

Recalage hybride des images médicales basé sur l'information mutuelle et l'ICP accéléré

Leila Benaissa Kaddar, Nacéra Benamrane

Département d'Informatique, Faculté des Sciences, USTOMB
B.P 1505, EL'Mnaouer 31000, Oran, Algérie
nabenamrane@yahoo.com, Bkleila_usto@yahoo.fr

Abstract. Le recalage d'images trouve de nombreuses applications médicales aussi bien dans le suivi thérapeutique que dans le diagnostic d'un patient. En s'inspirant des méthodes proposées dans la littérature nous proposons dans ce papier, une méthode de recalage d'images médicales basée sur une hybridation entre deux techniques géométrique et iconique afin de diminuer le temps de calcul et au même temps d'améliorer la qualité visuelle de l'image recalée en utilisant l'information mutuelle et l'ICP accéléré. Cette approche a été testée sur une panoplie d'images IRM et les résultats obtenus sont encourageants.

Keywords: Recalage iconique, recalage géométrique, information mutuelle, ICP accéléré, imagerie médicale.

1 Introduction

Avant d'étudier deux images (ou deux volumes) correspondant au même objet physique, il faut qu'il existe une concordance de position spatiale entre les deux images (ou volumes). Ceci est réalisé par une opération de recalage.

Le recalage consiste à mettre en correspondance, avec une grande précision, deux images d'un même patient qui n'ont pas été acquises simultanément. En effet, même si deux acquisitions d'images sont successives, le patient bouge de quelques millimètres dans une ou plusieurs des trois dimensions de l'espace. De plus, les mouvements physiologiques du patient sont responsables d'un décalage entre le plan programmé et les images obtenues lors de séquences successives. Cela rend difficile, voire impossible, de savoir si la variation de la taille d'une lésion est le fait d'une modification biologique de la lésion ou au contraire d'un changement de la position du patient entre les deux examens.

Le recalage s'avère indispensable pour la comparaison avec une image source normale (atlas), pour la comparaison par rapport au malade lui-même soit au cours d'une même séance d'acquisition soit par rapport à un examen précédent (notamment pour le suivi temporel). Le choix des attributs utilisés pour guider le recalage est crucial. Il est largement conditionné par la nature des images à traiter. Dès lors, quatre critères caractérisent une méthode de recalage : les attributs, le critère de similarité, le modèle de déformation et la stratégie d'optimisation [1]

En classifiant de manière simple les différentes méthodes de recalage d'images, on distingue deux grandes catégories. Les approches géométriques basées sur l'extraction

de primitives géométriques dans l'image et les approches iconiques basées sur la comparaison des valeurs d'intensité des voxels dans l'image. Par ailleurs, des méthodes hybrides [2] combinant ces différentes approches ont aussi été proposées.

Les méthodes hybrides, ce sont des méthodes qui reposent sur la combinaison de plusieurs types d'information différents. L'idée est d'améliorer la robustesse de l'algorithme de recalage en combinant les avantages liés à chaque type d'information utilisé. Trois cas peuvent être distingués : la combinaison de primitives géométriques de natures différentes, la combinaison de différentes informations issues des niveaux de gris et la combinaison des approches géométriques et iconiques.

Concernant la combinaison de primitives géométriques de natures différentes, on pourra se référer aux travaux de Yuille et al [3] pour la combinaison de points et de courbes, ceux de Loew et al [4] pour la combinaison de courbes et de surfaces et aux travaux de Maurer [5] pour la combinaison de différents types de surfaces. Concernant les exemples de combinaison de différentes informations iconiques, on peut citer les travaux de Plum et al [6] qui utilisent à la fois le gradient de l'image et l'information des niveaux de gris ainsi que ceux de Shen et al [7] qui proposent d'associer à chaque voxel un vecteur d'attributs composé de l'intensité du voxel en question, de différents moments géométriques invariants caractéristiques du voisinage du voxel et d'une information issue de la segmentation en trois classes de l'image (soit les probabilités d'appartenance à chacune des classes dans le cas d'une segmentation floue, soit une étiquette associée au type de frontière entre classes dans le cas d'une segmentation dure).

Enfin, concernant l'utilisation de certaines primitives géométriques pour contraindre des méthodes iconiques, on peut citer Sorzano & al [8] pour la contrainte par des amers ponctuels, et Cahier & al [9] pour la contrainte par des primitives courbes (sillons corticaux). Liu et al [10] proposent par ailleurs une méthode hybride de recalage volumique et surfacique pour la mise en correspondance des zones corticales.

Dans cet article, nous proposons une approche hybride basée sur l'information mutuelle et l'ICP accéléré. Le reste de cet article est organisé comme suit : Dans la section 2, nous exposons notre approche avec plus de détails. La section 3, présente les résultats expérimentaux obtenus. Une conclusion résumera notre contribution.

2. Approche proposée

Pour pallier aux différents inconvénients des techniques géométriques et iconiques telle que la qualité de l'image recalée qui est dégradée dans la méthode géométrique et le coût de calcul qui est considérable, ce qui est dû au fait que chacun des pixels ou voxels de l'image est pris en considération dans la méthode iconique [11], nous proposons une approche de recalage permettant d'hybrider les deux méthodes géométrique et iconique. Notre approche opère en deux passages, le passage global utilise l'ICP accéléré et le local l'information mutuelle comme mesure de similarité et la descente du gradient comme méthode d'optimisation. La figure 1 illustre les différentes étapes de notre approche.

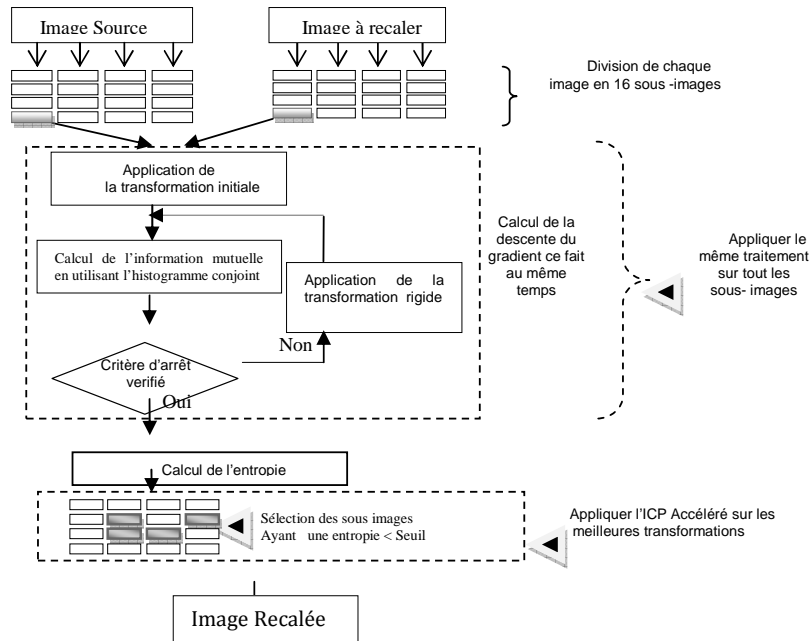


Fig. 1 : Schéma de l'approche proposée

2.1. Transformation rigide

Les transformations rigides sur une image sont des transformations qui s'effectuent sur toutes les coordonnées d'une image en suivant la même modalité. Nous pouvons ainsi translater une image et lui faire effectuer des rotations.

Les coordonnées de chaque point subissent des translations et des rotations suivant une matrice de transformation :

Soit P les coordonnées d'un point d'une image 2d. $P=(x y)$

Soit P' les coordonnées du point équivalent dans l'image modifiée. $P'=(x' y')$

M la matrice de transformation 3x3.

Alors :

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = M \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (1)$$

La forme d'une matrice de transformation pour une translation est:

$$M = \begin{pmatrix} 1 & 0 & x_{trans} \\ 0 & 1 & y_{trans} \\ 0 & 0 & 1 \end{pmatrix} \quad (2)$$

Pour une rotation de θ radians autour de l'axe des x , la matrice est la suivante:

$$M = \begin{pmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (3)$$

Une fois les coordonnées de tous les points de l'image modifiée, nous nous retrouvons avec une image ayant des coordonnées non entières. Il est donc nécessaire de la rééchantillonner. Ceci consiste à créer une nouvelle image ayant des coordonnées entières se rapprochant le plus possible de l'image ayant des coordonnées réelles. Dans notre implémentation, nous avons utilisé comme méthode de ré-échantillonnage l'interpolation bilinéaire. C'est une méthode simple pour éliminer le phénomène d'aliasing. Les 4 points les plus proches des coordonnées calculées dans l'image source sont utilisés en les pondérant par des coefficients inversement proportionnels à la distance et dont la somme vaut 1: Le poids affecté à chaque point est:

$$p(x, y) = \frac{1}{\sqrt{(x - ix)^2 + (y - iy)^2}} \quad (4)$$

La valeur du point obtenu par interpolation sur les 4 points les plus proches est:

$$v(i, j) = \frac{p(ix, iy)v(ix, iy) + p(ix+1, iy)v(ix+1, iy) + p(ix, iy+1)v(ix, iy+1) + p(ix+1, iy+1)v(ix+1, iy+1)}{p(ix, iy) + p(ix+1, iy) + p(ix, iy+1) + p(ix+1, iy+1)} \quad (5)$$

2.2 Maximisation de l'information mutuelle

L'information mutuelle est la quantité d'information d'une image contenue dans une seconde image. Ainsi, lorsque l'information mutuelle entre deux images est au maximum, elles sont identiques. Elle permet de transformer une image pour qu'elle ressemble le plus à une image donnée en maximisant l'information mutuelle des deux images concernées.

Le calcul d'information mutuelle de deux images est basé sur la densité conjointe de probabilité des niveaux de gris des images. Il est nécessaire pour estimer la densité conjointe de probabilité de calculer un histogramme conjoint des niveaux de gris entre ces deux images.

Concrètement, l'histogramme conjoint est un graphe tridimensionnel. Chaque point de l'histogramme représente le nombre de fois qu'une combinaison de niveau de gris entre les deux images est rencontrée.

Soit $g(x, y)$ la valeur de l'histogramme conjoint au point $[x, y]$.

On pose :

$$p_{1,2}(x, y) = \frac{g(x, y)}{\sum_{a,b} g(a, b)} \quad (6)$$

$$p_1(x) = \sum_b p_{1,2}(x, b)$$

$$p_2(y) = \sum_a p_{1,2}(a, y)$$

On a ainsi :

$$MI = \sum_{a,b} p_{1,2}(a, b) \log_2 \frac{p_{1,2}(a, b)}{p_1(a) \cdot p_2(b)} \quad (7)$$

Où :

1. $\sum_{a,b} g(a,b)$ représente le nombre de points utilisés pour créer l'histogramme conjoint.
2. $p_{1,2}$ est l'histogramme conjoint normalisé. En effet la somme de ces valeurs vaut 1. C'est ainsi une distribution de probabilité. $p_{1,2}(x,y)$ peut donc se lire comme la probabilité qu'un point pris au hasard dans l'image A soit la combinaison du niveau de gris x sur l'image A et du niveau de gris y sur l'image B .
3. $p_1(x)$ est aussi une distribution de probabilités. Pour un x donné c'est la probabilité que l'on trouve un point de niveau de gris x sur l'image A .
4. $p_2(y)$ est comme $p_1(x)$. Pour un y donné c'est la probabilité que l'on trouve un point de niveau de gris y sur l'image B .

2.3 La méthode d'optimisation (descente de gradient)

Il n'existe pas d'algorithme universel efficace pour minimiser toutes les fonctions. Pour notre approche nous avons opté pour la descente de gradient comme méthode d'optimisation. Le gradient d'une fonction de R^n dans R est un vecteur de dimension n . Chacun de ses termes contient la dérivée de la fonction par rapport à une variable. Le gradient a la propriété de donner la direction dans laquelle une fonction continue et dérivable varie le plus. La méthode du gradient consiste donc à utiliser la direction donnée par le gradient comme direction dans laquelle des tests vont être fait pour trouver un point où l'information mutuelle est maximale.

Le calcul du gradient se fait comme suit, On calcule tout d'abord la dérivée par rapport à chaque variable.

Soient $PTC = (decx, decy, rot)$ le point courant et eps une petite valeur, on note $MI(x)$ l'information mutuelle au point x .

La dérivée de l'information mutuelle en fonction du décalage en x sera estimée par :

$$der_{dec_x} = \frac{MI(PT_c + (eps, 0, 0)) - MI(PT_c - (eps, 0, 0))}{eps + eps} \quad (8)$$

La dérivée de l'information mutuelle en fonction du décalage en y sera estimée par

$$der_{dec_y} = \frac{MI(PT_c + (0, eps, 0)) - MI(PT_c - (0, eps, 0))}{eps + eps} \quad (9)$$

La dérivée de l'information mutuelle en fonction de la rotation autour de x sera estimée par :

$$der_{rot} = \frac{MI(PT_c + (0, 0, eps)) - MI(PT_c - (0, 0, eps))}{eps + eps} \quad (10)$$

Ensuite le gradient est rempli :

$$grad (MI (PT_c)) = \begin{pmatrix} der_{dec_x} \\ der_{dec_y} \\ der_{rot} \end{pmatrix} \quad (11)$$

Une fois la direction trouvée, le maximum dans la direction du gradient est recherché.

2.4 Entropie

Après l'obtention des meilleures transformations locales de chaque sous-image, l'entropie de chaque sous-image est calculée sachant que l'entropie est la quantité d'information contenue dans une série d'événements. Nous avons opté pour la définition de l'entropie [12] suivante :

$$H = \sum_i p_i \log \frac{1}{p_i} = - \sum_i p_i \log p_i \quad (12)$$

Après le calcul de l'entropie, un seuil S est fixé afin de choisir que les sous-images ayant une quantité d'information importante, ce qui permet de rendre le recalage plus rapide et plus robuste. En fait, toutes les sous-images de l'image A ne sont pas utilisées pour construire les appariements. Les sous-images sont en fait triées selon leur valeur de l'entropie et on ne considère que les sous-images ayant une entropie $< S$. En effet, les sous-images avec une forte entropie sont susceptibles d'être dans une zone homogène de l'image A . Elles pourront s'apparier avec n'importe quelle région homogène de l'image B , et les appariements ainsi construits ne pourront que gêner le calcul de la transformation.

2.5 ICP accéléré

L'ICP accéléré est appliqué sur les meilleures transformations locales des sous-images sélectionnée précédemment. Rappelons tout d'abord l'algorithme ICP.

Soient deux ensembles de points P et X . L'algorithme ICP produit une transformation rigide optimale \vec{q} pour un minimal local X_k .

Notre implémentation est basée sur les quaternions qui est seulement utilisé en 2D et 3D. Ce pendant l'algorithme de décomposition en valeur singulière SVD peut être utilisé afin de généraliser la méthode à N dimension [13][14].

Le centre de masse $\vec{\mu}_p$ pour l'ensemble de point mesuré P et le centre de masse $\vec{\mu}_x$ pour l'ensemble de point X est donné comme suit :

$$\vec{\mu}_p = \frac{1}{N_p} \sum_{i=1}^{N_p} \vec{P}_i \quad \text{et} \quad \vec{\mu}_x = \frac{1}{N_x} \sum_{i=1}^{N_x} \vec{x}_i \quad (13)$$

La matrice de covariance de l'ensemble de P et X est :

$$\sum_{px} = \frac{1}{N_p} \sum_{i=1}^{N_p} [(\vec{P}_i - \vec{\mu}_p)(\vec{x}_i - \vec{\mu}_p)^t] = \frac{1}{N_p} \sum_{i=1}^{N_p} [\vec{P}_i \vec{x}_i^t] - \vec{\mu}_p \vec{\mu}_x^t \quad (14)$$

Les composants cycliques de la matrice antisymétrique $A_{ij} = (\sum_{px} - \sum_{px}^t)_{ij}$ sont utilisés pour former un vecteur colonne $\Delta = [A_{23} A_{31} A_{12}]^T$. Ce vecteur est ainsi utilisé pour former la matrice $Q(\sum_{px})$ 4x4 symétrique.

$$Q(\sum_{px}) = \begin{bmatrix} \text{tr}(\sum_{px}) & & & \\ \Delta & \sum_{px} + \sum_{px}^t & & \\ & & \text{tr}(\sum_{px}) - \text{tr}(\sum_{px}^t) & \\ & & & I_3 \end{bmatrix} \quad (15)$$

Où I_3 est une matrice identité 3×3 . Le vecteur unitaire propre de la matrice \vec{q}_r correspondante au maximum des valeurs propres de la matrice $Q(\sum_{px})$ est sélectionné comme la rotation optimale. Le Jacobi peut être utilisé à la matrice $Q(\sum_{px})$ afin d'obtenir le vecteur propre [15][16].

Le vecteur de transformation optimale est donné par :

$$\vec{q}_T = \vec{\mu}_x - R(a_R) \vec{\mu}_p \quad (16)$$

Le vecteur d'espace 7 \vec{q}_k correspondant à la transformation rigide est construit comme la concaténation de \vec{q}_R et \vec{q}_T qui est : $\vec{q}_k = \left[\vec{q}_R \mid \vec{q}_T \right]^T$ (17)

Initialement \vec{q}_k est un vecteur identité et $P_k = P$.

Au début de chaque itération, un nuage X_k qui est une projection de P_k sur X , est construit. Ensuite le vecteur courant \vec{q}_k est calculé en fonction de P et X . Enfin \vec{q}_k est appliqué à P . La convergence est atteinte lorsque l'erreur quadratique soit en dessous de la valeur prédéfinie t .

L'algorithme :

1. Calculer la projection $X_k = CP(P_k, X)$
2. Calculer le recalage $\vec{q}_k = Q(P_0, X_k)$
3. Appliquer le recalage $P_{k+1} = \vec{q}_k(P_0)$
4. Répéter jusqu'à la convergence : $MSE(\vec{q}_k(P), X) < t$

L'ICP accéléré est originalement proposé par Besl et McKay [17], l'idée est de prédire l'évolution de \vec{q}_k , la transformation rigide.

Durant l'exécution de l'ICP une séquence des vecteurs de recalage est générée : $\vec{q}_0, \vec{q}_1, \vec{q}_2, \vec{q}_3, \vec{q}_4, \dots$, qui tracent le chemin dans la forme de l'espace des états du recalage à partir de la transformation identité jusqu'à la correspondance d'une forme optimale locale.

Considérant la séquence de vecteur de différence défini par :

$$\Delta \vec{q}_k = \vec{q}_k - \vec{q}_{k-1} \quad (18)$$

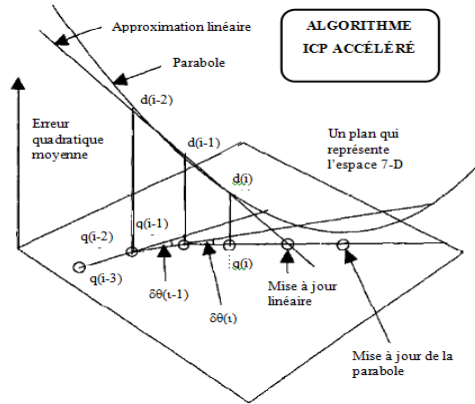


Fig.2 Variation de q_k dans l'espace d'état du recalage

Soit l'angle dans l'espace 7 entre les deux dernières directions :

$$\theta_k = \cos^{-1} \frac{\Delta \vec{q}_k \cdot \Delta \vec{q}_{k-1}}{\|\Delta \vec{q}_k\| \|\Delta \vec{q}_{k-1}\|} \quad (19)$$

Soit $\delta\theta$ une tolérance angulaire suffisamment petite :

$$\text{Si } \theta_k < \delta\theta \quad \text{et} \quad \theta_{k-1} < \delta\theta$$

Alors, il y a un bon alignement de direction pour les trois derniers vecteurs d'état de recalage \vec{q}_k, \vec{q}_{k-1} , et \vec{q}_{k-2} .

Soient d_k, d_{k-1} et d_{k-2} les erreurs quadratiques moyennes associées, soient v_k, v_{k-1} et v_{k-2} une approximation des valeurs d'argument de la longueur d'arcs associés :

$$\begin{aligned} v_k &= 0, \\ v_{k-1} &= -\|\Delta \vec{q}_k\|, \\ v_{k-2} &= -\|\Delta \vec{q}_{k-1}\| + v_{k-1} \end{aligned} \quad (20)$$

Après, une approximation linéaire et une interpolation parabolique, les trois derniers points de repères sont calculés :

$$\begin{aligned} d_1(v) &= a_1 v + b_1 \\ d_2(v) &= a_2 v^2 + b_2 v + c_2 \end{aligned} \quad (21)$$

Ce qui nous donne une modification linéaire possible basée sur le passage par zéro de la ligne et une modification de parabole possible basée sur le point extrême de la parabole.

$$v_1 = -b_1/a_1 > 0, v_2 = -b_2/2a_2 \quad (22)$$

Nous avons utilisé une valeur maximale possible v_{max} en fonction des modifications de \vec{q}_k .

1. Si $0 < v_2 < v_1 < v_{max}$ ou $0 < v_2 < v_{max} < v_1$, on utilise la parabole basé sur le vecteur de recalage modifié $\vec{q}'_k = \vec{q}_k + v_2 \Delta \vec{q}_k / \|\Delta \vec{q}_k\|$ à la place du vecteur usuel \vec{q}_k lorsqu'on effectue la mise à jour sur l'ensemble de points ce qui veut dire $C_{k+1} = \vec{q}'_k(C_0)$.
2. Si $0 < v_1 < v_2 < v_{max}$ ou $0 < v_1 < v_{max} < v_2$ ou $v_2 < 0$ et $0 < v_1 < v_{max}$, on utilise la ligne basée sur le vecteur de recalage modifié $\vec{q}'_k = \vec{q}_k + v_1 \Delta \vec{q}_k / \|\Delta \vec{q}_k\|$ à la place du vecteur usuel \vec{q}_k .
3. Si $v_1 > v_{max}$ et $v_2 > v_{max}$, on utilise un maximum de modification autorisé $\vec{q}'_k = \vec{q}_k + v_{max} \Delta \vec{q}_k / \|\Delta \vec{q}_k\|$ à la place du vecteur usuel \vec{q}_k .

3. Résultats expérimentaux

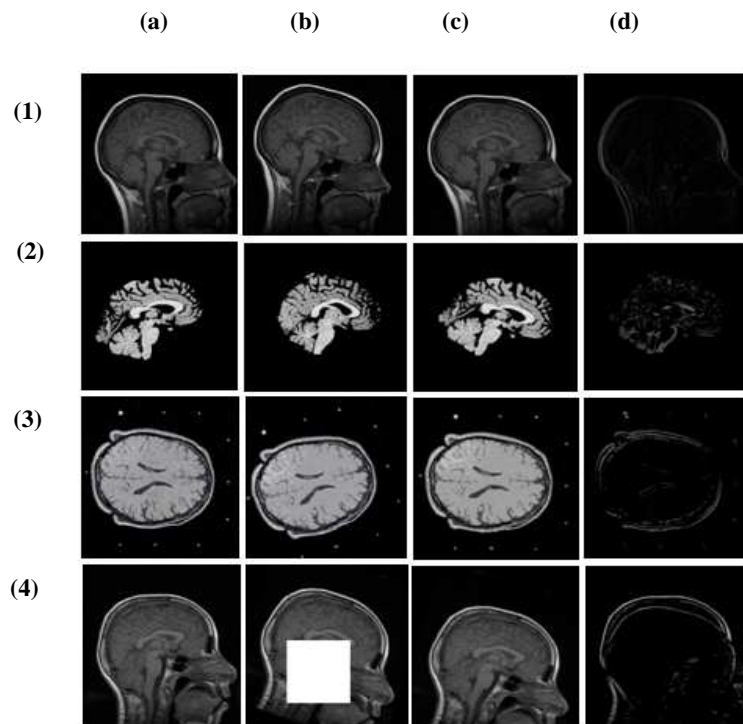
Notre méthode a été testée sur des images IRM de taille 256×256 . Le seuil S concernant l'entropie est fixé à 0.3 et le test d'arrêt (valeur de la norme de différence entre l'image résultat et l'image source) doit être inférieur à 0.3. Le temps de calcul global est aux environs de 1 seconde. Le tableau 1 montre quelques caractéristiques sur les bases d'images et pour ces bases le temps de calcul et l'information mutuelle obtenus.

La figure 3 montre les résultats obtenus, les exemples (1) et (2) c'est dans le cas où nous avons deux prises successives du même patient, les résultats du recalage sont encourageants vu que l'image de différence montre une légère différence entre l'image source et l'image résultat.

Notre méthode traite aussi le cas d'une image segmentée manuellement par un expert comme étant une image source et une autre image non segmentée comme étant une image cible, c'est ce qui est illustré dans l'exemple (3). Le dernier exemple (4), montre le cas d'une image ayant des données manquantes qu'on a pu compléter.

Table 1. Résultats obtenus par notre approche.

<i>N° de la base de données</i>	<i>Vecteur de la transformation (Trx, Try, Rot)</i>	<i>Information mutuelle</i>	<i>Temps d'exécution(s)</i>
(1)	(4, 10, 0.2)	0.9073	0.639
(2)	(20, 5,-0.4)	0.0578	0.661
(3)	(8, 0, -0.3)	0.3252	0.696
(4)	(5, 22, -0.28)	0.3911	0.694



(a) image source, (b) image cible (c) image résultat et (d) l'image de différence

Fig.3 Exemples des images recalées

4. Conclusion

Dans ce papier, nous proposons une approche hybride de recalage basée sur les deux techniques géométrique et iconique. Cette approche opère en deux passages, le premier est local en utilisant l'information mutuelle et le deuxième passage est global en utilisant la dernière extension de l'algorithme ICP, qui est l'ICP accéléré.

Les résultats obtenus sont encourageants et ont permis de diminuer le temps de calcul et au même temps d'améliorer la qualité visuelle de l'image recalée. L'intégration d'autres critères de similarité et l'utilisation d'un schéma multi-échelle peuvent améliorer les résultats obtenus.

Références

1. L.G. Brown. A survey of image registration techniques. *ACM Computing Surveys*, vol. 24, no. 4, pages 325-376, décembre 1992.
2. C. Barillot. Fusion de données et imagerie 3D en médecine. Habilitation à diriger des recherches, Université de Rennes I, 1999.
3. A.L. Yuille, P.W. Hallinan & D.S. Cohen. Feature Extraction from Faces Using Deformable Templates. *International Journal of Computer Vision*, vol. 8, no. 2, pages 99-111, 1992
4. L. Hsu, M. H. Loew & J. Ostuni. Automated Registration of Brain Images Using Edge and Surface Features. *IEEE Engineering in Medicine and Biology Magazine*, vol. 18, no. 6, pages 40-47, November 1999.
5. C.R. Jr Maurer, R.J. Maciunas & J.M. Fitzpatrick. Automated Registration of Brain Images Using Edge and Surface Features. *IEEE Transactions on Medical Imaging*, vol. 17, no. 5, pages 753-761, october 1998.
6. J.P.W. Pluim, J.B.A. Maintz & M.A. Viergever. Image registration by maximization of combined mutual information and gradient information. *IEEE Transactions on Medical Imaging*, vol. 19, no. 8, pages 809-814, août 2000.
7. D. Shen & C. Davatzikos. HAMMER : Hierarchical Attribute Matching Mechanism for Elastic Registration. *IEEE Transactions on Medical Imaging*, vol. 21, no. 11, pages 1421-1439, november 2002.
8. C. O.S. Sorzano, P. Thévenaz & M. Unser. Elastic Registration of Biological Images Using Vector-Spline Regularization. *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 4, pages 652-663, April 2005.
9. P. Cachier, E. Bardinnet, D. Dormont, X. Pennec & N. Ayache. Iconic feature based non rigid registration : the PASHA algorithm. *Computer Vision and Image Understanding*, vol. 89, no. 2-3, pages 272-298, 2003.
10. T. Liu, D. Shen & C. Davatzikos. Deformable registration of cortical structures via hybrid volumetric and surface warping. *NeuroImage*, vol. 22, no. 4, pages 1790-1801, 2004.
11. Vincent NOBLET, « Recalage non rigide d'images cérébrales 3D avec contrainte de conservation de la topologie », thèse doctorat, pages 29-30, 2006.
12. C. Shannon, "A mathematical theory of communication," technical report, Bell System.
13. Arun, K. S., Huang, T., and Blostein, S. D. (1987). Least square fitting of two 3-d point sets. *IEEE Trans. Pattern Anal. Machine Intelligence*. vol. PAMI-9
14. Gu, M., Demmel, J. W., and Dhillon, I. (1994). Efficient computation of the singular value decomposition with applications to least squares problems. Technical Report CS-94-257, institute Knoxville, TN, USA.
15. Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1992). *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, 2nd ed. Cambridge England: Cambridge University Press.
16. Golub, G. H. and Loan, C. F. V. (1989). *Matrix Computations*, 2nd Edition. Johns Hopkins University Press
17. Besl, P. J. and McKay, N. D. (1992). A method for registration of 3-d shapes. *IEEE transactions on pattern analysis and machine intelligence*, Vol. 14, NO. 2.