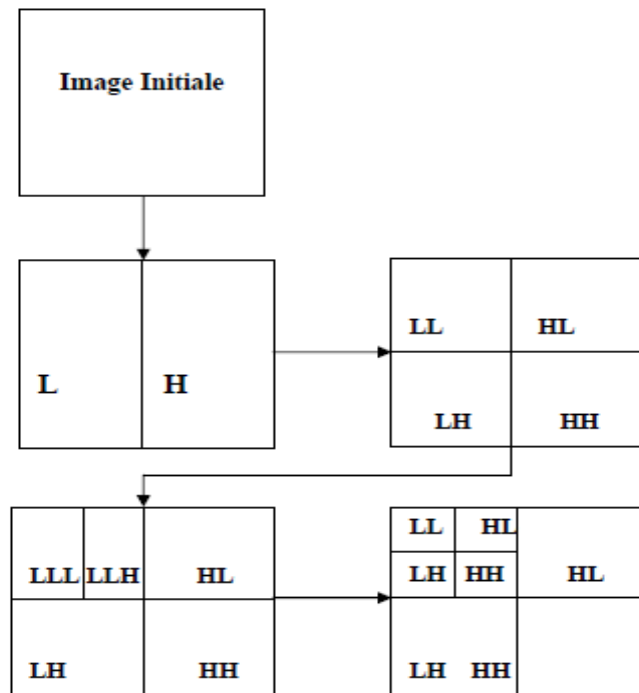


Figure (2.5) : Décomposition pyramidale d'une image

L: Low pass

H: High pass

2.5.2 Types d'ondelettes

Il existe plusieurs types d'ondelettes, les plus courantes sont :

- ✓ Ondelettes continues
- Ondelette de Morlet
- Ondelette Sombéro
- ✓ Ondelettes discrète
- Ondelettes orthogonales : Haar, Shanon, Meyer, Battle-Lemarié Daubechies, coiflets
- Ondelettes bi-orthogonales : Cohen-Daubechies-Feauveau, Chui
- Ondelettes semi-orthogonales : spline

Le choix de l'ondelette dépend surtout de l'application. En traitement d'image, on préfère utiliser des ondelettes bi-orthogonales qui permettent d'allier les propriétés de linéarité en phase.

2.6 Domaine d'application

La fusion d'images trouve sa place dès lors que plusieurs acquisitions du même « objet » (au sens large) sont effectuées et sont à analyser. Les premiers essais de fusion d'images ont été réalisés à partir d'images satellitaires et aériennes, puis pour des images dites « naturelles », en lumière visible, dans le milieu biomédical et industriel.

2.6.1 Imagerie aérienne et satellitaire

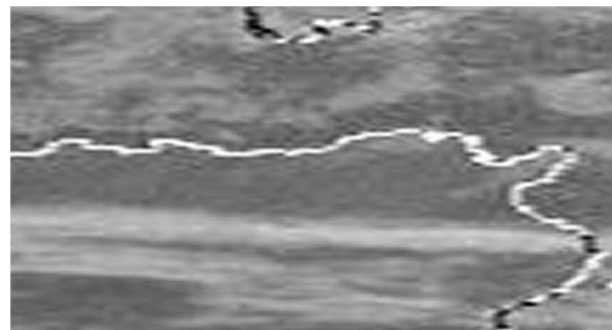
C'est le premier domaine où les méthodes de fusion numérique d'images ont été mises en place. Les premières applications furent militaires. La fusion a été développée pour permettre la reconnaissance d'objets, la poursuite de cibles ou encore la surveillance du champ de bataille. Le commandant en chef dispose de nombreuses sources d'informations (radar, laser, infrarouge, visible...) Le but de la fusion est de combiner de manière rapide et sûre toutes ces informations de manière à offrir au stratège une aide à la décision.

D'autres applications de la fusion d'images ont vu récemment le jour du fait du développement des techniques d'imagerie par satellite et par avion. Ces images sont souvent multi spectrales, c'est-à-dire que le satellite est capable de fournir des images de la même zone géographique, mais acquises dans des conditions différentes avec des capteurs différents. Citons par exemple les canaux XS1, XS2, et XS3 du satellite SPOT qui permet d'accéder de manière assez simple à des images susceptibles d'être fusionnées. Les applications de la fusion en imagerie satellitaire et aérienne permettent d'automatiser un certain nombre de tâches en climatologie, en prédiction des récoltes....

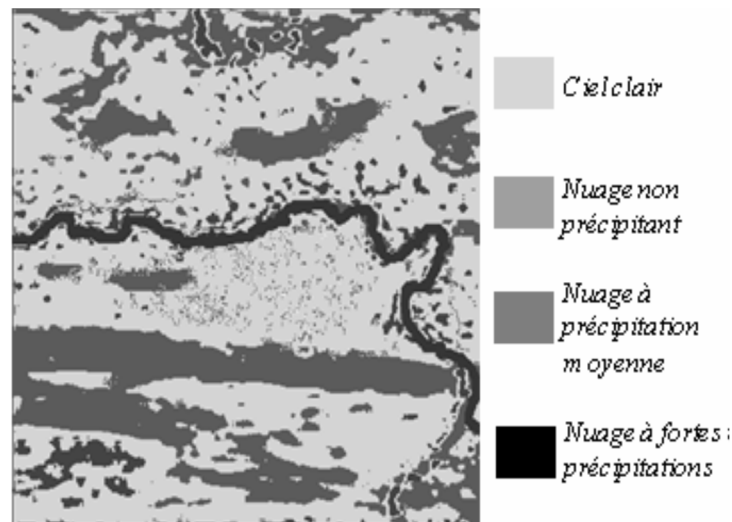
La figure (2.6) montre un exemple de fusion d'une image satellitaire avec une image Radar.



a



b



c

Figure (2.6) : Exemple de fusion d'images satellitaire et radar. a : Image satellitaire.
 b : Image Radar.
 c : Image fusionnée

2.6.2 Imagerie médicale

Dans le domaine médical, le médecin à sa disposition un nombre croissant d'images issues de modalités d'imagerie différentes (imagerie par rayons X, par ultrasons, par résonance magnétique, par émission de positons, ou encore imagerie nucléaire). Il s'agit pour lui de synthétiser toutes ces informations afin d'établir un diagnostic le plus fiable, le plus précis et le plus exact possible. En général, il les compare et les intègre mentalement pour se forger une représentation mentale de l'organe étudié afin de prendre sa décision.

Pourtant, les conditions de comparaison de ces images sont souvent difficiles, car ces images n'ont pas forcément la même taille, ni la même résolution spatiale, ni le même contraste, et ne représentent pas forcément l'organe sous le même point de vue. De plus, de par la nature physique de cette imagerie, les structures visualisées dans les images peuvent avoir des représentations bien différentes. Par exemple, une radiographie par rayons X représente une projection de l'organe dans un plan, alors qu'une échographie met en évidence les interfaces entre les tissus au niveau d'une coupe. Dans tous les cas, il est important, voire indispensable, de recalibrer toutes les données dans un même référentiel afin de pouvoir réaliser la comparaison ou l'intégration des informations. Ce n'est pas toujours facile. Nous détaillons ci-après les principales méthodes de recalage. D'une manière générale, le recalage permet déjà une meilleure comparaison des images, car il les place toutes dans un même repère.

La figure (2.7) illustre un exemple de fusion d'une image PET et une image CT :

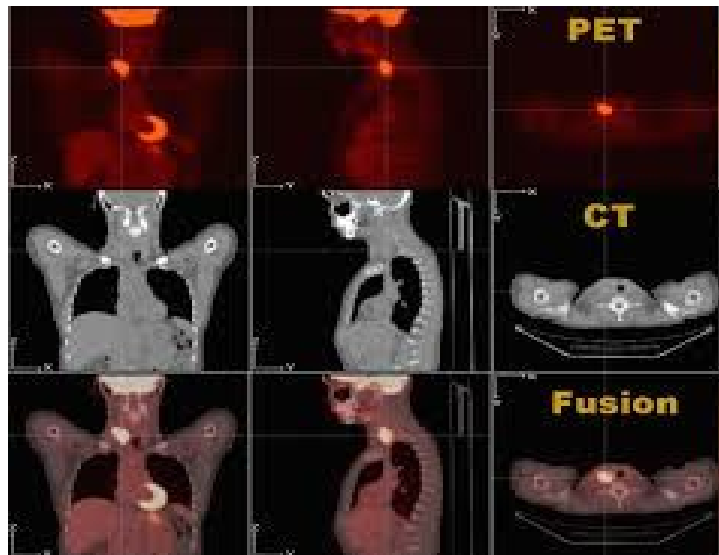


Figure (2.7) : Fusion d'une image PET et une image CT en médecine

2.6.3 Imagerie industrielle

Dans le domaine de l'industrie, et plus particulièrement en contrôle non destructif, on commence de plus en plus à utiliser différents types de modes de contrôle pour détecter de manière systématique les défauts. L'opérateur se retrouve alors dans la même situation que le médecin, avec souvent une contrainte de temps réel supplémentaire.

La plupart des applications consiste à détecter des défauts dans des pièces métalliques, et les imageries utilisées sont le plus souvent la radiographie sur film, la radioscopie numérique, les ultrasons et le contrôle par courants de Foucault. Les méthodes de fusion sont mises en place pour des pièces très précises dont il faut améliorer la fiabilité et la rapidité de contrôle.

2.7 Conclusion

Nous avons présenté dans ce chapitre un aperçu sur un domaine bien particulier du traitement d'images qui est celui de la fusion d'images. La fusion d'images apparaît indispensable dans certaines applications. Ceci explique le foisonnement des techniques utilisées dans la fusion. Parmi ces techniques, nous nous sommes particulièrement intéressé à celle basée sur la transformée en ondelettes. L'utilisation de cet outil pour fusionner les images issues du dispositif Kinect à savoir les images couleur et profondeur fera l'objet du prochain chapitre.

Chapitre 3

3.1 Introduction

Ce chapitre est consacré à la fusion des images de la Kinect (couleur, profondeur) par l'intermédiaire de la transformée en ondelettes dont le principe est donné dans le chapitre précédent. Sur la base de cette transformée, plusieurs stratégies de fusion d'une image couleur et d'une image de profondeur sont exposées. L'évaluation de ces fusions est effectuée moyennant plusieurs critères d'évaluation.

3.2 Fusion d'images de la Kinect par la transformée en ondelettes

L'idée fondamentale de la fusion d'images de la Kinect basée sur la T.O consiste à effectuer une décomposition multi-résolution sur chaque image source (image en niveau de gris et l'image de profondeur). Les coefficients de la bande basses fréquences (approximation) et des bandes hautes fréquences (détails) sont alors combinées selon une certaine règle de fusion.

Après celle, l'image fusionnée est obtenue en exécutant le IDWT (Inverse Discret Wavelet Transform qui correspond à la transformée en ondelettes discrète inverse) pour les coefficients combinés correspondants de l'ondelette.

Sept règles de fusion ont été appliquées dans notre travail :

- Fusion 1 : Elle consiste à combiner les coefficients d'approximation de l'image en niveau de gris avec les coefficients de détails de l'image de profondeur (Fig.3.1)

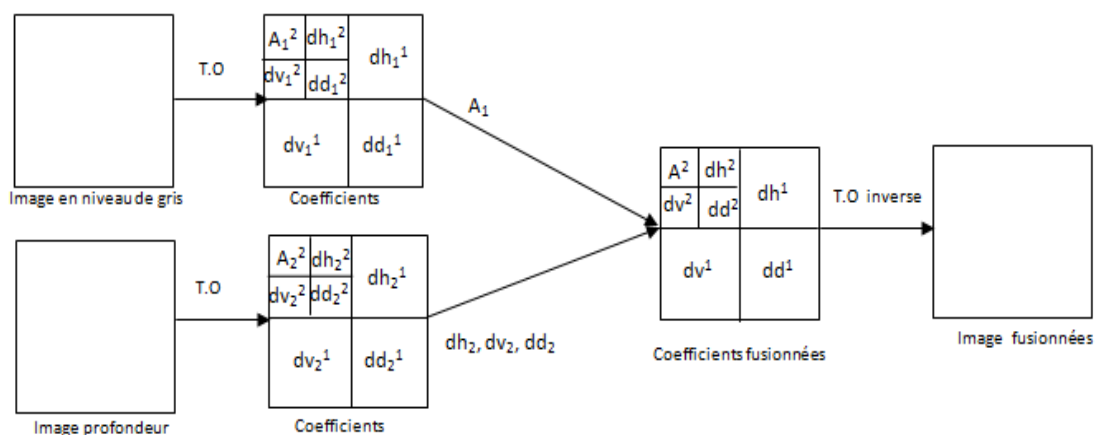


Figure (3.1) : Fusion des coefficients d'image de la Kinect par la T.O

- Fusion 2 : Elle combine les coefficients d'approximation de l'image de profondeur avec les coefficients de détails de l'image en niveau de gris (Fig.3.2)

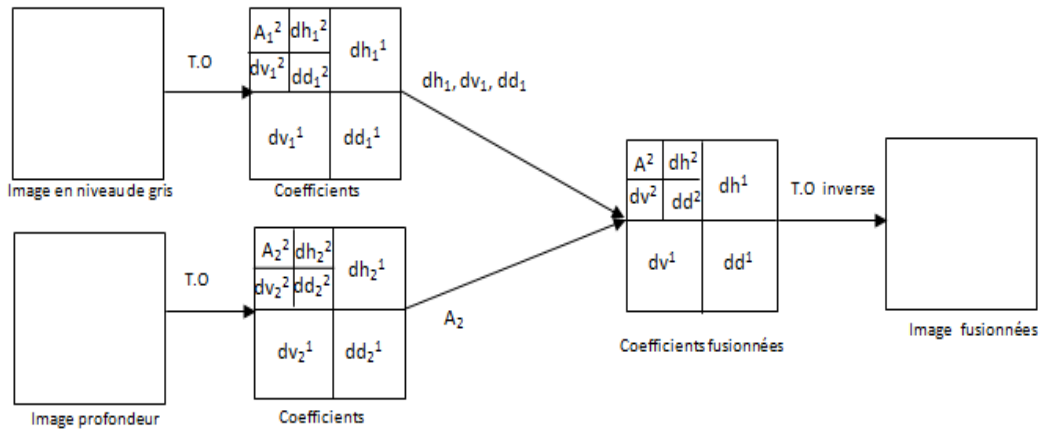


Figure (3.2) : Fusion des coefficients d'image de la Kinect par la T.O

- Fusion 3 : La moyenne de coefficients d'approximation et de détails des deux images sources (en niveau de gris et profondeur) est calculée (Fig.3.3)

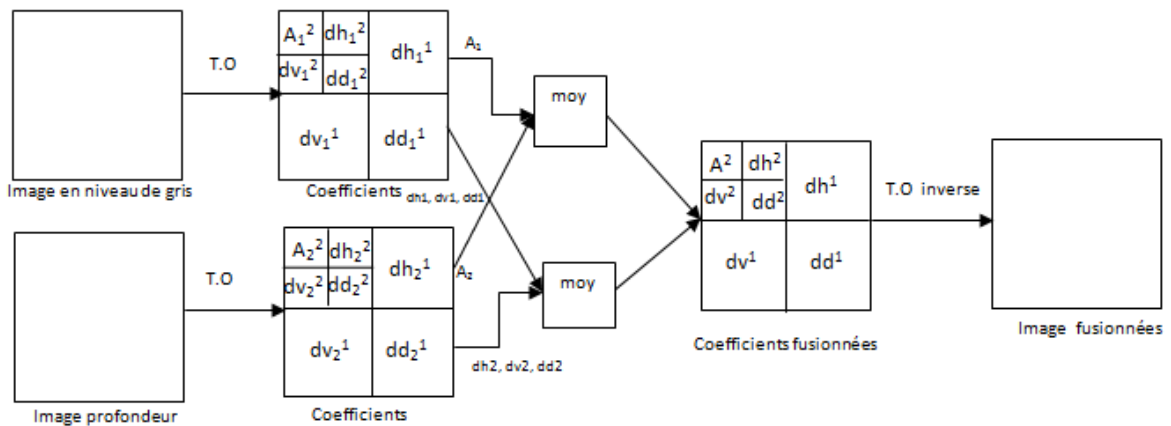


Figure (3.3) : Fusion de la moyenne des coefficients d'image de la Kinect par la T.O

Chapitre 3 Fusion d'image de la Kinect à base de la transformée en ondelettes

- Fusion 4 : elle sélectionne le minimum de coefficients d'approximation et le minimum des coefficients de détails entre les deux images sources (Fig.3.4)

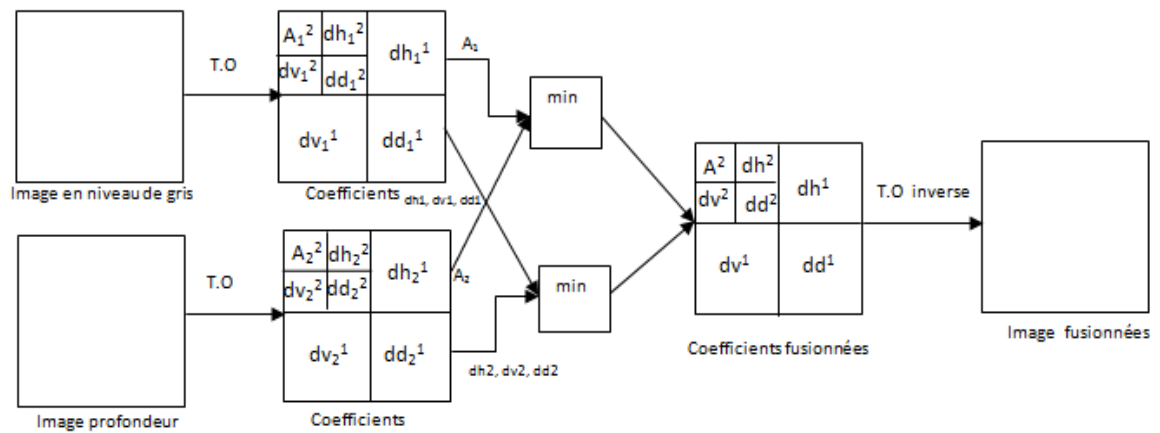


Figure (3.4) : Fusion du minimum des coefficients d'image de la Kinect par la T.O

- Fusion 5 : elle sélectionne le maximum de coefficients d'approximation et le maximum des coefficients de détails entre les deux images sources (Fig.3.5)

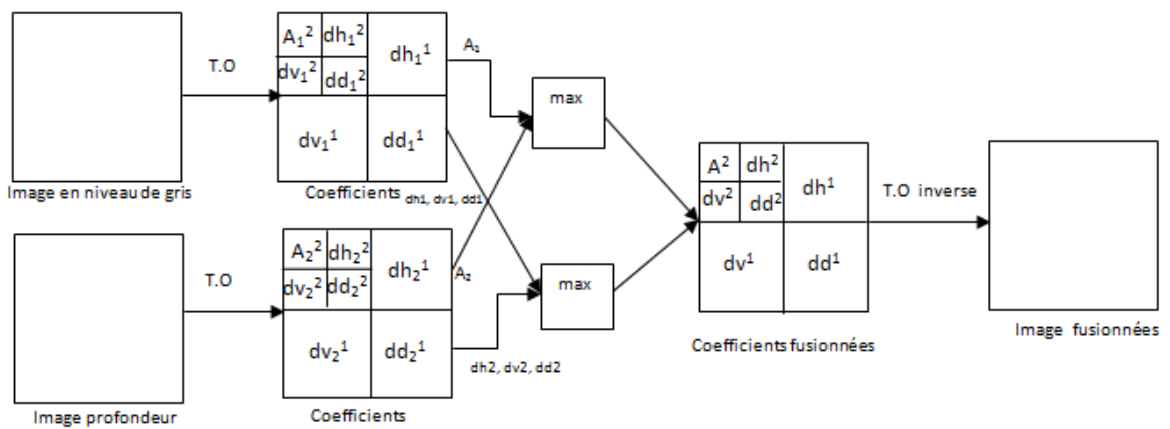


Figure (3.5) : Fusion du maximum des coefficients d'image de la Kinect par la T.O

Chapitre 3 Fusion d'image de la Kinect à base de la transformée en ondelettes

- Fusion 6 : elle sélectionne le maximum des coefficients d'approximation et le minimum des coefficients de détails entre les deux images sources (Fig.3.6)

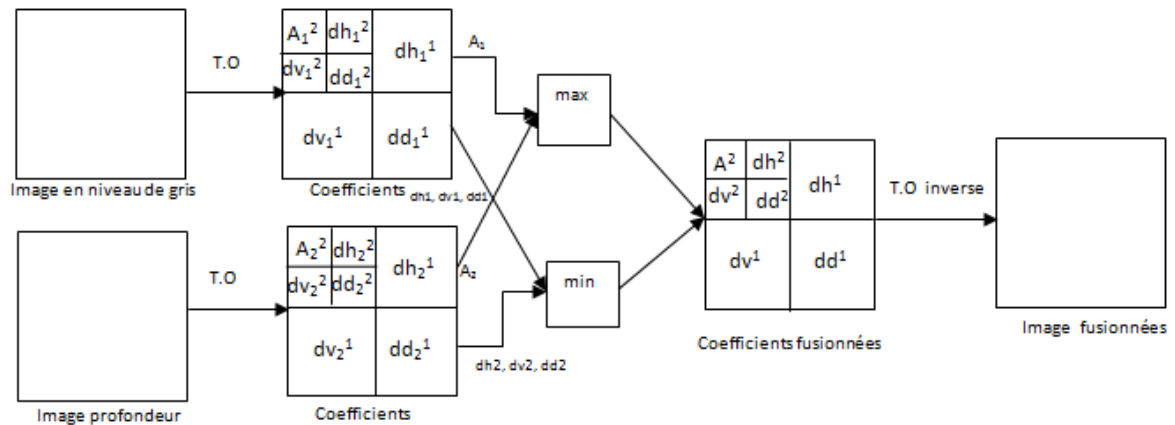


Figure (3.6) : Fusion du max-min des coefficients d'image de la Kinect par la T.O

- Fusion 7 : elle sélectionne le minimum de coefficients d'approximation et le maximum des coefficients de détails entre les deux images sources (Fig.3.7)

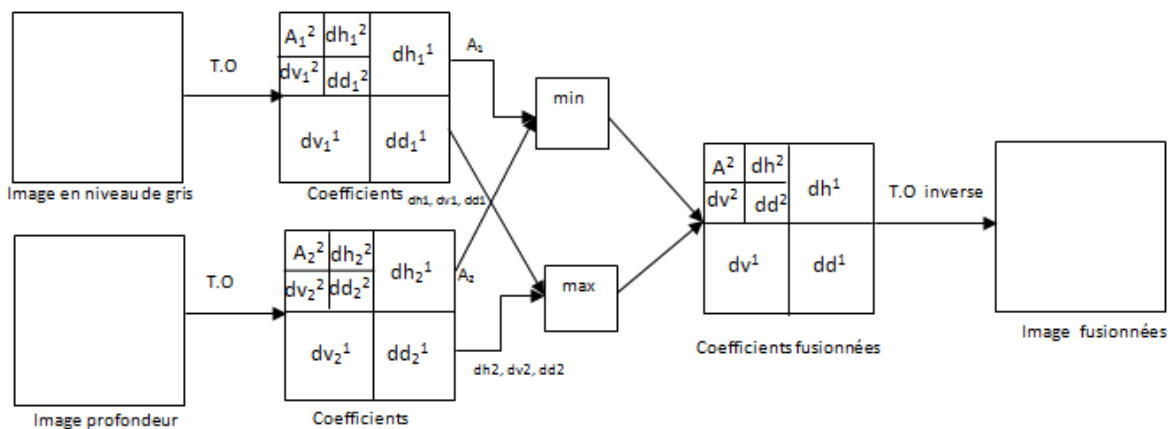


Figure (3.7) : Fusion du min-max des coefficients d'image de la Kinect par la T.O

3.3 Résultats de la Fusion d'images de la Kinect par la transformée en ondelettes

3.3.1 Critères d'évaluation des performances de la fusion

L'évaluation des performances de fusion est une tâche délicate car la vérité terrain n'est pas disponible dans la plupart des applications [11], [12]. Plusieurs paramètres d'évaluation ont été proposés dont les plus classiques sont:

Moyenne M :

L'intensité moyenne notée (\bar{F}) mesure le niveau de gris moyen de l'image fusionnée.

$$\bar{F} = \frac{\sum_{i=1}^m \sum_{j=1}^n f(i,j)}{m \times n} \quad (3.1)$$

avec $f(i, j)$ l'intensité du pixel à de coordonnées (i, j) et $m \times n$ est la taille de l'image

Ecart type SD :

L'écart type est la racine carrée des variations des niveaux de gris de l'image fusionnée. Il mesure l'étendu des niveaux de gris.

$$SD = \sqrt{\frac{\sum_{i=1}^m \sum_{j=1}^n (f(i,j) - \bar{F})^2}{m \times n}} \quad (3.2)$$

Gradient moyen AG:

Le gradient moyen ou Average Gradient noté (AG) mesure le degré de clarté de l'image fusionnée.

$$AG = \frac{\sum_i \sum_j ((f(i,j) - f(i+1,j))^2 + (f(i,j) - f(i,j+1))^2)^{\frac{1}{2}}}{m \times n} \quad (3.3)$$

Entropie H:

L'entropie notée (H) estime la quantité de l'information présentée dans l'image fusionnée !

$$H = - \sum_{k=0}^{255} p_f(k) \log_2(p_f(k)) \quad (3.4)$$

Avec p_k est la probabilité de la valeur k d'intensité dans une image codée sur 8 bits.

$p_k = \frac{h(k)}{m \times n}$ ou $h(k)$ est le nombre de pixels ayant le niveau de gris k

Information mutuelle MI :

L'information mutuelle ou Mutuel Information notée (MI) mesure l'information mutuelle globale entre les images de source A et B et l'image fusionnées F. Elle est donnée par :

$$MI = MI_{AF} + MI_{BF} \quad (3.5)$$

Où,

$MI_{AF} = \sum_{i=0}^{255} \sum_{j=0}^{255} p_{A,F}(i,j) \log_2 \left(\frac{p_{A,F}(k,l)}{p_A(k) \times p_F(l)} \right)$ → est la quantité d'information mutuelle entre l'image source A et l'image fusionnée F.

$MI_{BF} = \sum_{i=0}^{255} \sum_{j=0}^{255} p_{B,F}(i,j) \log_2 \left(\frac{p_{B,F}(k,l)}{p_B(k) \times p_F(l)} \right)$ → l'information réciproque entre l'image source B et l'image fusionnée F.

Et $p_{A,F}(k, l)$ et $p_{B,F}(k, l)$ représente la probabilité conjointe d'avoir le niveau de gris k dans l'image source A ou B et le niveau de gris l de l'image fusionnée.

Information symétrique FS :

L'information symétrique ou Symetry Fusion notée (FS) indique combien symétrique l'image fusionnée. Elle est donnée par :

$$FS = 2 - \left| \frac{MI_{AF}}{MI} - 0.5 \right| \quad (3.6)$$

Coefficient de corrélation CC :

Le coefficient de corrélation noté (cc) mesure la pertinence de l'image fusionnée par rapport aux images sources

$$CC = (r_{AF} + r_{BF})/2 \quad (3.7)$$

Avec

$$r_{AF} = \frac{\sum_i \sum_j (a(i,j) - \bar{A})(f(i,j) - \bar{F})}{\sqrt{(\sum_i \sum_j (a(i,j) - \bar{A})^2)(\sum_i \sum_j (f(i,j) - \bar{F})^2)}}$$

$$r_{BF} = \frac{\sum_i \sum_j (b(i,j) - \bar{B})(f(i,j) - \bar{F})}{\sqrt{(\sum_i \sum_j (b(i,j) - \bar{B})^2)(\sum_i \sum_j (f(i,j) - \bar{F})^2)}}$$

Fréquence spatiale SF:

La fréquence spatiale ou Spatial Frequency, notée (SF) mesure le niveau global de l'information dans les régions (le niveau d'activité) d'une image fusionnée

$$SF = \sqrt{RF^2 + CF^2} \quad (3.8)$$

Où:

$$RF = \sqrt{\frac{\sum_i \sum_j (f(i,j) + f(i,j-1))^2}{m \times n}}$$

$$CF = \sqrt{\frac{\sum_i \sum_j (f(i,j) - f(i,j-1))^2}{m \times n}}$$

3.3.2 Attributs de texture

En plus de paramètres classiques décrites précédemment, nous proposons dans ce mémoire d'étudier l'influence des méthodes de fusion sur l'information de texture.

La notion de texture a toujours été rattachée à la notion de perception. Elle renvoie à l'apparence et la consistance d'un objet. Plus précisément, une texture décrit la structure, l'agencement géométrique de celui-ci. Elle se manifeste par une information visuelle, et se décrit par les termes linguistiques qualitatifs comme la finesse, la granularité, la régularité, le contraste, etc.

Plusieurs méthodes d'analyse de la texture ont été développées. Elles ont pour but d'extraire un ensemble d'attributs ou paramètres pouvant décrire les caractéristiques de la texture. Ces attributs doivent être représentatifs, pertinents et discriminants de façon qu'on puisse discerner une texture parmi d'autres. Parmi ces méthodes, on trouve celle qui est basée sur les matrices de cooccurrence.

Les matrices de cooccurrence ont été introduites par Haralick en 1973 [13], Du fait de leur richesse en information de texture, les matrices de cooccurrences sont devenues les plus connues et les plus utilisées pour extraire ces caractéristiques de textures.

Son principe est basé sur la mesure de la distribution des niveaux de gris dans l'image tout en prenant en compte les interactions spatiales entre les pixels.

Chapitre 3 Fusion d'image de la Kinect à base de la transformée en ondelettes

Une matrice de cooccurrence mesure la probabilité d'apparition des paires de valeurs de pixels situés à une certaine distance dans l'image. Elle est basée sur le calcul de la probabilité $P(i, j, \delta, \theta)$ qui représente le nombre de fois où un pixel de niveau de gris i apparaît à une distance relative δ d'un pixel de niveau de gris j et selon une orientation θ donnée.

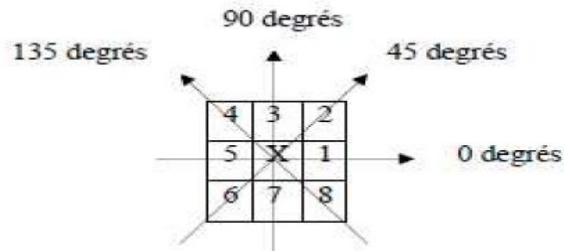


Figure (3.8): Plus proches voisins du pixel 'x' selon 4 directions

Les directions angulaires θ classiquement utilisées sont 0, 45, 90 et 135 degrés, par exemple, les plus proches voisins de 'x' selon la direction $\theta = 135$ degrés sont les pixels 4 et 8 (Fig.3.8).

Haralick a introduit quatorze attributs de texture extraits des matrices de cooccurrences. Les plus importants sont :

Energie : L'énergie mesure l'homogénéité de l'image. Plus cette valeur est faible, moins l'image est uniforme et dans ce cas, il existe beaucoup de transitions de niveaux de gris.

$$E = \sum_{i,j} p(i,j)^2 \quad (3.9)$$

Contraste : Le contraste (ou inertie) mesure les variations locales des niveaux de gris. Si ces variations sont importantes, alors le contraste sera élevé.

$$\text{Cont} = \sum_{i,j} |i - j|^2 p(i,j) \quad (3.10)$$

Corrélation : Ce paramètre permet de déterminer si certaines colonnes de la matrice si sont égales, c'est-à-dire s'il existe des dépendances linéaires dans l'image. Elle mesure la dépendance linéaire (relativement à (δ, θ)) des niveaux de gris de l'image.

$$\text{Cor} = \sum_{i,j} \frac{(i-\mu_i)(j-\mu_j)p(i,j)}{\sigma_i \sigma_j} \quad (3.11)$$

$\mu =$

Moment différencie inverse (Homogénéité) : Ce paramètre a un comportement inverse de celui du contraste. En effet, plus la texture possède de régions homogènes et plus le moment différentiel inverse est élevé.

$$H = \sum_{i,j} \frac{p(i,j)}{1+|i-j|} \quad (3.12)$$

3.4 Tests et Résultats

Pour évaluer les performances de la fusion des images issues de système Kinect (image couleur et de profondeur), nous avons utilisé 2 exemples d'images prises par la caméra Kinect. La figure (3.9) montre l'image profondeur, l'image couleur et l'image en niveau de gris de chaque exemple. Toutes ces images ont la même taille (640*480).

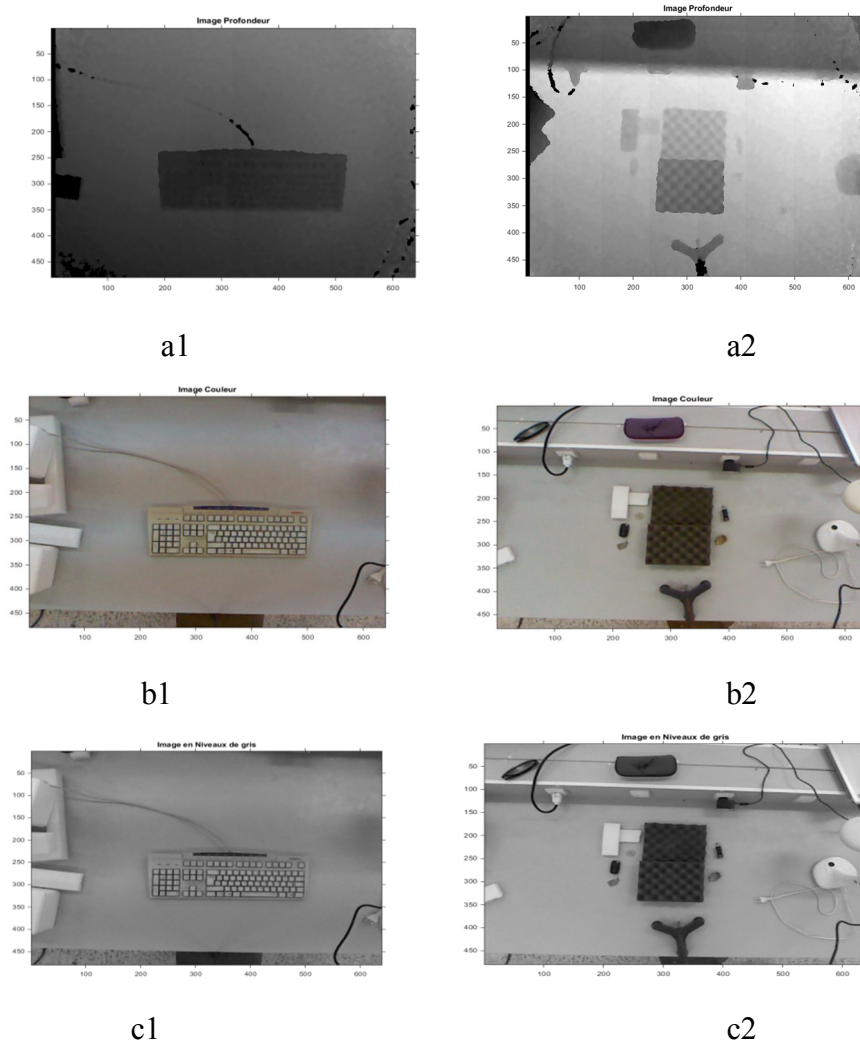


Figure (3.9) : Images tests. Colonne 1 : exemple 1, Colonne 2 : exemple 2.
Ligne 1 : image profondeur
Ligne 2 : image couleur.
Ligne 3 : image en niveau de gris

Chapitre 3 Fusion d'image de la Kinect à base de la transformée en ondelettes

Les figures (3.10) et (3.11) montrent le résultat de l'application de la T.O sur l'image en niveau de gris et l'image profondeur de l'exemple 1 respectivement avec l'ondelette Haar et sur 3 niveaux de décomposition.

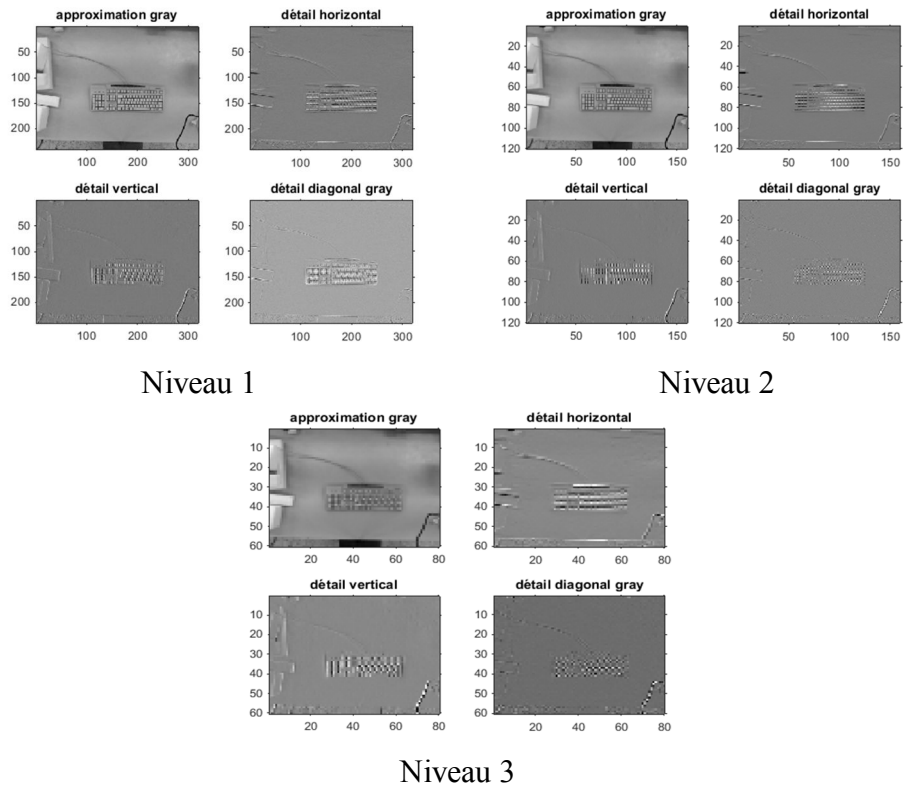


Figure (3.10) : Décomposition de l'image couleur en niveau de gris

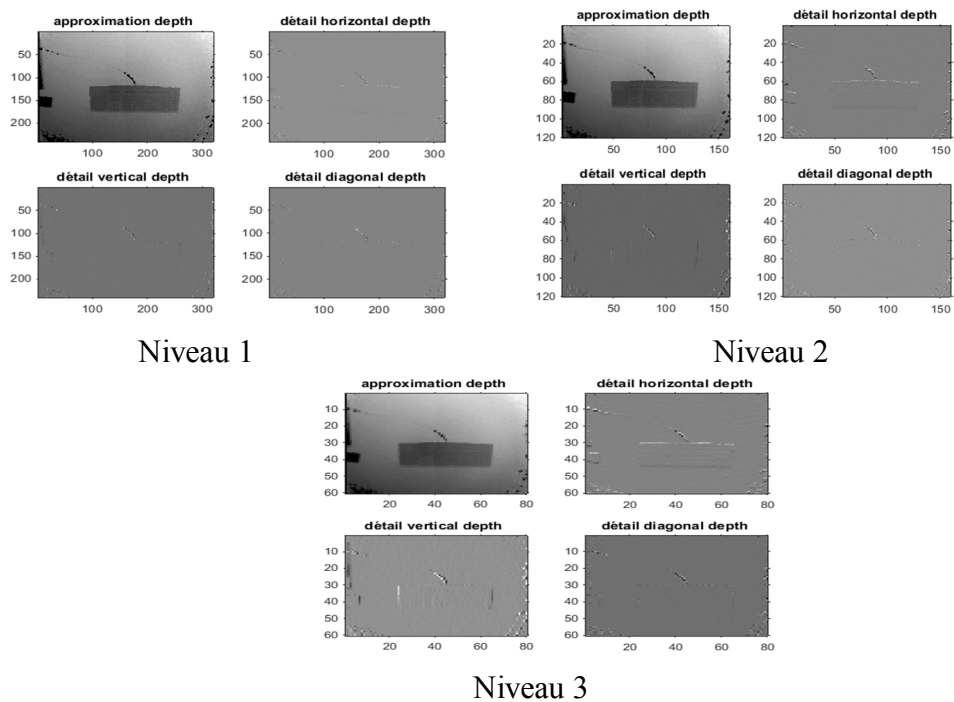


Figure (3.11): Décomposition de l'image de profondeur

Chapitre 3 Fusion d'image de la Kinect à base de la transformée en ondelettes

Ces deux figures laissent apparaître des informations différentes produites par l'image en niveaux de gris ou l'image profondeur grâce à leurs images approximation et aux images de détail horizontaux, verticaux et diagonaux.

Les figures (3.12) et (3.13) montrent les résultats des sept méthodes de fusion obtenues sur les 2 images test de la Kinect.

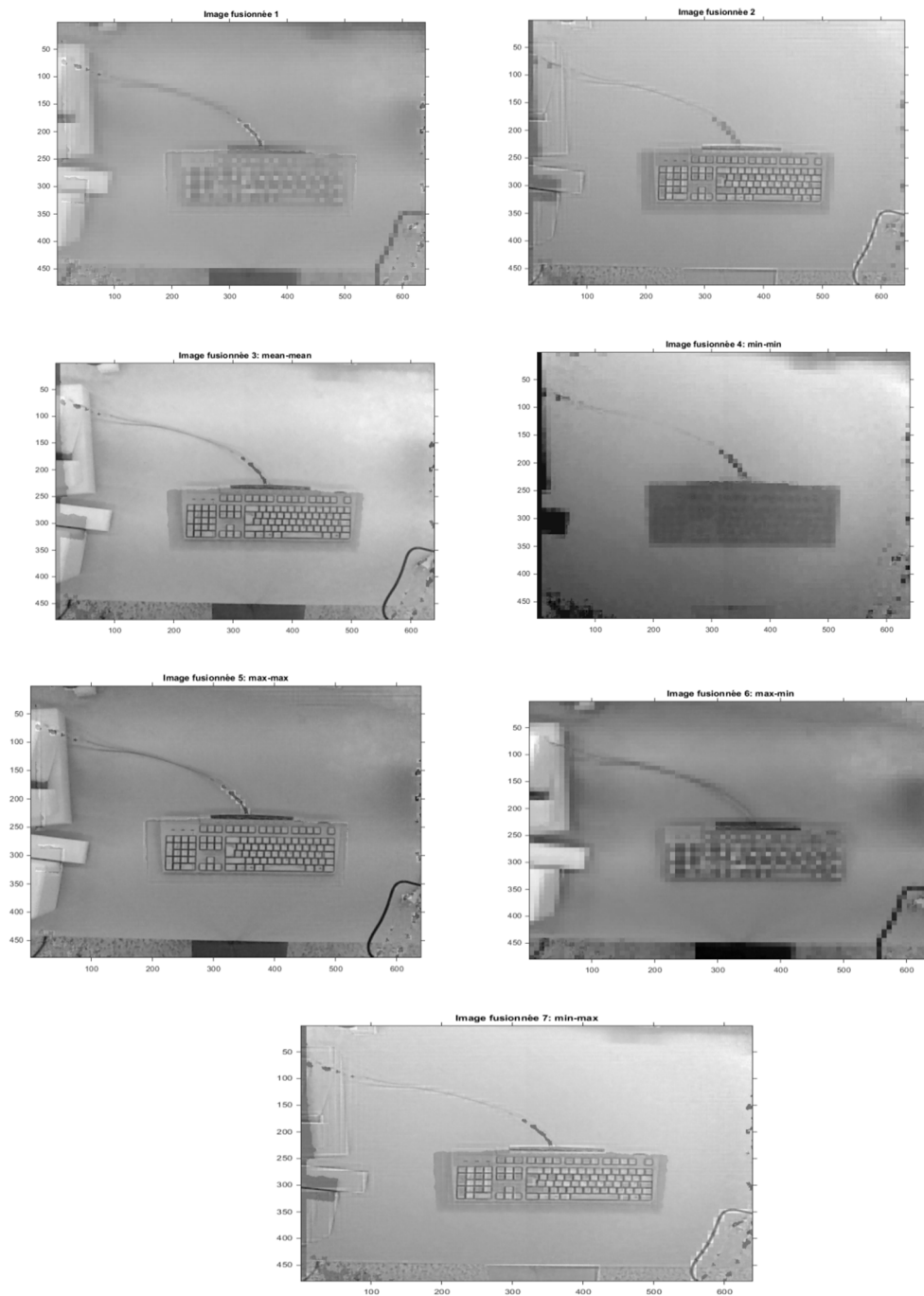


Figure (3.12) : Images fusionnées de la 1^{er} image test

Chapitre 3 Fusion d'image de la Kinect à base de la transformée en ondelettes

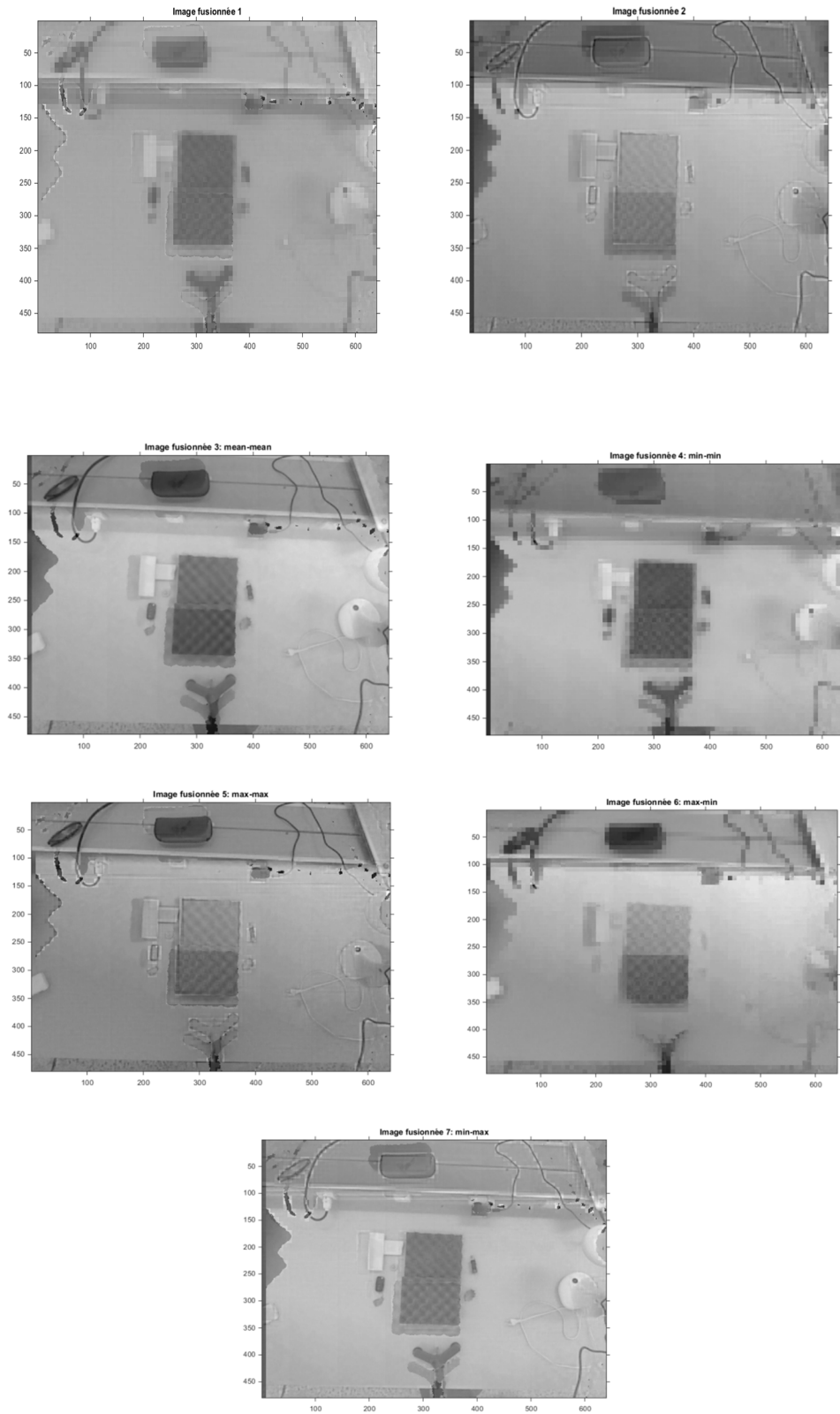


Figure (3.13) : Images fusionnées de la 2^{ème} image test

Chapitre 3 Fusion d'image de la Kinect à base de la transformée en ondelettes

En effet, on peut remarquer sur les images fusionnées que l'information apportée par l'image couleur en niveau de gris est prédominante dans les méthodes de fusion 2, 3, 5 et 7. Tandis que dans les méthodes 1, 4 et 6 sauf les contours sont préservés.

Les tableaux (3.1) et (3.2) donnent les valeurs des 8 critères d'évaluation de chaque fusion. Les images résultantes diffèrent d'une méthode de fusion à l'autre.

| | M_Fus | SD_Fus | H | AG | SF | CC | FS | MI |
|------|-----------------|----------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Fus1 | 160.5884 | 16.7997 | 3.9814 | 1.6879 | 4.4610 | 0.4556 | 1.779 | 1.8754 |
| Fus2 | 171.0934 | 26.9069 | 4.6446 | 3.1662 | 6.9069 | 0.5504 | 1.72206 | 2.3429 |
| Fus3 | 175.1959 | 35.4768 | 4.8527 | 3.5351 | 8.1748 | 0.7065 | 1.9397 | 2.4215 |
| Fus4 | 143.0027 | 49.4139 | 5.0495 | 1.8148 | 4.5729 | 0.5279 | 1.633 | 3.1201 |
| Fus5 | 137.7827 | 21.9887 | 4.1714 | 3.8159 | 8.6077 | 0.5352 | 1.6238 | 2.5948 |
| Fus6 | 129.0284 | 27.7357 | 4.4674 | 2.1507 | 5.3129 | 0.4761 | 1.7517 | 2.0973 |
| Fus7 | 183.4478 | 28.3213 | 4.6552 | 3.6811 | 8.1451 | 0.583 | 1.6937 | 2.464 |

Tableau (3.1) : Critères d'évaluation obtenus par les différentes méthodes de fusion sur les images de l'exemple 1

| | M_Fus | SD_Fus | H | AG | SF | CC | FS | MI |
|------|-----------------|----------------|---------------|---------------|--------------|---------------|---------------|---------------|
| Fus1 | 157.5675 | 19.8195 | 3.9554 | 2.1542 | 5.4806 | 0.4816 | 1.7769 | 1.8207 |
| Fus2 | 145.0804 | 29.9885 | 4.7532 | 2.9829 | 6.2009 | 0.5096 | 1.7442 | 2.4054 |
| Fus3 | 159.0844 | 34.979 | 4.8281 | 2.8138 | 6.6083 | 0.6994 | 1.977 | 2.7745 |
| Fus4 | 155.5514 | 35.8363 | 4.6265 | 2.208 | 5.4256 | 0.6022 | 1.9465 | 2.4182 |
| Fus5 | 140.8179 | 24.4853 | 4.4226 | 3.5975 | 7.902 | 0.5684 | 1.9435 | 2.2977 |
| Fus6 | 157.5762 | 32.3748 | 4.7082 | 2.2048 | 5.2954 | 0.5411 | 1.8879 | 2.4473 |
| Fus7 | 169.6303 | 25.9413 | 4.374 | 3.3137 | 7.3258 | 0.6287 | 1.9958 | 2.3135 |

Tableau (3.2) : Critères d'évaluation obtenus par les différentes méthodes de fusion sur les images de l'exemple 2

On remarque que la méthode de fusion qui a donnée des valeurs les plus élevées des paramètres de performances est la fusion 5 qui calcule le maximum des coefficients approximation et détails des deux images, suivi par la fusion 3 qui calcul la moyenne des coefficients approximation et de détails des deux images.

Chapitre 3 Fusion d'image de la Kinect à base de la transformée en ondelettes

Les tableaux (3.3) et (3.4) regroupent les valeurs des attributs de texture calculées sur les images test et fusionnées.

| | Contraste | Corrélation | Energie | Homogénéité |
|-------|-----------|-------------|---------|-------------|
| Depth | 22.9194 | 0.9871 | 0.0136 | 0.8181 |
| Gray | 12.3214 | 1 | 0.0159 | 0.3988 |
| Fus1 | 27.8925 | 0.9504 | 0.0124 | 0.7607 |
| Fus2 | 57.4481 | 0.9605 | 0.0028 | 0.5688 |
| Fus3 | 73.1575 | 0.9710 | 0.0013 | 0.5287 |
| Fus4 | 31.0685 | 0.9937 | 0.0029 | 0.6846 |
| Fus5 | 83.1266 | 0.914 | 0.004 | 0.5047 |
| Fus6 | 41.9888 | 0.9728 | 0.0054 | 0.6898 |
| Fus7 | 78.9894 | 0.9510 | 0.0017 | 0.5121 |

Tableau (3.3) : Attributs de textures obtenues sur les images de l'exemple 1

| | Contraste | Corrélation | Energie | Homogénéité |
|-------|-----------|-------------|---------|-------------|
| Depth | 110.902 | 0.979 | 0.0013 | 0.768 |
| Gray | 12.3214 | 1 | 0.0158 | 0.3988 |
| Fus1 | 45.3442 | 0.9427 | 0.0151 | 0.7312 |
| Fus2 | 50.8269 | 0.9718 | 0.0022 | 0.6038 |
| Fus3 | 51.1162 | 0.9792 | 0.0021 | 0.576 |
| Fus4 | 43.3977 | 0.9832 | 0.005 | 0.6967 |
| Fus5 | 78.6064 | 0.9349 | 0.0034 | 0.5205 |
| Fus6 | 41.4389 | 0.9804 | 0.0046 | 0.6875 |
| Fus7 | 68.3228 | 0.9494 | 0.0037 | 0.5402 |

Tableau (3.4) : Attributs de textures obtenues sur les images de l'exemple 2

On peut remarquer que la méthode de fusion n'influe pas sur l'attribut énergie et corrélation. De plus, ces attributs restent identiques à celui de l'image de profondeur.

Les méthodes de fusion préservent la corrélation

Quant à l'attribut homogénéité à des valeurs presque identiques, sauf la fusion 1 qui possède des valeurs proches de ceux de l'image de profondeur.

L'attribut de contraste présente des valeurs variables entre les images en niveaux de gris et des images de profondeur.

3.5 Conclusion

Dans ce chapitre, on a présenté la fusion d'image de la Kinect par la transformée en ondelettes qui se compose de 3 étapes. Dans la première étape, les images niveau de gris (couleur) et profondeur à fusionner sont décomposées en sous-images par la transformée en ondelettes. Dans la deuxième étape, les coefficients de la bande de basses fréquences et des bandes hautes fréquences sont combinés avec sept différentes stratégies. Dans la dernière étape, l'image fusionnée est construite par l'intermédiaire de la transformée en ondelettes inverse appliquée sur les coefficients composés.

L'évaluation objective de la fusion a été effectuée en utilisant des paramètres de performances et des attributs de texture. Les résultats montrent que la stratégie de la combinaison des coefficients approximation et détails influe sur ces critères.

Parmi ces sept stratégies utilisées, la fusion par moyenne des coefficients approximation et détails de l'image niveau de gris (couleur) et profondeur et la fusion par calcul du maximum des coefficients semblent les plus indiquées.

Conclusion générale

Conclusion générale

Ce travail nous a permis d'étudier le capteur Kinect de Microsoft, dans un premier temps nous avons présenté la caméra kinect, ses caractéristiques et son principe de fonctionnement.

Dans le but d'exploiter les images acquises par ce capteur à savoir les images couleur et images profondeur nous avons utilisée l'approche de fusion de ces images afin de créer une nouvelle image contenant le maximum d'informations afin d'améliorer leurs interprétations.

L'approche que nous avons adopté pour la fusion des images couleur et profondeur est la transformée en ondelette.

Les images d'entrée, sont premièrement, décomposées en sous images par la transformée en ondelette, par la suite les coefficients de la bande de basses fréquences et des bandes hautes fréquences sont combinés avec sept différentes stratégies. La dernière étape consiste à construire l'image fusionnée par l'intermédiaire de la transformée en ondelettes inverse qui est appliquée sur les coefficients composés.

A l'issue de ce travail, des outils d'évaluations appuyés par une analyse visuelle et une analyse quantitative nous ont permis de conclure sur notre approche.

Ainsi, les résultats obtenus montrent que la stratégie de la combinaison des coefficients approximation et détails influe considérablement sur les critères de performances et les attributs de texture.

Parmi les différentes d'ondelettes que nous avons utilisées, celle de Haar est la meilleure méthode qui nous a donné des images fusionnées proche de l'image originale. Cependant, le niveau de décomposition des images initiales, joue un rôle important sur la qualité de l'image fusionnée.

Enfin, on peut dire que notre travail présente un certain nombre d'améliorations et de perspectives qui peuvent être envisagées et ce en améliorant les techniques de fusion en introduisant de nouvelles méthodes. Ou encore, en tenant compte de la qualité des données acquises qui influe considérablement sur les résultats, on peut envisager en perspective l'utilisation de la Kinect pour Windows V2, cette nouvelle version de ce capteur a été lancé en 2014, elle s'appuie sur le principe de la mesure du temps de vol pour évaluer les distances

Conclusion générale

entre le système d'acquisition et la scène observée, contrairement à Kinect V1 qui s'appuie sur la lumière structurée.

Bibliographie

- [1] http://wn.com/kinect__fingertip_detection, “Kinect Fingertip Detection.”
- [2] Arras, K.O., Martinez, O., Burgard, M.W. Using boosted features for detection of people in 2D range scans. In: Proceedings of the IEEE International Conference on Robotics and Automation. 2007
- [3] Wojtaszak Jérôme. Rééducation de l'anticipation posturale par l'intermédiaire de l'outil « Kinect » au travers d'une étude de cas uniques à ligne de base multiple chez quatre enfants porteurs de troubles du spectre autistique .Mai 2013
- [4] Hangzhou. Depth camera in computer vision and computer graphics : an overview xiang Xueqin, PAN Zhigeng, TONG Jing. State Key Lab of Computer Aided Design and Computer Graphics, Zhejiang University, 310058, China
- [5] Bellotto.N, Hu.H,. Multisensor data fusion for joint peoples tracking and identification with a service robot. In: IEEE International Conference on Robotics and Biomimetics, ROBIO, pp. 1494–1499. 2007
- [6] C.BARILLOT. Fusion de donnée et Imagerie 3D en médecine. 1999
- [7] M. Sasikala and N. Kumaravel, “A comparative analysis of feature based image fusion methods,” *Information Technology Journal*, 6(8):1224- 1230, 2007.
- [8] A. Wang, H. Sun and Y. Guan, “The application of wavelet transform to multimodality medical image fusion” *Proc. IEEE International Conference on Networking, Sensing and Control (ICNSC)*, Ft. Lauderdale, Florida, 2006, pp.270-274.
- [9] N. OUARAB, M.ACHIR, N.KHETTAL, K.GOUMIRI et Y.SMARA. Application à la fusion et l'analyse multi-résolution et des méthodes floues pour la classification des images satellitaire. *Téledétection*, 2003, vol. 3, n° 1, p. 17–31.

Bibliographie

- [10] Shah, P., Merchant, S.N., Desai, U.B.: Fusion of surveillance images in infrared and visible band using curvelet, wavelet and wavelet packet transform. *Int. J. Wavelets. Multiresolution Inform. Process.* **8**(2), 271–292 (2010)
- [11] Shah, P., Srikanth, T.V., Merchant, S.N., Desai, U.B. ” A novel multifocus image fusion scheme based on pixel significance using wavelet transform’’. In: *Proceedings of Image, Video, and Multidimensional Signal Process (IVMSP)*, pp. 54–59, Aug (2011)
- [12] Haralick, R. M., K. Shanmugam, and I. H. Dinstein. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, **6**, 610–621. 1973
- [13] Statistical and structural approaches to textures, R. Haralick, *Proceedings IEEE*, vol.67, No5, 1979