Méthode non-additive intervalliste de super-résolution d'images, dans un contexte semi-aveugle

Farès Graba

ACADÉMIE DE MONTPELLIER
UNIVERSITÉ **D**E **M**ONTPELLIER
Sciences et **T**echniques du **L**anguedoc

THÈSE

présentée au Laboratoire d'Informatique de Robotique
et de Microélectronique de Montpellier pour
obtenir le diplôme de doctorat

Spécialité	:	**Génie Informatique, Automatique et Traitement du signal**
Formation Doctorale	:	**Systèmes automatiques**
École Doctorale	:	**Information, Structures, Systèmes**

Méthode Non-additive Intervalliste de Super-résolution d'Images, dans un contexte Semi-aveugle

par

Farès GRABA

Soutenance prévue pour le 17 avril 2015, devant le jury composé de :

Directeur de thèse
M. Olivier STRAUSS, Maître de conférences HDR . . . . . LIRMM, Université de Montpellier

Co-encadrant
M. Frédéric COMBY, Maître de conférences . . . . . LIRMM, Université de Montpellier

Rapporteurs
M. Didier DUBOIS, Directeur de recherche CNRS . . . . . Institut de Recherche en Informatique de Toulouse
M. Ali MOHAMMAD-DJAFARI, Directeur de recherche CNRS . . . . . École supérieur d'électricité - Université Paris sud 11

Examinateurs
Mme. Isabelle BLOCH, Professeur des universités . . . . . Télécom ParisTech
M. Jean François GIOVANNELLI, Professeur des universités . . . . . Université de Bordeaux
M. William PUECH, Professeur des universités . . . . . Université de Montpellier

Table des matières

Introduction générale

Nous avons tous déjà vu, au moins une fois, un épisode d'une série policière faisant la promotion d'un petit génie de l'informatique utilisant des logiciels sophistiqués lui permettant de lire le nom d'un suspect, sur sa carte d'identité, alors que celle-ci était posée négligemment sur une commode visible à travers une fenêtre dont l'image n'occupait que quelques pixels d'une image satellitaire de mauvaise qualité. Ce type de scène relève actuellement de la fiction, mais un tel exploit est-il possible dans le futur ?

La technique à la quelle ce genre de scène fait référence porte le nom de "super-résolution d'images". La super-résolution consiste en l'amélioration, ou l'augmentation de la résolution d'images acquises avec des imageurs faiblement résolus. Elle est déjà appliquée dans de nombreux domaines utilisant l'imagerie numérique comme l'imagerie médicale, la vidéo-surveillance et la microscopie.

Dans tous les domaines d'application de l'imagerie numérique, l'objectif final est d'extraire de l'information à partir des images acquises soit de manière visuelle, ou bien informatique.

Cependant, il existe de nombreuses situations où les dégradations introduites dans l'image durant l'acquisition sont trop importantes pour pouvoir extraire correctement l'information recherchée. Ces dégradations peuvent être dues, par exemple, aux conditions difficiles d'acquisition, ou simplement à la nature même de l'imageur utilisé. On peut citer trois types de dégradation dans les images numériques :

- Le bruit, dont les causes principales sont les erreurs de transmission, les moyens d'enregistrement, les bruits de mesure, et le bruit de quantification.
- La perte de résolution optique (ou réduction de la bande passante), dont les causes principales sont la Réponse Impulsionnelle (RI) de l'imageur, regroupant l'effet du système optique et l'intégration sur les détecteurs CCD.
- La perte de résolution numérique (ou spatiale), due à l'échantillonnage clairsemé

de l'image. Il se traduit par le phénomène de repliement de spectre (aliasage) dans l'image acquise.

Il arrive souvent qu'on ne puisse pas améliorer la résolution de l'imageur, soit pour des raisons de coût (les systèmes optiques sophistiqués, par exemple, coûtent très cher), ou bien à cause des limites physiques. Par exemple, une amélioration de la résolution numérique d'un imageur consisterait à augmenter son nombre de pixels, c.à.d. le nombre de détecteurs photosensibles du capteur CCD, afin de rendre l'échantillonnage de l'image plus dense. Malheureusement, cette augmentation implique une diminution de la taille de ces détecteurs, ce qui conduit à une diminution de la quantité de lumière collectée par chaque détecteur et donc une détérioration du rapport signal/bruit.

C'est justement là qu'intervient la SR pour essayer de pallier à l'impossibilité technologique ou économique de réaliser des images de résolution suffisante. Elle est généralement basée sur l'exploitation de la complémentarité de l'information spatiale présente dans les images d'une même scène, entre lesquelles il existe des mouvements. Le principe est basé sur l'idée suivante : *au lieu d'essayer d'augmenter le nombre de pixels d'un imageur, au risque de réduire le rapport signal/bruit, on augmente le nombre d'acquisitions d'une scène en bougeant légèrement l'imageur (ou le sujet) avant chaque acquisition. En veillant à ce que la position du capteur par rapport à la scène (ou au sujet), lors d'une acquisition, soit différente par rapport à toutes les acquisitions précédentes, l'information spatiale contenue dans l'image acquise sera différente de celle des images précédentes et donc complémentaire.* Afin de réaliser la restauration, ces techniques nécessitent une estimation préalable des mouvements dominants entre les images avec une précision sous-pixélique.

Le problème de super-résolution est souvent modélisé comme un problème inverse linéaire : on représente la relation entre l'image hautement résolue qu'on cherche à restaurer et les images bassement résolues acquises par un modèle de dégradation linéaire, appelé aussi modèle d'acquisition, comprenant les mouvements existants entre l'image hautement résolue recherchée et les images bassement résolues, l'effet de la RI de l'imageur lors de l'acquisition, le bruit ajouté et la perte de résolution numérique. La restauration se fait par inversion du modèle de dégradation.

Un des aspects caractérisant le problème de super-résolution est qu'il s'agit d'un problème inverse mal posé au sens de Hadamard [Bertero *et al.*, 1988], c.à.d. qu'il peut admettre plusieurs solutions ou aucune. En outre, la solution ne dépend pas continument des données, on parle alors de mauvais conditionnement. Ceci a pour conséquence qu'une petite erreur dans le modèle de dégradation considéré peut conduire à des erreurs importantes dans la solution estimée. Cette caractéristique rend le problème de super-résolution particulièrement difficile à résoudre. En effet, dans le modèle d'acquisition, l'effet de la RI de l'imageur et le modèle du bruit dépendent des caractéristiques intrinsèques de l'imageur. Leur estimation exacte et précise est très souvent impossible à réaliser car elle constitue, elle aussi, un problème mal posé.

Un certain nombre de méthodes ont été développées afin de réduire les artefacts induits par ces erreurs de modélisation dans la solution estimée. Elles sont généralement

basées sur des connaissances à priori (ou hypothèses) sur la statistique et/ou la structure de la solution recherchée. Ces méthodes ont été unifiées par la théorie de la régularisation des problèmes inverses mal posés. Malheureusement, ces méthodes peuvent parfois avoir un effet contradictoire avec l'objectif même de la super-résolution dans le sens où elles tendent à pénaliser les hautes fréquences tandis que la super-résolution a pour but de les recouvrer. De plus, aucune de ces méthodes ne modélise proprement le fait que la RI de l'imageur soit mal connue.

Dans ce manuscrit, nous proposons de réduire l'influence d'une mauvaise modélisation de la RI de l'imageur en utilisant une représentation imprécise de celle-ci. Cette représentation, dérivée de la théorie des probabilités imprécises, conduit à un modèle non linéaire et une méthode de super-résolution aboutissant à une solution imprécise (ou intervalliste) reflétant la connaissance imprécise sur la RI de l'imageur. La solution estimée, en utilisant une faible connaissance sur la RI, présente plusieurs caractéristiques intéressantes parmi lesquelles une certaine robustesse et une quantification de l'erreur d'estimation.

La suite du manuscrit est organisée comme suit : dans le chapitre 1 nous présentons un panorama des méthodes de restauration d'images par super-résolution et nous proposons une discussion concernant la modélisation et l'influence de la RI de l'imageur. Dans le chapitre 2, nous introduisons les concepts utilisés dans la méthode de super-résolution que nous proposons, dérivés de la théorie des probabilités imprécises, et quelques opérateurs associés. Parmi ces concepts, nous verrons notamment celui des noyaux maxitifs permettant de modéliser la connaissance imprécise de la RI de l'imageur et celui des partitions floues permettant d'obtenir des modèles continus des images discrètes manipulées. Dans le chapitre 3, nous présentons en détail la démarche que nous avons adoptée pour développer une méthode de super-résolution imprécise utilisant uniquement une connaissance partielle sur la RI de l'imageur. Dans le chapitre 4, nous présentons en détail une implémentation simple de la méthode que nous proposons et une étude sur la complexité algorithmique de celle-ci. Dans le chapitre 5, nous illustrons les performances de la méthode que nous proposons, et les comparons à celles de quelques autres méthodes de super-résolution, dans un contexte semi-aveugle, c.à.d. lorsque la RI de l'imageur est mal connue. Nous terminons ensuite par une conclusion générale et présentons quelques perspectives.

CHAPITRE

1 État de l'art

Depuis environ une trentaine d'années, la communauté scientifique s'intéresse à l'augmentation de la résolution des images dans le but d'outrepasser les limites imposées par les capteurs. Les premiers travaux sur la reconstruction d'images de haute résolution à partir de séquences d'images basse résolution ont été publiés en 1984 dans [Tsai et Huang, 1984]. Le terme de "Super-résolution", quant à lui, est apparu aux alentours des années 1990 dans les travaux de Irani et Peleg [Irani et Peleg, 1990]. Les auteurs définissent un procédé s'apparentant à la reconstruction d'images à partir de projections, utilisée en tomographie assistée par ordinateur.

L'intérêt porté à la super-résolution connait une croissance significative dans la communauté du traitement d'images. Un nombre conséquent travaux portant sur ce sujet ont été publiés durant ces deux dernières décennies. On retrouve des études assez complètes des techniques de super-résolution proposées par la littérature comme par exemple celle de Tian et Ma en 2011 [Tian et Ma, 2011] et l'ouvrage proposé par [Milanfar, 2010]. Ce dernier dresse un panorama quasi exhaustif des approches et techniques d'amélioration de la résolution, tant en imagerie classique qu'en vidéo, proposées avant l'année 2011. Nous nous devons de citer aussi deux études plus anciennes, [Borman et Stevenson, 1998] et [Park *et al.*, 2003], qui présentent les méthodes de super-résolution proposées sur la période couvrant des années 1984 au début des années 2000.

Les approches proposées par la communauté du traitement d'images diffèrent essentiellement par leur formulation du problème (mise en équation) et par les méthodes de résolution sur lesquelles elles s'appuient (outils utilisés et traitement du problème). Notre étude bibliographique nous a permis de distinguer 5 grandes familles de techniques de super-résolution. Elles sont présentées ci-dessous et sont détaillées dans la suite de ce chapitre. Nous les avons appelées respectivement :

1. Approches formulées dans le domaine fréquentiel,

2. Approches par interpolation puis restauration,
3. Approches par optimisation,
4. Projection sur des ensembles convexes,
5. Approches par apprentissage.

Toutes ces méthodes visent à reconstruire une image hautement résolue à partir d'une ou plusieurs images bassement résolues. Réaliser une telle reconstruction implique une bonne compréhension de ce qu'est la résolution d'un imageur, d'une part, et de ce qui est possible ou pas possible de réaliser, d'autre part. Le concept de résolution, pour un capteur d'images, s'appuie sur ce que l'on appelle communément un modèle d'observation. Un modèle d'observation est ce qui lie l'image numérique acquise à l'image réelle continue, c.à.d. l'illumination du capteur par la scène. Il fait généralement intervenir un opérateur modélisant le flou induit par le système optique de l'imageur et la réponse impulsionnelle de ses capteurs photosensibles, d'une part, et un opérateur d'échantillonnage d'autre part. En super-résolution, on tend à appeler "modèle d'observation" le lien existant entre l'image hautement résolue à reconstruire et et les images bassement résolues utilisées pour réaliser cette reconstruction. Au modèle initial se rajoute alors le positionnement relatif des images bassement résolues par rapport à l'image réelle continue (ou par rapport à l'image hautement résolue à reconstruire) et le changement de résolution. Enfin, on rajoute un modèle de bruit pour prendre en compte les erreurs de modélisation et les variations aléatoires des mesures de luminance. Le modèle le plus utilisé est linéaire [Elad et Feuer, 1997, Nguyen *et al.*, 2001a] et s'écrit :

$$\mathbf{Y}^k = DH^kW^k\mathbf{X} + \mathbf{B}^k, \tag{1.1}$$

où $\mathbf{Y}^k$ est le vecteur des valeurs d'illumination de la $k^{ième}$ image bassement résolue acquise ($k = 1 \dots K$), $\mathbf{X}$ est le vecteur des valeurs d'illumination de l'image hautement résolue recherchée, souvent supposée être échantillonnée en respectant le critère de Shannon-Nyquis, W^k est une matrice modélisant l'effet de mouvement de la caméra, qui recale l'image hautement résolue recherché sur la version interpolée de la $k^{ième}$ image bassement résolue, H^k est une matrice modélisant l'effet de flou induit par la RI de l'imageur, D est l'opérateur de décimation et $\mathbf{B}^k$ représente le bruit. $\mathbf{Y}^k$ et $\mathbf{X}$ sont généralement représentés dans l'ordre lexicographique. L'Expression (1.1) est souvent écrite sous la forme

$$\mathbf{Y}^k = A^k\mathbf{X} + \mathbf{B}^k, \tag{1.2}$$

où $A^k = DH^kW^k$.

1.1 Approches formulées dans le domaine fréquentiel

Tsai et Huang ont abordé le problème de la super-résolution pour des images Landsat en le formulant dans le domaine de Fourier [Tsai et Huang, 1984]. Leur approche est basée

sur l'hypothèse que le signal continu originel de la scène est à bande limitée. Elle utilise la propriété de translation de la transformée de Fourier et la relation d'aliasage entre la Transformée de Fourier Continue (TFC) du signal originel de la scène et la Transformée de Fourier Discrète (TFD) des images bassement résolues acquises.

Soit $f(x,y)$ une illumination continue 2D et $f^k(x,y)$, où $k \in \{1,\dots K\}$, cette même illumination translatée de (δ_x^k, δ_y^k) :

$$f^k(x,y) = f(x+\delta_x^k, y+\delta_y^k),\ pour\ tout\ k \in \{1,\dots K\}. \tag{1.3}$$

Soit F^k la $k^{ième}$ image discrète bassement résolue de $M \times N$ pixels, échantillonnée avec une période (T_x, T_y). L'intensité du pixel (i,j) avec $(i,j) \in \{0,\dots M-1\} \times \{0,\dots N-1\}$ est donnée par :

$$F_{i,j}^k = f^k(iT_x, jT_y) = f(iT_x+\delta_x^k, jT_y+\delta_y^k). \tag{1.4}$$

Soit $\mathcal{F}^k$ la TFD de l'image F^k et $\mathfrak{f}^k$ la TFC du signal continu f^k. $\mathcal{F}^k$ et $\mathfrak{f}^k$ sont liées par la relation d'aliasage :

$$\mathcal{F}_{m,n}^k = \frac{1}{T_x T_y} \sum_{l_1=-\infty}^{\infty} \sum_{l_2=-\infty}^{\infty} \mathfrak{f}^k\left(\frac{2\pi}{T_x}(\frac{m}{M}+l_1), \frac{2\pi}{T_y}((\frac{n}{N}+l_2))\right). \tag{1.5}$$

Si f est à bande limitée, alors f^k l'est aussi et il existe des entiers L_x et L_y tels que $|\mathfrak{f}^k(u,v)| = 0$ pour $|u| \geq \frac{2\pi L_x}{T_x}$ et $|v| \geq \frac{2\pi L_y}{T_y}$. La somme infinie de l'équation (1.5) devient alors une somme finie.

D'autre part, si $\mathfrak{f}$ est la TFC de f, la propriété de la translation de la transformée de Fourier donne :

$$\mathfrak{f}^k(u,v) = e^{j2\pi(\delta_x^k m + \delta_y^k n)} \mathfrak{f}(u,v). \tag{1.6}$$

En exploitant les relations (1.5) et (1.6) et en ordonnant lexicographiquement les données, on obtient un système d'équations de type

$$A\mathbf{X} = \mathbf{B}, \tag{1.7}$$

où $\mathbf{B}$ est le vecteur des coefficients de la TFD des images observées et $\mathbf{X}$ est un vecteur contenant les échantillons inconnus des coefficients de la TFC du signal originel de la scène. A est la matrice qui définit la relation entre les deux vecteurs. Si le système (1.7) peut être résolu, il suffit alors d'appliquer la TFD inverse pour reconstruire une image hautement résolue.

Bien qu'elle soit simple et élégante, cette formulation ne prend en compte ni le flou introduit par la RI de l'imageur ni les erreurs de modélisation et le bruit de mesure. Elle a été étendue par [Kim *et al.*, 1990] par l'introduction d'un algorithme récursif de moindres carrés pondérés combinant la reconstruction et le filtrage du bruit, puis par [Kim et Su, 1993]

en prenant en compte le flou introduit par la RI. Kim et Su ont également raffiné l'algorithme de [Kim *et al.*, 1990] en incluant une mise à jour itérative du terme de régularisation de Tikhonov (voir Section 1.3.1) qui sert à stabiliser la solution. Afin de réduire le coût en mémoire et en temps de calcul, Rhee et Kang [Rhee et Kang, 1999] ont proposé une approche régularisée dans le domaine de la transformée en cosinus discrète. Plus tard, Woods *et al.* [Woods *et al.*, 2006] ont proposé deux méthodes fréquentielles pour la SR. La première est basée sur une formulation Bayesienne et l'algorithme *Expectation maximisation* (EM) [Dempster *et al.*, 1977] et la deuxième est basée sur une formulation au sens du *Maximum a postériori* (MAP). Dans ces deux méthodes, le recalage, l'interpolation et la déconvolution sont effectués simultanément.

1.1.1 Discussion

Les principaux avantages des approches que nous venons de voir sont la simplicité de leur formulation et leur faible complexité de calcul. Cependant, il est impossible, avec ces approches, de prendre en compte des mouvements plus complexes que des translations globales entre les images, pas plus que des modèles de RI qui ne sont pas invariants par translation. D'autre part, il est très difficile d'introduire, dans des approches fréquentielles, d'autres termes de régularisation que des termes fréquentiels. Il est impossible, par exemple, d'utiliser une régularisation spatiale par voisinage. En raison de ces limitations, très peu d'intérêt a été porté aux approches fréquentielles et la plupart des méthodes de super-résolution publiées dans la littérature, ces deux dernières décennies, sont plutôt formulées dans le domaine spatial.

1.2 Approches par interpolation puis restauration

Les approches par interpolation puis restauration résolvent le problème de SR en deux étapes. La première consiste en l'interpolation des données des images bassement résolues dans une grille hautement résolue. La seconde étape consiste en la restauration des hautes fréquences de l'image hautement résolue par une technique de déconvolution. Ces deux étapes d'interpolation et de restauration sont précédées d'une étape d'alignement consistant à estimer les transformations géométriques à appliquer à chaque image bassement résolue pour l'aligner sur l'image hautement résolue à reconstruire. L'alignement s'effectue généralement en utilisant une technique de recalage [Brown, 1992, Zitova et Flusser, 2003].

Les approches de SR par interpolation puis restauration diffèrent essentiellement par l'étape d'interpolation. Certains auteurs affirment même que n'importe quelle technique classique de déconvolution d'images peut être appliquée à l'image interpolée pour restaurer les hautes fréquences [Ur et Gross, 1992].

1.2.1 Interpolation

Soit I_{HR} l'image hautement résolue à reconstruire et I_{LR}^{k} la $k^{ième}$ image bassement résolue disponible ($k = 1 \dots K$). L'étape d'interpolation ne permet pas directement d'estimer I_{HR} mais une version floutée de cette image que nous notons $\tilde{I}_{HR}$. La première partie de l'étape d'interpolation consiste à projeter, dans le repère de l'image hautement résolue, les données des images bassement résolues. Dans un second temps, il s'agit d'estimer, à partir de ces projections, les valeurs de $\tilde{I}_{HR}$ en chaque point de la grille hautement résolue. Deux problèmes sont généralement rencontrés à cette étape : (i) les données des images bassement résolues ne sont pas suffisantes pour remplir toute la grille hautement résolue, (ii) ces mêmes données ne sont pas dispersées de manière régulière rendant les techniques d'interpolation classiques basées sur des noyaux [Lehmann *et al.*, 1999] inapplicables pour pouvoir estimer les valeurs des pixels manquants de $\tilde{I}_{HR}$. Ce problème est illustré dans la Figure (1.1).

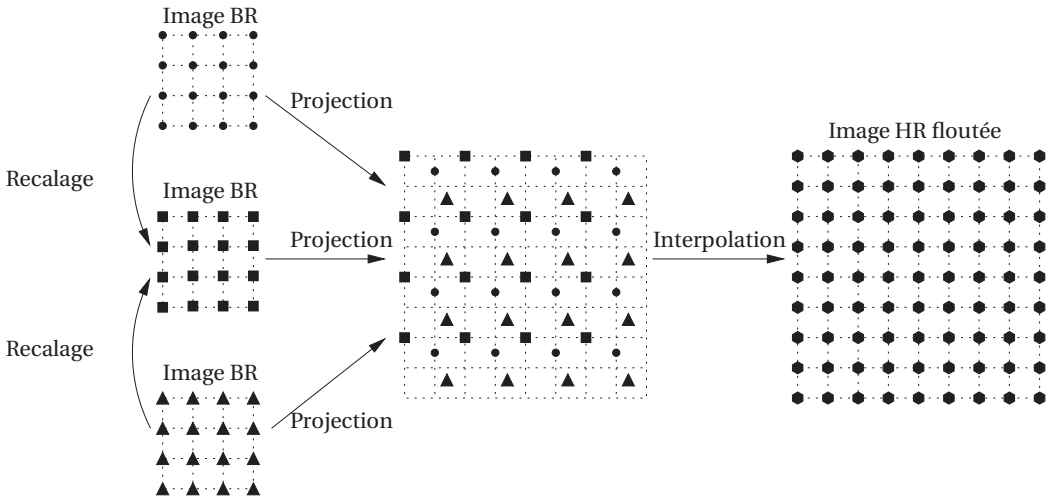


FIGURE 1.1 – Fusion des données bassement résolues dans une grille hautement résolue.

Plusieurs techniques d'interpolation ont été proposées pour résoudre ce problème. Ur et Gross [Ur et Gross, 1992] ont étendu le théorème de l'échantillonnage multi-canal généralisé de Papoulis [Papoulis, 1978, Papoulis, 1977] et Brown [Brown, 1981] pour reconstruire l'image $\tilde{I}_{HR}$. Alam *et al.* [Alam *et al.*, 2000] ont utilisé une méthode des plus proches voisins pondérés pour réaliser cette interpolation, dans le contexte de l'imagerie infrarouge. Dans [Nguyen et Milanfar, 2000], les auteurs ont étendu la technique d'interpolation basée sur les ondelettes présentée dans [Ford et Etter, 1998] au cas des signaux 2D. Elad et Hel-or [Elad et Hel-Or, 2001] ont montré les conditions sur le modèle d'observation pour que les méthodes par interpolation puis restauration soient applicables en SR et ont proposé une méthode d'inversion directe, exploitant les propriétés intéressantes des

matrices du modèle d'observation, pour effectuer une estimation optimale de $\tilde{I}_{HR}$. Dans [Lertrattanapanich et Bose, 2002] les auteurs ont proposé une méthode d'interpolation basée sur la triangulation de Delaunay, mais la méthode s'est avérée ne pas être robuste aux valeurs très bruitées. Pham *et al.* [Pham *et al.*, 2006] ont proposé une méthode d'interpolation robuste utilisant le principe de la convolution normalisée [Knutsson et Westin, 1993] adaptée à la structure locale de l'image. Bose et Ahuja [Bose et Ahuja, 2006] ont utilisé la méthode MLS (*Moving Least Square*) pour estimer l'intensité de chaque pixel de $\tilde{I}_{HR}$ via une approximation polynomiale utilisant les pixels d'un voisinage défini autour du pixel considéré. Takeda *et al.* [Takeda *et al.*, 2007] ont proposé une technique de régression adaptative à noyau orienté (*adaptive steering kernel regression*).

1.2.2 Restauration

L'étape de restauration doit augmenter les hautes fréquences de $\tilde{I}_{HR}$ pour produire l'estimation de I_{HR}. Elle est généralement réalisée via des techniques de déconvolution basées sur des filtres passe-hauts [Ur et Gross, 1992, Papoulis, 1978, Brown, 1981]. Par exemple, Alam *et al.* [Alam *et al.*, 2000] appliquent un filtrage de Wiener. Elad et Hel-or [Elad et Hel-Or, 2001], qui utilisent un filtre de Wiener également, affirment que n'importe quel algorithme classique de rehaussement des hautes fréquence peut être appliqué.

1.2.3 Discussion

Les méthodes basées sur une interpolation puis une restauration sont intuitives, simples et généralement rapides en temps de calcul. Cependant, elles sont implicitement basées sur l'hypothèse que les opérateurs de mouvement et de flou de l'équation (1.1) peuvent être permutés. Cette hypothèse n'est vraie que lorsque toutes les images bassement résolues ont été acquises avec le même imageur et que les mouvements entre elles ne sont que des translations [Elad et Hel-Or, 2001]. Lorsque les images sont acquises par des imageurs différents ou qu'il existe par exemple des rotations entre elles, la technique de déconvolution utilisée ne sera pas adéquate pour tous les pixels. D'autre part, les données bassement résolues étant généralement entachées de bruit, appliquer un filtre passe-haut à l'image interpolée $\tilde{I}_{HR}$ amplifie ce bruit et crée des artefacts. Enfin, ces approches ne garantissent pas l'optimalité si la technique de déconvolution utilisée ne prend pas en compte le modèle du flou de l'imageur. Les approches présentées dans la section suivante sont plus appropriées pour tenir compte de la méconnaissance de la RI.

1.3 Approches par optimisation

Les approches par optimisation essayent de résoudre l'ensemble des Equations (1.1) par des techniques d'optimisation. Nous pouvons reprendre l'Expression (1.1) et réécrire

l'ensemble des equations de la façon condensée suivante :

$$\mathbf{Y} = A\mathbf{X} + \mathbf{B}, \tag{1.8}$$

avec

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}^1 \\ \vdots \\ \mathbf{Y}^K \end{bmatrix}, A = \begin{bmatrix} DH^1W^1 \\ \vdots \\ DH^KW^K \end{bmatrix} \text{ et } \mathbf{B} = \begin{bmatrix} \mathbf{B}^1 \\ \vdots \\ \mathbf{B}^K \end{bmatrix}.$$

Les approches par optimisation sont basées sur la minimisation d'une fonction de coût de la forme :

$$\epsilon(\mathbf{X}) = \epsilon_1(\mathbf{Y}, A\mathbf{X}) + \lambda\epsilon_2(\mathbf{X}), \tag{1.9}$$

où le le premier terme (ϵ_1), appelé terme d'attache aux données, exprime à quel point l'image hautement résolue $\mathbf{X}$ est proche de la séquence $\mathbf{Y}$ via le modèle A. Il correspond généralement à une distance entre $\mathbf{Y}$ et $A\mathbf{X}$. Le terme ϵ_2, connu sous le nom de terme de régularisation, permet de rejeter les solution bruitées et de ne garder que les solutions vérifiant certaines connaissances *a priori* (ou hypothèses) sur la solution recherchée. $\lambda > 0$ est le paramètre de régularisation servant à contrôler le degré de régularisation de la solution. La solution optimale $\widehat{\mathbf{X}}$ est alors donnée par :

$$\widehat{\mathbf{X}} = \arg\min_{\mathbf{X}} \big(\epsilon(\mathbf{X})\big). \tag{1.10}$$

Les auteurs se revendiquent généralement de deux écoles différentes pour la définition de la fonction de coût ϵ : l'école des approches déterministes et celle des approches stochastiques.

1.3.1 Approches déterministes

Avec les approches déterministes, le terme d'attache aux données ϵ_1 est défini en utilisant une distance classique. Keren *et al.* [Keren *et al.*, 1988] ont proposé une méthode récursive pour minimiser la distance L_1 :

$$\epsilon_1(\mathbf{X}) = \|\mathbf{Y} - A\mathbf{X}\|_1, \tag{1.11}$$

connue sous le nom de *Simulation puis Correction*. L'algorithme part d'une solution initiale quelconque, pour chaque pixel de cette solution, si sa valeur est l, elle est remplacée par la valeur qui minimise la fonction de coût parmi l'ensemble des valeurs $\{l-1, l, l+1\}$. Ce procédé est itéré jusqu'à ce qu'un critère d'arrêt soit vérifié. L'algorithme s'est avéré très gourmand en temps de calcul et converge souvent vers un minimum local de la fonction de coût. Pour des raisons théoriques et calculatoires, on préfère généralement minimiser le carré de la distance L_2 :

$$\epsilon_2(\mathbf{X}) = \|\mathbf{Y} - A\mathbf{X}\|_2^2. \tag{1.12}$$

En fait, minimiser ϵ_2 revient à maximiser la vraisemblance de $\mathbf{X}$ par rapport à $\mathbf{Y}$ sous une hypothèse gaussienne. En dérivant cette fonction par rapport à $\mathbf{X}$ et en mettant la dérivée à zero, on obtient le résultat de la pseudo-inverse classique

$$\widehat{\mathbf{X}} = (A^T A)^{-1} A^T \mathbf{Y}.$$

$A^T A$ étant de très grande dimension et singulière de nature, le calcul de sa matrice inverse est généralement infaisable en pratique. Par exemple, si l'image hautement résolue recherchée est de 400×400 pixels, la matrice $A^T A$ est de dimension 160000×160000. Pour cette raison, on préfère généralement faire appel à des méthodes itératives comme la rétro-projection itérative (IBP pour iterative back-projection). Cette dernière a été introduite en SR par Irani et Peleg [Irani et Peleg, 1990, Irani et Peleg, 1991], dont chaque itération est donnée par :

$$\widehat{\mathbf{X}}^{i+1} = \widehat{\mathbf{X}}^i + A^{BP}(\mathbf{Y} - A\widehat{\mathbf{X}}^i), \tag{1.13}$$

où A^{BP}, appelé opérateur de rétro-projection, est un opérateur d'agrégation des données bassement résolue dans la grille hautement résolues. A^{BP} est généralement construit à partir de noyaux définissant le degré d'influence des pixels bassement résolus sur chaque pixel hautement résolu dans la rétro-projection et peut être choisi arbitrairement. Par exemple, on peu utiliser, pour A^{BP}, une des techniques d'interpolation vues dans le paragraphe 1.2 pour que l'algorithme converge. Ce type d'algorithmes est bien connu dans d'autres contextes (technique de reconstruction itérative simultanée en reconstruction tomographique [Gilbert, 1972], ou l'algorithme de Schultz ou Hotteling en déconvolution [Strauss et Rico, 2012], etc). En utilisant cet algorithme, Bannore [Bannore, 2009] a comparé une dizaine de techniques d'interpolation basées sur des noyaux invariants par translation, mais les résultats n'ont pas permis de dégager une technique d'interpolation qui domine les autres, pour tous les critères et dans toutes les conditions. Dans [Dong *et al.*, 2009] les auteurs ont proposé de pré-filtrer l'erreur rétro-projetée $A^{BP}(\mathbf{Y} - A\widehat{\mathbf{X}}^i)$ à chaque itération, avec un filtre utilisant les redondances non-locales, afin de réduire les effets d'oscillations et d'éviter les contours dentelés qui apparaissent généralement dans la solution reconstruite. L'itération de Landweber [Landweber, 1951] est un cas particulier de rétro-projection itérative. Il est équivalent à une descente du gradient de $\epsilon_2(\mathbf{X})$:

$$\widehat{\mathbf{X}}^{i+1} = \widehat{\mathbf{X}}^i + \beta A^T(\mathbf{Y} - A\widehat{\mathbf{X}}^i), \tag{1.14}$$

où A^T est la transposée de la matrice A et β est un paramètre de contrôle de la convergence de l'algorithme. Pour assurer la convergence, β doit être pris tel que $0 < \beta < 2/\|A^T A\|_2$. La méthode de Landweber a été utilisée par Aizawa *et al.* [Aizawa *et al.*, 1992] en SR en utilisant des images acquises par différentes caméras CCD et par Shah et Zakhor [Shah *et al.*, 1999] pour le cas des images couleurs. On peut remarquer l'équivalence entre la rétro-projection itérative de l'Expression (1.13) et l'algorithme de Landweber de l'Expression (1.14) lorsque l'opérateur A^{BP} est choisi comme étant βA^T.

Le problème de SR étant de nature mal posé, il peut s'avérer inapproprié de minimiser une fonction de coût dépourvue de terme de régularisation, notamment en présence d'erreurs dans le modèle d'observation supposé ou lorsque le système est sous-déterminé. En effet, la solution optimale d'une telle fonction peut dévier fortement de l'image hautement résolue recherchée. Le terme de régularisation permet d'écarter les solutions indésirables en introduisant des connaissances *a priori* sur la solution comme le degré de lissage. Il existe différents choix pour ce terme et le plus populaire est certainement le terme de régularisation de Tikhonov [Tikhonov et Arsenin, 1977] donné par

$$\epsilon_2(\mathbf{X}) = \|\Gamma\mathbf{X}\|_2^2,$$

où Γ peut être choisi comme la matrice identité, ou une matrice opérant comme un dérivateur du premier ou du second ordre sur l'image $\mathbf{X}$. La fonction de coût, connue sous le nom des moindres carrés contraints, est alors donnée par :

$$\epsilon_2(\mathbf{X}) = \|\mathbf{Y} - A\mathbf{X}\|_2^2 + \lambda\|\Gamma\mathbf{X}\|_2^2 \tag{1.15}$$

En dérivant cette fonction par rapport à $\mathbf{X}$ on obtient :

$$\frac{d\epsilon_2}{d\mathbf{X}} = 2\left(A^T A\mathbf{X} + \lambda\Gamma^T\Gamma\mathbf{X} - A^T\mathbf{Y}\right) = 2\left(A^T(A\mathbf{X} - \mathbf{Y}) + \lambda\Gamma^T\Gamma\mathbf{X}\right), \tag{1.16}$$

et la descente du gradient nous donne l'itération :

$$\widehat{\mathbf{X}}^{i+1} = \widehat{\mathbf{X}}^i + \beta\left[A^T(\mathbf{Y} - A\widehat{\mathbf{X}}^i) - \lambda\Gamma^T\Gamma\mathbf{X}^i\right], \tag{1.17}$$

où β est, là aussi, le paramètre de convergence.

Zomet et Peleg [Zomet et Peleg, 2000] ont utilisé la régularization de Tikhonov (1.15) pour résoudre le problème de SR combiné à un problème de mosaïque d'images. L'optimisation est réalisée en utilisant l'algorithme du gradient conjugué. Nguyen *et al.* [Nguyen *et al.*, 2001a] ont proposé des préconditionneurs efficaces circulants par block pour la minimisation de la fonction (1.15) avec l'algorithme du gradient conjugué. Ils ont utilisé la méthode de validation croisée généralisée pour calculer automatiquement le paramètre de régularisation λ. Ils ont ensuite étendu cette méthode dans [Nguyen *et al.*, 2001b] pour l'estimation des paramètres de la RI. Bose *et al.* [Bose *et al.*, 2001] ont mis l'accent sur l'importance du rôle du paramètre de régularisation λ en SR et ont proposé de le calculer de manière optimale en utilisant la méthode dite "L-curve" [Hansen et O'Leary, 1993].

La minimisation de la fonction ϵ_1 peut également être régularisée. En restauration et en débruitage d'images, la norme L_1 est utilisée pour définir le critère de la variation totale de l'image (TV pour Total Variation) [Rudin *et al.*, 1992, Li et Santosa, 1996, Chan *et al.*, 2001]. Le critère TV pénalise la quantité totale de la variation dans l'image et est mesuré par la norme L_1 de l'amplitude du gradient :

$$\epsilon_2(\mathbf{X}) = \|\nabla\mathbf{X}\|_1,$$

où ∇ est un opérateur de gradient. En se basant sur l'esprit du critère TV combiné avec la notion de filtre bilatéral [Tomasi et Manduchi, 1998], Farsiu *et al.* [Farsiu *et al.*, 2004b] ont proposé une technique de SR robuste aux erreurs de modélisation du mouvement, basée sur un terme de régularisation baptisé fonction de la variation totale bilatérale (BTV pour Bilateral Total Variation) qui est une généralisation du critère TV. Cette régularisation a comme vertu la préservation de la netteté des contours de l'image hautement résolue reconstruite. La fonction de coût est alors donnée par

$$\epsilon(\mathbf{X}) = \|\mathbf{Y} - A\mathbf{X}\|_1^1 + \lambda \underbrace{\sum_{l=-P}^{P} \sum_{m=0}^{P}}_{l+m\geq 0} \alpha^{|m|+|l|} \|\mathbf{X} - S_x^l S_y^m \mathbf{X}\|_1^1,$$

où les matrices S_x^l et S_y^m translatent $\mathbf{X}$ respectivement de l pixels horizontalement et de m pixels verticalement. Le scalaire α, $0 < \alpha < 1$, sert a donner des poids spatialement décroissants à la somme du terme de régularisation. P est un entier pris généralement entre 1 et 3. L'optimisation est réalisée grâce à un algorithme de descente du gradient.

Récemment, Zhang *et al.* [Zhang *et al.*, 2013] ont proposé une méthode de régularisation pour la SR basée sur une technique de régression à noyaux non-localement orientés. La technique exploite la régularité structurelle locale et la redondance de motifs dans certaines images extérieures comme les images représentant des paysages urbains.

1.3.2 Approches stochastiques

Les approches stochastiques diffèrent légèrement des approches déterministes. Tandis que les approches déterministes essayent de minimiser une fonction de coût déterministe basée signal, les approches stochastiques essayent de maximiser la probabilité a-posteriori $P(\mathbf{X}|\mathbf{Y})$ qui est la vraisemblance d'obtenir $\mathbf{X}$ sachant $\mathbf{Y}$, d'où leur nom d'approches MAP (pour Maximum A-Posteriori). Considérant $\mathbf{X}$ comme une variable stochastique, ce sont les approches les plus adéquates pour tenir compte des connaissances probabilistes *a priori* sur la solution recherchée $\mathbf{X}$ d'une part, et sur le vecteur du bruit $\mathbf{B}$ d'autre part. Les connaissances *a priori* sont modélisées sous la forme d'une fonction de densité de probabilité sur l'image haute résolution recherchée. La solution est donnée par :

$$\widehat{\mathbf{X}}_{MAP} = \arg\max_{\mathbf{X}} \big(P(\mathbf{X}|\mathbf{Y})\big). \tag{1.18}$$

En appliquant le théorème de Bayes on obtient

$$\widehat{\mathbf{X}}_{MAP} = \arg\max_{\mathbf{X}} \left(\frac{P(\mathbf{Y}|\mathbf{X})P(\mathbf{X})}{P(\mathbf{Y})}\right). \tag{1.19}$$

$\widehat{\mathbf{X}}_{MAP}$ étant indépendant de $\mathbf{Y}$, il vient :

$$\widehat{\mathbf{X}}_{MAP} = \arg\max_{\mathbf{X}} \big(P(\mathbf{Y}|\mathbf{X})P(\mathbf{X})\big). \tag{1.20}$$

Puisque le logarithme est une fonction croissante et monotone, on a

$$\widehat{\mathbf{X}}_{MAP} = \arg\max_{\mathbf{X}} \big(\log\big(P(\mathbf{Y}|\mathbf{X})\big) + \log\big(P(\mathbf{X})\big)\big), \tag{1.21}$$

où $\big(\log\big(P(\mathbf{Y}|\mathbf{X})\big)\big)$ est le log de la vraisemblance et $\big(\log\big(P(\mathbf{X})\big)\big)$ est le log de la densité de probabilité *a priori* de $\mathbf{X}$. D'après le modèle (1.8), la fonction de vraisemblance est déterminée par la densité de probabilité du bruit $\mathbf{B}$. Si on suppose que ce bruit suit une loi gaussienne $\mathcal{N}_\sigma$ de moyenne nulle d'un écart-type σ alors

$$\begin{aligned}\log\big(P(\mathbf{Y}|\mathbf{X})\big) &= \log\big(\mathcal{N}_\sigma(\mathbf{Y}-A\mathbf{X})\big)\\ &= \log\left(\frac{1}{\sigma\sqrt{2\pi}}\exp\big(-\frac{1}{2\sigma^2}\|\mathbf{Y}-A\mathbf{X}\|_2^2\big)\right)\\ &= -\frac{1}{2\sigma^2}\|\mathbf{Y}-A\mathbf{X}\|_2^2 + c,\end{aligned} \tag{1.22}$$

où $c = \log\left(\frac{1}{\sigma\sqrt{2\pi}}\right)$ est une constante.

Pour $P(\mathbf{X})$, il est très courant d'utiliser un champ aléatoire de Markov (MRF pour Markov Random Field) décrit par une densité de probabilité de Gibbs :

$$P(\mathbf{X}) = \frac{1}{\mathcal{Z}(\alpha)}\exp\big(-\alpha\mathcal{U}(\mathbf{X})\big), \tag{1.23}$$

où $\mathcal{U}$ est une fonction d'énergie, α est ce qu'on appelle le paramètre de température ou l'hyperparamètre qui peut être choisi arbitrairement et $\mathcal{Z}(\alpha)$ est un facteur de normalisation qui ne dépend que de α. On obtient alors :

$$\log\big(P(\mathbf{X})\big) = \log\left(\frac{1}{\mathcal{Z}(\alpha)}\right) - \alpha\mathcal{U}(\mathbf{X}), \tag{1.24}$$

le terme $\log\left(\frac{1}{\mathcal{Z}(\alpha)}\right)$ peut être considéré comme une constante lorsque la valeur de α est fixée. Les termes constants n'intervenant pas dans l'optimisation, la solution au sens du MAP est alors donnée par :

$$\begin{aligned}\widehat{\mathbf{X}}_{MAP} &= \arg\max_{\mathbf{X}}\big(-\frac{1}{2\sigma^2}\|\mathbf{Y}-A\mathbf{X}\|_2^2 - \alpha\mathcal{U}(\mathbf{X})\big)\\ &= \arg\max_{\mathbf{X}}\big(-\|\mathbf{Y}-A\mathbf{X}\|_2^2 - \lambda\mathcal{U}(\mathbf{X})\big)\\ &= \arg\min_{\mathbf{X}}\big(\|\mathbf{Y}-A\mathbf{X}\|_2^2 + \lambda\mathcal{U}(\mathbf{X})\big),\end{aligned} \tag{1.25}$$

où $\lambda = 2\sigma^2\alpha$, qui dépend uniquement de α, est un paramètre de régularisation. Nous pouvons noter l'analogie entre cette fonction de coût et celle donnée par l'expression (1.15) où le terme $\mathcal{U}(\mathbf{X})$ n'est qu'une notation généralisée du terme de régularisation. C'est pourquoi nous avons pris la précaution de dire que les auteurs se revendiquaient de telle ou telle école car, dans la plupart des cas, les méthodes stochastiques sont équivalentes à des méthodes déterministes.

Beaucoup de méthodes de SR ont adopté le formalisme décrit par l'Expression (1.25) et diffèrent essentiellement dans le modèle d'observation supposé (A) et dans le choix de la fonction d'énergie $\mathcal{U}$. Le choix de cette fonction est crucial : l'enjeu est de trouver une fonction capable de prévenir l'apparition de hautes fréquences dues au bruit dans les régions homogènes de l'image reconstruite tout en préservant les hautes fréquences des contours. Dans [Schultz et Stevenson, 1996, Capel et Zisserman, 1998, Borman et Stevenson, 1999, Pickup *et al.*, 2006, Pickup *et al.*, 2009], les auteurs ont utilisé un champ Markovien de Huber (HMRF pour Huber Markov Random Field) pour la régularisation. Dans [Tian et Ma, 2010], les auteurs ont utilisé un champ Markovien Gaussien (GMRF pour Gaussien Markov Random Field). Ce modèle a été initialement utilisé dans [Hardie *et al.*, 1997] où les auteurs font une estimation conjointe des paramètres des translations et des pixels de l'image hautement résolue, en considérant toutes ces quantités comme des variables stochastiques. Dans [Tipping et Bishop, 2002], les auteurs ont utilisé le même modèle GMRF où les paramètres des mouvements et de la RI sont conjointement estimés via une formulation Bayésienne. D'autres modèles ont également été utilisés en SR comme celui du champ aléatoire conditionnel [Kong *et al.*, 2006] (CRF pour Conditional Random field), du champ Markovien adaptatif aux discontinuités [Suresh et Rajagopalan, 2007] (DAMRF pour Discontinuity Adaptive MRF), du non-stationnaire Gaussien à deux niveaux [Belekos *et al.*, 2010] et d'autres encore [Shen *et al.*, 2007, Humblot et Mohammad-Djafari, 2006, Mohammad-Djafari, 2009, Capel et Zisserman, 2000, Chakrabarti *et al.*, 2007, Kim *et al.*, 2004]. Gunturk *et al.* [Gunturk et Gevrekci, 2006] ont proposé une approche basée MAP utilisant une séquence d'images différemment exposées où le modèle d'observation prend en compte la variation du temps d'exposition et les erreurs de quantification introduites durant le processus d'acquisition.

Il existe également des approches stochastiques pour la SR qui ne considèrent que la fonction de vraisemblance $P(\mathbf{Y}|\mathbf{X})$. Ce sont des techniques résolvant le problème au sens du maximum de vraisemblance (ML pour Maximum Likelihood) qui est un cas particulier de de la formulation MAP où la probabilité $P(\mathbf{X})$ est considérée comme étant uniforme (pas de connaissance *a priori* sur la solution). Dans l'hypothèse où le bruit présent dans les images bassement résolues est Gaussien de moyenne nulle, ces méthodes coïncident avec celles minimisant l'erreur quadratique $\|\mathbf{Y} - A\mathbf{X}\|_2^2$. Dans [Tom *et al.*, 1994, Tom et Katsaggelos, 1995], une formulation ML a été proposée pour la SR utilisant l'algorithme Expectation-Maximisation (EM) [Dempster *et al.*, 1977] avec une estimation conjointe des paramètres des mouvements et de l'image hautement résolue.

Les approches par optimisation formulées dans le domaine spatial ont été les plus adoptées pour la résolution du problème de SR dans la littérature durant ces deux dernières décennies, en raison de leur flexibilité. Ces approches sont en effet très flexibles. Elles permettent d'utiliser différents types de modèles d'observation et d'introduire des connaissances *a priori* sur la solution recherchée afin de régulariser la nature mal posée du problème de SR. De plus, le choix d'une fonction de coût convexe garantit l'unicité de la solution trouvée en utilisant par exemple un algorithme de descente du gradient. D'autre part, les formulations stochastiques permettent de prendre en compte le modèle du bruit présent dans les images et d'utiliser des modèles Markoviens, qui se sont avérés efficaces en restauration d'images, comme modèles *a priori* de la solution.

1.4 Projection sur des ensembles convexes

Les techniques de projection sur des ensembles convexes (POCS pour Projection Onto Convex Sets) essayent de résoudre le problème de SR dans le cadre de la théorie des ensembles. Elles ont été introduites dans [Youla et Webb, 1982] et [Sezan et Stark, 1982] en 1982. Dans [Stark, 1988] et [Stark, 1990], Stark explique comment appliquer la méthode à la restauration des images et dans [Stark et Oskoui, 1989], il collabore avec Oskoui pour l'appliquer au problème de SR.

Le modèle d'observation utilisé est le même que celui de l'Expression (1.1) mais en supposant le modèle exact ($\mathbf{B} = \mathbf{0}$) :

$$\mathbf{Y} = A\mathbf{X}, \tag{1.26}$$

où

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}^1 \\ \vdots \\ \mathbf{Y}^K \end{bmatrix} \text{ et } A = \begin{bmatrix} DH^1W^1 \\ \vdots \\ DH^KW^K \end{bmatrix}.$$

Comme précédemment, il s'agit d'estimer le vecteur $\mathbf{X}$ de M valeurs où $M \in \mathbb{N}$. On considère l'espace vectoriel $\mathbb{R}^M$ comme étant l'espace de toutes les solutions possibles. Chaque propriété connue de l'image recherchée restreint l'espace $\mathbb{R}^M$ à un sous-ensemble $\mathcal{C} \subseteq \mathbb{R}^M$ de solutions vérifiant cette propriété. En connaissant le modèle A, on peut définir pour chaque $\mathbf{Y}_i$, le $i^{ème}$ élément du vecteur $\mathbf{Y}$, un ensemble fermé et convexe $\mathcal{C}_i$ par :

$$\mathcal{C}_i = \{\mathbf{X} \in \mathbb{R}^M \mid A_i\mathbf{X} = \mathbf{Y}_i\}, \tag{1.27}$$

où A_i est la $i^{ème}$ ligne de la matrice A. On définit donc autant d'ensembles convexes que de pixels d'images bassement résolues disponibles. Ces ensembles sont aussi appelés les contraintes de cohérence des données.

Pour chaque ensemble $\mathcal{C}_i$, on définit un opérateur de projection T_i qui projette tout élément de $\mathbb{R}^M$ dans $\mathcal{C}_i$:

$$\forall \mathbf{F} \in \mathbb{R}^M,\ T_i\mathbf{F} \in \mathcal{C}_i.$$

Cet opérateur est donné par

$$T_i\mathbf{F} = \begin{cases} \mathbf{F} & \text{si } A_i\mathbf{F} = \mathbf{Y}_i, \\ \mathbf{F} + \dfrac{\mathbf{Y}_i - A_i\mathbf{F}}{A_i {A_i}^T} {A_i}^T & \text{sinon.} \end{cases}$$

Si le nombre d'ensembles convexes définis est P, l'itération de l'algorithme POCS est donnée par :

$$\widehat{\mathbf{X}}^{j+1} = T_P T_{P-1} \dots T_1 \widehat{\mathbf{X}}^j. \tag{1.28}$$

L'itération est répétée jusqu'à l'obtention d'une solution appartenant à

$$\mathcal{C}_0 = \bigcap_{i=1}^{P} \mathcal{C}_P.$$

La convergence de ce type d'algorithmes vers une telle solution, si elle existe, a été montrée dans [Youla et Webb, 1982]. Il est à noter que si les seules contraintes utilisées sont celles définies par l'Expression (1.27), la technique s'apparente à une méthode de rétro-projection itérative (Expression (1.13)).

La méthode POCS permet également d'introduire des contraintes exprimant des connaissances a-priori sur la solution recherchée. La plus simple est la contrainte de magnitude :

$$\mathcal{C}_m = \{\mathbf{X} \in \mathbb{R}^M \mid \alpha \leq \mathbf{X}_m \leq \beta\},$$

où $(\alpha, \beta) = (0, 255)$ pour les images en niveaux de gris par exemple. Une autre contrainte possible est la contrainte d'énergie :

$$\mathcal{C}_E = \{\mathbf{X} \in \mathbb{R}^M \mid \|\mathbf{X}\|^2 \leq E\},$$

qui permet de régulariser la solution.

Dans [Patti *et al.*, 1994, Patti *et al.*, 1997a, Patti *et al.*, 1997b], Les auteurs ont proposé des extensions de la méthodes POCS pour tenir compte de la variation de la RI par translation, du flou de mouvement et de l'effet d'aliasage. Eren *et al.* [Eren *et al.*, 1997] ont étendu la méthode de [Patti *et al.*, 1997b] et ont proposé une technique utilisant les mouvements des objets dans les séquences vidéos. Elad et Feuer [Elad et Feuer, 1997] ont analysé et comparé les méthodes ML, MAP et POCS et ont proposé une approche hybride. Patti et Altunbasak [Patti et Altunbasak, 2001] ont étendu les techniques précédentes en améliorant la discrétisation du modèle d'observation et ont modifié les contraintes (les ensembles convexes) afin de réduire le phénomène d'oscillation dans l'image reconstruite. Dans [Caner *et al.*, 2003], Caner *et al.* ont utilisé la méthode POCS en SR dans le contexte de la vidéo surveillance où la séquence d'images utilisée représente une scène dynamique et est acquise par des caméras différentes.

Le principal avantage des méthodes POCS réside dans la simplicité d'incorporer différents types de contraintes et de connaissances à priori afin de régulariser la solution.

De plus, la convergence vers une solution vérifiant toutes les contraintes est garantie, si cette dernière existe (si $\mathcal{C}_0 \neq \emptyset$). Cependant, les méthodes POCS ont un coût calculatoire très élevé et convergent très lentement, ce qui limite leur applicabilité dans beaucoup de domaines. D'autre part, la solution n'est pas unique et dépend fortement du choix de la solution initiale.

1.5 Approches par apprentissage

Les approches par apprentissage en SR ont émergé au début des années 2000 [Freeman *et al.*, 2000]. Elles ont été introduites suite au constat des limites des approches précédentes, dites classiques, vis à vis d'une insuffisance de données [Baker et Kanade, 2002]. Ces techniques sont basées sur des mécanismes d'apprentissage. Dans le cas de la SR, une donnée d'apprentissage, appelée aussi "exemple", est un couple de vecteurs $(\mathbf{x}_i^e, \mathbf{y}_i^e)$ où $\mathbf{x}_i^e$ représente un patch d'une image hautement résolue et $\mathbf{y}_i^e$ représente un patch d'une images bassement résolue tel que $\mathbf{y}_i^e = DH\mathbf{x}_i^e + \mathbf{B}_i$. Comme dans l'Expression (1.8), H modélise l'effet de la RI, D est un opérateur de décimation et $\mathbf{B}_i$ représente un bruit. On appelle l'ensemble de ces exemples la base d'apprentissage. Les approches par apprentissage se déroulent généralement en deux phases : la première, appelée phase d'apprentissage, s'effectue hors-ligne. Elle consiste en la création de la base d'apprentissage par génération de la séquence $\mathbf{y}_i^e$ à partir de la séquence $\mathbf{x}_i^e$. La deuxième phase est celle de la reconstruction de l'image hautement résolue. Elle consiste en l'extraction de patchs des images bassement résolues disponibles et en la prédiction de leurs patchs hautement résolus correspondants en utilisant la base d'apprentissage. La taille des patchs doit être proprement choisie. S'ils sont trop petits, la correspondance avec les patchs de la base ne sera pas significative. S'ils sont trop grands, il faudrait une quantité de données considérable pour que la phase d'apprentissage soit valable.

Une façon basique de faire de la reconstruction SR en utilisant la bases d'apprentissage est de trouver, pour chaque patch des images bassement résolues, son plus proche voisin $\mathbf{y}_i^e$ de la base, et de lui assigner le patch hautement résolu correspondant c.à.d. $\mathbf{x}_i^e$. Cependant, cette approche simpliste introduit des artefacts conséquents dans l'image reconstruite [Elad et Datsenko, 2009], dû à la mauvaise correspondance spatiale des patchs entre eux. Dans [Freeman *et al.*, 2000, Freeman *et al.*, 2002] les auteurs ont proposé une méthode utilisant les k plus proches voisins. Le meilleur patch hautement résolu est sélectionné à l'aide de l'algorithme de propagation de croyance [Yedidia *et al.*, 2000] basé sur un formalisme Markovien. Dans [Sun *et al.*, 2003], les auteurs ont étendu la méthode en utilisant les "primal sketch priors" [Marr et Vision, 1982] afin de préserver la netteté des contours. Ils appliquent ensuite l'algorithme de rétro-projection itérative [Irani et Peleg, 1990] comme post-traitement. Dans [Wang *et al.*, 2005], la même idée est reprise en intégrant un modèle statistique pour l'estimation de la RI via un modèle paramétrique. Chang *et al.* [Chang *et al.*, 2004] ont proposé une méthode plus performante basée sur la technique

d'intégration de voisins [Tenenbaum *et al.*, 2000] utilisant une seule image bassement résolue. Pour chaque patch de cette image, représenté par son vecteur $\mathbf{y}_j$, la méthode trouve $\mathcal{N}_j$, l'ensemble de ses k plus proches voisins ($k \in \mathbb{N}$) dans l'ensemble $\{\mathbf{y}_i^e\}$ et calcule les poids de reconstruction avec la technique d'intégration de voisins ("neighbor embedding") :

$$\hat{w}_s = \arg\min_{w_s} \left(\|\mathbf{y}_j - \sum_{\mathbf{y}_s^e \in \mathcal{N}_j} w_s \mathbf{y}_s^e\|_2^2 \right), \quad \text{tel que} \quad \sum_{\mathbf{y}_s^e \in \mathcal{N}_j} w_s = 1. \tag{1.29}$$

Ces poids de reconstruction sont alors utilisés pour générer le patch hautement résolu correspondant :

$$\mathbf{x}_j = \sum_{\mathbf{y}_s^e \in \mathcal{N}_j} \hat{w}_s \mathbf{x}_s^e. \tag{1.30}$$

Pour résoudre le problème de compatibilité entre patchs adjacents, une simple moyenne est appliquée aux zones de recouvrements. L'algorithme fonctionne bien même avec l'utilisation de patchs de taille relativement petite (patchs plus petits que ceux utilisés dans [Freeman *et al.*, 2000, Freeman *et al.*, 2002]). Yang *et al.* [Yang *et al.*, 2008] ont proposé une autre méthode où la base d'exemples est modélisée par deux dictionnaires $D_h = [\mathbf{x}_1^e\ \mathbf{x}_2^e \ldots \mathbf{x}_n^e]$ et $D_l = [\mathbf{y}_1^e\ \mathbf{y}_2^e \ldots \mathbf{y}_n^e]$. Les poids de reconstruction sont représentés par un vecteur calculé par une minimisation de la norme $\mathbf{w}$ L_1 du vecteur [Donoho, 2006] :

$$\hat{\mathbf{w}} = \arg\min_{\mathbf{w}} \left(\|\mathbf{w}\|_1^1 \right), \quad \text{tel que} \quad \|\mathbf{y}_j - D_l \mathbf{w}\|_2^2 \leq \alpha. \tag{1.31}$$

Les patchs hautement résolus sont alors reconstruits par :

$$\mathbf{x}_j = D_h \hat{\mathbf{w}}. \tag{1.32}$$

Cette méthode basée sur une minimisation de la norme L_1 est plus robuste aux données bruitées et aux valeurs aberrantes que toutes les techniques précédentes basées sur des patchs [Milanfar, 2010]. Dans une version ultérieure, Yang *et al.* [Yang *et al.*, 2010] ont étendu la méthode en utilisant une modélisation plus compacte des dictionnaires, rendant l'algorithme plus efficace.

Une critique habituellement formulée à l'égard de ces approches concerne leur caractère local qui va à l'encontre d'une estimation globalement optimale. Dans les travaux pionniers de Baker et Kanade [Baker et Kanade, 2002], une régularisation a été formulée de manière explicite assurant la continuité des dérivées spatiales de l'image hautement résolue globale. Dans [Datsenko et Elad, 2007], les auteurs ont proposé une méthode de régularisation basée sur les exemples. La méthode trouve un ensemble des plus proches voisins dans la base d'apprentissage, et cet ensemble est utilisé pour définir l'expression de l'à priori sur l'image recherchée, basée sur une formulation MAP. Dans [Glasner *et al.*, 2009], les auteurs ont proposé un cadre unifié pour combiner les techniques de SR dites classiques avec celles basées sur des exemples. Leur approche tire profit de la redondance des patchs (ou des motifs) dans les images naturelles.

Les approches par apprentissage connaissent un intérêt croissant dans la littérature depuis leur apparition au début des années 2000. De nombreuses évolutions ont été proposées. Leur performance a été montrée dans le cas de certains domaines d'application comme la reconstruction hautement résolue de caractères alphanumériques ou la reconnaissance faciale [Baker et Kanade, 2002]. De plus, en variant les modèles de RI durant la phase d'apprentissage, ces approches permettent, de manière implicite, de prendre en compte la méconnaissance de la RI de l'imageur dans la reconstruction. Il serait intéressant d'évaluer cette capacité. Cependant, quelques interrogations subsistent concernant l'application de ces méthodes. Par exemple, comment choisir la bonne taille des patchs ? Comment bien choisir la base d'exemples ? En effet, une base d'exemples contenant un type d'images ne sera pas adéquate pour la reconstruction d'images d'un autre type (statistiques différentes). Un autre inconvénient de ces approches est leur temps de calcul souvent très élevé notamment durant la phase de recherche des voisins les plus proches dans la base d'exemples.

1.6 Discussion

1.6.1 Représentation de la réponse impulsionnelle

La Réponse Impulsionnelle (RI) est une fonction mathématique décrivant la réponse d'un système d'imagerie à une source ponctuelle. C'est pourquoi on parle de réponse impulsionnelle. Elle est utilisée dans divers domaines pouvant relever de l'optique (astronomie, microscopie, ophtalmologie), ou d'autres techniques d'imagerie (radiographie, échographie, IRM).

En imagerie 2D, la RI combine généralement l'effet de l'optique de l'appareil de mesure et la façon dont le capteur de luminance réalise la transformation *illumination⟶electricité*. Elle peut être vue comme un filtre anti aliasage (ou anti repliement), mais elle est généralement considéré comme un défaut à cause de l'effet de flou qu'elle induit souvent dans l'image acquise. On modélise alors la RI d'un imageur, notée h, par la convolution de deux fonctions centrées à supports compacts h_o et h_i :

$$\forall (\omega_x, \omega_y) \in \mathbb{R}^2,\ h(\omega_x, \omega_y) = (h_o * h_i)(\omega_x, \omega_y),$$

où h_o est une fonction positive modélisant l'effet de l'optique induit par les lentilles et l'ouverture finie de l'imageur, et h_i est une fonction positive modélisant l'intégration spatiale effectuée par un détecteur CCD. La Figure (1.2) illustre ce phénomène.

En supposant qu'un détecteur photosensible CCD est carré de taille S et qu'il est uniformément sensible à la lumière, la fonction h_i peut raisonnablement être représentée par une fonction boite :

$$h_i(\omega_x, \omega_y) = \begin{cases} \frac{1}{S^2} & \text{si } \omega_x \in [\frac{-S}{2}, \frac{S}{2}] \text{ et } \omega_y \in [\frac{-S}{2}, \frac{S}{2}], \\ 0 & \text{sinon.} \end{cases}$$

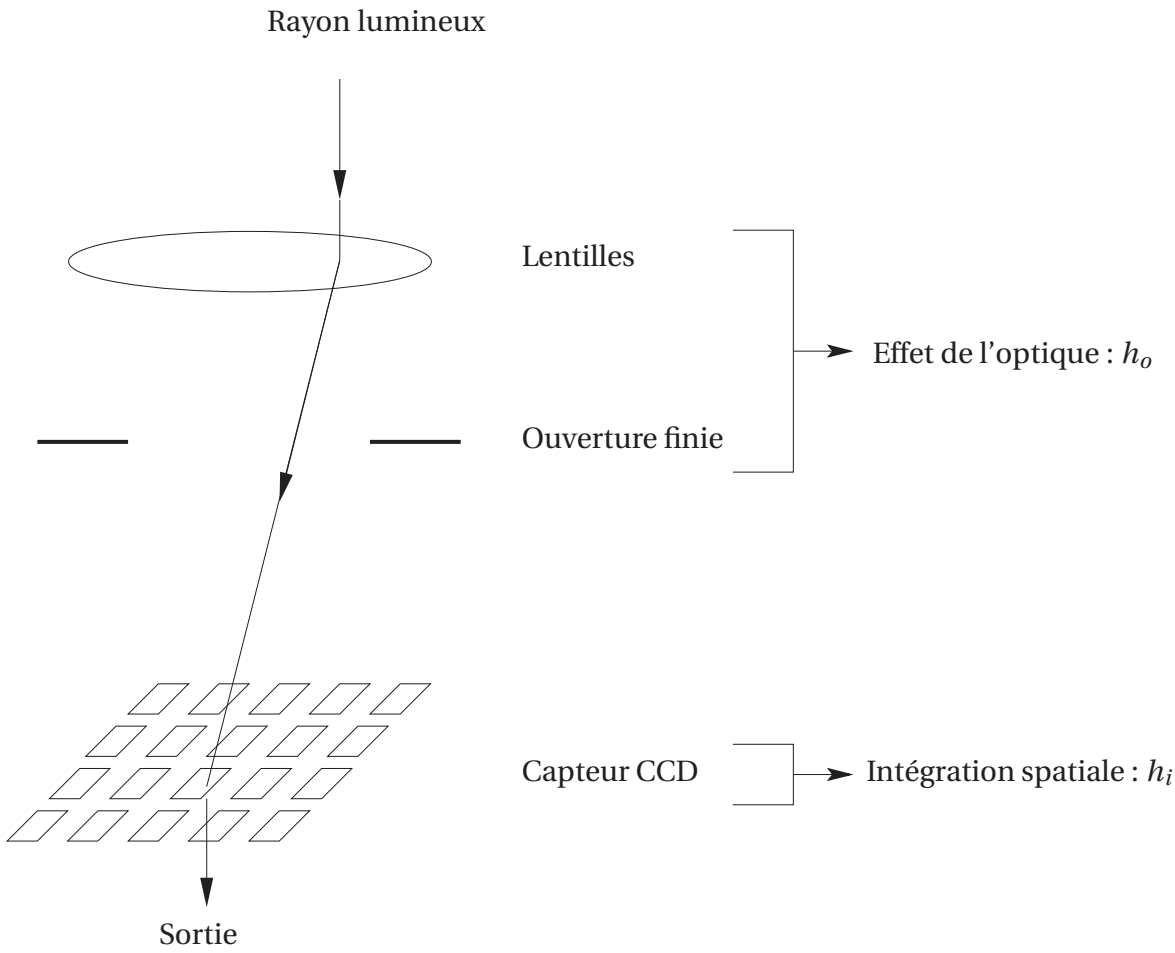


FIGURE 1.2 – Illustration de l'effet de la RI.

Cependant, cette représentation n'est qu'un modèle approximatif et rien ne garantit qu'il soit correct et que les détecteurs photosensibles se comportent tous de la même façon. Il en est de même pour la fonction h_o qui est généralement inconnue et non invariante par translation. Ceci rend la RI h impossible à connaître précisément.

Deux modèles de RI sont généralement utilisés en super-résolution dans la littérature. Certains auteurs modélisent la RI h elle-même par une fonction boite [Capel et Zisserman, 1998, Stark et Oskoui, 1989], et d'autres la modélisent par une gaussienne. Le choix d'une fonction boite peut être expliqué par la prise en compte uniquement de l'effet de l'intégration spatiale par les détecteurs CCD. En effet, la fonction h ne peut être une fonction boite si h_i en est une et h_o est une fonction positive différente du Dirac [Unser *et al.*, 1991]. Nous pensons donc que la fonction boite est un modèle incomplet. La convolution des deux fonctions h_i et h_o donne généralement une fonction unimodale

centrée, à support compact dont le mode est 0 et qui est croissante sur $\{\omega_x \in [-\infty, 0]\}$ et $\{\omega_y \in [-\infty, 0]\}$ et décroissante sur $\{\omega_x \in [0, +\infty]\}$ et $\{\omega_y \in [0, +\infty]\}$. C'est pour cette raison que d'autres auteurs modélisent la RI par une gaussienne. Mais la gaussienne n'est qu'une approximation de ce types de fonctions, et ne peut donc être considérée comme un modèle exact de la RI. Ces auteurs affirment généralement que le paramètre déterminant sa largeur (l'écart type σ) peut être estimé, soit

(i) en utilisant des images de calibration lorsque l'imageur est disponible,

(ii) ou bien à partir d'une ou plusieurs images acquises par un même imageur.

Dans les deux cas, une estimation exacte est très difficile, voire impossible, car le problème d'identification de ce paramètre est lui-même un problème inverse mal posé.

Le problème d'identification de la RI en super-résolution a été brièvement abordé dans [Farsiu *et al.*, 2004a]. Dans [Capel, 2004], l'auteur a mis l'accent sur l'importance de ce problème et a montré les artefacts qui peuvent être induits par l'utilisation d'un mauvais modèle de RI durant la reconstruction SR. Certains auteurs essayent de prévenir ces artefacts en utilisant des techniques classiques de régularisation (voir Section 1.3). Dans l'ouvrage [Milanfar, 2010], Chapitre 9, les auteurs ont montré expérimentalement que ces artefacts peuvent demeurer même en utilisant une bonne technique de régularisation. De plus, les techniques de régularisation tendent à pénaliser les hautes fréquences dans l'image reconstruite, alors que le but de la super-résolution est de les reconstruire. Certaines approches, appelées méthodes aveugles, essayent d'estimer conjointement l'image hautement résolue et la RI via des modèles paramétriques [Nguyen *et al.*, 2001b, Tipping et Bishop, 2002, Wang *et al.*, 2005] où le modèle de la RI est supposé être connu (fonction boite ou gaussienne), ce qui est une hypothèse assez forte, mais pas son étendue, qui doit donc être estimée.

Nous pensons qu'il serait intéressant, qu'au lieu d'utiliser un modèle erroné dont l'impact sur la reconstruction peut être important, d'utiliser un modèle imprécis, exprimant de façon plus réaliste, la connaissance qu'on a sur cette fonction. Le challenge est donc naturellement d'utiliser un tel modèle dans le cadre d'une reconstruction SR. Ce challenge est essentiellement l'objet des travaux de cette thèse.

1.6.2 Discussion sur le modèle d'acquisition

Comme nous venons de le voir, le modèle le plus utilisé pour lier les valeurs des pixels des images bassement résolues avec les valeurs des pixels de l'image hautement résolue recherchée est synthétisé par l'Expression (1.1). En supposant que toutes les images bassement résolues sont acquises par le même imageur, cette expression peut être réécrite par :

$$\mathbf{Y}^k = DHW^k\mathbf{X} + \mathbf{B}^k,$$

où D est une matrice modélisant l'opération de décimation, dont la structure et l'effet sont décrits dans [Mohammad-Djafari, 2009], H est une matrice creuse modélisant l'effet de la

RI de l'imageur, elle est généralement construite à partir des données d'un filtre numérique, noté h_d, qui est une version discrétisée de la RI h de l'imageur, échantillonnée sur la grille de l'image hautement résolue. W^k est la matrice de la transformation géométrique rigide modélisant l'effet du mouvement de la caméra, qui recale l'image hautement résolue recherché sur la version sur-échantillonnée de la $k^{ième}$ image bassement résolue, et $\mathbf{B}^k$ représente le bruit. L'image hautement résolue représentée par $\mathbf{X}$ est généralement considérée comme une version discrète du signal continu d'illumination de la scène, échantillonnée en respectant le critère de Shannon-Nyquist et avec un Dirac (sans introduction de flou).

Les détails concernant la construction des matrices D, H et W^k sont souvent passés sous silence dans la littérature, probablement parce qu'en pratique, elles ne sont généralement pas explicitement construites et servent juste à la modélisation et au développement des expressions (modèle d'observation, dérivée, inversion, descente du gradient etc.). En effet, en raison de leurs très grandes dimensions et du fait qu'elles soient creuses, ces matrices sont remplacées, dans l'implémentation des algorithmes, par des opérations linéaires équivalentes sur les pixels (convolution discrète, sous-échantillonnage etc.).

Cependant, la façon de construire la matrice W^k ou bien l'opération linéaire correspondante a une grande importance. En effet, notons par exemple I l'image hautement résolue représentée par $\mathbf{X}$ et I^k l'image représentée par le vecteur $\mathbf{X}^k = W^k\mathbf{X}$. I^k peut être vue comme une version hautement résolue de la $k^{ième}$ image bassement résolue. Si la transformation géométrique liant I à I^k est une simple translation à valeurs entières dans le référentiel hautement résolu, alors les points d'échantillonnage de l'image I^k correspondent à des points d'échantillonnage de l'image I qui correspondent aussi aux points d'échantillonnage du filtre h_d. Dans ce cas, W^k est une matrice binaire. Par contre, si les valeurs de la translation ne sont pas entières ou bien la transformation géométrique comprend une rotation, alors les points d'échantillonnage de l'image I^k ne correspondent généralement pas à des points d'échantillonnage de l'image I et donc ne correspondent pas aux points d'échantillonnage du filtre h_d. Dans ce cas la matrice W^k n'est pas binaire et l'opération linéaire correspondante doit comprendre une interpolation, la transformation géométrique puis un rééchantillonnage sur la grille de I.

Maintenant, selon la fonction d'interpolation utilisée, l'opération $HW^k\mathbf{X}$ peut avoir une interprétation différente. Notons γ_1 cette fonction d'interpolation, t^k la transformation géométrique liant I à I^k et γ_1^k la version de γ_1 ayant subi la transformation t^k. L'opération $HW^k\mathbf{X}$ est équivalente à l'opération :

$$h_d * \perp(\gamma_1^k) * I, \tag{1.33}$$

où $\perp(f)$ indique que la fonction f est discrétisée. Si, par exemple, γ_1 est une fonction d'interpolation d'ordre 1 (bilinéaire), alors elle peut s'écrire $\gamma_1 = \gamma_0 * \gamma_0$ où γ_0 est la fonction d'interpolation d'ordre 0 (plus proche voisin). En notant γ_0^k la version de γ_0 ayant subi la

transformation t^k, l'Expression (1.33) peut alors s'écrire

$$h_d * \perp(\gamma_0 * \gamma_0^k) * I = h_d * \perp(\gamma_0) * \perp(\gamma_0^k) * I$$

$$= g_d * \perp(\gamma_0^k) * I, \tag{1.34}$$

où $g_d = h_d * \perp(\gamma_0)$ peut être vue comme une version discrétisée d'une RI g plus large que la RI h. L'opération $HW^k\mathbf{X}$ peut donc correspondre au choix de la RI h et d'une interpolation bilinéaire, ou bien d'une RI g plus large que h et d'une interpolation par plus proche voisin. Cette remarque signifie que chacune des fonctions de la RI et d'interpolation doit être choisie judicieusement en fonction de l'autre.

1.7 Conclusion

Nous avons présenté, dans ce chapitre, un état de l'art des méthodes de SR en les classant en cinq grandes catégories. La catégorie qui a suscité le plus de travaux est celle des approches par optimisation, certainement en raison de leur flexibilité qui permet de traiter plusieurs problèmes rencontrées en SR. Par exemple, elles permettent de filtrer le bruit aléatoire, présent dans les images, par des techniques de régression, de tenir compte de la différence d'exposition entre les image bassement résolues en la prenant en compte dans le modèle d'observation, de prendre en compte différents types de mouvements entre les images (translations, rotations, transformations affines etc), ou encore, d'estimer de manière conjointe les valeurs des pixels de l'image hautement résolue recherchée et les paramètres des mouvements ou les paramètres de la RI de l'imageur. Ces approches diffèrent entre elles essentiellement par les techniques de régularisation qu'elles utilisent pour traiter le mauvais conditionnement du problème.

La deuxième catégorie d'approches qui a suscité le plus d'intérêt, notamment durant ces dernières années, est celle des approches par apprentissage apparues assez récemment (début des années 2000). Ces approches sont nées dans le but de surmonter les limites des approches précédentes dites classiques. Ces limites, qui sont essentiellement dues à l'insuffisance de données, ont été mises en évidence dans [Baker et Kanade, 2002]. Certaines de ces approches sont capables de faire de la reconstruction SR en utilisant une seule image bassement résolue comme celle présentée dans [Glasner *et al.*, 2009]. Quelques résultats de cette méthode sont visibles à cette URL [1].

Malgré le très grand nombre de méthodes de SR proposées dans la littérature et les diverses techniques de régularisation utilisées, très peu de travaux se sont focalisés sur le problème induit par une mauvaise identification de la RI de l'imageur. La plupart des méthodes de SR supposent que la RI est précisément connue ou bien qu'elle peut être

1. http ://www.wisdom.weizmann.ac.il/ vision/SingleImageSR.html

précisément estimée en utilisant des techniques d'identification, considérant ainsi ce problème comme un problème pré-SR. En effet, les auteurs supposent généralement que la RI de l'imageur peut être estimée soit en utilisant des images de calibration, lorsque l'imageur est disponible, ou bien à partir de la séquence d'images bassement résolues, lorsque l'imageur n'est pas disponible. Or, dans les deux cas, une estimation exacte et précise de la RI est impossible en pratique car cette estimation constitue, elle aussi, un problème mal posé.

D'autres auteurs considèrent que l'erreur induite par un mauvais modèle de RI peut être prise en compte dans le terme du bruit **B** de l'Expression (1.8), généralement considéré comme étant aléatoire et centré. Or, l'erreur induite par un mauvais modèle de RI n'est pas aléatoire mais systématique, et n'a aucune raison d'être toujours centrée. Les techniques classiques de régression ne prennent donc pas en compte cette erreur.

Pour ce qui est des approches par apprentissage, nous pensons qu'elles permettent implicitement de tenir compte de la méconnaissance de la RI en variant le modèle et les paramètres de la RI permettant d'obtenir les patchs bassement résolus de la base d'exemples. Mais cette capacité n'a pas encore été, à notre connaissance, évaluée quantitativement.

CHAPITRE

2 Notations et définitions préliminaires

Dans cette partie, nous introduisons les concepts utilisés dans la méthode de super-résolution que nous proposons et quelques opérateurs associés. Parmi ces concepts, nous verrons notamment celui des noyaux maxitifs permettant de modéliser la connaissance imprécise de la RI de l'imageur et celui des partitions floues permettant d'obtenir des modèles continus des images discrètes manipulées.

2.1 Notations

Afin de faciliter la compréhension à notre lecteur, et pour éviter toute confusion ou ambiguïté, nous fixons dans cette section quelques notations que nous utiliserons tout au long de ce chapitre. $\mathbb{R}$ désigne l'ensemble des nombres réels et $\mathbb{IR}$ l'ensemble de tous les intervalles de $\mathbb{R}$. Ω est un sous-ensemble convexe de $\mathbb{R}^2$ et $\mathcal{P}(\Omega)$ l'ensemble de tous les sous-ensembles Lebesgue-mesurables de Ω. Soit N un entier positif, nous notons Θ_N l'ensemble $\{1,\ldots,N\} \subset \mathbb{N}$ et $\mathcal{P}(\Theta_N)$ l'ensemble des parties de l'ensemble Θ_N (c.à.d. $\mathcal{P}(\Theta_N) = 2^{\Theta_N}$). Nous notons l_1 l'espace des fonctions $f : \Omega \to \mathbb{R}$ intégrables sur Ω et L_1 l'espace des fonctions discrètes $F : \Theta_N \to \mathbb{R}$ sommables.

2.2 Mesures de confiance et intégrales

Une mesure de confiance est une fonction d'ensembles exprimant le degré de certitude qu'on a qu'un évènement appartienne à un ensemble. Par exemple, la mesure de probabilité est la mesure la plus populaire et la plus utilisée. Une capacité est une fonction d'ensembles pouvant généraliser le concept de probabilité. La notion de capacité a été introduite dans [Choquet, 1953] dans le contexte de sa théorie des capacités. Un concept

similaire a été proposé par Sugeno [Sugeno, 1974] sous le nom de mesure floue et par Denneberg [Denneberg, 1994] sous le nom de mesure non-additive. Une capacité peut être définie dans le domaine discret ou continu.

Définition 1. *Une capacité continue ν est une fonction d'ensemble $\nu : \mathcal{P}(\Omega) \rightarrow [0,1]$ telle que*

$$\nu(\varnothing) = 0,\ \nu(\Omega) = 1 \ et\ \forall A, B \in \mathcal{P}(\Omega),\ A \subseteq B \Rightarrow \nu(A) \leq \nu(B),$$

où $\varnothing$ désigne l'ensemble vide dans $\mathbb{R}^2$.

La capacité conjuguée d'une capacité ν, notée ν^c [1], est définie par :

$$\forall A \in \mathcal{P}(\Omega),\ \nu^c(A) = 1 - \nu(A^c),$$

A^c étant le complémentaire de A dans Ω.

Une capacité ν est dite concave si

$$\forall A, B \in \mathcal{P}(\Omega),\ \nu(A \cup B) + \nu(A \cap B) \leq \nu(A) + \nu(B).$$

Elle est dite convexe si

$$\forall A, B \in \mathcal{P}(\Omega),\ \nu(A \cup B) + \nu(A \cap B) \geq \nu(A) + \nu(B).$$

Le cœur d'une capacité concave ν, noté $\mathcal{M}(\nu)$, est l'ensemble non vide de toutes les mesures de probabilités P sur $\mathcal{P}(\Omega)$ tel que

$$\forall A \in \mathcal{P}(\Omega),\ \nu(A) \geq P(A).$$

Remarque 1. *Si ν est une capacité concave, sa capacité conjuguée ν^c est convexe. En raison de cette relation de conjugaison entre ν et ν^c, le cœur de ν peut être réécrit :*

$$\mathcal{M}(\nu) = \{P \text{ probabilité sur } \mathcal{P}(\Omega), \forall A \in \mathcal{P}(\Omega), \nu^c(A) \leq P(A) \leq \nu(A)\}.$$

Remarque 2. *Une capacité concave égale à sa conjuguée est une mesure de probabilité.*

Définition 2. *Soit $f : \Omega \rightarrow \mathbb{R}^+$ une fonction positive dans l_1 bornée et ν une capacité définie sur $\mathcal{P}(\Omega)$. L'intégrale de Choquet de f par rapport à ν est la valeur réelle $\mathbb{C}_\nu(f)$ définie par :*

$$\mathbb{C}_\nu(f) = \int_0^\infty \nu\{x \in \Omega, f(x) \geq \alpha\} d\alpha. \tag{2.1}$$

Cette définition peut être facilement étendue au cas des fonction non positives en considérant l'intégrale de Choquet asymétrique.

1. Nous n'utiliserons pas la notation classique $\bar{\nu}$ pour éviter la confusion avec d'autre notations.

Définition 3. *Soit $f : \Omega \to \mathbb{R}$ une fonction dans l_1 bornée et ν une capacité concave définie sur $\mathcal{P}(\Omega)$. Soit f^+ (resp. f^-) la fonction définie par $\forall x \in \Omega$, $f^+(x) = \max(f(x), 0)$ (resp. $f^-(x) = \max(-f(x), 0)$). L'intégrale de Choquet asymétrique de f par rapport à ν est la valeur réelle $\check{\mathbb{C}}_\nu(f)$ définie par :*

$$\check{\mathbb{C}}_\nu(f) = \mathbb{C}_\nu(f^+) - \mathbb{C}_\nu(f^-). \tag{2.2}$$

Propriété 1. *Pour toutes fonctions f et g dans l_1, $\mathbb{C}_\nu(f+g) \leq \mathbb{C}_\nu(f) + \mathbb{C}_\nu(g)$. De plus, si f et g sont co-monotone sur Ω, alors $\mathbb{C}_\nu(f+g) = \mathbb{C}_\nu(f) + \mathbb{C}_\nu(g)$.*

Toutes ces définitions et propriétés sont valides aussi dans le domaine discret.

Définition 4. *Une capacité discrète ν est une fonction d'ensembles $\nu : \mathcal{P}(\Theta_N) \to [0,1]$ telle que*

$$\nu(\varnothing) = 0,\ \nu(\Theta_N) = 1 \text{ et } \forall A, B \in \mathcal{P}(\Theta_N),\ A \subseteq B \Rightarrow \nu(A) \leq \nu(B),$$

$\varnothing$ étant l'ensemble vide dans $\mathbb{N}$. Les définitions de la capacité conjuguée, de la concavité et du cœur des capacités concaves sont les mêmes que dans le cas continu.

Définition 5. *Soit $F : \Theta_N \to \mathbb{R}^+$ une fonction positive dans L_1 bornée et ν une capacité concave sur $\mathcal{P}(\Theta_N)$. L'intégrale de Choquet de F par rapport à ν est la valeur réelle $\mathbb{C}_\nu(F)$ définie par :*

$$\mathbb{C}_\nu(F) = \sum_{n \in \Theta_N} F_{(n)} \left(\nu(A_{(n)}) - \nu(A_{(n+1)}) \right), \tag{2.3}$$

où $F_{(n)}$ indique que les indices ont été permutés de manière à ce que

$$F_{(1)} \leq \ldots \leq F_{(N)},$$

et

$$A_{(n)} = \{(n), \ldots, (p)\} \text{ et } A_{(N+1)} = \varnothing.$$

Si F est non positive, alors l'intégrale de Choquet asymétrique peut aussi être définie dans le cas discret de la même façon que dans le cas continu. Soit $F : \Theta_N \to \mathbb{R}$ une fonction dans L_1 bornée et ν une capacité sur $\mathcal{P}(\Theta_N)$. L'intégrale de Choquet asymétrique de F par rapport à ν est la valeur réelle $\check{\mathbb{C}}_\nu(F)$ définie par :

$$\check{\mathbb{C}}_\nu(F) = \mathbb{C}_\nu(F^+) - \mathbb{C}_\nu(F^-), \tag{2.4}$$

avec F^+ (resp. F^-) la fonction définie par $\forall k \in \Theta_N$, $F_k^+ = \max(F_k, 0)$ (resp. $F_k^- = \max(-F_k, 0)$).

Les concepts de probabilités, de fonctions de croyance [Shafer *et al.*, 1976] et de possibilités [Zadeh, 1978, Dubois et Prade, 1983] sont des cas particuliers de capacités.

2.3 Noyaux sommatifs et noyaux maxitifs

Les noyaux sommatifs sont très souvent utilisés en traitement du signal pour définir des voisinages pondérés normalisés d'une position réelle. Ils sont utilisés en traitement d'images pour définir des opérateurs d'agrégation pour lisser ou interpoler des fonctions discrètes [Loquin et Strauss, 2008].

Un **noyau sommatif continu** [Loquin et Strauss, 2008] est une fonction continue $\kappa : \Omega \longrightarrow \mathbb{R}^+$ qui vérifie la propriété de sommativité, c.à.d.

$$\int_\Omega \kappa(x)dx = 1.$$

Une telle fonction est formellement équivalente à la fonction de densité d'une distribution de probabilités Lebesgue-measurable P_κ définie par :

$$\forall A \in \mathcal{P}(\Omega),\ P_\kappa(A) = \int_A \kappa(x)dx.$$

Nous notons $\mathcal{K}(\Omega)$ l'ensemble de tous les noyaux sommatifs définis sur Ω.

Un **noyau maxitif continu** [Loquin et Strauss, 2008] est une fonction continue $\pi : \Omega \longrightarrow [0,1]$ vérifiant la propriété de maxitivité, c.à.d.

$$\sup_{x\in\Omega} \pi(x) = 1.$$

Une telle fonction est équivalente à la densité d'une distribution de possibilités, définissant ainsi une mesure de possibilité (Π_π) et une mesure de nécessité (N_π) sur Ω :

$$\Pi_\pi(A) = \sup_{x\in A} \pi(x),$$

et

$$N_\pi(A) = 1 - \sup_{x\notin A} \pi(x).$$

Π_π est une capacité concave tandis que N_π, sa capacité conjuguée, est convexe.

Un noyau maxitif continu définit un sous-ensemble convexe de $\mathcal{K}(\Omega)$ [Loquin et Strauss, 2008]. Soit π un noyau maxitif continu, le sous-ensemble $\mathcal{M}(\pi)$ défini par

$$\mathcal{M}(\pi) = \{\kappa \in \mathcal{K}(\Omega) / \forall A \in \mathcal{P}(\Omega), N_\pi(A) \leq P_\kappa(A) \leq \Pi_\pi(A)\},$$

est appelé le cœur de π. Cette définition coïncide avec la définition donnée dans la Section 2.2.

Ces concepts peuvent facilement être étendu au cas d'une espace discret [Loquin et Strauss, 2008].

Un **noyau sommatif discret** [Loquin et Strauss, 2008] est une fonction discrète $\eta : \Theta_N \longrightarrow [0,1]$ vérifiant la propriété de sommativité, c.à.d. $\sum_{k\in\Theta_N} \eta_k = 1$. Une telle fonction définit une mesure de probabilités P_η par :

$$\forall A \in \mathcal{P}(\Theta_N),\ P_\eta(A) = \sum_{k\in A} \eta_k.$$

L'ensemble de tous les noyaux sommatifs discrets définis sur Θ_N est noté $\mathcal{K}(\Theta_N)$.

Un **noyau maxitif discret** [Loquin et Strauss, 2008] est une fonction discrète $\pi : \Theta_N \longrightarrow [0,1]$ vérifiant la propriété de maxitivité, c.à.d. $\sup_{k\in\Theta_N} \pi_k = 1$. Une telle fonction définit deux mesures de confiance duales sur Θ_N appelées mesure de possibilité (Π_π) et mesure de nécessité (N_π) :

$$\Pi_\pi(A) = \sup_{k\in A} \pi_k,$$

et

$$N_\pi(A) = 1 - \sup_{k\notin A} \pi_k.$$

Comme dans le cas continu, un noyau maxitif discret définit un sous-ensemble convexe de noyaux sommatifs discrets, noté $\mathcal{M}(\pi)$ [Rico et Strauss, 2010].

Le noyau maxitif triangulaire et symétrique a des propriétés très intéressantes. En effet, le cœur d'un noyau maxitive triangulaire symétrique dont le mode est $\tilde{m}$ et dont le support est $[-\Delta, \Delta]$ contient tous les noyaux sommatifs symétriques et monomodaux dont le mode est $\tilde{m}$ et dont support est $[-\delta, \delta]$ avec $\delta \leq \Delta$ [Loquin et Strauss, 2008].

2.4 Espérance précise et espérance imprécise

Soit $f : \Omega \to \mathbb{R}$ une fonction dans l_1 bornée et P une mesure de confiance additive (c.à.d. une probabilité). L'espérance précise de f par rapport à P est la valeur réelle $\mathbb{E}_P(f)$ définie par :

$$\mathbb{E}_P(f) = \int_0^\infty f\, dP. \tag{2.5}$$

Définir P est équivalent à la définition de sa fonction de densité κ :

$$\forall A \in \mathcal{P}(A),\ P(A) = P_\kappa(A) = \int_A \kappa(x)\, dx.$$

Ainsi, l'Expression (2.5) peut être réécrite par :

$$\mathbb{E}_P(f) = \mathbb{E}_{P_\kappa}(f) = \int_\Omega f(x)\kappa(x)\, dx. \tag{2.6}$$

Le concept d'espérance peut facilement être étendu aux capacités concaves (voir e.g. [Rico et Strauss, 2010]). Soit ν une capacité concave et $f : \Omega \to \mathbb{R}$ une fonction dans l_1 bornée. L'espérance imprécise de f par rapport à ν est l'intervalle réel $\overline{\underline{\mathbb{E}}}_\nu(f)$ défini par :

$$\overline{\underline{\mathbb{E}}}_\nu(f) = \left[\underline{\mathbb{E}}_\nu(f), \overline{\mathbb{E}}_\nu(f)\right] = \left[\check{\mathbb{C}}_{\nu^c}(f), \check{\mathbb{C}}_\nu(f)\right]. \tag{2.7}$$

Deux propriétés fondamentales viennent des travaux de Denneberg [Denneberg, 1994].

Propriété 2. *Si $f : \Omega \to \mathbb{R}$ est une fonction dans l_1 bornée et ν est une capacité concave définie sur Ω, alors*

$$\forall P \in \mathcal{M}(\nu),\ \mathbb{E}_P(f) \in \underline{\mathbb{E}}_\nu(f),$$

et

$$\forall y \in \underline{\mathbb{E}}_\nu(f),\ \exists P \in \mathcal{M}(\nu) \text{ telle que } y = \mathbb{E}_P(f).$$

Propriété 3. *Si $f, g : \Omega \to \mathbb{R}$ sont deux fonctions dans l_1 bornées et ν est une capacité concave définie sur Ω, alors*

$$\overline{\mathbb{E}}_\nu(f+g) \leq \overline{\mathbb{E}}_\nu(f) + \overline{\mathbb{E}}_\nu(g),$$

et

$$\underline{\mathbb{E}}_\nu(f+g) \geq \underline{\mathbb{E}}_\nu(f) + \underline{\mathbb{E}}_\nu(g).$$

Les espérances précise et imprécise coïncident lorsqu'on considère une mesure de probabilité, c.à.d. si P est une mesure de probabilité sur $\mathcal{P}(\Omega)$, alors

$$\underline{\mathbb{E}}_P(f) = \mathbb{E}_P(f).$$

Les espérances précise et imprécise peuvent également être définies dans le domaine discret. Soit $F : \Theta_N \to \mathbb{R}$ une fonction bornée dans L_1. Soit P_η une mesure de probabilité discrète définie sur $\mathcal{P}(\Theta_N)$ générée par le noyau sommatif discret η. L'espérance précise de F par rapport à P_η est la valeur précise $\mathbb{E}_{P_\eta}(F)$ définie par :

$$\mathbb{E}_{P_\eta}(F) = \sum_{k \in \Theta_N} F_k \eta_k. \tag{2.8}$$

Soit ν une capacité concave définie sur $\mathcal{P}(\Theta_N)$. L'espérance imprécise de F par rapport à ν est l'intervalle réel $\underline{\mathbb{E}}_\nu(F)$ défini par :

$$\underline{\mathbb{E}}_\nu(F) = \left[\underline{\mathbb{E}}_\nu(F), \overline{\mathbb{E}}_\nu(F)\right] = \left[\check{\mathbb{C}}_{\nu^c}(F), \check{\mathbb{C}}_\nu(F)\right]. \tag{2.9}$$

Les propriétés 2 et 3 sont vraies dans le domaine discret également [Schmeidler, 1989].

Remarque 3. *Les espérances supérieures et les capacités concaves coïncident lorsqu'on considère la fonction caractéristique. Soit ν une capacité concave,*

$$\forall A \in \mathcal{P}(\Omega),\ \nu(A) = \overline{\mathbb{E}}_\nu(\chi_A),$$

χ_A étant la fonction caractéristique de A.

Enfin, les opérateurs d'espérance peuvent facilement être étendus aux fonctions intervallistes [Dubois, 2006]. Soit $\overline{\underline{f}} : \Omega \to \mathbb{IR}$ une fonction intervalliste, c.à.d. $\forall x \in \Omega$, $\overline{\underline{f}}(x) = [\underline{f}(x), \overline{f}(x)]$. Soit P une mesure de probabilité :

$$\mathbb{E}_P(\overline{\underline{f}}) = [\mathbb{E}_P(\underline{f}), \mathbb{E}_P(\overline{f})] = \{\mathbb{E}_P(g) / g \in \overline{\underline{f}}\}.$$

Cette extension s'applique aussi à l'opérateur d'espérance imprécise. Soit ν une capacité concave :

$$\overline{\underline{\mathbb{E}}}_\nu : \overline{\underline{\mathbb{E}}}_\nu(\overline{\underline{f}}) = [\underline{\mathbb{E}}_\nu(\underline{f}), \overline{\mathbb{E}}_\nu(\overline{f})] = \{\mathbb{E}_P(g) / g \in \overline{\underline{f}} \text{ and } P \in \mathcal{M}(\nu)\}.$$

Ces extensions s'appliquent aussi aux fonctions discrètes (voir [Strauss et Rico, 2012]).

2.5 Les partitions floues

2.5.1 Ensembles flous

Les ensembles flous sont une généralisation du concept classique des ensembles. Ils ont été introduit par Zadeh [Zadeh, 1965] afin de représenter mathématiquement le fait que le passage d'une classe d'objets à une autre est graduel. Un ensemble flou A de Ω est défini par sa fonction d'appartenance $\mu_A : \Omega \to [0,1]$. Soient A et B deux sous-ensembles flous de Ω définis respectivement par leurs fonctions d'appartenance μ_A et μ_B. La fonction d'appartenance de l'intersection de deux ensembles flous A et B est généralement définie par une T-norme $\top$ et la fonction d'appartenance de leur intersection est définie par la T-conorme de $\top$ notée $\bot$, c.à.d. :

$$\forall \omega \in \Omega,\ \mu_{A\cap B}(\omega) = \top(\mu_A(\omega), \mu_B(\omega)),$$

et

$$\mu_{A\cup B}(\omega) = \bot(\mu_A(\omega), \mu_B(\omega)) = 1 - \top(1-\mu_A(\omega), 1-\mu_B(\omega)).$$

Parmi les T-normes qu'on peut considérer, on peut citer :

- La T-norme du minimum : $\top_{\min}(\mu_A, \mu_B) = \min\{\mu_A, \mu_B\}$, dont la T-conorme, notée $\bot_{\max}$, est l'opérateur maximum : $\bot_{\max}(\mu_A, \mu_B) = \max\{\mu_A, \mu_B\}$.
- La T-norme du produit : $\top_{\text{prod}}(\mu_A, \mu_B) = \mu_A \cdot \mu_B$, dont la T-conorme, notée $\bot_{\text{sum}}$, est la somme probabiliste : $\bot_{\text{sum}}(\mu_A, \mu_B) = \mu_A + \mu_B - \mu_A \cdot \mu_B$.
- La T-norme du Łukasiewicz : $\top_{\text{Łuk}}(\mu_A, \mu_B) = \max\{0, \mu_A + \mu_B - 1\}$, dont la T-conorme, notée $\bot_{\text{Łuk}}$, est la somme bornée : $\bot_{\text{Łuk}}(\mu_A, \mu_B) = \min\{\mu_A + \mu_B, 1\}$.
- La T-norme drastique :

$$\top_{\text{drast}}(\mu_A, \mu_B) = \begin{cases} \mu_B & \text{si } \mu_A = 1, \\ \mu_A & \text{si } \mu_B = 1, \\ 0 & \text{sinon.} \end{cases}$$

dont la T-conorme, notée $\perp_{\text{drast}}$, est définie par :

$$\perp_{\text{drast}}(\mu_A, \mu_B) = \begin{cases} \mu_B & \text{si } \mu_A = 0, \\ \mu_A & \text{si } \mu_B = 0, \\ 1 & \text{sinon.} \end{cases}$$

– etc.

2.5.2 Partitions floues

Nous considérons ici Ω comme étant le plan d'une image, c.à.d. une boite de $\mathbb{R}^2$. Une partition floue *à la Ruspini* [Ruspini, 1973] de Ω est un ensemble de N sous-ensembles flous $\{C_n\}_{n\in\Theta_N}$ tels que :

(i) $\forall \omega \in \Omega,\ \sum_{n=1}^{N} \mu_{C_n}(\omega) = 1$,

(ii) μ_{C_n} est continue,

(iii) $\forall n \in \Theta_N,\ \exists \omega \in \Omega$ tel que $\mu_{C_n}(\omega) = 1$.

Habituellement, les sous-ensembles C_n sont monomodaux, symétriques et régulièrement espacés, générés par une fonction générique simple E. Soit $\{\omega_n\}_{n\in\Theta_N}$ un ensemble de N positions de Ω régulièrement espacées. La génération des C_n avec E est obtenue par

$$\forall \omega \in \Omega,\ \mu_{C_n}(\omega) = \mu_E(\omega_n - \omega).$$

Toutes les partitions floues utilisées dans ce manuscrit sont des partitions *à la Ruspini*.

Les partitions floues sont très utiles pour réaliser des interpolations [Perfilieva, 2006]. Soit $\{F_n\}_{n\in\Theta_N}$ une fonctions réelles discrète. L'interpolation de F conduit à la définition d'une fonction continue $\hat{F} : \Omega \to \mathbb{R}$ telle que

$$\forall \omega \in \Omega,\ \hat{F}(\omega) = \sum_{n\in\Theta_N} F_n \mu_{C_n}(\omega).$$

Définition 6. *Soit $A \subseteq \Theta_N$. On définit Υ_A comme étant la fonction d'appartenance de $\bigcup_{n\in A} C_n$, où l'union est définie par la T-conorme de Łukasiewicz :*

$$\forall \omega \in \Omega,\ \Upsilon_A(\omega) = \min(1, \sum_{n\in A} \mu_{C_n}(\omega)).$$

En raison de la Propriété (i) de la partition floue *à la Ruspini*, on a :

$$\Upsilon_A(\omega) = \sum_{n\in A} \mu_{C_n}(\omega).$$

Cet opérateur possède une propriété importante :

$$\forall A, B \in \mathcal{P}(\Theta_N),\ \forall \omega \in \Omega,\ \Upsilon_{A\cup B}(\omega) + \Upsilon_{A\cap B}(\omega) = \Upsilon_A(\omega) + \Upsilon_B(\omega). \tag{2.10}$$

L'Equation (2.10) découle directement de l'équation générale suivante qui est vrai pour toute séquence de nombres $(u_n)_{n=1,\ldots,N}$:

$$\sum_{n\in A\cup B} u_n = \sum_{n\in A} u_n + \sum_{n\in B} u_n - \sum_{n\in A\cap B} u_n.$$

2.6 Les intervalles réels généralisés

Les intervalles réels sont souvent utilisés pour exprimer l'imprécision sur la connaissance de la valeur d'un paramètre, d'une variable etc. Un intervalle réel généralisé est noté $[a] = [\underline{a}, \overline{a}]$ avec $\underline{a}, \overline{a} \in \mathbb{R}$. Il existe deux types d'intervalles : si $\underline{a} \leq \overline{a}$, alors l'intervalle est dit propre, et cette définition correspond aux intervalles réels dit ordinaires (ou classiques). Si $\underline{a} \geq \overline{a}$, alors l'intervalle est dit impropre.

Soit $[a]$ et $[b]$ deux intervalles réels généralisés.

Définition 7. *L'addition de Minkowski* $\oplus$ *et la soustraction de Minkowski* $\ominus$ *de* $[a]$ *et* $[b]$ *sont respectivement définies par :*

$$[a] \oplus [b] = [\underline{a} + \underline{b}, \overline{a} + \overline{b}],$$

et

$$[a] \ominus [b] = [\underline{a} - \overline{b}, \overline{a} - \underline{b}].$$

L'addition de Minkowski des deux intervalles $[a]$ et $[b]$ est l'intervalle de toutes les valeurs pouvant être obtenues en additionnant un élément de $[a]$ avec un élément de $[b]$. De même, la soustraction de Minkowski de $[a]$ et $[b]$ est l'intervalle de toutes les valeurs pouvant être obtenues en soustrayant un élément de $[b]$ d'un élément de $[a]$.

Définition 8. *L'addition duale de Minkowski* $\boxplus$*(ou de Hukuhara [Hukuhara, 1967]) de* $[a]$ *et* $[b]$ *est définie comme étant la solution de la soustraction de Minkowski* $[x] \ominus [b] = [a]$*. Elle est donnée par :*

$$[a] \boxplus [b] = [\underline{a} + \overline{b}, \overline{a} + \underline{b}].$$

La soustraction duale de Minkowski $\boxminus$ *(ou de Hukuhara) de* $[a]$ *et* $[b]$ *est définie comme étant la solution de l'addition de Minkowski* $[x] \oplus [b] = [a]$*. Elle est donnée par :*

$$[a] \boxminus [b] = [\underline{a} - \underline{b}, \overline{a} - \overline{b}].$$

Il est à noter que si $[a]$ et $[b]$ sont des intervalles propres, alors $[a] \oplus [b]$ donne un intervalle propre également, alors que $[a] \boxplus [b]$ ne donne pas toujours un intervalle propre.

Les equations $[x] \oplus [b] = [a]$ et $[x] \ominus [b] = [a]$ n'ont pas toujours une solutions dans le cadre des intervalles réels ordinaires, alors qu'elles ont toujours une solution dans le cadre des intervalles réels généralisés.

Toute opération de Minkowski ou opération duale de Minkowski peut être facilement étendue aux vecteurs intervallistes. Soit $\mathbb{IR}^p$ l'ensemble des vecteurs intervallistes de dimension p. Soit $[A], [B] \in \mathbb{IR}^p$ deux vecteurs intervallistes tels que $[A] = ([a_1], \ldots, [a_p])$ et $[B] = ([b_1], \ldots, [b_p])$, on a alors

$$[A] \bowtie [B] = ([a_1] \bowtie [b_1], \ldots, [a_p] \bowtie [b_p]),$$

avec $\bowtie \in \{\oplus, \ominus, \boxplus, \boxminus\}$.

Le cadre des intervalles généralisés est très commode pour la généralisation des opérations arithmétiques (voir e.g. [Gardenes *et al.*, 2001]). Soit $[a]$ et $[b]$ deux intervalles propres (ordinaires) et $[x]$ un intervalle inconnu. La généralisation de Minkowski de l'équation $x - a = b$, dont la solution est $x = b + a$, peut conduire à deux équations intervallistes possibles :

(i) $[x] \ominus [a] = [b]$, si $[a]$ est plus spécifiques que $[b]$,

(ii) $[x] \ominus [b] = [a]$, dans le cas contraire.

Si la solution $[x] = [b] \boxplus [a]$ est un intervalle propre, alors c'est la solution propre de l'équation $[x] \ominus [a] = [b]$. Si $[x] = [b] \boxplus [a]$ est un intervalle impropre, alors $[\overline{x}, \underline{x}]$ est la solution propre de l'équation $[x] \ominus [b] = [a]$.

CHAPITRE

3 Traitement d'un modèle imprécis de la RI en Super-Résolution

Dans ce chapitre, nous présentons en détails la démarche que nous avons adoptée pour développer une méthode de super-résolution utilisant une connaissance imprécise sur la RI de l'imageur. Ce chapitre est organisé comme suit : dans la Section 3.1 nous introduisons notre modèle de super-résolution. Dans la Section 3.2 nous revisitions le modèle de super-résolution via une approche par passage continu/discret qui nous permet de mieux définir les opérateurs de projection et de rétro-projection (voir Section 1.3.1). Dans la Section 3.3 nous proposons une réinterprétation de l'algorithme de rétro-projection itérative (IBP) basée sur ce modèle de super-résolution. Dans la Section 3.4 nous introduisons le concept de fonctions de pondération de voisinage que nous proposons de construire à partir de partitions floues du plan image. Nous montrons que ces fonctions permettent de modéliser une connaissance imprécise de la RI d'un imageur. Dans la Section 3.5 nous proposons des opérateurs de projection et de rétro-projection additifs précis pour le cas où la RI de l'imageur est précisément connue. Dans la Section 3.6 nous proposons une extension non-additive imprécise de ces deux opérateurs pour le cas où la RI de l'imageur est connue de manière imprécise. Enfin, dans la Section 3.7 nous présentons une adaptation de l'algorithme de rétro-projection itérative aux deux opérateurs imprécis de projection et de rétro-projection, c.à.d. au modèle imprécis de RI que nous avons proposé.

3.1 Notre modèle de Super-Résolution

Dans la plupart des articles de la littérature, la réalisation d'une image super-résolue s'appuie sur le modèle d'une relation directe entre l'image hautement résolue recherchée et les images bassement résolues. Dans ce modèle, les images bassement résolues sont supposées être obtenues en mesurant l'image hautement résolue recherchée dont elles seraient des versions floutées par la RI de l'imageur, sous-échantillonnées et dont la superpo-

sition pourrait être obtenue par une transformation rigide. Ce modèle n'est pas correct. Le rééchantillonnage de l'image hautement résolue ne peut aboutir à une image bassement résolue ne serait-ce que parce que leurs points d'échantillonnage ne coïncident généralement pas. D'ailleurs, comme nous l'avons fait remarquer dans la Section 1.6.2, le modèle de l'Expression (1.1) doit utiliser de façon implicite un retour vers une image continue. Nous proposons d'expliciter ce retour vers le continu en créant une modélisation reliant l'image hautement résolue à l'ensemble des images bassement résolues, se basant sur la simple hypothèse que toutes ces images sont des mesures discrètes d'une même image continue z. Une autre façon d'énoncer cette modélisation est la suivante : *l'image hautement résolue recherchée est l'image qui aurait été acquise par un imageur avec une grille de pixels plus hautement résolue et ayant une RI plus spécifique.* La Figure (3.1) illustre ce modèle. Pour mettre en équation ce modèle, nous gardons les notations établies dans le Chapitre 1. Le vecteur des valeurs de l'image hautement résolue à reconstruire est noté $\mathbf{X} = [x_1, x_2, \ldots, x_M]^T$, M étant le nombre de pixels de cette image et les pixels étant rangés suivant un ordre lexicographique. g est la RI de l'imageur hautement résolue. Pour reconstruire cette image hautement résolue, on dispose de K images bassement résolues. Le vecteur des valeurs d'illumination de la $k^{ième}$ image bassement résolue ($k \in \Theta_K$) est noté $\mathbf{Y}^k = [y_1^k, y_2^k, \ldots y_N^k]^T$, N étant le nombre de pixels de l'imageur bassement résolu et h sa RI. Les pixels sont rangés dans un ordre lexicographique. $t^k : \Omega \to \Omega$, $v \mapsto t^k(v)$ est la transformation rigide associée à la $k^{ième}$ image bassement résolue et $\mathfrak{t}^k : \Omega \to \Omega$ est la transformation inverse de t^k (c.à.d. $\forall v \in \Omega, \mathfrak{t}^k(t^k(v)) = t^k(\mathfrak{t}^k(v)) = v$). On suppose bien sur que $M \gg N$ et que g est plus spécifique que h. Enfin, nous notons $z : \Omega \to \mathbb{R}$ la fonction d'illumination continue dont $\mathbf{X}$ et $\{\mathbf{Y}^k\}_{k\in\Theta_K}$ sont des mesures.

Nous modélisons la relation entre l'image continue z et l'image discrète bassement résolue par :

$$\forall (n,k) \in \Theta_N \times \Theta_K,\ y_n^k = \int_{\mathbb{R}^2} z(t^k(v)) h(\omega_n - v)\, dv + b_n^k, \tag{3.1}$$

où $\omega_n \in \Omega$ est la position du $n^{ième}$ pixel de l'imageur et b_n^k représente un bruit de mesure.

La relation entre l'image continue z et l'image discrète hautement résolue est modélisée par :

$$\forall m \in \Theta_M,\ x_m = \int_{\mathbb{R}^2} z(v) g(\varpi_m - v)\, dv + b'_m, \tag{3.2}$$

où ϖ_m est la position du $m^{ième}$ pixel hautement résolu et b'_m représente un bruit de mesure.

Une autre raison qui a motivé le choix de ce modèle est que l'hypothèse, largement adoptée dans la littérature, que l'image hautement résolue recherchée est une version discrétisée de z échantillonnée par un Dirac n'est pas consistante. En effet, il est important de se rappeler que le fait de limiter les hautes fréquences d'un signal échantillonné réduit l'apparition d'effets d'aliasage lors de la manipulation de cette information échantillonnée. Le fait de vouloir estimer une image dont l'étendue en fréquences n'est pas limitée

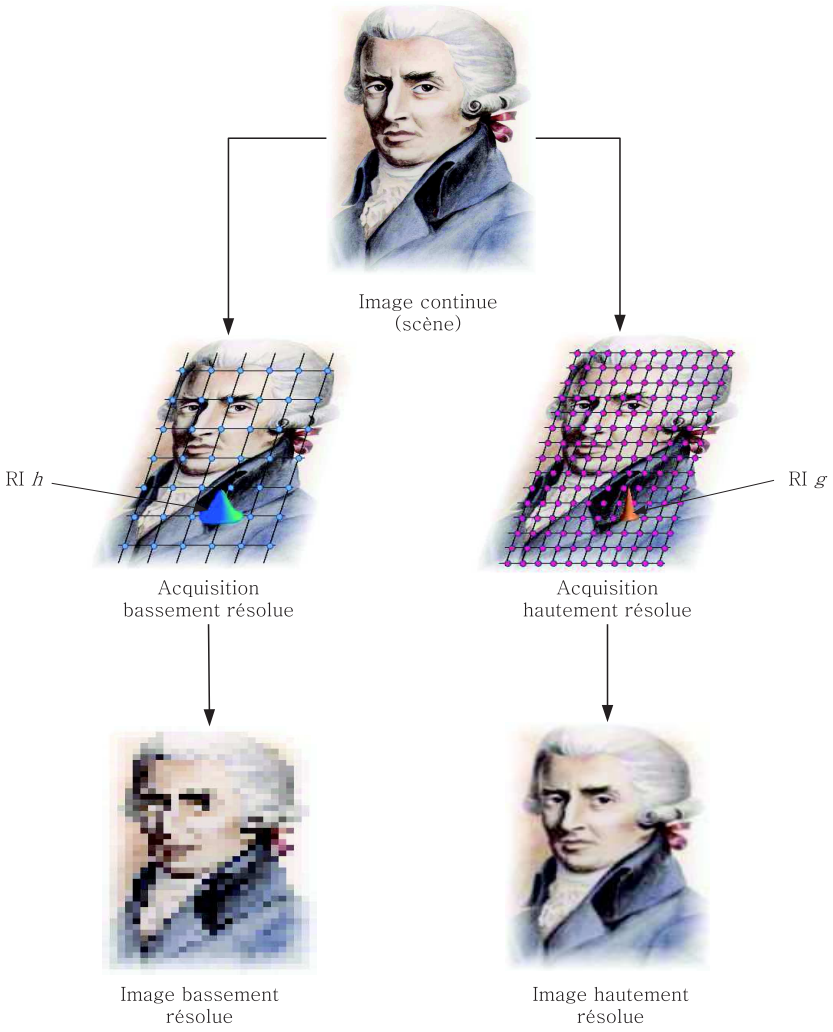


FIGURE 3.1 – Acquisitions bassement résolue et hautement résolue.

est une source de ces effets. Dans [Michaeli et Irani, 2013], les auteurs ont utilisé ce même modèle, ce qui nous conforte dans notre choix.

3.2 Approche par passage continu/discret pour la définition des projection et rétro-projection

La méthode de rétro-projection itérative proposée par Irani et Peleg [Irani et Peleg, 1991] est essentiellement une technique de descente du gradient minimisant l'erreur quadratique $\|\mathbf{Y} - A\mathbf{X}\|_2^2$, où A est l'opérateur de projection défini dans l'Expression (1.8). L'algorithme est basé sur l'utilisation répétée de deux opérateurs appelés projection et rétro-projection notés respectivement A et R et est donné par l'Equation (3.3)

$$\widehat{\mathbf{X}}^{i+1} = \widehat{\mathbf{X}}^{i} + R(\mathbf{Y} - A\widehat{\mathbf{X}}^{i}). \tag{3.3}$$

Nous présentons dans cette section une approche par passage continu/discret pour la définition des opérateurs de projection et de rétro-projection basées sur notre modèle introduit dans la Section 3.1.

3.2.1 Projection

La projection est l'opération permettant d'estimer les vecteurs $\{\mathbf{Y}^k\}_{k\in\Theta_K}$ contenant les valeurs d'illumination des K images bassement résolues à partir du vecteur $\mathbf{X}$ contenant les valeurs d'illumination de l'image hautement résolue. Soit $\hat{z}_g$ l'image continue obtenue en interpolant l'image hautement résolue en utilisant une fonction d'interpolation γ :

$$\forall \omega \in \Omega,\ \ \hat{z}_g(\omega) = \sum_{m\in\Theta_M} \gamma(\omega - \varpi_m) x_m. \tag{3.4}$$

En définissant une fonction γ_m par

$$\forall \omega \in \Omega \text{ et } \forall m \in \Theta_M,\ \gamma_m(\omega) = \gamma(\omega - \varpi_m),$$

l'Expression (3.4) peut être réécrite par :

$$\forall \omega \in \Omega,\ \ \hat{z}_g(\omega) = \sum_{m\in\Theta_M} \gamma_m(\omega) x_m. \tag{3.5}$$

Cette réécriture nous permettra de simplifier le développement de certaines expressions par la suite. L'image $\hat{z}_g$ peut être vue comme une approximation de l'image z convoluée avec la RI g, c.à.d. :

$$\hat{z}_g \simeq z * g. \tag{3.6}$$

L'image $\hat{z}_g$ nous permettra de simuler l'acquisition de la séquence d'images bassement résolues. Notons f la distribution qu'il faut convoluer avec g pour obtenir la RI h, c.à.d. la solution de l'équation :

$$h = g * f. \tag{3.7}$$

f peut donc être définie par :

$$f = g^{-*} * h, \tag{3.8}$$

ou g^{-*} désigne la fonction inverse de g, si elle existe, par rapport à l'opération de convolution $*$. f est une distribution qui n'est pas nécessairement une fonction, mais on peut, sans perte de généralité, supposer dans la suite que c'est une fonction. On utilisera donc, par analogie avec ce qui se fait en traitement du signal classique, la notation abusive mais pratique $f(\omega)$.

En convoluant l'image $\hat{z}_g$ avec f, on obtient une approximation de l'image z convoluée avec la RI h. En effet, $\hat{z}_g * f \simeq z * g * f = z * h$. La simulation de l'acquisition de la séquence d'images bassement résolues est alors donnée en remplaçant z par $\hat{z}_g$ et h par f dans l'Equation (3.1) :

$$\forall (n,k) \in \Theta_N \times \Theta_K,\ \hat{y}_n^k = \int_{\mathbb{R}^2} \hat{z}_g(t^k(v)) f(\omega_n - v)\, dv + b_n^k. \tag{3.9}$$

En injectant l'Equation (3.5) dans l'Equation (3.9) et en négligeant le terme du bruit on obtient :

$$\forall (n,k) \in \Theta_N \times \Theta_K,\ \hat{y}_n^k = \int_{\mathbb{R}^2} \left(\sum_{m \in \Theta_M} \gamma_m(t^k(v)) x_m \right) f(\omega_n - v)\, dv \tag{3.10}$$

$$= \sum_{m \in \Theta_M} x_m \int_{\mathbb{R}^2} \gamma_m(t^k(v)) f(\omega_n - v)\, dv. \tag{3.11}$$

En définissant la fonction γ_m^k, par

$$\forall k \in \Theta_K \text{ et } \forall v \in \Omega,\ \gamma_m^k(v) = \gamma_m(t^k(v)),$$

l'Expression (3.11) devient

$$\forall (n,k) \in \Theta_N \times \Theta_K,\ \hat{y}_n^k = \sum_{m \in \Theta_M} x_m \int_{\mathbb{R}^2} \gamma_m^k(v) f(\omega_n - v)\, dv \tag{3.12}$$

$$= \sum_{m \in \Theta_M} x_m (\gamma_m^k * f)(\omega_n) \tag{3.13}$$

$$= \sum_{m \in \Theta_M} \eta_{m,n}^k x_m, \tag{3.14}$$

où

$$\forall (n,m,k) \in \Theta_N \times \Theta_M \times \Theta_K,\ \eta_{m,n}^k = (\gamma_m^k * f)(\omega_n). \tag{3.15}$$

η_m^k est un noyau de reconstruction discret obtenu en échantillonnant les fonctions $(\gamma_m^k * f)_{((m,k) \in \Theta_M \times \Theta_K)}$ sur les grilles bassement résolues. La projection peut être exprimée sous la forme matricielle :

$$\mathbf{Y}^k = A^k \mathbf{X},$$

comme dans l'Expression (1.2) où

$$A^k = \begin{bmatrix} \eta^k_{1,1} & \eta^k_{2,1} & \cdots & \eta^k_{M,1} \\ \eta^k_{1,2} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \eta^k_{1,N} & \cdots & \cdots & \eta^k_{M,N} \end{bmatrix},$$

lorsque le bruit est ignoré. On peut aussi écrire la projection sous la forme

$$\mathbf{Y} = A\mathbf{X},$$

où

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}^1 \\ \mathbf{Y}^2 \\ \vdots \\ \mathbf{Y}^K \end{bmatrix} \text{ et } A = \begin{bmatrix} A^1 \\ A^2 \\ \vdots \\ A^K \end{bmatrix}.$$

3.2.2 Rétro-projection

La rétro-projection est une opération duale de la projection. C'est une opération qui consiste en la projection et l'agrégation des données des images bassement résolues dans la grille hautement résolue. Lorsqu'elle est appliquée aux K vecteurs $\mathbf{Y}^k$, on obtient un vecteur $\hat{\mathbf{X}}^h = \left[\hat{x}^h_1, \hat{x}^h_2, \ldots \hat{x}^h_M\right]^T$ contenant les valeurs d'illumination d'une image hautement résolue en nombre de pixels mais dont les hautes fréquences on été fortement atténuées. Cette idée est également utilisée par les méthodes dites d'interpolation puis restauration (voir Section 1.2) sans que la fonction de rehaussement des hautes fréquences ne soit correctement spécifiée.

Avec la $k^{ième}$ image bassement résolue ($k \in \Theta_K$), on peut utiliser une fonction d'interpolation générique β et la transformation rigide correspondante $\mathfrak{t}^k$ pour reconstruire une image continue $\hat{z}^k_h$ qui est une approximation très peu informée de l'image z convoluée avec la RI h. Cette interpolation est donnée par l'Expression (3.16) :

$$\forall \omega \in \Omega, \quad \hat{z}^k_h(\omega) = \sum_{n \in \Theta_N} \beta(\mathfrak{t}^k(\omega) - \omega_n)\, y^k_n. \tag{3.16}$$

Naturellement, on peut obtenir une meilleure estimation (car plus informée) de $(z * h)$, en utilisant les K images bassement résolues. C'est justement là qu'on tire profit de la complémentarité de l'information entre les images bassement résolues. Une façon de réaliser cette estimation est d'agréger les K interpolations $\{\hat{z}^k_h\}$. La technique la plus triviale

consiste à utiliser l'opérateur moyenne. On obtient alors une estimation $\hat{z}_h$ par :

$$\forall \omega \in \Omega,\ \ \hat{z}_h(\omega) = \frac{1}{K} \sum_{k \in \Theta_K} \hat{z}_h^k(\omega) \tag{3.17}$$

$$= \frac{1}{K} \sum_{k \in \Theta_K} \sum_{n \in \Theta_N} \beta(\mathfrak{T}^k(\omega) - \omega_n) y_n^k. \tag{3.18}$$

En définissant la fonction β_n^k par

$$\forall n, k \in \Theta_N \times \Theta_K \text{ et } \forall \omega \in \Omega,\ \beta_n^k(\omega) = \beta(\mathfrak{T}^k(\omega) - \omega_n),$$

l'Expression (3.18) devient

$$\forall \omega \in \Omega,\ \ \hat{z}_h(\omega) = \frac{1}{K} \sum_{k \in \Theta_K} \sum_{n \in \Theta_N} \beta_n^k(\omega) y_n^k. \tag{3.19}$$

Échantillonner $\hat{z}_h$ sur la grille hautement résolue nous donne alors une estimation d'une image qui aurait été obtenue avec un imageur hautement résolu en nombre de pixels mais dont la RI est h. Cet échantillonnage est donné par :

$$\forall m \in \Theta_M,\ \ \hat{x}_m^h = \hat{z}_h(\varpi_m). \tag{3.20}$$

En injectant l'Expression (3.19) dans l'Equation (3.20) on obtient l'expression de rétro-projection :

$$\forall m \in \Theta_M,\ \ \hat{x}_m^h = \frac{1}{K} \sum_{k \in \Theta_K} \sum_{n \in \Theta_N} \beta_n^k(\varpi_m) y_n^k \tag{3.21}$$

$$= \frac{1}{K} \sum_{k \in \Theta_K} \sum_{n \in \Theta_N} \varphi_{n,m}^k y_n^k, \tag{3.22}$$

où

$$\varphi_{n,m}^k = \beta_n^k(\varpi_m),\ \forall (m,n,k) \in \Theta_M \times \Theta_N \times \Theta_K.$$

φ_n^k est un noyau de reconstruction discret obtenu en échantillonnant les $(\beta_n^k)_{((n,k) \in \Theta_N \times \Theta_K)}$ sur la grille hautement résolue. La rétro-projection peut aussi être exprimée sous la forme matricielle :

$$\hat{\mathbf{X}}^h = R\mathbf{Y}$$

où :

$$R = \frac{1}{K} \left[R^1, R^2, \ldots R^K \right],$$

et

$$\forall k \in \Theta_K,\ R^k = \begin{bmatrix} \varphi_{1,1}^k & \varphi_{2,1}^k & \cdots & \varphi_{N,1}^k \\ \varphi_{1,2}^k & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \varphi_{1,M}^k & \cdots & \cdots & \varphi_{N,M}^k \end{bmatrix}.$$

$\hat{\mathbf{X}}^h$ peut être vu comme un vecteur contenant les données d'une approximation de l'image continue z échantillonnée sur la grille hautement résolue, mais dont la mesure est obtenue par une convolution avec la RI h. Cette image peut également être vue comme une approximation de l'image hautement résolue recherchée convoluée avec une version discrète de f.

3.3 Réinterprétation de l'algorithme de rétro-projection itérative IBP

Comme nous l'avons souligné précédemment, le vecteur $\hat{\mathbf{X}}^h$ obtenu en rétro-projetant les $\{\mathbf{Y}^k\}_{k\in\Theta_k}$ peut être vu comme un vecteur contenant les données d'une approximation de l'image hautement résolue recherchée convoluée avec une version discrétisée de f. L'algorithme de rétro-projection itérative essaye de retirer le flou introduit par cette convolution en recherchant le vecteur $\mathbf{X}$ qui rend $\hat{z}_h$ le plus proche possible de $(f * \hat{z}_g)$ lorsque ces deux images sont échantillonnées sur chaque grille bassement résolue. Ceci est réalisé en minimisant l'erreur quadratique $\|\mathbf{Y} - A\mathbf{X}\|_2^2$ avec la méthode de Schultz basée sur l'utilisation répétée des deux opérateurs de projection et de rétro-projection A et R (Expression (3.3)).

Dans cette section, nous proposons une réinterprétation de cet algorithme. Notre problème de départ était de faire en sorte de trouver $\mathbf{X}$ tel que $\hat{z}_h$ soit le plus proche possible de $(\hat{z}_g * f)$. Ce problème peut être transformé avantageusement en remarquant qu'il est équivalent à trouver $\mathbf{X}$ tel que $(\hat{z}_h * g)$ soit le plus proche possible de $(\hat{z}_g * h)$, en considérant leur échantillonnage sur les K grilles bassement résolues. Soit G la matrice carrée de taille $KN \times KN$ obtenue en échantillonnant la RI g sur la grille hautement résolue. Soit H la matrice de taille $KN \times M$ obtenue en échantillonnant les fonctions $(\gamma_m^k * h)_{(m,k\in\Theta_M\times\Theta_K)}$ sur les grilles bassement résolues. La minimisation du critère $\|\mathbf{Y} - A\mathbf{X}\|_2^2$ peut être remplacée par la minimisation de

$$\|\mathbf{Y}' - H\mathbf{X}\|_2^2,$$

où $\mathbf{Y}' = G\mathbf{Y}$. L'application de la minimisation itérative de Schultz de ce critère donne un algorithme itératif de la forme

$$\widehat{\mathbf{X}}^{i+1} = \widehat{\mathbf{X}}^i + R(\mathbf{Y}' - H\widehat{\mathbf{X}}^i).$$

En remplaçant $\mathbf{Y}'$ par $G\mathbf{Y}$ on obtient

$$\widehat{\mathbf{X}}^{i+1} = \widehat{\mathbf{X}}^i + R(G\mathbf{Y} - H\widehat{\mathbf{X}}^i).$$

Soit G^+ la matrice pseudo-inverse de G. On a donc

$$\widehat{\mathbf{X}}^{i+1} = \widehat{\mathbf{X}}^i + RG(\mathbf{Y} - G^+H\widehat{\mathbf{X}}^i).$$

Réinterprétons ce résultat. R est obtenue en échantillonnant les $(\beta_n^k)_{((n,k)\in\Theta_N\times\Theta_K)}$ sur la grille hautement résolue, et dont les valeurs sont normalisées (multipliées par $\frac{1}{K}$). G étant obtenue en échantillonnant la RI g sur cette même grille, RG peut donc être obtenue par l'échantillonnage des $(\beta_n^k * g)_{((n,k)\in\Theta_N\times\Theta_K)}$ sur la grille hautement résolue, suivi d'une normalisation. En modifiant l'Expression (3.21) pour tenir compte de ce résultat, on obtient un nouvel opérateur de rétro-projection donné par :

$$\forall m \in \Theta_M, \quad \hat{x}_m^{h*g} = \frac{1}{K} \sum_{k\in\Theta_K} \sum_{n\in\Theta_N} (\beta_n^k * g)(\varpi_m) y_n^k, \tag{3.23}$$

où $\hat{x}_m^{h*g}$ est l'échantillon à la position ϖ_m d'une approximation de l'image $(\hat{z}_h * g)$.

G^+ peut être considérée comme étant l'échantillonnage de g^{-*}, définie par l'Equation (3.8), sur les grilles bassement résolues. Donc, G^+H est équivalent à l'échantillonnage des $(g^{-*} * \gamma_m^k * h)_{((m,k)\in\Theta_M\times\Theta_K)} = (\gamma_m^k * f)_{((m,k)\in\Theta_M\times\Theta_K)}$ sur les grilles bassement résolues. L'opérateur de projection reste donc le même que celui défini par l'Equation (3.13).

3.4 Les Fonctions de Pondération de Voisinages (FPV)

En traitement du signal, et plus particulièrement en traitement d'images, on utilise le concept de voisinage pour définir des opérations d'agrégation linéaires et non linéaires. Parmi les opérations non linéaires utilisant la notion de voisinage, on peut citer les érosions, dilatations, ouvertures et fermetures utilisées en morphologie, et le filtrage médian. Les opérations linéaires, quant à elles, associent à cette idée de voisinage une fonction de pondération modélisant ainsi le rôle de chaque pixel du voisinage considéré dans l'opération d'agrégation réalisée. Dans de nombreux cas, ces opérations servent à définir des équivalents numériques de filtres linéaires continus (passe haut, passe bas, dérivation etc.). On peut aussi utiliser des voisinages pondérés en morphologie [Bloch et Maitre, 1994].

Dans cette section, nous proposons d'étendre cette notion de voisinages pondérés en introduisant le concept de Fonctions de Pondération de Voisinages (FPV).

3.4.1 Définition

Nous appelons *fonction de pondération de voisinage* (FPV) une fonction d'ensembles qui associe, à toute position $\omega \in \Omega$, la potentialité d'un sous-ensemble A ($A \in \mathcal{P}(\Omega)$ pour le cas continu et $A \in \mathcal{P}(\Theta_N)$ pour le cas discret) d'être un voisinage de ω. Comme nous le verrons, ces fonctions généralisent les voisinages pondérés utilisés classiquement pour définir des opérations d'agrégation dans les techniques de traitement d'images. Une FPV est formellement équivalente à une capacité de Choquet. Une FPV associée à une position ω sera notée ρ_ω.

3.4.2 FPV continue

En topologie, on dit que $A \in \mathcal{P}(\Omega)$ est un voisinage de $\omega \in \Omega$ si $\omega \in \mathcal{I}(A)$, où $\mathcal{I}(A)$ est l'intérieur du sous-ensemble A. Dans ce cas, $\rho_\omega(A) \in \{0,1\}$. Si on note $\chi_{\mathcal{I}(A)}$ la fonction caractéristique de l'intérieur de A, ce prédicat de voisinage peut s'écrire sous une forme de FPV de la façon suivante :

$$\rho_\omega(A) = \chi_{\mathcal{I}(A)}(\omega). \tag{3.24}$$

Une première extension pondérée de cette notion est obtenue en considérant la position de ω comme régie par un processus aléatoire associé à une fonction de probabilité P_ω. Dans ce cas, la potentialité de A d'être un voisinage de ω est une valeur de l'intervalle $[0,1]$ définie par :

$$\rho_\omega(A) = \int_A dP_\omega = \mathbb{E}_{P_\omega}(\chi_{\mathcal{I}(A)}). \tag{3.25}$$

Dans ce cas, ρ_ω est additive par construction, c.à.d. :

$$\forall A, B \in \mathcal{P}(\Omega),\ \rho_\omega(A \cup B) = \rho_\omega(A) + \rho_\omega(B) - \rho_\omega(A \cap B). \tag{3.26}$$

L'Equation (3.24) est un cas particulier de l'Equation (3.25) où P_ω est une mesure de Dirac centrée sur ω.

Si A est un sous-ensemble flou, la notion d'intérieur ne s'applique plus puisque la frontière est floue. La fonction caractéristique χ_A est alors remplacée par une fonction d'appartenance μ_A appartenant à $[0,1]$ et on peut étendre l'Equation (3.25) au cas où A est un sous-ensemble flou en considérant l'extension de l'espérance proposée par Zadeh [Zadeh, 1968] :

$$\rho_\omega(A) = \int_\Omega \mu_A(\omega) dP_\omega = \mathbb{E}_{P_\omega}(\mu_A). \tag{3.27}$$

Dans ce cas, ρ_ω est additive si la T-norme $\top$ considérée pour définir l'intersection de deux sous-ensembles flous vérifie la propriété suivante :

$$\forall A, B \in \mathcal{P}(\Omega),\ \bot(\mu_A, \mu_B) = \mu_A + \mu_B - \top(\mu_A, \mu_B). \tag{3.28}$$

où $\bot$ désigne la T-conorme de $\top$.

Démonstration. Soient A et B deux sous-ensembles flous de Ω. On a :

$$\begin{aligned}
\rho_\omega(A \cup B) &= \mathbb{E}_{P_\omega}(\mu_{A \cup B}) && (3.29)\\
&= \mathbb{E}_{P_\omega}\left(\bot(\mu_A, \mu_B)\right) && (3.30)\\
&= \mathbb{E}_{P_\omega}\left(\mu_A + \mu_B - \top(\mu_A, \mu_B)\right) && (3.31)\\
&= \mathbb{E}_{P_\omega}\left(\mu_A + \mu_B - \mu_{A \cap B}\right) && (3.32)\\
&= \mathbb{E}_{P_\omega}(\mu_A) + \mathbb{E}_{P_\omega}(\mu_B) - \mathbb{E}_{P_\omega}(\mu_{A \cap B}) && (3.33)\\
&= \rho_\omega(A) + \rho_\omega(B) - \rho_\omega(A \cap B). && (3.34)
\end{aligned}$$

□

Parmi les T-normes utilisées habituellement, plusieurs vérifient la propriété de l'Expression (3.28). C'est le cas pour la T-norme du minimum $\top_{\min}$, celle du produit $\top_{\text{prod}}$ et celle de Łukasiewicz $\top_{\text{Łuk}}$.

Démonstration. Soient A et B deux sous-ensembles flous de Ω.

1. $\perp_{\max}(\mu_A,\mu_B) = \max\{\mu_A,\mu_B\} = \mu_A+\mu_B-\min\{\mu_A,\mu_B\} = \mu_A+\mu_B-\top_{\min}(\mu_A,\mu_B)$.
2. $\perp_{\text{sum}}(\mu_A,\mu_B) = \mu_A+\mu_B-\mu_A\cdot\mu_B = \mu_A+\mu_B-\top_{\text{prod}}(\mu_A,\mu_B)$ (trivial, voir Section 2.5.1).
3. $\perp_{\text{Łuk}}(\mu_A,\mu_B) = \min\{\mu_A+\mu_B,1\} = \mu_A+\mu_B+\min\{0,1-\mu_A-\mu_B\}$
   $= \mu_A+\mu_B+\max\{0,\mu_A+\mu_B-1\} = \mu_A+\mu_B-\top_{\text{Łuk}}(\mu_A,\mu_B)$.

□

Enfin, si la position de ω est régie par un processus aléatoire associé à une capacité concave v_ω, l'espérance classique de l'Equation (3.27) ne peut plus être appliquée en raison de la non-additivité de v_ω. Elle doit donc être remplacée par l'espérance supérieure (voir Equation (2.7)) basée sur l'intégrale de Choquet :

$$\rho_\omega(A) = \overline{\mathbb{E}}_{v_\omega}(\mu_A). \tag{3.35}$$

Dans ce cas, la FPV ρ_ω est concave si la T-norme $\top$ considérée pour définir l'intersection de deux sous-ensembles floues vérifie la propriété de l'Expression (3.28).

Démonstration. Soient A et B deux sous ensembles flous de Ω. On a

$$\begin{aligned}
\rho_\omega(A\cup B)+\rho_\omega(A\cap B) &= \overline{\mathbb{E}}_{v_\omega}(\mu_{A\cup B})+\overline{\mathbb{E}}_{v_\omega}(\mu_{A\cap B}) && (3.36)\\
&= \overline{\mathbb{E}}_{v_\omega}\left(\perp(\mu_A,\mu_B)\right)+\overline{\mathbb{E}}_{v_\omega}(\mu_{A\cap B}) && (3.37)\\
&= \overline{\mathbb{E}}_{v_\omega}\left(\mu_A+\mu_B-\top(\mu_A,\mu_B)\right)+\overline{\mathbb{E}}_{v_\omega}(\mu_{A\cap B}) && (3.38)\\
&= \overline{\mathbb{E}}_{v_\omega}\left(\mu_A+\mu_B-\mu_{A\cap B}\right)+\overline{\mathbb{E}}_{v_\omega}(\mu_{A\cap B}). && (3.39)
\end{aligned}$$

En raison de la concavité de v_ω, on a :

$$\overline{\mathbb{E}}_{v_\omega}\left(\mu_A+\mu_B-\overline{\mathbb{E}}_{v_\omega}(\mu_{A\cap B})\right) \le \overline{\mathbb{E}}_{v_\omega}(\mu_A)+\overline{\mathbb{E}}_{v_\omega}(\mu_B)-\overline{\mathbb{E}}_{v_\omega}(\mu_{A\cap B}),$$

d'où

$$\overline{\mathbb{E}}_{v_\omega}\left(\mu_A+\mu_B-\overline{\mathbb{E}}_{v_\omega}(\mu_{A\cap B})\right)+\overline{\mathbb{E}}_{v_\omega}(\mu_{A\cap B}) \le \overline{\mathbb{E}}_{v_\omega}(\mu_A)+\overline{\mathbb{E}}_{v_\omega}(\mu_B).$$

Donc

$$\rho_\omega(A\cup B)+\rho_\omega(A\cap B) \le \rho_\omega(A)+\rho_\omega(B).$$

□

3.4.3 FPV discrète

La construction des FPVs discrètes se fait sur le même modèle que celui permettant de définir des opérateurs discrets à partir d'opérateurs continus (filtrage, dérivation, transformation géométrique etc.). Ces techniques sont basées sur l'utilisation d'une paire de noyaux, l'un assurant le passage du discret vers le continu (le noyau d'interpolation ou plus généralement de reconstruction) et l'autre modélisant la mesure, c.à.d. assurant le passage du continu au discret. Par exemple, dans l'Expression (3.15) de projection, γ_m^k est le noyau d'interpolation, f est le noyau d'échantillonnage et η_m^k est l'opérateur discret obtenu en convoluant γ_m^k avec f. Ces opérateurs peuvent être additifs [K. Loquin, 2009] ou non-additifs [Jacquey *et al.*, 2007].

Nous proposons ici une construction de FPVs discrètes à partir de FPVs continues et de partitions floues du plan image Ω. Soit N le nombre de pixels de l'image considérée, $\{C_n\}_{n\in\Theta_N}$ les N sous-ensembles flous de la partition floue *à la Ruspini* du plan de l'image Ω tels que chaque C_n soit associé au $n^{ième}$ pixel et $\omega \in \Omega$.

Propriété 4. *Si la position de ω est régie par un processus aléatoire associé à une fonction de probabilité P_ω, la FPV discrète ν_ω définie par :*

$$\forall n \in \Theta_N,\ \ \nu_\omega(\{n\}) = \rho_\omega(C_n) = \mathbb{E}_{P_\omega}(\mu_{C_n}),\ et$$
$$\forall A \in \mathcal{P}(\Theta_N),\ \ \nu_\omega(A) = \mathbb{E}_{P_\omega}(\Upsilon_A),$$

est une FPV additive. Υ_A a été introduit dans la Définition 6, Section 2.5.

Démonstration. Soit $A \in \mathcal{P}(\Theta_N)$.

$$\nu_\omega(A) = \mathbb{E}_{P_\omega}(\Upsilon_A) = \mathbb{E}_{P_\omega}(\sum_{n\in A} \mu_{C_n}) \tag{3.40}$$
$$= \sum_{n\in A} \mathbb{E}_{P_\omega}(\mu_{C_n}) = \sum_{n\in A} \rho_\omega(C_n) \tag{3.41}$$
$$= \sum_{n\in A} \nu_\omega(\{n\}). \tag{3.42}$$

□

Propriété 5. *Si la position de ω est régie par un processus aléatoire associé à une capacité concave v_ω, la FPV ν_ω discrète définie par :*

$$\forall n \in \Theta_N,\ \ \nu_\omega(\{n\}) = \rho_\omega(C_n) = \overline{\mathbb{E}}_{v_\omega}(\mu_{C_n}),\ et$$
$$\forall A \in \mathcal{P}(\Theta_N),\ \ \nu_\omega(A) = \overline{\mathbb{E}}_{v_\omega}(\Upsilon_A),$$

est une FPV concave.

Démonstration. Soit $A, B \in \mathcal{P}(\Theta_N)$. $\Upsilon_{A\cup B}$ et $\Upsilon_{A\cap B}$ sont comonotones dans Ω. En effet, en raison de la partition floue *à la Ruspini*, on a soit :

(i) $A \cap B = \emptyset$, alors $\Upsilon_{A\cap B}(\omega) = 0,\ \forall \omega \in \Omega$ et donc $\Upsilon_{A\cup B}$ et $\Upsilon_{A\cap B}$ sont comonotones.

(ii) ou bien $A \cap B \neq \emptyset$, dans ce cas, pour tout ω dans Ω, on a soit $\Upsilon_{A\cup B}(\omega) = \Upsilon_{A\cap B}(\omega)$ ou bien $\Upsilon_{A\cup B}(\omega) = 1$. Dans les deux cas, $\Upsilon_{A\cup B}$ et $\Upsilon_{A\cap B}$ sont comonotones.

Ainsi, en raison de la Propriété 1 on a

$$\overline{\mathbb{E}}_{v_\omega}(\Upsilon_{A\cup B} + \Upsilon_{A\cap B}) = \overline{\mathbb{E}}_{v_\omega}(\Upsilon_{A\cup B}) + \overline{\mathbb{E}}_{v_\omega}(\Upsilon_{A\cap B}).$$

De l'Expression (2.10), on a aussi

$$\overline{\mathbb{E}}_{v_\omega}(\Upsilon_{A\cup B} + \Upsilon_{A\cap B}) = \overline{\mathbb{E}}_{v_\omega}(\Upsilon_A + \Upsilon_B),$$

donc

$$\overline{\mathbb{E}}_{v_\omega}(\Upsilon_{A\cup B}) + \overline{\mathbb{E}}_{v_\omega}(\Upsilon_{A\cap B}) = \overline{\mathbb{E}}_{v_\omega}(\Upsilon_A + \Upsilon_B).$$

Comme v_ω est concave, on a

$$\overline{\mathbb{E}}_{v_\omega}(\Upsilon_A + \Upsilon_B) \leq \overline{\mathbb{E}}_{v_\omega}(\Upsilon_A) + \overline{\mathbb{E}}_{v_\omega}(\Upsilon_B),$$

d'où

$$\overline{\mathbb{E}}_{v_\omega}(\Upsilon_{A\cup B}) + \overline{\mathbb{E}}_{v_\omega}(\Upsilon_{A\cap B}) \leq \overline{\mathbb{E}}_{v_\omega}(\Upsilon_A) + \overline{\mathbb{E}}_{v_\omega}(\Upsilon_B),$$

Donc

$$v_\omega(A \cup B) + v_\omega(A \cap B) \leq v_\omega(A) + v_\omega(B).$$

□

Les FPVs discrètes que nous avons proposées ici serviront par la suite à définir des opérateurs de projection et de rétro-projection basés sur des partitions floues du plan image.

3.5 Projection et rétro-projection additives basées sur les partitions floues

Nous considérons ici un cas particulier des opérateurs de projection et de rétro-projection (voir Section 3.2) où l'interpolation est effectuée en utilisant des partitions floues. Nous supposons, comme c'est souvent le cas en traitement du signal et des images, que les RI h et g sont des noyaux sommatifs. f s'apparente donc à un noyau sommatif.

Reprenons l'Expression (3.11) de projection et l'Expression (3.23) de rétro-projection. En définissant respectivement deux fonctions f_n^k et g_m^k par :

$$\forall (n,k) \in \Theta_N \times \Theta_K \text{ et } \forall \omega \in \Omega,\ f_n^k(\omega) = f(\omega_n - \imath^k(\omega)), \text{ et}$$

$$\forall (m,k) \in \Theta_M \times \Theta_K \text{ et } \forall \omega \in \Omega,\ g_m^k(\omega) = g(\varpi_m - t^k(\omega)),$$

les Expressions (3.11) et (3.23) peuvent être respectivement réécrites par :

$$\forall (n,k) \in \Theta_N \times \Theta_K,\ \hat{y}_n^k = \sum_{m \in \Theta_M} x_m \int_{\mathbb{R}^2} \gamma_m(\omega) f_n^k(\omega) d\omega, \text{ et} \tag{3.43}$$

$$\forall m \in \Theta_M,\ \ \hat{x}_m^{h*g} = \frac{1}{K} \sum_{k \in \Theta_K} \sum_{n \in \Theta_N} y_n^k \int_{\mathbb{R}^2} \beta_n(\omega) g_m^k(\omega)) d\omega. \tag{3.44}$$

Il suffit pour cela d'effectuer les changements de variables suivants : $\omega = t^k(v)$ pour l'Expression (3.11) et $\omega = \wr^k(v)$ pour l'Expression (3.23), et d'utiliser la propriété que la transformation rigide t^k préserve les distances, c.à.d. $dv = dt^k(v) = d\wr^k(v)$.

Soit $\{B_m\}_{m \in \Theta_M}$ une partition floue hautement résolue de Ω et $\{C_n\}_{n \in \Theta_N}$ une partition floue bassement résolue de Ω (voir Section 2.5). En considérant respectivement ces deux partitions pour réaliser l'interpolation de l'image hautement résolue et les images bassement résolues, les Expressions (3.43) et (3.44) deviennent respectivement :

$$\forall (n,k) \in \Theta_N \times \Theta_K,\ \hat{y}_n^k = \sum_{m \in \Theta_M} x_m \int_{\mathbb{R}^2} \mu_{B_m}(\omega) f_n^k(\omega) d\omega, \text{ et} \tag{3.45}$$

$$\forall m \in \Theta_M,\ \ \hat{x}_m^{h*g} = \frac{1}{K} \sum_{k \in \Theta_K} \sum_{n \in \Theta_N} y_n^k \int_{\mathbb{R}^2} \mu_{C_n}(\omega) g_m^k(\omega)) d\omega. \tag{3.46}$$

Considérons maintenant la FPV discrète P_n^k définie sur Θ_M par :

$$\begin{aligned}
\forall A \subseteq \Theta_M,\ P_n^k(A) &= \sum_{m \in A} \int_{\mathbb{R}^2} \mu_{B_m}(\omega) f_n^k(\omega) d\omega && (3.47)\\
&= \int_{\mathbb{R}^2} \sum_{m \in A} \mu_{B_m}(\omega) f_n^k(\omega) d\omega && (3.48)\\
&= \int_{\mathbb{R}^2} \Upsilon_A(\omega) f_n^k(\omega) d\omega && (3.49)\\
&= \mathbb{E}_{P_{f_n^k}}(\Upsilon_A), && (3.50)
\end{aligned}$$

où $P_{f_n^k}$ est la mesure de probabilité induite par le noyau sommatif f_n^k. P_n^k est une FPV discrète additive (voir Propriété 4), l'Expression (3.45) de projection peut donc être réécrite par :

$$\forall (n,k) \in \Theta_N \times \Theta_K, \hat{y}_n^k = \mathbb{E}_{P_n^k}(\mathbf{X}), \tag{3.51}$$

et la projection complète obtenue en concaténant les NK valeurs projetées dans un seul vecteur $\mathbf{Y}$ sera notée

$$\mathbf{Y} = \mathcal{A}(\mathbf{X}). \tag{3.52}$$

Il est à noter que l'opérateur $\mathcal{A}$ peut être étendu au cas des vecteurs intervallistes (voir Section 2.4) :

$$[\mathbf{Y}] = \mathcal{A}([\mathbf{X}]) = [\mathcal{A}(\underline{\mathbf{X}}), \mathcal{A}(\overline{\mathbf{X}})], \tag{3.53}$$

où $\underline{\mathbf{X}}$ représente le vecteur de la borne inférieure de $[\mathbf{X}]$ et $\overline{\mathbf{X}}$ représente le vecteur de la borne supérieure de $[\mathbf{X}]$.

De même pour la rétro-projection, considérons la FPV discrète Q_m^k définie sur Θ_N par :

$$\forall A \subseteq \Theta_N,\ Q_m^k(A) = \sum_{n\in A} \int_{\mathbb{R}^2} \mu_{C_n}(\omega) g_m^k(\omega) d\omega \tag{3.54}$$

$$= \int_{\mathbb{R}^2} \sum_{n\in A} \mu_{C_n}(\omega) g_m^k(\omega) d\omega \tag{3.55}$$

$$= \int_{\mathbb{R}^2} \Upsilon_A(\omega) g_m^k(\omega) d\omega \tag{3.56}$$

$$= \mathbb{E}_{P_{g_m^k}}(\Upsilon_A), \tag{3.57}$$

où $P_{g_m^k}$ est la mesure de probabilité induite par le noyau sommatif g_m^k. Q_m^k est une FPV discrète additive (voir Propriété 4), l'Expression (3.46) de rétro-projection peut donc être réécrite par :

$$\forall m \in \Theta_M, \hat{x}_m^{h*g} = \frac{1}{K} \sum_{k\in\Theta_K} \mathbb{E}_{Q_m^k}(\mathbf{Y}^k), \tag{3.58}$$

et la rétro-projection complète obtenue en projetant les NK valeurs du vecteur $\mathbf{Y}$ sera notée

$$\mathbf{X} = \mathcal{B}(\mathbf{Y}), \tag{3.59}$$

où $\mathbf{X} = [\hat{x}_1^{h*g}, \hat{x}_2^{h*g}, \ldots \hat{x}_M^{h*g}]^T$. L'opérateur $\mathcal{B}$ peut être étendu au cas des vecteurs intervalliste :

$$[\mathbf{X}] = \mathcal{B}([\mathbf{Y}]) = [\mathcal{B}(\underline{\mathbf{Y}}), \mathcal{B}(\overline{\mathbf{Y}})], \tag{3.60}$$

où $\underline{\mathbf{Y}}$ représente le vecteur de la borne inférieure de $[\mathbf{Y}]$ et $\overline{\mathbf{Y}}$ représente le vecteur de la borne supérieure de $[\mathbf{Y}]$.

Enfin, l'algorithme de rétro-projection itérative basé sur cette construction est donné par :

$$\widehat{\mathbf{X}}^{i+1} = \widehat{\mathbf{X}}^i + \mathcal{B}(\mathbf{Y} - \mathcal{A}(\widehat{\mathbf{X}}^i)). \tag{3.61}$$

3.6 Projection et rétro-projection non-additives basées sur des partitions floues

Comme nous l'avons fait remarquer précédemment, l'approche de rétro-projection itérative et la plupart des approches par optimisation sont basées sur l'hypothèse que f et g sont connues de manière précise. Bien que g puisse être fixée de façon arbitraire, f est généralement inconnue car l'identification de la RI h est très difficile. Dans [Michaeli et Irani, 2013], les auteurs font une estimation non paramétrique de f basée sur

l'expolatation de la redondance de motifs à différentes échelles dans une image bassement résolue. Nous proposons ici d'utiliser la capacité des noyaux maxitifs à représenter un ensemble de noyaux sommatifs pour définir des opérateurs de projection et de rétroprojection utilisant un modèle imprécis de la RI de l'imageur.

Soit κ un noyau sommatif sur lequel la seule connaissance que l'on ait est qu'il est symétrique, uninomodal et séparable, et que sa demi largeur dans chaque direction est bornée par $q \in \mathbb{R}$. Cette connaissance imprécise sur f peut être facilement représentée par un noyau maxitif 2D π. Nous allons montrer cela en utilisant la propriété qu'un noyau maxitif 1D triangulaire et symétrique modélise la famille de tous les noyaux 1D sommatifs symétriques et monomodaux dont le support est inclus dans le support de ce noyau maxitif [Loquin et Strauss, 2008].

Ω étant une boite de $\mathbb{R}^2$, il peut être défini par $\Omega_1, \Omega_2 \subseteq \mathbb{R}$, c.à.d. $\Omega = \Omega_1 \times \Omega 2$. Nous considérons uniquement les ensembles A qui sont des boites de Ω, c.à.d. $A = A_1 \times A_2$ avec $A_1 \subseteq \Omega_1$ et $A_2 \subseteq \Omega_2$.

Propriété 6. *Soit π un noyau maxitif 2D construit en multipliant deux noyaux maxitifs 1D triangulaires π_1 et π_2 définis respectivement sur Ω_1 et Ω_2. Soit κ_1 et κ_2 deux noyaux sommatifs 1D définis respectivement sur Ω_1 et Ω_2, et κ le noyau sommatif 2D défini en multipliant κ_1 et κ_2. Si $\kappa_1 \in \mathcal{M}(\pi_1)$ et $\kappa_2 \in \mathcal{M}(\pi_2)$ alors $\kappa \in \mathcal{M}(\pi)$.*

Démonstration. Soit P_κ la mesure de probabilité définie sur Ω par le noyau κ, et P_{κ_1} et P_{κ_2} les mesures de probabilités définies respectivement sur Ω_1 et Ω_2 par les noyaux κ_1 et κ_2. on a :

$$\forall A \subseteq \Omega,\ P_\kappa(A) = \int_A \kappa(\omega) d\omega \tag{3.62}$$
$$= \int_{A_1} \int_{A_2} \kappa_1(r_1)\kappa_2(r_2) dr_1 dr_2 \tag{3.63}$$
$$= \int_{A_1} \kappa_1(r_1) dr_1 \cdot \int_{A_2} \kappa_2(r_2) dr_2 \tag{3.64}$$
$$= P_{\kappa_1}(A_1) \cdot P_{\kappa_2}(A_2). \tag{3.65}$$

Soit Π_π la mesure de possibilité définie sur Ω par le noyau maxitif π, et Π_{π_1} et Π_{π_2} les deux mesures de possibilités définies respectivement sur Ω_1 et Ω_2 par les noyaux π_1 et π_2. On a

$$\forall A \subseteq \Omega,\ \Pi_\pi(A) = \sup_{\omega \in A}(\pi(\omega)) \tag{3.66}$$
$$= \sup_{r_1 \in A_1, r_2 \in A_2} (\pi_1(r_1)\pi_2(r_2)) \tag{3.67}$$
$$= \sup_{r_1 \in A_1} (\pi_1(r_1)) \cdot \sup_{r_2 \in A_2} (\pi_2(r_2)) \tag{3.68}$$
$$= \Pi_{\pi_1}(A_1) \cdot \Pi_{\pi_2}(A_2). \tag{3.69}$$

Si

$$\kappa_1 \in \mathcal{M}(\pi_1) \text{ et } \kappa_2 \in \mathcal{M}(\pi_2),$$

alors

$$P_{\kappa_1}(A_1) \leq \Pi_{\pi_1}(A_1) \text{ et } P_{\kappa_2}(A_2) \leq \Pi_{\pi_2}(A_2).$$

Donc

$$P_{\kappa_1}(A_1) \cdot P_{\kappa_2}(A_2) \leq \Pi_{\pi_1}(A_1) \cdot \Pi_{\pi_2}(A_2).$$

Ainsi,

$$P_\kappa(A) \leq \Pi_\pi(A). \tag{3.70}$$

En utilisant le même raisonnement, on peut montrer que

$$\forall A \subseteq \Omega,\ P_\kappa(A) \geq N_\pi(A), \tag{3.71}$$

où N_π est la mesure de nécessité associée à π. Il découle des Expressions (3.70) et (3.71) que

$$\kappa \in \mathcal{M}(\pi).$$

□

Supposons que la seule connaissance que nous ayons sur les RI h et g est que ce sont des noyaux sommatifs séparables, symétriques et unimodaux et que la demi largeur de h est bornée par $\Delta \in \mathbb{R}^+$ dans chaque direction et la demi largeur de g est bornée par $\delta \in \mathbb{R}^+$ dans chaque direction, avec $\delta < \Delta$. Cette hypothèse est raisonnable et communément admise dans la communauté du traitement d'images, où Δ et δ sont souvent considérés respectivement comme étant le pas d'échantillonnage bassement résolu et le pas d'échantillonnage hautement résolu. f étant la solution de l'équation $h = g * f$, elle s'apparente à un noyau sommatif unimodal et symétrique, dont la demi largeur est bornée par Δ dans chaque direction. Cette connaissance imprécise peut donc être modélisée par un noyau maxitif π séparable dont la demi largeur est Δ dans chaque direction.

L'Expression (3.50) peut donc être modifiée, pour tenir compte de cette nouvelle modélisation, en :

$$\forall A \subseteq \Theta_M, \forall k \in \Theta_K, \forall n \in \Theta_N, \nu_n^k(A) = \overline{\mathbb{E}}_{\Pi_{\pi_n^k}}(\Upsilon_A), \tag{3.72}$$

où $\Pi_{\pi_n^k}$ est la mesure de possibilité induite par le noyau maxitif π_n^k défini par

$$\forall \omega \in \Omega,\ \pi_n^k(\omega) = \pi(\omega_n - \jmath^k(\omega)).$$

ν_n^k est une FPV discrète concave (non-additive) (voir Propriété 5).

Propriété 7. *Soit f un noyau sommatif et π un noyau maxitif définis sur Ω, et P_n^k et ν_n^k les deux FPV additive et non-additive définies respectivement par les Expressions (3.50) et (3.72). Si $f \in \mathcal{M}(\pi)$ alors $P_n^k \in \mathcal{M}(\nu_n^k)$.*

La Propriété 7 découle directement de la Propriété 2.

La généralisation de l'Expression (3.51) donne un opérateur de projection imprécis :

$$\forall k \in \Theta_K, \forall n \in \Theta_N, [y]_n^k = \overline{\underline{\mathbb{E}}}_{\nu_n^k}(\mathbf{X}). \tag{3.73}$$

Comme dans le cas additif, nous notons $\overline{\underline{\mathcal{A}}}$ l'opérateur de projection conduisant à un vecteur intervalliste $[\mathbf{Y}]$ obtenu en concaténant les NK intervalles $[y]_n^k$:

$$[\mathbf{Y}] = \overline{\underline{\mathcal{A}}}(\mathbf{X}). \tag{3.74}$$

L'opérateur de projection $\overline{\underline{\mathcal{A}}}$ peut être étendu pour traiter un vecteur intervalliste $[\mathbf{X}]$:

$$[\mathbf{Y}] = \overline{\underline{\mathcal{A}}}([\mathbf{X}]). \tag{3.75}$$

La RI g peut également être représentée de manière imprécise en utilisant un noyau maxitif séparable τ d'une demi largeur δ dans chaque direction. L'Expression (3.57) peut alors être modifiée, pour tenir compte de cette modélisation, en :

$$\forall A \subseteq \Theta_N, \forall k \in \Theta_K, \forall m \in \Theta_M, \upsilon_m^k(A) = \overline{\mathbb{E}}_{\Pi_{\tau_m^k}}(\Upsilon_A), \tag{3.76}$$

où $\Pi_{\tau_m^k}$ est la mesure de possibilité induite par le noyau maxitif τ_m^k défini par

$$\forall \omega \in \Omega,\ \tau_m^k(\omega) = \tau(\varpi_m - t^k(\omega)).$$

υ_m^k est une FPV discrète concave (non-additive) (voir Propriété 5).

Propriété 8. *Soit g un noyau sommatif et τ un noyau maxitif définis sur Ω, et Q_m^k et υ_m^k les deux FPV additive et non-additive définies respectivement par les Expressions (3.57) et (3.76). Si $g \in \mathcal{M}(\tau)$ alors $Q_m^k \in \mathcal{M}(\upsilon_m^k)$.*

La Propriété 8 découle directement de la Propriété 2.

La généralisation de l'Expression (3.58) donne un opérateur de rétro-projection imprécis :

$$\forall m \in \Theta_M, [x]_m = \frac{1}{K} \bigoplus_{k \in \Theta_K} \overline{\underline{\mathbb{E}}}_{\upsilon_m^k}(\mathbf{Y}^k), \tag{3.77}$$

où $\bigoplus$ désigne la somme de Minkowski.

Nous notons $\overline{\underline{\mathcal{B}}}$ l'opérateur de rétro-projection conduisant au vecteur intervalliste $[\mathbf{X}]$ obtenu en rétro-projetant les NK valeurs y_n^k concaténées en un seul vecteur $\mathbf{Y}$:

$$[\mathbf{X}] = \overline{\underline{\mathcal{B}}}(\mathbf{Y}). \tag{3.78}$$

L'opérateur $\overline{\underline{\mathcal{B}}}$ peut être étendu au cas où des vecteurs intervallistes $[\mathbf{Y}]$:

$$[\mathbf{X}] = \overline{\underline{\mathcal{B}}}([\mathbf{Y}]). \tag{3.79}$$

Un résultat important découle des Propriétés 7 et 8 est que si $f \in \mathcal{M}(\pi)$ et $g \in \mathcal{M}(\tau)$, alors pour tous vecteurs intervallistes $[\mathbf{X}]$ et $[\mathbf{Y}]$ on a :

(i) $\mathcal{A}([\mathbf{X}]) \in \overline{\underline{\mathcal{A}}}([\mathbf{X}])$, et

(ii) $\mathcal{B}([\mathbf{Y}]) \in \overline{\underline{\mathcal{B}}}([\mathbf{Y}])$.

Ce résultat est valable également pour le cas des vecteurs précis (non intervallistes) puisque ce ne sont que des cas particuliers des vecteurs intervallistes, où les bornes supérieures et inférieures coïncident.

3.7 La technique de rétro-projection itérative imprécise

Maintenant que nous avons présenté comment construire les opérateurs de projection et de rétro-projection capables de tenir compte de la faible connaissance de la RI de l'imageur, il faut définir une façon de construire l'image hautement résolue recherchée.. La technique que nous proposons ici est basée sur une réinterprétation de l'algorithme de Schultz, comme cela a été présenté dans [Strauss et Rico, 2012].

D'abord, considérons le schéma de descente du gradient de l'Expression (3.3) :

$$\widehat{\mathbf{X}}^{i+1} = \widehat{\mathbf{X}}^{i} + R(\mathbf{Y} - A\widehat{\mathbf{X}}^{i}).$$

Cet algorithme essaye de trouver la solution $\widehat{\mathbf{X}}$ qui fasse tendre la quantité $(\mathbf{Y} - A\widehat{\mathbf{X}})$ vers $\mathbf{0}$, c.à.d. qui minimise l'erreur de projection par rapport aux données $\mathbf{Y}$. La mise à jour de l'estimation de l'itération précédente $\widehat{\mathbf{X}}^{i}$ est obtenue en lui ajoutant la rétro-projection de l'erreur de reconstruction $\mathbf{D}^{i}$ donnée par :

$$\mathbf{D}^{i} = R(\mathbf{Y} - A\widehat{\mathbf{X}}^{i}),$$

c.à.d.

$$\widehat{\mathbf{X}}^{i+1} - \widehat{\mathbf{X}}^{i} = \mathbf{D}^{i}.$$

Ensuite, réécrivons le calcul de $\mathbf{D}^{i}$ en nous basant sur les opérateurs de projection et de rétro-projection définis dans la Section 3.5 :

$$\mathbf{D}^{i} = \mathcal{B}(\mathbf{Y} - \mathcal{A}(\widehat{\mathbf{X}}^{i})).$$

Maintenant, l'extension de cet algorithme aux données et aux opérateurs intervallistes est triviale. Le calcul de $[\mathbf{D}]^{i}$, la valeur utilisée pour la mise à jour de la $i^{ème}$ estimation de l'algorithme itératif est donnée par :

$$[\mathbf{D}]^{i} = \overline{\underline{\mathcal{B}}}(\mathbf{Y} \ominus \overline{\underline{\mathcal{A}}}([\widehat{\mathbf{X}}]^{i})). \tag{3.80}$$

$[\mathbf{D}]^{i}$ représente l'ensemble convexe de toutes les corrections qui auraient pu être utilisées pour mettre à jour $[\widehat{\mathbf{X}}]^{i}$. Il peut être écrit par :

$$[\mathbf{D}]^{i} = [\widehat{\mathbf{X}}]^{i+1} \ominus [\widehat{\mathbf{X}}]^{i}.$$

La valeur mise à jour $[\widehat{\mathbf{X}}]^{i+1}$ est la solution de cette équation, et peut donc être calculée par :

$$[\widehat{\mathbf{X}}]^{i+1} = [\mathbf{D}]^{i} \boxplus [\widehat{\mathbf{X}}]^{i} = \overline{\underline{\mathcal{B}}}(\mathbf{Y} \ominus \overline{\underline{\mathcal{A}}}([\widehat{\mathbf{X}}]^{i})) \boxplus [\widehat{\mathbf{X}}]^{i}, \tag{3.81}$$

où $\boxplus$ désigne l'addition duale de Minkowski (voir Section 2.6). L'Expression (3.81) décrit finalement un algorithme imprécis de rétro-projection itérative, qui est une extension de l'algorithme de l'Expression (3.61), capable de tenir compte de la connaissance imprécise sur la RI de l'imageur. Cet algorithme essaye de trouver une solution intervalliste $[\widehat{\mathbf{X}}]$ qui fasse tendre $(\mathbf{Y} \ominus \overline{\underline{\mathcal{A}}}([\widehat{\mathbf{X}}]))$ vers $[\mathbf{0}]$, c.à.d un ensemble convexe de solutions qui, lorsqu'on lui applique l'ensemble convexe de projections $\overline{\underline{\mathcal{A}}}$, minimise l'erreur intervalliste $(\mathbf{Y} \ominus \overline{\underline{\mathcal{A}}}([\widehat{\mathbf{X}}]))$. On peut choisir, comme solution initiale $[\widehat{\mathbf{X}}]^{0}$, n'importe quelle approximation, intervalliste ou non. Cependant, la vitesse de convergence dépend du choix de cette solution. Dans nos travaux, nous avons choisi la solution intervalliste $[\widehat{\mathbf{X}}]^{0} = \overline{\underline{\mathcal{B}}}(\mathbf{Y})$ qui est équivalent au choix de la solution $[\mathbf{0}]$ mais avec une itération de moins dans l'algorithme. Nous avons remarqué expérimentalement que le choix d'une agrégation des images bassement résolues, comme solution initiale, fait converger plus rapidement l'algorithme, que lorsqu'on choisit la version interpolée d'une des images bassement résolues, comme il est souvent fait dans la littérature. Nous avons remarqué également que l'erreur quadratique entre $\mathbf{Y}$ et la médiane de $\overline{\underline{\mathcal{A}}}([\widehat{\mathbf{X}}]^{i})$ diminue tout au long des itérations, convergeant vers 0, ce qui montre que l'Expression (3.81) est bien une extension de l'algorithme classique de rétro-projection itérative.

CHAPITRE

4 Implémentation

L'implémentation de la technique de rétro-projection itérative imprécise est très simple : les opérations sont décrites dans la Section 2.6. De même, les opérations de projection et de rétro-projection basées sur l'intégrale de Choquet discrète donnée par l'Expression (2.3) se programme facilement. La seule difficulté concerne l'implémentation des deux FPV non-additives ν_n^k (Expression de projection (3.72)) et υ_m^k (Expression de rétro-projection (3.76)) basée sur l'intégrale de Choquet continue (Expression (2.1)).

L'implémentation que nous proposons ici est basée sur l'hypothèse que les transformations $\{t^k\}_{k\in\Theta_K}$ sont de simples translations et que les noyaux maxitifs 2D qui définissent les capacités 2D sont séparables, c.à.d. obtenus en multipliant deux noyaux maxitifs 1D définis le long de chaque axe. Dans ce cas, les opérateurs de projection et de rétro-projection peuvent aussi être décomposés sur chaque axe : une projection 2D (additive ou non-additive) consiste en deux projections successives 1D le long de chaque axe. C'est l'implémentation que nous proposons ici. Cette simplification nécessite aussi que les partitions floues soient séparables.

Afin d'alléger les notation, nous changeons chaque notation 2D en une notation 1D. Nous notons Ω un intervalle de $\mathbb{R}$ (un des axes), N le nombre d'échantillons de la grille bassement résolue de Ω et M le nombre d'échantillons de la grille hautement résolue de Ω. Nous appelons $\{\omega_n\}_{n\in\Theta_N}$ les N positions d'échantillonnage bassement résolus et $\{\varpi_m\}_{m\in\Theta_M}$ les M positions d'échantillonnage hautement résolus. K est le nombre d'images. $\{C_n\}_{n\in\Theta_N}$ est la partition floue linéaire bassement résolue de Ω et $\{B_m\}_{m\in\Theta_M}$ est la partition floue linéaire hautement résolue de Ω (voir Section 2.5). t^k est donc une simple translation monodimentionnelle qui s'écrit :

$$a_k \in \mathbb{R},\ \forall \omega \in \Omega,\ t^k(\omega) = \omega + a_k,$$

où $a_k \in \mathbb{R}$ est la valeur de la translation. $\jmath^k$ est alors donné par :

$$\forall \omega \in \Omega,\ \jmath^k(\omega) = \omega - a_k.$$

4.1 Calcul de la FPV non-additive 1D de projection

Dans cet algorithme, nous supposons que la connaissance imprécise sur le noyau sommatif f est représentée par le noyau maxitif triangulaire, symétrique et centré π, de support $[-\Delta, \Delta]$ ($\Delta \in \mathbb{R}^+$). Nous notons π_n^k le noyau maxitif associé à la $k^{ième}$ image bassement résolue ($k \in \Theta_K$) et au $n^{ième}$ pixel ($n \in \Theta_N$) défini par :

$$\forall \omega \in \Omega,\ \pi_n^k(\omega) = \pi(\omega_n - \jmath^k(\omega)) = \pi(\omega_n - \omega + a_k).$$

La distance unité est fixée par le pas d'échantillonnage de l'image hautement résolue :

$$\forall m \in \Theta_{M-1},\ \varpi_{m+1} - \varpi_m = 1.$$

4.1.1 Calcul de $\nu_n^k(\{m\})$, pour tout $m \in \Theta_M$

Le calcul de $\nu_n^k(\{m\})$, $\forall m \in \Theta_M$, basé sur l'intégrale de Choquet continue (Expression (2.1)), est donné par [Dubois, 2006] :

$$\nu_n^k(\{m\}) = \overline{\mathbb{E}}_{\Pi_{\pi_n^k}}(\mu_{B_m}) = \int_0^\infty \Pi_{\pi_n^k}(\{\omega \in \Omega, \mu_{B_m}(\omega) \geq \alpha\}) d\alpha. \tag{4.1}$$

Comme l'ensemble flou μ_{B_m} est normalisé, $\sup_{\omega \in \Omega} \mu_{B_m}(\omega) = 1$. L'intégrale de l'Expression (4.1) se réduit alors à :

$$\nu_n^k(\{m\}) = \int_0^1 \Pi_{\pi_n^k}(\{\omega \in \Omega, \mu_{B_m}(\omega) \geq \alpha\}) d\alpha. \tag{4.2}$$

Comme le pas d'échantillonnage hautement résolu est égal à 1 et en raison de la linéarité et de la symétrie de μ_{B_m}, l'ensemble $\{\omega \in \Omega, \mu_{B_m}(\omega) \geq \alpha\}$, pour tout $\alpha \in [0,1]$, est l'intervalle $[\varpi_m - 1 + \alpha, \varpi_m + 1 - \alpha]$ (voir Figure 4.1) :

$$\forall \alpha \in [0,1],\ \{\omega \in \Omega, \mu_{B_m}(\omega) \geq \alpha\} = [\varpi_m - 1 + \alpha, \varpi_m + 1 - \alpha].$$

ϖ_m étant le mode de μ_{B_m}. Notons ω_n^k le mode de π_n^k. Sans perte de généralité, en raison de la symétrie des fonctions concernées, on peut étudier uniquement le cas où $\varpi_m \leq \omega_n^k$. Comme π_n^k est croissante sur $[-\infty, \omega_n^k]$, on a :

$$\Pi_{\pi_n^k}([\varpi_m - 1 + \alpha, \varpi_m + 1 - \alpha]) = \Pi_{\pi_n^k}([\varpi_m, \varpi_m + 1 - \alpha]).$$

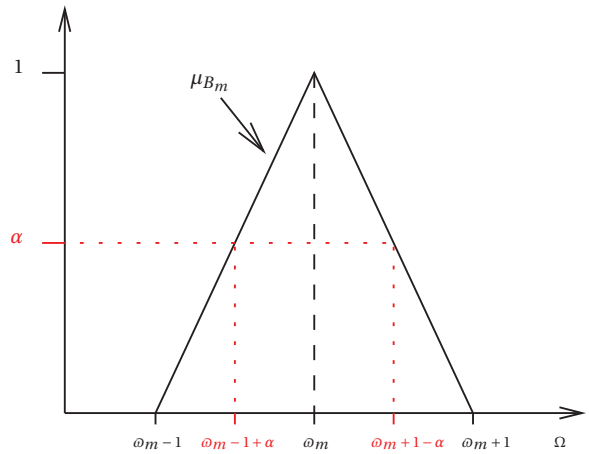


FIGURE 4.1 – Illustration de $\{\omega \in \Omega, \mu_{B_m}(\omega) \geq \alpha\} = [\varpi_m - 1 + \alpha, \varpi_m + 1 - \alpha]$

l'Expression (4.2) peut donc s'écrire :

$$v_n^k(\{m\}) = \int_0^1 \Pi_{\pi_n^k}([\varpi_m, \varpi_m + 1 - \alpha])d\alpha. \tag{4.3}$$

Nous proposons dans ce qui suit, de guider notre lecteur dans le calcul de $\Pi_{\pi_n^k}([\varpi_m, \varpi_m + 1 - \alpha])$ donné par :

$$\Pi_{\pi_n^k}([\varpi_m, \varpi_m + 1 - \alpha]) = \sup_{\omega \in [\varpi_m, \varpi_m + 1 - \alpha]} \pi_n^k(\omega).$$

L'idée est, ici, de permettre à tous de se familiariser avec ce type de calcul. Soit β la distance entre ω_n^k et ϖ_m, c.à.d. $\beta = |\omega_n^k - \varpi_m| = \omega_n^k - \varpi_m$ puisque on se place dans le cas où $\varpi_m \leq \omega_n^k$. On distingue 3 cas, illustrés par la Figure (4.2) :

(i) $\underline{\beta \in [1 - \alpha + \Delta, +\infty]} \iff \underline{\varpi_m + 1 - \alpha < \omega_n^k - \Delta}$ (Figure (4.2)(a)). Dans ce cas, par construction :

$$\Pi_{\pi_n^k}([\varpi_m, \varpi_m + 1 - \alpha]) = 0.$$

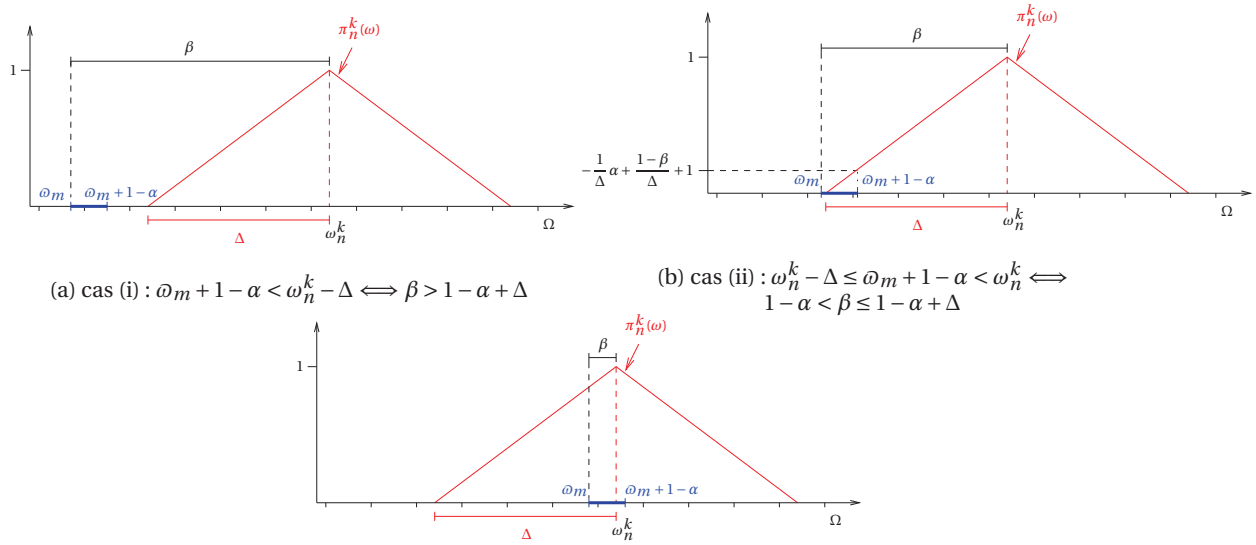
(ii) $\underline{\beta \in [1 - \alpha, 1 - \alpha + \Delta]} \iff \underline{\omega_n^k - \Delta \leq \varpi_m + 1 - \alpha < \omega_n^k}$ (Figure (4.2)(b))

$$\Pi_{\pi_n^k}([\varpi_m, \varpi_m + 1 - \alpha]) = \frac{1}{\Delta}(\varpi_m + 1 - \alpha) + 1 - \frac{1}{\Delta}\omega_n^k \tag{4.4}$$

$$= -\frac{1}{\Delta}\alpha + \frac{1 - \beta}{\Delta} + 1. \tag{4.5}$$

(iii) $\underline{\beta \in [-\infty, 1 - \alpha]} \iff \underline{\varpi_m + 1 - \alpha \geq \omega_n^k}$ (Figure (4.2)(c))

$$\Pi_{\pi_n^k}([\varpi_m, \varpi_m + 1 - \alpha]) = 1.$$



(a) cas (i) : $\varpi_m + 1 - \alpha < \omega_n^k - \Delta \iff \beta > 1 - \alpha + \Delta$

(b) cas (ii) : $\omega_n^k - \Delta \leq \varpi_m + 1 - \alpha < \omega_n^k \iff 1 - \alpha < \beta \leq 1 - \alpha + \Delta$

(c) cas (iii) : $\varpi_m + 1 - \alpha \geq \omega_n^k \iff \beta \leq 1 - \alpha$

FIGURE 4.2 – Calcul de $\Pi_{\pi_n^k}([\varpi_m, \varpi_m + 1 - \alpha]) = \sup_{\omega \in [\varpi_m, \varpi_m + 1 - \alpha]} \pi_n^k(\omega)$.

Le calcul de l'intégrale de l'Expression (4.3) nécessite alors de distinguer 4 cas en fonction de la valeur de β, illustrés par la Figure (4.3) :

1. $\beta \in [\Delta + 1, +\infty]$ (Figure(4.3)(a))
   Dans ce cas, pour toutes les valeurs de α dans $[0,1]$, on a $\beta \in [1 - \alpha + \Delta, +\infty]$. Le calcul de $\Pi_{\pi_n^k}([\varpi_m, \varpi_m + 1 - \alpha])$ coïncide donc avec le cas (i) pour tout $\alpha \in [0,1]$:
   $$\nu_n^k(\{m\}) = 0.$$

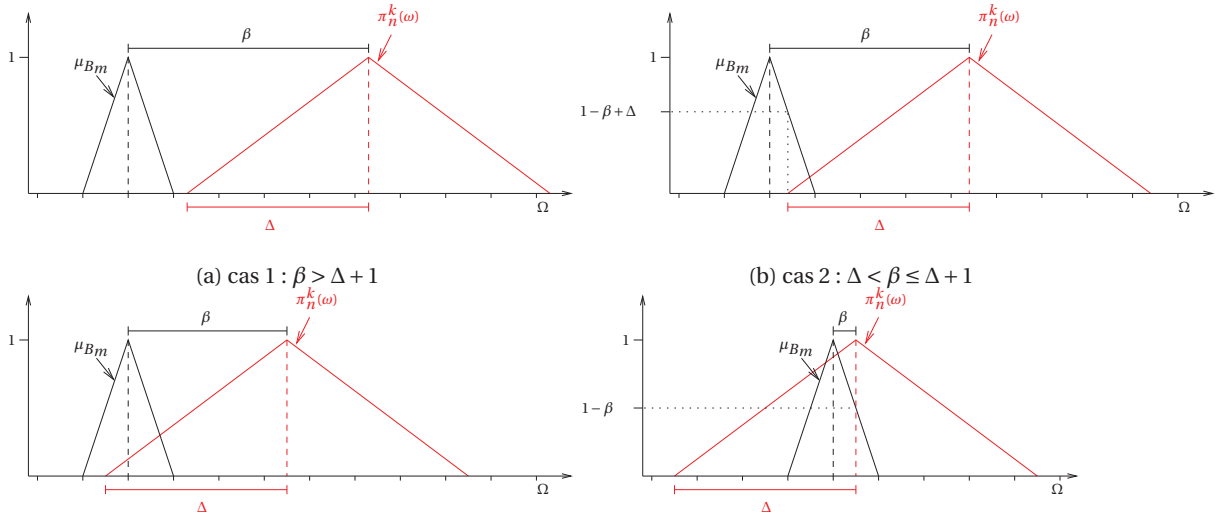
2. $\beta \in [\Delta, \Delta + 1]$ (Figure(4.3)(b))
   Dans ce cas, le calcul de $\Pi_{\pi_n^k}([\varpi_m, \varpi_m + 1 - \alpha])$ coïncide avec la cas (ii) pour $\alpha \in [0, 1 - \beta + \Delta]$ et avec le cas (i) pour $\alpha \in [1 - \beta + \Delta, 1]$. Donc :
   $$\nu_n^k(\{m\}) = \int_0^{1-\beta+\Delta} \left(-\frac{1}{\Delta}\alpha + \frac{1-\beta}{\Delta} + 1\right) d\alpha \tag{4.6}$$
   $$= \frac{-(\Delta - \beta + 1)^2}{2\Delta} + \frac{(1-\beta)(\Delta - \beta + 1)}{\Delta} + \Delta - \beta + 1. \tag{4.7}$$

3. $\beta \in [1, \Delta]$ (Figure(4.3)(c))
   Le calcul de $\Pi_{\pi_n^k}([\varpi_m, \varpi_m + 1 - \alpha])$ coïncide avec le cas (ii) pour $\alpha \in [0,1]$. Donc :
   $$\nu_n^k(\{m\}) = \int_0^1 \left(-\frac{1}{\Delta}\alpha + \frac{1-\beta}{\Delta} + 1\right) d\alpha \tag{4.8}$$
   $$= \frac{-1}{2\Delta} + \frac{1-\beta}{\Delta} + 1. \tag{4.9}$$

4. $\beta \in [-\infty, 1]$ (Figure(4.3)(d))
   Le calcul de $\Pi_{\pi_n^k}([\varpi_m, \varpi_m + 1 - \alpha])$ coïncide avec le cas (iii) pour $\alpha \in [0, 1-\beta]$ et avec le cas (ii) pour $\alpha \in [1-\beta, 1]$. Donc :

$$\nu_n^k(\{m\}) = \int_0^{1-\beta} 1 \, d\alpha + \int_{1-\beta}^{1} \left(-\frac{1}{\Delta}\alpha + \frac{1-\beta}{\Delta} + 1\right) d\alpha \tag{4.10}$$

$$= 1 - \frac{1-(1-\beta)^2}{2\Delta} + \frac{\beta(1-\beta)}{\Delta}. \tag{4.11}$$



(a) cas 1 : $\beta > \Delta + 1$

(b) cas 2 : $\Delta < \beta \leq \Delta + 1$

(c) cas 3 : $1 < \beta \leq \Delta$

(d) cas 4 : $\beta \leq 1$

FIGURE 4.3 – Calcul de $\nu_n^k(\{m\}) = \overline{\mathbb{E}}_{\Pi_{\pi_n^k}}(\mu_{B_m})$

Ces résultats sont valables pour le cas où $\varpi_m > \omega_n^k$ et les formules ne dépendent que de la distance $\beta = |\omega_n^k - \varpi_m|$.

L'implémentation du calcul de $\nu_n^k(\{m\}),\ \forall m \in \Theta_M$, est donné par l'Algorithme 1.

4.1.2 Calcul de $\nu_n^k(A)$, pour tout $A \subseteq \Theta_M$

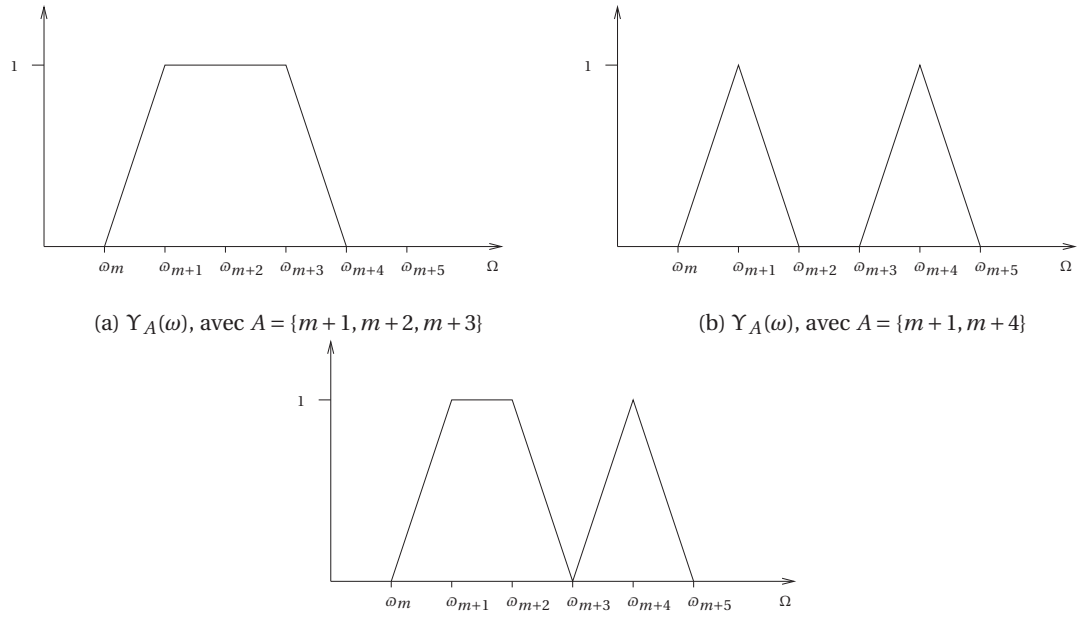
Les propriétés intéressantes de la partition floue *à la Ruspini* et de l'union définie par la T-conorme de Łukasiewicz (voir Section 2.5) conduisent à une expression simple de $\nu_n^k(A) = \overline{\mathbb{E}}_{\Pi_{\pi_n^k}}(\Upsilon_A),\ \forall A \subseteq \Theta_M$, en fonction des valeurs de $\nu_n^k(\{m\}),\ m \in A$. En effet, la fonction d'appartenance Υ_A a une forme :

– trapézoïdale si tous les éléments de A sont voisins (successifs) (voire exemple dans Figure (4.4)(a)),

Algorithm 1: Calcul de $v_n^k(m) = \overline{\mathbb{E}}_{\Pi_{\pi_n^k}}(\mu_{B_m}),\ \forall m \in \Theta_M$

Require: Δ, $m \in \Theta_M$.
  $\beta \leftarrow |\omega_n - \bar{\omega}_m|$;
  if $\beta \leq 1$ **then**
    $output \leftarrow 1 - \frac{1-(1-\beta)^2}{2\Delta} + \frac{\beta(1-\beta)}{\Delta}$;
  else
    if $\beta \leq \Delta$ **then**
      $output \leftarrow \frac{-1}{2\Delta} + \frac{1-\beta}{\Delta} + 1$;
    else
      if $\beta \leq \Delta + 1$ **then**
        $output \leftarrow \frac{-(\Delta-\beta+1)^2}{2\Delta} + \frac{(1-\beta)(\Delta-\beta+1)}{\Delta} + \Delta - \beta + 1$;
      else
        $output \leftarrow 0$;
      end if
    end if
  end if
  return $output$;

– d'un ensemble de triangles s'il n'existe pas d'éléments voisins dans A (voire exemple dans Figure (4.4)(b)),
– ou d'une combinaisons de triangles et de trapèzes s'il existe dans A des éléments qui sont voisins et d'autres qui ne le sont pas (voire exemple dans Figure (4.4)(c)).



(a) $\Upsilon_A(\omega)$, avec $A = \{m+1, m+2, m+3\}$

(b) $\Upsilon_A(\omega)$, avec $A = \{m+1, m+4\}$

(c) $\Upsilon_A(\omega)$, avec $A = \{m+1, m+2, m+4\}$

FIGURE 4.4 – Exemples de formes de la fonction d'appartenance Υ_A.

Dans les trois cas, parce que π_n^k est un noyau maxitif, le calcul de $\nu_n^k(A)$ se déduit aisément du calcul de la valeur de ν_n^k pour le point $m \in A$, dont le mode ϖ_m de l'ensemble flou B_m associé est le plus proche du mode du noyau maxitif π_n^k. Seule exception à cette règle : si le mode de π_n^k se trouve entre deux modes consécutifs ϖ_m et ϖ_{m+1} tels que $\{m, m+1\} \subseteq A$, alors $\nu_n^k(A) = 1$.

Posons $p \in \Theta_M$ tel que $\varpi_p \leq \omega_n < \varpi_{p+1}$. Soit L le nombre de sous-ensembles flous que le noyau π_n^k intersecte, c.à.d. le nombre de B_m tel que

$$\Pi_{\pi_n^k}(B_m) > 0,$$

et $l \in \mathbb{N}$ tel que[1] $l = \dfrac{L}{2} - 1$. La Figure (4.5) présente une configuration dans laquelle on cherche à calculer $\nu_n^k(A),\ \forall A \subseteq \Theta_M$. Ce calcul est effectué grâce à l'Algorithme 2 qui s'appuie sur le calcul des $\nu_n^k(\{m\})$ présenté dans l'Algorithme 1.

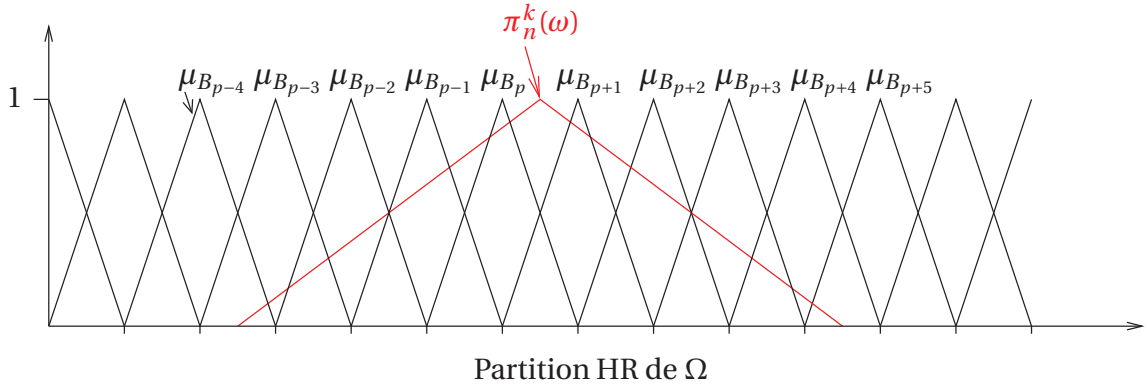


FIGURE 4.5 – Calcul de $\nu_n^k(A) = \overline{\mathbb{E}}_{\Pi_{\pi_n^k}}(\Upsilon_A),\ \forall A \subseteq \Theta_M$

4.2 Calcul de la FPV non-additive 1D de rétro-projection

De façon analogue à ce que nous avons proposé dans le cas de la projection, nous supposons que la connaissance imprécise sur la RI 1D de l'imageur hautement résolu est représentée par τ, le noyaux maxitif triangulaire, symétrique et centré, dont le support est $[-\delta, \delta]$ ($\delta \in \mathbb{R}^+$). Nous notons τ_m^k le noyau maxitif associé à la $k^{ième}$ image ($k \in \Theta_K$) et au $m^{ième}$ pixel ($m \in \Theta_M$), défini par :

$$\forall \omega \in \Omega,\ \tau_m^k(\omega) = \tau(\varpi_m - t^k(\omega)) = \tau(\varpi_m - \omega - a_k).$$

1. Par construction, L est un nombre pair, sauf dans le cas où le mode du noyau π_n^k coïncide avec un des modes de la partition. Si tel est le cas, on peut, sans perte de généralité, ajouter 1 à L pour en faire un nombre pair.

Algorithm 2: Calcul de $v_n^k(A) = \overline{\mathbb{E}}_{\Pi_{\pi_n^k}}(\Upsilon_A),\ \forall A \subseteq \Theta_M$

Require: $A \subseteq \Theta_M, l \in \mathbb{N}$ and Algorithm 1
  for $i = 0, \dots l$ **do**
    if $\{p-i, p+i+1\} \subseteq A$ **then**
      if $i \neq 0$ **then**
        $output \leftarrow \max(v_n^k(\{p-i\}), v_n^k(\{p+i+1\}));$;
      else
        $output \leftarrow 1$;
      end if
    else
      if $\{p-i\} \subseteq A$ **then**
        $output \leftarrow v_n^k(\{p-i\})$;
      else
        if $\{p+i+1\} \subseteq A$ **then**
          $output \leftarrow v_n^k(\{p+i+1\})$;
        else
          $output \leftarrow 0$;
        end if
      end if
    end if
  end for
  return $output$;

4.2.1 Calcul de $v_m^k(\{n\})$, pour tout $m \in \Theta_M$

Le calcul de $v_m^k,\ \forall m \in \Theta_M$, basé sur l'intégrale de Choquet continue (Expression (2.1)) est donné par

$$v_m^k(\{n\}) = \overline{\mathbb{E}}_{\Pi_{\tau_m^k}}(\mu_{C_n}) = \int_0^{\infty} \Pi_{\tau_m^k}(\{\omega \in \Omega, \mu_{C_n}(\omega) \geq \alpha\})\, d\alpha. \tag{4.12}$$

Soit β la distance entre le mode du noyau τ_m^k et celui du sous-ensemble flou μ_{C_n}. Le calcul de l'intégrale de l'Expression (4.12) nécessite de distinguer 4 cas en fonction de la valeur de β, illustrés par la Figure(4.6) :

1. $\underline{\beta \in [\delta+1, +\infty]}$ (Figure(4.6)(a))
$$v_m^k(\{n\}) = 0.$$
2. $\underline{\beta \in [1, \delta+1]}$ (Figure(4.6)(b))
$$v_m^k(\{n\}) = \int_0^{1-\beta+\delta} \left(-\frac{1}{\delta}\alpha + \frac{1-\beta}{\delta} + 1\right) d\alpha \tag{4.13}$$
$$= \frac{-(\delta-\beta+1)^2}{2\delta} + \frac{(1-\beta)(\delta-\beta+1)}{\delta} + \delta - \beta + 1. \tag{4.14}$$

3. $\beta \in [\delta, 1]$ (Figure(4.6)(c))

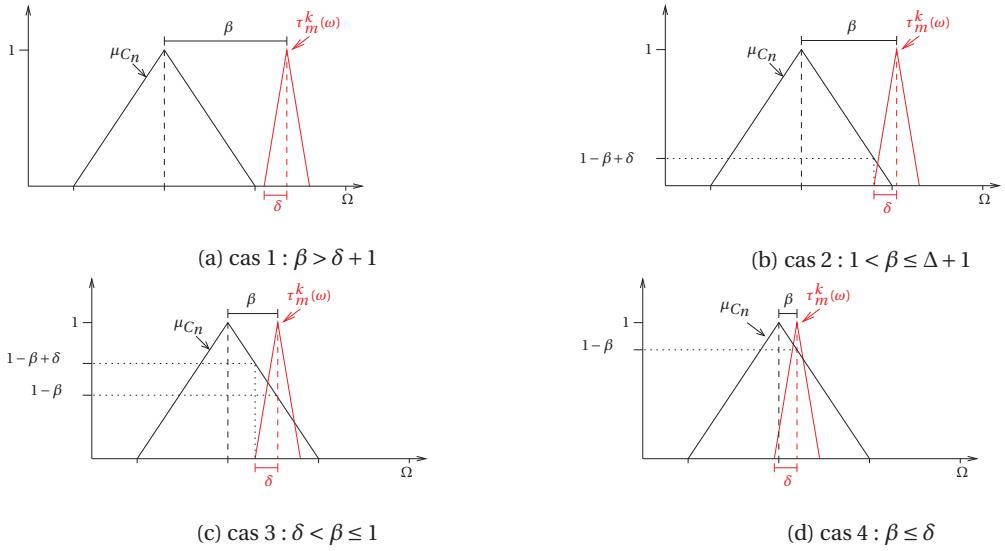
$$v_m^k(\{n\}) = \int_0^{1-\beta} 1 \, d\alpha + \int_{1-\beta}^{1-\beta+\delta} \left(-\frac{1}{\delta}\alpha + \frac{1-\beta}{\delta} + 1\right) d\alpha \tag{4.15}$$

$$= 2 - 2\beta - \frac{(\delta-\beta+1)^2 - (1-\beta)^2}{2\delta} + \delta. \tag{4.16}$$

4. $\beta \in [-\infty, \delta]$ (Figure(4.6)(d))

$$v_m^k(\{n\}) = \int_0^{1-\beta} 1 \, d\alpha + \int_{1-\beta}^{1} \left(-\frac{1}{\delta}\alpha + \frac{1-\beta}{\delta} + 1\right) d\alpha \tag{4.17}$$

$$= 1 - \frac{1-(1-\beta)^2}{2\delta} + \frac{\beta(1-\beta)}{\delta}. \tag{4.18}$$



(a) cas 1 : $\beta > \delta + 1$

(b) cas 2 : $1 < \beta \leq \Delta + 1$

(c) cas 3 : $\delta < \beta \leq 1$

(d) cas 4 : $\beta \leq \delta$

FIGURE 4.6 – Calcul de $v_m^k(\{n\}) = \overline{\mathbb{E}}_{\Pi_{\tau_m^k}}(\mu_{C_n})$

L'implémentation du calcul de $v_m^k(\{n\})$, $\forall n \in \Theta_N$, est donné par l'Algorithme 3.

4.2.2 Calcul de $v_m^k(A)$, pour tout $A \subseteq \Theta_N$

Le calcul de $v_m^k(A)$ se déduit aisément du calcul de la valeur de v_m^k pour le point $n \in A$, dont le mode ω_n de l'ensemble flou C_n associé est le plus proche du mode du noyau maxitif τ_m^k. Seule exception à cette règle : si le mode de τ_m^k se trouve entre deux modes consécutifs ω_n et ω_{n+1} tels que $\{n, n+1\} \subseteq A$, alors $v_m^k(A) = 1$.

Algorithm 3: Calcul de $v_m^k(\{n\}) = \overline{\mathbb{E}}_{\Pi_{\tau_m^k}}(\mu_{C_n}),\ \forall n \in \Theta_N$

Require: $0 < \delta \leq 1,\ n \in \Theta_N$.
  $\beta \leftarrow |\omega_n - \varpi_m|$
  if $\beta \leq \delta$ **then**
    $output \leftarrow 1 - \frac{1-(1-\beta)^2}{2\delta} + \frac{\beta(1-\beta)}{\delta}$;
  else
    if $\beta \leq 1$ **then**
      $output \leftarrow 2 - 2\beta - \frac{(\delta-\beta+1)^2-(1-\beta)^2}{2\delta} + \delta$;
    else
      if $\beta \leq \delta + 1$ **then**
        $output \leftarrow \frac{-(\delta-\beta+1)^2}{2\delta} + \frac{(1-\beta)(\delta-\beta+1)}{\delta} + \delta - \beta + 1$;
      else
        $output \leftarrow 0$;
      end if
    end if
  end if
  return $output$;

Soit $p \in \Theta_N$ tel que $\omega_p \leq \varpi_m < \omega_{p+1}$. La Figure (4.7) présente une configuration dans laquelle on cherche à calculer $v_m^k(A),\ \forall A \subseteq \Theta_N$. Ce calcul est effectué grâce à l'Algorithme 4 qui s'appuie sur le calcul des $v_m^k(\{n\})$ présenté dans l'Algorithme 3.

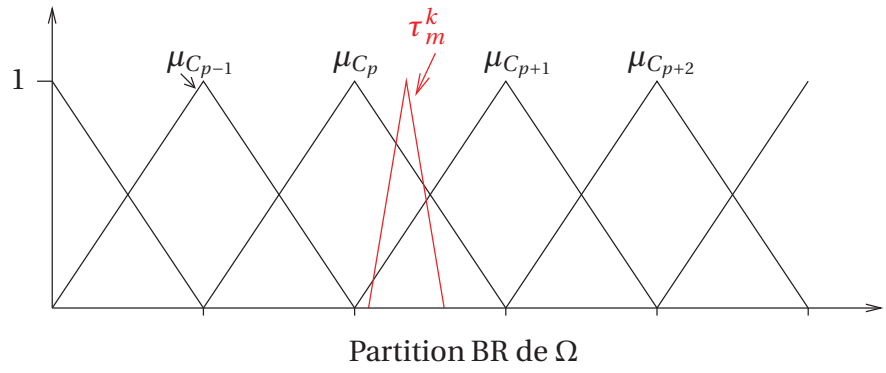


FIGURE 4.7 – Calcul de $v_m^k(A) = \overline{\mathbb{E}}_{\Pi_{\tau_m^k}}(\Upsilon_A), \forall A \subseteq \Theta_N$

4.3 Complexité algorithmique

Nous proposons dans cette section une analyse de la complexité algorithmique temporelle de la technique de super-résolution imprécise que nous proposons, en fonction de plusieurs paramètres, à savoir, le facteur de rehaussement, que nous notons e, le nombre de pixels dans chaque image bassement résolue N, le nombre d'images K et le nombre

Algorithm 4: Calcul de $v_m^k(A) = \overline{\mathbb{E}}_{\Pi_{\tau_m^k}}(\Upsilon_A), \forall A \subseteq \Theta_N$

Require: $A \subseteq \Theta_N$ and Algorithm 3
  for $i = 0, \ldots 1$ **do**
    if $\{p-i, p+i+1\} \subseteq A$ **then**
      if $i \neq 0$ **then**
        $output \leftarrow \max(v_m^k(\{p-i\}), v_m^k(\{p+i+1\}))$;
      else
        $output \leftarrow 1$;
      end if
    else
      if $\{p-i\} \subseteq A$ **then**
        $output \leftarrow v_m^k(\{p-i\})$;
      else
        if $\{p+i+1\} \subseteq A$ **then**
          $output \leftarrow v_m^k(\{p+i+1\})$;
        else
          $output \leftarrow 0$;
        end if
      end if
    end if
  end for
  return $output$;

d'itérations de l'algorithme, que nous notons J. Nous ne traitons pas la complexité spatiale car l'exécution de l'algorithme ne requiert pas plus d'espace que celui des données d'entrée.

Pour ce faire, nous commençons par analyser les deux opérations principales de cette technique qui sont la projection et la rétro-projection imprécises. Nous commençons par l'opération de projection donnée par l'Expression (3.73). Le calcul de cette projection est basé sur l'intégrale de Choquet discrète (Expression (2.3)) et le calcul de la FPV ν_n^k :

$$\forall k \in \Theta_K, \forall n \in \Theta_N, [y]_n^k = \overline{\underline{\mathbb{E}}}_{\nu_n^k}(\mathbf{X}) \tag{4.19}$$

$$= [\underline{\mathbb{E}}_{\nu_n^k}(\mathbf{X}), \overline{\mathbb{E}}_{\nu_n^k}(\mathbf{X})] \tag{4.20}$$

$$= [\mathbb{C}_{\nu_n^{k\,c}}(\mathbf{X}), \mathbb{C}_{\nu_n^k}(\mathbf{X})], \tag{4.21}$$

où $\nu_n^{k\,c}$ est la FPV conjuguée de ν_n^k. Le calcul de $\mathbb{C}_{\nu_n^k}(\mathbf{X})$ est donné par :

$$\mathbb{C}_{\nu_n^k}(\mathbf{X}) = \sum_{m \in \Theta_M} \mathbf{X}_{(m)\uparrow} \left(\nu_n^k(A_{(m)}) - \nu_n^k(A_{(m+1)}) \right),$$

où $\mathbf{X}_{(m)\uparrow}$ indique que les indices ont été permutés de manière à ce que

$$\mathbf{X}_{(1)} \leq \ldots \leq \mathbf{X}_{(M)},$$

et

$$A_{(m)} = \{(m), \ldots, (p)\} \text{ et } A_{(M+1)} = \varnothing.$$

Le calcul de $\mathbb{C}_{\nu_n^{k^c}}(\mathbf{X})$ peut, quant à lui, être donné par :

$$\mathbb{C}_{\nu_n^{k^c}}(\mathbf{X}) = \sum_{m\in\Theta_M} \mathbf{X}_{(m)\downarrow}\left(\nu_n^k(A_{(m)}) - \nu_n^k(A_{(m+1)})\right),$$

où $\mathbf{X}_{(m)\downarrow}$ indique que les indices ont été permutés de manière à ce que

$$\mathbf{X}_{(1)} \geq \ldots \geq \mathbf{X}_{(M)}.$$

Notons $\mathcal{V}^k(n)$, le sous-ensemble de Θ_M défini par :

$$\mathcal{V}^k(n) = \{m \in \Theta_M | \nu_n^k(m) > 0\}.$$

Le calcul de la projection imprécise peut être réduit à l'ensemble $\mathcal{V}^k(n)$, c.à.d :

$$\mathbb{C}_{\nu_n^k}(\mathbf{X}) = \sum_{m\in\mathcal{V}^k(n)} \mathbf{X}_{(m)\uparrow}\left(\nu_n^k(A_{(m)}) - \nu_n^k(A_{(m+1)})\right), \tag{4.22}$$

et

$$\mathbb{C}_{\nu_n^{k^c}}(\mathbf{X}) = \sum_{m\in\mathcal{V}^k(n)} \mathbf{X}_{(m)\downarrow}\left(\nu_n^k(A_{(m)}) - \nu_n^k(A_{(m+1)})\right). \tag{4.23}$$

La Figure (4.5) montre qu'en 1D, $|\mathcal{V}^k(n)|$ (le cardinal de $\mathcal{V}^k(n)$), noté L dans la Section 4.1.2, est généralement égal à $2(\Delta+1)$ excepté le cas où le mode de Π_n^k coïncide avec le mode d'un sous-ensemble flou de la partition, dans ce cas, il est égal à $2\Delta+1$. Nous considérons donc le pire des cas, c.à.d. $L = 2(\Delta+1)$. Δ étant borné par le pas d'échantillonnage bassement résolu qui est égal au facteur de rehaussement [2] $e = \frac{M}{N}$, on peut écrire $L = 2(e+1)$. Chacune des Expressions (4.22) et (4.23) nécessite un tri des éléments $\{\mathbf{X}_m | m \in \mathcal{V}^k(n)\}$ et $(2L = 4(e+1))$ calculs de la FPV $\nu_n^k(A)_{A\in\Theta_M}$. Le calcul de tout $\nu_n^k(A)$ se fait avec l'Algorithme 2 qui, dans le pire des cas, fait appel à l'Algorithme 1 $\left(\frac{L}{2}+1\right)$ fois. l'Algorithme 1 ayant une complexité constante, il peut être considéré comme une instruction élémentaire. La complexité de l'Algorithme 2 est donc de $\frac{L}{2}+1 = e+2$. Le tri peut se faire en un temps $L\,log(L) = 2(e+1)log(2(e+1))$ en utilisant par exemple un algorithme de *tri fusion* ou de *tri par tas*.

L'exécution de chacune des Expressions (4.22) et (4.23) se fait donc en un temps

$$4(e+1)(e+2) + 2(e+1)log(2(e+1)) = 4e^2 + 12e + 8 + 2(e+1)log(2(e+1)),$$

l'exécution des deux Expressions (4.22) et (4.23), c.à.d. la projection 1D d'un pixel bassement résolu se fait en un temps

$$8e^2 + 24e + 16 + 4(e+1)log(2(e+1)),$$

2. On peut fixer le pas d'échantillonnage hautement résolu à 1 et ensuite fixer le pas d'échantillonnage bassement résolu au facteur de rehaussement choisi.

et la projection 2D multiplie par 2 cette complexité, c.à.d. qu'elle s'exécute en un temps

$$\phi(e) = 16e^2 + 48e + 32 + 8(e+1)log(2(e+1)).$$

La projection de tous les pixels des K images bassement résolues de taille N se fait alors en un temps

$$K \cdot N \cdot \phi(e).$$

En utilisant le même procédé et le fait que δ est borné par le pas d'échantillonnage hautement résolue qui est égal à 1, on peut montrer que la rétro-projection d'un pixel hautement résolu s'exécute en un temps constant, et la rétro-projection de toute l'image hautement résolue s'exécute en un temps $M = N \cdot e$. En considérant une itération de l'algorithme comme étant une projection suivie d'une une rétro-projection, une itération s'exécute en un temps

$$K \cdot N \cdot \phi(e) + N \cdot e.$$

Enfin, si J est le nombre total d'itérations de l'algorithme imprécis de super-résolution, sa complexité est de

$$J \cdot K \cdot N \cdot \phi(e) + J \cdot N \cdot e = O\left(J \cdot K \cdot N \cdot e^2\right).$$

La complexité de notre méthode est du même ordre que celle de la rétro-projection itérative classique dont elle est une extension, mais son temps d'exécution est environ deux fois plus important. Ceci est certainement dû au traitement, à chaque itération, de deux fois plus de données (à cause de la nature intervalliste des images manipulées) et à l'utilisation de l'intégrale de Choquet discrète qui nécessite un tri des valeurs de la fonction considérée. Mais le temps d'exécution de notre méthode peut être divisé par deux en remarquant qu'à chaque projection et rétro-projection, le traitement des données représentant les bornes supérieures peut être réalisé en parallèle avec le traitement des données représentant les bornes inférieures.

4.4 Conclusion

Nous avons proposé dans ce chapitre, une implémentation des principaux opérateurs de notre méthode de super-résolution imprécise qui sont les deux FPV v_n^k et v_m^k.

Cette implémentation est grandement simplifiée par les partitions floues *à la Ruspini* linéaires que nous avons utilisées et les hypothèses que les RI sont séparables et que les mouvements entre les images bassement résolues sont de pures translations.

Cette implémentation est générique, au sens où elle est donnée pour des valeurs quelconques des demi-largeurs des noyaux maxitifs π et τ, qui sont respectivement Δ et δ. Ces deux derniers sont les seuls paramètres à fixer dans cette implémentation.

Dans le cas de présence de rotations entre les images bassement résolues, l'implémentation n'est plus la même. En effet, les opérateurs de projection et de rétro-projection ne

sont plus décomposables, même dans l'hypothèse que les RI sont séparables. Dans ce cas, le calcul des FPV v_n^k et v_m^k nécessite le calcul d'intégrales de Choquet 2D qui est plus compliquée que le calcul des intégrales de Choquet 1D.

CHAPITRE

5 Expérimentations

Dans ce chapitre, nous illustrons les performances de la méthode que nous proposons, et nous comparons ces performances à celles de quelques autres méthodes de super-résolution, dans un contexte semi-aveugle, c.à.d. lorsque la RI de l'imageur est mal connue. Cette comparaison se base sur des séquences d'images synthétiques et des séquences d'images réelles. Dans ces expérimentations, les RI sont supposées être séparables et avoir la même résolution dans les deux directions. Le facteur de rehaussement de résolution est fixé à 4×4, c.à.d. 4 dans chaque direction, pour toutes les méthodes considérées dans ces expérimentations. Dans la suite, nous utiliserons l'acronyme SRNA (Super-Résolution Non-Additive) pour désigner la méthode imprécise de super-résolution que nous proposons.

Conformément à l'argumentation développée dans la Section 3.6, nous modélisons la mauvaise connaissance sur les RI, dans la SRNA, par des noyaux maxitifs 2D où chacun est obtenus en multipliant deux noyaux maxitifs triangulaires dont la demi largeur des supports est Δ pour la projection, et δ pour la rétro-projection (voir Section 4). Comme le facteur de rehaussement de résolution est fixé à 4×4, nous fixons δ à $\frac{\Delta}{4}$.

Les méthodes considérées dans ces expérimentations étant toutes itératives, se pose le problème du nombre d'itérations à considérer pour chacune d'elles lors d'une comparaison. En effet, le nombre d'itérations a un impact certain sur les résultats, sachant qu'il n'existe pas de méthode standard fiable pour détecter le nombre optimal d'itérations. Ce problème se pose naturellement aussi pour la SRNA. C'est pourquoi nous proposons la solution suivante : lorsqu'une image de référence hautement résolue est disponible, nous considérons, pour chaque méthode, l'itération qui donne le meilleur résultat en terme de distance quadratique entre l'image reconstruite et l'image de référence. Lorsqu'aucune image de référence n'est disponible, nous fixons le nombre d'itérations à 8, car les diffé-

rentes expérimentations que nous avons menées ont majoritairement abouti à un nombre optimal d'itérations compris entre 7 et 9, et ce quelle que soit la méthode utilisée.

Un autre problème se pose pour comparer la méthode SRNA aux méthodes issues de la littérature, qui est la nature intervalliste de l'image reconstruite. Afin de rendre la comparaison la plus objective possible, nous proposons d'utiliser l'image médiane, c.à.d. l'image minimisant la distance de Hausdorff à l'image intervalliste reconstruite.

La première expérimentation illustre la sensibilité de la méthode IBP, qui utilise une modélisation précise de la RI de l'imageur considéré à cette modélisation. Elle met en évidence la robustesse de la SRNA à cette modélisation, c.à.d. sa capacité à mettre à profit une connaissance imprécise sur la RI de l'imageur. La deuxième expérimentation met en évidence l'apport informatif véhiculé par l'imprécision d'une image intervalliste reconstruite par la SRNA. On montre ainsi d'une part, que l'imprécision de cette image est fortement corrélée à l'erreur de reconstruction et, d'autre part, que l'image qui aurait été reconstruite avec une méthode analogue précise utilisant le bon modèle de la RI est majoritairement incluse dans l'image imprécise reconstruite par SRNA. La troisième expérimentation illustre la robustesse de la SRNA aux erreurs de recalage. La quatrième expérimentation vise à comparer les performances de la SRNA avec celles de quelques autres méthodes issues de la littérature lorsqu'on considère des séquences d'images réelles.

5.1 Utilisation d'une connaissance imprécise de la RI

Le but de cette expérimentation est, d'une part, de montrer la sensibilité de la méthode IBP à la modélisation de la RI de l'imageur et, d'autre part, d'illustrer la performance de la SRNA utilisant une connaissance imprécise de cette RI. Nous avons utilisé l'image hautement résolue 'eia' de 452×452 pixels, souvent utilisée pour mettre en évidence les performances de méthodes de super-résolution [Farsiu *et al.*, 2004b], donnée par la Figure (5.3.a). Nous avons fait un zoom sur un détail sensible de l'image car il contient du text et un faisceau convergent de droites. C'est sur ce détail qu'on se basera pour faire une comparaison qualitative. Nous avons généré synthétiquement une séquence de 16 images bassement résolues en utilisant la projection définie dans la Section 3.5 et en appliquant uniquement des translations. L'effet de la RI est simulée par la convolution avec un noyau quadratique à bande limitée dont la demi largeur de son support Δ est égale à 2.5. Le facteur de décimation est de 4×4. Les images obtenues ont ensuite été quantifiées c.à.d. les valeurs d'intensité des pixels ont été codées sur une échelle entière ($[0, 255]$). La Figure (5.1) montre les 6 premières images de la séquence.

Nous utilisons la méthode IBP pour effectuer 4 reconstructions : deux reconstructions basées sur une modélisation du noyau de convolution ayant la forme appropriée (quadratique) mais dont un ayant une demi-largeur Δ égale à 1.75 (sous-estimée) et l'autre une demi-largeur Δ égale à 3 (sur-estimée), une reconstruction basée sur un noyau de convolution cubique ayant la taille de support appropriée (demi largeur égale à 2.5) et une recons-

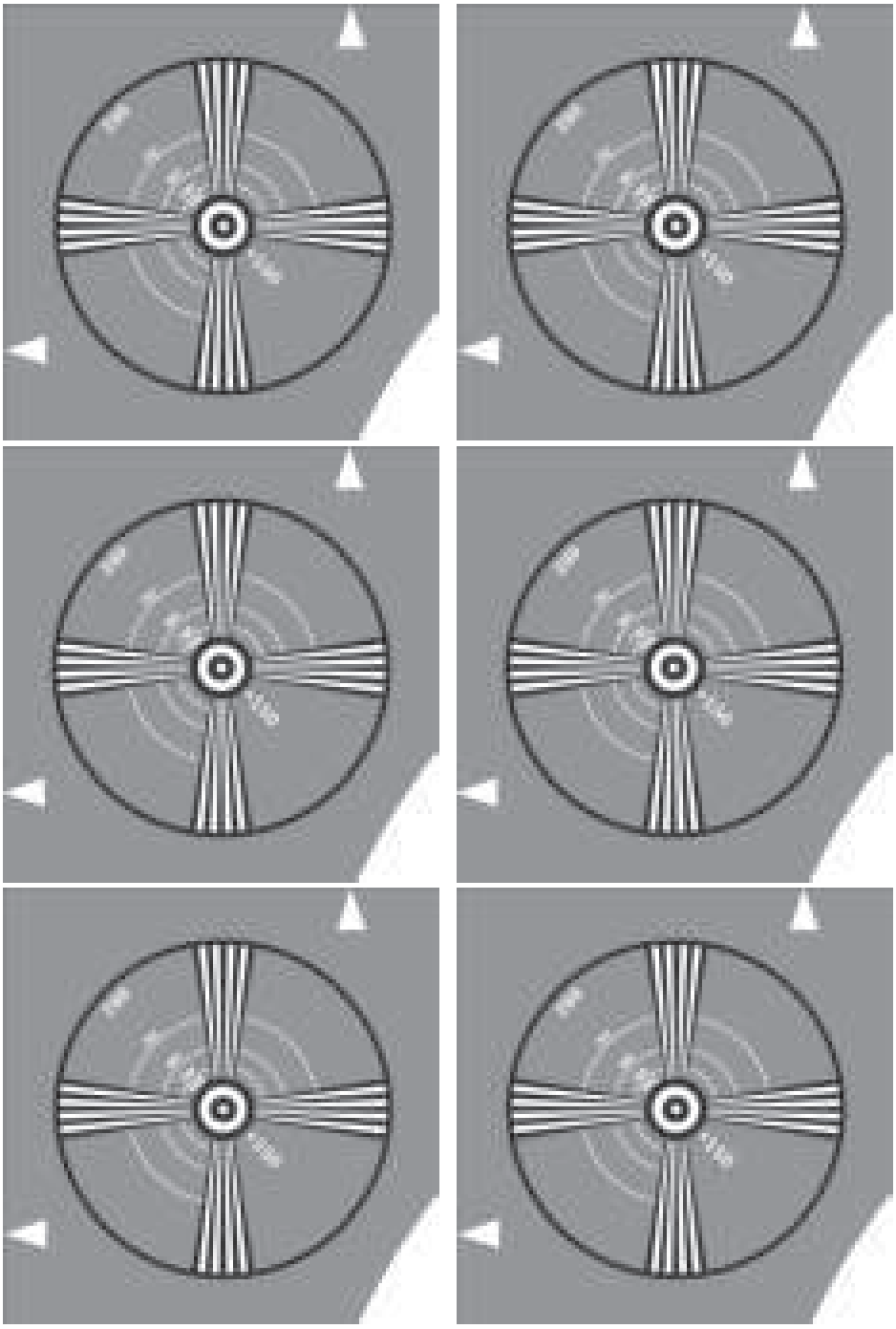


FIGURE 5.1 – 6 des 16 images bassement résolues générées synthétiquement

truction que, par la suite, nous appellerons optimale, basée sur un noyau de convolution ayant la forme et la demi-largeur appropriées, c.à.d. celles que nous avons utilisées pour créer la séquence synthétique. Nous nous mettons ensuite dans un contexte semi-aveugle, c.à.d. nous considérons que la seule connaissance que nous avons sur la RI de l'imageur est d'être un noyau sommatif, symétrique et monomodal, de demie-largeur $\Delta \leq 3$. Nous modélisons cette connaissance imprécise par un noyau maxitif bidimensionnel dont la demi-largeur du support Δ suivant chaque axe est égale à 3 et nous effectuons une reconstruction avec la méthode SRNA.

Afin d'éviter l'influence des erreurs de recalage, les valeurs des translations n'ont pas été estimées, nous avons utilisé les valeurs ayant servi pour générer la séquence.

Les images (b) à (f) de la Figure (5.3) présentent une vue subjective des résultats. Pour permettre à notre lecteur de se faire une opinion sur la qualité des résultats obtenus, nous avons fait un zoom sur le même détail que sur l'image originale (image (a)). Notons que les images (c) à (f) ont été obtenues après le même nombre d'itérations (8). Seule la reconstruction optimale (b) a été obtenue après 15 itérations, car l'adéquation entre modèle utilisé pour simuler l'acquisition et modèle utilisé pour reconstruire permet de pousser les itérations plus loin alors que le biais de modèle, dans les autres cas, introduit une divergence de la reconstruction après 8 itérations. Une présentation plus objective des résultats est donnée par un calcul de la distance quadratique, exprimée par le PSNR, entre les images reconstruites et l'image originale (Tableau (5.1)).

Dans un premier temps, on peut noter sur le Tableau (5.1), que la reconstruction de l'image hautement résolue en utilisant IBP avec le noyau utilisé pour générer la séquence (Expérience $quad_{2.5}$) donne l'image qui est la plus proche, en PSNR, de l'image originale. Par contre, un autre choix de taille du support du noyau (Expériences $quad_3$ et $quad_{1.75}$), ou de forme du noyau (Expérience $cub_{2.5}$) aboutit à une dégradation du PSNR, montrant ainsi la sensibilité de la méthode IBP à la modélisation de la RI. Par opposition, la SRNA donne un score de PSNR proche de celui obtenu par la reconstruction optimale. Ce score de PSNR est supérieur à ceux obtenus avec des modèles de RI proches mais erronées.

En regardant de près les Figures (5.3) (b à f), on peut constater que toutes les reconstructions présentent des artefacts de Gibbs (oscillations près des zones de changement brusque de l'intensité). Il apparait que les détails sont mieux reconstruits dans les images

Méthode de SR	IBP				SRNA
Noyau de la RI	$cub_{2.5}$	$quad_3$	$quad_{1.75}$	$quad_{2.5}$	
PSNR	22.24	23.40	23.19	24.03	23.84

TABLE 5.1 – Valeurs du PSNR de la méthode IBP avec différentes formes et tailles de support du noyau de convolution, et de la SRNA.

(b) et (f) que dans les autres reconstructions, ce que l'on peut constater aisément en comparant les caractères des nombres 150 et 200.

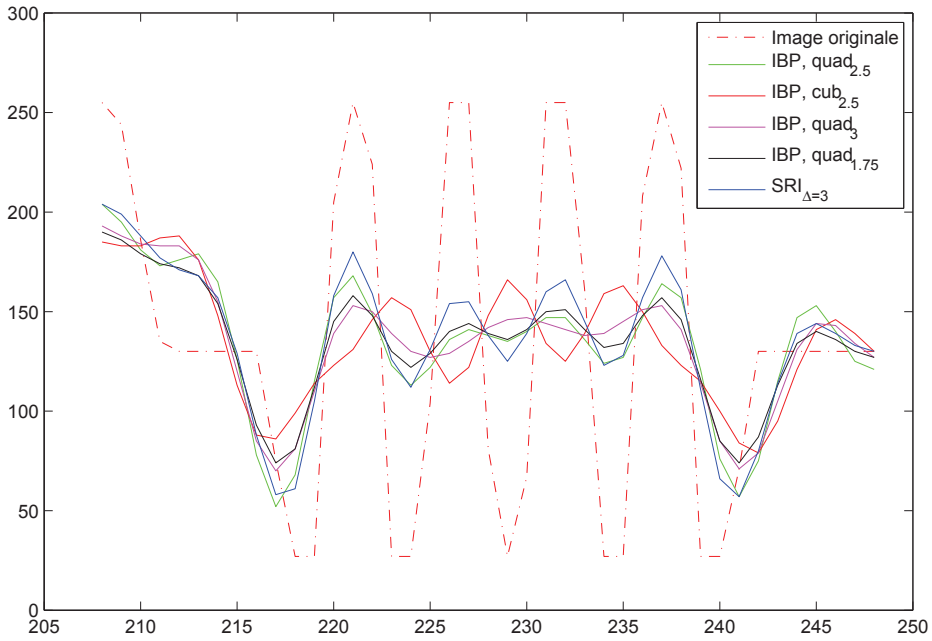


FIGURE 5.2 – Niveaux de gris d'un segment de la ligne 270 des différentes reconstructions et de l'image originale de la Figure 5.3.

La Figure (5.2) trace, pour chacune des reconstructions, les niveaux de gris d'un segment de la ligne 270 de l'image. Ce segment correspond à la zone encadrée illustrée par la Figure (5.4). On peut constater que les deux signaux reconstruits les plus proches du signal original sont ceux obtenus par la reconstruction optimale (en vert) et la SRNA (en bleu) confirmant ainsi les résultats obtenus par PSNR. On peut noter que le signal obtenu par la reconstruction IBP en utilisant le noyau cubique avec $\Delta = 3$ (ligne continue rouge) va dans le sens opposé du signal original entre les positions 223 et 235. Ceci confirme la forte sensibilité de la méthode IBP au choix du modèle de la RI.

	Seq1	Seq2	Seq3	Seq4	Seq5	Seq6	Seq7	Seq8	Seq9	Seq10	Seq11	Seq12
Recons. optimales	95.48	95.35	95.59	95.11	94.33	93.22	94.34	94.33	94.67	94.38	93.93	93.43
Image originale	75.27	74.14	72.71	70.85	68.86	66.98	77.10	75.99	74.96	73.73	72.30	70.79

TABLE 5.2 – Ligne 1 : taux d'inclusion en % de la reconstruction optimale dans la reconstruction SRNA. Ligne 2 : taux d'inclusion en % de l'image originale dans la reconstruction SRNA.

5.2 L'apport d'information véhiculée par l'imprécision de l'image reconstruite

Cette expérimentation vise à illustrer deux propriétés intéressantes de la SRNA. La première propriété est l'aptitude de la reconstruction intervalliste à inclure l'image originale et l'image qui aurait été obtenue en utilisant le modèle opportun de la RI, qu'on appellera par la suit, la reconstruction optimale. La deuxième est une propriété classique du filtrage non-additif [] qui est la corrélation entre l'imprécision des images intervallistes obtenues par la SRNA d'une part, et l'erreur de reconstruction d'autre part.

Nous avons utilisé l'image benchmark classique 'Lena' de 512×512 pixels comme référence hautement résolue et simulé son acquisition synthétiquement par 12 caméras différentes. Cette simulation est obtenue en générant 12 séquences de 16 images bassement résolues en niveaux de gris en utilisant la projection basée sur les partitions floues (voir Section 3.5) avec un facteur de décimation de 4×4. Nous avons considéré uniquement des translations. Chaque séquence a été générée en utilisant un modèle de RI différent via 6 noyaux cubiques et 6 noyaux quadratiques, avec des demie-largeurs différentes $\Delta \in [1.75, 3]$.

Pour chacune des 12 séquences, nous effectuons une reconstruction utilisant IBP avec le bon noyau (reconstruction optimale). Nous nous mettons ensuite dans un contexte semi-aveugle et supposons que la seule connaissance que nous avons sur le noyau de la RI est qu'il est sommatif, symétrique et monomodal dont la demi largeur $\Delta \leq 3$ et nous effectuons une reconstruction en utilisant la SRNA avec un noyau maxitif dont la demi largeur $\Delta = 3$. La méthode IBP a été itérée 20 fois et la SRNA 8 fois. Ces valeurs ont été choisies pour obtenir la solution la plus proche, en distance quadratique, de l'image originale. Le facteur de rehaussement de résolution a été fixé à 4×4 pour toutes les reconstructions. Les valeurs des translations utilisées pour les reconstructions sont celles ayant servi à générer les séquences.

Le Tableau 5.2 montre, pour chacune des 12 séquences, le taux d'inclusion de, la reconstruction optimale d'une part (ligne 1), et l'image originale d'autre part (ligne 2), dans l'image intervalliste de la reconstruction SRNA, c.à.d. le pourcentage des intensités de pixels (reconstruites ou originales) se trouvant entre les bornes des pixels correspondants de la reconstruction avec la SRNA. Le Tableau 5.3 dresse les statistiques du Tableau 5.2. On

Statistique des taux d'inclusion	Min	Mean	Max	Std.
Recons. optimales	93.22%	94.51%	95.59%	0.77
Image originale	66.98%	72.81%	77.10%	3.01

TABLE 5.3 – Ligne 1 : statistiques des taux d'inclusion en % de la reconstruction optimale dans la reconstruction SRNA. Ligne 2 : statistiques des taux d'inclusion en % de l'image originale dans la reconstruction SRNA.

peut voir que la reconstruction optimale est presque complètement incluse dans notre reconstruction SRNA. Cela veut dire que la SRNA inclut presque complètement la meilleure reconstruction que l'on peut obtenir. Il est important de noter qu'à la première itération des deux méthodes (IBP et SRNA) c.à.d. à la première rétro-projection des images bassement résolue, le taux d'inclusion est de 100% puisque pour chaque séquence, l'opérateur de rétro-projection imprécis inclut celui de la rétro-projection précise. Ce taux diminue ensuite lentement au fur des itérations.

Cependant, le taux d'inclusion est moins important lorsqu'on considère l'image originale. Nous pensons que ceci est normal car l'image originale contient des hautes fréquences impossible à restaurer correctement. Ce résultat est compensé par l'aptitude de l'imprécision de chaque pixel reconstruit par la SRNA à quantifier l'erreur de reconstruction. Pour mettre en évidence cette aptitude, nous calculons, pour chacun des $12 \times 512 \times 512 = 3145728$ pixels intervallistes reconstruits, l'écart entre ses deux bornes et la distance entre sa médiane et la valeur du pixel correspondant, dans l'image de la reconstruction optimale d'une part, et dans l'image originale d'autre part. Nous présentons graphiquement par la Figure 5.5 l'imprécision versus l'erreur de reconstruction calculée avec l'image de la reconstruction optimale (image (a)), et l'imprécision versus l'erreur de reconstruction calculée avec l'image originale. Ces graphiques montrent de fortes corrélations entre l'imprécision et les erreurs de reconstructions : le coefficient de corrélation de Pearson est de, 80% entre l'imprécision et l'erreur de reconstruction calculée avec l'image de la reconstruction optimale, et de 77% entre l'imprécision et l'erreur de reconstruction calculé avec l'image originale.

Nous avons montré précédemment un très fort taux d'inclusion des images obtenues par la reconstruction optimale dans les images intervallistes obtenues par la SRNA. Ce résultat est intéressant lorsque la valeur de l'imprécision n'est pas trop importante. En effet, si les écarts entre les bornes inférieures et supérieures d'une image intervaliste reconstruite par la SRNA sont très grands, l'imprécision n'apporte pas d'information supplémentaire significative. Nous illustrons par l'histogramme normalisé de la Figure 5.6 la distribution de l'imprécision des images intervallistes reconstruites par la SRNA. Il est donné pour 3145728 données, c.à.d. 12 séquences $\times$ 512 $\times$ 512 pixels. Cet histogramme montre que 0.44% des écarts pixel-à-pixel entre entre les bornes supérieures et inférieures sont infé-

Translations	Méthode	Seq1	Seq2	Seq3	Seq4	Seq5	Seq6	Seq7	Seq8	Seq9	Seq10	Seq11	Seq12
Valeurs vraies	IBP	34.34	33.77	33.81	34.13	33.19	34.12	32.27	34.86	33.5	32.52	31.61	31.29
	SRNA	31.91	30.92	30.92	31.92	30.23	31.92	29.94	31.9	31.57	30.23	29.14	28.8
Valeurs bruitées	IBP	29.78	29.73	29.73	29.48	29.87	29.48	27.25	27.28	27.3	27.32	27.22	27.15
	SRNA	30.21	30.28	30.27	30.77	29.17	30.77	28.98	30.71	30.01	29.18	28.42	28.15

TABLE 5.4 – PSNR des images reconstruites pour les 12 séquences.

rieurs ou égaux à 5 niveaux de gris et que 58% d'entre eux sont inférieurs ou égaux à 10 niveaux de gris.

5.3 Robustesse aux erreurs de recalage

Le but de l'expérimentation suivante est d'illustrer la robustesse de la SRNA aux erreurs de recalage. Lorsqu'on considère des séquences d'images réelles, les mouvements entre les images ne sont généralement pas connus et doivent être estimés, introduisant ainsi des erreurs qui peuvent être critiques pour les algorithmes de super-résolution.

Pour cette expérimentation, nous utilisons les mêmes 12 séquences de l'expérimentation précédente et considérons les mêmes méthodes de reconstruction : la méthode IBP avec les bonnes modélisations de la RI et la SRNA avec les mêmes paramètres. Deux cas sont traités : dans le premier, nous utilisons les bonnes valeurs des translations, c.à.d. celles qui ont servi pour générer les séquences. Dans le deuxième, ces valeurs sont altérées par un bruit Gaussien de moyenne nulle avec un écart-type de 0.3 dans l'échelle bassement résolue. Nous utilisons le PSNR pour quantifier la déviation des images reconstruites par rapport à l'image originale. Dans le cas de la SRNA, nous considérons l'image médiane pour calculer le PSNR. Les résultats sont dressés dans le tableau 5.4.

Le Tableau 5.4 montre que, en utilisant les vraies valeurs des translations, la SRNA donne de bons résultats. Le PSNR n'est jamais inférieur à 28 et est même supérieur à 30 dans 9 cas sur les 12. Évidemment, en utilisant la méthode IBP avec la bonne modélisation de la RI, on obtient les meilleurs résultats pour toutes les séquences. En fait, ce sont les meilleurs résultats que l'on puisse obtenir en utilisant cette méthode puisqu'il n'y a ni erreur de modélisation ni erreur de recalage. Cependant, lorsqu'on utilise les valeurs de translation bruitées, la dégradation du PSNR avec la SRNA est beaucoup moins importante que la dégradation du PSNR avec la méthode IBP pour les 12 séquences. De plus, la SRNA donne de meilleurs résultats dans 11 cas sur les 12.

5.4 Tests sur des séquences réelles

Dans cette dernière expérimentation, nous étudions le comportement la SRNA lorsqu'on considère des séquences d'images réelles et nous comparons ses performances, qualitativement en utilisant une première séquence d'images, et quantitativement en uti-

lisant une deuxième séquence, avec quelques autres méthodes très compétitives issues de la littérature.

Pour la première séquence, nous prenons les 12 premières images de la séquence nommée 'disk' d'une résolution de 49×57 pixels de Peyman Milanfar[1] pour reconstruire des images hautement résolues avec un facteur de rehaussement de 4×4. La Figure 5.7 montre les 6 premières images de la séquence. Nous effectuons des reconstructions en utilisant 4 méthodes : la SRNA, la méthode IBP, la méthode régularisée basée sur la norme L_1 présentée dans [Farsiu *et al.*, 2004b] et l'approche MAP présentée dans [Gunturk et Gevrekci, 2006] mais sans reconstruction haute dynamique. Pour les deux dernières, l'effet de la RI est représenté, comme le préconisent leurs auteurs, par la convolution avec un noyau Gaussien. Dans la suite, la taille de la fenêtre de ce noyau sera notée s et son écart-type, σ. Pour chaque méthode, nous varions les paramètres de la RI pour étudier sa sensibilité à ces derniers.

Nous avons expérimentalement constaté que ces méthodes itératives convergent après environ 8 itérations. Nous les avons donc arrêtées à la $8^{ième}$ itération. Les mouvements sont supposés être des pures translations, et sont estimés par un algorithme de recalage basé sur le flot optique [Bouguet, 2000].

La Figure 5.8 montre les images hautement résolues obtenues avec les différentes méthodes de reconstruction considérées. On peut noter que la méthode de [Gunturk et Gevrekci, 2006] est très sensible aux choix des paramètres du noyau de la RI. De plus, même en utilisant le noyau qui semble donner le meilleur résultat ($s = 9 \times 9$ et $\sigma = 2$), les écritures demeurent floutées et certains caractères ne sont pas distinguables. La méthode présentée dans [Farsiu *et al.*, 2004b] est moins sensible au choix des paramètres de la RI et donne de bons résultats : il n y a presque pas d'effet de Gibbs dans les images reconstruites. Cependant, les écritures sont légèrement floutées et il manque un peu de texture dans les zones uniformes. La méthode IBP donne également de bons résultats. Les écritures sont plus nettes que dans les résultats de la méthode de [Farsiu *et al.*, 2004b] mais on y voit aussi un apparaitre des artefacts de Gibbs près des contours. La SRNA semble donner le meilleur compromis entre la netteté et ces artefacts. De plus, les reconstructions de la SRNA sont très peu affectés par le choix de la taille du noyau maxitif utilisé.

Pour la deuxième séquence, nous utilisons 25 images réelles en niveaux de gris acquises d'une scène représentant une mire, qu'on nommera la séquence 'mire', avec une caméra FL2-08S2C POINT GREY ayant une résolution de 1024×768 pixels. La caméra a été fixée sur un robot contrôlé et les mouvements effectués sont des pures translations. La Figure 5.9 montre deux photos montrant les conditions d'acquisition. La RI de cette caméra n'est pas connue. La Figure 5.10 montre les 6 premières images de 160×160 pixels représentant la région d'intérêt de la scène, que nous voulons restaurer.

Nous effectuons les mêmes reconstructions que pour la séquence précédente, avec les

1. http://users.soe.ucsc.edu/~milanfar/software/sr-datasets.html

mêmes paramètres, en utilisant les 25 images représentant la région d'intérêt. Les translations ont été estimées avec la même technique également.

Les images reconstruites avec les différentes méthodes sont données par la Figure 5.11. Visuellement, presque les mêmes remarques que pour la séquence précédente peuvent être faites pour chaque méthode, à l'exception de la méthode IBP qui donne un moins bons résultats en utilisant le noyau quadratique qu'avec le noyau cubique. Ceci peut s'expliquer par le fait que le noyau cubique soit certainement plus proche du noyau de la RI de la caméra que le noyau quadratique.

Afin d'essayer d'avoir une image de référence à laquelle on peut quantitativement comparer ces résultats, nous avons effectué une acquisition hautement résolue de la même scène. Nous avons donc rapproché la caméra de la mire, avec le robot, de manière à ce que la région d'intérêt soit représentée par le nombre de pixels voulu sur l'image, c.à.d. 640×640, qui correspond à un rehaussement de résolution 4×4. L'image de la région d'intérêt de la scène est donc acquise avec une plus grande densité de pixels que précédemment, et la RI couvre une zone moins importante ce qui correspond précisément à notre modèle de super-résolution introduit dans la Section 3.1. Les conditions de cette acquisition hautement résolue sont illustrées par les photos de la Figure 5.12 et l'image hautement résolue acquise de la région d'intérêt est donnée par la Figure 5.13.

Nous utilisons le PSNR afin d'évaluer la fidélité des différentes reconstructions par rapport à l'image de référence hautement résolue. Cette mesure étant très sensible au bruit et au changement d'exposition durant l'acquisition, nous utilisons en plus le SSIM qui est un critère plus cohérent avec la perception humaine. En effet, le SSIM mesure la similarité de structure entre deux images, plutôt qu'une différence pixel-à-pixel comme le fait le PSNR. L'image de référence hautement résolue étant moins exposée que les images bassement résolues, ces mesures sont calculées après correction l'exposition de l'image de référence. Soit $\mathbf{X}$ le vecteur représentant l'image de référence et $\hat{\mathbf{X}}$ le vecteur représentant l'image dont on veut estimer la fidélité de reconstruction. Comme dans [Gunturk et Gevrekci, 2006], la différence d'exposition entre les deux images est modélisée par :

$$\hat{\mathbf{X}} = e\mathbf{X} + \mathbf{C}, \tag{5.1}$$

où e est le facteur d'exposition et $\mathbf{C}$ un terme de compensation. La correction est réalisé en résolvant l'Equation (5.1) par les moindres carrés.

Le Tableau 5.5 dresse les résultats obtenus en PSNR et en SSIM, pour les différentes méthodes de reconstruction considérées. Ce tableau montre que la méthode IBP, basée sur notre réinterprétatio, et la SRNA donnent des résultats meilleurs que ceux des méthodes proposées dans [Farsiu *et al.*, 2004b] et [Gunturk et Gevrekci, 2006], et ce en terme de PSNR et de SSIM. Nous pouvons noter que la méthode IBP donne les meilleurs résultats que toutes les autres en utilisant des noyaux cubiques (22.35 et 22.59 en PSNR et 0.87 et 0.88 en SSIM) mais des résultats moins bons en utilisant un noyau quadratique (20.39 en PSNR et 0.78 en SSIM), tandis que notre méthode donne de bons résultats et ce pour les trois

tailles considérées du noyau maxitif utilisés (21.90, 22.11 et 22.36 en PSNR et 0.86 et 0.87 en SSIM).

Enfin, nous illustrons pour une seconde fois et confirmons l'aptitude de l'imprécision de la SRNA à quantifier l'erreur de reconstruction. Une image de référence hautement résolue étant disponible, nous pouvons calculer l'erreur entre la médiane de la reconstruction avec la SRNA et l'image de référence. Les courbes de la Figure 5.14 représentent graphiquement l'imprécision versus l'erreur de reconstruction avec la SRNA pour les trois différentes tailles de support du noyau maxitif considérées. Ces graphiques montrent une forte corrélation entre l'imprécision et l'erreur de reconstruction. En effet, le coefficient de corrélation est de 84% dans les premier et deuxième cas et de 85% dans le troisième.

5.5 Conclusion

Nous avons, dans ce chapitre, présenté les résultats d'expérimentations intensives qui montrent les performances de la méthode de super-résolution imprécise que nous avons proposée.

En utilisant des séquences d'images synthétiques, nous avons commencé par illustrer l'importance de la bonne connaissance de la RI lorsqu'on considère certaines techniques existantes nécessitant une modélisation précise de celle-ci, en montrant les effets que peut avoir une mauvaise modélisation de la RI sur l'image reconstruite. Nous avons également montré les bonnes performances de notre SRNA utilisant une modélisation imprécise de la RI dans un contexte semi-aveugle.

Toujours en utilisant des séquences d'images synthétiques, nous avons expérimentalement montré qu'il existe une forte corrélation entre l'imprécision de la reconstruction de notre SRNA et l'erreur commise par la médiane de celle-ci. Cette corrélation montre que l'imprécision est capable de quantifier l'erreur de reconstruction et peut donc nous fournir

	Méthode de reconstruction					
	[Farsiu *et al.*, 2004b]			[Gunturk et Gevrekci, 2006]		
	$s = 5\times5, \sigma = 1$	$s = 7\times7, \sigma = 1.5$	$s = 9\times9, \sigma = 2$	$s = 5\times5, \sigma = 1$	$s = 7\times7, \sigma = 1.5$	$s = 9\times9, \sigma = 2$
PSNR	20.71	20.82	20.94	20.30	20.82	20.91
SSIM	0.82	0.82	0.83	0.81	0.84	0.85
	Méthode de reconstruction					
	IBP			SRNA		
	cub_2	$cub_{2.5}$	$quad_{2.5}$	$maxitif_2$	$maxitif_{2.5}$	$maxitif_3$
PSNR	22.35	22.59	20.39	21.90	22.11	22.36
SSIM	0.87	0.88	0.78	0.86	0.86	0.87

TABLE 5.5 – Valeurs du PSNR et du SSIM des reconstructions de la séquence 'mire' avec les différentes méthodes.

une mesure de confiance sur les valeurs reconstruites, ce qui peut être très utile pour des traitements ultérieurs de l'image, comme la fusion pondérées de données (images) etc.

Dans le même cadre, nous avons illustré la robustesse de notre SRNA aux erreurs de recalage. Le recalage est une étape cruciale pour la super-résolution et est généralement basé sur des techniques d'estimation de mouvements. Les résultats de ces techniques sont souvent entachés d'erreurs qui ont une incidence sur l'image hautement résolue reconstruite par la super-résolution. Nous avons expérimentalement montré que notre SRNA est moins sensible à ces erreurs que la méthode IBP et qu'elle peut même, dans certains cas, donner de meilleurs résultats, même lorsque cette dernière utilise la bonne modélisation de la RI.

Nous avons montré les performances de la SRNA lorsqu'on considère des séquences d'images réelles et nous les avons comparées avec les performances de trois autres méthodes très compétitives issues de la littérature. La première, présentée dans [Farsiu *et al.*, 2004b] et basée sur la norme l_1 et la régularisation BTV (voir Section 1.3.1), est robuste aux erreurs de modélisation du bruit et du mouvement. La deuxième, présentée dans [Gunturk et Gevrekci, 2006] est basée sur une approche MAP (voir Section 1.3.2) utilisant un modèle d'observation tenant compte de la différence d'exposition entre les images et des erreurs de quantification. La troisième méthode est la méthode IBP [Irani et Peleg, 1991] basée sur notre réinterprétation. En utilisant une première séquence d'images, tirée de la base de Peyman Milanfar [2], nous avons qualitativement montré que notre SRNA donne des résultats très satisfaisants et qu'elle offre le meilleur compromis entre la netteté et les artefacts de Gibbs.

Nous avons également effectué une série d'acquisitions bassement résolues d'une scène représentant une mire, et une acquisition hautement résolue de la même scène, conformément à notre modèle de super-résolution introduit dans la Section 3.1, afin d'avoir une image de référence et donc, d'utiliser des mesures de quantitatives de comparaison. Les résultats ont montré que notre SRNA donne des résultats très satisfaisants : meilleurs que ceux des méthodes proposées dans [Farsiu *et al.*, 2004b] et [Gunturk et Gevrekci, 2006], et qu'elle offre le meilleur compromis entre stabilité, vis-à-vis du choix des paramètres de la RI, et de fidélité de reconstruction.

La disponibilité d'une image de référence nous a permis également de confirmer la corrélations entre l'imprécision de notre SRNA et l'erreur commise par la médiane de celle-ci.

2. http ://users.soe.ucsc.edu/ milanfar/

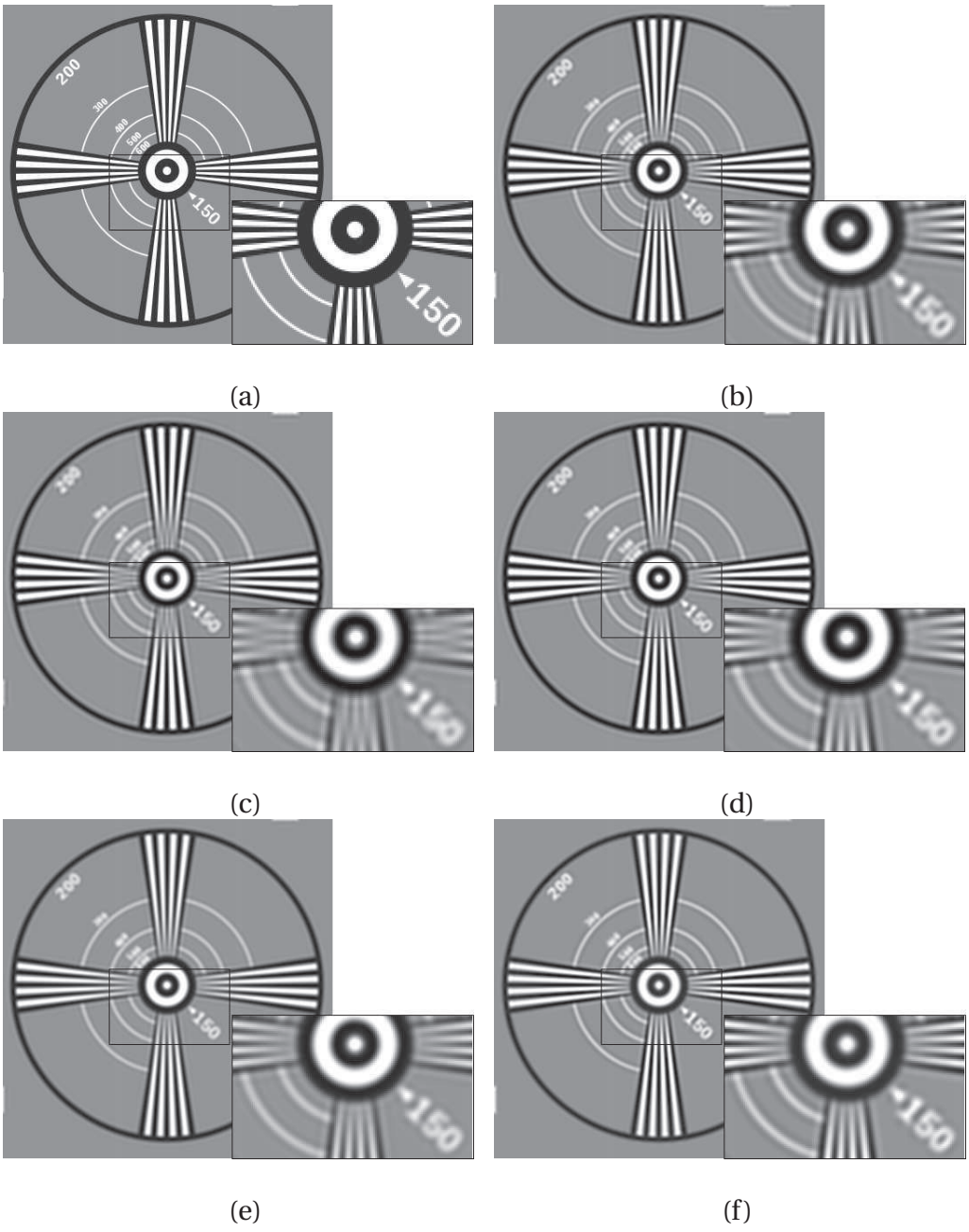


FIGURE 5.3 – (a) Image originale. Images reconstruites : (b) IBP avec un noyau quadratique d'une demi largeur de 2.5, (c) IBP avec un noyau cubique d'une demi largeur de 2.5, (d) IBP avec un noyau quadratique d'une demi largeur de 3, (e) IBP avec un noyau quadratique d'une demi largeur de 1.75, (f) médiane la reconstruction SRNA.

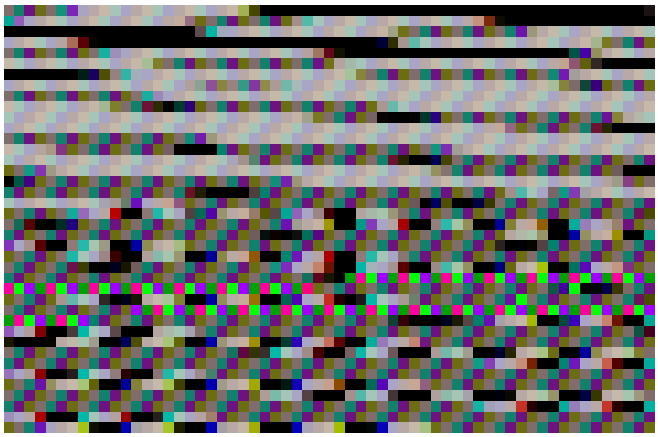
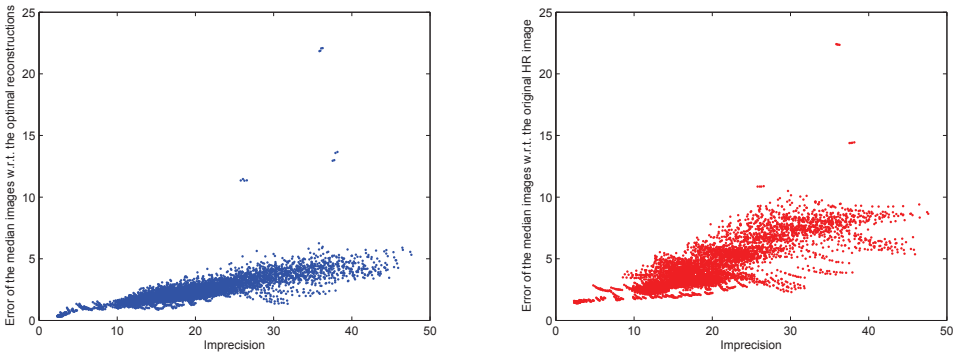


FIGURE 5.4 – Segment de la ligne 270 de l'image originale.



(a) (b)

FIGURE 5.5 – (a) Imprécision versus l'erreur entre la médiane et la reconstruction optimale, coefficient de corrélation = 80%. (b) Imprécision versus l'erreur entre la médiane et l'image originale, coefficient de corrélation = 77%.

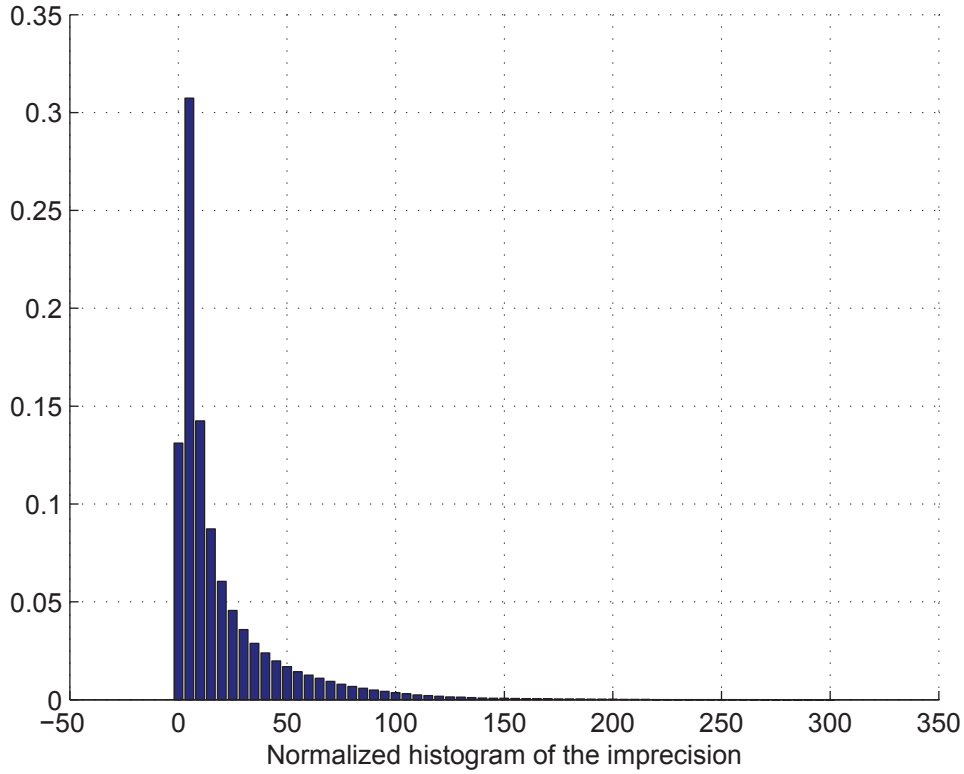


FIGURE 5.6 – Histogramme normalisé de l'imprécision des reconstructions intervallistes de la SRNA

FIGURE 5.7 – Les 6 premières images de la séquence 'disk'.

(a1) $s = 5 \times 5$, $\sigma = 1$ (a2) $s = 7 \times 7$, $\sigma = 1.5$ (a3) $s = 9 \times 9$, $\sigma = 2$

(b1) $s = 5 \times 5$, $\sigma = 1$ (b2) $s = 7 \times 7$, $\sigma = 1.5$ (b3) $s = 9 \times 9$, $\sigma = 2$

(c1) cub_2 (c2) $cub_{2.5}$ (c3) $quad_{2.5}$

(d1) $maxitif_2$ (d2) $maxitif_{2.5}$ (d3) $maxitif_3$

FIGURE 5.8 – Différentes reconstructions de la séquence 'disk'. De a1 à a3 : méthode de [Farsiu *et al.*, 2004b], de b1 à b3 : méthode de [Gunturk et Gevrekci, 2006], de c1 à c3 : IBP, de d1 à d3 : médiane de la SRNA.

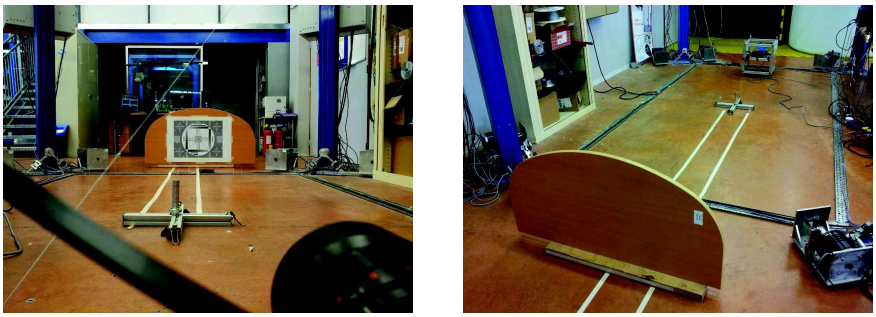


FIGURE 5.9 – Conditions des acquisitions bassement résolues de la séquence 'mire'.

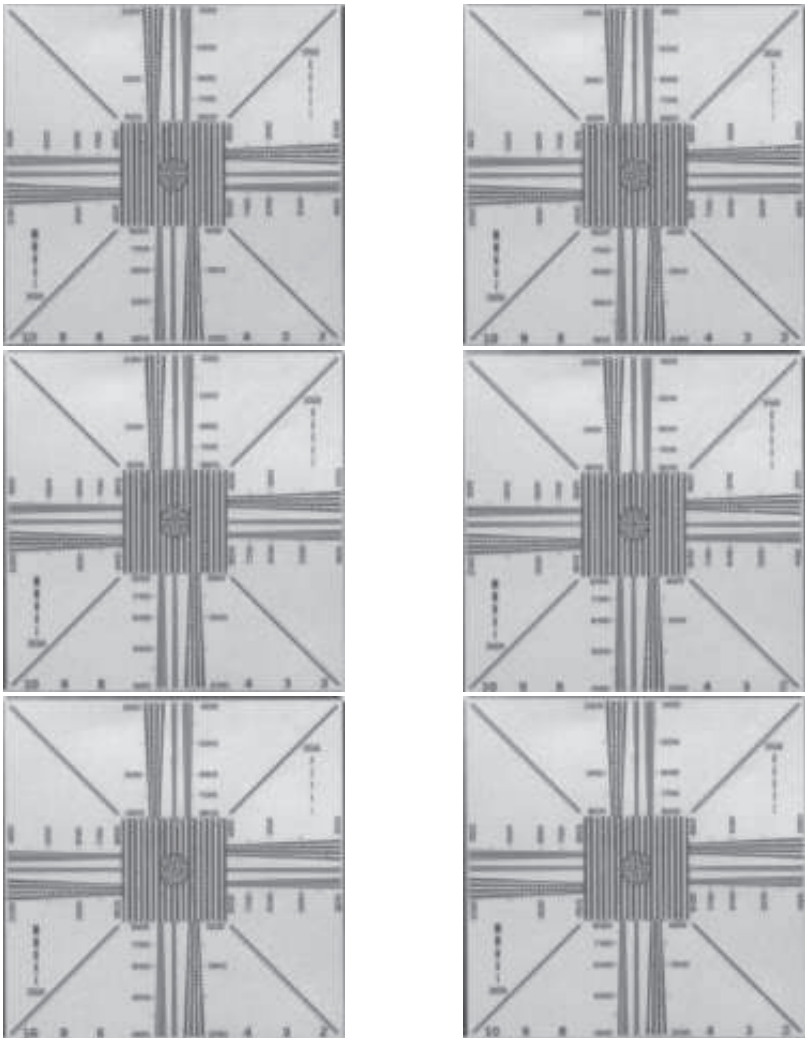


FIGURE 5.10 – Les 6 premières images de la séquence 'mire'.

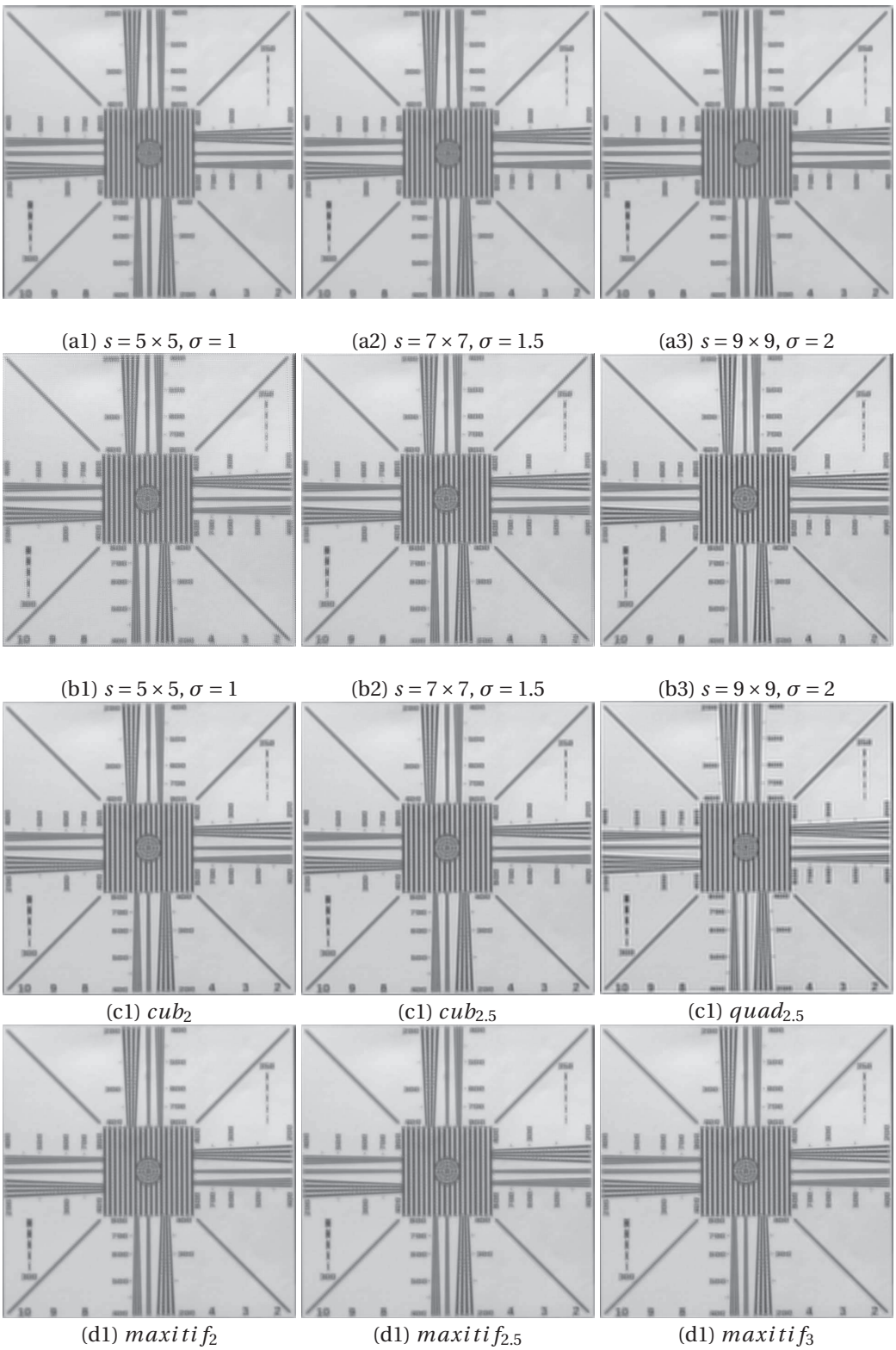


FIGURE 5.11 – Différentes reconstructions de la séquence 'mire'. De a1 à a3 : méthode de [Farsiu *et al.*, 2004b], de b1 à b3 : méthode de [Gunturk et Gevrekci, 2006], de c1 à c3 : IBP, de d1 à d3 : médiane la SRNA.

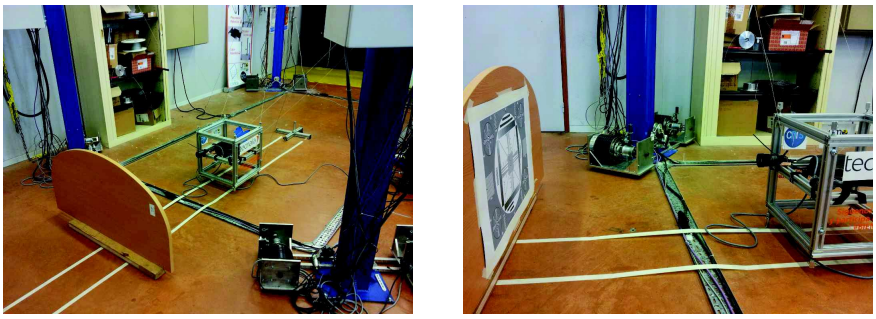


FIGURE 5.12 – Conditions de l'acquisition hautement résolue de l'image 'mire'.

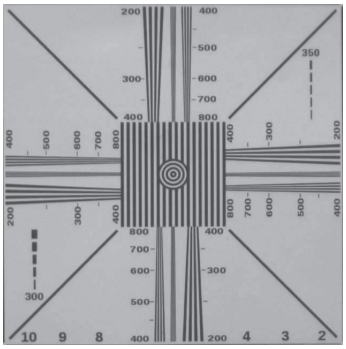
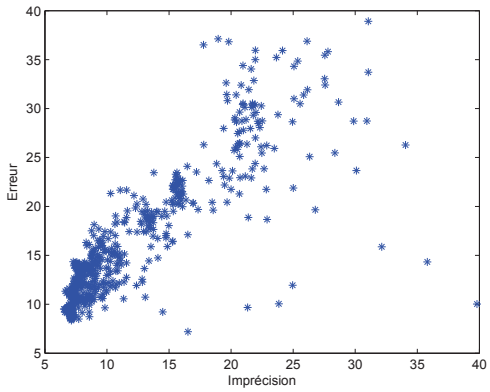
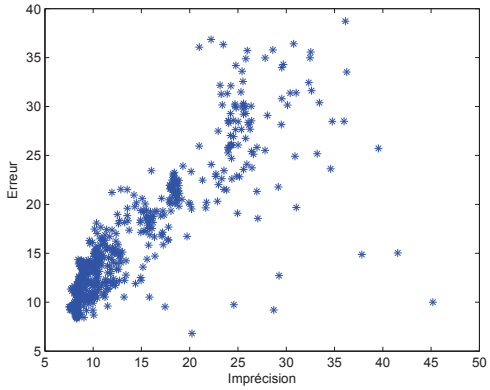


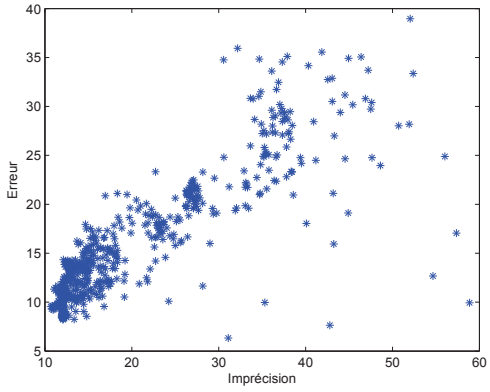
FIGURE 5.13 – Image 'mire' acquise en haute résolution.



(a)



(b)



(c)

FIGURE 5.14 – Corrélation entre l'imprécision de la SRNA et l'erreur de reconstruction entre l'image de référence et la médiane de la SRNA, pour différente taille de support du noyau maxitif employé : (a) $\Delta = 2$, (b) $\Delta = 2.5$, (c) $\Delta = 3$.

CHAPITRE

6 Conclusion générale et perspectives

Nous avons abordé, dans cette thèse, le problème de la reconstruction d'une image hautement résolue à partir de plusieurs images bassement résolues, problème connu sous le nom de super-résolution. La super-résolution est connue pour être un problème mal posé et mal conditionné. Ceci rend la qualité de l'image reconstruite par super-résolution très sensible à la modélisation de l'acquisition, notamment à la modélisation des mouvements reliant, entre elles, les images bassement résolues, mais aussi à la modélisation de la réponse impulsionnelle (RI) de l'imageur utilisé lors de l'acquisition. Notre travail s'est focalisé sur la question de la connaissance de la RI de l'imageur. Les problèmes liés à une modélisation inadéquate des mouvements n'ont pas été directement abordés dans cette thèse.

Une étude comparative des méthodes de super-résolution proposées par la littérature scientifique, que nous avons réalisée, a montré que la plupart de ces méthodes ne tiennent pas compte du fait que la RI de l'imageur est mal connue, ou bien modélisent cette méconnaissance en supposant qu'elle participe d'un phénomène aléatoire alors qu'aucun aléa n'existe une fois un modèle choisi.

Partant de ce constat, nous avons proposé de modéliser la connaissance imprécise de cette RI par une capacité de Choquet concave. Plus particulièrement, nous avons utilisé l'aptitude des noyaux maxitifs à représenter un ensemble convexe de RI pour modéliser le fait que la RI de l'imageur est connue de façon imprécise. Une telle modélisation nous a semblé plus appropriée au contexte semi-aveugle.

Modéliser la RI par une capacité concave induit une modélisation du phénomène d'acquisition un peu particulière car de nature intervalliste. L'utilisation de ce modèle pour représenter l'opérateur reliant l'image hautement résolue à l'ensemble des images bassement résolues mesurées, appelé projection, doit être modifié à cet effet. On obtient alors un nouvel opérateur de projection dont le résultat est un ensemble d'images imprécises

(intervallistes). Afin de résoudre le problème de super-résolution, on peut soit développer une méthode dédiée à l'inversion de ce modèle, soit étendre une des méthodes existantes de super-résolution à cette nouvelle modélisation. Nous avons choisi la seconde option en étendant la méthode de rétro-projection itérative. Ce choix a été motivé par le fait que la rétro-projection itérative fasse partie des méthodes les plus performantes de super-résolution, d'une part, et le fait que la rétro-projection itérative puisse être vue comme la base d'un grand nombre des techniques basées optimisation présentées dans la Section 1.3, d'autre part. En effet, les techniques d'optimisation sont généralement des améliorations, souvent par régularisation, de la rétro-projection itérative.

L'utilisation de notre modèle dans l'algorithme de rétro-projection itérative conduit à une modification de l'opérateur de rétro-projection également. On utilise alors deux nouveaux opérateurs imprécis de projection et de rétro-projection conduisant chacun à des données imprécises (intervallistes). L'image hautement résolue reconstruite par l'algorithme étendu est de nature intervalliste, reflétant la connaissance imprécise sur la RI de l'imageur.

Nous avons présenté en détail une implémentation simple de cet algorithme, basée sur des opérateurs unidimensionnels. Cette implémentation ne convient cependant qu'au cas où les images bassement résolues sont reliées entre elles par des transformations qui sont de pures translations. Pour les mouvements plus généraux, l'implémentation nécessite des calculs de l'intégrale de Choquet 2D qui sont plus complexes à réaliser. L'étude de la complexité algorithmique temporelle a montré que la complexité de notre méthode est quadratique vis-à-vis du facteur de rehaussement de la résolution et linéaire vis-à-vis des autres paramètres qui sont le nombre d'images bassement résolues utilisées, le nombre de pixels dans chacune de ces images et le nombre d'itérations. Bien que la complexité de notre méthode soit du même ordre que celle de l'algorithme classique de rétro-projection itérative, dont elle est une extension, son temps d'exécution est environ deux fois plus important. Ceci est certainement dû au traitement, à chaque itération, de deux fois plus de données (à cause de la nature intervalliste des images manipulées) et à l'utilisation de l'intégrale de Choquet discrète qui nécessite un tri des valeurs de la fonction considérée.

La nature intervalliste de l'image obtenue par notre méthode pose un problème quand il s'agit de comparer ses performances à celles d'autres méthodes existantes. Pour y remédier, et rendre la comparaison la plus objective possible, nous avons considéré l'image médiane, c.à.d. l'image minimisant la distance de Hausdorff à l'image intervalliste reconstruite. Les résultats expérimentaux ont mis en évidence plusieurs propriétés intéressantes de notre méthode : en premier lieu, sa capacité à pouvoir reconstruire correctement une image super-résolue en utilisant uniquement une connaissance imprécise sur la RI, ce qui est une forme de robustesse à la méconnaissance de cette RI ; ensuite, sa capacité à quantifier les erreurs de reconstruction ; l'inclusion de la quasi-totalité de l'image qui aurait été reconstruite en utilisant le bon modèle de RI entre les bornes de la solution imprécise obtenue ; et enfin, sa robustesse aux erreurs d'estimation de mouvements. Si l'on s'intéresse maintenant simplement à la qualité des images reconstruites, il ressort de nos expérimen-

tations que la méthode que nous proposons se range dans les plus performantes tant en terme de qualité des détails reconstruits (critères visuels) qu'en terme de distance à l'image de référence (PSNR et SSIM).

Les résultats expérimentaux montrent que l'image médiane de la reconstruction obtenue par notre méthode offre un bon compromis entre la reconstruction de détails qu'on ne voit pas dans les images bassement résolues, la netteté des contours, et l'atténuation des artefacts de reconstruction comme le rehaussement du bruit et le phénomène de Gibbs. Cependant, ce choix est arbitraire et n'exploite pas la totalité de l'information véhiculée par la variation de l'imprécision sur l'ensemble des pixels de l'image intervalliste obtenue. Une des améliorations, qui pourrait être apportée, est de développer une technique permettant de choisir, de manière renseignée, une image précise se trouvant entre les bornes inférieure et supérieure de la solution obtenue, au lieu de choisir arbitrairement l'image médiane. Cette technique pourrait être basée sur la recherche d'une image optimisant un critère défini au préalable comme par exemple un des critères utilisés pour la régularisation (voir Section 1.3). L'hypothèse que l'image recherchée soit comprise entre les bornes inférieure et supérieure de la solution obtenue par notre méthode pourrait alors réduire considérablement l'espace de recherche d'une telle image. Cette amélioration peut alors être vue comme une forme de régularisation de notre méthode qui est à la base l'extension d'une méthode classique non régularisée. Cela permettrait peut-être d'exécuter l'algorithme jusqu'à convergence au lieu de le régulariser par un arrêt précoce des itérations.

Une des prolongations possibles de notre méthode serait de l'adapter au cas où les mouvements reliant les images bassement résolues sont plus complexes que des transformations rigides (exemples : transformations affines, projectives etc), voire aux mouvements non globaux (exemples : objets en mouvement, images avec différents plans, etc). On pourrait même envisager d'utiliser des modèles de mouvements imprécis lorsque ceux-ci ne peuvent être estimés de manière précise.

Enfin, il serait intéressant d'évaluer les performances de notre méthode dans des cas d'applications réelles. Parmi les applications les plus évidentes, on peut citer la vidéosurveillance, l'imagerie satellitaire et l'imagerie médicale, ou encore des applications multimédia comme la construction de mosaïques super-résolues, de vidéos super-résolues, et le développement d'applications smartphone permettant l'acquisition rapide de photos super-résolues avec un capteur intégré bassement résolu.

Bibliographie

[Aizawa *et al.*, 1992] Aizawa, K., Komatsu, T. et Saito, T. (1992). A scheme for acquiring very high resolution images using multiple cameras. *In IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 3, pages 289–292. IEEE. Cité page 12.

[Alam *et al.*, 2000] Alam, M. S., Bognar, J. G., Hardie, R. C. et Yasuda, B. J. (2000). Infrared image registration and high-resolution reconstruction using multiple translationally shifted aliased video frames. *IEEE Transactions on Instrumentation and Measurement*, 49(5):915–923. Cité pages 9 et 10.

[Baker et Kanade, 2002] Baker, S. et Kanade, T. (2002). Limits on super-resolution and how to break them. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(9): 1167–1183. Cité pages 19, 20, 21 et 25.

[Bannore, 2009] Bannore, V. (2009). Iterative-interpolation super-resolution (iisr). *Iterative-Interpolation Super-Resolution Image Reconstruction*, pages 19–50. Cité page 12.

[Belekos *et al.*, 2010] Belekos, S. P., Galatsanos, N. P. et Katsaggelos, A. K. (2010). Maximum a posteriori video super-resolution using a new multichannel image prior. *IEEE Transactions on Image Processing*, 19(6):1451–1464. Cité page 16.

[Bertero *et al.*, 1988] Bertero, M., Poggio, T. A. et Torre, V. (1988). Ill-posed problems in early vision. *Proceedings of the IEEE*, 76(8):869–889. Cité page 2.

[Bloch et Maitre, 1994] Bloch, I. et Maitre, H. (1994). Fuzzy mathematical morphology. *Annals of Mathematics and Artificial Intelligence*, 10(1-2):55–84. Cité page 45.

[Borman et Stevenson, 1998] Borman, S. et Stevenson, R. (1998). Spatial resolution enhancement of low-resolution image sequences-a comprehensive review with directions for

future research. *Lab. Image and Signal Analysis, University of Notre Dame, Tech. Rep.* Cité page 5.

[Borman et Stevenson, 1999] Borman, S. et Stevenson, R. L. (1999). Simultaneous multi-frame map super-resolution video enhancement using spatio-temporal priors. *In International Conference on Image Processing, Proceedings.*, volume 3, pages 469–473. IEEE. Cité page 16.

[Bose et Ahuja, 2006] Bose, N. K. et Ahuja, N. A. (2006). Superresolution and noise filtering using moving least squares. *IEEE Transactions on Image Processing*, 15(8):2239–2248. Cité page 10.

[Bose *et al.*, 2001] Bose, N. K., Lertrattanapanich, S. et Koo, J. (2001). Advances in superresolution using l-curve. *In Circuits and Systems, 2001. ISCAS 2001. The 2001 IEEE International Symposium on*, volume 2, pages 433–436. IEEE. Cité page 13.

[Bouguet, 2000] Bouguet, J.-Y. (2000). Pyramidal implementation of the affine lucas kanade feature tracker description of the algorithm. *Intel Corporation, Microprocessor Research Labs*. Cité page 79.

[Brown, 1981] Brown, J. L. (1981). Multi-channel sampling of low-pass signals. *IEEE Transactions on Circuits and Systems*, 28(2):101–106. Cité pages 9 et 10.

[Brown, 1992] Brown, L. G. (1992). A survey of image registration techniques. *ACM computing surveys (CSUR)*, 24(4):325–376. Cité page 8.

[Caner *et al.*, 2003] Caner, G., Tekalp, A. M. et Heinzelman, W. (2003). Super resolution recovery for multi-camera surveillance imaging. *In International Conference on Multimedia and Expo ICME, Proceedings*, volume 1, pages I–109. IEEE. Cité page 18.

[Capel, 2004] Capel, D. (2004). *Image mosaicing and super-resolution*. Springer. Cité page 23.

[Capel et Zisserman, 1998] Capel, D. et Zisserman, A. (1998). Automated mosaicing with super-resolution zoom. *In IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Proceedings.*, pages 885–891. IEEE. Cité pages 16 et 22.

[Capel et Zisserman, 2000] Capel, D. et Zisserman, A. (2000). Super-resolution enhancement of text image sequences. *In International Conference on Pattern Recognition, Proceedings.*, volume 1, pages 600–605. IEEE. Cité page 16.

[Chakrabarti *et al.*, 2007] Chakrabarti, A., Rajagopalan, A. et Chellappa, R. (2007). Super-resolution of face images using kernel pca-based prior. *IEEE Transactions on Multimedia*, 9(4):888–892. Cité page 16.

[Chan *et al.*, 2001] Chan, T. F., Osher, S. et Shen, J. (2001). The digital tv filter and nonlinear denoising. *IEEE Transactions on Image Processing*, 10(2):231–241. Cité page 13.

[Chang *et al.*, 2004] Chang, H., Yeung, D.-Y. et Xiong, Y. (2004). Super-resolution through neighbor embedding. *In IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR, Proceedings*, volume 1, pages I–I. IEEE. Cité page 19.

[Choquet, 1953] Choquet, G. (1953). Theory of capacities. *Annales de l'Institut Fourier*, 5:131–295. Cité page 27.

[Datsenko et Elad, 2007] Datsenko, D. et Elad, M. (2007). Example-based single document image super-resolution : a global map approach with outlier rejection. *Multidimensional Systems and Signal Processing*, 18(2-3):103–121. Cité page 20.

[Dempster *et al.*, 1977] Dempster, A. P., Laird, N. M. et Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 1–38. Cité pages 8 et 16.

[Denneberg, 1994] Denneberg, D. (1994). *Non-Additive Measure and Integral*. Kluwer Academic Publishers. Cité pages 28 et 32.

[Dong *et al.*, 2009] Dong, W., Zhang, D., Shi, G. et Wu, X. (2009). Nonlocal back-projection for adaptive image enlargement. *In IEEE International Conference on Image Processing ICIP*, pages 349–352. IEEE. Cité page 12.

[Donoho, 2006] Donoho, D. L. (2006). For most large underdetermined systems of linear equations the minimal l_1-norm solution is also the sparsest solution. *Communications on pure and applied mathematics*, 59(6):797–829. Cité page 20.

[Dubois, 2006] Dubois, D. (2006). Possibility theory and statistical reasoning. *Computational Statistics and Data Analysis*, 51(1):47–69. Cité pages 33 et 58.

[Dubois et Prade, 1983] Dubois, D. et Prade, H. (1983). Ranking fuzzy numbers in the setting of possibility theory. *Information sciences*, 30(3):183–224. Cité page 29.

[Elad et Datsenko, 2009] Elad, M. et Datsenko, D. (2009). Example-based regularization deployed to super-resolution reconstruction of a single image. *The Computer Journal*, 52(1):15–30. Cité page 19.

[Elad et Feuer, 1997] Elad, M. et Feuer, A. (1997). Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images. *IEEE Transactions on Image Processing*, 6(12):1646–1658. Cité pages 6 et 18.

[Elad et Hel-Or, 2001] Elad, M. et Hel-Or, Y. (2001). A fast super-resolution reconstruction algorithm for pure translational motion and common space-invariant blur. *IEEE Transactions on Image Processing*, 10(8):1187–1193. Cité pages 9 et 10.

[Eren *et al.*, 1997] Eren, P. E., Sezan, M. I. et Tekalp, A. M. (1997). Robust, object-based high-resolution image reconstruction from low-resolution video. *IEEE Transactions on Image Processing*, 6(10):1446–1451. Cité page 18.

[Farsiu *et al.*, 2004a] Farsiu, S., Robinson, D., Elad, M. et Milanfar, P. (2004a). Advances and challenges in super-resolution. *International Journal of Imaging Systems and Technology*, 14(2):47–57. Cité page 23.

[Farsiu *et al.*, 2004b] Farsiu, S., Robinson, M. D., Elad, M. et Milanfar, P. (2004b). Fast and robust multiframe super resolution. *IEEE Transactions on Image processing*, 13(10): 1327–1344. Cité pages 14, 72, 79, 80, 81, 82, 87, 89 et 108.

[Ford et Etter, 1998] Ford, C. et Etter, D. (1998). Wavelet basis reconstruction of nonuniformly sampled data. *Circuits and Systems II : IEEE Transactions on Analog and Digital Signal Processing*, 45(8):1165–1168. Cité page 9.

[Freeman *et al.*, 2002] Freeman, W. T., Jones, T. R. et Pasztor, E. C. (2002). Example-based super-resolution. *Computer Graphics and Applications, IEEE*, 22(2):56–65. Cité pages 19 et 20.

[Freeman *et al.*, 2000] Freeman, W. T., Pasztor, E. C. et Carmichael, O. T. (2000). Learning low-level vision. *International journal of computer vision*, 40(1):25–47. Cité pages 19 et 20.

[Gardenes *et al.*, 2001] Gardenes, E., Sainz, M., Jorba, I., Calm, R., Estela, R., Mielgo, H. et Trepat, A. (2001). Modal intervals. *Reliable Computing*, 7(2):77–111. Cité page 36.

[Gilbert, 1972] Gilbert, P. (1972). Iterative methods for the three-dimensional reconstruction of an object from its projections. *J. Theor. Bio.*, 36:105 – 117. Cité page 12.

[Glasner *et al.*, 2009] Glasner, D., Bagon, S. et Irani, M. (2009). Super-resolution from a single image. *In IEEE International Conference on Computer Vision*, pages 349–356. IEEE. Cité pages 20 et 25.

[Gunturk et Gevrekci, 2006] Gunturk, B. K. et Gevrekci, M. (2006). High-resolution image reconstruction from multiple differently exposed images. *Signal Processing Letters, IEEE*, 13(4):197–200. Cité pages 16, 79, 80, 81, 82, 87, 89 et 108.

[Hansen et O'Leary, 1993] Hansen, P. C. et O'Leary, D. P. (1993). The use of the l-curve in the regularization of discrete ill-posed problems. *SIAM Journal on Scientific Computing*, 14(6):1487–1503. Cité page 13.

[Hardie *et al.*, 1997] Hardie, R. C., Barnard, K. J. et Armstrong, E. E. (1997). Joint map registration and high-resolution image estimation using a sequence of undersampled images. *IEEE Transactions on Image Processing*, 6(12):1621–1633. Cité page 16.

[Hukuhara, 1967] Hukuhara, M. (1967). Intégration des applications mesurables dont la valeur est un compact convexe. *In Funkcialaj Ekvacioj*, volume 10, pages 205–223. Cité page 35.

[Humblot et Mohammad-Djafari, 2006] Humblot, F. et Mohammad-Djafari, A. (2006). Super-resolution using hidden markov model and bayesian detection estimation framework. *EURASIP Journal on Advances in Signal Processing*, 2006. Cité page 16.

[Irani et Peleg, 1990] Irani, M. et Peleg, S. (1990). Super resolution from image sequences. *In International Conference on Pattern Recognition, Proceedings.*, volume 2, pages 115–120. IEEE. Cité pages 5, 12 et 19.

[Irani et Peleg, 1991] Irani, M. et Peleg, S. (1991). Improving resolution by image registration. *CVGIP : Graphical models and image processing*, 53(3):231–239. Cité pages 12, 40 et 82.

[Jacquey *et al.*, 2007] Jacquey, F., Loquin, K., Comby, F. et Strauss, O. (2007). Non-additive approach for gradient-based detection. *In ICIP'07*, pages 49–52, San Antonio, Texas. Cité page 48.

[K. Loquin, 2009] K. Loquin, O. S. (2009). Linear filtering and mathematical morphology on an image : a bridge. *In IEEE International Conference on Image Processing*, page 3965Ð3968, Le Caire, Egypt. Cité page 48.

[Keren *et al.*, 1988] Keren, D., Peleg, S. et Brada, R. (1988). Image sequence enhancement using sub-pixel displacements. *In Computer Society Conference on Computer Vision and Pattern Recognition CVPR, Proceedings*, pages 742–746. IEEE. Cité page 11.

[Kim *et al.*, 2004] Kim, K. I., Franz, M. et Schölkopf, B. (2004). Kernel hebbian algorithm for single-frame super-resolution. Cité page 16.

[Kim *et al.*, 1990] Kim, S., Bose, N. et Valenzuela, H. (1990). Recursive reconstruction of high resolution image from noisy undersampled multiframes. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 38(6):1013–1027. Cité pages 7 et 8.

[Kim et Su, 1993] Kim, S. et Su, W.-Y. (1993). Recursive high-resolution reconstruction of blurred multiframe images. *IEEE Transactions on Image Processing*, 2(4):534–539. Cité page 7.

[Knutsson et Westin, 1993] Knutsson, H. et Westin, C.-F. (1993). Normalized and differential convolution. *In IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR, Proceedings*, pages 515–523. IEEE. Cité page 10.

[Kong *et al.*, 2006] Kong, D., Han, M., Xu, W., Tao, H. et Gong, Y. (2006). A conditional random field model for video super-resolution. *In International Conference on Pattern Recognition, ICPR*, volume 3, pages 619–622. IEEE. Cité page 16.

[Landweber, 1951] Landweber, L. (1951). An iteration formula for fredholm integral equations of the first kind. *American journal of mathematics*, pages 615–624. Cité page 12.

[Lehmann *et al.*, 1999] Lehmann, T. M., Gonner, C. et Spitzer, K. (1999). Survey : Interpolation methods in medical image processing. *IEEE Transactions on Medical Imaging*, 18(11):1049–1075. Cité page 9.

[Lertrattanapanich et Bose, 2002] Lertrattanapanich, S. et Bose, N. K. (2002). High resolution image formation from low resolution frames using delaunay triangulation. *IEEE Transactions on Image Processing*, 11(12):1427–1441. Cité page 10.

[Li et Santosa, 1996] Li, Y. et Santosa, F. (1996). A computational algorithm for minimizing total variation in image restoration. *IEEE Transactions on Image Processing*, 5(6):987–995. Cité page 13.

[Loquin et Strauss, 2008] Loquin, K. et Strauss, O. (2008). On the granularity of summative kernels. *Fuzzy Sets and Systems*, 159(15):1952–1972. Cité pages 30, 31 et 52.

[Marr et Vision, 1982] Marr, D. et Vision, A. (1982). A computational investigation into the human representation and processing of visual information. *WH San Francisco : Freeman and Company*. Cité page 19.

[Michaeli et Irani, 2013] Michaeli, T. et Irani, M. (2013). Nonparametric blind super-resolution. *In IEEE International Conference on Computer Vision ICCV*, pages 945–952. IEEE. Cité pages 39 et 51.

[Milanfar, 2010] Milanfar, P. (2010). *Super-resolution imaging*, volume 1. CRC Press. Cité pages 5, 20 et 23.

[Mohammad-Djafari, 2009] Mohammad-Djafari, A. (2009). Super-resolution : a short review, a new method based on hidden markov modeling of hr image and future challenges. *The Computer Journal*, 52(1):126–141. Cité pages 16 et 23.

[Nguyen et Milanfar, 2000] Nguyen, N. et Milanfar, P. (2000). A wavelet-based interpolation-restoration method for superresolution (wavelet superresolution). *Circuits, Systems and Signal Processing*, 19(4):321–338. Cité page 9.

[Nguyen *et al.*, 2001a] Nguyen, N., Milanfar, P. et Golub, G. (2001a). A computationally efficient superresolution image reconstruction algorithm. *IEEE Transactions on Image Processing*, 10(4):573–583. Cité pages 6 et 13.

[Nguyen *et al.*, 2001b] Nguyen, N., Milanfar, P. et Golub, G. (2001b). Efficient generalized cross-validation with applications to parametric image restoration and resolution enhancement. *IEEE Transactions on Image Processing*, 10(9):1299–1308. Cité pages 13 et 23.

[Papoulis, 1977] Papoulis, A. (1977). Generalized sampling expansion. *IEEE Transactions on Circuits and Systems*, 24(11):652–654. Cité page 9.

[Papoulis, 1978] Papoulis, A. (1978). *Signal analysis*, volume 191. McGraw-Hill New York. Cité pages 9 et 10.

[Park *et al.*, 2003] Park, S. C., Park, M. K. et Kang, M. G. (2003). Super-resolution image reconstruction : a technical overview. *Signal Processing Magazine, IEEE*, 20(3):21–36. Cité page 5.

[Patti *et al.*, 1994] Patti, A., Sezan, M. et Tekalp, A. (1994). High-resolution image reconstruction from a low-resolution image sequence in the presence of time-varying motion blur. *In Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference*, volume 1, pages 343–347 vol.1. Cité page 18.

[Patti et Altunbasak, 2001] Patti, A. J. et Altunbasak, Y. (2001). Artifact reduction for set theoretic super resolution image reconstruction with edge adaptive constraints and higher-order interpolants. *IEEE Transactions on Image Processing*, 10(1):179–186. Cité page 18.

[Patti *et al.*, 1997a] Patti, A. J., Sezan, M. I. et Tekalp, A. M. (1997a). Robust methods for high-quality stills from interlaced video in the presence of dominant motion. *IEEE Transactions on Circuits and Systems for Video Technology*, 7(2):328–342. Cité page 18.

[Patti *et al.*, 1997b] Patti, A. J., Sezan, M. I. et Tekalp, A. M. (1997b). Superresolution video reconstruction with arbitrary sampling lattices and nonzero aperture time. *IEEE Transactions on Image Processing*, 6(8):1064–1076. Cité page 18.

[Perfilieva, 2006] Perfilieva, I. (2006). Fuzzy transforms : Theory and applications. *Fuzzy Sets and Systems*, 157(8):993–1023. Cité page 34.

[Pham *et al.*, 2006] Pham, T. Q., Van Vliet, L. J. et Schutte, K. (2006). Robust fusion of irregularly sampled data using adaptive normalized convolution. *EURASIP Journal on Advances in Signal Processing*, 2006. Cité page 10.

[Pickup *et al.*, 2006] Pickup, L. C., Capel, D. P., Roberts, S. J. et Zisserman, A. (2006). Bayesian image super-resolution, continued. *In NIPS*, volume 19. Cité page 16.

[Pickup *et al.*, 2009] Pickup, L. C., Capel, D. P., Roberts, S. J. et Zisserman, A. (2009). Bayesian methods for image super-resolution. *The Computer Journal*, 52(1):101–113. Cité page 16.

[Rhee et Kang, 1999] Rhee, S. et Kang, M. G. (1999). Discrete cosine transform based regularized high-resolution image reconstruction algorithm. *Optical Engineering*, 38(8): 1348–1356. Cité page 8.

[Rico et Strauss, 2010] Rico, A. et Strauss, O. (2010). Imprecise expectations for imprecise linear filtering. *International Journal of Approximate Reasoning*, 51(8):933–947. Cité page 31.

[Rudin *et al.*, 1992] Rudin, L. I., Osher, S. et Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D : Nonlinear Phenomena*, 60(1):259–268. Cité page 13.

[Ruspini, 1973] Ruspini, E. (1973). New experimental results in fuzzy. *Information Sciences*, 6:273Ð284. Cité page 34.

[Schmeidler, 1989] Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica*, 57(3):571–587. Cité page 32.

[Schultz et Stevenson, 1996] Schultz, R. R. et Stevenson, R. L. (1996). Extraction of high-resolution frames from video sequences. *IEEE Transactions on Image Processing*, 5(6): 996–1011. Cité page 16.

[Sezan et Stark, 1982] Sezan, M. et Stark, H. (1982). Image restoration by the method of convex projections : Part 2-applications and numerical results. *IEEE Transactions on Medical Imaging*, 1(2):95–101. Cité page 17.

[Shafer *et al.*, 1976] Shafer, G. *et al.* (1976). *A mathematical theory of evidence*, volume 1. Princeton university press Princeton. Cité page 29.

[Shah *et al.*, 1999] Shah, N. R., Zakhor, A. *et al.* (1999). Resolution enhancement of color video sequences. *IEEE Transactions on Image Processing*, 8(6):879–885. Cité page 12.

[Shen *et al.*, 2007] Shen, H., Zhang, L., Huang, B. et Li, P. (2007). A map approach for joint motion estimation, segmentation, and super resolution. *IEEE Transactions on Image Processing*, 16(2):479–490. Cité page 16.

[Stark, 1988] Stark, H. (1988). Theory of convex projection and its application to image restoration. *In Circuits and Systems, 1988., IEEE International Symposium on*, pages 963–964. IEEE. Cité page 17.

[Stark, 1990] Stark, H. (1990). Convex projections in image processing. *In Circuits and Systems, 1990., IEEE International Symposium on*, pages 2034–2036. IEEE. Cité page 17.

[Stark et Oskoui, 1989] Stark, H. et Oskoui, P. (1989). High-resolution image recovery from image-plane arrays, using convex projections. *JOSA A*, 6(11):1715–1726. Cité pages 17 et 22.

[Strauss et Rico, 2012] Strauss, O. et Rico, A. (2012). Towards interval-based non-additive deconvolution in signal processing. *Soft computing*, 16(5):809–820. Cité pages 12, 33 et 55.

[Sugeno, 1974] Sugeno, M. (1974). *Theory of fuzzy integrals and its applications*. Thèse de doctorat, Tokyo Institute of technology. Cité page 28.

[Sun *et al.*, 2003] Sun, J., Zheng, N.-N., Tao, H. et Shum, H.-Y. (2003). Image hallucination with primal sketch priors. *In IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Proceedings.*, volume 2, pages II–729. IEEE. Cité page 19.

[Suresh et Rajagopalan, 2007] Suresh, K. V. et Rajagopalan, A. N. (2007). Robust and computationally efficient superresolution algorithm. *JOSA A*, 24(4):984–992. Cité page 16.

[Takeda *et al.*, 2007] Takeda, H., Farsiu, S. et Milanfar, P. (2007). Kernel regression for image processing and reconstruction. *IEEE Transactions on Image Processing*, 16(2):349–366. Cité page 10.

[Tenenbaum *et al.*, 2000] Tenenbaum, J. B., De Silva, V. et Langford, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323. Cité page 20.

[Tian et Ma, 2010] Tian, J. et Ma, K.-K. (2010). Stochastic super-resolution image reconstruction. *Journal of Visual Communication and Image Representation*, 21(3):232–244. Cité page 16.

[Tian et Ma, 2011] Tian, J. et Ma, K.-K. (2011). A survey on super-resolution imaging. *Signal, Image and Video Processing*, 5(3):329–342. Cité page 5.

[Tikhonov et Arsenin, 1977] Tikhonov, A. N. et Arsenin, V. Y. (1977). *Solutions of Ill-posed problems*. W.H. Winston. Cité page 13.

[Tipping et Bishop, 2002] Tipping, M. E. et Bishop, C. M. (2002). Bayesian image super-resolution. *In NIPS*, volume 15, pages 1279–1286. Cité pages 16 et 23.

[Tom et Katsaggelos, 1995] Tom, B. C. et Katsaggelos, A. K. (1995). Reconstruction of a high-resolution image by simultaneous registration, restoration, and interpolation of low-resolution images. *In International Conference on Image Processing, Proceedings.*, volume 2, pages 539–542. IEEE. Cité page 16.

[Tom *et al.*, 1994] Tom, B. C., Katsaggelos, A. K. et Galatsanos, N. P. (1994). Reconstruction of a high resolution image from registration and restoration of low resolution images. *In Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference*, volume 3, pages 553–557. IEEE. Cité page 16.

[Tomasi et Manduchi, 1998] Tomasi, C. et Manduchi, R. (1998). Bilateral filtering for gray and color images. *In International Conference on Computer Vision*, pages 839–846. IEEE. Cité page 14.

[Tsai et Huang, 1984] Tsai, R. et Huang, T. S. (1984). Multiframe image restoration and registration. *Advances in computer vision and Image Processing*, 1(2):317–339. Cité pages 5 et 6.

[Unser *et al.*, 1991] Unser, M., Aldroubi, A. et Eden, M. (1991). Fast b-spline transforms for continuous image representation and interpolation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(3):277–285. Cité page 22.

[Ur et Gross, 1992] Ur, H. et Gross, D. (1992). Improved resolution from subpixel shifted pictures. *CVGIP : Graphical Models and Image Processing*, 54(2):181–186. Cité pages 8, 9 et 10.

[Wang *et al.*, 2005] Wang, Q., Tang, X. et Shum, H. (2005). Patch based blind image super resolution. *In IEEE International Conference on Computer Vision ICCV*, volume 1, pages 709–716. IEEE. Cité pages 19 et 23.

[Woods *et al.*, 2006] Woods, N. A., Galatsanos, N. P. et Katsaggelos, A. K. (2006). Stochastic methods for joint registration, restoration, and interpolation of multiple undersampled images. *IEEE Transactions on Image Processing*, 15(1):201–213. Cité page 8.

[Yang *et al.*, 2008] Yang, J., Wright, J., Huang, T. et Ma, Y. (2008). Image super-resolution as sparse representation of raw image patches. *In IEEE Conference on Computer Vision and Pattern Recognition CVPR*, pages 1–8. IEEE. Cité page 20.

[Yang *et al.*, 2010] Yang, J., Wright, J., Huang, T. S. et Ma, Y. (2010). Image super-resolution via sparse representation. *IEEE Transactions on Image Processing*, 19(11):2861–2873. Cité page 20.

[Yedidia *et al.*, 2000] Yedidia, J. S., Freeman, W. T., Weiss, Y. *et al.* (2000). Generalized belief propagation. *In NIPS*, volume 13, pages 689–695. Cité page 19.

[Youla et Webb, 1982] Youla, D. C. et Webb, H. (1982). Image restoration by the method of convex projections : Part 1-theory. *IEEE Transactions on Medical Imaging*, 1(2):81–94. Cité pages 17 et 18.

[Zadeh, 1965] Zadeh, L. (1965). Fuzzy sets. *Information and Control*, 8(3):338–353. Cité page 33.

[Zadeh, 1968] Zadeh, L. (1968). Fuzzy sets. *Journal of Mathematical Analysis and Applications*, 23:421–427. Cité page 46.

[Zadeh, 1978] Zadeh, L. (1978). Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1(1):3 – 28. Cité page 29.

[Zhang *et al.*, 2013] Zhang, K., Gao, X., Tao, D. et Li, X. (2013). Image super-resolution via non-local steering kernel regression regularization. *In IEEE International Conference on Image Processing (ICIP)*. Cité page 14.

[Zitova et Flusser, 2003] Zitova, B. et Flusser, J. (2003). Image registration methods : a survey. *Image and vision computing*, 21(11):977–1000. Cité page 8.

[Zomet et Peleg, 2000] Zomet, A. et Peleg, S. (2000). Efficient super-resolution and applications to mosaics. *In International Conference on Pattern Recognition, Proceedings.*, volume 1, pages 579–583. IEEE. Cité page 13.

Table des figures

Liste des tableaux

Abstract

Super-resolution is an image processing technique that involves reconstructing a high resolution image based on one or several low resolution images. This technique appeared in the 1980's in an attempt to artificially increase image resolution and therefore to overcome, algorithmically, the physical limits of an imager. Like many reconstruction problems in image processing, super-resolution is known as an ill-posed problem whose numerical resolution is ill-conditioned. This ill-conditioning makes high resolution image reconstruction quality very sensitive to the choice of image acquisition model, particularly to the model of the imager Point Spread Function (PSF).

In the panorama of super-resolution methods that we draw, we show that none of the methods proposed in the relevant literature allows properly modeling the fact that the imager PSF is, at best, imprecisely known. At best the deviation between model and reality is considered as being a random variable, while it is not: the bias is systematic.

We propose to model scant knowledge on the imager's PSF by a convex set of PSFs. The use of such a model challenges the classical inversion methods. We propose to adapt one of the most popular super-resolution methods, known under the name of "iterative back-projection", to this imprecise representation. The super-resolved image reconstructed by the proposed method is interval-valued, i.e. the value associated to each pixel is a real interval. This reconstruction turns out to be robust to the PSF model and to some other errors. It also turns out that the width of the obtained intervals quantifies the reconstruction error.

Keywords: *Super-resolution, maxitive kernels, Choquet integral, non-additive measures, imprecise expectation, intervals and imprecision.*

Résumé

La super-résolution est une technique de traitement d'images qui consiste en la reconstruction d'une image hautement résolue à partir d'une ou plusieurs images bassement résolues. Cette technique est apparue dans les années 1980 pour tenter d'augmenter artificiellement la résolution des images et donc de pallier, de façon algorithmique, les limites physiques des capteurs d'images. Comme beaucoup des techniques de reconstruction en traitement d'images, la super-résolution est connue pour être un problème mal posé dont la résolution numérique est mal conditionnée. Ce mauvais conditionnement rend la qualité des images hautement résolues reconstruites très sensible au choix du modèle d'acquisition des images, et particulièrement à la modélisation de la réponse impulsionnelle de l'imageur.

Dans le panorama des méthodes de super-résolution que nous dressons, nous montrons qu'aucune des méthodes proposées par la littérature ne permet de modéliser proprement le fait que la réponse impulsionnelle d'un imageur est, au mieux, connue de façon imprécise. Au mieux l'écart existant entre modèle et réalité est modélisé par une variable aléatoire, alors que ce biais est systématique.

Nous proposons de modéliser l'imprécision de la connaissance de la réponse impulsionnelle par un ensemble convexe de réponses impulsionnelles. L'utilisation d'un tel modèle remet en question les techniques de résolution. Nous proposons d'adapter une des techniques classiques les plus populaires, connue sous le nom de rétro-projection itérative, à cette représentation imprécise. L'image super-résolue reconstruite est de nature intervalliste, c'est à dire que la valeur associée à chaque pixel est un intervalle réel. Cette reconstruction s'avère robuste à la modélisation de la réponse impulsionnelle ainsi qu'à d'autres défauts. Il s'avère aussi que la largeur des intervalles obtenus permet de quantifier l'erreur de reconstruction.

Mots clefs : *Super-résolution, noyaux maxitifs, intégrale de Choquet, mesures non additives, espérance imprécise, intervalles et imprécision.*

LIRMM — 860 rue de St Priest — 34095 Montpellier cedex 5 — France