Fusion d'Image Médicale Multimodal (IRM, TEP)

S. Cherik, H. Mouhadjer, A. Belghoraf
Université Mohamed Boudiaf d'ORAN (USTO). BP 1505 M'naouer .Oran, Algérie
Laboratoire de Communication
Département d'Electronique, Faculté de Génie électrique
E-mail : cherik_eln @yahoo.fr

Résumé- **Le recalage d'images trouve de nombreuses applications médicales aussi bien dans le diagnostic que dans le traitement. La combinaison d'images multimodales a même conduit à de nouvelles techniques de chirurgies. Pour ce faire, le succès des différentes mesures de recalage fut plus ou moins reconnu jusqu'à ce qu'une méthode se démarque en utilisant la mesure de l'information mutuelle . Contrairement aux méthodes courantes se basant sur la mesure des intensités de l'image, l'information mutuelle est basée sur la quantité d'information que contiennent les images. Cette méthode s'avère être plus robuste au recalage d'images multimodales. Dans cet article, nous nous intéresserons à la qualité de la mesure de l'information mutuelle en la comparant à une mesure classique de recalage, la corrélation. Nous critiquerons également une amélioration à la méthode de recalage combinant l'information du gradient à l'information 31/10/20007mutuelle.**

Mots clés- **Corrélation, Gradient, Information mutuelle, Recalage d'image.**

I. INTRODUCTION

L'alignement d'images est un problème courant en vision par ordinateur. Ses applications médicales sont diverses : suivre l'évolution d'un patient, la localisation de sites malins et l'aide à la visualisation chirurgicale. L'exactitude d'un alignement offre même de nouvelles possibilités de traitements. Par exemple, l'imagerie par résonance magnétique (IRM) et l'imagerie tomographique (TEP : tomographie par mission de positon).

Le recalage d'images consiste à trouver la transformation qui aligne exactement deux images. Ces images présentent en général des différences qui peuvent rendre le recalage difficile. Des structures anatomiques clairement visibles sur une image peuvent par exemple être altérées ou même inexistantes sur l'autre image. Ces différences sont encore plus flagrantes dans le cas d'un recalage multimodal (eg. IRM-TEP) où les images sont très peu similaires les unes des autres.

Traditionnellement, des marqueurs, ou des points préalablement mis en évidence étaient utilisés pour recaler des images multimodales. Les approches basées sur les intensités des images furent introduites par Woods en 1992 [4]. Le recalage multimodal était alors effectué en minimisant la variance des ratios d'intensités des images. Tout de suite après, Hill [5] introduit la notion d'histogramme joint. Dans le même groupe, Van den Elsen publie une méthode de recalage par corrélation [6]. Collignon [7] et Studholme [8] proposent d'utiliser l'entropie comme mesure de recalage. La même année, Viola [9] et Collignon [1] trouvent simultanément une approche basée sur l'information mutuelle. Cette méthode s'avère être plus robuste au recalage multimodal. Elle continue néanmoins à faire aujourd'hui l'objet d'améliorations [10].

Nous avons utilisée comme première méthode la corrélation pour le recalage d'images. Cette corrélation est basée uniquement sur les intensités des images (voir ci-dessous). Leur inconvénient pour le recalage d'images multimodales introduira ensuite le recalage par l'information mutuelle. . Egalement, une amélioration combinant l'information du gradient [3] sera examinée. En dernière partie, l'efficacité de ces méthodes sera finalement testée et une conclusion résumera nos résultats.

II. METHODES

A. *Critère de corrélation*

Parallèlement à l'analyse théorique, la minimisation d'un critère de corrélation semble être une solution intuitive au problème de recalage de deux images. En effet, si deux images identiques A et B sont correctement alignées l'une avec l'autre, les différences d'intensités $A(x) - B(x')$ de chaque paire de pixels x et x' seront nulles. Si les deux images ne sont pas correctement alignées, ces différences $PA(x) - B(x')$ augmenteront. Recherché le maximum de vraisemblance $\hat{T}$ en minimisant le paramètre $J(T)$ permet donc de recaler deux images A et B.

$$J(T) = \sum_{x \in A \cap B} \left| A(x) - B(T(x)) \right|^2$$

De plus, en développant l'expression de $J(T)$, il est possible de prouver que cette méthode est en réalité équivalente à rechercher le maximum d'une fonction de corrélation entre les images A et B :

$$\begin{aligned} J(T) &= \sum_{x \in A \cap B} \left| A(x) - B(T(x)) \right|^2 \\ &= \sum_{x \in A \cap B} A^2(x) - 2A(x)B(T(x)) + B^2(T(x)) \end{aligned}$$

Si l'image B contient toujours les mêmes pixels, le terme $B^2(T(x))$ devient constant. Les termes au carré sont donc indépendants de la minimisation de T . Il ne reste alors que $J(T) = -2PA(x)B(T(x))$.

Cette expression est la généralisation de la fonction de corrélation à n'importe quelle transformation T . La fonction de corrélation, définie par :

$$\gamma_{AB(T)} = \sum_{x \in A \cap B} A(x)B(T(x)) \quad (1)$$

Dans le cas particulier d'un décalage sur un axe, $T(x) = x - t$. Dans le cas d'une translation sur deux axes, $T(x) = x + [tx, ty]^T$. Par spéculation , seule les déformations rigides seront considérées. La transformation T comportera donc un premier terme pour une rotation d'angle θ et un second terme de translation tx, ty :

$$T(x) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} x + \begin{bmatrix} t_x \\ t_y \end{bmatrix}$$

La fonction de corrélation varie en fonction de la surface commune des deux images A ∩ B. Plus cette surface est grande, plus la corrélation sera importante, mais si les images se superposent sur une petite surface, la corrélation sera faible. Afin d'éviter ce biais, il existe une normalisation à la fonction de corrélation. Elle consiste à diviser la fonction de corrélation par les valeurs maximales que peuvent prendre les deux images, soient par les fonctions d'autocorrélation :

$$\gamma_N(A,B) = \frac{\sum_{x \in A \cap B} A'(x)B'(T(x))}{\sqrt{\sum_{x \in A \cap B} (A'(x))^2 \sum_{x \in A \cap B} (B'(T(x)))^2}} \quad (2)$$

Ou A' et B' sont les intensités centrées sur les moyennes des images A et B, soit $A'(x) = A(x) - \overline{A(x)}$, et $B'(x) = B(x) - \overline{B(x)}$.
Il y aura maximum de la fonction de corrélation (1,2) lorsque les images $A(x)$ et $B(T(x))$ seront les plus semblables. La vraisemblance à rechercher est donc la transformation T recalant au mieux les images $A(x)$ et $B(T(x))$. Pour maximiser la fonction de cout $J(T) = \gamma AB(T)$, il faut théoriquement annuler sa dérivée $d\gamma/dT$. En pratique, la recherche du maximum peut se faire par des méthodes numériques ne faisant pas appel à la dérivée $d\hat{\gamma}/dT$. Par exemple la méthode de Powell [12] utilise cette stratégie. Chaque dimension est optimisée tour à tour jusqu'à ce qu'il y ait convergence.

B. Maximisation de l'information mutuelle

La méthode de maximisation de la fonction de corrélation telle quelle sera vue par la suite, suppose que les images à recaler ont des intensités similaires. En effet, si les intensités d'une même structure anatomique sont inversées dans les deux images (ie. les os dans des images IRM-TEP), les différences d'intensité des images ne refléteront plus une mesure de la différence des images et ne sera plus maximale au recalage. La valeur de corrélation sera maximale là ou les pixels des deux images ont des intensités similaires. En inversant des intensités, les régions les plus claires ne sont plus aux mêmes positions. Ces méthodes qui suivent sont donc adaptées pour un recalage intra modal, où les intensités des deux images sont supposées similaires. Pour des images multimodales, d'intensités inversées, ces méthodes se montreront inefficaces. Pour le recalage multimodal, les méthodes basées sur l'information mutuelle sont mieux adaptées.

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 considérées comme identiques. Tenter de transformer une image de manière à ce qu'elle ressemble le plus à une seconde image peut donc se faire en maximisant leur information mutuelle. Avant d'aborder l'information mutuelle et son amélioration, il est très important de parler de la notion d'entropie et par la suite, de la maximisation de l'information mutuelle.

B.1. Entropie

L'entropie est la quantitée d'information contenue dans une série d'événements. Une image A est par exemple constituée d'une série d'événements, des pixels, ayant tous une probabilité pi et une intensité i .
Plus une image est complexe, plus son entropie $H(A)$ est grande. Shannon propose une définition de l'entropie telle que :

$$H = \sum_i p_i \log \frac{1}{p_i} = -\sum_i p_i \log p_i \quad (3)$$

La première expression montre bien que plus un élément est rare, plus il a de signification. Si une image est constituée de pixels de plusieurs tons de gris, l'image transporte une information plus importante qu'une image d'un ton unique. La définition de Shannon de l'entropie indique l'information moyenne que l'on peut s'attendre de chaque élément de l'image.

B.2. Information mutuelle

L'information mutuelle $IM(A,B)$ de deux images A, B, possède 3 définitions équivalentes. Chacune d'elles permet d'expliquer différemment l'information mutuelle.

La première utilise la différence de l'entropie d'une image A et de l'entropie de la même image A sachant une autre image B :

$$\begin{aligned} IM(A,B) &= H(A) - H(A/B) \\ &= H(B) - H(B/A) \end{aligned} \quad (4)$$

Ici, $H(A)$ mesure l'information contenu dans l'image A, tandis que $H(A/B)$ mesure la quantité d'information contenu dans l'image A lorsque l'image B est connue. L'information mutuelle correspond donc à la quantité d'information que l'image B possède sur l'image A, similairement, la quantité d'information que l'image A possède sur l'image B.
La seconde définition évoque la distance de Kullback-Leibler $\sum_i p_i \log p_i / q_i$, qui mesure la distance entre deux distributions p et q :

$$IM(A,B) = \sum_{a,b} p_{ab} \log \frac{p_{ab}}{p_a p_b} \quad (5)$$

soit la mesure entre la distribution $p_{a,b}$ des images *A* et *B* et la distribution $p_a p_b$ où les images *A* et *B* sont indépendantes. Cette définition de l'information mutuelle mesure donc la dépendance des images *A* et *B*. Il y aura recalage lorsque les images *A* et *B* sont le plus semblables [12].

La troisième définition de l'information mutuelle est une combinaison des entropies de deux images, séparées et jointes :

$$IM(A,B) = H(A) + H(B) - H(A,B) \quad (6)$$

Les entropies séparées *H (A)* et *H (B)* mesurent la complexité des images *A* et *B*. L'entropie jointe *H (A, B)* mesure la quantité d'information que les images *A* et *B* apportent en même temps. Si les images *A* et *B* sont proches, la première image explique bien la seconde, et l'entropie jointe est minimale. L'entropie jointe est plus précisément l'entropie de l'histogramme joint des images *A* et *B*. Cet histogramme répertorie la correspondance des intensités des pixels entre deux images. Maximiser l'information mutuelle selon cette troisième définition revient donc à dire que le plus d'information est recherchée dans chaque image, tout en cherchant à faire ressembler les images *A* et *B*.

Nous nous intéressent dans cet article à la dernière définition de l'information mutuelle (6) sera utilisée. Néanmoins, cette fonction est biaisée par la surface commune des images recalées. Tout comme la fonction de corrélation, une normalisation de l'information mutuelle est proposée en utilisant une fraction de deux termes dépendant de la surface commune des deux images [1] :

$$IMN(A,B) = \frac{H(A) + H(B)}{H(A,B)} \quad (7)$$

La recherche de la transformation *T* recalant les images *A(x)* et *B (T(x))* correspond au maximum de la fonction de coût *J (T) = IM (A, B)*. Elle se fait théoriquement en annulant sa dérivée dIM/dT. En pratique, une méthode numérique ne faisant pas appel à la dérivée permettra d'estimer la vraisemblance $\hat{T}$.

C. Gradient

Le recalage par maximisation de l'information mutuelle en utilisant la définition de Shannon pour l'entropie ne tient pas compte de l'information spatiale des images. En effet, en réorganisant aléatoirement des paires de pixels dans deux images, la valeur de l'information mutuelle est la même que celle des deux images originales. Chercher à maximiser l'information mutuelle des gradients des images permettrait de tenir compte de l'information spatiale. Par contre, le maximum de l'information mutuelle survient lorsque les gradients des deux images sont recalés. Les maximums à rechercher sont donc plus étroits, et par conséquent plus difficiles à rechercher en pratique.

Ce problème à amené l'introduction de l'information des gradients des images dans la fonction à maximiser [13]. Il est en effet attendu que les gradients des images représentant un même objet soient similaires, du moins dans les mêmes directions. Ainsi, pour tout point x de l'image *A* et sa correspondance *x' = T (x)* dans l'image B, on cherche à ce que les gradients $\nabla A(x)$ et $\nabla B(x')$ aient les mêmes directions.

Pour un angle α, l'angle entre les gradients $\nabla A(x)$ et $\nabla B(x')$, la fonction $cos(2\alpha)$ favorise des gradiants de même orientation. Cette fonction est en effet maximale lorsque l'angle α est proche de 0 ou de α. Ainsi, pour un même point *(x, x')* dans deux images *(A,B)*, les gradients de même orientation sont obtenus en maximisant la fonction suivante :

$$W(\alpha) = \frac{1}{2}(\cos(2\alpha) + 1)$$

Egalement, afin de favoriser les gradients apparaissant dans les deux images, et de pénaliser les cas où le gradient est fort que dans une image, il sera possible de multiplier la fonction $W(\alpha)$ par le minimum des normes des deux gradients. La sommation de cette fonction sur chaque pixel de l'image donne le terme de l'information du gradient :

$$G(A,B) = \sum W(\alpha) \min(|\nabla_{A(x)}|, |\nabla_{B(x')}|)$$

En introduisant l'information spatiale des images A et B, la mesure du recalage est améliorée. Maximiser ce terme favorisera le recalage des contours des images. Il est à noter que dans cette expression, le terme min rend cette fonction discontinue. L'information du gradient telle que définie ici ne possède donc pas de dérivée. Pour maximiser ce terme, il faudra utiliser des méthodes numériques.

En combinant l'information du gradient *G(A,B)* l'information mutuelle *IM(A,B),* maximiser la vraisemblance *T* est donnée par :

$$\hat{T} = \arg\max_{T} G(A,B) IM(A,B)$$

Cette expression, non continue, ne peut être résolue théoriquement en annulant sa dérivée. La recherche du maximum de vraisemblance $\hat{T}$ se fera donc en utilisant une méthode numérique, telle que celle de Powell.

L'utilisation de l'information mutuelle n'est pas affectée par l'utilisation d'intensités différentes dans le cadre d'un recalage multimodal. En effet, si deux images représentent une même structure anatomique, une image contiendra forcement de l'information sur la seconde, il y aura de l'information mutuelle entre les deux images. L'utilisation d'intensités différentes dans les deux images se traduira dans une transformation de la distribution de l'entropie jointe. La valeur de l'entropie restera semblable, l'utilisation d'intensités différentes n'augmentera, ou ne diminuera pas la quantité d'information. La maximisation de l'information mutuelle est donc bien adaptée au recalage multimodal.

D. Interpolation

En transformant un pixel *x* d'une image de référence *A*, il n'est pas trivial de déterminer l'intensité du pixel transformé *x' = T (x)* sur une image *B*. En effet, du à la discrétisation des images, l'intensité du pixel *x'* doit être interpolé. Les erreurs

dues à cette interpolation peuvent perturber la fonction à maximiser.
Une interpolation simple consiste à prendre l'intensité du point discret le plus proche. Cela provoquera par contre la présence de maximum local dans la fonction à maximiser. En utilisant une interpolation linéaire entre les intensités des points discrets voisins, les erreurs de perturbation seront moindres. Elles seront encore moindres en utilisant une interpolation de plus haut ordre, telle que linéaire ou cubique.

III. RESULTATS

Nous avons générés des courbes et des surfaces pour présenter les différentes mesures de recalage étudiées (équations 2, 8,9). Ces courbes et surfaces sont de très bons outils pour prévoir le comportement d'un algorithme de recherche de maximum. Les faiblesses des mesures de recalage par corrélation au recalage multimodal seront montrées. Par la suite, l'amélioration apportée par l'information mutuelle sera mise en évidence. Un exemple pratique avec notre méthode d'optimisation sera finalement présenté.

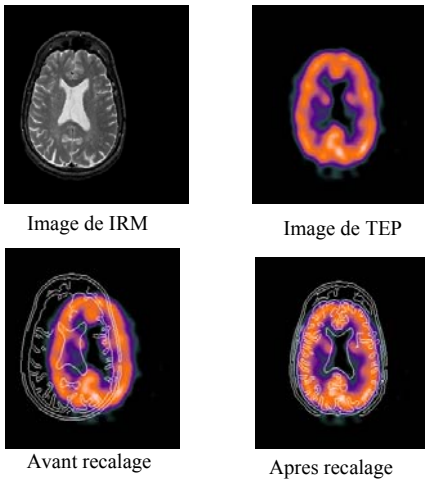


Fig. 1- Résultat de la recherche du maximum de coefficient de corrélation

A. La corrélation

Nous avons tenté de recaler des images multimodales, à l'aide de la corrélation. Une image IRM et une image TEP d'un crâne (Fig.1) sont utilisées pour ce test. La transformation idéale est trouvée ici empiriquement en translatant de [x =29 ; y = 53]. En faisant varier les paramètres de translation et en reportant la mesure de corrélation sur une surface (Fig.2), le maximum de corrélation n'indique pas la transformation idéale. Il indique plutôt un anneau de valeurs maximales autour du recalage idéal. Cet anneau de corrélation maximale correspond aux mesures pour lesquelles la région du cerveau sur l'image IRM chevauche la région de l'image TEP. Cet essai montre très bien la difficulté reliée au recalage par maximum de corrélation.

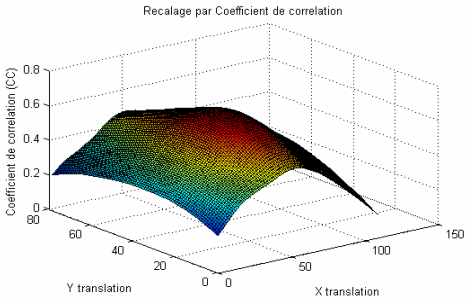


Fig. 2- Mesure de corrélation en fonction de la translation

B. Corrélation normalisée

Les essais sur la corrélation normalisée ont montré des résultats semblables à ceux de la corrélation non-normalisée. Toutefois, certaines différences ont pu être remarquées. Pour un même test faisant varier les paramètre de translation sur l'image, la fonction de corrélation normalisée (Fig.3) montre que le pic de corrélation maximale est beaucoup plus effilé que celui utilisant la fonction de corrélation non-normalisée (Fig.2), ce qui fait en sorte a diminuer la zone de corrélation maximale. Cet effet peut éventuellement rendre difficile la recherche de cette zone de maximum de corrélation.

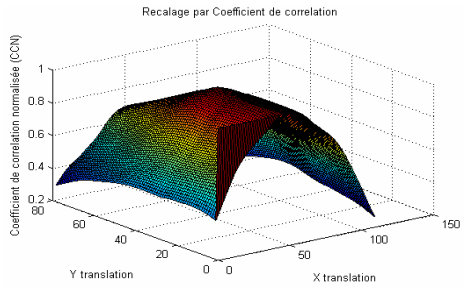


Fig. 3– Mesure de corrélation normalisée en fonction de la translation

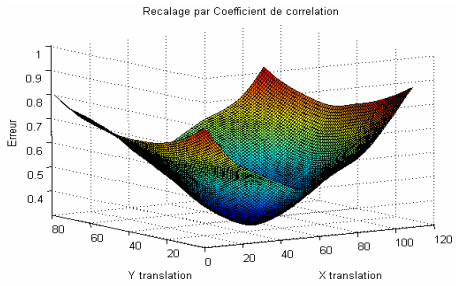


Fig. 4 – Mesure d'erreur pour chaque translation

Dans le graphe au dessus nous distinguons la parfaite concordance entre ce que mesure la corrélation et le calcul de l'erreur entre l'image sujet de traitement pour chaque translation, nous remarquons clairement que le maximum de la corrélation correspond à la valeur minimale de l'erreur de valeur 0.1243.

C. *Information mutuelle*

Le recalage par information mutuelle donne de bons résultats. Pour ce test, l'image de l'IRM est recalée avec une image de TEP translatée de [x = 32; y = 53] sans rotation. En reportant la valeur de information mutuelle de chaque translation [x; y] sur une surface, le maximum de information mutuelle indique bien le bon recalage a la transformation [x=32, y=52] (fig.5). On remarque que la progression de la valeur de information mutuelle pour les différentes translations est très régulière.

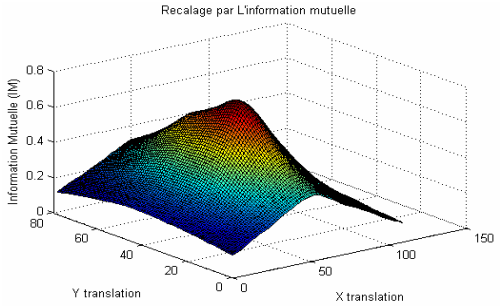


Fig. 5– Mesure de l'information mutuelle en fonction de la translation

D. *Information mutuelle normalisée*

L'utilisation de l'information mutuelle normalisée à la place de l'information mutuelle n'a pas montré d'amélioration apparente (fig.6). En effectuant un teste sur l'image IRM et TEP, le maximum des mesures de recalage indiquent la même transformation, et la largeur des pics maximaux est semblable.

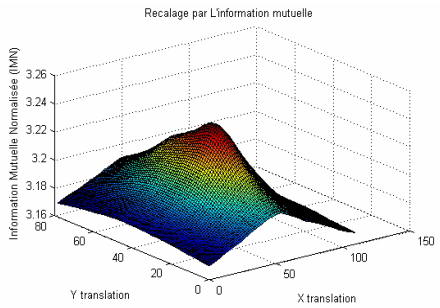


Fig. 6– Mesure de l'information mutuelle normalisée en fonction de la translation

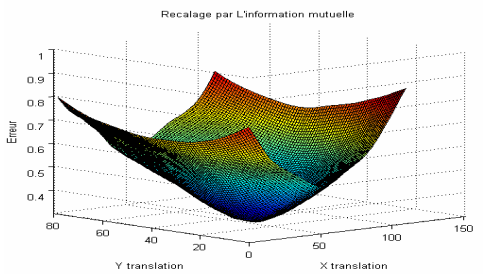


Fig. 7 – Mesure d'erreur pour chaque translation

Comme nous l'avions prédit la valeur de l'erreur entre les images sujet de recalage est minimale lorsque la valeur de l'information mutuelle est maximale, ainsi d'après la lecture numérique des graphes, on remarque que la valeur maximale de information mutuelle (IM) correspond à une valeur de translation [x=32, y=52] fig.5 avec l'erreur de 0.1045 (fig. 7).

E. *Information mutuelle avec gradient*

L'information du gradient a été combinée à l'information mutuelle non normalisée. La multiplication de l'information mutuelle avec l'information du gradient tend à effiler le pic du maximum de la fonction de recalage.
L'utilisation de l'information mutuelle combinée à l'information du gradient sur les scans TEP et IRM (Fig.8) fait ressortir 2 pics maximaux. Il y a donc deux translations qui rendent un recalage en viseageable, soit [x = 32; y = 52] et [x = 32; y = 50], mais aucune de ces deux solutions correspondent avec notre translation idéale de [x = 32; y = 53].

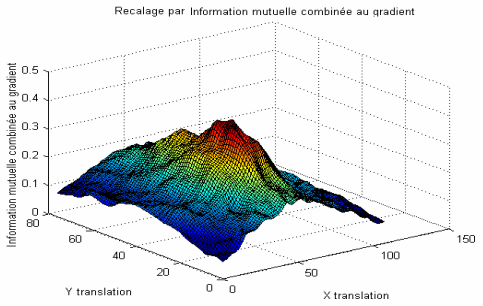


Fig. 8 – Mesure de IM combinée au gradient en fonction de la translation

Ces translations verticales mettent respectivement en évidences les contours frontaux du cerveau.
Le cerveau du scan IRM est donc plus long que celui du scan TEP. Il est ainsi possible d'émettre l'hypothèse que soit les scans TEP et IRM ne sont pas à la même échelle, soit qu'il s'agit de deux coupes différentes du cerveaux, soit de deux patients différents, l'un ayant un crane plus long. Par la suite, l'exemple pratique du recalage des images TEP et IRM (Fig.11) montre l'alignement des contours arrières du cerveau.

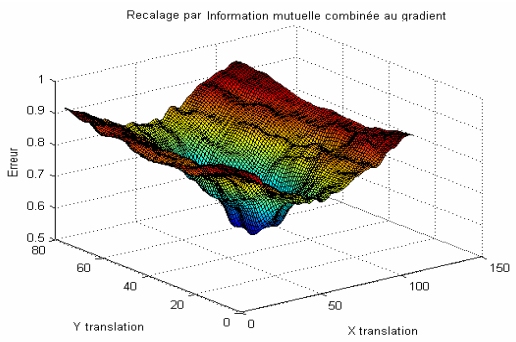
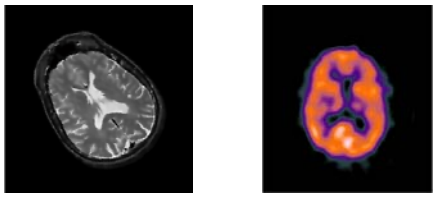


Fig. 9 – Mesure d'erreur pour chaque translation

Le même test que précédemment faisant varier le paramètre de rotation a été fait sur les scans IRM et TEP du crâne, le maximum de l'information mutuelle indique toujours la transformation idéale de [$\theta = 40^{\circ}$] voir (fig. 10).



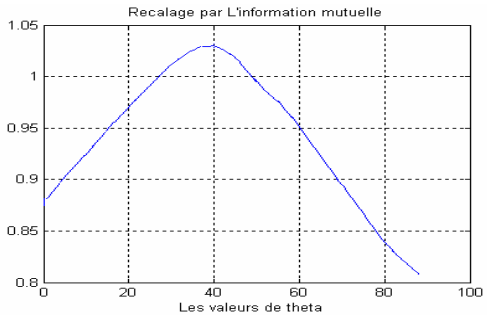
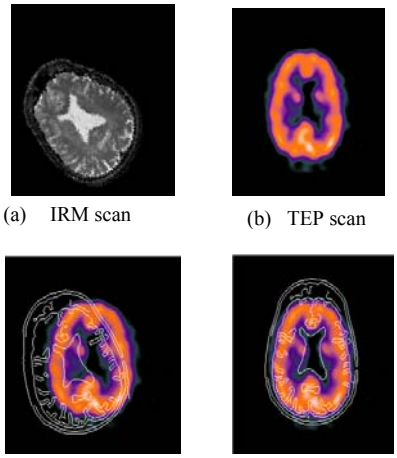


Fig. 10 – Mesure de l'information mutuelle pour une rotation de l'image IRM sur l'image TEP.

F. Optimisation

Notre méthode d'optimisation pour la recherche du maximum nous a permis de déterminer rapidement un recalage idéal dans la plupart des cas. La méthode permet de déterminer le maximum de l'indice utilisé (corrélation, information mutuelle avec ou sans gradient) pour un recalage à trois degrés de liberté (translation en x, en y et rotation). Il est à noter que le nombre d'itérations est connu à l'avance.

Ce nombre d'itération est le logarithme en base deux de la plus grande valeur entre la largeur de l'image (translation en x), la hauteur de l'image (translation en y) et 360 (rotation). Sur les scans IRM et TEP du crâne, une transformation de [x = 32; y = 51; $\theta = 20^{\circ}$] est appliquée au scan IRM. L'optimisateur est initialisé avec des paramètres nuls qui sont relativement éloignés de la solution finale. Après 8 itérations, l'erreur de la solution proposée par l'optimisateur de [x = 0; y = 2; $\theta = 2$], ce qui amène le scan IRM aligné avec le scan TEP. La fig.11 montre le résultat de l'optimisation. Le module des gradients des deux images est affiché afin de mieux apprécier le résultat du recalage.



(a) IRM scan (b) TEP scan

(c) Avant recalage (d) Apres recalage

Fig. 11 - Résultat de la recherche du maximum d'information mutuelle avec gradient après huit itérations

IV. DISCUSSIONS

A. Mise en contexte des résultats

L'utilisation de différentes mesure de recalage nous a permis de comprendre la difficulté relative au recalage des images, et particulièrement au recalage d'image médicale multimodales. Il ressort toutefois de nos essais que le recalage par corrélation d'image multimodales n'est pas mesure fiable. En effet, la sensibilité de cette mesure à l'intensité faite en sorte que la corrélation ne donne pas une bonne mesure de recalage dans ce cas.

Dans ce contexte, l'utilisation de la mesure de l'information mutuelle s'est montrée comme un meilleur choix pour les images multimodales. Malgré le fait que les images à recalage avaient une échelle différente.

Les résultats montrent que le maximum de l'indice correspond au recalage le plus proche. Par ailleurs, l'information mutuelle normalisée n'a pas montré de différences marquée avec la mesure non normalisé.

La combinaison de l'information du gradient augmente la précision des pics maximaux. Des images multimodales, l'information du gradient indique même de nouveaux pics maximaux.

En effet, la recherche des pics maximaux plus étroits par optimisation devient plus difficile. Les pics larges sont par contre plus faciles à trouver et leur recherche est d'autant plus rapide.

V. CONCLUSION

Nous avons présenté dans cet article la qualité et la puissance de différentes mesures de recalage. Nous avons commencé par utiliser une méthode intuitive qui tend à superposer des régions de tons de gris similaires. Cette méthode utilisant la fonction de corrélation s'est avérée etre inefficace dans le cadre du recalage multimodal. En mesurant la quantité d'information qu'une image contient sur une autre image ; l'introduction de l'information mutuelle établie par Viola et al. [1] ainsi que Collignon et al. [2], a été d'une contribution majeure au problème du recalage multimodal. D'autant plus, là ou l'information mutuelle n'indique qu'une unique solution. Nous avons pu constater ce phénomène en effectuant des tests concluants sur des scans IRM et TEP d'un crâne., la combinaison de l'information du gradient et de l'information mutuelle a permis de faire ressortir différents recalages possibles. La complexité des algorithmes nous a fait prendre conscience de l'importance de l'optimisation dans le processus de recalage.

BIBLIOGRAPHIE

[1] P. Viola and W. M.W. III, "Alignment by maximization of mutual information," in Internation Conference on Computer Vision, pp. 16–23,June 1995.

[2] A. Collignon, F. Maes, D. Delaere, D. Vander- meulen, P. Suetens, and G. Marchal, "Automated multi-modality image registration based on information theory," in Information Precessing in Medical Imaging, pp. 263–174, 1995.

[3] A. Roche, Recalage d'image médicale par inférence statistique, Thèse de doctorat de Université de Nice Sophia Antipolis, 2001.

[4] R. Woods, S. Cherry, and J. Mazziotta, "Rapid automated algorithm for aligning and reslicing pet images," Journal of Computer Assisted Tomography, vol. 16, pp. 620–633, July-August 1992.

[5] D. Hill, D. Hawkes, N. Harrison, and C. Ruff, "A strategy for automated multimodality registration incorporating anatomical knowledge and imager characteristics," Information Processing in Medical Imaging, pp. 182–196, 1993.

[6] P. A. V. den Elsen, J. A. Mints, E.-J. D. Pol, and M. A. Viergever, "Automatic registration using correlation of CT and MR brain images," IEEE Transactions on Medical Imaging, vol. 14, June 1995.

[7] A. Collignon, D. Vandermeulen, P. Suetens, and G. Marchal, "3d multi-modality medical image registration," Computer Vision, Virtual Reality, and Robotics in Medicine, vol. 905, pp. 195–204, 1995.

[8] C. Studholme, D. Hill, and D. Hawkes, "Multiresolution voxel similarity measures for mr-pet registration," Information Processing in Medical Imaging, pp. 287–298, 1995.

[9] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," IEEE Transactions on Medical Imaging, vol. 16, no. 2, pp. 187–198, 1997.

[10] J. P. Pluim, J. A. Maintz, and M. A. Viergever, "Mutual information based registration of medical images : a survey," IEEE Transactions on Medical Imaging, vol. 22, no. 8, pp. 986–1004, 2003.

[11] S.A. Inria, Recalage rigide d'image médicale volumiques contributions aux approches iconiques et géométriques, thèse de doctorat de l'université de France ,2002.

[12] CH. Barillot, Fusion de Données et Imagerie 3D en Médecine, thèse de doctorat, université de Rennes 1, institut de formation supérieur en informatique et en communication ,1999.

[13] Pluim et al , " Image Registration by maximisation of combined mutual information and gradient information ," IEEE Transaction on Medical Imaging, Vol 19, No 8, pp 809-814 August 2000.