

Reconstruction d'images irrégulièrement échantillonnées: Application à la reconstruction de vues intermédiaires et au codage vidéo

**Reconstruction d'images irrégulièrement échantillonnées:
Application à la reconstruction de vues intermédiaires et au
codage vidéo**

Carlos Vázquez



Université du Québec

Institut national de la recherche scientifique

INRS-Télécommunications

16, place du Commerce, Verdun

Québec, Canada, H3E 1H6

23 septembre 1999

Rapport technique de l'INRS-Télécommunications no. 99-22

Résumé

La reconstruction d'images à partir des échantillons sur une grille irrégulière a plusieurs applications dans le domaine du traitement numérique. Pour les méthodes de codage vidéo non récursives, avoir la possibilité de reconstruire des images à partir des échantillons sur une grille irrégulière peut devenir un grand pas. L'utilisation de précisions en dessous du pixel pour l'estimation du mouvement, permettra d'améliorer de façon considérable la performance de ces méthodes de codage. La reconstruction de vues intermédiaires d'images stéréoscopiques profitera aussi des progrès dans ce domaine. Une méthode de reconstruction d'images irrégulièrement échantillonnées aura de multiples applications dont l'ajustement de la parallaxe, le suivi de la position de l'observateur et le codage de vidéo stéréoscopique.

Dans le présent travail nous proposons une telle méthode de reconstruction. La méthode utilise un algorithme itérative pour obtenir des approximations à l'image cherchée. Un espace d'ondelettes est utilisé pour modéliser les images et la reconstruction est faite en projetant une approximation de haute résolution obtenue avec une méthode simple sur un espace de plus basse résolution dont on suppose appartient l'image originale. La possibilité d'utiliser des bancs de filtres pour calculer les coefficients de la transformation ondelettes fait que le processus soit rapide et efficace.

Des résultats préliminaires sont aussi présentés. Des expériences ont été faites avec des champs de mouvement synthétiques et avec des paires d'images stéréoscopiques pour la reconstruction des vues intermédiaires. Les résultats obtenus permettent de valider les possibilités d'application de l'algorithme, même s'il faut faire des améliorations pour son application.

Table des matières

1	Introduction	1
2	Reconstruction des images à partir des échantillons irrégulièrement espacés	3
2.1	La reconstruction des vues intermédiaires.	3
2.1.1	Approches	5
2.2	Le codage vidéo	8
2.3	Formulation du problème	11
3	Théorie de l'échantillonnage irrégulier	13
3.1	Bases d'ondelettes	14
3.2	Les splines pour la représentation des images	18
3.3	L'échantillonnage irrégulier	19
4	Méthode proposée	24
4.1	Modèle d'image	24
4.2	Échantillonnage irrégulier	26
4.3	Algorithme de reconstruction	27
4.3.1	Les occlusions et les zones nouvellement exposées	33
5	Résultats expérimentaux	35
5.1	Algorithme implanté	35
5.2	Images et champs utilisés	35
5.3	Résultats	37
6	Travail futur	41
7	Conclusions	44

Table des figures

1	Représentation des plans d'images des caméras	4
2	\mathbf{x}_G et \mathbf{x}_D représentent la projection sur I_G et I_D respectivement du point P , \mathbf{x}_α est la projection de P sur une caméra virtuelle à la position α et B représente la distance entre les axes des caméras	6
3	Estimation de disparité à la position α	7
4	Transformations en arrière (a) et en avant (b)	9
5	Trajectoires du mouvement en avant (a) et en arrière (b)	10
6	Banc de filtres utilisé pour l'analyse a) et la synthèse b) dans un espace d'ondelettes	17
7	Zones des points cachés a) et nouveaux b)	26
8	Échantillonnage irrégulier sur une grille plus dense	29
9	Opérateur S pour l'obtention de $I_r[\mathbf{m}_i]$ à partir de I	31
10	Diagramme de l'opérateur A	31
11	Diagramme de l'algorithme itératif	32
12	Images originales pour la compensation du mouvement	36
13	Champs de mouvement utilisés	36
14	Images gauche et droite de la paire stéréoscopique "flowers"	37
15	Reconstruction de l'image PCS après rotation	49
16	Reconstruction de l'image PCS après "zoom"	50
17	Reconstruction de l'image Bouquet après rotation	51
18	Reconstruction de l'image Bouquet après "zoom"	52
19	Reconstruction de l'image droite de Flower avec compensation de disparité	53
20	Reconstruction d'une vue intermédiaire ($\rho = 0.25$)	54

Liste des tableaux

1	PSNR des images reconstruites (db)	39
2	PSNR de l'image reconstruite (db)	39

1 Introduction

Dans les dernières années, avec le développement des techniques de télécommunications et l'expansion des réseaux d'ordinateurs et particulièrement de l'Internet, les progrès dans les champs de l'informatique, la construction des ordinateurs chaque fois plus puissantes et l'amélioration des techniques de visualisation des images; les travaux orientés vers le traitement numérique et codage d'images fixes et en mouvements ont connu aussi une croissance importante.

Les sujets visant l'amélioration des méthodes de codage vidéo et les problèmes reliés au vidéo tri-dimensionnel, en particulier, sont devenus des sujets de recherche très actifs. Dans la plupart des méthodes de codage vidéo utilisées aujourd'hui, l'information de mouvement est prise en compte pour améliorer l'efficacité de la méthode de codage à partir de la compensation du mouvement ou du filtrage compensé par le mouvement. La méthode la plus utilisée pour coder l'information temporelle des séquences vidéo est le DPCM, une méthode récursive qui est presque à la limite de sa performance. Pour de meilleurs résultats il faut se tourner vers des méthodes non récursives pour coder plusieurs images en utilisant les trajectoires du mouvement. De plus, il est bien connu [13], que pour améliorer la qualité de l'estimation du mouvement on doit la faire avec une précision en dessous du pixel. Si ces deux conditions sont mises ensemble, les images obtenues dans le processus de compensation seront des images échantillonnées sur des grilles non uniformes. Pour la visualisation ou le traitement, ces images doivent être échantillonnées sur une grille uniforme. Il faut donc reconstruire des images échantillonnées de façon irrégulière sur des grilles régulières. Jusqu'à présent, la plupart des travaux dans cette direction ont évité cette reconstruction en utilisant des précisions d'un pixel ou des méthodes récursives pour le codage.

On trouve cette même situation dans les applications d'images stéréoscopiques. Dans ce cas, l'information de profondeur est contenue dans la disparité entre les images gauche et droite. La compensation de la disparité est nécessaire pour plusieurs applications dans ce domaine. Encore une fois, les précisions en dessous du pixel pour l'estimation de la disparité pourrait nous conduire vers un échantillonnage irrégulier et la reconstruction à partir de ces échantillons est par conséquent nécessaire. La reconstruction des vues intermédiaires particulièrement est un sujet qui a reçu beaucoup d'attention à cause de ses multiples applications.

La plupart des travaux à ce sujet ont été orientés vers la reconstruction des fonctions limitées en bande et à déterminer les conditions que la grille d'échantillonnage doit respecter pour une reconstruction parfaite. Ils sont basés sur des algorithmes itératifs qui permettent d'obtenir de meilleures approximations à chaque itération.

Dans ce projet nous analysons cette situation et proposons une méthode de solution avec un algorithme itératif pour la reconstruction des images à partir des échantillons sur une grille irrégulière dans le but de l'appliquer au problème de la compensation du mouvement ou de la disparité. La méthode est basée sur un modèle d'ondelettes et utilise des bancs de filtres pour améliorer de façon successive une

première approximation de l'image désirée.

Dans un premier temps, nous présenterons le problème et les raisons pour l'étudier. Les théories d'échantillonnage irrégulier, les espaces d'ondelettes et les splines sont présentés dans la section 3 ainsi que les méthodes trouvées dans la littérature. Dans la section 4 on décrit la méthode proposée et à la section suivante nous présentons quelques résultats préliminaires dans la compensation du mouvement avec des champs synthétiques et la reconstruction des vues intermédiaires qui montrent les possibilités d'application de la méthode proposée. En dernier, nous verrons le travail qu'il nous reste à faire.

2 Reconstruction des images à partir des échantillons irrégulièrement espacés

L'échantillonnage irrégulier a été un sujet de recherche actif depuis longtemps. Dans le domaine du traitement des images, il y a plusieurs applications où l'échantillonnage irrégulier est présent. Les algorithmes pour la reconstruction d'images à partir de leurs échantillons sur une grille irrégulière deviennent donc essentiels.

Nous présenterons deux applications de ce type pour montrer l'importance de travailler sur ce sujet et pour présenter le problème et ses principales caractéristiques: la reconstruction des vues intermédiaires et le codage vidéo.

2.1 La reconstruction des vues intermédiaires.

Les systèmes de vision stéréoscopiques ont connu des avancements significatifs au cours des dix dernières années, en facilitant l'apparition de multiples applications dans les domaines de la médecine, de la télé-opération et de l'industrie de l'amusement (entertainment industry). Les systèmes de type analogiques ont dominé le champ d'application des images stéréoscopiques, mais la migration vers les systèmes numériques est inévitable étant donné les avantages pour le traitement, le stockage et la transmission.

Dans le domaine numérique on trouve des nouveaux défis, mais aussi des nouvelles possibilités d'amélioration de la perception de la profondeur, de la commodité des observateurs pendant la visualisation, et la création des applications impossibles d'imaginer dans le champ analogique tel que les effets spéciaux par ordinateurs à partir des images réelles.

Le système de vision stéréoscopique est basé sur le fait que notre vision est formée par deux images: une obtenue par l'oeil gauche et l'autre par l'oeil droit, de manière à ce que la représentation en 3D de la scène soit formée par le cerveau en utilisant la différence de position entre les points et les objets homologues des deux images (disparité). La différence de position des points dans les images qu'on voit, définit alors la sensation de profondeur.

Pour obtenir des images stéréoscopiques, on utilise une paire d'images prises sur la même scène avec deux caméras dont les caractéristiques optiques sont pareilles et leurs axes optiques sont parallèles ou convergents. Dans la Fig. 1 on voit un schéma qui représente les plans d'images de deux caméras dans le cas des caméras avec des axes parallèles [2].

Avec cette géométrie chaque point de la scène qui est capturé par les caméras a une projection dans le plan d'image correspondant. Les coordonnées de ces projections sont différentes pour les deux plans. Le vecteur obtenu par la différence de position, lorsqu'on fait coïncider les deux plans, est appelé vecteur de disparité.

Un sujet particulièrement intéressant est la reconstruction des vues intermédiaires, c'est-à-dire des images qui seraient obtenues par une caméra virtuelle située entre les caméras gauche et droite. Une telle reconstruction serait très utile pour plusieurs

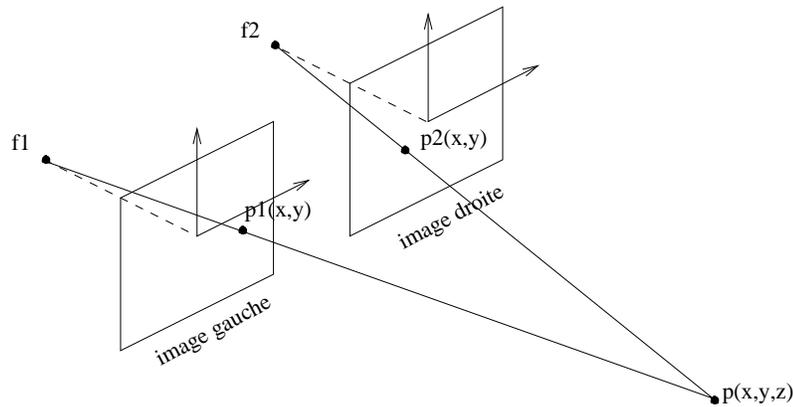


FIG. 1 – Représentation des plans d'images des caméras

applications:

1. *Ajustement de la parallaxe des caméras*[16]: La distance entre les axes optiques des caméras définit la parallaxe de la paire d'images stéréoscopiques. Si cette distance est très différente de celle qu'il y a entre nos yeux, on peut avoir des difficultés à fusionner les deux images dans une seule scène 3D, ou ne pas percevoir de profondeur. Les différences entre personnes sont significatives [30] et le système de visualisation est aussi important. On peut ajuster la parallaxe au moment de l'acquisition des images, mais ce n'est possible que si on visualise au même moment de l'acquisition, ce qui n'est pas le cas de façon générale. Une meilleure solution est de trouver à partir des images déjà acquises, l'image qu'on obtiendrait dans une caméra virtuelle située au bon endroit, c'est-à-dire reconstruire une vue différente de celles qu'on a. L'ajustement de la parallaxe deviendra essentiel pour la transmission de TV-3D où il y a un seul transmetteur et plusieurs récepteurs avec des différents systèmes de visualisation.
2. *Suivi de la position de l'observateur*: Lorsque l'observateur change légèrement de position en face d'un système de visualisation 3D son champ de vision change, mais si on visualise toujours les mêmes images, ce changement de position n'aura pas l'effet désiré. Il faudra suivre avec les caméras les changements de position de l'observateur, mais une fois encore ce n'est pas possible de façon générale. La solution est de reconstruire des vues sur des caméras virtuelles.
3. *Codage des images stéréoscopiques*: Étant donné le haut niveau de redondance qui existe entre les images gauche et droite d'une paire d'images stéréoscopiques il n'est pas nécessaire de transmettre ou stocker les deux images. Généralement, on code la paire avec une des images et un champ de disparité pour ne reconstruire la deuxième image qu'au moment de la visualisation. Il faut alors reconstruire une image dans une position qui peut ou non être la même que l'originale.

2.1.1 Approches

Nous avons vu que pour plusieurs applications, la manipulation des caméras virtuelles est nécessaire. Mais comment le faire?

On peut classifier les méthodes utilisées pour obtenir des images provenant des caméras virtuelles en deux grands groupes [16]:

1. Celles basées sur un modèle 3D de la scène [4]. L'idée de base est de construire une représentation 3D de la scène en utilisant les images gauche et droite, et ensuite projeter cette représentation sur une caméra virtuelle située à l'endroit où on veut reconstruire l'image. Cette méthode est flexible parce qu'elle permet que les caméras virtuelles et réelles soient positionnées de façon arbitraire, mais elle est très limitée du point de vue de la complexité de la scène et demande normalement des caméras calibrées.
2. Celles basées sur des algorithmes de traitement de signaux 2D. Ces méthodes n'essaient pas de construire une représentation 3D de la scène, mais elles essaient de reconstruire l'image désirée à travers le traitement de signaux 2D. La façon classique consiste à estimer un champ de disparité pour établir une correspondance entre les images gauche et droite et faire ensuite une interpolation compensée en disparité. Cette méthode fonctionne mieux pour des scènes plus complexes, lorsqu'elle n'est pas basée sur un modèle 3D. La méthode fonctionne mieux pour des disparités limitées et pour des caméras avec des axes qui ne sont pas très séparés (près de la distance entre les yeux) et presque parallèles.

Pour les applications en TV-3D ou en photographie-3D, les images sont acquises avec des caméras proches (distance \approx distance inter-oculaire humaine) et avec des axes presque parallèles. Ces conditions en plus du fait de que les images sont assez complexes, nous amènent à la conclusion que la méthode basée sur le traitement de signaux est plus pratique [16].

On a vu que la méthode basée sur le traitement de signaux a comme idée de base l'estimation de la disparité entre les images gauche et droite pour faire une interpolation compensée en disparité à la position qu'on désire avoir l'image. Si on désigne l'image gauche comme $I_G(\mathbf{x})$, l'image droite comme $I_D(\mathbf{x})$ avec $\mathbf{x} \in \mathcal{R}^2$, B la distance entre les axes optiques des caméras ($B = 1$, pour simplicité des équations) et P un point sur la scène 3D, on peut définir (comme on peut voir à la Fig. 2 pour le cas des axes parallèles) la disparité $\mathbf{d} = \mathbf{x}_G - \mathbf{x}_D$ où $\mathbf{x}_{G(D)}$ est la position de la projection du point P sur l'image gauche(droite).

Il faut préciser quelques caractéristiques des images stéréoscopiques, de l'estimation de disparité et de la géométrie des caméras utilisés: [17]

1. Les caméras ont des axes presque parallèles et alignés, alors les disparités verticales sont petites et la projection des points sur des caméras virtuelles est située sur la ligne qui passe par les projections sur les caméras réelles Fig. 2.

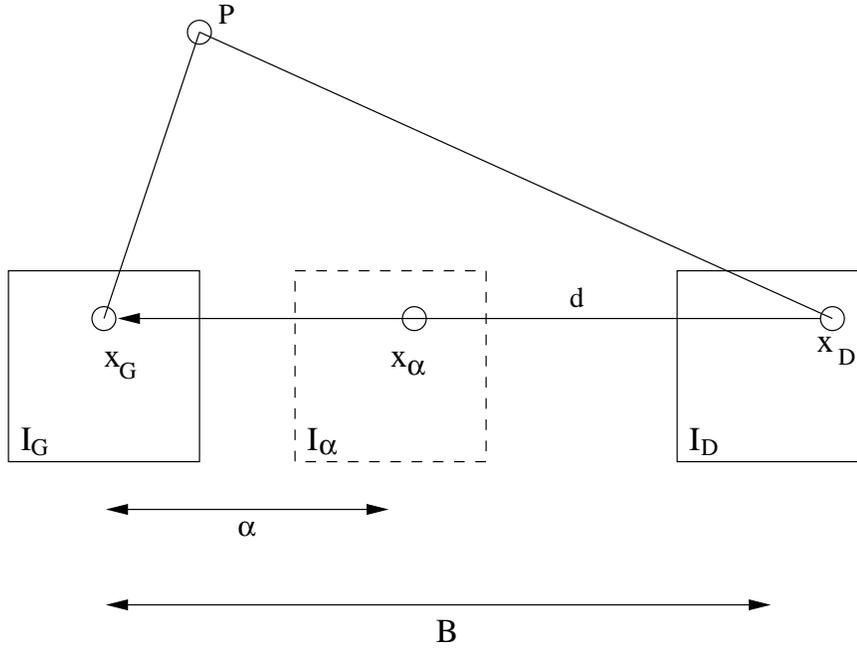


FIG. 2 – \mathbf{x}_G et \mathbf{x}_D représentent la projection sur I_G et I_D respectivement du point P , \mathbf{x}_α est la projection de P sur une caméra virtuelle à la position α et B représente la distance entre les axes des caméras

2. Les points homologues sur les images gauche et droite sont reliés par la disparité: $\mathbf{x}_G = \mathbf{x}_D + \mathbf{d}$ et en considérant l'énoncé antérieur, $\mathbf{x}_\alpha = \mathbf{x}_D + (1 - \alpha)\mathbf{d}$, avec $\alpha \in [0, 1]$, pour $\alpha = 0(1)$, $\mathbf{x}_\alpha = \mathbf{x}_{G(D)}$.
3. Les points homologues ont des intensités similaires, $I_G(\mathbf{x}_G) \approx I_D(\mathbf{x}_D) \approx I_\alpha(\mathbf{x}_\alpha)$. Ce n'est pas exactement vrai, parce que les caméras ne sont pas identiques et les changements de position pourraient générer des changements d'illumination sur les mêmes points de la scène, mais de façon générale la condition est respectée.

Ces considérations nous permettent de reconstruire la projection d'un point à la position α à partir de ces projections aux positions \mathbf{x}_G et \mathbf{x}_D en utilisant la disparité estimée et les valeurs d'intensité dans les images connues:

$$I_\alpha(\mathbf{x}_\alpha) = f(I_G(\mathbf{x}_\alpha + \alpha\mathbf{d}), I_D(\mathbf{x}_\alpha - (1 - \alpha)\mathbf{d})) \quad (1)$$

où f est une certaine fonction d'interpolation, qui pourrait être linéaire comme dans [17] ou non linéaire comme dans [21].

Jusqu'à présent nous n'avons rien dit sur les positions \mathbf{x}_α et sur l'estimation du vecteur \mathbf{d} . Lorsqu'on est dans le domaine numérique, on veut reconstruire une image I_α échantillonnée sur une grille régulière. \mathbf{x}_α doit alors appartenir à une grille régulière et le vecteur \mathbf{d} doit donc passer par la grille régulière à la position α . Dans ce cas, les extrêmes du vecteur \mathbf{d} ne se retrouveront pas nécessairement sur la grille régulière aux

images I_G et I_D . Il faudra donc calculer les valeurs à ces extrêmes avec un processus d'interpolation. On estime les vecteurs de disparité en faisant "bouger" les vecteurs aux points \mathbf{x}_α (Fig. 3). Dans cette figure les cercles représentent les positions sur la grille régulière, les gros points noirs représentent les échantillons des images gauche et droite et les cercles blancs les échantillons sur la grille régulière aux position α et ρ .

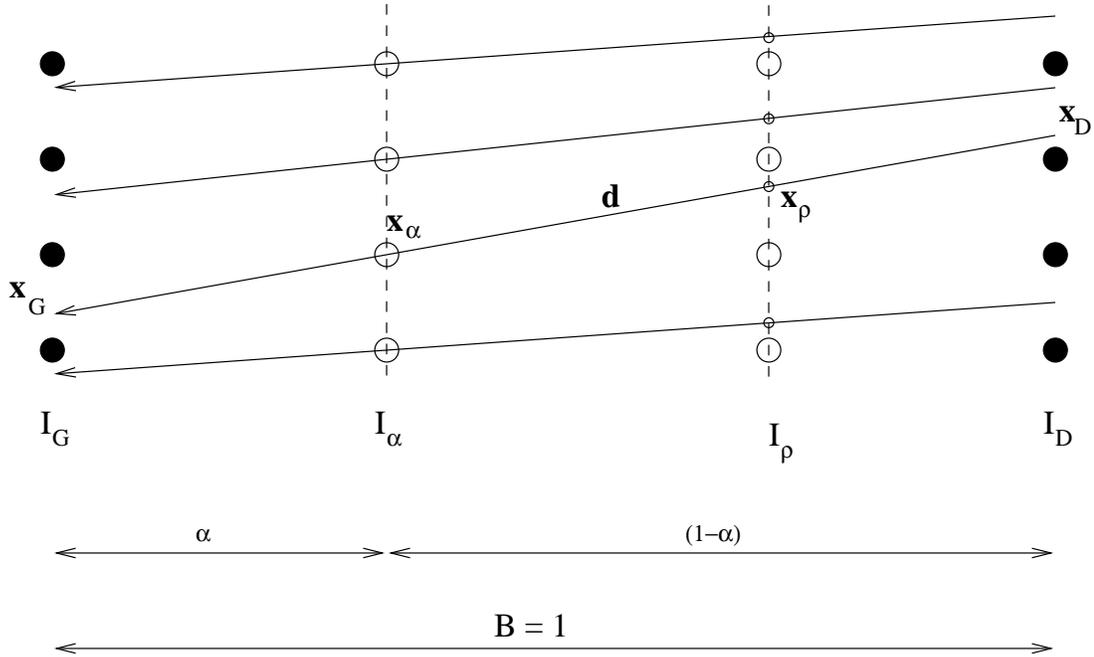


FIG. 3 – Estimation de disparité à la position α

La contrainte de calculer le champ de disparité en faisant bouger les vecteurs aux points de la grille régulière à la position α ne constitue pas un problème lorsqu'on essaie de reconstruire une seule vue et qu'on peut faire l'estimation de disparité au moment de la reconstruction. Mais ces conditions ne sont pas toujours satisfaites. Si on veut reconstruire plusieurs vues à partir d'une seule paire d'images stéréoscopiques, il faudra estimer un champ de disparité pour chaque position à reconstruire afin de garantir que les vecteurs passent par les positions de la grille d'échantillonnage. Autrement, si on fait l'estimation de disparité à la position α et on veut reconstruire à la position $\rho \neq \alpha$ (Fig. 3), le nouveau point \mathbf{x}_ρ ne tombera pas sur la grille régulière et il sera donc nécessaire de reconstruire l'image à partir des points qui n'appartiennent pas à la grille régulière.

Ce problème se présente souvent dans le cas de reconstruction de vues intermédiaires, et non pas seulement lorsqu'on veut reconstruire plusieurs vues. Dans le cas du codage d'une séquence vidéo stéréoscopique pour la transmettre à certains récepteurs, l'information de disparité est utilisée pour réduire la redondance de l'information à transmettre en réduisant la quantité de bits utilisé pour représenter les images [27].

Ce qui est alors transmis est une image, le champ de disparité et peut-être de l'information sur l'erreur d'estimation. Si on veut faire un ajustement de la parallaxe au récepteur, on voudrait utiliser le champ de disparité qui a été transmis. Cependant, avec ce champ, il arrive que les points obtenus ne soient pas sur la grille régulière.

On envisage alors trois solutions:

1. Estimer un champ de disparité pour chaque position à reconstruire. C'est une solution très coûteuse si le nombre de vues à reconstruire est élevé. On pourra dire que celle-ci est une solution optimale [17].
2. La solution proposée par Konrad [17]. Utiliser un seul champ de disparité mais au lieu de faire une compensation de disparité propager les vecteurs d'une position à une autre suposant que si les vecteurs sont semblables dans une voisinage spatial, ils doivent l'être également pour une voisinage de α . Ce modèle est valide à l'intérieur des objets mais n'est pas exact au frontières.
3. Utiliser un seul champ de disparité et faire une compensation de disparité. L'interception du champ de disparité avec le plan d'image qu'on veut reconstruire nous donne une représentation de l'image, mais de façon générale cette représentation n'est pas sur un grille régulière et il est donc nécessaire d'utiliser une méthode d'interpolation non uniforme.

La troisième solution devrait donner des meilleures résultat que la deuxième au prix d'une plus grande complexité de calcul. Cependant la complexité de calcul ne sera pas si grande que pour la première solution. On se propose donc d'aborder le problème de la *reconstruction des vues intermédiaires à n'importe quelle position en utilisant un seul champ de disparité et reconstruire l'image à partir de sa représentation sur une grille irrégulière.*

2.2 Le codage vidéo

On a vu que pour la reconstruction des vues intermédiaires en paires d'images stéréoscopiques il faut reconstruire des images à partir des échantillons sur une grille irrégulière. On verra dans la suite que dans certaines méthodes de codage vidéo, ce problème peut aussi apparaître.

Le codage de l'information temporelle d'une séquence vidéo est fait en estimant la transformation spatiale subie par une image afin de prédire l'image suivante. Cette transformation spatiale contient l'information de mouvement, où se trouve la plupart de la nouvelle information dans les images d'une séquence. Les processus d'estimation du mouvement et de la compensation du mouvement jouent alors un rôle très important dans le codage vidéo. [31]

La compensation du mouvement (CM) est un processus de prédiction dans lequel l'image actuelle est prédite à partir d'une ou plusieurs images connues de la séquence. De façon générale cette approche aide à réduire la quantité de données nécessaire pour représenter la séquence par rapport à une approche dans laquelle on n'utilise

que de l'information spatiale [13] et elle est basée sur l'estimation de la transformation spatiale qui définit le mouvement dans la séquence.

Cette transformation peut être décrite de deux façons: comme une transformation en avant ou comme une transformation en arrière. La dernière est la plus utilisée car elle est plus facile à mettre en oeuvre lorsqu'on peut estimer une transformation pour chaque image à reconstruire. Elle est basée sur l'identification du point homologue de chaque point de l'image à reconstruire dans une image connue de la séquence. Ce point homologue pourrait ne pas se situer sur la grille régulière et pour le trouver il est souvent nécessaire d'interpoler. Par contre, la première technique mentionnée a été beaucoup moins étudiée [25]. Le problème principal est relié au fait que l'image résultante de la transformation spatiale est généralement formée par des échantillons repartis non uniformément, ce qui génère des difficultés pour sa reconstruction sur une grille régulière.

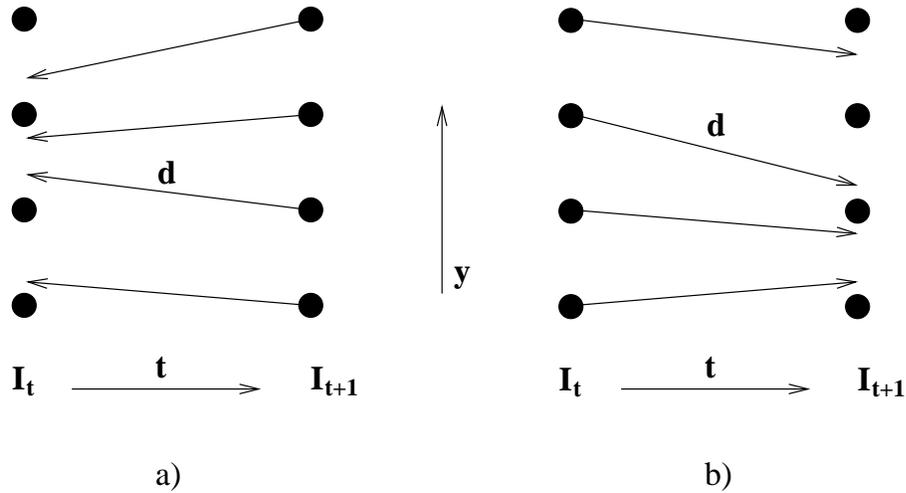


FIG. 4 – Transformations en arrière (a) et en avant (b)

Dans le cas de la transformation en arrière, (Fig. 4 a) la correspondance s'établit entre une grille régulière où l'image à estimer $\tilde{I}_{t+1}(\mathbf{n})$, $\mathbf{n} = (u, v)^T \in Z^2$ est définie et une grille continue où l'image source $I_t(\mathbf{x})$, $\mathbf{x} = (x, y)^T \in \mathfrak{R}^2$ est considérée définie. On construit alors les points dans l'image à estimer à partir des points dans une image connue de la façon suivante:

$$\tilde{I}_{t+1}(\mathbf{n}) = \phi(I_t(f(\mathbf{n}))) \quad (2)$$

où $f : Z^2 \rightarrow \mathfrak{R}^2 \mid \mathbf{x} = f(\mathbf{n})$ est la fonction qui définit la transformation spatiale et ϕ est une fonction qui définit les changements d'intensité le long de la trajectoire du mouvement. Dans ce cas $I(\mathbf{x})$ est obtenu par interpolation parce que I est définie sur Z^2 . Alors pour reconstruire les points dans l'image estimée il suffit d'interpoler l'image originale pour trouver la valeur de la fonction dans les points transformés \mathbf{x} .

Pour atteindre cet objectif on pourra utiliser n'importe quel schéma d'interpolation connu.

Par contre, dans le cas de la transformation en avant représentée à la Fig. 4 b, la correspondance s'établit entre une grille régulière où l'image connue ou de référence $I_t(\mathbf{n})$ est définie, et un plan continu où l'image à estimer $\tilde{I}_{t+1}(\mathbf{x})$ est considérée définie. Les points de l'image à estimer sont alors obtenus à travers de la transformation φ qui définit les changements d'intensité le long des trajectoires de mouvement et de la transformation spatiale g qui définit le mouvement proprement dit:

$$\tilde{I}_{t+1}(\mathbf{x}) = \varphi(I_t(g(\mathbf{x}))) \quad (3)$$

où $g : \mathfrak{R}^2 \rightarrow Z^2 \mid \mathbf{n} = g(\mathbf{x})$. Dans ce cas pour trouver les valeurs d'intensité de l'image à estimer, il faut interpoler les points non uniformément repartis pour connaître les valeurs aux points \mathbf{n} .

Évidemment ces deux types de transformations peuvent être étendus à l'estimation des trajectoires du mouvement sur plusieurs images dans une séquence vidéo. Pour le faire, il faudra relâcher la contrainte qui suppose qu'une des grilles, celle de départ ou celle d'arrivée est régulière, parce que la grille d'arrivée d'une transformation devient la grille de départ de la suivante. On peut voir à la Fig. 5 une trajectoire formée par des transformations en avant (a) et une autre formée par des transformations en arrière (b). Clairement pour l'estimation de la transformation entre les images en $t+1$ et $t+2$, les deux grilles, celle de départ et celle finale, sont formées par des points irrégulièrement espacés. Si on utilise pour le calcul une précision en dessous du pixel, ce qui est le cas si on désire avoir une bonne estimation [13] et qui est représenté à la Fig. 5, les points ne seront pas sur la grille régulière. Si la précision utilisée est un pixel, on peut avoir un cas spécial d'échantillonnage irrégulier, celui où il y a des points de la grille régulière qui ne sont pas sur la grille finale.

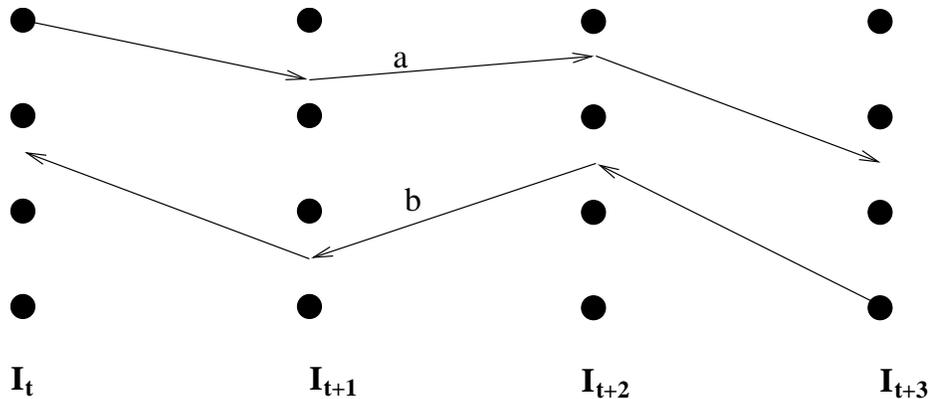


FIG. 5 – Trajectoires du mouvement en avant (a) et en arrière (b)

L'intérêt pour les trajectoires de mouvement vient du fait que les méthodes de codage vidéo les plus utilisées à présent sont de type récursif par rapport à l'information temporelle. La technique la plus utilisée est le codage DPCM ("Differential Pulse

Code Modulation”) qui utilise la dernière image décodée pour reconstruire l’image actuelle. Avec cette méthode, les erreurs de reconstruction (dues à la quantification pour la transmission) sont cumulées, affectant l’efficacité du codage.

Dans les dernières années quelques méthodes alternatives non récursives ont été proposées, parmi lesquelles on trouve les travaux de Choi et Woods [6] et Ohm [26] en codage sous-bande 3D avec compensation du mouvement et ce de Foodeei [12] en quantification vectorielle des trajectoires de mouvement.

Ces travaux profitent de la redondance le long des trajectoires du mouvement et ouvrent une nouvelle étape pour le codage vidéo. Mais comme on a déjà vu, les trajectoires de mouvement pour être précises doivent avoir une précision en dessous du pixel, ce qui rend la reconstruction plus difficile parce que l’image doit être récupérée à partir des échantillons sur une grille irrégulière. La solution à ce problème est souvent d’éviter les précisions en bas d’un pixel pour l’estimation de mouvement [7], mais les résultats sont dominés par la qualité de l’estimation de mouvement qui introduit une erreur qui fait tomber l’efficacité de la méthode.

D’autres solutions possibles comprennent l’interpolation des échantillons non uniformément repartis en utilisant un des algorithmes d’interpolation connus tel qu’une interpolation d’ordre zéro (Nearest Neighbor) ou une interpolation linéaire. Des interpolations d’ordre supérieur sont beaucoup plus difficiles étant donné l’irrégularité des échantillons. Des interpolateurs non linéaires ont été aussi proposés [29].

La reconstruction des images à partir des échantillons non uniformément repartis demeure un sujet de recherche actif où il reste encore un long chemin à parcourir. Nous allons nous concentrer alors sur le problème de la reconstruction des images compensées en mouvement (ou en disparité). La solution à ce problème a plusieurs applications pratiques immédiates dans les domaines du codage vidéo non récursif et dans le domaine de l’interpolation des images stéréoscopiques pour l’obtention des nouvelles vues.

2.3 Formulation du problème

On a vu qu’on doit reconstruire une image à partir d’un ensemble d’échantillons irrégulièrement espacés, obtenus comme résultat d’un processus de compensation du mouvement ou de la disparité.

Soit $\Gamma = \{\mathbf{x}_i, i \in [0, P-1]\}$ la grille irrégulière obtenue à partir de la compensation du mouvement ou disparité (désormais “compensation”), où P est le nombre de points dans la grille. On a alors $\{I(\mathbf{x}_i), \mathbf{x}_i \in \Gamma\}$, l’image définie sur la grille irrégulière, $\{I(\mathbf{n}), n \in \Lambda\}$ la vraie image qu’on veut reconstruire (inconnue) définie sur la grille régulière Λ . Notre objectif est de trouver une approximation $\{\tilde{I}(\mathbf{n}), n \in \Lambda\}$ le plus proche possible de l’image originale I .

Si on utilise une mesure de distance $dist$ définie sur l’espace auquel appartiennent les images, on a:

$$\tilde{I} = arg \min_{\tilde{I}}(dist(I(\mathbf{n}), \tilde{I}(\mathbf{n}))) \quad (4)$$

Trouver une telle approximation n'est pas une tâche facile. Notre objectif se réduit alors à trouver une approximation qui soit suffisante pour les applications auxquelles nous sommes intéressés, en tenant compte des caractéristiques des images et des grilles irrégulières pour ces applications.

Un modèle pour les images, un modèle pour la grille irrégulière et un algorithme de reconstruction doivent être définis au cours de ce travail. Une estimation de l'erreur finale et des comparaisons avec d'autres méthodes doivent aussi être faites. Enfin, un système pour la reconstruction des vues intermédiaires ou de codage vidéo avec l'utilisation de cette algorithme doit être développé pour valider les résultats.

3 Théorie de l'échantillonnage irrégulier

La reconstruction des images à partir des échantillons sur une grille irrégulière est un problème peu traité dans la littérature. Les systèmes qui font la reconstruction des images par compensation de façon générale évitent d'utiliser des échantillons irrégulièrement espacés en faisant la compensation pour les points d'une grille régulière. Les raisons pour le faire sont évidentes: le processus est beaucoup plus simple, bien connu et la théorie d'échantillonnage de Shannon donne une base théorique très forte. Mais cette méthode est très limitée par rapport à ces possibilités. Il y a plusieurs applications où il n'est pas possible de l'utiliser, par conséquent une reconstruction à partir des échantillons sur une grille irrégulière est nécessaire.

Quelques travaux ont été dédiés à résoudre ce problème et surtout le problème plus général de l'échantillonnage irrégulier et la reconstruction des fonctions non uniformément échantillonnées.

Depuis longtemps, le modèle qui considère les images comme des échantillons d'une fonction continue, limitée en bande et échantillonnée selon le critère de Nyquist est le plus utilisé et amplement accepté. Ce modèle est basé sur la très bien connue théorie d'échantillonnage de Shannon qui décrit l'équivalence entre une fonction limitée en bande et ses échantillons réguliers pris à une fréquence plus grande ou au moins égale à celle de Nyquist [28].

Mais même si cette théorie a eu un impact très important dans le domaine de la représentation et le traitement des images, on retrouve plusieurs problèmes dans son application [33]:

1. La fonction de base ($sinc(x)$) a une décroissance très lente.
2. L'élimination des hautes fréquences par le filtrage passe-bas idéal génère des oscillations de Gibbs, qui ne sont pas désirées en images.
3. La condition de limitation en bande est incompatible avec l'idée des signaux avec support fini.
4. Le filtre passe-bas idéal, qui est nécessaire pour la reconstruction, n'est pas réalisable.

Les deux derniers problèmes pourraient être résolus avec certaines conditions imposés aux signaux, mais les deux premiers sont impossibles d'éviter sans changer les fonctions de base. C'est à partir de cette conclusion que des études ont été faites pour trouver des espaces de fonctions plus pratiques pour la représentation et le traitement des images.

Les espaces d'ondelettes sont une alternative aux modèles des fonctions limitées en bande. Ils ont trouvé des applications dans le codage, l'amélioration et plusieurs autres sujets dans le domaine des images numériques.

Dans cette section on verra les bases d'ondelettes qui peuvent être utilisés pour la représentation des images, les espaces de splines qui conduisent à la construction de

bases d'ondelettes avec caractéristiques de régularité très utiles pour les images et les travaux en échantillonnage irrégulier et reconstruction des fonctions à partir de ses échantillons sur une grille irrégulière.

3.1 Bases d'ondelettes

Les bases d'ondelettes ont été de plus en plus utilisées au cours des dernières années dans le domaine du traitement numérique des images. Les applications au codage, l'amélioration, la segmentation et plusieurs autres sujets ont été étudiées. Leurs caractéristiques d'être reliées à l'approximation multirésolution permet d'obtenir des approximations à différentes résolutions d'une fonction continue et d'énergie finie.

On peut obtenir une base de l'espace $L^2(\mathfrak{R})$ à partir de la famille des fonctions obtenues par translation et dilatation d'une "ondelette" ψ :

$$\left\{ \psi_{j,n}(t) = \frac{1}{\sqrt{2^j}} \psi \left(\frac{t - 2^j n}{2^j} \right) \right\}_{(j,n) \in Z^2} \quad (5)$$

Les dilatations de la fonction ψ par 2^j contiennent l'information relative aux variations à la résolution 2^{-j} . Une approximation multirésolution d'une fonction $f \in L^2(\mathfrak{R})$ à la résolution 2^{-j} est définie comme la projection orthogonale de f sur l'espace $V_j \subset L^2(\mathfrak{R})$. Cet espace fait partie de l'ensemble $\{V_j\}_{j \in Z}$ de sous-espaces fermés de $L^2(\mathfrak{R})$ qu'on appelle approximation multirésolution et qui doit satisfaire les propriétés suivantes: [19], Ch. VII.

$$\forall (j, k) \in Z^2, \quad f(t) \in V_j \Leftrightarrow f(t - 2^j k) \in V_j \quad (6)$$

$$\forall j \in Z, \quad V_{j+1} \subset V_j \quad (7)$$

$$\forall j \in Z, \quad f(t) \in V_j \Leftrightarrow f\left(\frac{t}{2}\right) \in V_{j+1} \quad (8)$$

$$\lim_{j \rightarrow +\infty} V_j = \bigcap_{j=-\infty}^{+\infty} V_j = 0 \quad (9)$$

$$\lim_{j \rightarrow -\infty} V_j = \text{Closure} \left(\bigcup_{j=-\infty}^{+\infty} V_j \right) = L^2(\mathfrak{R}) \quad (10)$$

$$\exists \theta \mid \{\theta(t - n)\}_{n \in Z} \text{ est une base Riesz de } V_0 \quad (11)$$

Il est toujours possible de construire une base orthonormée $\{\phi_{j,n}\}_{n \in Z}$ pour chaque espace V_j à partir de θ , avec:

$$\phi_{j,n} = \frac{1}{\sqrt{2^j}} \phi \left(\frac{t - 2^j n}{2^j} \right) \quad (12)$$

La fonction ϕ qui est utilisé pour générer les bases est appelée fonction d'échelle.

Alors l'approximation de $f \in L^2(\mathfrak{R})$ à la résolution 2^{-j} est calculée en utilisant cette base orthonormée:

$$f_j = P_{V_j} f = \sum_{n=-\infty}^{+\infty} \langle f, \phi_{j,n} \rangle \phi_{j,n} \quad (13)$$

et le produit scalaire:

$$a_j[n] = \langle f, \phi_{j,n} \rangle \quad (14)$$

représente une approximation numérique de f à la résolution 2^{-j} .

Une approximation multirésolution est complètement caractérisée par la fonction d'échelle ϕ qui génère une base orthonormée de chaque espace V_j , et celle-ci est spécifié par le filtre numérique $h[n]$ appelé filtre miroir conjugué qu'on trouve à l'équation d'échelle:

$$\frac{1}{\sqrt{2}} \phi\left(\frac{t}{2}\right) = \sum_{n=-\infty}^{+\infty} h[n] \phi(t-n) \quad (15)$$

avec:

$$h[n] = \left\langle \frac{1}{\sqrt{2}} \phi\left(\frac{t}{2}\right), \phi(t-n) \right\rangle. \quad (16)$$

D'autre part, si on dénote W_j comme le complément orthogonal de V_j dans V_{j-1} ,

$$V_{j-1} = V_j \oplus W_j \quad (17)$$

la projection de f sur V_{j-1} peut être décomposé comme:

$$P_{V_{j-1}} f = P_{V_j} f + P_{W_j} f \quad (18)$$

et le complément $P_{W_j} f$ contient les détails de f qui apparaissent à la résolution 2^{-j+1} mais qui disparaît à la résolution 2^{-j} .

On peut construire une base orthonormée de l'espace W_j de la même façon que pour la fonction d'échelle et définir un filtre qui caractérise la fonction ψ qui génère la base et qu'on appelle ondelette:

$$\psi_{j,n} = \frac{1}{\sqrt{2^j}} \psi\left(\frac{t - 2^j n}{2^j}\right) \quad (19)$$

Alors pour chaque résolution 2^{-j} , l'ensemble $\{\psi_{j,n}\}_{n \in \mathbb{Z}}$ est une base orthonormée de l'espace W_j et la projection orthogonale de f sur cet espace sera:

$$P_{W_j} f = \sum_{n=-\infty}^{+\infty} \langle f, \psi_{j,n} \rangle \psi_{j,n} \quad (20)$$

et le produit scalaire:

$$d_j[n] = \langle f, \psi_{j,n} \rangle \quad (21)$$

permet de calculer les coefficients d'ondelettes $d_j[n]$ qui constituent une représentation numérique des nouveaux détails de f à la résolution 2^{-j} .

On a aussi l'équation:

$$\frac{1}{\sqrt{2}}\psi\left(\frac{t}{2}\right) = \sum_{n=-\infty}^{+\infty} g[n]\phi(t-n) \quad (22)$$

avec:

$$g[n] = \left\langle \frac{1}{\sqrt{2}}\psi\left(\frac{t}{2}\right), \phi(t-n) \right\rangle. \quad (23)$$

Les filtres h et g sont reliés par:

$$g[n] = (-1)^n h[1-n] \quad (24)$$

et ils jouent un rôle essentiel dans le calcul de la transformée rapide en ondelettes. En fait, ces filtres sont utilisés dans un banc de filtres pour obtenir la transformée en ondelette et la transformée inverse.

Si on dénote $\bar{h}[n] = h[-n]$ et

$$\check{x}[n] = \begin{cases} x[p] & n = 2p \\ 0 & n = 2p + 1 \end{cases} \quad (25)$$

le processus d'analyse sera:

$$a_{j+1}[p] = \sum_{n \in \mathbb{Z}} h[n-2p]a_j[n] = a_j * \bar{h}[2p] \quad (26)$$

$$d_{j+1}[p] = \sum_{n \in \mathbb{Z}} g[n-2p]a_j[n] = a_j * \bar{g}[2p] \quad (27)$$

et le processus de synthèse:

$$a_j[p] = \sum_{n \in \mathbb{Z}} h[n-2p]a_{j+1}[n] + \sum_{n \in \mathbb{Z}} g[n-2p]d_{j+1}[n] \quad (28)$$

$$= \check{a}_{j+1} * h[n] + \check{d}_{j+1} * g[n] \quad (29)$$

La théorie des bancs de filtres a été assez bien étudiée et les conditions pour avoir une reconstruction parfaite sont bien connues. Si on a un banc de filtres à deux canaux avec filtres d'analyse $\bar{h}[n] = h[-n]$ et $\bar{g}[n] = g[-n]$ et filtres de reconstruction duals \tilde{h} et \tilde{g} , et si les filtres ont une réponse impulsionnelle finie, la condition pour une reconstruction parfaite devient:

$$g[n] = (-1)^{1-n} \tilde{h}[1-n] \quad (30)$$

$$\tilde{g}[n] = (-1)^{1-n} h[1-n] \quad (31)$$

Les ensembles $\{\tilde{h}[n-2l], \tilde{g}[n-2l]\}$ et $\{h[n-2l], g[n-2l]\}$ sont des bases Riesz biorthogonales de l'espace $l^2(\mathbb{Z})$ et une cascade infinie de ces filtres permet de construire des bases biorthogonales d'ondelettes de l'espace $L^2(\mathfrak{R})$. Si les filtres sont à réponse

impulsionnelle finie, les fonctions d'échelle et ondelettes auront un support fini aussi. Les équations pour la décomposition seront :

$$a_{j+1}[n] = \langle f, \phi_{j,n} \rangle = a_j * \bar{h}[2n] \quad (32)$$

$$d_{j+1}[n] = \langle f, \psi_{j,n} \rangle = a_j * \bar{g}[2n] \quad (33)$$

et pour la reconstruction avec les filtres duals

$$a_j[p] = \check{a}_{j+1} * \check{h}[n] + \check{d}_{j+1} * \check{g}[n] \quad (34)$$

À la Fig. 6 on voit une représentation graphique de ce processus.

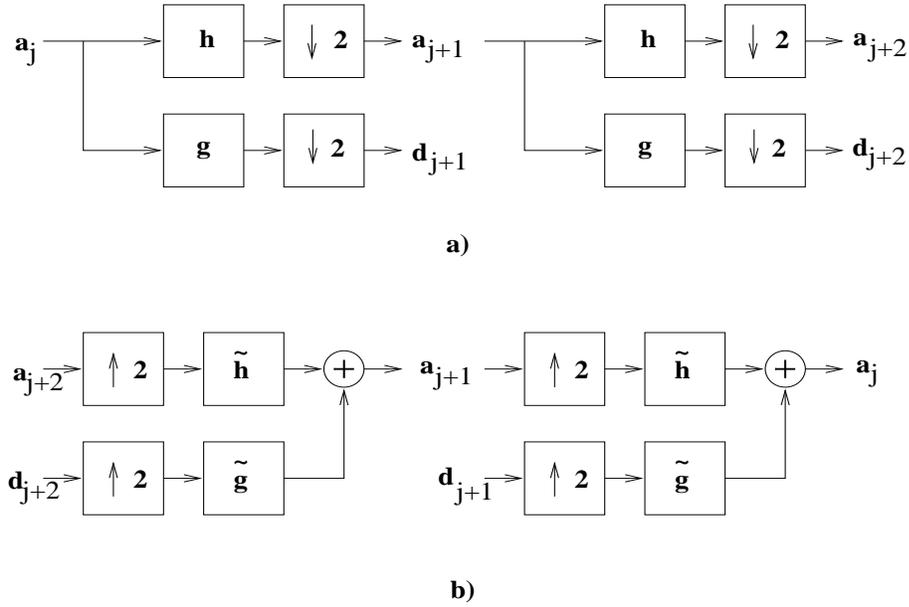


FIG. 6 – Banc de filtres utilisé pour l'analyse a) et la synthèse b) dans un espace d'ondelettes

La représentation d'une fonction $f \in L^2(\mathfrak{R})$ dans une telle base sera :

$$f(t) = \sum_{j=-\infty}^J \sum_{n \in \mathbb{Z}} d_{j,n} \psi_{j,n}(t) + \sum_{n \in \mathbb{Z}} a_{J,n} \phi_{J,n}(t) \quad (35)$$

où $d_{j,n} = \langle f, \psi_{j,n} \rangle$ et $a_{J,n} = \langle f, \phi_{J,n} \rangle$.

Les bases d'ondelettes à support fini où les fonctions d'échelle et ondelettes sont du type spline sont très utilisées en traitement d'images parce qu'elles sont continues et dérivables et les erreurs qui pourraient être introduites seront donc plus régulières et par conséquent moins visibles. On verra à la prochaine sous-section quelques caractéristiques et définitions des splines.

Les bases dans plus de dimensions sont construites comme des produits séparables des bases d'une dimension. Alors, pour le cas bi-dimensionnel on construit une base à partir du produit séparable des bases unidimensionnelles.

3.2 Les splines pour la représentation des images

Les splines sont des fonctions polynômiales par morceaux. Un spline de degré n est formé par des segments de polynômes de degré n entre les points connus comme nœuds où on impose une contrainte de continuité et continuité de la dérivé jusqu'à l'ordre $n - 1$, ce qui limite le nombre de degrés de liberté à seulement 1 par segment, c'est-à-dire qu'on peut représenter un spline avec le même nombre de coefficients que de nœuds. Le résultat le plus important est alors la représentation unique d'un spline à partir de son expansion en B-splines:

$$s(x) = \sum_{k \in \mathbb{Z}} c(k) \varphi^n(x - k) \quad (36)$$

où $\varphi^n(x)$ est la fonction B-spline de degré n qui constitue la fonction de base de l'espace des splines et les $c(k)$ sont les coefficients de l'expansion dans cette base.

Les B-splines ont plusieurs caractéristiques qui facilitent leur application dans le domaine de traitement numérique des signaux, nous allons en voir quelques unes [32]:

1. Ils ont des formules explicites dans les domaines du temps,

$$\varphi^n(x) = \frac{1}{n!} \sum_{k=0}^{n+1} (-1)^k \binom{n+1}{k} (x-k)_+^n \quad (37)$$

où $(x)_+^n = \max(0, x)^n$ et de la fréquence,

$$\hat{\varphi}^n(\omega) = \left(\hat{\varphi}^0(\omega) \right)^{n+1} = \left(\frac{1 - e^{-j\omega}}{j\omega} \right)^{n+1} \quad (38)$$

où $\hat{\varphi}^n(\omega)$ est la transformée de Fourier de $\varphi^n(x)$.

2. Ils sont faciles à manipuler. Comme les B-splines sont des polynômes par morceaux, il sont très facile a manipuler, incluant la dérivation et l'intégration.
3. Les B-splines satisfont une relation d'échelle qui permet de construire facilement des ondelettes:

$$\varphi^m(x/m) = m \sum_{k \in \mathbb{Z}} h_m^n(k) \varphi^n(x - k) \quad (39)$$

4. Le B-spline $\varphi^n(x)$ est la fonction d'échelle d'ordre $L = n + 1$ d'approximation dans des espaces d'ondelettes avec le plus petit support.
5. Ils ont la plus grande régularité pour un ordre d'approximation L donné.
6. Le B-spline est l'interpolateur qui oscille le moins.
7. Ils ont des propriétés d'approximation classés comme les meilleure entre les interpolateurs.

8. Le B-spline $\varphi^n(x)$ converge vers le filtre passe-bas idéal lorsque n tend vers l'infini.
9. Les filtres pour réaliser les opérations avec des splines sont parfaitement réalisables.

À partir de ces considérations, on trouve des applications des espaces de splines à l'interpolation des images[34], la transformation géométrique des images[33] où la construction des ondelettes pour la représentation et le codage des images[19].

Les opérations telles que l'interpolation et la projection dans des espaces de splines peuvent être réalisées à travers du filtrage numérique sans aucune difficulté. Unser[35] propose des algorithmes rapides de filtrage numérique pour l'interpolation et pour la projection sur des espaces de splines. Ces algorithmes se basent dans la propriété de convolution des versions échantillonnées des B-splines: la fonction de base de l'espace.

3.3 L'échantillonnage irrégulier

Plusieurs méthodes ont été proposées pour la reconstruction des fonctions à partir de leurs échantillons sur une grille irrégulière. On trouve dans la littérature des méthodes d'interpolation basées sur des modèles très simples et des méthodes basées sur la théorie de signaux avec une base théorique plus forte.

Les solutions les plus simples consistent à utiliser des modèles des fonctions polynômiaux par morceaux et faire une interpolation pour calculer la valeur aux points sur la grille régulière. La complexité et la qualité de ces méthodes sont en relation directe avec le degré du polynôme utilisé. Par exemple, si on utilise des polynômes de degré 0, l'interpolation devient un processus relativement simple, la valeur du pixel dans une position de la grille régulière est égale à la valeur du pixel sur la grille irrégulière qui est le plus proche, mais la qualité de la reconstruction n'est pas suffisante pour les applications envisagées, les effets sur les lignes et les frontières étant très visibles [29]. Des interpolateurs d'ordre supérieur donnent de meilleurs résultats mais au coût d'une plus grande complexité de calcul. Même les interpolateurs d'ordre 1 où la valeur du pixel est calculée comme la combinaison linéaire de ses voisins les plus proches, devient une tâche très complexe et les résultats ne sont pas encore satisfaisants. Les interpolations d'ordre supérieur sont trop complexes étant donné la nature irrégulière de la grille d'échantillonnage.

La théorie d'échantillonnage irrégulier développée pour la reconstruction des fonctions à partir de ses échantillons non uniformément repartis est l'approche la plus courante dans la littérature et celle qu'on verra dans ce travail.

On verra les différentes méthodes proposées sous l'optique de la théorie de frames, développé par Duffin et Schaeffer dans ses travaux sur l'échantillonnage irrégulier [8].

Voyons la définition de frame: [3]

Définition 1 (Frame) 1. Une séquence $\{e_n\}_{n \in \mathbb{N}}$ dans un espace de Hilbert H est

une frame si existent les constantes $0 < A \leq B$ tel que:

$$A\|f\|^2 \leq \sum_{n=1}^{\infty} |\langle f, e_n \rangle|^2 \leq B\|f\|^2, \quad \forall f \in H \quad (40)$$

2. L'opérateur de frame pour la frame $\{e_n\}_{n \in \mathbb{N}}$ est la fonction $S : H \rightarrow H$ définie comme:

$$S(f) = \sum_i \langle f, e_n \rangle e_n \quad (41)$$

3. L'opérateur S est inversible et l'ensemble $\{S^{-1}e_n\}$ est une frame de H ,

$$f = \sum_i \langle f, e_n \rangle S^{-1}e_n \quad (42)$$

$$= \sum_i \langle f, S^{-1}e_n \rangle e_n \quad (43)$$

Cette définition garantie que si les échantillons sur la grille irrégulière définissent une frame de l'espace dont appartient la fonction à reconstruire il est toujours possible de trouver une frame duale en utilisant l'opérateur S^{-1} pour reconstruire la fonction de façon parfaite. L'idée de base dans la plupart des méthodes d'échantillonnage irrégulier est justement de trouver quand le schéma d'échantillonnage définit une frame de l'espace et faire la reconstruction en considérant les échantillons comme les coefficients d'expansion dans la frame.

Après la découverte du théorème d'échantillonnage de Shannon qui permet de reconstruire une fonction continue et limitée en bande à partir de ses échantillons réguliers en utilisant une formule d'interpolation, plusieurs travaux ont été consacrés à la recherche d'un mécanisme similaire pour reconstruire la fonction à partir des échantillons non uniformes.

Dans les efforts pour trouver une solution au problème de l'échantillonnage irrégulier, un premier résultat a été apporté par le théorème de Paley-Wiener [9]. L'idée est de trouver quelle perturbation il est possible d'introduire dans une grille régulière, de telle façon qu'une base orthonormée de l'espace soit encore une base de Riesz lorsque les fonctions de la base sont déplacées aux endroits des échantillons irréguliers. Le théorème obtenu donne une fonction d'interpolation pour un certain schéma d'échantillonnage, à savoir quand les échantillons sur la grille irrégulière sont éloignés des échantillons réguliers une distance δ_k , c'est-à-dire les échantillons sont $f(k + \delta_k)$, avec les contraintes que $\sup_k (|\delta_k|) < 1/\pi^2$ et $\delta_k = -\delta_{-k}$ en considérant normalisé à 1 l'intervalle d'échantillonnage régulier ($|x_n - x_{n-1}| = 1$).

Une extension de ce résultat a été fait par Kadec qui a montré la validité du théorème pour $\sup_k (|\delta_k|) < 1/4$ [1].

Le résultat est très limité pour l'appliquer à notre problème. Les contraintes imposées à la grille irrégulière pour garantir la reconstruction ne sont pas respectées dans le cas de la compensation. Une déviation de 1/4 de pixel est trop petite, les vecteurs de mouvement ou de disparité représentent des déviations plus grandes.

Les travaux de Chen et al.[5] ont été orientés à reconstruire un signal $f(t) \in V_0$ où V_0 appartient à une analyse multirésolution avec fonction d'échelle $\varphi(t)$ à partir des échantillons non réguliers mais avec certaine contrainte:

$$f(t) = \sum_k f(k + \delta_k)S_k(t) \quad (44)$$

où $\sup_k |\delta_k| < \delta_\varphi$, $\delta_\varphi \in (0, 1]$ et $\{S_k(t)\} \subset V_0$.

Les travaux sont orientés vers la recherche d'un ensemble de fonctions interpolatrices $\{S_k(t)\} \subset V_0$ pour reconstruire la fonction f à partir des échantillons $f(k + \delta_k)$ et d'un écart maximal δ_φ . L'écart permis aux échantillons sur la grille irrégulière par rapport aux positions dans la grille régulière est trop petit: de l'ordre de $\delta_\varphi < 1/3\sqrt{3}$ sans restrictions sur le schéma d'échantillonnage. L'analyse est faite pour des espaces orthogonaux et biorthogonaux d'ondelettes. Néanmoins, la condition $\delta_\varphi \in (0, 1]$, qui ne peut pas être violé, est incompatible avec notre problème ainsi que le besoin d'avoir un échantillon sur la grille irrégulière par chaque échantillon sur la grille régulière.

Un autre résultat très important a été apporté par Duffin et Schaeffer [8]. Ils ont montrée que si l'échantillonnage irrégulier est tel que l'ensemble de fonctions $\{sinc_\omega(x - x_i)\}$, $i \in Z$ définit une frame de l'espace B_ω , la fonction originale $f(x) \in B_\omega$ peut être reconstruite à partir de ses échantillons en utilisant un algorithme itératif et l'opérateur de frame. Alors la tâche est maintenant de trouver quand le schéma d'échantillonnage définit une frame de l'espace. Dans cette situation une reconstruction parfaite est possible.

Quelques travaux ont suivi cette approche. Marvasti et al.[22] ont proposé un algorithme itératif pour la reconstruction des fonctions limitées en bande $f(x) \in B_\omega$ à partir de ses échantillons sur une grille irrégulière en utilisant l'opérateur de frame:

$$f_{k+1}(t) = \lambda PSf(t) + (P - \lambda PS)f_k(t) \quad (45)$$

$$f_0 = PSf \quad (46)$$

où λ est une constante de convergence, P est un opérateur de limitation en bande et S l'opérateur d'échantillonnage irrégulier. Plus récemment le même auteur [23] a proposé un opérateur non linéaire pour la limitation en bande, basé sur des études du spectre des signaux irrégulièrement échantillonnés. Sharaf et al. [29] utilisent ces méthodes pour la reconstruction des images obtenues par compensation du mouvement avec une transformation en avant. Les meilleurs résultats sont obtenus avec la méthode itérative et l'opérateur non linéaire et suggèrent une performance proche à celle obtenue avec la transformation en arrière classique et une interpolation bilinéaire. Un autre élément important dans cette approche est que pour garantir la convergence de la méthode non linéaire et donc de la méthode itérative, l'image doit être sur-échantillonnée. Le premier pas est alors un filtrage passe-bas avec une fréquence de coupure de 80% de la fréquence maximale de l'image originale, ce qui génère une perte d'information dès le début.

D'autres variations à l'idée originale ont été proposées par Feichtinger et Gröchenig [9]. Ces auteurs ont proposé une méthode adaptative (ADPW "Adaptive Weights

Method”) qui fait une pondération des fonctions qui forment la frame de reconstruction en tenant compte des poids relatifs des échantillons dans la grille irrégulière. Ils ont construit des frames de l’espace des fonctions limitées en bande $f(x) \in B_\omega$ avec des fonctions sinc_ω déplacées aux endroits des échantillons sur la grille irrégulière et pondéré par un facteur w_n qui dépend de la grille d’échantillonnage.

Des travaux pour accélérer la convergence de l’algorithme de frames, (celui utilisé par Marvasti dans ses travaux ou même le ADPW) ont été proposés par Gröchenig [14] et Feichtinger et Strohmer [11] en utilisant la méthode de gradient conjugué. L’avantage de ces méthodes d’accélération est qu’elles n’ont pas besoin de connaître les bornes A et B de la frame qui interviennent dans la constante de convergence de l’algorithme classique. Le prix est une plus grande complexité de calcul.

Un autre algorithme qu’on trouve aussi dans [9] propose de faire une interpolation des échantillons sur la grille irrégulière et ensuite faire la projection sur l’espace considéré. La méthode propose d’utiliser cet opérateur comme opérateur de limitation en bande dans B_ω et l’intégrer dans un algorithme itératif du type de l’algorithme de frame. L’algorithme itérative utilisé est:

$$f_{k+1} = f_k + A(f - f_k) \quad (47)$$

et l’opérateur :

$$Af = P_{B_\omega} \left(\sum_{i \in Z} f(x_i) \chi_i \right) \quad (48)$$

où P est un projecteur sur l’espace B_ω et $\{\chi_i\}_{i \in Z}$ est un ensemble des fonctions interpolatrices, le plus souvent la fonction caractéristique de la zone de Voronoïdes échantillons $f(x_i)$.

Définition 2 (Zones de Voronoï) *La zone de Voronoï V_i d’un échantillon x_i est défini comme*

$$V_i := \{x : |x_i - x| < |x_j - x|, \forall j \neq i\} \quad (49)$$

Définition 3 (Fonction caractéristique) *La fonction caractéristique de la zone V_i est défini comme,*

$$\chi_{V_i}(x) = \begin{cases} 1 & \forall x \in V_i \\ 0 & \text{autrement} \end{cases} \quad (50)$$

En suivant cette dernière méthode, Liu [18] et Aldroubi [1] proposent un algorithme de reconstruction pour des fonctions dans des espaces du type spline à partir de ses échantillons sur une grille irrégulière. Particulièrement Liu a travaillé sur des espaces d’ondelettes de splines. Aldroubi a montré les conditions dans la grille d’échantillonnage irrégulière pour avoir une reconstruction parfaite dans des espaces du type spline.

La condition de reconstruction parfaite dans ce dernier cas est liée à l’oscillation- γ de la fonction, définie comme,

$$(\text{osc}_\gamma f)(x) = \sup_{|y| \leq \gamma} |f(x) - f(x + y)| \quad (51)$$

ce qui donne une idée des variations de la fonction dans une intervalle d'ampleur γ .

Une autre méthode qui a été utilisée pour la reconstruction des fonctions non uniformément échantillonnées est la méthode POCS (“Projection Onto Convex Set”). L'idée a été utilisée par Yeh et Stark [36] pour l'interpolation des fonctions limitées en bande à partir d'un nombre fini d'échantillons sur une grille irrégulière. Si on a N échantillons $(x_n, y_n)_{n \in [1, N]}$, on peut chercher une fonction $f \in B_\omega$ tel que $f(x_n) = y_n$. On peut toujours trouver une solution à ce problème et les solutions forment un ensemble fermé et convexe de B_ω . L'idée est donc de projeter une première approximation de la fonction sur ces espaces de façon successive pour se rapprocher de la solution. Dans [10] les auteurs proposent différentes stratégies pour exécuter une méthode POCS du type “un échantillon à la fois”.

Dans l'approche POCS classique on a une ensemble d'espaces convexes et on projette de façon successive sur les espaces pour améliorer l'approximation. Les trois éléments importants dans cette approche sont:

1. Le choix des familles des espaces convexes.
2. La construction de l'opérateur de projection.
3. La stratégie d'évolution de l'algorithme à travers les familles des espaces convexes.

4 Méthode proposée

Nous décrivons dans cette section la méthode de reconstruction qu'on propose. D'abord on verra le modèle choisi pour représenter les images, ensuite quelques caractéristiques de la grille obtenue comme résultat du processus de compensation et à la fin un algorithme de reconstruction.

4.1 Modèle d'image

Les images numériques sont obtenues par échantillonnage horizontal, vertical et temporel de la projection sur le plan d'une caméra d'une scène de la vie réelle. Pour faire ce processus d'échantillonnage on suppose un modèle limité en bande. Le modèle s'ajuste très bien au système visuel humain qui est lui même limité en bande, mais ce modèle n'est pas toujours respecté pour les images réelles. Dans la nature on trouve des fréquences plus hautes que les fréquences maximales pouvant être perçues par le système visuel humain. Si ces fréquences sont présentes au moment de l'échantillonnage elles deviendront visibles à cause de l'effet de recouvrement. Le système optique de la caméra et l'ouverture du capteur font un filtrage passe-bas qui limite l'effet de recouvrement en réduisant les hautes fréquences. Le modèle est amplement utilisé pour l'acquisition des images numériques.

Une fois dans le domaine numérique les modèles utilisés pour le traitement des images sont très variés et dépendent de l'application envisagée. Pour la compensation (du mouvement ou de la disparité) avec une précision en dessous du pixel il faut trouver les valeurs des images aux points qui ne sont pas sur la grille d'échantillonnage. L'idéal serait de respecter le modèle limité en bande utilisé pour l'échantillonnage, mais les filtres nécessaires pour le faire sont irréalisables et donc d'autres modèles sont utilisés.

Parmi les plus utilisés on trouve les espaces des fonctions polynômiales par morceaux, avec le degré du polynôme pouvant aller de façon générale de 0 à 7 [24] mais les plus communs sont les polynômes de degré entre 1 (linéaire) et 3 (cubique). En bas de l'interpolation linéaire, les résultats ne sont pas très bons et il y a des distorsions visibles. En dessus de la cubique, l'amélioration ne justifie pas l'augmentation de la complexité de calcul. Les espaces de splines cubiques sont très utilisés en interpolation.

Les modèles basés sur des espaces d'ondelettes ont aussi beaucoup d'acceptation dans le milieu des images. Ils sont utilisés pour le codage et plusieurs autres applications telles que la segmentation, l'amélioration des images, etc. Une base d'ondelettes est une base de l'espace $L^2(\mathfrak{R}^2)$ des fonctions d'énergie finie qui permet de construire des approximations multirésolution des fonctions qui appartiennent à cet espace. [19] p.221.

Si $\{V_j\}_{j \in (-\infty, +\infty)}$ est une approximation multirésolution de $L^2(\mathfrak{R}^2)$, l'espace des fonctions de carré intégrable sur \mathfrak{R}^2 , on peut construire une base d'ondelettes $\psi_{j,\mathbf{n}}(\mathbf{x})$ associée à cet approximation multirésolution, avec fonction d'échelle $\phi_{j,\mathbf{n}}(\mathbf{x})$. Toute

fonction $f \in L^2(\mathbb{R}^2)$ peut être décrite comme: [19]

$$f(\mathbf{x}) = \sum_{j=-\infty}^J \sum_{\mathbf{n} \in Z^2} w_{j,\mathbf{n}} \psi_{j,\mathbf{n}}(\mathbf{x}) + \sum_{\mathbf{n} \in Z^2} c_{J,\mathbf{n}} \phi_{J,\mathbf{n}}(\mathbf{x}) \quad (52)$$

où $w_{j,\mathbf{n}}$ sont les coefficients d'ondelettes et $c_{J,\mathbf{n}}$ sont les coefficients d'expansion dans l'espace V_J avec les fonctions d'échelles:

$$\phi_{J,\mathbf{n}}(\mathbf{x}) = \frac{1}{\sqrt{2^J}} \phi\left(\frac{\mathbf{x} - 2^J \mathbf{n}}{2^J}\right) \quad (53)$$

On peut toujours obtenir les coefficients d'expansion dans les espaces V_j de la projection orthogonale sur cet espace de toute fonction $f \in L^2(\mathbb{R}^2)$. Ces coefficients représentent des approximation à la résolution 2^{-j} de la fonction f . Alors les coefficients $c_{J,\mathbf{n}}$ sont la représentation numérique de l'approximation $f_J = P_{V_J} f \in V_J$ de la fonction $f \in L^2(\mathbb{R}^2)$.

Si $Ic_J = P_{V_J} Ic$ est l'approximation à la résolution 2^{-J} d'une certaine image continue Ic , et si $I[\mathbf{n}]$ est sa représentation numérique (expansion avec les bases $\phi_{J,\mathbf{n}}$), alors

$$Ic_J(\mathbf{x}) = \sum_{\mathbf{n}} I[\mathbf{n}] \phi_{J,\mathbf{n}}(\mathbf{x}) \quad (54)$$

Par les propriétés des approximations multirésolution (équation (7)), $Ic_J \in V_{J-r}$ et:

$$Ic_J(\mathbf{x}) = \sum_{\mathbf{m}} Ir(\mathbf{m}) \phi_{J-r,\mathbf{m}}(\mathbf{x}) \quad (55)$$

où Ir est la représentation numérique dans l'espace V_{J-r} de Ic_J .

Dans un espace d'ondelettes où la transformation entre les différentes échelles peut être faite en utilisant des bancs de filtres, on peut obtenir Ir à partir de I et vice versa par filtrage étant donné que pour Ic_J tous les coefficients ondelettes $w_{j,\mathbf{n}} = 0, \forall j \leq J$.

Les bases d'ondelettes avec fonction d'échelle et ondelettes formées par des splines sont les plus utilisées en images. Une base biorthogonale de ce type a un support fini et les filtres nécessaires sont donc des filtres RIF facilitant le processus de transformation.

En résumé, on propose d'utiliser pour modéliser les images une approximation multirésolution $\{V_j\}_{j \in Z}$ de l'espace des fonctions de carré intégrable avec une base d'ondelettes biorthogonale de type spline. Les fonctions d'échelle ϕ et duale $\tilde{\phi}$ et ondelettes ψ et duale $\tilde{\psi}$ constituent la base. La décomposition d'une fonction dans cette base sera:

$$f = \sum_{j=-\infty}^J \sum_{\mathbf{n} \in Z^2} \langle f, \psi_{j,\mathbf{n}} \rangle \tilde{\psi}_{j,\mathbf{n}} + \sum_{\mathbf{n} \in Z^2} \langle f, \phi_{J,\mathbf{n}} \rangle \tilde{\phi}_{J,\mathbf{n}} \quad (56)$$

On considère que l'image originale appartient à l'espace V_J et elle est donc parfaitement décrite par ses coefficients d'expansion avec la base $\{\phi_{J,\mathbf{n}}\}_{\mathbf{n} \in Z^2}$ et on peut obtenir les représentations numériques pour des approximation dans des espaces de plus haute résolution, $V_{J-r}, r \in Z^+$ où les coefficients d'ondelettes seront tous zéros.

4.2 Échantillonnage irrégulier

Les points obtenus comme résultat de la compensation sont sur une grille irrégulière. Les caractéristiques de cette grille dépendent du champ de vecteurs qui décrit la transformation (disparité ou mouvement). Ce champ de vecteurs est obtenu à partir de l'estimation de la transformation en utilisant plusieurs images. Il y a quelques situations dans le processus d'estimation qui sont intéressants si on pense à la grille irrégulière:

1. On pense toujours à un champ de vecteurs dense, c'est-à-dire un vecteur par chaque point de l'image. Il est toujours possible d'obtenir un tel champ même si l'estimation a été faite avec un algorithme de blocs. On aura donc presque la même quantité de points dans les grilles régulières et irrégulières.
2. Il y a des points qui disparaissent avec la transformation. Pour ces points, il est impossible de trouver un vecteur qui décrit la transformation parce qu'ils n'ont pas de correspondants dans l'autre image. Dans ce cas, on aura des vecteurs qui décrivent une fausse transformation.
3. Il y a des points qui apparaissent dans la nouvelle image, donc ces points ne seront pas représentés par le champ de vecteurs qui décrit la transformation. Aucun vecteur tombera sur ces points et on aura donc des trous sur la grille irrégulière.

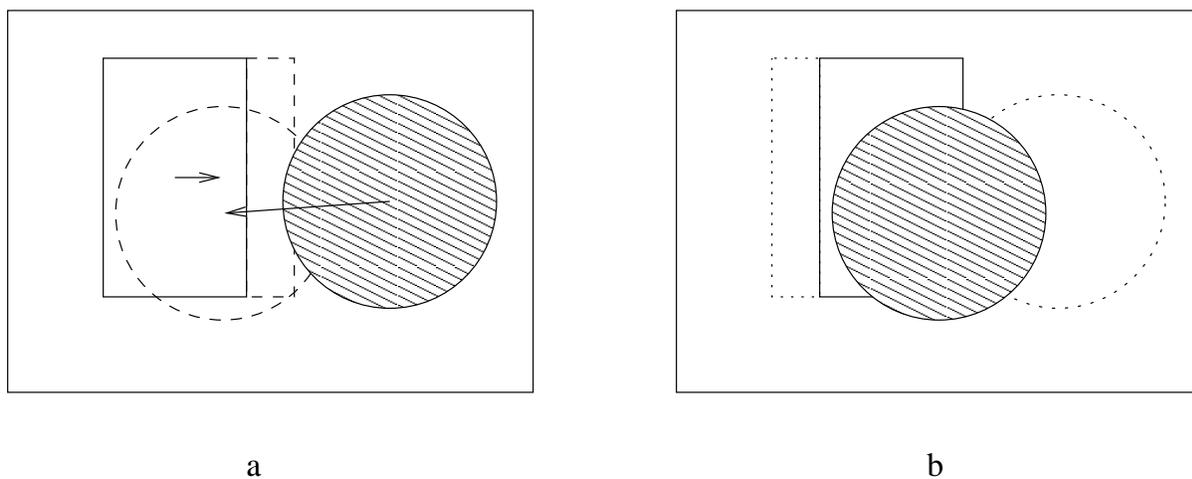


FIG. 7 – Zones des points cachés a) et nouveaux b)

À la Fig. 7 on voit une zone de points cachés définie par la ligne discontinue dans l'image a) et des points nouvellement exposés définis par la ligne de points à l'image b) lorsque le rectangle bouge vers la droite et le cercle bouge vers la gauche. Dans ces zones le processus d'estimation ne sera pas exact et la compensation donnera donc

des mauvais résultats. Il est impossible de récupérer les valeurs de la vraie image dans ces zones en utilisant seulement l'information obtenue par compensation de la transformation.

Les zones où la compensation ne donne pas de résultats exacts doivent être traitées de façon différente des zones où les échantillons obtenus par compensation donnent des vrais échantillons de l'image à reconstruire.

4.3 Algorithme de reconstruction

Nous proposons donc une méthode pour la reconstruction sur une grille régulière d'images obtenues à partir de l'application d'une transformation géométrique spatiale à une image dans le cas où la transformation nous conduit à une grille irrégulière.

D'abord, il faut considérer les caractéristiques de cet échantillonnage irrégulier, on a vu qu'il y a des régions avec des échantillons irrégulièrement espacés mais avec une densité relativement constante à l'intérieur des objets où la transformation est relativement uniforme, et des régions où il n'y a pas d'échantillons où même ces échantillons ne sont pas fidèles à cause des erreurs dans l'estimation de la transformation. Nous allons nous concentrer d'abord sur la reconstruction des premiers types de régions. Ensuite on verra des solutions possibles pour les autres régions.

On a proposé un modèle d'ondelettes spline pour décrire les images. Si on considère que l'image à reconstruire $Ic_J \in V_J$ et qu'elle est décrite par les coefficients d'expansion dans cet espace $I[\mathbf{n}] = \langle Ic_J, \phi_{J,\mathbf{n}} \rangle$, on peut construire une version de plus haute résolution dans l'espace V_{J-r} décrite par les coefficients $Ir[\mathbf{n}] = \langle Ic_J, \phi_{J-r,\mathbf{n}} \rangle$:

$$Ic_J(\mathbf{x}) = \sum_{\mathbf{n} \in Z^2} I[\mathbf{n}] \tilde{\phi}_{J,\mathbf{n}}(\mathbf{x}) = \sum_{\mathbf{m} \in Z^2} Ir[\mathbf{m}] \tilde{\phi}_{J-r,\mathbf{m}}(\mathbf{x}) \quad (57)$$

La différence entre les variables indépendantes des deux sommes n'est que pour signaler qu'il n'y a pas de correspondance entre les valeurs des images pour les mêmes valeurs de la variable indépendante. À la Fig. 8 on a noté les valeurs de \mathbf{n} et de \mathbf{m} pour les points dans les deux grilles. Il est clair que la correspondance entre les deux images est établie quand $\mathbf{m} = 2^r \mathbf{n}$.

Ces deux jeux de coefficients sont reliés par les filtres associés aux fonctions d'échelles, $h[\mathbf{n}]$ et son dual $\tilde{h}[\mathbf{n}]$ dans le cas qui nous occupe où les coefficients d'ondelettes sont zéros. Les convolutions successives de ces filtres génèrent des bases biorthogonales $\{\theta_{r,\mathbf{n}}\}$ et $\{\tilde{\theta}_{r,\mathbf{n}}\}$ de l'espace de coefficients à la résolution 2^{-J+r} des fonctions $Ic_J \in V_J$.

Si on définit l'espace des coefficients des fonctions $Ic_j \in V_j$ comme U_j , on aura que $I \in U_J$ et $Ir \in U_{J-r}^J \subset U_{J-r}$. L'espace U_{J-r}^J est formé par les coefficients d'expansion des fonctions $Ic_J \in V_J$ à la résolution 2^{-J+r} . Évidemment cet espace est un sous-espace de l'espace U_{J-r} formé par les coefficients d'expansion à la résolution 2^{-J+r} des fonctions $Ic_{J-r} \in V_{J-r}$, étant donné que $V_J \subset V_{J-r}$. Les espaces U_J et U_{J-r}^J ont la même dimension et décrivent le même ensemble des fonctions continues, ce qui veut dire qu'on n'a pas besoin de tous les échantillons de Ir pour reconstruire I Il

suffit d'avoir autant d'échantillons qu'en I et que ces échantillons soient pris de façon adéquate.

Les bases biorthogonales $\{\theta_{r,\mathbf{n}}\}$ et $\{\tilde{\theta}_{r,\mathbf{n}}\}$ nous permettent de construire l'approximation Ir à partir de I .

$$Ir[\mathbf{m}] = \sum_{\mathbf{n} \in Z^2} I[\mathbf{n}] \tilde{\theta}(\mathbf{m} - 2^r \mathbf{n}) \quad (58)$$

Dans le cas des images numériques de dimension $N \times M$, la représentation de plus haute résolution aura une dimension $2^r N \times 2^r M$, alors on aura 2^{2r} points par chaque point de l'image originale. En fait on a une approximation de l'image originale sur une grille plus dense que la grille originale.

Si les échantillons sont pris sur cette grille plus dense, ce qui est toujours possible étant donné qu'on peut faire r aussi grand qu'on veut et que de façon générale les vecteurs qui décrivent la transformation sont quantifiés à une certaine précision, les échantillons seraient sur une grille régulière plus dense que l'originale. À la Fig. 8 on voit comment les vecteurs qui décrivent la transformation tombent sur des points d'une grille G_d , plus dense ($r = 2$) que la grille de référence G_r . Les gros points noirs, représentent la grille moins dense; les petits points noirs, la grille plus dense et les points encadrés, les positions des échantillons irrégulièrement repartis sur la grille plus dense. Dans ce cas, les vecteurs ont été quantifiés à une précision d'un quart de pixel.

Définissons l'ensemble des échantillons sur la grille irrégulière par $\Gamma = \{\mathbf{m}_i\}, i \in Z$. Les valeurs de $Ir[\mathbf{m}_i]$ peuvent être obtenues par:

$$Ir[\mathbf{m}_i] = \sum_{\mathbf{n} \in Z^2} I[\mathbf{n}] \tilde{\theta}[\mathbf{m}_i - 2^r \mathbf{n}] = \langle I, \tilde{\theta}_i \rangle \quad (59)$$

où

$$\tilde{\theta}_i[\mathbf{n}] = \tilde{\theta}[\mathbf{m}_i - 2^r \mathbf{n}] \quad (60)$$

L'ensemble des fonctions $\{\tilde{\theta}_i\}$ constitue une frame de l'espace qu'elle génère et les échantillons sur la grille irrégulière $Ir[\mathbf{m}_i], \mathbf{m}_i \in \Gamma$ seront les coefficients d'expansion de l'image I régulièrement échantillonnée dans cet espace. Les caractéristiques de cet espace, qui sont déterminées par la transformation utilisée pour la compensation sont très importantes pour la reconstruction de I à partir des échantillons sur une grille irrégulière de Ir . La référence qu'on a fait à une frame et non pas à une base est basée sur l'argument qu'on ne peut pas affirmer que cet ensemble soit une base; les éléments pourraient être linéairement dépendants.

Les fonctions $\Theta = \{\tilde{\theta}_i\}$ sont à support fini, parce qu'elles sont formées à partir de convolutions successives de filtres RIF. Ce qui implique que la où les échantillons sur la grille irrégulière soient assez nombreux et sa distribution soit assez uniforme, la reconstruction sera bonne. Dans les zones des trous provoqués par les régions nouvelles, la reconstruction dépendra de l'opérateur choisi et de la première approximation. Cet ensemble génère l'espace des coefficients U_J lorsque tous les vecteurs sont linéairement indépendants et qu'on a autant de vecteurs que de dimensions dans l'espace.

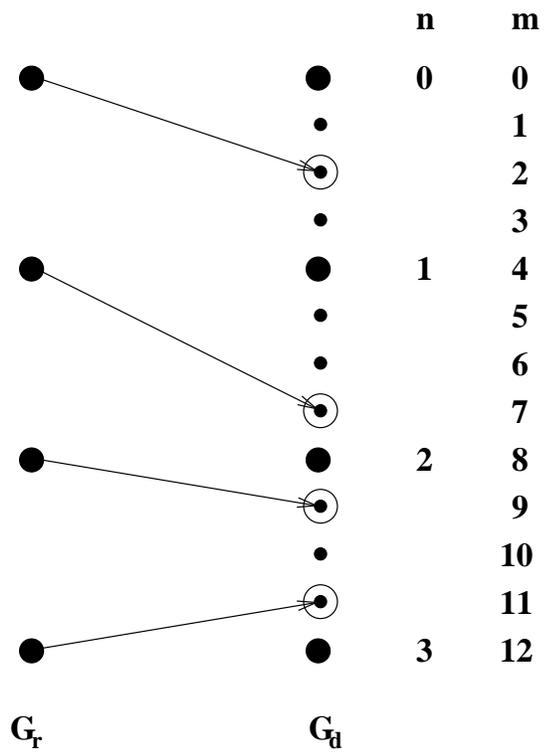


FIG. 8 – Échantillonnage irrégulier sur une grille plus dense

Cette situation est rarement rencontrée. On a vu que la quantité des vecteurs est normalement plus petite que la dimension de l'espace (la quantité des échantillons de I) et qu'il est possible d'avoir des vecteurs linéairement dépendants dans l'ensemble. Même dans le cas où ces vecteurs ne soient pas une frame de l'espace U_J , ils seront une frame d'un sous-espace de U_J et on peut trouver la projection orthogonale de I sur cet espace.

Si Θ est une frame de U_J , il existe une frame duale Θ^d qui permet de reconstruire les fonctions de U_J :

$$I = \sum_i \langle I, \tilde{\theta}_i \rangle \tilde{\theta}_i^d. \quad (61)$$

Trouver la frame duale est une tâche difficile. Elle dépend du schéma d'échantillonnage et doit donc être calculée pour chaque image. La façon normale d'aborder le problème est d'utiliser des algorithmes itératifs.

Si Θ n'est pas une frame de U_J , elle est une frame d'un sous-espace $U'_J \subset U_J$ et on peut obtenir la projection orthogonale sur ce sous-espace:

$$\tilde{I} = P_{U'_J} I = \sum_i \langle I, \tilde{\theta}_i \rangle \tilde{\theta}_i^d \quad (62)$$

Dans le cas de compensation qui nous occupe, on aura la deuxième situation le plus souvent. La reconstruction parfaite ne serait alors pas possible, étant donné que dans le processus de compensation il y a une perte d'information qui n'est pas récupérable en utilisant seulement l'information correspondante à l'image de référence et le champ de vecteurs. Il faudra utiliser d'autres techniques pour récupérer cette information.

La méthode proposée est basée sur la reconstruction de l'image I en utilisant des algorithmes itératifs. L'idée de base est de considérer les échantillons sur la grille irrégulière comme les coefficients d'expansion de l'image à reconstruire avec la frame Θ et reconstruire la projection orthogonale de cette image sur l'espace U'_J . Cette projection sera la meilleure approximation \tilde{I} de l'image originale I qu'on peut trouver sur l'espace U'_J . Quand $U'_J = U_J$, on aura une reconstruction parfaite.

L'algorithme itératif en question est:

$$f_{k+1} = f_k + T(f - f_k) \quad (63)$$

avec f_0 une première approximation à la fonction f et T un opérateur sur l'espace qui satisfait pour une certaine constante positive $\gamma < 1$:

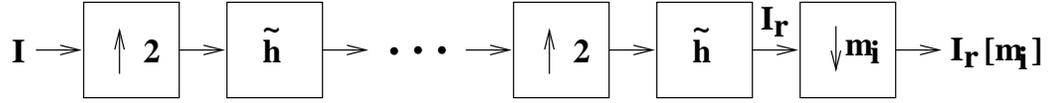
$$\|f - Tf\| \leq \gamma \|f\| \quad (64)$$

et avec une estimation de l'erreur après k itérations:

$$\|f - f_k\| \leq \gamma^{k+1} \|f\| \quad (65)$$

L'opérateur T peut être décomposé en deux parties:

1. L'obtention des échantillons sur la grille irrégulière $Ir_\Gamma = Ir[\mathbf{m}_i]_{\mathbf{m}_i \in \Gamma}$ dont on a le diagramme à la Fig. 9 et qu'on appelle S ($Ir_\Gamma = SI$).

FIG. 9 – Opérateur S pour l’obtention de $I_r[m_i]$ à partir de I

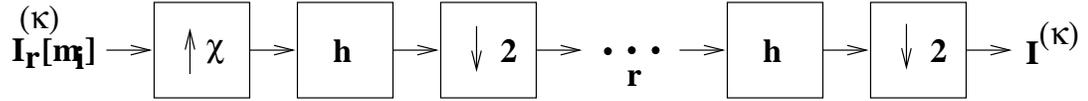
2. La construction de l’approximation \tilde{I} à partir des échantillons sur la grille irrégulière Ir_Γ , dénoté par A ($\tilde{I} = A Ir_\Gamma$).

L’obtention des échantillons sur la grille irrégulière (avec l’opérateur S) est simplement $Ir[m_i] = \langle I, \tilde{\theta}_i \rangle$. L’opérateur A qu’on a choisi est la projection d’une interpolation des échantillons irrégulièrement repartis aux points de la grille plus dense:

$$A(Ir) = P_{U_J} \left(\sum_i Ir[m_i] \chi_i \right) \quad (66)$$

où χ_i est une fonction d’interpolation, qui pourrait être la fonction caractéristique de la zone de Voronoï associée à l’échantillon $Ir[m_i]$ pour une interpolation d’ordre zéro.

Le premier élément dans la définition de l’opérateur est la fonction interpolatrice χ_i . On a choisi une interpolation d’ordre 0, mais on pourrait utiliser d’autres interpolateurs plus simples ou plus complexes. On applique l’interpolation sur la grille plus dense et ensuite on projette cette fonction sur la grille moins dense. Cette dernière opération peut être fait aussi par filtrage avec les filtres h et g . Le diagramme à la Fig. 10 montre une possibilité pour l’application de l’opérateur, avec une projection orthogonale sur l’espace U_J . D’autres projections sont aussi possibles.

FIG. 10 – Diagramme de l’opérateur A

Alors pour appliquer l’algorithme itératif il ne reste qu’à définir la première approximation. Cette première approximation pourrait être simplement AI : le résultat d’appliquer l’opérateur à l’image originale irrégulièrement échantillonnée. Cette méthode donne de bons résultats pour commencer le processus itératif.

$$\tilde{I}^0 = A(Ir_\Gamma) \quad (67)$$

Le diagramme pour le processus itératif est montré à la Fig. 11. L’algorithme est décrit comme:

1. Obtention de l’image compensée en mouvement irrégulièrement échantillonnée. (Cette démarche n’est pas vraiment exécutée avec cette formule. Les échantillons sont le résultat de la compensation.)

$$Ir_\Gamma = \left\{ Ir[m_i] = \langle I, \tilde{\theta}_i \rangle \right\}_{m_i \in \Gamma} \quad (68)$$

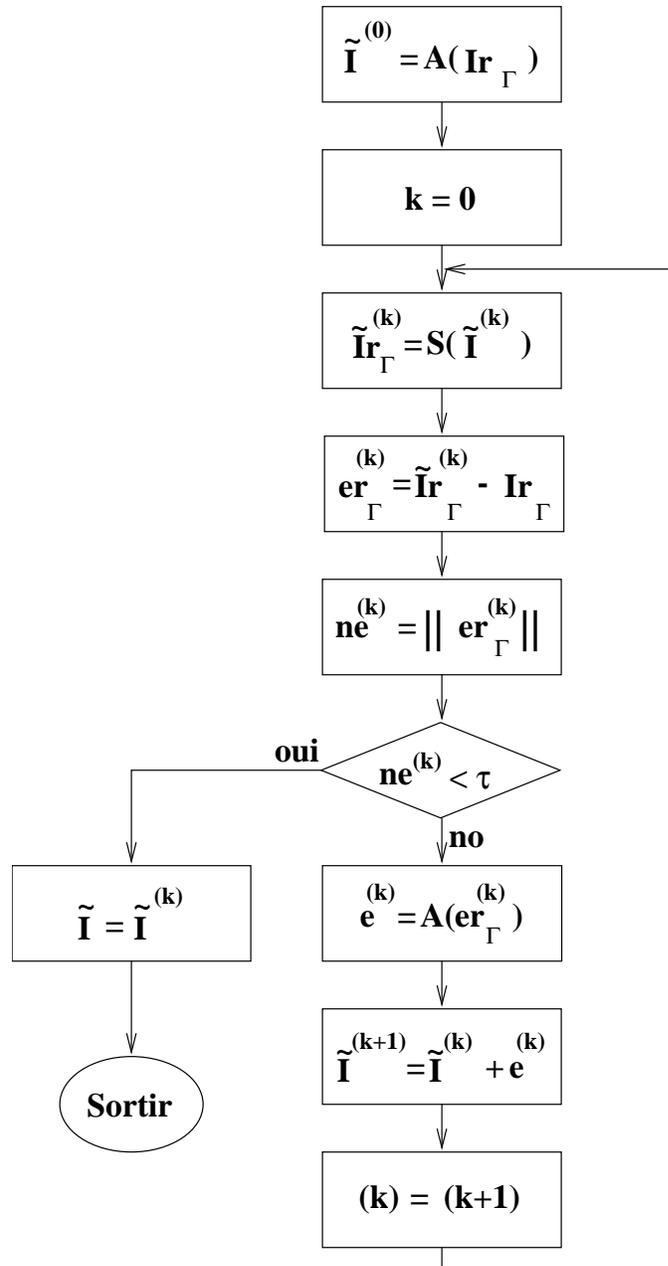


FIG. 11 – Diagramme de l'algorithme itératif

2. Obtention de la première approximation \tilde{I}_0

$$\tilde{I}^0 = A(Ir_\Gamma) \quad (69)$$

3. Initialisation de la variable $k = 0$
4. Estimation des échantillons sur la grille irrégulière à partir de l'approximation antérieure

$$\tilde{I}_\Gamma^{k+1} = \left\{ \tilde{I}^{k+1}[\mathbf{m}_i] = \langle \tilde{I}^k, \tilde{\theta}_i \rangle \right\}_{\mathbf{m}_i \in \Gamma} \quad (70)$$

5. Calcul de l'erreur et sa norme

$$er_\Gamma^k = \left\{ er^k[\mathbf{m}_i] = Ir[\mathbf{m}_i] - \tilde{I}^{k+1}[\mathbf{m}_i] \right\}_{\mathbf{m}_i \in \Gamma} \quad (71)$$

6. Application de l'opérateur A à l'erreur et somme avec l'approximation antérieure

$$\tilde{I}_\Gamma^{k+1} = \tilde{I}_\Gamma^k + A(er_\Gamma^k) \quad (72)$$

7. Répéter le processus défini pour les 3 dernières étapes jusqu'à ce que la condition d'arrêt soit satisfaite.

Quelques variations pourraient être introduites pour améliorer la vitesse de convergence de l'algorithme, comme celle proposée par Gröchenig [14] d'utiliser un algorithme de type de gradient conjugué ou celle proposée par Feichtinger [9] avec une pondération des échantillons. Ces variations pourront être considérées pour accélérer l'algorithme et augmenter son efficacité.

4.3.1 Les occlusions et les zones nouvellement exposées

La façon de traiter les occlusions et les zones nouvellement exposées dépend de l'application. Les effets introduits par les erreurs liés à ces phénomènes ne sont pas les mêmes et ils doivent être traités de façon différentes.

Dans le cas de la reconstruction des vues intermédiaires des paires stéréoscopiques, les erreurs d'estimation générées par les occlusions peuvent nous conduire à une mauvaise valeur dans un point si on utilise les deux points aux extrêmes du vecteur pour estimer la valeur à une position intermédiaire. On a vu à l'équation (1) que pour trouver la valeur du point correspondant à la nouvelle position on utilise une fonction d'interpolation. Si cette fonction est une fonction linéaire comme proposé dans la littérature et les points utilisés ne sont pas des vrais correspondants, la nouvelle valeur sera très différente des deux points et l'erreur sera donc notable. Il sera alors préférable d'utiliser une fonction qui tient compte de cette situation et aux endroits où l'estimation ne soit pas bonne, prendre une seule valeur pour la reconstruction à la position intermédiaire. Ce problème est relié à la méthode utilisée pour l'estimation de la disparité et doit être analysé conjointement avec cette méthode. Cette analyse fait partie du travail à faire pour améliorer la performance de l'algorithme.

Les zones nouvellement exposées provoquent des trous dans l'image compensée. À cause de ces trous, la reconstruction de ces zones ne sera pas exacte et dépendra beaucoup de la première approximation et des points voisins. Il faudra alors identifier ces zones et une possible solution est d'introduire quelques points à ces endroits ou faire une estimation de disparité réduite pour ces zones dans l'autre image en tenant compte de l'estimation déjà connue.

Pour le codage, la situation est différente: les erreurs d'estimation peuvent être évaluées de façon exacte. Si on code les trajectoires de mouvement, il n'y aura pas d'erreur dans les échantillons parce que les trajectoires tiennent compte des erreurs d'estimation qui sont codées et envoyées au récepteur. Seulement les erreurs introduites par la quantification et la méthode de codage doivent être considérées. De plus, les zones nouvellement exposées, constituent un défi: on n'aura pas d'information sur ces zones et, comme dans le cas précédent, elles provoqueront des trous dans l'image compensée. La solution pour ce cas doit être plus spécifique et dépendra du schéma de codage à employer.

5 Résultats expérimentaux

Dans le but de vérifier les possibilités de l'approche proposée on a fait des expériences avec des images irrégulièrement échantillonnées obtenues à partir d'un processus de compensation. Des paires d'images stéréoscopiques et des images synthétiques et réelles avec des champs de mouvement synthétiques ont été utilisées.

Pour comparer les résultats dans l'interpolation des vues intermédiaires on a utilisé une des images originales qu'on a reconstruite en utilisant un champ de disparité réelle. On a aussi reconstruit une séquence d'images intermédiaires pour comparer avec la propagation des vecteurs tel que proposé par Konrad [16]. Pour les images compensées par le mouvement l'approche classique utilisant une transformation en arrière a été le choix pour la comparaison.

5.1 Algorithme implanté

Nous avons utilisé pour les expériences l'algorithme itératif proposé (Fig. 11). Nous décrivons les principaux éléments qui interviennent dans l'algorithme.

- L'espace utilisé est un espace d'ondelettes splines avec une base biorthogonale. Les filtres h, \tilde{h}, g et \tilde{g} sont définis par les fonctions d'échelles et ondelettes. Dans ce cas ces fonctions sont des splines et les fonctions de l'espace sont donc aussi des splines. On a utilisé les filtres calculés par Cohen, Daubechies et Feauveau [19] page 271. Ces filtres correspondent aux splines de degré 1 et 2 pour l'analyse.
- La fonction χ_i dans l'équation (66) qu'on a choisi est la fonction caractéristique de la zone de Voronoï associée à l'échantillon $Tr[\mathbf{m}_i]$. On a alors une interpolation d'ordre 0 des échantillons comme premier pas dans le processus de reconstruction.
- La projection P dans l'équation (66) qu'on a implantée est la projection orthogonale sur l'espace V_J . Cette projection est obtenue en calculant la transformée en ondelette de l'image sur la grille plus dense en mettant à zéro tous les coefficients ondelettes $d_{j,n}$ tel que $J - r < j \leq J$. De cette façon, on assure que la fonction résultante appartient à l'espace V_J et qu'elle est la projection orthogonale sur cet espace de la fonction originale. Le diagramme de la projection est représenté à la Fig. 10.

5.2 Images et champs utilisés

On a fait deux types d'expériences: la reconstruction des images intermédiaires dans les paires d'images stéréoscopiques et la reconstruction des images compensées par le mouvement.

Pour la compensation de mouvement on a utilisé deux images, une synthétique et une autre réelle, toujours avec des champs de mouvement synthétiques pour éviter

le problème des régions nouvelles ou des régions cachées qui peuvent générer des difficultés additionnelles pour la reconstruction.

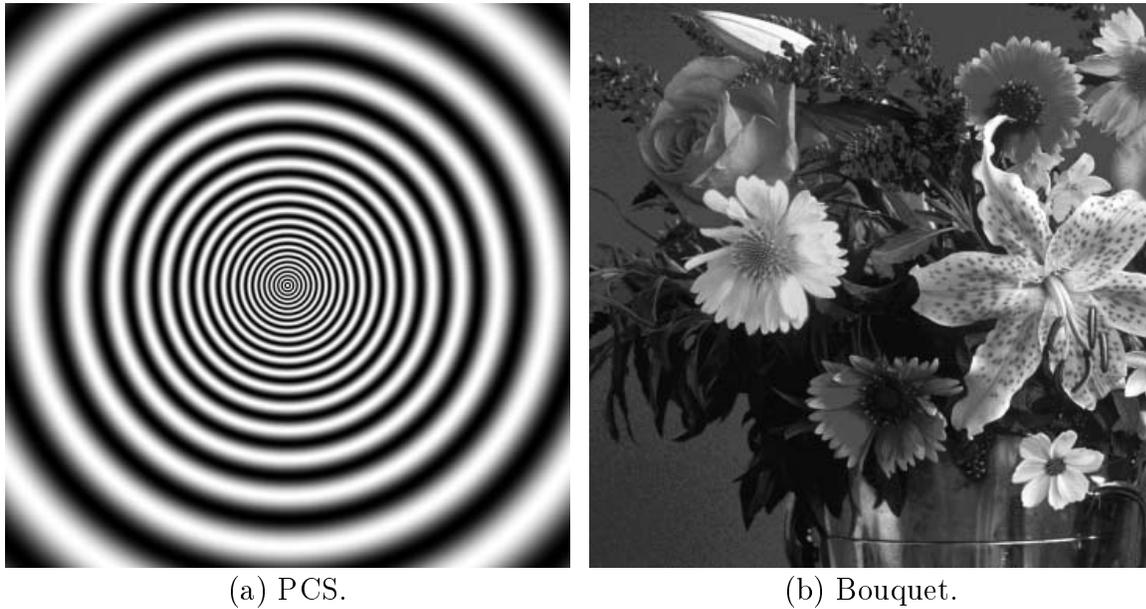


FIG. 12 – Images originales pour la compensation du mouvement

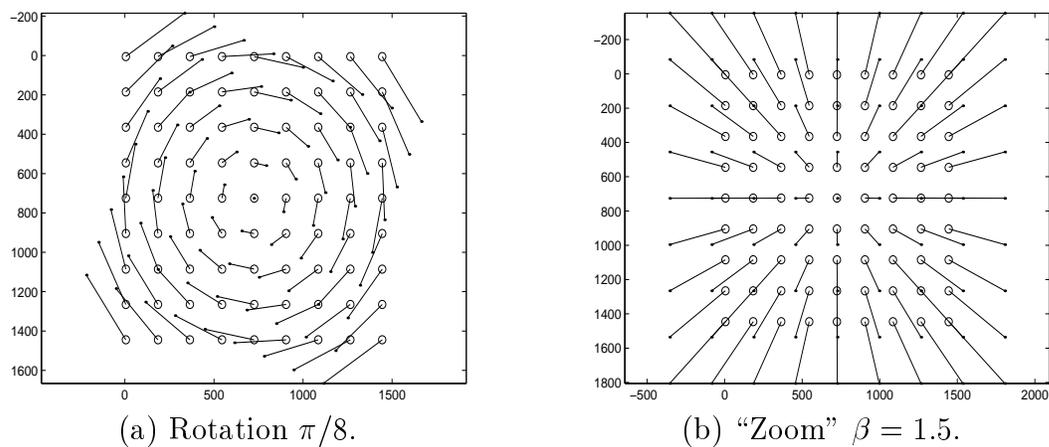


FIG. 13 – Champs de mouvement utilisés

L'image synthétique utilisée est formée par un patron circulaire sinusoïdal (PCS) avec des plus hautes fréquences au centre, dans le but de vérifier le comportement de l'algorithme face aux différentes fréquences trouvées dans les images. On peut voir à la Fig. 12 (a) ce patron (fichier `pcs.tif`). L'image réelle utilisée pour les expériences est l'image "Bouquet", qui est montrée à la Fig. 12 (b) (fichier `bouquet.tif`). On a appliqué à ces images deux types de transformation: un mouvement de rotation

d'un angle de $\alpha = \pi/8$ ou un "zoom" au centre avec un facteur $\beta = 1.5$. Ensuite on a appliqué les transformations inverses pour comparer avec les images originales. Les deux transformations sont montrées aux Fig. 13 (a) et Fig. 13 (b) respectivement, où les cercles sont identifiés avec les points originaux sur la grille uniforme et les points avec les positions des points irrégulièrement espacés.

Pour la compensation de la disparité, la paire utilisée est la première de la séquence stéréoscopique "flower", dont on montre à la Fig. 14 les images gauche (a) (fichier `flow_left.tif`) et droite (b) (fichier `flow_right.tif`). Un champ de disparité dense a été estimé pour cette paire avec la méthode décrite dans par Mancini [20].



FIG. 14 – Images gauche et droite de la paire stéréoscopique "flowers"

5.3 Résultats

Les épreuves de compensation du mouvement ont consisté à appliquer l'algorithme proposé aux échantillons sur la grille irrégulière obtenus comme résultat de la compensation du mouvement de ces deux images avec les transformations mentionnées. Après avoir appliqué la transformation, on applique la transformation inverse au résultat pour revenir sur l'image originale et avoir un patron de comparaison. Le même processus est appliqué avec une interpolation d'ordre zéro et avec une transformation en arrière qui n'a pas besoin de l'interpolation des échantillons sur la grille irrégulière et qui est la façon classique de faire la compensation du mouvement. Pour la rotation on a obtenu ces images avec l'aide du logiciel "Image Magic", un logiciel de traitement des images. Pour le zoom, les images ont été obtenues à l'aide de MATLAB, utilisant l'interpolation bicubique décrite par Keys dans [15]

On montre aux Fig. 15, 16, 17 et 18 des détails des images et leurs reconstructions avec les méthodes mentionnées, et entre parenthèses le nom du fichier qui contient l'image:

1. Rotation $\pi/8$

- (a) Fig. 15 (a) Détail de PCS original (`pcs.tif`)

- (b) Fig. 15 (b) Reconstruction avec interpolation d'ordre zéro (`pcs_r_NN_b.tif`).
- (c) Fig. 15 (c) Reconstruction avec transformation en arrière (`pcs_r_b_b.tif`).
- (d) Fig. 15 (d) Reconstruction avec l'algorithme proposé (`pcs_r_fe_p37_r4_b.tif`).
- (e) Fig. 17 (a) Détail de Bouquet original (`bouquet.tif`).
- (f) Fig. 17 (b) Reconstruction avec interpolation d'ordre zéro (`bouquet_r_NN_b.tif`).
- (g) Fig. 17 (c) Reconstruction avec transformation en arrière (`bouquet_r_b_b.tif`).
- (h) Fig. 17 (d) Reconstruction avec l'algorithme proposé (`bouquet_r_fe_p44_i3_b.tif`).

2. Zoom de rapprochement et d'éloignement

- (a) Fig. 16 (a) Détail de l'image PCS original (`pcs.tif`).
- (b) Fig. 16 (b) Reconstruction avec interpolation d'ordre zéro (`pcs_e_NN_b.tif`).
- (c) Fig. 16 (c) Reconstruction avec transformation en arrière (`pcs_e_b_b.tif`).
- (d) Fig. 16 (d) Reconstruction avec l'algorithme proposé (`pcs_e_fe_p3744_r4_b.tif`).
- (e) Fig. 18 (a) Détail de l'image Bouquet original (`bouquet.tif`).
- (f) Fig. 18 (b) Reconstruction avec interpolation d'ordre zéro (`bouquet_e_NN_b.tif`).
- (g) Fig. 18 (c) Reconstruction avec transformation en arrière (`bouquet_e_b_b.tif`).
- (h) Fig. 18 (d) Reconstruction avec l'algorithme proposé (`bouquet_e_fe_p3744_r4_b.tif`).

Les résultats de calculer la PSNR par rapport à l'image original sont montrés au tableau Tab. 1.

On voit que les résultats obtenus même s'ils correspondent à des situations très simples avec des images synthétiques ou champs de mouvement synthétiques, sont très encourageants. Dans le cas de la rotation les résultats sont mieux que ceux obtenus avec l'approche qui est utilisée aujourd'hui pour le codage vidéo et qui ne peut pas être utilisée avec les nouveaux schémas de codage qui utilisent les trajectoires du mouvement. Pour le zoom les résultats sont beaucoup mieux que ceux obtenus avec

		"Ordre 0"	Arrière	Méthode proposée
PCS	Rotation	26.74	75.48	83.17
	Zoom	32.66	104.21	88.60
Bouquet	Rotation	43.22	80.00	87.89
	Zoom	48.44	109.22	90.45

TAB. 1 – PSNR des images reconstruites (db)

une interpolation d'ordre 0 et du point de vue visuel ils sont assez proches de ceux obtenus avec l'approche de la transformation en arrière.

Dans le cas de la reconstruction des vues intermédiaires, on a reconstruit l'image droite à partir d'un champ de disparité estimé avec un $\alpha = 0.5$. En utilisant ce champ de disparité on obtient des points de l'image droite obtenus par interpolation mais sur une grille irrégulière. Reconstruire l'image droite donne la possibilité de comparer avec une image connue et minimiser les effets des erreurs dans l'estimation de la disparité. On a fait aussi des expériences pour reconstruire des images à d'autres positions. Dans la Fig. 19 on voit:

1. Fig. 19 (a) Détail de l'image droite de Flower (`flow_right.tif`)
2. Fig. 19 (b) Interpolation d'ordre zéro (`flow_NN.tif`).
3. Fig. 19 (c) Algorithme proposé $r = 4$ (`flow_fe_p24_r4.tif`).
4. Fig. 19 (d) Algorithme proposé $r = 8$ (`flow_fe_p24_r8.tif`).

Au tableau 2 on montre les résultats pour le calcul du PSNR et le nombre des itérations utilisées.

	"Ordre 0"			Méthode proposée $r = 4$			Méthode proposée $r = 8$		
	Y	C1	C2	Y	C1	C2	Y	C1	C2
PSNR	49.08	74.57	71.93	95.07	113.79	111.68	87.85	106.63	102.81
Itérations	0	0	0	9	10	13	18	16	16

TAB. 2 – PSNR de l'image reconstruite (db)

Ces résultats montrent l'amélioration donnée par l'approche proposée par rapport à l'interpolation d'ordre 0. On voit que les images obtenues sont proches de l'image originale. Dans la reconstruction avec $r = 8$ on voit des oscillations à quelques endroits.

La reconstruction d'une séquence d'images intermédiaires pour cette paire d'images stéréoscopiques peut être visualisée avec le système VIEWSTORE au laboratoire du groupe VISCOM. On a fait la reconstruction avec la méthode proposée et la méthode de propagation des vecteurs. À la Fig. 20 a (fichier `flow_seq_reg_025.tif`) on voit

un détail de l'image reconstruite avec la méthode de propagation des vecteurs proposée par Konrad et à la Fig. 20 b (fichier `flow_seq_fe_p24_r4_025.tif`) un détail de l'image reconstruite avec la méthode proposée.

Toutes les images sont dans le répertoire:

```
INRS /home/viscom/vazquez/Rep_Images/TIF
-rwxr--r--  1 viscom   131308 Aug 19 14:15 bouquet.tif
-rwxr--r--  1 viscom   131260 Aug 19 16:50 bouquet_e_NN_b.tif
-rwxr--r--  1 viscom   131308 Aug 19 16:46 bouquet_e_b_b.tif
-rwxr--r--  1 viscom   131260 Sep  6 15:20 bouquet_e_fe_p3744_r4_b.tif
-rwxr--r--  1 viscom   131260 Aug 19 16:53 bouquet_r_NN_b.tif
-rwxr--r--  1 viscom   134155 Aug 19 17:14 bouquet_r_b_b.tif
-rwxr--r--  1 viscom   131260 Sep  4 21:52 bouquet_r_fe_p44_i3_b.tif
-rwxr--r--  1 viscom    38718 Aug 19 17:34 flow_NN.tif
-rwxr--r--  1 viscom    38718 Aug 19 17:30 flow_fe_p24_r4.tif
-rwxr--r--  1 viscom    38718 Aug 19 17:32 flow_fe_p24_r8.tif
-rwxr--r--  1 viscom   518734 Aug 26 18:48 flow_left.tif
-rwxr--r--  1 viscom   518736 Aug 26 18:49 flow_right.tif
-rwxr--r--  1 viscom   518716 Aug 26 18:22 flow_seq_fe_p24_r4_025.tif
-rwxr--r--  1 viscom   518716 Aug 26 18:22 flow_seq_reg_025.tif
-rwxr--r--  1 viscom   131310 Aug 19 14:07 pcs.tif
-rwxr--r--  1 viscom   131260 Aug 19 17:15 pcs_e_NN_b.tif
-rwxr--r--  1 viscom   131310 Aug 19 17:19 pcs_e_b_b.tif
-rwxr--r--  1 viscom   131260 Sep  6 15:21 pcs_e_fe_p3744_r4_b.tif
-rwxr--r--  1 viscom   131260 Aug 19 17:18 pcs_r_NN_b.tif
-rwxr--r--  1 viscom   138175 Aug 19 17:29 pcs_r_b_b.tif
-rwxr--r--  1 viscom   131260 Aug 19 17:17 pcs_r_fe_p37_r4_b.tif
```

Les fichiers sont nommés selon la convention (nom)-(transformation)-(méthode) et sont en format TIFF. Le nom peut être *pcs* pour le patron circulaire sinusoïdal, *bouquet* pour l'image Bouquet ou *flow* pour l'image Flower. La transformation pourrait être *r* pour la rotation, *e* pour le "zoom", rien pour la compensation en disparité de l'image droite et *seq* pour la reconstruction de la vue avec $\rho = 0.25$. La méthode pourrait être *b* pour la méthode classique de transformation en arrière, *NN* pour une interpolation d'ordre 0, *fe* pour la méthode proposée. La *r* suivie d'un numéro indique le rapport de densité entre les grilles plus dense et moins dense.

6 Travail futur

Les résultats obtenus jusqu'à présent permettent de prévoir la possibilité d'utilisation de la méthode proposée dans les applications envisagées. Les expériences ont été très limitées et, dans le futur, il faudra faire beaucoup plus d'expériences et trouver des solutions aux problèmes actuels et à ceux qui vont se présenter. Il y a plusieurs éléments qui peuvent intervenir dans la performance de l'algorithme; on devra dans le futur faire des expériences pour trouver la meilleure combinaison de ces éléments du point de vue de la performance.

D'abord, il faudra améliorer l'algorithme de reconstruction en cherchant la base qui donne les meilleurs résultats. On a proposé des bases d'ondelettes biorthogonales de type spline. Plusieurs paramètres peuvent être changés pour obtenir la meilleure performance. Le degré des splines, la régularité et le support des fonctions d'échelles et ondelettes et les filtres h, \tilde{h} , g et \tilde{g} dans le banc de filtres doivent être choisis avec soin. On doit faire des expériences avec des différents filtres et fonctions de base pour déterminer lesquels donnent les meilleurs résultats.

Un autre sujet très important est la sélection de l'opérateur de projection. Le processus de projection peut être fait en tenant compte des caractéristiques des images. Les images sont plutôt régulières avec des hautes fréquences aux frontières des objets. Les coefficients d'ondelettes seront grands sur les frontières et relativement petits dans les plages régulières à l'intérieur des objets. On pourra modifier les coefficients ondelettes pour améliorer la qualité de la solution. Les coefficients ondelettes responsables de l'introduction des hautes fréquences dans les zones régulières, peuvent être enlevés pour obtenir une solution plus régulière à l'intérieur des objets, en conservant les coefficients aux frontières. Il faudra définir quels coefficients modifier et quelle modification introduire.

Un sujet qui doit recevoir une attention spéciale est le traitement des zones d'occlusion et nouvellement exposées dans la compensation. Si le processus de compensation n'est pas bon, les données de départ ne correspondront pas à la vraie image qu'on veut reconstruire. Le processus de compensation et d'estimation de la transformation sont alors des sujets à traiter. Ici, plusieurs approches sont possibles, mais elles dépendent de l'application. Par exemple, pour la reconstruction des vues intermédiaires, la qualité de l'estimation de la disparité est très importante parce que l'information de profondeur est justement contenue dans les vecteurs de disparité. Par contre, pour le codage vidéo une estimation qui donne la plus petite erreur sera peut être la meilleure. Pour les occlusions, le problème est de faire une bonne estimation de la transformation subie par ces points en utilisant l'information de leurs voisins. La solution est donc au niveau de l'estimation de la transformation, un sujet bien traité dans la littérature et qui reste ouvert à la recherche. Les zones nouvelles sont un peu différentes. La conséquence est qu'il y aura des trous dans l'image à reconstruire mais la grandeur de ces trous dépend aussi de l'algorithme d'estimation de la transformation. Si la contrainte de régularité dans l'estimation est forte, les trous seront petits, mais pour des estimations avec des frontières bien définies les trous seront plus grands. On

pourrait penser à mettre quelques points de la vraie image à ces zones pour garantir la reconstruction. Pour la reconstruction des vues intermédiaires, une solution est de faire une estimation de la disparité dans le sens inverse pour toute l'image ou seulement pour les zones nouvelles. Avec deux champs de disparité on aura plus de points pour la reconstruction et en plus on pourra faire l'estimation de ces deux champs de façon simultanée pour avoir une meilleure approximation du vrai champ de disparité. Ça coûterait, bien sûr, plus cher du point de vue du calcul et donc du temps.

Le plan à suivre pour les prochains mois:

1. Optimisation de la base à utiliser: Il y a plusieurs éléments qui interviennent dans la qualité de la reconstruction. La régularité des filtres, leur support, le degré d'approximation des fonctions de base sont des caractéristiques à tenir en compte au moment de la sélection des filtres. Il faudra faire des expériences pour trouver la meilleure combinaison pour l'algorithme de point de vue de la performance.

Durée estimée: 2 mois

2. Définition de l'opérateur de projection. L'opérateur de projection est très important dans le processus itératif. L'utilisation des coefficients ondelettes pourra aider à éliminer les distorsion de haute fréquence qu'on a maintenant. Il faut faire des épreuves pour déterminer les valeurs optimales des paramètres qui interviennent dans le processus.

Durée estimée: 2 mois

3. Étude des méthodes d'accélération et d'amélioration de la mise en oeuvre pour réduire le temps de reconstruction. Le temps est aussi un élément important. Des méthodes pour augmenter la vitesse de convergence de l'algorithme pourraient être testées.

Durée estimée: 3 mois

4. Traitement des zones d'occlusion et nouvellement exposées. Les caractéristiques des champs de disparité et du mouvement importantes pour ce type d'algorithmes doivent être étudiées. Trouver des solutions aux problèmes de zones cachées et nouvelles est essentiel pour appliquer l'algorithme dans le codage vidéo et la reconstruction des vues intermédiaires.

Durée estimée: 3 mois

5. Amélioration du processus de compensation de la disparité. Dans le cas de la reconstruction des vues intermédiaires l'estimation de disparité et la compensation de la disparité sont essentiels pour garantir la qualité de la reconstruction. L'utilisation de deux champs de disparité, la possibilité d'améliorer les champs de disparité en utilisant les deux images et la méthode d'interpolation pour calculer les valeurs compensées doivent être étudiés.

Durée estimée: 3 mois

6. Mise en œuvre d'un système de reconstruction des vues intermédiaires ou de codage vidéo pour valider les résultats.

Durée estimée: 3 mois

7. Rédaction de la thèse, correction et soutenance.

Durée estimée: 6 mois

7 Conclusions

Nous avons présenté une méthode de reconstruction des images à partir des échantillons sur une grille irrégulière obtenus comme résultat de la compensation d'une transformation spatiale. La méthode est basée sur un algorithme itératif qui agit sur des espaces d'ondelettes du type spline en améliorant une première approximation obtenue par un processus simple d'interpolation sur une grille régulière plus dense que la grille originale.

Quelques éléments peuvent être mentionnés en faveur de cette approche:

1. L'utilisation d'un algorithme itératif permet de contrôler la qualité de la reconstruction. Avec plus d'itérations il est possible d'obtenir des meilleures approximations de l'image.
2. La transformée en ondelettes peut être calculée à l'aide des bancs de filtres de façon rapide. Le processus est alors réduit à des filtrages numériques successifs des images, un sujet bien connu et avec des algorithmes efficaces.
3. L'utilisation des bases biorthogonales à support fini permet d'utiliser filtres RIF facilitant le processus de filtrage.
4. L'utilisation des splines garantit la régularité de la solution et des bonnes caractéristiques d'approximation.

Les solutions à ce problème trouvées dans la littérature utilisent de façon générale un modèle d'image limité en bande. Ce modèle a plusieurs problèmes qu'on a déjà vu et qui ne sont pas présentes avec un modèle basé en ondelettes.

Cet algorithme va nous permettre de développer des nouvelles méthodes pour la reconstruction des vues intermédiaires d'images stéréoscopiques. Les méthodes proposées dans la littérature ont plusieurs limitations pour leur application. Les méthodes basées sur la construction d'un modèle 3D de la scène deviennent trop complexes pour les images tirées de la vie réelle. La méthode d'estimer un champ de disparité pour chaque position à reconstruire n'est pas toujours applicable et la complexité de l'estimation de la disparité le rend peu pratique si on doit reconstruire plusieurs vues. Enfin, la méthode de propagation des vecteurs de disparité ne donne pas une bonne qualité si le champ de disparité a des frontières bien définies ou si les disparités sont grandes. En utilisant l'algorithme proposé on pourra reconstruire plusieurs vues intermédiaires avec l'estimation d'un ou deux champs de disparité. La complexité de calcul ne sera pas si grande que dans le cas de l'estimation d'un champ de disparité par position à reconstruire.

On pourra aussi améliorer la performance des méthodes de codage vidéo non récursives qui utilisent les trajectoires de mouvement. Étant donné qu'en utilisant des précision en dessous du pixel les trajectoires du mouvement sont plus précises, le codage de ces trajectoires peut être fait avec plus d'efficacité, réduisant le débit nécessaire pour la transmission. Les méthodes de codage vidéo non récursives avec

compensation du mouvement qu'on trouve dans la littérature sont limités par l'utilisation d'une précision dans l'estimation du mouvement d'un pixel. La conséquence est que le nombre d'images qu'on peut utiliser dans le codage demeure petite (seulement 2 dans les travaux de Choi et Woods). Avec une bonne méthode de reconstruction pour les images irrégulièrement échantillonnée obtenues comme résultat de la compensation, ce nombre pourra devenir plus grand et donc la performance de la méthode pourrait être meilleure.

Il reste encore beaucoup de travail à faire. Les résultats obtenus même si encourageants ne sont pas suffisants pour les applications envisagés. Dans le futur on propose de travailler sur l'optimisation de la méthode et son application et si les résultats sont comme prévus on envisage des contributions aux sujets de la reconstruction des vues intermédiaires et du codage vidéo au niveau de la performance.

Références

- [1] A. Aldroubi and H. Feichtinger, “Complete iterative reconstruction algorithms for irregularly sampled data in spline-like spaces,” in *ICASSP 97*, pp. 1857–1860, 1997.
- [2] H. Aydinoglu and M. Hayes, “Stereo image coding: A projection approach,” *IEEE Transaction on Image Processing*, vol. 7, pp. 506–516, Apr 1998.
- [3] J. Benedetto, “Frame decompositions, sampling and uncertainty principle inequalities,” in *Wavelet: Mathematics and Applications* (J. Benedetto and M. Frazier, eds.), ch. 7, pp. 247–304, Boca Raton FL: CRC Press, 1993.
- [4] N. Chang and A. Zakhor, “View generation for three-dimensional scenes from video sequences,” *IEEE Trans. Image Process.*, vol. 6, pp. 584–598, Apr. 1997.
- [5] W. Chen, S. Itoh, and J. Shiki, “Irregular sampling theorems for wavelet subspaces,” *IEEE Trans. Inf. Theory*, vol. 44, pp. 1131–1142, May 1998.
- [6] S.-J. Choi and J. Woods, “Three-dimensional subband/wavelet coding of video with motion compensation,” in *SPIE Visual Communications and Images Processing*, (San Jose, CA), pp. 96–104, Feb. 1997.
- [7] S.-J. Choi and J. Woods, “Motion compensated 3-d subband coding of video,” *IEEE Trans. Image Process.*, vol. 8, pp. 155–167, Feb. 1999.
- [8] R. Duffin and A. Schaeffer, “A class of nonharmonic fourier series,” *Trans. Amer. Math. Soc.*, vol. 72, pp. 341–366, 1952.
- [9] H. Feichtinger and K. Gröchenig, “Theory and practice of irregular sampling,” in *Wavelet: Mathematics and Applications* (J. Benedetto and M. Frazier, eds.), ch. 8, pp. 305–363, Boca Raton FL: CRC Press, 1994.
- [10] H. Feichtinger, C. Cenker, M. Mayer, H. Steier, and T. Strohmer, “New variants of the pocs method using affine subspaces of finite codimension, with applications to irregular sampling,” in *Proc. SPIE-92*, (Boston), pp. 299–310, 1992.
- [11] H. Feichtinger and T. Strohmer, “Fast iterative reconstruction of band-limited images from irregular sampling values,” in *Computer Analysis of Images and Patterns*, pp. 82–91, Budapest: Conf.CAIP Budapest 93, 1993.
- [12] M. Foodei, *Entropy-Constrained Recursive Vector Quantization and Application to Image Sequence Coding*. PhD thesis, Department of Electrical Engineering, McGill University, Montreal, Canada, 1995.
- [13] B. Girod, “Motion compensated prediction with fractional-pel accuracy,” *IEEE Trans. Commun.*, vol. 42, pp. 604–612, Apr 1993.

-
- [14] K. Gröchenig, "Acceleration of the frame algorithm," *IEEE Trans. Signal Process.*, vol. 41, pp. 3331–3340, Dec. 1993.
- [15] R. Keys, "Cubic convolution interpolation for digital images processing," *IEEE-T-ASSP*, vol. ASSP-29, pp. 1153–1160, Dec. 1981.
- [16] J. Konrad, "Enhancement of viewer comfort in stereoscopic viewing: parallax adjustment," in *Proc. IS&T/SPIE Symposium on Electronic Imaging Stereoscopic Displays and Virtual Reality Syst.*, vol. 3639, (San José, CA, USA), Jan 1999.
- [17] J. Konrad, "View reconstruction for 3-d video entertainment: Issues, algorithms and applications," in *Proc. Int. Conf. on Image Process. and its Applications*, (Manchester, UK), Jul 1999.
- [18] Y. Liu, "Irregular sampling for spline wavelet," *IEEE Trans. Inf. Theory*, vol. 42, pp. 623–627, Feb. 1996.
- [19] S. Mallat, *A Wavelet Tour of Signal Processing*. Academic Press, 1998.
- [20] A. Mancini and J. Konrad, "Robust quadtree-based disparity estimation for the reconstruction of intermediate stereoscopic images," in *Proc. SPIE Stereoscopic Displays and Virtual Reality Systems*, vol. 3295, (San José, CA, USA), pp. 53–64, Jan 1998.
- [21] A.-R. Mansouri and J. Konrad, "Block based winner-takes-all reconstruction of intermediate stereoscopic images," in *Proc. SPIE Visual Communication and Image Processing*, vol. 3309, (San José, CA, USA), pp. 922–933, Jan 1998.
- [22] F. Marvasti, M. Analoui, and M. Gamshadzahi, "Recovery of signals from nonuniform samples using iterative methods," *IEEE Trans. Signal Process.*, vol. 39, pp. 872–878, Apr. 1991.
- [23] F. Marvasti, C. Liu, and G. Adams, "Analysis and recovery of multidimensional signals from irregular samples using nonlinear and iterative techniques," *Signal Process.*, vol. 36, pp. 13–30, 1994.
- [24] E. Meijering, K. Zuiderveld, and M. Viergever, "Image reconstruction by convolution with symmetrical piecewise n th-order polynomial kernels," *IEEE Trans. Image Process.*, vol. 8, pp. 192–201, Feb. 1999.
- [25] Y. Nakaya and H. Harashima, "Motion compensation based on spatial transformation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 4, pp. 339–356, Jun 1994.
- [26] J.-R. Ohm, "Three-dimensional subband coding with motion compensation," *IEEE Transaction on Image Processing*, vol. 3, pp. 559–571, Sep 1994.

-
- [27] M. Perkins, "Data compression of stereopairs," *IEEE Transactions on Communications*, vol. 40, pp. 684–696, Apr 1992.
 - [28] J. Proakis and D. Manolakis, *Digital Signal Processing*. Prentice Hall, third ed., 1996.
 - [29] A. Sharaf and F. Marvasti, "Motion compensation using spatial transformations with forward mapping," *Signal Processing: Image Communication*, vol. 14, pp. 209–227, Jan. 1999.
 - [30] W. Tam and L. Stelmach, "Stereo depth perception in a sample of young television viewers," in *Int. Workshop on Stereoscopic and 3D Imaging*, (Santorini, Greece), 1995.
 - [31] M. Tekalp, *Digital Video Processing*. Prentice Hall PTR, 1996.
 - [32] M. Unser, "Ten good reasons for using spline wavelets," in *SPIE Proceedings*, vol. 3169 of *Wavelets Applications in Signal and Images Processing V*, pp. 422–431, 1997.
 - [33] M. Unser, "Splines: A perfect fit for signal processing." Submitted to *IEEE Signal Processing Magazin*, Nov. 1998.
 - [34] M. Unser, A. Aldroubi, and M. Eden, "Fast B-spline transforms for continuous image representation and interpolation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, pp. 277–285, Mar. 1991.
 - [35] M. Unser, A. Aldroubi, and M. Eden, "B-spline signal processing: Part–II efficient design and application," *IEEE Trans. Signal Process.*, vol. 41, pp. 834–848, Feb. 1993.
 - [36] S. Yeh and H. Stark, "Iterative and one-step reconstruction from nonuniform samples by convex projections," *J. Opt. Soc. Amer. A*, vol. 7, no. 3, pp. 491–499, 1990.

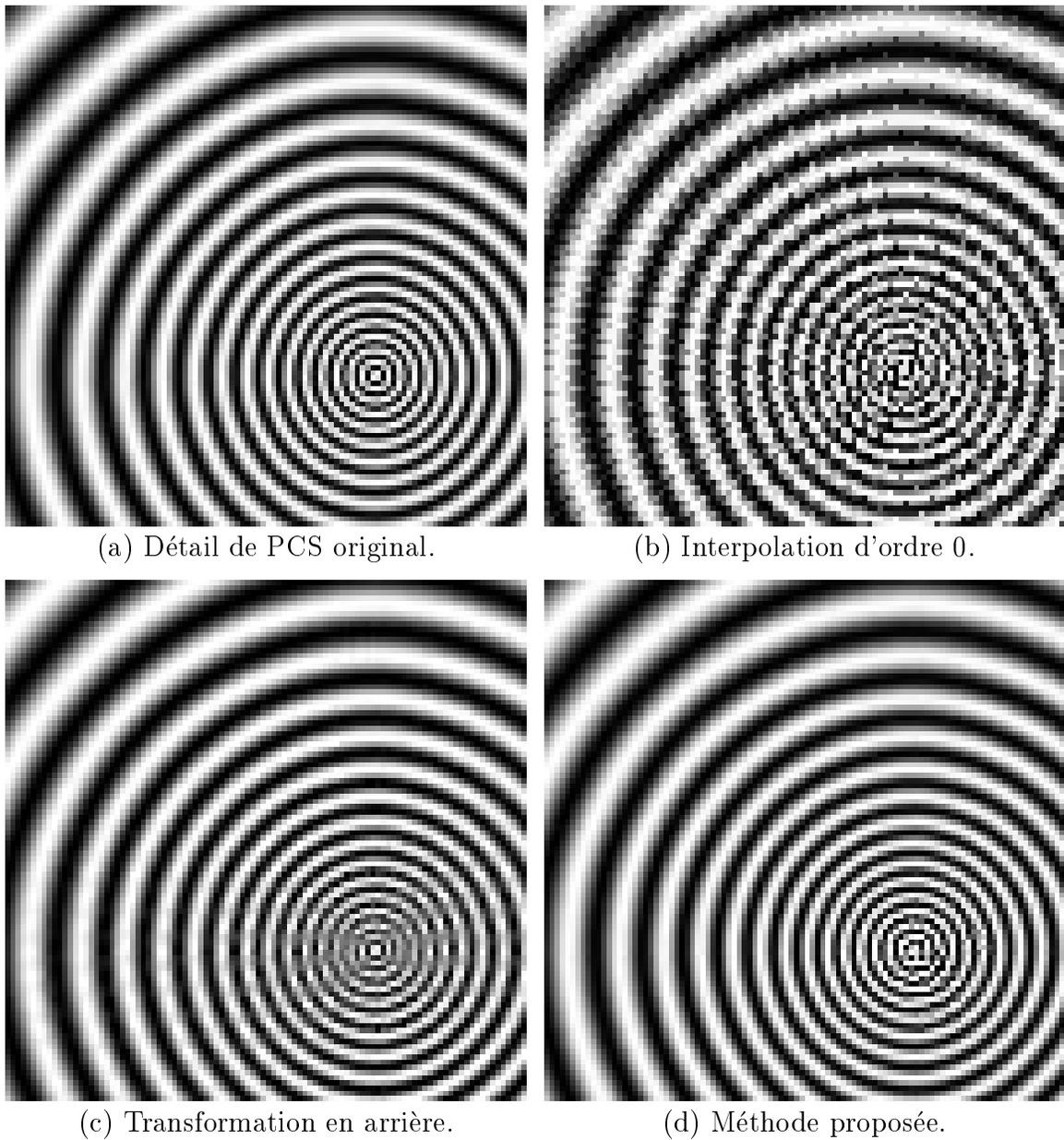


FIG. 15 – Reconstruction de l'image PCS après rotation

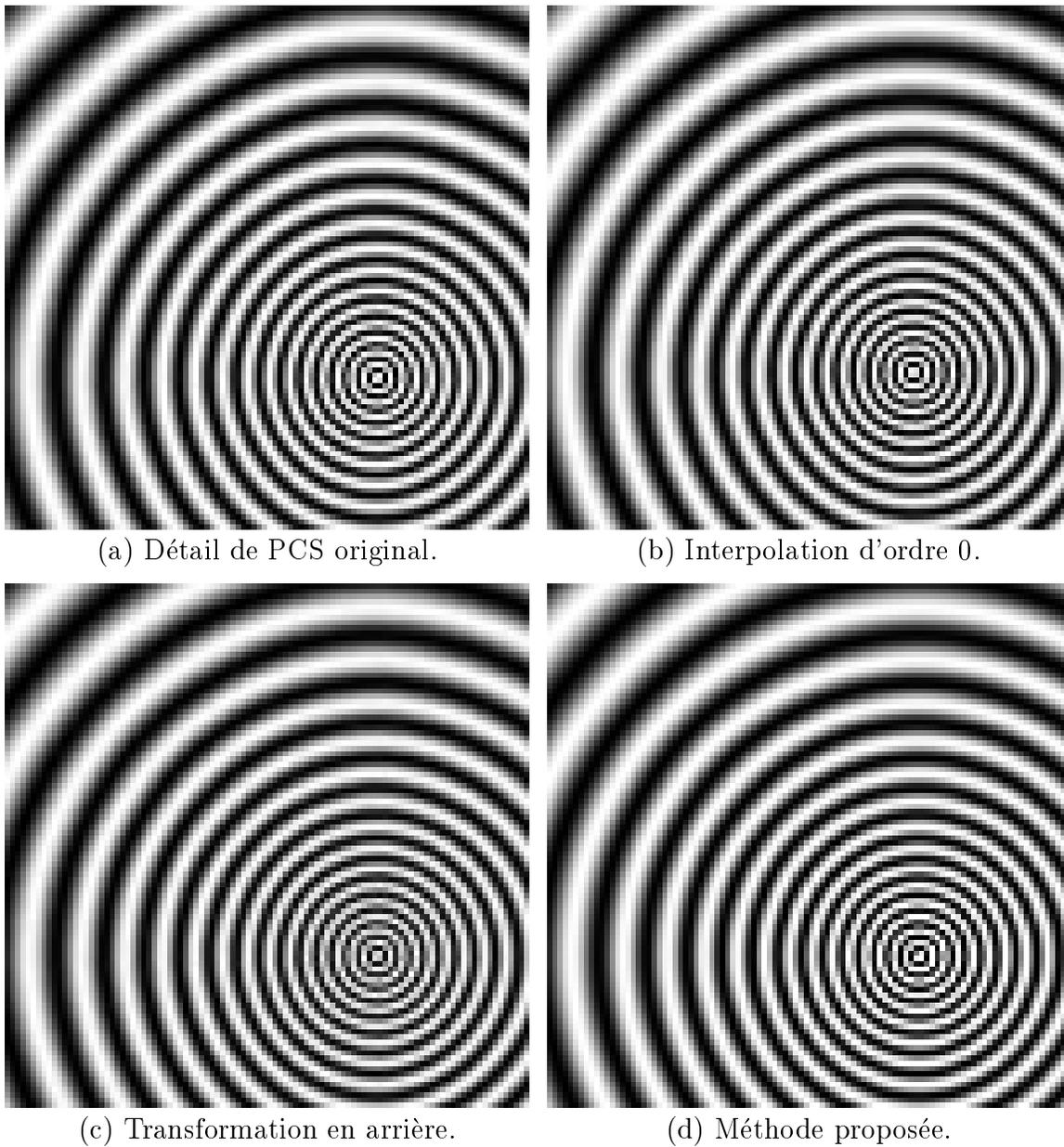


FIG. 16 – *Reconstruction de l'image PCS après "zoom"*



(a) Détail de l'image Bouquet originale .



(b) Interpolation d'ordre 0.



(c) Transformation en arrière.



(d) Méthode proposée.

FIG. 17 – *Reconstruction de l'image Bouquet après rotation*



(a) Détail de l'image Bouquet original.



(b) Interpolation d'ordre 0.



(c) Transformation en arrière.



(d) Méthode proposée.

FIG. 18 – *Reconstruction de l'image Bouquet après "zoom"*



(a) Détail de l'image droite.



(b) Interpolation d'ordre 0.

(c) Méthode proposée $r = 4$.(d) Méthode proposée $r = 8$.

FIG. 19 – Reconstruction de l'image droite de Flower avec compensation de disparité



(a) Propagation des vecteurs.

(b) Méthode proposé.

FIG. 20 – *Reconstruction d'une vue intermédiaire ($\rho = 0.25$)*