

# Prise en compte des discontinuités dans un algorithme de super-résolution

A. Lorette, H. Shekarforoush et J. Zerubia

INRIA

2004 route des Lucioles - BP93  
06902 Sophia-Antipolis Cedex, France

## RÉSUMÉ

Dans la plupart des problèmes de reconstruction utilisant une régularisation classique, les discontinuités sont lourdement pénalisées. C'est notamment le cas pour les modèles de reconstruction d'une image haute résolution à partir d'une séquence d'images basse résolution présentant un décalage sous-pixélique. Notre but est ici de tenir compte des discontinuités dans les images afin de remédier à l'aspect contradictoire de la super-résolution qui vise à retrouver des hautes fréquences tout en régularisant. Nous nous sommes placés dans un cadre probabiliste et les images ont été modélisées par un champ de Markov. Lors de notre étude, nous avons comparé différents  $\phi$ -modèles a priori eu égard à leur influence sur la préservation des discontinuités.

## ABSTRACT

Multi-channel super-resolution is a means of recovering high frequency information by trading off the temporal bandwidth. Almost all the methods proposed in the literature are based on optimizing a cost function. But since the problem is usually ill-posed, one needs to impose some regularity constraints. However, regularity constraints tend to attenuate high frequency contents of data (usually present in the form of discontinuities). This inherent contradiction between regularization and super-resolution has not been addressed in the literature, despite the availability of off the shelf tools. In this paper, we have investigated this issue in the context of adaptive regularization, using  $\phi$ -functions.

## 1 Introduction

Plusieurs modèles plus ou moins complets ont été développés afin de reconstruire une image à une haute résolution à partir d'une série d'images basse résolution présentant un décalage sous-pixélique [5][8][9]. Le problème consiste alors à optimiser une fonction de coût sous contraintes de régularité (pour plus de références se reporter à [6]). Cependant, le modèle a priori est très souvent le plus classique à savoir la régularisation d'Arsenin-Tikhonov qui tend à surlisser les discontinuités. Les deux aspects antagonistes du problème apparaissent alors. En effet, d'un côté la super-résolution a pour but de récupérer les hautes fréquences perdues à cause du phénomène de sous échantillonnage et de flou, d'un autre côté la régularisation tend à lisser les discontinuités et donc à perdre ces hautes fréquences.

Récemment, H. Shekarforoush et al. [1][7] ont proposé un modèle pour les données qui diffère des précédents dans la mesure où il s'agit d'un modèle de reconstruction 3D tenant compte du bruit et du flou. Cependant, le problème des discontinuités n'a pas été considéré. Notre travail a donc consisté à améliorer la qualité de l'image haute résolution en modifiant le modèle a priori du travail précédemment cité.

## 2 Modèle des données

Le modèle de prise de vue considéré est identique à celui décrit dans [7] :

$$I(k, l) = \sum_{(i,j) \in W_{k,l}} g(i, j) R(i, j) H(i - i_0, j - j_0) \quad (1)$$

$I(k, l)$  est l'intensité de l'image au niveau du pixel  $(k, l)$ .

$g$  est l'albédo et  $R$  la carte de réflectance.

$W_{k,l}$  est le support du masque de convolution.

$(i_0, j_0)$  est le décalage de la convolution discrète.

Les images observées ont été modélisées à l'aide d'un champ de Markov. Nous noterons  $G$  et  $Z$  respectivement les champs aléatoires représentant l'albédo et l'altitude de la scène.  $G$  et  $Z$  sont supposés indépendants. Nous noterons  $I_1, \dots, I_m$  les  $m$  images basse résolution observées. Elles sont supposées être détériorées par un bruit blanc gaussien de variance  $\sigma_e$ , par  $H$  la réponse impulsionnelle de la caméra qui agit comme un filtre passe-bas et par le phénomène de sous-échantillonnage.

Le problème consiste à trouver le meilleur estimateur de  $G$  et  $Z$  étant donné  $I_1, \dots, I_m$ .

Nous nous sommes placés dans un cadre bayésien et avons utilisé le critère du MAP. Ainsi, le problème consiste à maximiser la probabilité suivante :

$$P(G, Z | I_1, \dots, I_m) = P(I_1, \dots, I_m | G, Z) * P(G, Z) \quad (2)$$

$P(I_1, \dots, I_m | G, Z)$  est caractéristique des données.

$P(G, Z)$  caractérise l'a priori.

Comme  $I_1, \dots, I_m$  sont supposés indépendants :

$$P(I_1, \dots, I_m | G, Z) = \prod_{r=1}^m \frac{1}{\sigma_e^2} \exp \left\{ -\frac{\|I_r - \hat{I}_r\|^2}{2\sigma_e^2} \right\} \quad (3)$$

$\hat{I}_r$  représentent les images basse résolution estimées et obtenues en simulant le processus de formation des images donné par l'équation 1.

Notons que l'équation 1 est linéaire par rapport à  $g$  mais non linéaire par rapport à  $z$  qui apparait dans le terme  $R(i, j)$  [1]. Par la suite et pour simplifier le problème, nous supposons  $z$  connu afin de rester dans un cadre linéaire.

La fonction d'énergie à minimiser est la suivante :

$$\begin{aligned}
 E = & \sum_{i,j} \left( \sum_{r=1}^m \sum_{(k,l) \in \nu_r(i,j)} \frac{(\hat{I}^r(k,l) - I^r(k,l))^2}{2\sigma_e^2} \right. \\
 & + \sum_{(i',j') \in V(i,j)} \frac{\phi\left(\frac{g(i,j) - g(i',j')}{\Delta_g}\right)}{2\sigma_g^2} \\
 & \left. + \sum_{(i',j') \in V(i,j)} \frac{\phi\left(\frac{z(i,j) - z(i',j')}{\Delta_z}\right)}{2\sigma_z^2} \right) \quad (4)
 \end{aligned}$$

$\nu_r(i, j)$  représente l'ensemble des pixels de l'image basse résolution dont l'intensité dépend du pixel  $(i,j)$  de l'image haute résolution.

$V(i, j)$  est le voisinage du pixel  $(i,j)$  (nous avons utilisé un champ de Markov d'ordre 1 en ne tenant compte que des cliques de 2 pixels).

$g$  est l'albédo et  $z$  l'altitude de la scène.

$\Delta_g$  et  $\Delta_z$  sont des paramètres d'échelle qui caractérisent la hauteur de discontinuité minimale que l'on souhaite préserver.

### 3 Modèle a priori

La résolution multi-canal a pour but essentiel de retrouver les hautes fréquences perdues lors de la prise de vue à cause du sous-échantillonnage et de la réponse impulsionnelle de la caméra qui agit comme un filtre passe-bas. Les modèles développés jusqu'à présent utilisaient un a priori classique, purement quadratique. C'est pourquoi nous avons introduit la notion de régularisation adaptative dans le modèle de super-résolution.

Notre but étant de préserver les discontinuités, notre modèle a priori ne sera pas le modèle gaussien classique. Nous avons utilisé les deux fonctions de potentiel données ci-dessous (équations (5)(6)), proposées respectivement par P. Charbonnier [2], Hebert et Leahy [4]. Pour le terme a priori, nous nous sommes limités à une approximation d'ordre un, c'est à dire à une différence entre l'intensité de deux pixels voisins comme le montrent les deux derniers termes de l'équation (4). Il est toujours possible d'augmenter cet ordre en utilisant la différence des composantes du vecteur gradient ou bien des éléments de la matrice hessienne, ceci au prix d'une complication de la fonction d'énergie à minimiser.

$$\phi(u) = 2 * \sqrt{(1 + u^2)} - 2 \quad (5)$$

$$\phi(u) = \log(1 + u^2) \quad (6)$$

Les  $\phi$ -fonctions (cf. équations (5) et (6)) ont pour caractéristiques d'être quadratiques autour de zéro (ceci est valable pour les fonction étudiées mais n'est pas une condition nécessaire), paires et croissantes (cf. Figure 1).

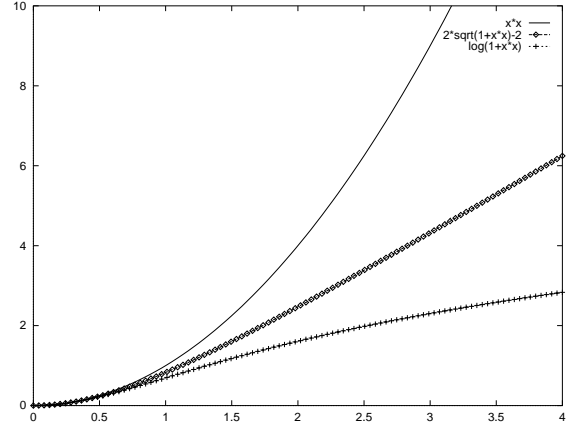


FIG. 1 — Diverses  $\phi$ -fonctions

## 4 Résultats

### 4.1 Conditions de réalisation des tests

Nous avons donc introduit successivement ces deux modèles a priori. Pour les tests, la réponse impulsionnelle de la caméra est supposée connue (filtre gaussien). Nous disposons de 4 images basse résolution synthétiques représentant une succession d'échelons ainsi que de 4 images basse résolution du Grand Palais (Paris) bruitées à 15 dB. Pour faciliter la phase de minimisation, nous avons supposé que nous connaissions l'altitude de la scène et nous nous sommes donc intéressés à la reconstruction de l'albédo afin de rester dans un cadre de traitement linéaire.

L'algorithme en lui même est relativement simple. Partant d'une image haute résolution initiale, l'algorithme simule le processus de formation des images afin d'obtenir une série d'images basse résolution estimées. La fonction d'énergie est alors calculée et minimisée. Le traitement des deux modèles diffère par la phase de minimisation de l'énergie.

Pour la fonction (5) nous avons utilisé l'algorithme déterministe de Newton, pour la fonction (6) l'algorithme du recuit simulé.

Dans les deux cas, l'algorithme a été initialisé avec une image aléatoire. Dans le premier cas, nous minimisons une fonction convexe avec un algorithme déterministe et dans le second, une fonction non convexe avec un algorithme stochastique. Les différents paramètres du problème ont été choisis de manière empirique.

### 4.2 Analyse des résultats

De manière générale, nous remarquons une amélioration des résultats, plus ou moins nette, quelque soit le modèle a priori choisi par rapport à une simple régularisation de type Arsenin-Tikhonov (Fig. 2 et 5). Cette remarque vaut aussi bien dans le cas des images synthétiques que dans le cas des images réelles.

Toutefois, les différences entre les 2 modèles sont plus ou moins marquées suivant le type d'images que l'on traite. Nous pouvons légitimement penser que le modèle non convexe

apportera les meilleurs résultats car c'est celui qui pénalise le moins les forts gradients.

En analysant les résultats, nous avons remarqué une très nette amélioration des résultats dans le cas des images synthétiques en utilisant le modèle (6) par rapport au modèle convexe (5). Ce résultat est visible sur les images haute résolution reconstruites : la figure 4 apparaît plus nette au niveau des discontinuités que la figure 3. Toutefois, cette remarque n'est plus valable dans le cas des images réelles, cas pour lequel les deux modèles donnent des résultats plus proches. La fonction non convexe offre cependant une légère amélioration. Cette dernière est visible en comparant les images d'erreur des figures 6 et 7 ainsi que les rapports S/B (36.9 dB->37.4 dB). Dans le second cas, les zones homogènes sont plus lisses et les erreurs sur les discontinuités moins marquées.

Les images synthétiques sont des images constantes par morceaux. Donc, une approximation du premier ordre est bien adaptée, ce que confirment les résultats. Cependant, en ce qui concerne les images réelles l'apport du modèle non convexe n'est pas aussi marqué. Dans ce cas, il aurait fallu augmenter l'ordre ce qui pratiquement n'est pas "rentable" compte tenu du temps de calcul. Le meilleur résultat d'un point de vue qualitatif est obtenu à l'aide de la  $\phi$ -fonction non convexe associée à l'algorithme du recuit simulé. Mais le temps de calcul est alors accru. Dans le cas des images réelles, le meilleur compromis entre temps de calcul et qualité des résultats est sans nul doute obtenu à l'aide de fonctions convexes. Cependant, il faut souligner le fait que dans le cas de l'algorithme du recuit simulé une parallélisation est toujours possible.

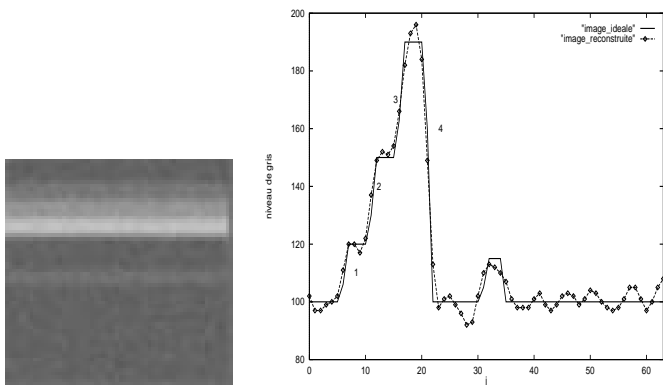


FIG. 2 — image reconstruite et un profil de l'image (S/B=44.7 dB)

## 5 Conclusion

Notre travail a consisté en une étude comparative de différents  $\phi$ -modèles pour la régularisation adaptative d'un modèle de super-résolution multi-canal ceci en vue de conserver les hautes fréquences liées à la présence de discontinuités et perdues lors de la prise de vue par des phénomènes de recouvrement de spectre liés au sous-échantillonnage. En effet, jusqu'à présent on faisait de la super-résolution tout en régularisant à

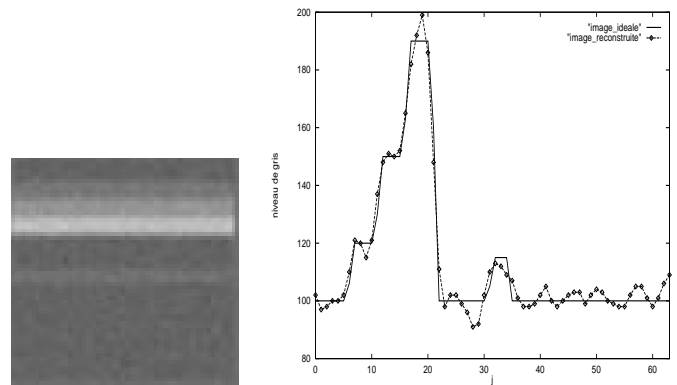


FIG. 3 — image reconstruite et un profil de l'image (S/B=46.3 dB)

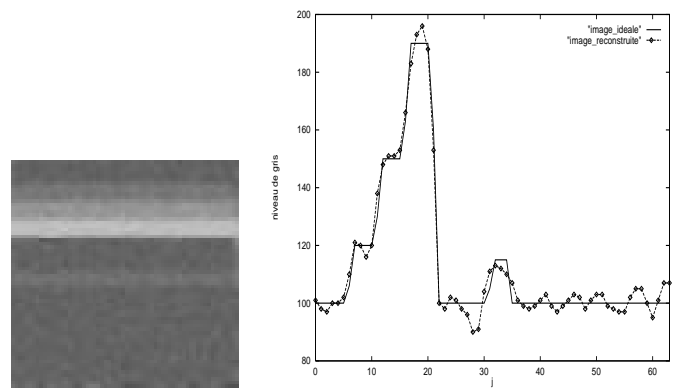


FIG. 4 — image reconstruite et un profil de l'image (S/B=49 dB)

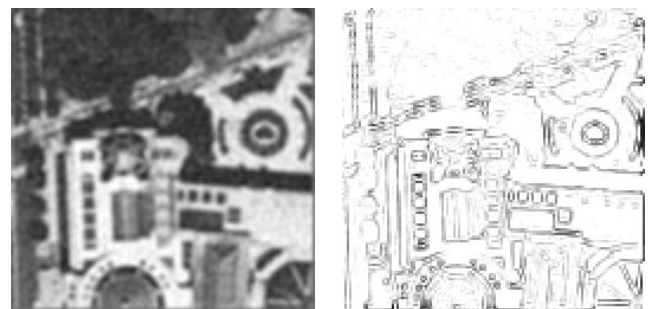


FIG. 5 — images reconstruite et d'erreur (S/B=34.5 dB)



FIG. 6 — images reconstruite et d'erreur (S/B=36.9 dB)

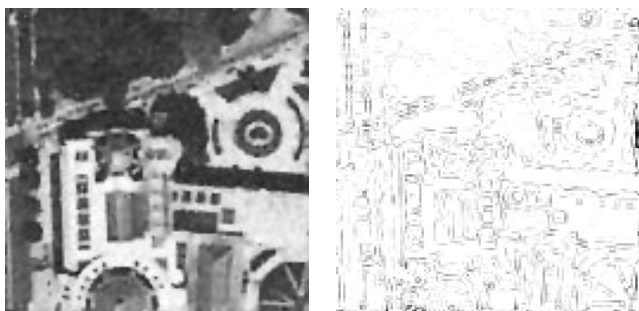


FIG. 7 — image reconstruites et d'erreur (S/B=37.4 dB)

l'aide d'une fonction purement quadratique, ce qui est contradictoire.

Si dans le cas d'images constantes par morceaux le modèle non convexe donne de bien meilleurs résultats, ceci est à nuancer dans le cas d'images réelles. Pour des questions algorithmiques nous nous sommes limités à une approximation à l'ordre 1, qui n'est pas la plus adaptée au cas d'images réelles.

Dans la suite de l'étude, nous nous intéresserons à un troisième modèle correspondant à une  $\phi$ -fonction non convexe et admettant une asymptote à l'infini [3].

## Références

- [1] M. Berthod, H. Shekarforoush, M. Werman, and J. Zerubia. Reconstruction of high resolution 3d visual information. In *CVPR*, pages 654–657, Seattle, June 1994.
- [2] P. Charbonnier. Reconstruction d'image : régularisation avec prise en compte des discontinuités. Thèse de doctorat, UNSA, Septembre 1994.
- [3] D. Geman and G. Reynolds. Constrained restoration and the recovery of discontinuities. *IEEE Trans.*, PAMI-14(3) :367–383, March 1992.
- [4] T. Hebert and R. Leahy. A generalized em algorithm for 3d bayesian reconstruction from poisson data using gibbs priors. *IEEE Trans. Med. Imaging*, MI-8 :194–202, June 1989.
- [5] M. Irani and S. Peleg. Improving resolution by image registration. *CVGIP*, 53(3) :231–239, May 1991.
- [6] H. Shekarforoush. Super-résolution en vision par ordinateur, Thèse de doctorat, UNSA, Novembre 1996.
- [7] H. Shekarforoush, M. Berthod, J. Zerubia, and M. Werman. Sub-pixel bayesian estimation of albedo and height. *IJCV*, 3(19) :289–300, 1996.
- [8] R. Y. Tsai and T. S. Huang. Multiframe image restoration and registration. *ACVIP*, 1 :317–339, 1984.
- [9] H. Ur and D. Gross. Improved resolution from subpixel shifted pictures. *CVGIP*, 54(2) :181–186, March 1992.