

Mesures de similarité robustes pour le recalage d'images médicales volumiques multimodales

Robust similarity metrics for the registration of 3D multimodal medical images

par Christophoros NIKOU^{†‡}, Fabrice HEITZ[†], Jean-Paul ARMSPACH[‡], Izzie-Jacques NAMER[‡]

[†] Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédétection
UPRESA CNRS 7005 / Université Strasbourg I 4, Bd. Sébastien Brant, 67400 Illkirch, France

[‡] Institut de Physique Biologique
UPRESA CNRS 7004 / Université Strasbourg I
Faculté de Médecine 4 rue Kirschleger, 67085 Strasbourg, France
e-mail : nikou@picabia.u-strasbg.fr

résumé et mots clés

Le recalage non supervisé d'images médicales volumiques reste un problème difficile en raison de l'importante variabilité et des grandes différences d'information pouvant apparaître dans des séquences d'images de même modalité ou dans des couples d'images multimodales. Nous présentons dans cet article des méthodes robustes de recalage rigide d'images 2D et 3D monomodales et multimodales, reposant sur la minimisation de mesures de similarité inter-images. Les méthodes proposées s'appuient sur la théorie de l'estimation robuste et mettent en œuvre des M -estimateurs associés à des techniques d'optimisation stochastique multigrilles rapides. Ces estimateurs robustes sont évalués à travers le recalage d'images médicales volumiques monomodales (IRM/IRM) et multimodales (IRM/TEMP). Ils sont comparés aux autres fonctions de similarité classiques, proposées dans la littérature. Les méthodes de recalage robustes ont, en particulier, été validées dans le cadre d'un protocole comparatif mis en place par l'Université de Vanderbilt. Elles sont actuellement utilisées en routine clinique et conduisent, tant pour les images de même modalité que pour les images multimodales à une précision sous-voxel, comparable aux meilleures méthodes actuelles. Elles permettent de plus de recalcer des couples d'images sur lesquels les méthodes classiques échouent.

Recalage d'images monomodales et multimodales, mesures de similarité, estimation robuste, optimisation stochastique, évaluation de la précision du recalage.

abstract and key words

In this paper, we develop data driven registration algorithms, relying on pixel similarity metrics, that enable an accurate rigid registration of dissimilar single or multimodal 2D/3D medical images. Gross dissimilarities are handled by considering similarity measures related to robust M -estimators. Fast stochastic multigrid optimization algorithms are used to minimize these similarity metrics. The proposed robust similarity metrics are compared to the most popular standard similarity metrics on real MRI/MRI and MRI/SPECT image pairs showing gross dissimilarities. A blinded evaluation of the algorithm was performed, using as gold standard a prospective, marker-based registration method, by participating in a registration evaluation project (Vanderbilt University). Our robust similarity measures compare favourably with all standard (non robust) techniques.

Single and multimodal image registration, dissimilar image registration, similarity metrics, robust estimation, stochastic optimization, registration accuracy, registration evaluation.

1. introduction

Le but du recalage d'images est de mettre en correspondance des images 2D ou 3D afin de superposer les pixels ou voxels représentant les mêmes structures. Dans le cas des images médicales, la mise en correspondance des structures anatomiques est recherchée. Ces deux dernières décennies, les progrès de l'imagerie neurologique ont révolutionné la recherche clinique dans les domaines de la neurologie et de la neurochirurgie. Les techniques modernes d'imagerie fournissent ainsi des informations anatomiques (IRM, CT), fonctionnelles (TEMP, TEP, IRM fonctionnelle) ou métaboliques (TEP, spectroscopie RMN). L'exploitation de séquences temporelles d'images et/ou d'images multimodales, dans l'objectif de détecter des changements ou d'étudier des évolutions de pathologie chez un même patient, rend ainsi nécessaire un recalage fiable et précis.

De nombreuses méthodes de recalage ont été proposées dans ce domaine. Une comparaison entre différentes méthodes de recalage d'images médicales, sur une base commune d'images 3D multimodales est menée dans [34]. Plusieurs synthèses récentes proposent des typologies des différentes approches en imagerie médicale [19, 30], ou dans un contexte plus général [5]. Une première famille d'approches, supervisées, implique une intervention manuelle de l'opérateur [25]. Dans cet article nous nous plaçons dans le cadre de méthodes non supervisées de recalage, guidées uniquement par les données. Les approches s'appuyant sur la minimisation d'une fonction de coût ou d'énergie, associée à une mesure de similarité inter-image, ont été largement utilisées dans ce contexte [5, 19], aussi bien dans le cas d'images de même modalité [12, 30] que dans le cas d'images multimodales [35].

Dans le cas d'images monomodales, outre les classiques mesures de corrélation [5], la fonction de coût utilisée est souvent quadratique, associée à un modèle d'observation gaussien ou de façon équivalente, à une formulation au sens des moindres carrés [8]. D'autres mesures de similarité, s'appuyant sur diverses grandeurs statistiques estimées dans l'image, ont été introduites dans le cas (notoirement plus difficile) du recalage d'images multimodales. Les variances intra-image, associées à un critère d'uniformité inter-image, ont par exemple été utilisées par Woods *et al.* [35], de même que l'information mutuelle, liée à des critères entropiques, introduite récemment par Maes *et al.* [16] ainsi que par Wells *et al.* [33].

Les techniques de recalage s'appuyant sur de telles mesures de similarité se sont montrées très efficaces dans les problèmes de recalage usuels, mais présentent toutefois des limitations, qui peuvent s'avérer rédhibitoires lorsque se pose le problème du recalage d'images présentant des différences d'informations importantes (non prises en compte dans le modèle), dues par exemple à l'évolution d'une pathologie, à des données manquantes ou à des phénomènes de diffusion étendue comme en imagerie TEMP.

Plus précisément, dans le cas d'images de même modalité, le modèle d'observation gaussien suppose que les deux images à

recaler ne diffèrent, après recalage, que par un bruit additif, supposé blanc et gaussien. Ce modèle ne prend pas en compte des situations usuelles telles que l'évolution d'une lésion dans une séquence d'images ou le cas d'acquisitions multiples ne se recouvrant pas complètement.

Dans le cas du recalage d'images multimodales, un modèle maintenant classiquement utilisé est le modèle proposé par Woods *et al.* [35]. La fonction de similarité de Woods *et al.* suppose qu'à une région uniforme dans l'image IRM, correspond une région également uniforme dans l'image TEP ou TEMP, lorsque les deux images sont recalées (critère d'uniformité inter-image UI). La méthode de recalage décrite dans [35] consiste donc à minimiser, pour une segmentation préalable de l'image IRM en zones uniformes, la variance des régions projetées dans l'image TEP ou TEMP. L'hypothèse d'uniformité inter-image n'est toutefois qu'une approximation, qui peut se révéler grossière dans les situations difficiles décrites ci-dessus : évolution significative d'une lésion, etc. Notons également que l'intérêt clinique d'un couple d'images de même modalité ou multimodales réside précisément dans les informations complémentaires (non redondantes), qu'il peut apporter au médecin. L'innovation portée par l'une des images par rapport à l'autre n'est généralement pas prise en compte par les modèles et conduit donc à des erreurs de recalage, qui, même légères, peuvent être très gênantes, lorsqu'une détection précise de changements inter-images est requise (voir par exemple figure 6). Lorsque la différence d'information entre images est très importante, elle peut même rendre impossible tout recalage par les méthodes classiques.

Nous proposons dans cet article une approche s'appuyant sur la théorie de l'estimation robuste [4, 21], permettant un recalage rigide, entièrement automatique, d'images 3D monomodales et multimodales. L'objet de l'estimation robuste est de permettre une détermination précise des paramètres d'un modèle (ici la transformation rigide utilisée pour le recalage), en présence de pollution par des données « aberrantes » (« outliers » en anglais). Le recalage est basé sur la minimisation de fonctions de coût prenant en compte, de façon sélective (et robuste) l'ensemble des voxels des deux volumes à recaler. L'utilisation d'estimateurs robustes, dans ce contexte, permet de recaler des couples d'images qu'il n'était pas possible de recaler avec les techniques classiques, et permet de mettre en évidence ainsi des changements temporels ou des évolutions de pathologie dans des cas difficiles.

Notons qu'une approche robuste a été proposée récemment et indépendamment dans [1] pour le recalage d'images IRM/IRM 2D. Contrairement aux approches décrites dans ce travail, la méthode citée s'appuie sur des primitives d'images discrètes et elle ne s'applique pas à des images multimodales 3D. L'approche décrite dans [1] est d'autre part très coûteuse, car reposant sur l'estimateur robuste des moindres carrés médians [21]. Venot *et al.* [12, 32] ont également étudié, dans un cadre qui ne se rattachait pas à l'estimation robuste, des critères « robustes » (s'appuyant sur le nombre de changements de signe dans la différence inter-image) pour le recalage d'images de mêmes modalités présentant

des évolutions de lésions. Quoique efficace, leur approche n'a pas été, à notre connaissance, étendue aux images multimodales.

L'article est organisé comme suit : dans la section 2, nous rappelons le formalisme du recalage par fonctions de similarité, ainsi que les fonctions de coût classiquement utilisées en imagerie monomodale et multimodale. Ces méthodes serviront de base de comparaison dans la suite. Nous motivons ensuite, dans la section 3, l'usage des fonctions de coût robustes que nous proposons dans le cas des images monomodales et multimodales. Nous y décrivons également les techniques algorithmiques mises en œuvre pour la minimisation de ces fonctions de coût (relaxation stochastique multigrille rapide). Des résultats expérimentaux portant à la fois sur des problèmes de recalage simulés, sur le recalage d'un objet test et sur le recalage de séquences d'images IRM/IRM et d'images IRM/TEMP de patients, sont présentés et commentés dans la section 4. Une comparaison des principales méthodes citées précédemment est menée sur des cas où les paramètres de recalage sont connus (soit par simulation, soit par recalage manuel préalablement effectué par un médecin). Nous présentons enfin une évaluation des algorithmes développés sur la base d'images de l'Université de Vanderbilt [34], qui permet une comparaison « objective » avec les autres approches qui se sont soumises à ce protocole. Les approches robustes conduisent aux meilleurs résultats ou à des résultats comparables aux meilleures méthodes actuelles [34], avec une précision sous-voxel. De plus, ces approches permettent de traiter des cas pour lesquels les méthodes classiques donnent des résultats erronés.

2. fonctions de similarité classiques

2.1. fonctions de similarité pour le recalage d'images monomodales

Dans le cas d'images monomodales, la fonction de similarité quadratique [8] est utilisée dans de nombreux domaines d'applications. Le recalage consiste ainsi à déterminer les paramètres de translation et de rotation 3D¹ :

$$\Theta = (T_x, T_y, T_z, \theta_x, \theta_y, \theta_z)^T$$

de la transformation rigide T_Θ qui minimise la fonction de coût suivante :

$$E(T_\Theta(x)) = \sum_x [I_{ref}(x) - I_{rec}(T_\Theta(x))]^2 \quad (1)$$

où $I_{ref}(x)$ désigne l'image de référence, $I_{rec}(x)$ l'image à recalcer, et x est le vecteur 3D des coordonnées spatiales. Cette fonction de

similarité est liée à une estimation non linéaire classique au sens des moindres carrés ou à des modèles de bruit blanc gaussien. La corrélation normalisée, très efficace en présence de bruit blanc gaussien a également été très largement mise en œuvre dans ce contexte [5, 19, 30]. Elle est directement liée à la fonction de similarité quadratique [5].

Une première approche robuste du recalage d'images monomodales bidimensionnelles présentant des dissimilarités a été proposée en 1989 par Herbin *et al.* [12]. Cette approche consiste à minimiser le nombre de changements de signe lors d'un balayage de l'image de différence :

$$D(x) = I_{ref}(x) - I_{rec}(T_\Theta(x)) \quad (2)$$

La méthode est appelée stochastique car elle ne s'applique que lorsqu'il y a du bruit dans les images. Elle s'est en particulier révélée performante pour le recalage d'images scintigraphiques polluées par du bruit poissonnien [12]. Des variantes déterministes de ce critère sont également proposées dans [12].

2.2. fonctions de similarité pour le recalage d'images multimodales

2.2.1. l'uniformité inter-image

Dans le cas d'images multimodales, le problème du recalage est beaucoup plus délicat, en raison des grandes différences d'informations existant entre les deux images. Le critère proposé par Woods *et al.* [35] est actuellement l'un des critères les plus souvent référencés. La fonction de coût proposée par Woods *et al.* s'appuie sur l'hypothèse d'uniformité inter-image pour un couple d'images multimodales. L'hypothèse centrale est qu'à une zone homogène de l'image de référence correspond, après recalage, une zone également homogène dans la seconde image. On réalise donc tout d'abord une partition de l'image de référence (en régions non forcément connexes) par une simple classification au sens des niveaux de gris des voxels de cette image. Le nombre de niveaux de gris G dans l'image est en général réduit (nous considérons $G = 256$ classes dans notre implantation). La partition (spatiale) ainsi obtenue est ensuite simplement projetée sur l'image à recalcer. On calcule alors les valeurs moyennes μ_g , $g = 1, \dots, G$ ainsi que les variances σ_g^2 , $g = 1, \dots, G$ des G régions ainsi définies dans l'image à recalcer. S'appuyant sur l'hypothèse d'uniformité inter-image, Woods propose de minimiser dans l'image à recalcer les variances (normalisées) des régions de la partition, ce qui se traduit par la fonction de coût suivante :

$$E(T_\Theta(x)) = \sum_{g=1}^G \frac{N_g}{N} \frac{\sigma_g(T_\Theta(x))}{\mu_g(T_\Theta(x))}, \quad (3)$$

1. Le facteur d'échelle est obtenu à partir de la taille connue des voxels. Il pourrait également être intégré dans l'estimation, au détriment du coût calculatoire.

où :

$$\sigma_g(T_\Theta(x)) = \sqrt{\sum_{x|I_{ref}(x)=g} [I_{rec}(T_\Theta(x)) - \mu_g(T_\Theta(x))]^2} \quad (4)$$

et :

$$\mu_g(T_\Theta(x)) = \frac{1}{N_g} \sum_{x|I_{ref}(x)=g} I_{rec}(T_\Theta(x)) \quad (5)$$

Dans l'équation (3), N représente le nombre de voxels de l'image volumique et N_g correspond au nombre de points possédant le niveau de gris g dans l'image de référence.

Notons à nouveau que l'hypothèse d'uniformité inter-image n'est qu'une approximation, puisque les images qu'il est intéressant de recalcr présentent généralement des différences significatives (assimilables à une « innovation » ou un « événement » violant cette hypothèse d'uniformité). En effet, l'intérêt du recalage de deux images réside dans la comparaison d'informations différentes mais complémentaires obtenues soit en utilisant des modalités d'acquisition différentes, soit à la suite d'un événement temporel.

Notons également que la fonction de coût de Woods *et al.*, *a priori* destinée au recalage d'images multimodales s'applique *a fortiori* au recalage d'images de mêmes modalités. Il apparaît toutefois, comme nous le verrons dans la section 4, que le recalage obtenu est moins précis, dans ce cas, qu'avec une simple fonction de similarité quadratique.

Alpert *et al.* [2] ont proposé une version modifiée de la fonction de similarité de Woods. La mesure qu'ils proposent s'écrit :

$$E(T_\Theta(x)) = \sum_{g=1}^G w_g^2 \sum_{x|I_{ref}(x)=g} \left(1 - \frac{I_{rec}(T_\Theta(x))}{\mu_g(T_\Theta(x))}\right)^2 \quad (6)$$

où w_g est une fonction de pondération qui tient compte du nombre de voxels dans la partition g . Contrairement à la fonction de Woods (3), les variances ne sont pas normalisées par le nombre de points dans une partition, ce qui ne donne pas davantage d'importance aux régions peuplées. Dans l'équation (6) le choix de la fonction de pondération est laissé libre afin de l'adapter à l'application. Alpert *et al.* [2] ont adopté :

$$w_g = \frac{N_g^2}{N(N_g - 1)} \quad (7)$$

ce qui nous ramène à une variante de l'équation (3).

Exploitant le même principe, Ardekani *et al.* [3] ont modifié la méthode en utilisant l'algorithme de *K-means* pour réaliser le partitionnement en régions et en minimisant directement la variance :

$$E(T_\Theta(x)) = \sum_{g=1}^G \sum_{x|I_{ref}(x)=g} |I_{rec}(T_\Theta(x)) - \mu_g(T_\Theta(x))|^2 \quad (8)$$

2.3. mesures de similarité basées sur l'entropie

D'autres approches, reposant sur le partitionnement de l'image de référence puis l'exploitation de l'histogramme conjoint ont été conçues plus récemment. Ainsi, la minimisation de l'entropie conjointe a été proposée [16,29] :

$$E(I_{ref}(\cdot), I_{rec}(T_\Theta(\cdot))) = \mathcal{H}(I_{ref}(\cdot), I_{reg}(T_\Theta(\cdot))) \quad (9)$$

$$= - \sum_{g=1}^G \sum_{k=1}^K p(g, k) \log(p(g, k)) \quad (10)$$

La plus performante des techniques référencées (d'après plusieurs études récentes [34]) est la maximisation de l'information mutuelle proposée dans [16, 33], qui revient à minimiser :

$$E(I_{ref}(\cdot), I_{rec}(T_\Theta(\cdot))) = -\mathcal{I}(I_{ref}(\cdot), I_{reg}(T_\Theta(\cdot))) \quad (11)$$

$$= - \sum_{g=1}^G \sum_{k=1}^K p(g, k) \log \frac{p(g, k)}{p(g)p(k)} \quad (12)$$

où $p(g)$ est la probabilité de l'événement un voxel de l'image de référence a le niveau de gris g , $p(k)$ est la probabilité de l'événement un voxel de l'image à recalcr a le niveau de gris k et $p(g, k)$ est la probabilité de l'événement un voxel de l'image de référence a le niveau de gris g et le même voxel de l'image à recalcr a le niveau de gris k .

Il a été établi que l'information mutuelle fournit une meilleure mesure que l'entropie conjointe [29] et que le critère d'uniformité inter-images de Woods est équivalent au critère de l'information mutuelle sous certaines conditions [16].

La minimisation des fonctions de coût décrites précédemment s'appuie habituellement sur des techniques d'optimisation déterministes (algorithmes du gradient et ses variantes) qui nécessitent une initialisation proche de la solution désirée. Dans notre cas nous utilisons un algorithme d'optimisation stochastique multi-grille, décrit dans la partie 3.3, qui permet dans une large mesure de s'affranchir de cette contrainte.

3. recalage par fonctions de similarité robustes

3.1. motivations

Dans le cas d'images monomodales, la fonction de similarité quadratique (Eq. 1) est optimale lorsque les deux images ne diffèrent, après recalage, que par un bruit blanc gaussien additif. Ses limitations deviennent patentes lorsque les deux images

présentent des dissimilarités importantes ou sont entachées de bruits non gaussiens. Cette remarque est également valable pour le critère d'uniformité inter-image [35] qui s'appuie sur des statistiques non robustes (moyennes, variances) extraites des deux images. Afin de vérifier l'hypothèse sous-jacente à la méthode de Woods *et al.* [35] (critère d'uniformité inter-image), nous avons recalé de façon manuelle une image TEMP du cerveau d'un patient avec un examen IRM du même patient (figure 1) en ayant au préalable extrait le cerveau en IRM et nous avons calculé l'histogramme conjoint des deux images.

Comme on peut le constater sur la figure 2, l'histogramme conjoint ne présente pas un ensemble de pics, comme le voudrait l'hypothèse de l'uniformité inter-image. Ceci apparaît clairement sur la figure 3(a) où l'histogramme monodimensionnel d'une ligne de l'histogramme conjoint (figure. 2) a été tracé. Même lorsque les images sont correctement recalées et malgré le fait que l'image TEMP présente une certaine uniformité des niveaux de gris due à la diffusion de Compton, on remarque que l'histogramme de la

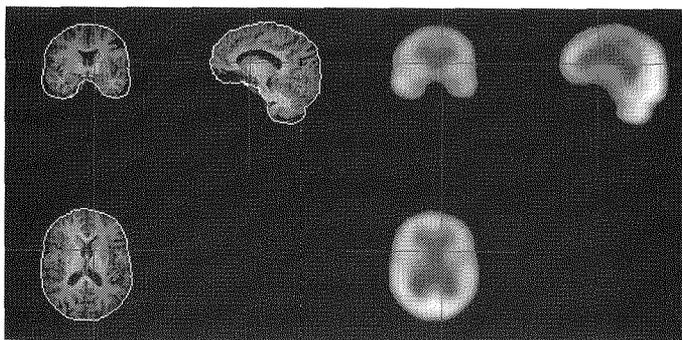


Figure 1. – A droite : image TEMP présentée en coupes multi-planaires. A gauche : image IRM avec, en superposition, le contour extérieur de l'image TEMP après recalage manuel.

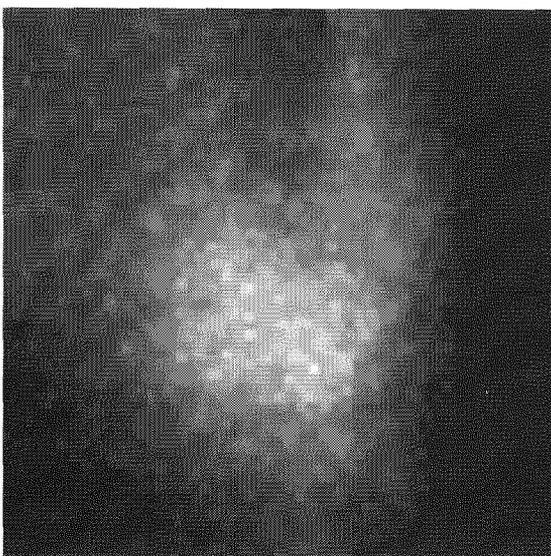


Figure 2. – L'histogramme conjoint du couple d'images de la figure 1. Les niveaux de gris de l'image IRM sont représentés en ordonnée et ceux de l'image TEMP en abscisse.

figure 3(a) diffère du pic supposé par l'hypothèse d'uniformité inter-image (figure. 3(b)). L'histogramme s'apparente plutôt à une loi de mélange (caractéristiques de plusieurs populations). Dans ce type d'histogramme, le calcul de la valeur moyenne et de l'écart-type par les équations (4) et (5) introduit des erreurs importantes en raison du mélange de plusieurs populations.

Ce critère est donc sensible aux différences d'information pouvant apparaître dans un couple d'images multimodales. Lorsqu'un nombre important de points ne suivent pas l'hypothèse d'uniformité inter-image, des recalages imprécis, voire erronés peuvent être obtenus, comme nous le verrons par la suite. Pour améliorer la robustesse du recalage à des données non conformes, la fonction de coût utilisée doit donc tolérer un certain pourcentage de points ne suivant pas le modèle dominant. C'est précisément l'objet des estimateurs robustes.

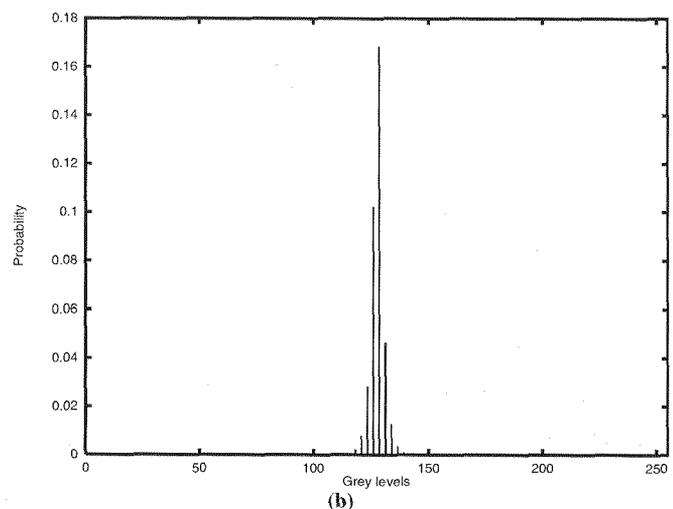
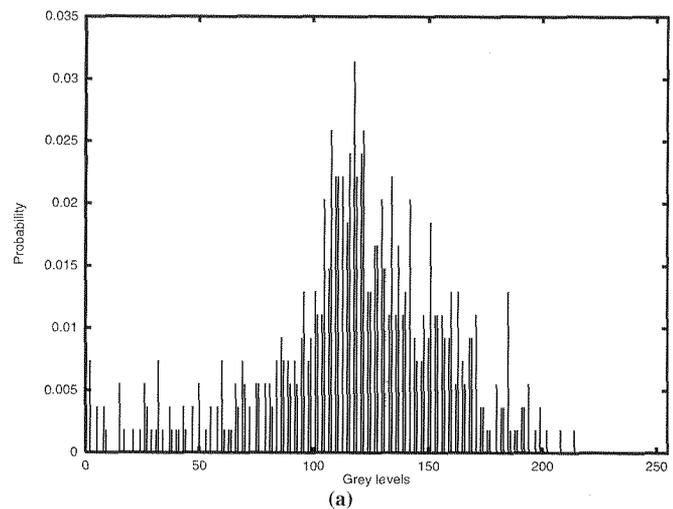


Figure 3. – a) Un exemple d'une ligne de l'histogramme conjoint de la figure 2. b) L'histogramme « théorique » des voxels d'une image recalée dont les voxels correspondant sur l'image de référence ont le même niveau de gris g . L'hypothèse fondamentale de la méthode de Woods est qu'après recalage, tous les voxels de l'image de référence ayant le même niveau de gris g sont représentés par une région uniforme dans l'image recalée (ici le pic).

3.2. fonctions de similarité robustes

L'objet de l'estimation robuste est précisément de permettre une détermination précise des paramètres d'un modèle, en présence de pollution par des données « aberrantes » (« outliers » en anglais), mais néanmoins précieuses ici, puisque ce sont souvent elles qui portent l'information intéressante. Les techniques d'estimation robuste font désormais partie des outils couramment utilisés en analyse d'images et en vision par ordinateur, et nous renvoyons le lecteur à [4, 21, 28] pour une synthèse sur le sujet ainsi qu'une présentation des différentes classes d'estimateurs et de leurs propriétés.

Une caractéristique importante des estimateurs robustes est leur *point de rupture*, qui correspond à la proportion de données pouvant être arbitrairement polluées, sans que cela n'ait d'incidence sur le résultat de l'estimation. L'estimateur quadratique par exemple a un point de rupture égal à 0 puisque une seule mesure qui ne suit pas le modèle peut complètement perturber le résultat. L'estimateur des moindres carrés médians [21], utilisé dans [1] pour le recalage d'images monomodales, est basé sur la minimisation de la valeur médiane des carrés des erreurs résiduelles. Cet estimateur est très performant, puisqu'il est robuste à une contamination des données qui peut atteindre 50%. C'est également le cas de l'algorithme MINPRAN [28]. Ces deux estimateurs appartiennent à la classe des R-estimateurs qui nécessitent un tri des erreurs résiduelles et des techniques d'échantillonnage aléatoire des données. Ceci conduit à des temps de calcul qui sont prohibitifs sur des problèmes de grande taille, tel que le nôtre (rappelons que les données correspondent dans notre cas à l'ensemble des voxels des images volumiques).

Une autre caractéristique importante des estimateurs robustes est leur *complexité algorithmique*. En traitement d'images médicales, où le nombre de voxels à traiter est très élevé, la complexité algorithmique doit être aussi faible que possible. Généralement un gain en robustesse s'accompagne d'une plus grande complexité algorithmique. Il est donc nécessaire de réaliser un compromis entre ces caractéristiques. Les *M-estimateurs*, possèdent des propriétés attractives (points de ruptures relativement élevés et faible sur-coût calculatoire), qui ont favorisé leur large utilisation en analyse d'images [4, 21]. Le M-estimateur conduit à remplacer, dans la fonction de coût à minimiser, la norme d'erreur quadratique (associée au modèle gaussien) par une norme d'erreur ρ robuste (associée à des modèles statistiques non linéaires) [21]. Le principe est de minimiser :

$$\min_{x_i} \sum_i \rho(x_i, C) \quad (13)$$

où C est un paramètre d'échelle. Le comportement d'un M-estimateur est caractérisé par sa fonction d'influence :

$$\psi(x, C) = \frac{d\rho}{dx}(x, C) \quad (14)$$

qui caractérise l'influence des résiduels. On montre [4, 21] que cette technique d'estimation robuste est équivalente à la minimisation d'un critère des moindres carrés pondérés, dont les poids dépendent précisément de l'influence des résiduels. Le point de rupture théorique des M-estimateurs est de $\frac{1}{p+1}$, où p désigne le nombre de paramètres à estimer [21]. Les points de rupture observés en pratique, sont toutefois souvent plus élevés, et Black a par exemple observé, dans des problèmes d'estimation du mouvement apparent, des points de rupture de 35% – 45% pour ce type d'estimateur [4].

Nous avons considéré les fonctions de similarité robustes suivantes pour le recalage d'images monomodales ou multimodales présentant des différences significatives d'information :

– pour des images de même modalité :

$$E(T_\Theta(x)) = \sum_x \rho \{I_{ref}(x) - I_{rec}(T_\Theta(x)), C\} \quad (15)$$

– pour des images multimodales :

$$E(T_\Theta(x)) = \sum_{g=1}^G \frac{N_g}{N} \widetilde{\sigma}_g(T_\Theta(x)) \quad (16)$$

avec :

$$\widetilde{\sigma}_g(T_\Theta(x)) = \sqrt{\sum_{x|I_{ref}(x)=g} \rho \{I_{rec}(T_\Theta(x)) - \widetilde{\mu}_g(T_\Theta(x)), C\}} \quad (17)$$

et :

$$\widetilde{\mu}_g(T_\Theta(x)) = \arg \min_{\mu_g} \sum_{x|I_{ref}(x)=g} \rho \{I_{rec}(T_\Theta(x)) - \mu_g, C\} \quad (18)$$

où rappelons-le, x désigne un point de l'espace 3D, $I_{ref}(x)$ et $I_{rec}(x)$ sont respectivement les valeurs de l'image de référence et de l'image à recalcr, ρ est une norme d'erreur robuste et C est un paramètre d'échelle permettant d'ajuster le seuil de rejet des points ne suivant pas le modèle [4]. G correspond au nombre de niveaux de gris de l'image, N_g est le nombre de voxels ayant le niveau de gris g et N désigne le nombre total de voxels.

La première fonction de similarité (Equ. 15) est simplement une version robuste classique de l'estimateur quadratique. Pour traiter le cas d'images multimodales (Equ. 16), la fonction de similarité robuste est construite sur le même principe que celle de Woods *et al.*, en admettant toutefois des violations de l'hypothèse d'uniformité inter-image. Nous minimisons donc les « variances robustes » $\widetilde{\sigma}_g$ des régions de l'image à recalcr, définies par une partition de l'image de référence en classes uniformes (Equ. 17). Notons que la valeur moyenne des régions, $\widetilde{\mu}_g$ est également calculée par un estimateur robuste (Equ. 18)). Le fait de remplacer les statistiques classiques, par des statistiques robustes permet de traiter des situations où l'hypothèse d'uniformité inter-image est inexacte pour une population significative de points. Notons que

si l'on adopte

$$\rho(x, C) = \frac{x^2}{2C^2} \quad (19)$$

on retrouve, comme attendu, des variantes des estimateurs non robustes présentés dans la section précédente.

Le calcul de la valeur moyenne robuste $\tilde{\mu}_g$ dans l'équation (18) peut être mené en évaluant la fonction robuste pour chaque valeur entière de μ_g et en définissant $\tilde{\mu}_g$ comme la valeur minimisant (18). Pour accélérer l'algorithme, la valeur μ_g peut être approchée par la valeur médiane dans chacun des G histogrammes. Dans tous les cas, une perte de précision est inévitable mais elle ne perturbe pas la solution finale.

Parmi les M-estimateurs proposés dans la littérature [4], nous avons comparé trois estimateurs robustes (l'estimateur de Geman-McClure, l'estimateur de Tukey et l'estimateur quadratique tronqué). Les meilleurs résultats ont été obtenus avec l'estimateur robuste de Geman-McClure, également utilisé comme fonction de potentiel dans les modèles statistiques markoviens non linéaires. Cet estimateur est défini par (figure 4) :

$$\rho(x, C) = \frac{x^2}{x^2 + C^2} \quad (20)$$

L'allure de la fonction $\rho(x, C)$ permet de diminuer l'influence des points associés à des erreurs résiduelles importantes. La fonction d'influence de l'estimateur de Geman-McClure (figure 4) :

$$\psi(x, C) = \frac{2xC^2}{(x^2 + C^2)^2} \quad (21)$$

caractérise l'influence des erreurs résiduelles.

On constate qualitativement que lorsque l'amplitude des erreurs résiduelles augmente, leur influence sur la solution commence à diminuer, alors que la valeur de $\rho(x, C)$ devient approximativement constante. Le paramètre C détermine le point au-delà duquel une mesure est considérée comme aberrante. Ainsi pour l'estimateur (15), les données pour lesquelles :

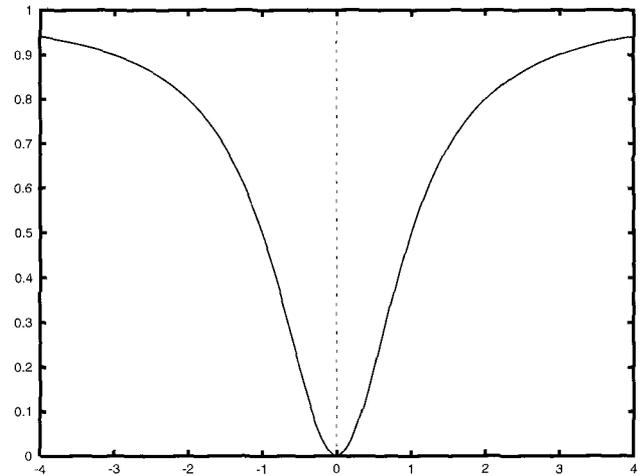
$$|I_{rec}(T_{\Theta}(x)) - I_{ref}(x)| \geq \frac{C}{\sqrt{3}} \quad (22)$$

sont considérées comme aberrantes car le rejet commence pour

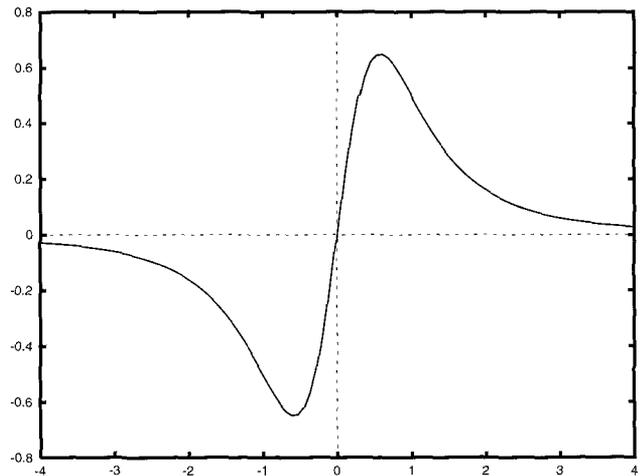
$$\frac{d^2\rho}{dx^2} = \frac{2C^2(x^2 + C^2)(C^2 - 3x^2)}{(x^2 + C^2)^4} = 0 \quad (23)$$

Le paramètre C qui fixe le point de rejet des données peut être calculé de différentes façons :

- en l'identifiant à la variance du bruit, estimée sur une région d'intérêt appartenant au fond de l'image médicale. On peut ensuite le diminuer au fur et à mesure des itérations;
- en l'optimisant simultanément avec les paramètres de la transformation rigide;
- en le calculant en fonction des erreurs résiduelles;



$$\rho(x, C) = \frac{x^2}{x^2 + C^2}$$



$$\psi(x, C) = \frac{2xC^2}{(x^2 + C^2)^2}$$

Figure 4. - Estimateur robuste de Geman-McClure : fonction ρ (en haut) et fonction d'influence ψ (en bas).

La première technique est la moins coûteuse. La deuxième approche est plus précise mais introduit un coût calculatoire supplémentaire qui n'est pas justifié par un gain au niveau de résultats. La troisième stratégie peut se décliner des deux façons différentes pour une itération de l'algorithme :

- en l'identifiant à l'écart-type des poids des erreurs résiduelles [27] :

$$C = \frac{1}{n} \sum_{i=1}^n \left(\frac{\psi(x_i)}{x_i} - \mu_r \right)^2 \quad (24)$$

où

$$\mu_r = \frac{1}{n} \sum_{i=1}^n \frac{\psi(x_i)}{x_i} \quad (25)$$

– en le définissant comme la valeur médiane des valeurs absolues des erreurs résiduelles :

$$C = 1,4826 \operatorname{median}_i(|x_i|) \quad (26)$$

– L'expression (26) vient du fait que la valeur médiane des valeurs absolues d'un large ensemble d'échantillons suivant une loi normale de valeur moyenne nulle et d'écart-type 1 est $\frac{1}{1,4826} = 0,6745$ [27].

Nous avons développé les deux premières méthodes mais privilégié la première car elle est beaucoup plus rapide.

3.3. description de l'algorithme de recalage

Soulignons que les fonctions de coût à minimiser sont fortement non convexes, en raison d'une dépendance implicite par rapport aux paramètres de recalage. Ceci est vrai également pour l'estimateur quadratique (non robuste) qui sera considéré dans la suite à titre de comparaison. Afin d'éviter une initialisation manuelle, tout en se prémunissant des minima locaux, la fonction de coût est minimisée par un algorithme d'optimisation stochastique multigrille décrit dans cette section.

L'algorithme de recalage s'appuie sur un schéma d'optimisation multigrille de type descendant (partant d'une grille grossière pour aller vers des grilles de résolutions croissantes). L'image volumique de référence est ainsi sous-échantillonnée régulièrement d'un facteur 81, 27, 9, 3 et 1 (grille initiale). L'optimisation est menée successivement sur les différentes grilles réduites. Notons qu'aucun sous-échantillonnage n'est appliqué à l'image à recalier, pour conserver toute l'information utile pour l'interpolation. Les paramètres rigides obtenus, après optimisation sur une grille réduite, sont transmis à la grille adjacente plus fine.

L'optimisation multigrille est usuellement motivée par le gain calculatoire qu'elle apporte. Elle conduit également à des algorithmes significativement moins sensibles aux minima locaux de la fonction de coût initiale [10]. Dans notre cas, pour assurer une bonne robustesse aux minima locaux, et permettre ainsi un recalage entièrement non supervisé, nous avons discrétisé l'espace des paramètres de recalage rigide et appliqué un algorithme d'optimisation stochastique, à chaque échelle. L'algorithme d'optimisation s'appuie sur un échantillonneur de Gibbs [7], associé à une descente en température exponentielle. L'excursion dans l'espace des configurations visitées est progressivement réduite lorsque l'on se déplace vers les grilles de résolution croissante, ce qui permet de ne pas remettre entièrement en cause les solutions obtenues aux échelles grossières (les premières grilles visitées conduisent rapidement à des solutions très proches de la solution désirée). Par ailleurs la discrétisation des paramètres est de plus en plus fine lorsque l'on va vers les résolutions croissantes. Pour accélérer la convergence, l'algorithme peut être initialisé par un

recalage préalable des centres de gravité des structures d'intérêt des deux images (il s'agit ici d'images du cerveau humain) ou par un recalage de leurs axes principaux d'inertie. En pratique, le recalage des centres de gravité, qui fournit une estimée grossière des paramètres de translation 3D est suffisant. Nous n'avons volontairement pas eu recours ici au recalage préalable par les axes d'inertie, pour montrer la robustesse de l'approche. Un tel recalage n'est par ailleurs pas toujours possible sur des structures anatomiques.

Les étapes de l'algorithme se décrivent comme suit :

Prétraitement

1. Segmentation rapide de la structure d'intérêt (ici le cerveau), par rapport au fond de l'image.
2. Normalisation en moyenne des deux images.
3. Recalage initial des barycentres des deux volumes pour fournir une première estimée des paramètres de recalage.

Pour chaque grille de résolution

1. Recalage sur l'image de référence par optimisation stochastique rapide de la fonction de similarité robuste.
2. Transmission des paramètres obtenus après convergence vers la grille plus fine. Réduction de l'espace de configurations des paramètres de recalage rigide, raffinement de la discrétisation des paramètres.

La séparation de l'objet d'intérêt du fond de l'image est réalisée par une technique de seuillage et de croissance de région pour le volume TEMP. Pour la segmentation du cerveau en IRM, l'algorithme décrit dans [22] a été utilisé.

Les niveaux de gris de l'image de référence sont simplement normalisés par rapport aux niveaux de gris de l'image à recalier, suivant :

$$I_{ref}(x) = \frac{\bar{I}_{rec}}{\bar{I}_{ref}} I_{ref}(x) , \quad \forall x$$

où \bar{I}_{ref} et \bar{I}_{rec} sont les valeurs moyennes des deux images. Le nombre de niveaux de gris des deux images est également normalisé et réduit à $G = 256$.

Notons qu'un grand nombre d'interpolations 3D est nécessaire pour mener à bien le processus de recalage. La précision finalement obtenue pour les paramètres de rotation et de translation 3D est fortement conditionnée par la qualité de l'interpolation utilisée. Des approches simples, comme l'interpolation au plus proche voisin, sont classiquement utilisées, en raison de leur faible coût calculatoire, bien qu'elles produisent des artefacts importants. Des résultats plus satisfaisants peuvent être obtenus par interpolation trilineaire, par l'utilisation de fonctions splines, ou de noyaux de convolution plus complexes [5]. Nous avons, à titre de compromis, tout d'abord utilisé une interpolation rapide par le plus proche voisin sur les grilles grossières, pour enfin raffiner la solution au niveau le plus fin par une interpolation trilineaire qui préserve mieux la qualité des images à recalier. Cette approche a

conduit à des précisions très satisfaisantes (*i.e.*, sous-voxel, dans tous les cas où les paramètres de recalage étaient connus).

Nous analysons dans la section suivante les performances des méthodes robustes, par rapport à celles des principales approches classiques, présentées précédemment.

4. résultats expérimentaux

Nous avons comparé l'approche des moindres carrés robuste (MCR) (Equ. 15) et la fonction de coût robuste traduisant l'uniformité inter-image (UIR) (Equ. 16) aux fonctions de similarité classiques : norme quadratique (moindres carrés : MC) (Equ. 1), fonction de similarité de Woods *et al.* non robuste reposant sur l'uniformité inter-image (UI) (Equ. 3). Nous avons également implanté le critère de l'information mutuelle (IM) [16,33], considéré actuellement comme présentant les « meilleures » performances [34]. Une évaluation sur la base d'images de l'Université de Vanderbilt [34] est par ailleurs proposée. L'ensemble des développements, des calculs et l'affichage des résultats a été réalisé sur une station de travail Hewlett Packard 9000/C200 (256 MO RAM), sous l'environnement logiciel de traitement d'images médicales 3D MED-IMAX, développé à l'IPB. Les algorithmes ont été écrits en langage C et le programme s'exécute sous le système d'exploitation Unix en utilisant l'interface graphique standard X11/R6. L'environnement logiciel MEDIMAX est présenté et peut être obtenu gracieusement, à des fins de recherche, sur le serveur web de l'IPB à l'adresse <http://alsace.u-strasbg.fr>.

4.1. acquisition et traitement des images

Afin de comparer les différentes fonctions de similarité proposées, les données suivantes ont été acquises à l'IPB :

- des images IRM 2D d'un objet test, ainsi que des images IRM 2D de patients atteints de sclérose en plaques ont été acquises sur un système Bruker de 0,28 Tesla. Les images IRM ont été obtenues en utilisant une séquence multi-coupes, multi-échocs (TE = 15 ms, TR = 2000 ms). La dimension des images est de 256 × 256 (FOV = 25,6 cm), la taille des pixels de 1 mm × 1 mm avec une épaisseur de coupe de 5 mm.
- des images IRM 3D de patients souffrant d'épilepsie partielle ont été acquises sur un système Bruker 2 Tesla, en utilisant une séquence écho de gradient (TE = 40 ms, TR = 8,5 ms, angle = 45 deg). La dimension des images est 128 × 128 × 128 (FOV = 25,6 cm), et la taille des voxels 2 mm × 2 mm × 2 mm.
- des images TEMP ont été acquises par une caméra Elscint Helix deux têtes, avec des collimateurs à basse énergie et haute

résolution, en utilisant les isotopes radioactifs 700MBq de $^{99m}\text{Tc} - \text{HMPAO}$ ou $^{99m}\text{Tc} - \text{ECD}$. Les projections, au nombre de 120, ont été obtenues avec des intervalles de 3 deg, pour des matrices de taille 64 × 64.

La comparaison des algorithmes de recalage a été obtenue à partir de deux ensembles d'images IRM dont la rotation et la translation 3D étaient connues. La translation a été obtenue en modifiant la fréquence de démodulation de l'acquisition, et la rotation en changeant la direction du gradient de lecture.

4.2. images de même modalité

4.2.1. données « aberrantes » simulées

Des expériences de recalage ont été réalisées dans ce cas, à la fois sur des images 2D et 3D. Une première série d'expériences a consisté à appliquer une transformation connue à une image IRM 2D ou 3D, servant de référence. 25% de l'image ainsi transformée a ensuite été dégradée par du bruit poivre et sel, pour simuler la présence de données aberrantes. Ce pourcentage de données erronées correspond approximativement aux cas réels les plus défavorables susceptibles d'être rencontrés en pratique. Les paramètres de recalage (connus dans ce cas) ont été comparés aux paramètres estimés pour juger de la précision du recalage. La moyenne et l'écart-type de l'erreur de recalage ont ainsi été calculés, pour chaque méthode, sur 20 transformations différentes, correspondant à des translation entre -20 et +20 voxels et à des rotations entre -30 et +30 degrés.

Le tableau 1 présente des résultats pour des images IRM 2D et le tableau 2 pour des images IRM 3D. Les algorithmes robustes MCR et UIR permettent d'atteindre une précision sous-voxel dans le recalage, contrairement aux approches non robustes (MC ou UI). La technique de maximisation de l'information mutuelle (IM) fournit également de bons résultats, quoique ses performances soient légèrement inférieures à celle de l'estimateur robuste MCR.

La figure 5(c) présente un exemple pour lequel l'estimateur des moindres carrés (MC) échoue dans le recalage des coupes

Tableau 1. – Recalage d'images 2D de même modalité. Une image IRM 2D a subi 20 transformations rigides, avec différentes valeurs de rotation et de translation. 25% des pixels ont été dégradés par du bruit poivre et sel (voir figure 5). Le tableau donne la moyenne et l'écart-type de l'erreur de recalage exprimée en pixels pour la translation et en degrés pour la rotation.

Recalage IRM/IRM 2D			
	Δt_x	Δt_y	$\Delta \theta$
MC	0,77 ± 0,44	0,99 ± 1,09	3,90 ± 2,87
UI	2,12 ± 3,25	2,54 ± 1,91	4,34 ± 2,25
IM	0,29 ± 0,22	0,19 ± 0,12	0,15 ± 0,14
MCR	0,27 ± 0,26	0,20 ± 0,07	0,14 ± 0,77
UIR	0,18 ± 0,12	0,29 ± 0,23	0,66 ± 0,67

Tableau 2. – Recalage d'images 3D de même modalité. Une image IRM 3D a subi 20 transformations rigides, avec différentes valeurs de rotation et de translation. 25% des voxels ont été dégradés par du bruit poivre et sel. Le tableau donne la moyenne et l'écart-type de l'erreur de recalage exprimée en voxels pour la translation et en degrés pour la rotation.

Recalage IRM/IRM 3D					
	MC	UI	IM	MCR	UIR
Δt_x	2,30 ± 1,75	1,49 ± 1,40	0,05 ± 0,06	0,04 ± 0,07	0,09 ± 0,05
Δt_y	2,53 ± 1,56	1,56 ± 1,41	0,22 ± 0,15	0,16 ± 0,11	0,18 ± 0,14
Δt_z	2,77 ± 1,83	1,93 ± 1,63	0,09 ± 0,14	0,06 ± 0,10	0,10 ± 0,05
$\Delta \theta_x$	4,71 ± 2,89	3,75 ± 2,03	0,35 ± 0,35	0,41 ± 0,21	0,22 ± 0,34
$\Delta \theta_y$	5,33 ± 3,40	3,65 ± 2,54	0,27 ± 0,32	0,16 ± 0,22	0,24 ± 0,17
$\Delta \theta_z$	5,05 ± 3,51	2,99 ± 3,06	0,44 ± 0,69	0,33 ± 0,24	0,40 ± 0,59

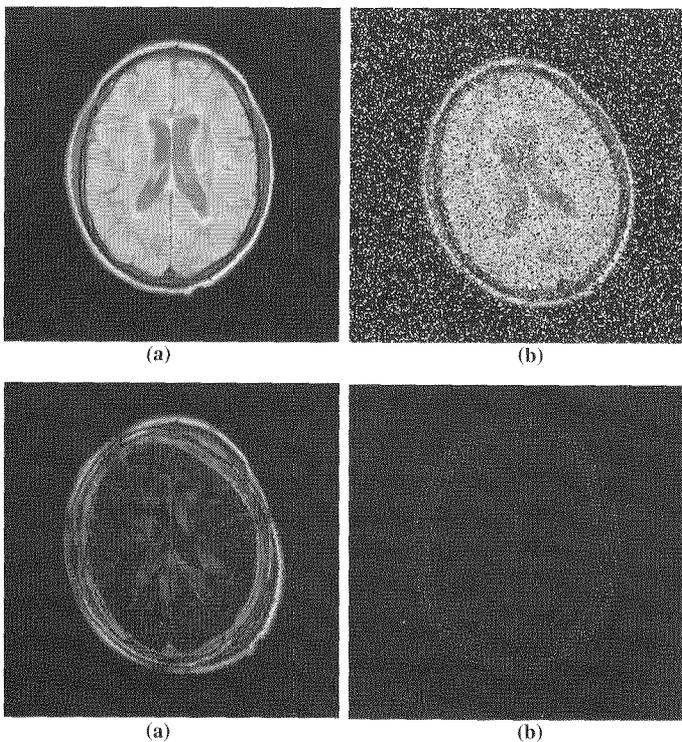


Figure 5. – Recalage robuste d'images 2D de même modalité. (a) Image de référence. (b) Image transformée (rotation : 20 deg, translation : 10 voxels, dégradation par du bruit poivre et sel). (c) Erreur de recalage pour l'estimateur quadratique classique. (d) Erreur de recalage pour l'estimateur robuste de Geman-McClure

IRM présentées figures 5(a) et 5(b), contrairement à l'estimateur robuste qui réalise un recalage très précis (figure 5(d)) Notons que les normes des erreurs de recalage présentées figures 5(c) et 5(d) ont été amplifiées pour la visualisation.

4.2.2. image d'objet test

Ces performances ont été confirmées en appliquant l'algorithme au recalage d'images IRM 2D d'un objet test réel, pour lequel des translations et des rotations ont été réalisées directement par modification des gradients de lecture et de la fréquence de démodulation de l'acquisition, comme expliqué précédemment.

Tableau 3. – Recalage d'un objet-test 2D. Des images IRM 2D d'un objet-test ont été acquises en simulant, par le système d'acquisition, différentes transformations rigides. Le tableau donne la moyenne et l'écart-type de l'erreur de recalage exprimée en voxels pour la translation et en degrés pour la rotation.

Recalage objet test IRM 2D			
	Δt_x	Δt_y	$\Delta \theta$
MC	0,29 ± 0,22	0,18 ± 0,18	0,26 ± 0,25
UI	0,43 ± 0,51	0,45 ± 0,88	0,75 ± 0,71
IM	0,25 ± 0,15	0,22 ± 0,29	0,29 ± 0,30
MCR	0,23 ± 0,20	0,19 ± 0,21	0,17 ± 0,23
UIR	0,33 ± 0,24	0,27 ± 0,29	0,27 ± 0,27

Le tableau 3 présente les erreurs de recalage mesurées dans ce cas. En l'absence de bruit significatif et d'évolution de lésion, toutes les techniques de recalage atteignent une précision sous-voxel, mais comme on peut le constater c'est encore l'estimateur des moindres carrés robustes qui conduit à la meilleure précision. Il apparaît donc comme un bon choix pour le recalage d'images de mêmes modalités. Notons également ici que les approches s'appuyant sur la maximisation de l'uniformité inter-image (UI et UIR) apparaissent moins adaptées. Rappelons que ces méthodes ont été conçues pour le recalage d'images multimodales, et qu'elles ne peuvent, s'appuyant sur des statistiques globales calculées sur des régions, atteindre les mêmes précisions que les méthodes MC et MCR.

4.2.3. application clinique

Nous présentons enfin un exemple de recalage d'une séquence de coupes IRM d'un patient souffrant de sclérose en plaques, présentant des évolutions de lésion au cours du temps (figure 6(a)-(b)) [26]. L'estimateur quadratique classique MC n'est pas en mesure, dans ce cas, de fournir un recalage suffisamment précis pour faire ressortir de façon satisfaisante l'évolution des lésions dans l'image des erreurs résiduelles seuillée (figure 6(c)). L'algorithme robuste fournit dans ce cas un recalage de précision sous-voxel permettant une meilleure détermination de l'évolution des lésions par simple seuillage de l'image des erreurs résiduelles (figure 6(d)). Un examen de ces résultats par des médecins spécialistes a permis de confirmer cette interprétation.

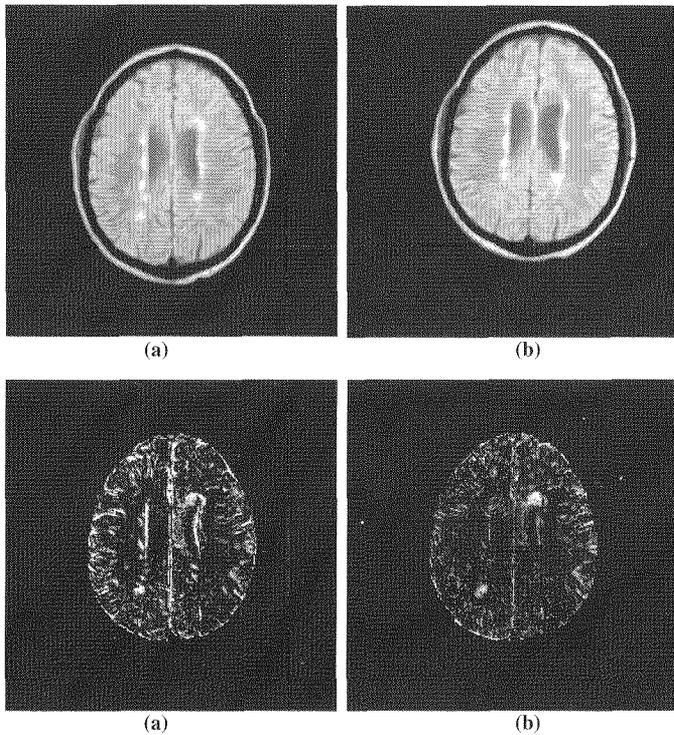


Figure 6. – (a) Image IRM 2D d'un patient atteint de sclérose en plaques. (b) Image du même patient acquise quelques mois plus tard. (c) Norme de l'erreur de recalage de l'image 6-b par rapport à l'image 6-a obtenue par l'estimateur non robuste des moindres carrés (MC). (d) Erreur de recalage de l'image 6-b par rapport à l'image 6-a en utilisant l'estimateur robuste de Geman-McClure (MCR). L'image des erreurs met clairement en évidence une évolution des lésions.

4.3. images multimodales

4.3.1. données « aberrantes » simulées

Pour évaluer les algorithmes de recalage d'images multimodales, une image volumique TEMP du cerveau a été recalée manuellement, avec l'aide d'un médecin, sur l'image IRM correspondante. L'image TEMP ainsi obtenue a ensuite subi une transformation rigide, en utilisant les mêmes paramètres de rotation et de translation 3D que ceux utilisés dans les expériences simulées sur les images de même modalité. Nous avons de même corrompu 25% des données de l'image TEMP par du bruit poivre et sel pour simuler des différences importantes d'information. La technique robuste de maximisation de l'uniformité inter-image (UIR) a été comparée, dans ce cas, à la technique non robuste UI [35] et à l'approche par maximisation de l'information mutuelle IM [33]. Le tableau 4 illustre la robustesse de l'estimateur UIR aux données aberrantes. L'erreur obtenue pour la méthode UIR est de l'ordre du voxel pour la translation et de l'ordre du degré pour la rotation, conduisant à une précision significativement supérieure à celle de la méthode UI. Nous remarquons également les assez bonnes performances de la méthode IM, qui fournit des résultats toutefois inférieurs à ceux de l'estimateur robuste UIR.

Tableau 4. – Recalage d'images volumiques IRM/TEMP. Une image TEMP 3D préalablement recalée de façon manuelle avec l'IRM par un neurologue a subi 20 transformations rigides, avec différentes valeurs de rotation et de translation. 25% des voxels de l'image ont été dégradés par du bruit poivre et sel. Le tableau donne la moyenne et l'écart-type de l'erreur de recalage exprimée en voxels pour la translation et en degrés pour la rotation.

Recalage IRM/TEMP 3D			
	UI	IM	UIR
Δt_x	$3,85 \pm 5,59$	$1,41 \pm 0,74$	$0,82 \pm 0,53$
Δt_y	$3,02 \pm 4,78$	$1,38 \pm 1,23$	$0,61 \pm 0,50$
Δt_z	$4,16 \pm 4,38$	$2,06 \pm 1,29$	$0,83 \pm 0,60$
$\Delta \theta_x$	$8,33 \pm 4,51$	$0,94 \pm 1,58$	$0,21 \pm 0,48$
$\Delta \theta_y$	$6,23 \pm 3,52$	$1,04 \pm 1,15$	$1,14 \pm 0,26$
$\Delta \theta_z$	$6,80 \pm 4,15$	$1,36 \pm 4,77$	$0,71 \pm 0,94$

4.3.2. performances sur des images réelles présentant des structures non modélisées

En général, le recalage d'une image fonctionnelle (TEMP, TEP) sur l'IRM correspondante, est réalisé avec segmentation du cerveau en IRM car les structures n'appartenant pas au cerveau (os, peau) ne sont pas présentes sur l'image fonctionnelle et introduisent, dans la plupart des cas, des erreurs importantes de recalage. Dans la littérature, la plupart de méthodes de recalage nécessitent donc un prétraitement qui supprime les structures non cérébrales. C'est le choix qui a été fait systématiquement dans l'étude comparative de West *et al.* [34]. En routine clinique, ce prétraitement est contraignant, le résultat de la segmentation du cerveau n'est pas toujours parfait et il dépend souvent de l'opérateur. Ainsi, l'utilisation d'un algorithme de recalage ne nécessitant aucun prétraitement permettrait d'automatiser entièrement la procédure de recalage et apporterait un gain de temps très appréciable à l'utilisateur.

L'algorithme original basé sur l'uniformité inter-images (UI) n'arrive en aucun cas à gérer des couples IRM/TEMP sans prétraitement. La mesure de l'information mutuelle en théorie est capable de le faire mais la pratique a montré que pour atteindre un résultat acceptable l'algorithme doit être initialisé près de la solution [16]. Avec l'utilisation d'estimateurs robustes, nous avons constaté qu'il est possible de recalcer des images multimodales (IRM/TEMP), sans segmentation préalable du cerveau en IRM, même si elles présentent un décalage initial important. Ceci est dû d'une part à l'estimateur robuste et d'autre part à la procédure d'optimisation qui parcourt un grand espace de recherche. Nos expériences ont montré que 70% des couples d'images non segmentées préalablement peuvent être recalés de façon satisfaisante. Un cas typique de cette situation est illustré par la figure 7 où on peut voir que malgré le fait que les images sont relativement proches, la fonction de similarité UI ne permet pas de recalcer correctement les deux images (figure 7(a)). L'information mutuelle est plus proche d'un résultat satisfaisant mais un décalage résiduel est visible (figure 7(b)). Finalement, le meilleur résultat est obtenu par la mesure UIR (figure 7(c)).

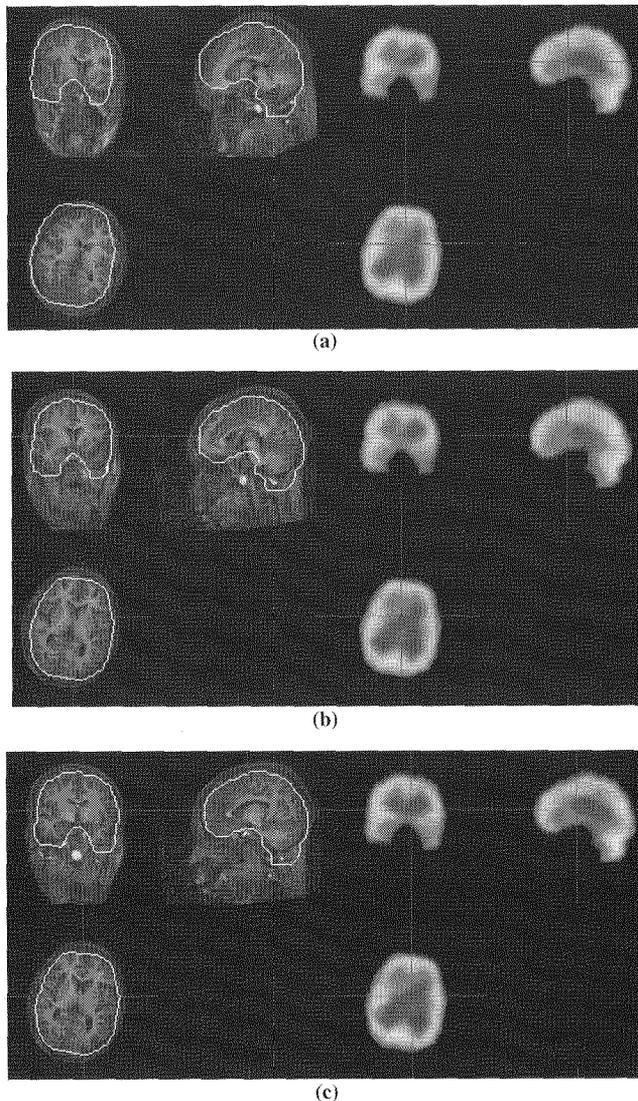


Figure 7. – Recalage IRM/TEMP sans segmentation préalable du cerveau en IRM. (a) Approche UI. (b) Approche IM. (c) Approche UIR.

4.3.3. application clinique

Nous présentons enfin (figure 8) le résultat du recalage, par l'approche robuste, d'images fonctionnelles obtenues en TEMP sur une image morphologique RM d'un patient, réalisée dans le cadre d'un protocole clinique concernant des patients souffrant d'épilepsie.

L'image IRM, servant de référence, porte l'information anatomique. Deux images TEMP ont été réalisées chez le même patient, une image critique et une inter-critique. L'image TEMP critique, réalisée pendant la crise d'épilepsie, donne des informations de perfusion sur l'évolution d'une crise d'épilepsie. L'image TEMP intercritique fournit des informations sur l'état normal (en dehors de la crise). L'image TEMP critique plus l'image TEMP inter-critique ont été recalées sur l'image IRM, mettant ainsi en

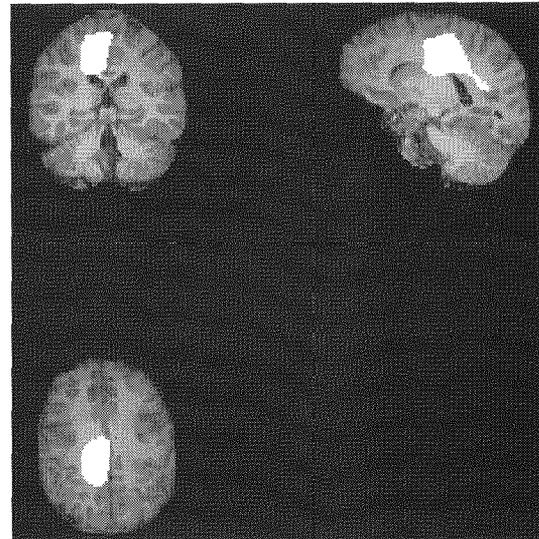


Figure 8. – Superposition des images volumiques IRM/TEMP d'un patient présentant une épilepsie partielle complexe d'origine temporale droite : la soustraction des images critique et intercritique TEMP, superposée à l'IRM correspondante par recalage robuste, montre la zone irritative caractérisée par une hyperperfusion locale.

correspondance informations anatomiques et fonctionnelles. La différence entre les images TEMP critique et inter-critique permet de connaître la zone irritative. La région d'activité hyperintense (zone irritative) obtenue par simple différence des images TEMP recalées a enfin été superposée à l'information anatomique de l'image IRM (figure 8) pour permettre au médecin de localiser avec une plus grande précision la zone irritative, ce qui sans le recalage des images n'était pas possible. Le résultat a été validé par des médecins spécialistes sur des critères cliniques et SEEG. Plus de 50 cas ont déjà été traité dans le cadre de ce protocole.

4.4. évaluation sur la base d'images de l'université de Vanderbilt

Nous avons procédé à une évaluation complète de l'algorithme de recalage d'images multimodales robuste sur la base d'images de l'Université de Vanderbilt, Nashville, Tennessee, USA dans le cadre du projet « Evaluation of Retrospective Image Registration » auquel ont déjà contribué 20 équipes de par le monde. Il s'agit d'un projet qui a été mis en place en 1995 par le Dr M. J. Fitzpatrick afin de comparer et d'évaluer les algorithmes de recalage rigide d'images multimodales du cerveau.

Le projet consiste à recalculer un ensemble d'images multimodales 3D de 7 patients comprenant les modalités suivantes : IRM-T1 (temps de relaxation T1), IRM-T2 (temps de relaxation T2), IRM-DP (densité de protons), IRM-T1 rectifiée, IRM-T2 rectifiée, IRM-DP rectifiée, CT et TEP. Le terme *rectifié* signifie que les images ont été transformées afin de corriger la distorsion géométrique due à l'acquisition.

Les images CT et TEP ont été recalées par rapport à chaque image IRM à l'aide de marqueurs externes par les responsables du projet d'évaluation, pour établir une « vérité terrain ». Les véritables paramètres de recalage ne sont pas diffusés aux équipes participant au projet [34]. Seuls sont connus, après évaluation, les erreurs de recalage, par rapport à des régions d'intérêt dans l'image. Les traces de marqueurs ont soigneusement été effacées et remplacées par du bruit similaire au bruit de l'acquisition pour qu'elles ne soient pas détectables. Toute équipe participant au projet doit transmettre les résultats de ses recalages avec un certain protocole pour faciliter leur évaluation. Ce protocole consiste à calculer la transformation rigide associée à chaque recalage et à l'appliquer aux huit points définissant les sommets du cube de l'image à recalé. Les nouvelles coordonnées des sommets du cube sont transmises aux responsables du projet afin d'évaluer la précision du recalage. Pour cela, les responsables du projet appliquent la transformation rigide qui leur a été communiquée à des volumes d'intérêt (VOI) prédéfinis dans les images et calculent l'erreur de la position du barycentre de ces volumes d'intérêt par rapport à la vérité terrain. Le nombre de VOI varie de 6 à 10 selon le cas. Les détails concernant les acquisitions, les recalages manuels initiaux et la précision sur le calcul de l'erreur de recalage sont décrits dans [34]. Les principales caractéristiques des images à recalé sont résumées dans le tableau 5.

Nous avons évalué deux variantes de notre algorithme robuste reposant sur l'uniformité inter images pour le recalage d'images CT/IRM et TEP/IRM. Dans la première variante (NI1), l'image de référence (chaque IRM) est interpolée afin que ses voxels aient la même taille dans les trois directions de l'espace (voxel isotrope). Ensuite, l'image à recalé (CT ou TEP) a également été interpolée pour qu'elle ait la même résolution que l'image de référence. Dans la deuxième variante (NI2), l'image à recalé a simplement été interpolée pour qu'elle ait la même résolution que l'image de référence sans homogénéiser les voxels des deux images (voxel anisotrope) pour éviter les minima locaux dus à la double interpolation.

Nous reproduisons dans cette section, à titre de comparaison, les tableaux publiés dans [34] et présentant les résultats des autres méthodes évaluées (à la date de la publication). Les tableaux 6, 8, 10 et 12 donnent l'erreur médiane et maximale pour les recalages CT/IRM et TEP/IRM pour les équipes ayant participé au projet initial. Soulignons que depuis la première évaluation, d'autres équipes ont fait évaluer leurs résultats par l'Université de Vanderbilt. Les tableaux suivants ne sont donc pas exhaustifs. Le lecteur est invité à se reporter au site web du projet pour des résultats complets qui sont régulièrement mis à jour. Les abréviations concernant les équipes et les variantes des méthodes utilisées sont les mêmes que celles dans la publication d'origine : BA : Barillot *et al.* [15], CO : Collignon *et al.* [6, 16], EL : Van den Elsen *et al.* [31], HA : Harkness *et al.* [24], HE : Hemler *et al.* [11], HI : Hill *et al.* [13, 29], MAI : Maintz *et al.* [18], MAL : Malandain *et al.* [20], NO : Noz *et al.* [17], PE : Pelizzari *et al.* [24], RO : Robb *et al.* [14], WO : Woods *et al.* [35].

Tableau 5. – Les caractéristiques des images fournies par l'Université de Vanderbilt. Les niveaux de gris de toutes les images sont codés sur 16 bit.

Modalité	Nombre des voxels			Dimension de voxels (mm)		
	<i>x</i>	<i>y</i>	<i>z</i>	<i>x</i>	<i>y</i>	<i>z</i>
IRM	256	256	20-26	1,25-1,28	1,25-1,28	4,0
CT	512	512	28-34	0,65	0,65	4,0
TEP	128	128	15	2,59	2,59	8,0

Dans le cadre de notre évaluation, pour le recalage CT/IRM, le crâne a été segmenté du fond par simple seuillage, dans les deux modalités, et l'algorithme décrit dans le chapitre précédent a été appliqué. Notons que nous avons été amené à réduire la taille de l'image CT de 512×512 à 256×256 , pour des raisons techniques (la gestion de la mémoire ne permet pas de traiter des images 512×512). Ceci a bien entendu une incidence sur la précision du recalage. Nous avons effectué 41 recalages CT/IRM pour les 7 patients dont les images CT et IRM étaient disponibles. Les résultats de nos recalages sont résumés dans les tableaux 7 et 9 où l'erreur médiane et maximale sont présentées pour chaque recalage du volume CT vers la modalité IRM correspondante. Soulignons que ces erreurs correspondent aux distances par rapport aux positions des barycentres des VOI (6 à 10 par couple d'images).

En comparant (tableaux 7 et 9) les deux variantes de notre méthode par rapport aux méthodes présentées dans les tableaux 6 et 8 nous constatons que nos techniques donnent des résultats satisfaisants. Les erreurs médianes varient de 1,6 à 2,9 mm, ce qui situe nos performances entre le milieu et le haut du classement et les erreurs maximales sont comprises entre 3,8 et 6,4 mm ce qui classe nos méthodes vers le haut du classement. Dans tous les cas, l'erreur médiane de notre technique est inférieure à la médiane des médianes des erreurs fournies par les autres équipes. Les faibles erreurs maximales montrent la robustesse de notre approche.

En ce qui concerne le recalage TEP/IRM, le crâne en IRM et le cerveau en TEP ont été segmentés de leurs fonds par simple seuillage. *Nous n'avons pas, contrairement aux autres équipes, segmenté le cerveau en IRM* et ce pour deux raisons. Premièrement, parce que cette tâche nécessite un travail important, délicat et difficilement reproductible et deuxièmement car nous souhaitons vérifier la robustesse de l'algorithme aux données manquantes (crâne apparaissant en IRM et non représenté en TEP). Nous avons effectué 35 recalages de ce type pour les 7 patients dont les images TEP étaient disponibles. Les résultats de nos recalages sont résumés dans les tableaux 11 et 13 où l'erreur médiane et maximale sont présentées pour chaque recalage du volume TEP vers la modalité IRM correspondante ainsi que des statistiques comparant les performances de notre approche par rapport aux autres méthodes (tableau 10 et 12).

En termes d'erreur médiane, nos techniques obtiennent des erreurs entre 1,9 et 5,1 mm, ce qui les classe dans la première moitié des méthodes. La plus grande erreur médiane observée est pour le recalage d'images TEP/IRM-T1 (5,1 mm dernière

Tableau 6. – Erreurs médianes des différentes techniques évaluées par le projet initial (d'après [34]). L'étiquette « r » signifie que les images ont été corrigées pour la distortion géométrique avant recalage. * Un patient a été omis. † Transformations non-rigides. ‡ Résultats resoumis après communication de la vérité terrain. Toutes les erreurs sont exprimées en mm.

Erreurs médianes pour le recalage CT/IRM														
Mod.	Technique													
IRM	BA	CO	EL	HA	HE	HI [‡]	MAI	MAL	NO [†]	PE	RO1	RO2	RO3	RO4
T1	1,6	1,5	1,6	3,4	1,4	1,2	5,1	4,3	3,3	2,7	4,2	5,2	5,7	5,4
DP	1,9	1,5	2,0	3,1	2,4	1,9	4,1	4,0	7,8	1,9	4,5	5,5	4,9	4,8
T2	2,5	1,5	1,6	4,2	4,7	1,5	3,9	5,0	3,9	2,5	4,5	4,5	5,4	4,7
T1r	1,4*	0,7	0,9	3,3	1,0	0,7	4,9	5,4	3,4	2,2	5,9	5,9	6,3	5,9
DPr	1,7*	0,8	1,1	3,0	1,7	0,7	3,0	4,0	4,6	2,1	5,9	5,7	5,5*	5,5*
T2r	2,1*	0,8	1,6	3,5	1,6	0,8	4,3	5,3	4,2	2,9	5,5	5,3	5,3	5,3

Tableau 7. – Les erreurs médianes en millimètres pour les 41 recalages CT/IRM, obtenues par les deux versions de notre algorithme. Le chiffre entre parenthèses correspond au classement de cette erreur médiane par rapport aux erreurs médianes des 15 autres équipes. N est le nombre de recalages effectué. Les colonnes « Autres » correspondent aux résultats présentés tableau 6.

Erreurs médianes pour le recalage CT/IRM					
Modalité	Technique				N
IRM	NI1	NI2	Autres (méd)	Autres ($\mu \pm \sigma$)	
T1	2,6 (6)	1,5 (3)	3,3	3,3 \pm 1,6	7
DP	2,6 (7)	2,3 (6)	3,5	3,6 \pm 1,7	7
T2	1,6 (3)	1,9 (4)	4,0	3,6 \pm 1,3	7
T1 rect.	2,7 (7)	1,4 (5)	3,3	3,4 \pm 2,1	6
DP rect.	2,9 (7)	1,9 (6)	3,0	3,2 \pm 1,9	7
T2 rect.	2,5 (6)	2,3 (6)	3,8	3,5 \pm 1,7	7

Tableau 8. – Erreurs maximales des différentes techniques évaluées par le projet initial (d'après [34]). L'étiquette « r » signifie que les images ont été corrigées pour la distortion géométrique avant recalage. * Un patient a été omis. † Transformations non-rigides. ‡ Résultats resoumis après communication de la vérité terrain. Toutes les erreurs sont exprimées en mm.

Erreurs maximales pour le recalage CT/IRM														
Mod.	Technique													
IRM	BA	CO	EL	HA	HE	HI [‡]	MAI	MAL	NO [†]	PE	RO1	RO2	RO3	RO4
T1	6,4	6,7	6,0	51,8	11,0	2,8	12,8	61,4	10,4	7,3	26,0	21,8	17,8	18,8
DP	6,9	3,6	6,6	49,6	10,4	4,1	19,0	59,0	13,9	4,3	25,9	22,2	24,0	20,1
T2	9,1	3,4	4,1	50,6	13,6	4,2	6,3	59,5	9,7	7,2	26,7	22,0	19,4	19,8
T1r	5,8*	3,8	2,6	48,2	2,1	2,3	14,2	60,9	9,6	5,9	27,8	22,2	18,1	18,2
DPr	5,9*	2,5	5,3	45,9	3,7	2,3	9,9	62,7	11,5	4,6	27,5	22,1	24,9*	20,2*
T2r	7,4*	4,3	5,2	49,1	14,3	3,0	6,5	63,2	10,2	9,0	27,1	22,5	19,9	21,6

Tableau 9. – Les erreurs maximales en millimètres pour les 41 recalages CT/IRM, obtenues par les deux versions de notre algorithme. Le chiffre entre parenthèses correspond au classement de cette erreur maximale par rapport aux erreurs maximales des 15 autres équipes. N est le nombre de recalages effectué. Les colonnes « Autres » correspondent aux résultats présentés tableau 8.

Erreurs maximales pour le recalage CT/IRM					
Modalité	Technique				N
IRM	NI1	NI2	Autres (méd)	Autres ($\mu \pm \sigma$)	
T1	6,4 (3)	4,4 (2)	11,9	18,6 \pm 16,9	7
DP	5,8 (4)	6,2 (4)	16,4	19,3 \pm 16,2	7
T2	4,3 (4)	3,8 (2)	11,7	17,4 \pm 16,4	7
T1 rect.	4,6 (5)	4,5 (5)	11,9	17,3 \pm 17,3	6
DP rect.	5,3 (5)	5,1 (5)	10,7	17,8 \pm 17,4	7
T2 rect.	6,0 (4)	4,3 (2)	12,2	18,8 \pm 17,1	7

Tableau 10. – Erreurs médianes des différentes techniques évaluées par le projet initial (d'après [34]). L'étiquette « r » signifie que les images ont été corrigées pour la distortion géométrique avant recalage. * Un patient a été omis. † Transformations non-rigides. ‡ Résultats resoumis après communication de la vérité terrain. Toutes les erreurs sont exprimées en mm.

Erreurs médianes pour le recalage TEP/IRM												
Mod. IRM	Technique											
	BA	CO	HA	HI [‡]	MAI	MAL	NO [†]	PE	RO3	RO4	WO1	WO2
T1	4,6	3,6	2,8	3,2	3,5	4,2	3,6	2,9	4,0	3,4	2,3	3,1
DP	5,2	2,9	4,2	3,1	4,7	4,0	4,1	3,3	4,3	3,3	2,9	3,1
T2	4,7	2,8	2,7	2,4	5,3	4,9	4,6	3,3	4,0	3,6	3,6	3,4
T1r	3,2	2,8*	3,6	2,5	3,9	3,6	3,9	2,8	3,8	3,6	2,0	2,0
DPr	4,5	3,0	3,2	3,0	4,7	3,6	4,4	2,8	3,6*	4,1	2,5	2,3
T2r	3,9	2,0	3,3	2,2	4,0	3,6	5,2	2,9	3,8*	3,4	2,5	2,4

Tableau 11. – Les erreurs médianes en millimètres pour les 35 recalages TEP/IRM, obtenues par les deux versions de notre algorithme. Le chiffre entre parenthèses correspond au classement de cette erreur médiane par rapport aux erreurs médianes des 13 autres équipes. N est le nombre de recalages effectué. Les colonnes « Autres » correspondent aux résultats présentés tableau 10.

Erreurs médianes pour le recalage TEP/IRM					
Modalité IRM	Technique				N
	NI1	NI2	Autres (méd)	Autres ($\mu \pm \sigma$)	
T1	4,3 (12)	5,1 (13)	3,4	3,4 \pm 0,6	7
DP	3,2 (5)	3,6 (7)	3,6	3,8 \pm 0,7	7
T2	3,3 (4)	3,6 (6)	3,6	3,8 \pm 0,9	7
T1 rect.	4,1 (13)	3,0 (6)	3,4	3,1 \pm 0,6	4
DP rect.	1,9 (1)	3,1 (5)	3,4	3,4 \pm 0,8	5
T2 rect.	2,9 (5)	3,1 (6)	3,3	3,3 \pm 0,9	5

Tableau 12. – Erreurs maximales des différentes techniques évaluées par le projet initial (d'après [34]). L'étiquette « r » signifie que les images ont été corrigées pour la distortion géométrique avant recalage. * Un patient a été omis. † Transformations non-rigides. ‡ Résultats resoumis après communication de la vérité terrain. Toutes les erreurs sont exprimées en mm.

Erreurs maximales pour le recalage TEP/IRM												
Mod. IRM	Technique											
	BA	CO	HA	HI [‡]	MAI	MAL	NO [†]	PE	RO3	RO4	WO1	WO2
T1	11,5	12,7	12,1	9,3	10,6	8,5	11,4	10,0	9,4	5,9	5,8	6,0
DP	11,2	9,2	10,3	8,1	9,8	9,3	8,9	11,3	8,8	7,1	6,9	6,3
T2	12,3	7,5	17,4	8,3	15,0	12,3	7,3	13,4	9,0	7,3	8,4	7,3
T1r	6,0	3,7*	17,7	6,0	7,7	8,4	14,2	7,9	7,3	8,9	4,2	5,0
DPr	11,0	7,3	10,1	7,5	9,2	9,4	7,4	11,0	6,6*	6,6*	5,5	5,4
T2r	9,8	7,1	10,2	9,3	10,9	12,4	11,2	15,2	5,8	7,1	6,0	6,1

Tableau 13. – Les erreurs maximales en millimètres pour les 35 recalages TEP/IRM, obtenues par les deux versions de notre algorithme. Le chiffre entre parenthèses correspond au classement de cette erreur maximale par rapport aux erreurs maximales des 13 autres équipes. N est le nombre de recalages effectué. La colonne « Autres » correspondent aux résultats présentés tableau 12.

Erreurs maximales pour le recalage TEP/IRM					
Modalité IRM	Technique				N
	NI1	NI2	Autres (méd)	Autres ($\mu \pm \sigma$)	
T1	8,6 (4)	12,3 (12)	9,7	9,4 \pm 2,3	7
DP	6,0 (1)	6,3 (1)	8,0	8,9 \pm 1,5	7
T2	7,3 (1)	8,2 (5)	8,7	10,5 \pm 3,3	7
T1 rect.	9,0 (11)	8,0 (9)	7,6	8,1 \pm 3,9	4
DP rect.	5,2 (1)	10,0 (10)	7,4	8,1 \pm 1,9	5
T2 rect.	5,9 (2)	10,2 (8)	9,5	9,2 \pm 2,9	5

position au classement) mais dans le cas du recalage TEP/IRM-DT rectifié on fournit la meilleure erreur médiane (1,9 mm). Les autres cas sont classés entre le milieu et le haut du tableau. Dans tous les cas, l'erreur obtenue est acceptable pour la routine clinique et reste sous-voxel. Nos erreurs médianes sont, à part trois cas (T1 (NI1, NI2) et T1 rectifié (NI1)) inférieures à la médiane des médianes des erreurs fournies par les autres équipes. En ce qui concerne l'erreur maximale, nos techniques fournissent une erreur entre 5,2 et 12,3 mm. La méthode NI1 fournit dans trois cas la meilleure erreur maximale et dans cinq cas sur six l'erreur maximale est très inférieure aux erreurs des autres techniques. En revanche, la méthode NI2 fournit des erreurs maximales plus importantes. Rappelons que ces résultats en TEP/IRM sont obtenus *sans segmentation du cerveau en IRM* ce qui montre la robustesse de notre approche et la rend particulièrement attractive en routine clinique. Finalement, en comparant les deux méthodes NI1 et NI2 on constate qu'il n'y a pas de différence significative entre elles pour le recalage CT/IRM. La méthode NI1 (voxels isotropes) fournit des résultats légèrement meilleurs que la méthode NI2 (voxel anisotrope) pour le cas du recalage TEP/IRM. Ceci s'explique par la faible résolution des images TEP (8 mm en profondeur) dans ce cas, rendant utile une interpolation en z .

Les résultats détaillés pour les différents VOI, davantage de statistiques, ainsi que les résultats d'autres équipes sont mis à jour régulièrement sur le site web du projet :

<http://cswwww.vuse.vanderbilt.edu/~image/registration>

4.5. temps de calcul

Les estimateurs robustes induisent un sur-coût calculatoire tout à fait acceptable puisque, dans notre implantation, sur une station de travail HP 9000/C200 (256 MO RAM), les temps de calcul moyens constatés pour des images volumiques $128 \times 128 \times 128$ sont d'environ 10mn cpu pour les moindres carrés classiques (MC) et robustes (MCR), de 30 mn pour l'algorithme de maximisation de l'information mutuelle (IM) et l'estimateur non robuste maximisant l'uniformité inter-image (UI) et de 50 mn pour l'estimateur robuste (UIR)². Comme on peut le constater la complexité calculatoire supplémentaire introduite peut se justifier dans de nombreux cas critiques, où les techniques de recalage classiques ne donnent pas satisfaction. Le tableau 14 résume les temps

Tableau 14. – Temps d'exécution des différentes techniques (station HP 9000/C200).

Temps de calcul					
	MC	MCR	IM	UI	UIR
2D	30 sec	40 sec	60 sec	60 sec	100 sec
3D	10 min	10 min	30 min	30 min	50 min

2. Ces temps de calcul ont d'ores et déjà été réduit d'un facteur 3 sur des stations de la génération suivante.

d'exécution pour chacune des méthodes (les images 2D ont une taille de 256×256 pixels).

5. discussion et conclusion

Les algorithmes de recalage décrits dans cet article ont été inspirés par les algorithmes reposant sur des fonctions de similarité classiques décrits dans [8] pour les images de même modalité, ainsi que par le critère proposé par Woods *et al.* [35] dans le cas d'images multimodales. Les performances de ces approches classiques ont été améliorées de façon significative par les extensions, non triviales, que nous avons décrites ici. La technique d'estimation robuste introduite dans ce contexte présente deux avantages majeurs par rapport aux méthodes précédentes :

- aucun recalage manuel préalable, proche de la solution désirée, n'est requis, ce qui conduit à un algorithme entièrement non supervisé actuellement testé en routine clinique au CHU de Strasbourg;
- les différences d'informations importantes existant dans des couples d'images IRM/IRM ou IRM/TEMP sont explicitement prises en compte par l'algorithme. Jusqu'alors ces différences d'information n'avaient bénéficié d'aucune modélisation spécifique, dans le cas difficile des images multimodales.

La comparaison des estimateurs robustes avec les autres approches classiques de la littérature a montré l'intérêt de ces estimateurs, dans le cas de données simulées et réelles (évolution de lésion, structures non modélisées, etc.). La méthode de recalage robuste a également été évaluée sur la base de données commune de l'Université de Vanderbilt[34] et les résultats se sont révélés très satisfaisants en considérant qu'aucune segmentation du cerveau n'a été appliquée avant le recalage TEP/IRM. Les résultats expérimentaux présentés dans cet article ne comparent que les différentes fonctions de coût (MC, MCR, IM, UI, UIR) et pas les autres paramètres intervenant dans le recalage (technique d'optimisation, qualité de la segmentation du cerveau en IRM, initialisation). Nous avons également dû limiter, pour des raisons évidentes, le nombre de méthodes comparées.

Soulignons enfin que l'approche proposée ici pour le recalage d'images multimodales est générale et ne se limite pas aux seules images médicales : l'imagerie satellitaire (détection de changements), l'imagerie militaire, ainsi que l'analyse multisource des œuvres d'art (analyse des repentirs) [9] peuvent par exemple tirer profit de la robustesse de cette approche. Des résultats de recalage dans d'autres domaines d'application sont présentés dans [23].

6. remerciements

Les images et les transformations de la section 4.4 font partie du projet « Evaluation of Retrospective Image Registration », National Institutes of Health (NIH), No 1 RO1 NS33926-01, responsable J. Michael Fitzpatrick, Vanderbilt University, Nashville, Tennessee, USA.

The images and the standard transformations were provided as part of the project « Evaluation of Retrospective Image Registration », National Institutes of Health (NIH), No 1 RO1 NS33926-01, principal investigator J. Michael Fitzpatrick, Vanderbilt University, Nashville, Tennessee, USA.

Cette étude a été financée par la Commission de la Communauté Européenne DG XII, dans le cadre du programme TMR (Training and Mobility of Researchers), contrat Nr ERBFMIBCT960701 ainsi que par le « Groupement d'Intérêt Scientifique » GIS (CNRS, CEA, INRIA, MENESR) « Sciences de la Cognition ».

BIBLIOGRAPHIE

- [1] M.E Alexander and R. L. Somorjai. The registration of MR images using multiscale robust methods. *Magnetic Resonance Imaging*, 14(5) :453–468, 1996.
- [2] N. M. Alpert, D. Berdichevsky, Z. Levin, E. D. Morris, and A. J. Fischman. Improved methods for image registration. *Neuroimage*, 3 :10–17, 1996.
- [3] B. A. Ardekani, M. Braun, B. F. Hutton, I. Kanno, and H. Iida. A fully automatic multimodality image registration algorithm. *Journal of Computer Assisted Tomography*, 19(4) :615–623, 1995.
- [4] M. J. Black and A. Rangarajan. On the unification of line processes, outliers rejection and robust statistics in early vision. *International Journal of Computer Vision*, 19(1) :57–91, 1996.
- [5] L. G. Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4) :325–376, 1992.
- [6] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal. Automated multimodality image registration using information theory. In *Proceedings of the International Conference on Information Processing in Medical Imaging*, pages 263–274, 1995.
- [7] S. Geman and D. Geman. Stochastic relaxation, Gibbs distribution and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(6) :721–741, 1984.
- [8] J. Hajnal, N. Saeed, E. J. Soar, A. Oatridge, I. R. Young, and G. M. Blyder. A registration and interpolation procedure for subvoxel matching of serially acquired MR images. *Journal of Computer Assisted Tomography*, 19(2) :289–296, 1995.
- [9] F. Heitz, H. Maître, and C. de Couessin. Event detection in multisource imaging : application to fine arts painting analysis. *IEEE Transactions on Acoustic, Speech and Signal Processing*, 38(4) :695–704, 1990.
- [10] F. Heitz, P. Perez, and P. Bouthemy. Multiscale minimization of global energy functions in some visual recovery problems. *Computer Vision, Graphics and Image Processing : Image Understanding*, 59(1) :125–134, 1994.
- [11] P. F. Hemler, T. S. Sumanaweera, P. van den Elsen, S. Napel, and J. R. Adler. A versatile system for multimodality image fusion. *Journal of Image Guided Surgery*, 1 :35–45, 1995.
- [12] M. Herbin, A. Venot, J. Y. Devaux, E. Walter, F. Lebruchec, L. Dubertet, and J. C. Roucayrol. Automated registration of dissimilar images : application to medical imagery. *Computer Vision, Graphics and Image Processing*, 47 :77–88, 1989.
- [13] A. Hill, D. Hawkes, N. Harrison, and C. Ruff. A strategy for automated multimodality registration incorporating anatomical knowledge and imager characteristics. In *Proceedings of the International Conference on Information Processing in Medical Imaging*, pages 182–196. Springer-Verlag, Berlin, 1993.
- [14] H. Jiang, A. Robb, and K. S. Holton. A new approach to 3D registration of multimodality medical images by surface matching. In *SPIE Proceedings of Visualization in Biomedical Imaging*, volume 1808, pages 196–213, 1992.
- [15] D. Lemoine, D. Liegeard, E. Lussot, and C. Barillot. Multimodal registration system for the fusion of MRI, CT, MEG and 3D or stereotactic angiographic data. In *Proceedings of SPIE International Conference on Medical Imaging*, volume 2164, pages 46–56, 1994.
- [16] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16(2) :187–198, 1997.
- [17] G. Q. Maguire Jr, M. E. Noz, H. Rusinek, J. Jaeger, E. L. Kramer, J. J. Sanger, and G. Smith. Graphics applied to medical image registration. *IEEE Transactions on Computer Graphics and Applications*, 11 :20–29, 1991.
- [18] J. B. A. Maintz, P. A. Van den Elsen, and M. Viergever. Evaluation of ridge seeking operators for multimodality medical image matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(4) :353–365, 1996.
- [19] J. B. A. Maintz and M. A. Viergever. A survey of medical image registration techniques. *Medical Image Analysis*, 2(1) :1–36, 1998.
- [20] G. Malandain, S. Fernández-Vidal, and J. M. Rocchisani. Improving registration of 3D medical images using a mechanical-based method. In *Lecture Notes in Computer Science. Proceedings of the 3rd European Conference on Computer Vision (ECCV'94)*, volume 801, pages 131–136, 1994.
- [21] P. Meer, D. Mintz, A. Rosenfeld, and D. Y. Kim. Robust regression methods for computer vision : a review. *International Journal of Computer Vision*, 6(1) :59–70, 1990.
- [22] O. Musse, J. P. Armspach, I. J. Namer, F. Heitz, F. Hennel, and D. Grucker. Data-driven curvilinear reconstruction of 3D MRI : application to cryptogenic extratemporal epilepsy. *Magnetic Resonance Imaging*, 16(10) :1227–1235, 1998.
- [23] C. Nikou, F. Heitz, and J. P. Armspach. Robust registration of dissimilar single and multimodal images. In *Lecture Notes in Computer Science. Proceedings of the 5th European Conference on Computer Vision (ECCV'98)*, volume 2, pages 51–65, Freiburg, Germany, 2-6 June 1998.
- [24] C. A. Pelizzari, G. T. Y. Chen, D. R. Spelbring, R. R. Weichelbaum, and C. T. Chen. Accurate three-dimensional registration of CT, PET and/or MR images of the brain. *Journal of Computer Assisted Tomography*, 13(1) :20–26, 1989.
- [25] U. Pietrzyk, K. Herholz, A. Jacobs, R. Mielke, I. Slansky, M. Wurker, and W. D. Heiss. An interactive technique for three-dimensional image registration : validation for PET, SPECT, MRI and CT brain images. *Journal of Computer Assisted Tomography*, 35 :2011–2018, 1994.
- [26] L. Rumbach, E. Racadot, J. P. Armspach, I. J. Namer, J. F. Bonneville, J. Wijdenes, C. Marescaux, P. Herve, and J. Chambon. Biological assessment and MRI monitoring of the therapeutic efficacy of a monoclonal anti-T CD4 antibody in multiple sclerosis patients. *Multiple Sclerosis*, pages 207–212, 1996.
- [27] H. Sawhney, S. Ayer, and M. Gorkani. Model-based 2D-3D dominant motion estimation for mosaicing and video representation. In *Proceedings of the*

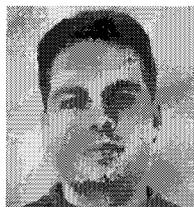
Mesures de similarité robustes

- International Conference on Computer Vision*, pages 583–590, Cambridge, USA, June 1995.
- [28] C. Stewart. Bias in robust estimation caused by discontinuities and multiple structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(8) :818–833, 1997.
- [29] C. Studholme, D. Hill, and D. Hawkes. Automated three-dimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures. *Medical Physics*, 24(1) :25–35, 1997.
- [30] P. Van den Elsen, E. J. D. Paul, and M. A. Viergever. Medical image matching - a review with classification. *IEEE Engineering in Medicine and Biology*, 12(1) :26–39, 1993.
- [31] P. Van den Elsen, E. J. D. Pol, T. S. Sumanaweera, P. F. Hemler, S. Napel, and J. R. Adler. Grey value correlation techniques used for automatic matching of CT and MR brain and spine images. In *SPIE International Conference on Visualization in Biomedical Computing*, volume 2359, pages 227–237, 1994.
- [32] A. Venot, J. F. Lebruchec, and J. C. Roucaÿrol. A new class of similarity measures for robust image registration. *Computer Vision, Graphics and Image Processing*, 28 :176–184, 1984.
- [33] W. Wells III, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis. Multimodal volume registration by maximization of mutual information. *Medical Image Analysis*, 1(1) :33–51, 1996.
- [34] J. West, M. Fitzpatrick, M. Wang, B. Dawan, C. Maurer Jr., R. Kessler, R. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, P. Suetens, D. Vandermeulen, P. Van den Elsen, S. Napel, T. Sumanaweera, B. Harkness, P. Hemler, D. Hill, D. Hawkes, C. Studholme, A. Mainz, M. Viergever, G. Malandain, X. Pennec, M. Noz, G. Maguire, Jr. M. Pollack, C. Pelizzari, R. Robb, D. Hanson, and R. Woods. Comparison and evaluation of retrospective intermodality brain image registration techniques. *Journal of Computer Assisted Tomography*, 21(4) :554–566, 1997.
- [35] R. P. Woods, J. C. Mazziota, and S. R. Cherry. MRI-PET registration with automated algorithm. *Journal of Computer Assisted Tomography*, 17(4) :536–546, 1993.

Manuscrit reçu le 2 avril 1998.

LES AUTEURS

Christophoros NIKOU



Christophoros Nikou est né en 1971 à Thessalonique (Grèce). Il a obtenu le diplôme d'ingénieur (Electrical Engineering) de l'Ecole Polytechnique de l'Université Aristote de Thessalonique (Grèce) en 1994. Il a effectué son stage de fin d'études à l'Institut de Recherche et d'Enseignement Supérieurs Techniques de l'Électronique (IRESTE), Nantes, dans le cadre du programme ERASMUS de la Communauté Européenne. Il a obtenu le DEA «Photonique et Image» de l'Université Louis Pasteur (ULP-Strasbourg I) en

1995. Actuellement, il est doctorant à l'ULP au Laboratoire des Sciences de l'Image de l'Informatique et de la Télédétection (LSIIT, CNRS UPRES-A 7005) et à l'Institut de Physique Biologique, Hôpitaux Universitaires de Strasbourg (IPB, CNRS UPRES-A 7004) sous contrat avec la Commission Européenne (DG XII). Ses domaines de recherche incluent le traitement d'images biomédicales (recalage, segmentation), la vision par ordinateur et les modèles déformables statistiques.

Jean-Paul ARMSPACH



Jean-Paul Armspach a obtenu le diplôme d'ingénieur et le doctorat respectivement en 1978 et 1982 à l'INSA de Lyon. Il est actuellement ingénieur de recherche à l'Institut de Physique Biologique (CNRS UPRES-A 7004), Faculté de Médecine, Université Louis Pasteur (Strasbourg I) où il est responsable du groupe d'informatique et de traitement d'images médicales. Ses domaines de recherche incluent l'imagerie par Résonance Magnétique (IRM) et le traitement d'images biomédicales.

Fabrice HEITZ



Fabrice Heitz est né en 1961. Diplômé de Télécom Bretagne en 1984, il a soutenu sa thèse à Télécom Paris en 1988. Il a été de 1988 à 1994 Chargé de Recherches à l'INRIA de Rennes (dans le projet TEMIS). Il est actuellement Professeur à l'Ecole Nationale Supérieure de Physique de Strasbourg et anime un groupe de recherche en analyse multi-images au sein du LSIIT (Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédétection, UPRES-A CNRS 7005). F. Heitz est éditeur associé

pour la revue *IEEE Transactions on Image Processing* depuis 1996. Ses domaines d'intérêt incluent la modélisation statistique et les modèles déformables appliqués à l'analyse multi-images et à l'imagerie médicale.

Izzie-Jacques NAMER



Izzie-Jacques Namer est né en 1959 à Ankara (Turquie). Il a obtenu le diplôme de docteur en médecine de l'Université d'Istanbul (Turquie) en 1981, le diplôme de spécialiste en neurologie de l'Université d'Hacettepe d'Ankara (Turquie) en 1987, le DEA de neurosciences de l'Université Pierre et Marie Curie-Paris VI en 1991 et le diplôme de doctorat d'université (sciences) de l'Université Louis Pasteur de Strasbourg en 1994. Actuellement, il est assistant hospitalier-universitaire à l'Institut de Physique Biologique, Faculté de Médecine de Strasbourg (UPRES-A 7004ULP/CNRS) et anime un groupe de recherche sur les maladies du système nerveux central. Ses domaines de recherche sont la sclérose en plaques et son modèle animal, l'EAE, et l'épilepsie.