

# Un Algorithme Multirésolution de Recalage de Signaux et d'Images

Martin Lefebure et Laurent D. Cohen

CEREMADE, U.R.A. CNRS 749

Université Paris 9 - Dauphine

75775 Paris cedex 16, France

Email: lefebure et cohen@ceremade.dauphine.fr

## RÉSUMÉ

Nous traitons le problème de la convergence d'un algorithme d'estimation du mouvement pour le recalage ou la mise en correspondance de signaux ou d'images. Nous introduisons un algorithme de recalage multirésolution, basé sur une pyramide de Fourier. Dans le cas particulier de translation, nous prouvons la convergence en moyenne quadratique de l'algorithme, sous certaines hypothèses raisonnables. L'algorithme s'effectue à la résolution la plus fine vérifiant l'hypothèse requise, puis par raffinement jusqu'à pleine résolution. Nous illustrons cette méthode par des résultats sur des signaux simples en 1D et une image IRM fonctionnelle 3D.

## ABSTRACT

We address the problem of convergence of a motion estimation algorithm for signal matching or image registration. We introduce a new multiresolution algorithm, based on the Fourier pyramid. In the case of a translation motion, we prove mean quadratic convergence of our algorithm under some reasonable assumptions. The algorithm starts at the finest resolution for which the assumption is satisfied, and then refines until the highest resolution is reached. We illustrate the method by showing results on simple 1D signals and a 3D functional MR image.

## 1 Introduction

Le recalage d'images est maintenant un problème fondamental en analyse d'images médicales (voir par exemple [1]). Le problème trouve son origine dans l'utilisation de techniques dynamiques en imagerie médicale : la diversité des formes de crâne des différents patients, ou encore leur mouvement au cours d'une séquence d'images, rend délicate l'utilisation des données recueillies pour le diagnostic et l'élaboration de statistiques. Dans le premier cas, où il faut recalibrer des images sur une image de référence, on demande un recalage non rigide. Quand le but est d'analyser l'évolution des niveaux de gris dans une séquence dynamique d'images, il n'est pas possible d'éviter un léger mouvement du patient. Il est donc nécessaire d'opérer un recalage affine sur la séquence d'images pour comparer les bons pixels dans les images successives.

Bien que notre algorithme puisse s'appliquer à toute sorte de signal ou d'image, notre objectif initial était de réaliser le recalage d'images 3-D en IRM fonctionnelle. Ces images permettent d'explorer les liens entre les structures du cerveau et leurs fonctions. Afin de tenir compte d'un grand nombre d'expérimentations, il est fondamental de développer des algorithmes automatiques de traitement d'images destinés à extraire l'information utile. Un recalage affine est donc nécessaire pour suivre l'évolution des images et détecter les activations dans celles-ci.

Des méthodes générales de recalage [4, 9] ont été mises en oeuvre en vision par ordinateur ou en imagerie médicale pour l'ajustement de surfaces 3-D, prenant en compte diverses contraintes géométriques ou différentielles [13, 6, 7]. Nous

donnons ici une bonne formulation mathématique du problème, dans le sens où nous garantissons la convergence de l'algorithme multirésolution.

Le recalage de deux images  $u$  et  $v$  est obtenu par une transformation  $T(X)$  minimisant une énergie

$$E(T) = \int \|u(X) - v(T(X))\|^2 \quad (1)$$

Une première utilisation de ce type d'énergie avec des modèles déformables a été présentée dans [8] dans le cadre de mise en correspondance pour une paire stéréo. Notre contribution concerne deux points. Premièrement, nous avons introduit un algorithme multirésolution, basé sur une approche de flot optique de Odobez et Bouthemy [10]. Cela nous permet de trouver la bonne transformation entre deux images 3-D (cf figure 3). Nous supposons que la transformation est une combinaison linéaire de fonctions de base. Ceci représente une connaissance a priori en termes de localisation, de régularité ou d'oscillation sur la solution. Comme deuxième contribution, nous montrons, pour des cas particuliers, que notre algorithme converge. Ceci suppose que certaines hypothèses sur la variation et l'amplitude du mouvement soient vérifiées.

## 2 Algorithme Multirésolution

L'idée principale est de considérer la pyramide multirésolution pour les signaux ou images engendrée par les premiers éléments de la base de Fourier. En effet, la présence de nombreux minima locaux à pleine échelle rend difficile la minimisation directe de l'énergie. La minimisation à une échelle plus grossière permet de simplifier l'énergie. On peut interpréter ce

résultat comme un lissage de la fonction d'énergie, qui garantit la convergence de la descente de gradient vers l'unique minimum global (cf figures 1,2). Cette idée existe déjà sous une forme différente dans des travaux antérieurs [16, 15, 14, 5].

Sous certaines hypothèses (voir section suivante) sur la variation et l'amplitude du mouvement, nous prouvons que l'algorithme converge rapidement vers les paramètres réels du mouvement. Lorsqu'on choisit une résolution suffisamment basse, l'hypothèse est satisfaite et la minimisation peut être appliquée. Ensuite, l'algorithme multirésolution recherche le minimum global en estimant les paramètres puis en recalant, et en ajoutant les termes successifs dans la décomposition de Fourier. Voici un schéma général de l'algorithme :

1. Trouver la résolution la plus grande  $j$  pour laquelle l'hypothèse est satisfaite.
2. Minimiser l'énergie sur les signaux à la résolution  $j$ , par itération des moindres carrés dans l'espace des paramètres définissant la transformation  $T$  (nous supposons cet espace de dimension finie).
3. Tant que  $j$  est plus petit que la résolution maximale, raffiner le résultat en augmentant la résolution et en minimisant la nouvelle énergie,

Pour l'exemple test donné dans la figure 1, le paramètre de translation exact est trouvé. Nous montrons également un exemple de recalage affine d'images 3-D en IRM fonctionnelle (voir commentaires sur la figure 3).

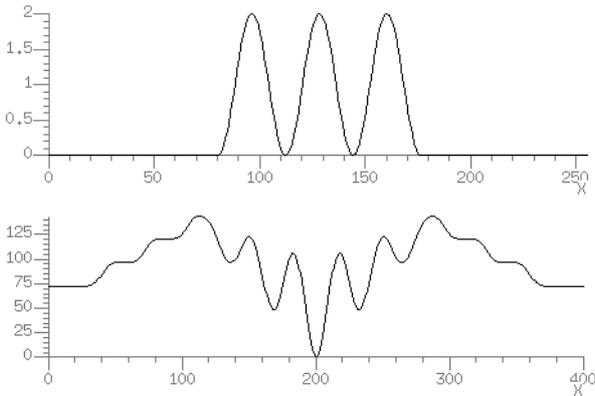


FIG. 1 — Signal Test en haut. Le second signal est le même signal translaté de 200. En bas, l'énergie en fonction du paramètre de translation. Il y a de nombreux minima locaux autour du minimum global en  $x = 200$ .

### 3 Condition de Convergence

Nous détaillons ici l'étape 2 de l'algorithme dans le cas particulier où l'espace de transformation admissible est réduit aux translations et pour le recalage de deux signaux 1D,  $u$  et  $v$ . Nous donnons une condition suffisante de convergence, ainsi que l'algorithme détaillé correspondant.

Supposons que  $u$  et  $v$  appartiennent à  $C_0^1(D)$ , l'espace des fonctions continuellement dérivables à support compact, avec  $D = [-M, M]$ . Supposons également que  $u(x) = v(x + h^*)$ ,

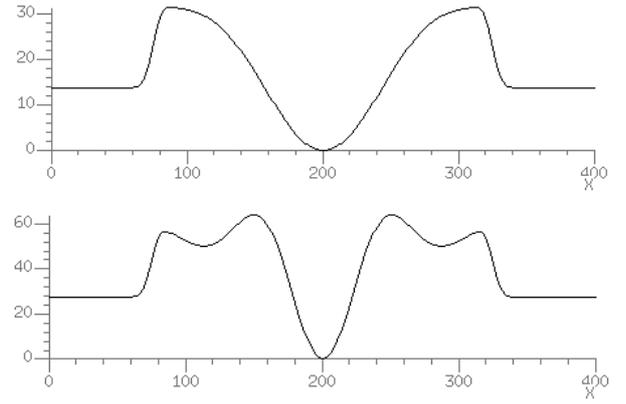


FIG. 2 — Fonction d'énergie obtenue à partir des signaux à l'échelle 2 et 3. Elle doit être comparée à celle du bas dans la figure précédente, où elle est obtenue à pleine résolution (échelle 8). A l'échelle 2, la solution est l'unique minimum local.

$x \in D$ ,  $h^* \in \mathbb{R}$ , et que  $dist(Supp(v), \partial D) \geq |h^*|$ , où  $\partial D$  désigne le bord de  $D$ , et  $Supp(v)$  le support de la fonction  $v$ . Le problème de recalage s'écrit :

$$(P) : \text{Trouver } \hat{h} = \arg \min_h \|u(x) - v(x+h)\|_{L^2(D)}^2$$

Notre objectif est d'analyser les performances théoriques d'un algorithme multirésolution simple. Donnons-nous la suite de sous-espaces vectoriels emboîtés  $V_j$ ,  $j \geq 0$  de  $L^2(D)$  définis par :

$$V_j = \text{vect}\{e_k(x) = \frac{1}{\sqrt{2M}} e^{-i\pi kx/M}, -2^j \leq k \leq 2^j\}.$$

Pour chaque sous-espace  $V_j$ , on note  $\Pi_j$  l'opérateur de projection de  $L^2(D)$  dans ce sous-espace. En pratique, où nous disposerons de  $2M$  échantillons,  $j$  est implicitement majoré par l'inégalité  $1 + 2^{j+1} \leq 2M$ . Appellons  $(PL_j)$  le problème linéarisé et plongé dans  $V_j$ ,  $j \geq 0$  :

$$(PL_j) : \text{Trouver } \hat{h}_j = \arg \min_h \|u(x) - v(x) - v'(x)h\|_{V_j}^2.$$

Notons enfin  $E_j = \|u - v - v'h\|_{V_j}^2$  l'énergie associée. C'est une fonction quadratique de  $h$ . Alors un résultat immédiat est donné par le

**Lemme 3.1** — Si  $\|v'\|_{V_j} \neq 0$ , alors

$$\hat{h}_j = \frac{\langle \Pi_j(v'), \Pi_j(u - v) \rangle_{V_j}}{\|v'\|_{V_j}^2}, \quad (2)$$

et si

$$|h^*| \leq \frac{M}{2^{j+1}}, \quad (3)$$

alors on a :

$$|\hat{h}_j - h^*| \leq \frac{|h^*|}{2}. \quad (4)$$

**Preuve :** en annulant la dérivée de  $E_j$  par rapport à  $h$ , on obtient (2). Par hypothèse,  $v$  s'annule au bord du domaine. On en déduit, via une intégration par parties, que  $c_k(v') = \frac{i\pi k}{M} c_k(v)$ , où  $c_k(v) = \langle v, e_k \rangle_{L^2(D)}$ . On a alors, pour le dénominateur,

$$\|v'\|_{V_j}^2 = \sum_{|k| \leq 2^j} |c_k(v')|^2 = \sum_{|k| \leq 2^j} \left(\frac{\pi k}{M}\right)^2 |c_k(v)|^2.$$

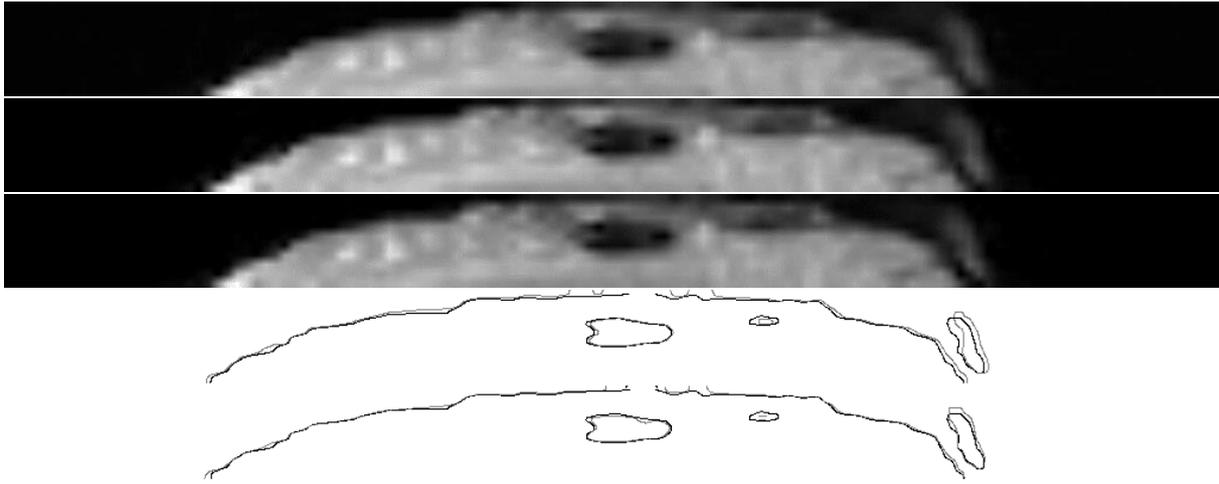


FIG. 3 — Recalage affine d'images 3-D en IRM fonctionnelle : nous le visualisons ici sur des coupes sagittales. L'image du haut est donc une coupe sagittale extraite d'une première image 3-D. En dessous, nous montrons la même section, extraite de l'image 3-D suivante dans la séquence temporelle, après un léger mouvement du patient, avant et après recalage 3D. Comme il est difficile d'apprécier visuellement la différence entre les deux images fixes, nous montrons plus bas les bords détectés à l'aide d'un filtre de Canny-Deriche avant puis après recalage. Les contours noirs représentent la forme originale. Remarquer la précision du recalage sur la petite région à l'extrémité droite (i.e. le sinus longitudinal supérieur).

Comme  $v$  est réelle, ses coefficients de Fourier  $c_k(v)$  et  $c_{-k}(v)$  sont complexes conjugués. Par un simple changement de variable, on a  $c_k(u) = e^{i\pi kh^*/M} c_k(v)$ . On en déduit que, pour le numérateur,

$$\begin{aligned} & \langle \Pi_j(v'), \Pi_j(u - v) \rangle_{V_j} = \sum_{|k| \leq 2^j} c_k(v') \overline{c_k(u - v)} \\ &= \sum_{|k| \leq 2^j} \frac{i\pi k}{M} c_k(v) \overline{(e^{i\pi kh^*/M} - 1) c_k(v)} \\ &= \frac{\pi}{M} \sum_{|k| \leq 2^j} |c_k(v)|^2 k \sin(\pi kh^*/M), \end{aligned}$$

après simplification. On trouve donc :

$$\hat{h}_j = \frac{M}{\pi} \frac{\sum_{|k| \leq 2^j} k |c_k(v)|^2 \sin(\pi kh^*/M)}{\sum_{|k| \leq 2^j} k^2 |c_k(v)|^2}.$$

Maintenant formons la différence

$$\begin{aligned} |\hat{h}_j - h^*| &= \left| \frac{M}{\pi} \frac{\sum_{|k| \leq 2^j} k |c_k(v)|^2 \sin(\pi kh^*/M)}{\sum_{|k| \leq 2^j} k^2 |c_k(v)|^2} - h^* \right| \\ &\leq |h^*| \frac{\sum_{|k| \leq 2^j} k^2 |c_k(v)|^2 \left| \frac{\sin(\pi kh^*/M)}{\pi kh^*/M} - 1 \right|}{\sum_{|k| \leq 2^j} k^2 |c_k(v)|^2}. \end{aligned}$$

Or, pour tout  $x$ , on a  $1 - \frac{x^2}{6} \leq \frac{\sin(x)}{x} \leq 1$ . Par hypothèse,  $|h^*| \leq \frac{M}{2^{j+1}}$  et  $|k| \leq 2^j$ , donc  $\frac{\pi |kh^*|}{M} \leq \frac{\pi |k|}{2^{j+1}} \leq \frac{\pi}{2}$ . On en déduit que :

$$\left| \frac{\sin(\pi kh^*/M)}{\pi kh^*/M} - 1 \right| \leq \frac{1}{6} \left( \frac{\pi kh^*}{M} \right)^2 \leq \frac{\pi^2}{24} \leq \frac{1}{2},$$

d'où le résultat annoncé.  $\square$

Par suite du Lemme 3.1, il est facile de voir que si on remplace  $v$  par  $v_1(x) = v(x + \hat{h}_j)$ , alors les hypothèses du lemme sont à nouveau vérifiées puisque  $\|v'_1\|_{V_j} = \|v'\|_{V_j} > 0$ , la valeur de la translation est maintenant  $h^* - \hat{h}_j$  et :

$$|h^* - \hat{h}_j| \leq \frac{|h^*|}{2} \leq \frac{M}{2^{j+2}} \leq \frac{M}{2^{j+1}}.$$

On peut alors trouver

$$\hat{h}_j^1 = \arg \min_h \|u(x) - v_1(x) - v'_1(x)h\|_{V_j}^2,$$

et le Lemme 3.1 nous garantit que

$$|\hat{h}_j^1 - (h^* - \hat{h}_j)| \leq \frac{|\hat{h}_j - h^*|}{2} \leq \frac{|h^*|}{4} \leq \frac{M}{2^{j+1+2}}.$$

C'est à dire que

$$|\hat{h}_j^1 + \hat{h}_j - h^*| \leq \frac{M}{2^{j+1+2}}.$$

Posons  $\hat{h}_j^0 = \hat{h}_j$ . En itérant le processus, on a par récurrence, à l'étape  $L$  :

$$|\sum_{l=0}^L \hat{h}_j^l - h^*| \leq \frac{M}{2^{j+L+2}}.$$

La série de terme général  $\hat{h}_j^l$  est donc convergente vers  $h^*$ . Enfin, comme

$$\begin{aligned} \|u(x) - v(x + \sum_{l=0}^L \hat{h}_j^l)\|_{\infty} &= \|v(x + h^*) - v(x + \sum_{l=0}^L \hat{h}_j^l)\|_{\infty} \\ &\leq \|v'\|_{\infty} |h^* - \sum_{l=0}^L \hat{h}_j^l| \leq \frac{M \|v'\|_{\infty}}{2^{j+L+2}}, \end{aligned}$$

on en conclut que la suite de fonctions  $v_L(x) = v(x + \sum_{l=0}^L \hat{h}_j^l)$  converge uniformément vers  $u$ . Ce qui prouve le théorème suivant :

**Théorème 3.1** — Sous les conditions du Lemme 3.1, l'algorithme converge au sens où, quand  $L \rightarrow \infty$ ,

$$\sum_{l=0}^L \hat{h}_j^l \longrightarrow h^*,$$

et

$$v_L \longrightarrow u \text{ uniformément.}$$

*Remarque.* Il n'est pas surprenant que  $h^*$  puisse être calculé avec très peu de coefficients de Fourier de  $u$  et  $v$ . En effet, en théorie comme en pratique, la détermination du paramètre de translation  $h^*$  ne pose pas de problème. Si on passe dans Fourier, on a :  $\hat{u}(\xi) = e^{i\pi h^*/M} \hat{v}(\xi)$ , et on peut en

déduire la valeur de  $h^*$  de manière très fiable numériquement [12]. Dans ce cas, il est clair que notre objectif n'est pas de construire une méthode alternative, mais d'illustrer les performances théorique d'un algorithme multirésolution par un cas particulier simple. Notre but est de prolonger ce travail à l'estimation de modèles paramétriques de translation, où celle-ci est décomposée dans une base  $\{\psi_i\}_{i=1..n}$  de fonctions admissibles.

## 4 Justification de l'algorithme

Comme conséquence pratique du théorème, nous proposons un algorithme multirésolution. En effet, il est intéressant de voir que si  $h^*$  n'est pas petit devant  $M$ , c'est à dire que la condition n'est pas satisfaite à pleine résolution ( $j$  maximal), alors on doit commencer l'algorithme à la résolution la plus fine possible pour laquelle la condition suffisante de convergence de l'algorithme est satisfaite ( $j$  maximal tel que  $h^* \leq \frac{M}{2^{j+1}}$ ). Vu que  $h^*$  est justement l'inconnue du problème, il vaut mieux commencer l'algorithme à la résolution la plus basse possible ( $j \geq 0$  minimal pour lequel  $\|v'\|_{V_j} \neq 0$ ).

Nous proposons donc, dans le cas de la translation, l'**Algorithme** :

Soit deux fonctions  $u$  et  $v$ , et une précision désirée  $\varepsilon_0$ .

- Initialisation :  $h \leftarrow 0, w \leftarrow v,$   
 $j \leftarrow \text{Min}\{j \geq 1 \text{ avec } \|v'\|_{V_j} \neq 0\}, \varepsilon \leftarrow \frac{M}{2^{j+1}}$
- Boucle : Tant que ( $\varepsilon \geq \varepsilon_0$ ), Faire
  - $\hat{h} \leftarrow \frac{\langle \Pi_j(w'), \Pi_j(u-w) \rangle_{V_j}}{\|w'\|_{V_j}^2}$
  - $h \leftarrow h + \hat{h}$
  - $w \leftarrow v(x + h)$
  - $\varepsilon \leftarrow \frac{\varepsilon}{2}$
  - éventuellement  $j \leftarrow j + 1$  (améliore la stabilité numérique)

Le dernier point vient du fait que l'erreur d'estimation étant divisée par 2 à chaque itération (equation (4)), l'hypothèse (3) devient vérifiée pour la résolution supérieure ( $j + 1$ ).

## 5 Conclusion

Nous avons proposé un algorithme multirésolution de recalage de signaux ou d'images. Nous avons donné une justification mathématique de la convergence de cet algorithme dans un cas particulier. Cela explique le bon comportement de cette méthode en pratique, que nous avons illustré par des exemples réels et synthétiques. Nous améliorons et étendons actuellement ces résultats à des classes plus générales de transformations.

## Références

- [1] N. Ayache. Medical computer vision, virtual reality and robotics. *Image and Vision Computing*, August 1995.
- [2] R.K. Bajcsy and C. Broit. Matching of deformed images. In *Proc. ICPR'82*, pages 351–353, 1982.
- [3] Ruzena Bajcsy and Stane Kovacic. Multiresolution elastic matching. *Computer Vision, Graphics, and Image Processing*, 46 :1–21, 1989.
- [4] Paul Besl and Neil McKay. A method for registration of 3–D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2) :239–256, February 1992.
- [5] Laurent D. Cohen and Anne Gorre. On the convexity of the active contour energy. In *Proceedings of GRETSI*, Juan-les-Pins, September 1995.
- [6] C. Davatzikos. Non-linear registration of brain images using deformable models. In *Proc. IEEE MMBIA*, San Francisco, USA, June 1996.
- [7] S. Fletcher, A. Bulpitt, and D. Hogg. Global alignment of mr images using a scale based hierarchical. In *Proc. ECCV'96*, pages II :283–292, Cambridge, U. K., 1996.
- [8] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. Snakes : Active contour models. *International Journal of Computer Vision*, 1(4) :321–331, 1988.
- [9] J.F. Mangin, V. Frouin, I. Bloch, and B. Bendriem. Fast nonsupervised 3d registration of pet and mr images of the brain. Technical Report ENST-93-C-006, ENST-Paris, 1993.
- [10] J. M. Odobez and P. Bouthemy. Robust multiresolution estimation of parametric motion models. *Journal of Visual Communication and Image Representation*, 6(4) :348–365, 1995.
- [11] M. Okutomi and T. Kanade. A locally adaptive window for signal matching. *International Journal of Computer Vision*, 7(2) :143–162, January 1992.
- [12] Muriel Roth. Irmf : Correction des mouvements 2d inter-images dans l'espace k. In *Méthodologie en Imagerie Cérébrale Fonctionnelle*, Lyon, Décembre 1996. GdR ISIS - Projet 11.2.
- [13] G. Subsol, J.P. Thirion, and N. Ayache. Non-rigid registration for building 3d anatomical atlases. In *Proc. ICPR94*, pages A :576–578, Jerusalem, Israel, 1994. IEEE.
- [14] Demetri Terzopoulos. Multiresolution algorithms in computational vision. In *Image Understanding*, pages 225–262. S. Ullman, W. Richards, 1986.
- [15] Gary Whitten. A framework for adaptive scale space tracking solutions to problems in computational vision. In *Proc. Third IEEE International Conference on Computer Vision*, pages 210–220, Osaka, Japan, December 1990.
- [16] A. Witkin., D. Terzopoulos, and M. Kass. Signal matching through scale space. *International Journal of Computer Vision*, 1(2) :133–144, 1987.