

# Etude Comparative d'Algorithmes de Super-Résolution

Antoine LÉTIENNE<sup>1</sup>, Frédéric CHAMPAGNAT<sup>1</sup>, Guy LE BESNERAIS<sup>1</sup>, Caroline KULCSÁR<sup>2</sup>,

<sup>1</sup>Office National d'Etudes et de Recherches Aérospatiales (ONERA), DTIM/EVS  
BP-72, 92322 Chatillon Cedex, France

<sup>2</sup>Laboratoire de Traitement et Transport de l'information, Université Paris 13  
99 av. J.-B. Clément 93430, Villetaneuse, France

aletien@onera.fr, fchamp@onera.fr, lebesner@onera.fr  
kulcsar@l2ti.univ-paris13.fr

**Résumé** – Cet article propose une étude expérimentale comparative entre des algorithmes de super-résolution (SR) basés sur une reconstruction régularisée (RR) et sur une interpolation non uniforme (INU) dans le cas plus général que la translation. Les performances de méthodes à bas coût sont évaluées dans le but de quantifier la perte de qualité des reconstructions SR. Un classement entre les méthodes comparées est établi à partir de résultats expérimentaux réalisés sur des données de synthèse.

**Abstract** – We present a super-resolution reconstruction comparison between reconstruction based and non-uniform interpolation based SR algorithms in the non-translational motion case. We evaluate the performance of low cost methods in order to quantify the loss of reconstruction quality. The experimental results obtained on synthetic data provide a ranking of tested methods.

## 1 Introduction

Les techniques de super-résolution (SR) visent à améliorer la qualité d'une image en exploitant les variations entre les images d'une même scène dues au mouvement et au repliement spatial. Ces techniques intéressent les applications dans lesquelles la résolution nominale des images est faible relativement à l'utilisation que l'on veut en faire, comme le traitement de séquences vidéo en vidéo-surveillance, ou l'observation aérienne à l'aide de capteurs embarqués à bas-coût.

Dans ce papier, nous nous intéressons à l'évaluation d'algorithmes de SR à complexité réduite capables de gérer des mouvements plus généraux que la simple translation. Le recalage, paramétrique, est effectué préalablement à la reconstruction SR. Nous évaluons plusieurs de ces approches sous optimaux dans le but de quantifier la perte de qualité image.

Notre étude se concentre sur deux familles majeures, présentées dans [1] : les méthodes basées sur la reconstruction régularisée (RR) et celles sur l'interpolation non uniforme (INU). La RR, représentée sur la branche gauche de la figure 2, consiste en la construction puis en l'inversion d'un grand système linéaire. Quand le modèle de mouvement n'est pas une translation, cette inversion ne peut être effectuée que par des techniques itératives et coûteuses [2][3]. Dans le cas de la translation, cette opération peut être réalisée en utilisant des algorithmes rapides [1][4]. La branche droite de la figure 2 présente les autres approches possibles basées sur une INU suivie d'une étape de restauration d'image [5] [6] [7].

Les résultats présentés en figure 1 sont obtenus à partir de quatre de ces méthodes, présentées dans la section 2, en utili-



(a) Image BR



(b) Hardie & al. (12.33 sec) (c) TVSS (0.84 sec)



(d) Takeda & al. (6.9 sec) (e) SAWT (0.27 sec)

FIG. 1 – Reconstruction SR pour vingt images, un bruit de variance  $\sigma = 5$  et un facteur SR de trois.

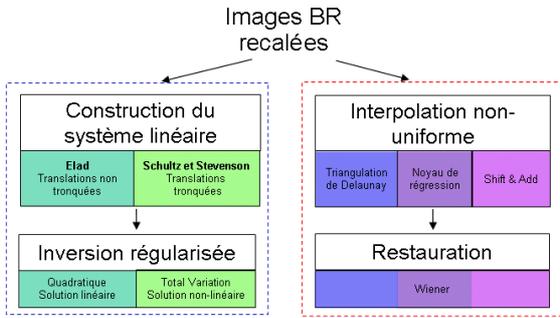


FIG. 2 – Méthodes comparées.

sant vingt images basse-résolution (BR) de taille  $43 \times 40$  extraites des données « CAR »<sup>1</sup> et un facteur de SR égal à trois. Des différences entre les reconstructions sont visibles, particulièrement sur le contraste et la netteté des images. Cependant, notons que Hardie & al. et SAWT sont visuellement très proches pour une grande différence du temps de calcul (d'un facteur 45). Cet exemple conduit à la question : « faut-il investir dans une méthode coûteuse en SR ? » L'étude de performance qui suit a pour but d'apporter des éléments de réponse à cette question.

Nous présentons les principes des méthodes comparées dans la section 2. La section 3 décrit notre méthode de génération de données synthétiques. Finalement, dans la section 4, une évaluation de performance de ces méthodes est proposée sur données synthétiques.

## 2 Méthodes comparées

Nous supposons que le recalage (paramétrique ou dense) entre les images BR est effectué avant la reconstruction SR, et que nous disposons de  $K$  images BR recalées, notées  $\mathbf{f} = \{\mathbf{f}^{(1)}, \mathbf{f}^{(2)}, \dots, \mathbf{f}^{(K)}\}$ . L'image SR reconstruite est notée  $\hat{\mathbf{x}}$ .

### 2.1 Reconstruction régularisée

La première étape de la RR consiste en la construction d'un grand système linéaire  $\mathbf{f} = \mathbf{A}\mathbf{x}$  qui modélise la formation des images BR observées, par l'application de mouvements, d'un flou optique et d'un échantillonnage à image SR. Comme montré dans [2], deux classes principales de modèles de formation d'images sont disponibles dans la littérature, la première peut être associée à l'article de Elad & Feuer (E&F) [8], la seconde à Schultz & Stevenson (S&S) [2]. Le modèle S&S est moins précis et suppose essentiellement que les opérateurs de mouvement et de dégradation de l'optique de mesure peuvent être permutés, réduisant ainsi le nombre d'opérations pour construire  $\mathbf{A}$ . Une estimation de  $\hat{\mathbf{x}}$  peut ensuite être obtenue par minimisation du critère des moindres carrés régularisé :

$$\mathcal{J}(\mathbf{x}) = \|\mathbf{f} - \mathbf{A}\mathbf{x}\|^2 + \lambda\mathcal{R}(\mathbf{x}). \quad (1)$$

<sup>1</sup>disponibles sur <http://www.ee.ucsc.edu/~milanfar/software/sr-datasets.html>

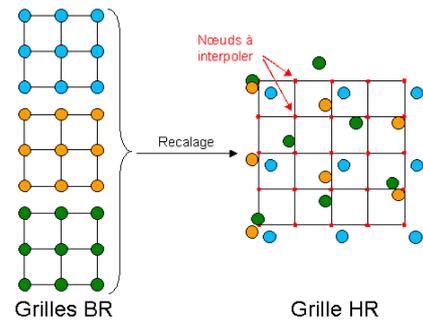


FIG. 3 – Interpolation non uniforme.

Nous avons sélectionné deux termes de régularisations différents. Le premier est quadratique et conduit donc à une résolution linéaire :  $\mathcal{R}$  est la somme des carrés des différences finies du second ordre de l'image  $\mathbf{x}$  calculés sur les 8 plus proches voisins. Le second est un terme non quadratique de type variation totale (VT). Cette approche préserve les forts gradients mais lisse fortement les zones homogènes (« effet cartoon » visible sur la figure 1b).

Afin de restreindre le nombre de méthodes testées et le temps de calcul, nous avons associé le modèle de formation d'images le plus coûteux E&F [8] avec la régularisation quadratique. L'approche est alors équivalente à celle décrite par Hardie *et al.* [3], et sera donc nommée « Hardie & al. ». Le critère (1) est minimisé par une descente de gradient conjugué. Cette méthode sert de référence à l'ensemble de nos tests. La seconde méthode est nommée « Total Variation Schultz & Stevenson » (TVSS). Le critère (1) est minimisé par une descente de gradient à pas fixe, comme proposé dans [4].

### 2.2 Interpolation non uniforme

Après l'étape de recalage entre les images BR sur une image de référence commune, une grille haute-résolution (HR) irrégulière de points est obtenue, comme illustré sur la figure 3. Une INU permet alors d'obtenir une grille régulièrement échantillonnée. Le bruit et les dégradations du capteur sont traités dans une étape ultérieure de restauration, usuellement un simple filtrage de Wiener.

Deux approches basées sur l'INU ont été retenues. La première, nommée SAWT (Shift & Add plus Wiener avec mouvements Tronqués) dans la suite, a été proposée dans [6] et est une généralisation du « Shift & Add » [4] aux mouvements non restreints à la translation. Après le recalage, chaque pixel BR est associé au noeud de la grille HR le plus proche. Lorsque plusieurs pixels BR sont associés au même pixel HR, on calcule la moyenne. Les pixels HR non-remplis sont comblés par interpolation linéaire.

La seconde méthode INU utilise des techniques d'interpolation plus récentes utilisant des noyaux de régression [5][9]. A chaque pixel HR, un voisinage spatial est associé, contenant un nombre variable de pixels BR. Un modèle polynômial plaqué sur ces points fournit une valeur estimée à la position du

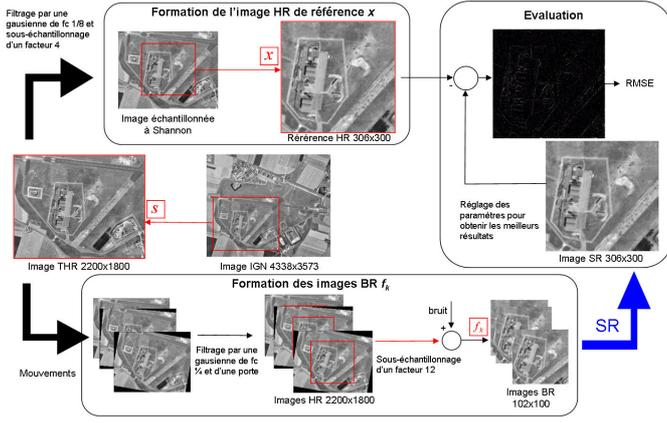


FIG. 4 – Génération des données simulées.

pixel HR. Dans une seconde passe, le voisinage spatial est déformé selon le gradient image et la densité de pixels BR. Cette interpolation est plus précise dans les régions texturées et plus robuste au bruit. Nous utilisons le code fourni par Takeda basé sur [5], et la méthode est nommée « Takeda & al. ».

### 3 Méthodologie de comparaison

Nous présentons une méthode originale de génération de données de synthèse permettant la comparaison des méthodes. Une image HR de référence  $x$  et des images BR observées  $f$  sont générées en deux étapes indépendantes à partir d'une image très haute résolution (THR)  $s$  modélisant la scène continue. Le procédé est décrit sur la figure 4. L'image HR  $x$  est obtenue par application sur  $s$  d'un filtre gaussien passe-bas de fréquence de coupure  $1/8$  pour éviter les effets d'aliasing, puis par sous-échantillonnage d'un facteur quatre. Cette image est quatre fois moins résolue que  $s$  et trois fois plus que les images BR. Pour la génération des images BR, nous utilisons le modèle de E&F

$$f_k = DHW_k s + v_k, k = 1, \dots, K \quad (2)$$

où  $W_k$  est la transformation géométrique entre  $s$  et  $f_k$  interpolée dans la géométrie THR.  $H$  est la fonction d'étalement de point du capteur, correspondant à une convolution d'un filtre gaussien passe-bas par une fonction rectangulaire de largeur 12 pixels.  $D$  est l'opérateur de sous-échantillonnage.  $v_k$  est un bruit de mesure gaussien additif.

Dans la littérature, les images BR  $f$  sont générées usuellement à partir de l'image HR  $x$  servant de référence à la SR. Comme montré sur la figure 4, avec notre méthode, ces images sont générées indépendamment à partir de l'image THR  $s$ , et ne sont donc pas liées entre elles par le modèle E&F. Ainsi, nous n'avons pas introduit de biais en faveur de la méthode Hardie & al..

A partir de l'ensemble des images BR, la reconstruction SR a pour but de fournir une image HR  $\hat{x}$  en utilisant un facteur SR égal à trois. Les méthodes SR sont comparées quantitativement et chaque image  $\hat{x}$  est évaluée par le calcul de la racine carrée de l'écart quadratique moyen (ou *root-mean-squared error*

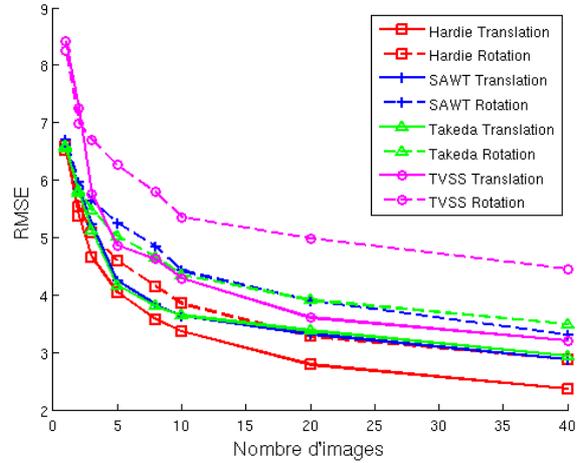
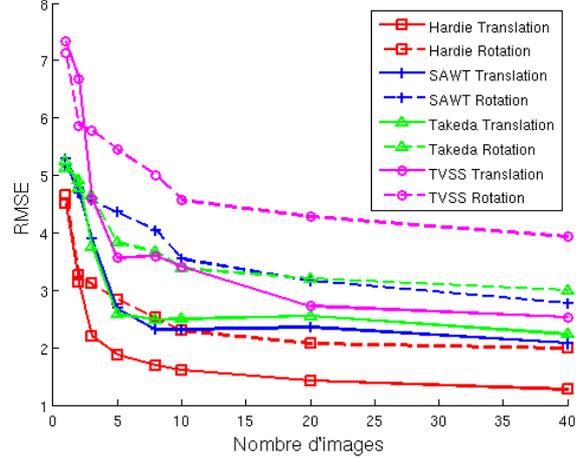


FIG. 5 – RMSE en fonction du nombre d'images BR utilisées, avec un bruit de variance  $\sigma = 1$  (haut),  $\sigma = 5$  (bas).

(RMSE) en anglais) avec l'image HR  $x$  de référence.

### 4 Résultats

Nous présentons dans ce papier les résultats avec les mouvements exacts. Les méthodes comparées ont chacune un ensemble de paramètres (paramètres de régularisation, d'interpolation, nombre d'itérations, etc.). Ces paramètres ont été optimisés pour minimiser le RMSE.

Nous avons considéré deux types de mouvements : translation et translation plus rotation (5 degrés maximum). La figure 5 présente la variation du RMSE en fonction du nombre d'images utilisées, pour un bruit de variance faible  $\sigma = 1$  (haut), ou forte  $\sigma = 5$  (bas).

En général sur l'ensemble des tests réalisés, Hardie & al. donne les meilleurs résultats, SAWT et Takeda & al. sont similaires, et TVSS est le plus mauvais. Nous pouvons noter que toutes les méthodes perdent en qualité de reconstruction entre la translation et la rotation et que TVSS en souffre le plus. La raison peut être liée au modèle de S&S, qui n'est pas

valide dans le cas d'un mouvement affine [2]. De plus, la régularisation VT produisant des « effets cartoon » marqués, ne semble pas être adéquate pour ce type d'images. L'interpolation robuste de Takeda & al. n'améliore pas les résultats par rapport à SAWT sauf dans les cas limites d'un SNR faible et d'un nombre d'images restreint.

Ces résultats confirment la comparaison visuelle, où SAWT fournit des résultats comparables à la méthode Hardie & al. dans le cas d'un SNR et d'un nombre d'images suffisant. Dans cette configuration, il n'apparaît pas nécessaire d'investir dans des méthodes optimales et cela même dans le cas de la rotation. D'autre part, les approches non linéaires TVSS et Takeda & al. n'apportent pas un gain sur la qualité de la reconstruction SR suffisant pour justifier un coût de calcul nettement plus élevé qu'une méthode linéaire comme SAWT.

## 5 Conclusion

Notre contribution sur l'évaluation de méthodes à bas coût doit être complétée dans le cas de rotations plus fortes, et en considérant des déplacements estimés. La méthode Hardie & al., qui utilise le modèle direct le plus précis, fournit les meilleures reconstructions. TVSS donne les résultats les plus mauvais, mais cette conclusion doit être étayée par des tests sur d'autres types de données, comme des mires. SAWT et Takeda & al. donnent des reconstructions similaires, sauf pour un SNR faible où Takeda & al. est meilleur. Sur les exemples traités, le recours à des approches non linéaires ou optimales, coûteuses en temps de calcul n'apporte donc pas d'amélioration déterminante dans les reconstructions. Notons cependant que les écarts entre les erreurs mesurées de ces méthodes avec Hardie & al. se traduisent visuellement par des effets limités : les reconstructions apparaissent un peu moins contrastées et plus floues, comme illustré sur la figure 6. En ce qui concerne l'écart entre la méthode SAWT et Hardie & al., nous avons identifié deux sources de dégradations potentielles. Premièrement, SAWT utilise des translations tronquées au pixel HR le plus proche. Deuxièmement, SAWT repose sur l'approche INU qui est une approximation du problème dans le cas de mouvements non translationnels. Nous comptons dans la suite de nos travaux étudier ces deux types d'approximations pour déterminer leur influence relative dans l'erreur de reconstruction.

## Références

- [1] Sung Cheol Park, Min Kyu Park, and Moon Gi Kang, "Super-resolution image reconstruction : A technical overview," *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 21–36, May 2003.
- [2] Gilles Rochefort, Frédéric Champagnat, Guy Le Besnerais, and Jean-Francois Giovannelli, "An improved observation model for super-resolution under affine motion," *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3325–3337, November 2006.

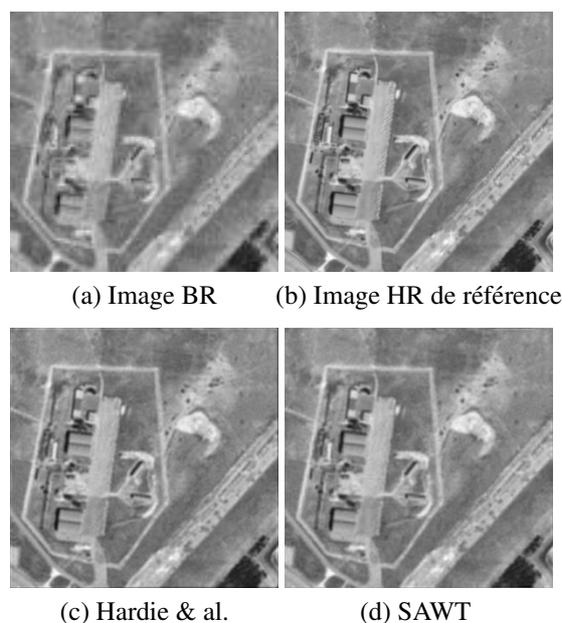


FIG. 6 – Reconstruction SR pour vingt images, un bruit de variance  $\sigma = 5$  et un facteur SR de trois.

- [3] Russel C. Hardie, Kenneth J. Barnard, John G. Bognar, Ernest E. Armstrong, and Edward A. Watson, "High-resolution image reconstruction from a sequence of rotated and translated frames and its application to an infrared imaging system," *Optical Engineering*, vol. 37, no. 1, pp. 247–260, January 1998.
- [4] Sina Farsiu, M. Dirk Robinson, Michael Elad, and Peyman Milanfar, "Fast and robust multiframe super-resolution," *IEEE Transactions on Image Processing*, vol. 13, no. 10, pp. 1327–1343, October 2004.
- [5] Hiroyuki Takeda, Sina Farsiu, and Peyman Milanfar, "Kernel regression for image processing and reconstruction," *IEEE Transactions on Image Processing*, vol. 16, pp. 349–366, 2007.
- [6] Antoine Létienne, Frédéric Champagnat, Guy Le Besnerais, Caroline Kulcsár, and Patrick Viaris De Lesegno, "Fast super-resolution on moving objects in video sequences," in *EUSIPCO European Signal Processing Conference*, August 2008.
- [7] S. Lertrattanapanich and N. K. Bose, "High resolution image formation from low resolution frames using delaunay triangulation," *IEEE Transactions on Image Processing*, vol. 12, no. 11, pp. 1427–1441, 2002.
- [8] Michael Elad and Arie Feuer, "Restoration of a single super-resolution image from several blurred, noisy, and under-sampled measured images," *IEEE Transactions on Image Processing*, vol. 6, no. 12, pp. 1646–1658, December 1997.
- [9] Tuan Q. Pham, Lucas J. Van Vliet, and Klammer Schutte, "Robust fusion of irregularly sampled data using adaptive normalized convolution," in *EURASIP Journal on Applied Signal Processing*, 2006, p. 2006.