

Méthode supervisée de labellisation d'images par utilisation de similarités semi-locales: Application aux images IRM cérébrales.

François ROUSSEAU¹, Piotr HABAS², Colin STUDHOLME²

¹Laboratoire des Sciences de l'Images, de l'Informatique et de la Télédétection
Pôle API, Bd S. Brant, 67412, Illkirch, France

²Department of Pediatrics, Division of Neonatology
University of Washington, Box 356320, Seattle, WA 98195 USA
rousseau@unistra.fr, p.habas@gmail.com
studholm@uw.edu

Résumé – Nous proposons dans ce travail une nouvelle méthode de segmentation d'images reposant sur le principe de propagation de labels. En utilisant les similarités semi-locales entre les images d'une base d'apprentissage et l'image à segmenter, une stratégie ne requérant aucune étape de recalage non-rigide est présentée. Dans la continuité des méthodes de moyennes non-locales proposées pour le débruitage d'images, les similarités entre images sont représentées par un graphe pondéré calculé à partir de distance entre patches. Les expériences sur données IRM in-vivo ont montré l'intérêt de l'approche proposée afin d'obtenir une labellisation automatique du cerveau humain.

Abstract – We propose in this work a patch-based segmentation method relying on a label propagation framework. Based on image intensity similarities between the input image and a learning dataset, an original strategy which does not require any non-rigid registration is presented. Following recent developments in non-local image denoising, the similarity between images is represented by a weighted graph computed from intensity-based distance between patches. Experiments on in-vivo MR images show that the proposed method is very successful in providing automated human brain labeling.

1 Introduction

La labellisation automatique d'images IRM cérébrales est une étape importante pour les études de neuro-imagerie car elle permet d'obtenir une carte de multiples régions contiguës du cerveau. Un des outils pour obtenir ces cartes repose sur la propagation de labels. Ceci s'effectue en recalant une ou plusieurs images (dont on a une segmentation) sur l'image à segmenter. Les cartes de segmentation (ou labels) sont ensuite propagées en appliquant les transformations des recalages estimés. Le principe général est illustré par la figure 1. Cette approche est très flexible puisqu'elle nécessite seulement quelques exemples (i.e. un ensemble d'images anatomiques et leurs segmentations correspondantes). Les points clés de cette approche concernent la précision de l'algorithme de recalage non-rigide, les règles de fusion des cartes de segmentation, la sélection des images de l'ensemble d'apprentissage et les erreurs contenues dans les segmentations de cet ensemble.

De par l'utilisation d'algorithmes de recalage non-rigides, il est souvent fait l'hypothèse de transformations bijectives. Cependant, cette contrainte peut mener à des erreurs de propagation lorsque cette hypothèse n'est pas vérifiée. De plus, la qualité et la précision du recalage va influencer la segmentation finale. Enfin, il a été montré récemment qu'il y a une grande variabilité en terme de qualité entre les algorithmes de recalage

existants [8]. Ainsi, même si de multiples algorithmes de recalage sont disponibles à travers Internet, cette étape demeure complexe à mettre en œuvre et est relativement gourmande en temps de calcul. Dans ce travail, nous proposons une stratégie alternative qui ne nécessite pas d'étape de recalage non-rigide. Pour ce faire, nous utilisons des similarités semi-locales entre l'image à segmenter et l'ensemble d'apprentissage, dans l'esprit des méthodes non-locales développées récemment pour le débruitage d'images [4]. Le principe est donc de relâcher la contrainte de transformation bijective en utilisant une mise en correspondance floue.

2 Propagation de labels par similarité entre blocs

Soit \mathcal{T} , l'ensemble d'apprentissage contenant un ensemble d'images anatomiques \mathcal{I} et leurs segmentations correspondantes $\mathcal{L} : \mathcal{T} = \{(\mathcal{I}_i, \mathcal{L}_i), i = 1, \dots, n\}$. Considérons un graphe w_i sur le domaine image Ω , qui relie les voxels \mathbf{x} de l'image à segmenter I et les voxels \mathbf{y} de l'image \mathcal{I}_i , avec le poids $w_i(\mathbf{x}, \mathbf{y})$, $(\mathbf{x}, \mathbf{y}) \in \Omega^2$. Ces poids sont calculés de la façon suivante :

$$w_i(\mathbf{x}, \mathbf{y}) = f \left(\sum_{\mathbf{x}' \in \mathcal{P}_I(\mathbf{x}), \mathbf{y}' \in \mathcal{P}_{\mathcal{I}_i}(\mathbf{y})} (I(\mathbf{x}') - \mathcal{I}_i(\mathbf{y}'))^2 \right), \quad (1)$$

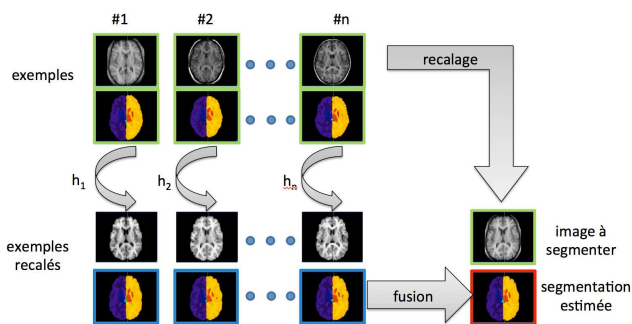


FIGURE 1 – Principe des méthodes de propagation de labels reposant sur un recalage non-rigide. Les données d’entrée de l’algorithme (bordure verte) sont un ensemble d’images segmentées, et l’image I que l’on souhaite segmenter. L’ensemble d’images est recalé sur l’image I et les images de labels sont déformées suivant les transformations non-rigides estimées h_i . La segmentation estimée (bordure rouge) est obtenue en fusion des images de labels déformées (bordure bleue).

où $P_I(\mathbf{x})$ est un bloc (ou patch 3D) de l’image I centré au voxel \mathbf{x} ; f est une fonction noyau. Ici, $f(x) = e^{-\frac{x}{2N\beta\sigma^2}}$, où N est le nombre de voxels contenus dans le bloc 3D; σ est l’écart-type du bruit et β est un paramètre de lissage. En supposant que le bruit soit gaussien, β peut être choisi égal à 1 (voir [4] pour les justifications théoriques) et l’écart-type du bruit σ est estimé à l’aide des pseudo-résidus [6].

Le graphe pondéré w_i représente les similarités non-locales entre l’image à segmenter I et l’image \mathcal{I}_i de l’ensemble d’apprentissage. La méthode de segmentation par propagation de labels que nous proposons se fonde sur le principe suivant : *si les blocs de l’image à segmenter I sont localement similaires à des blocs de l’ensemble d’apprentissage \mathcal{T} , alors ils doivent avoir un label similaire*. Dans le contexte du débruitage d’images [4], Buades *et coll.* supposent que pour chaque bloc de l’image, il existe des blocs similaires dans cette même image. Ce principe de redondance est à l’origine de la méthode des moyennes non-locales. Le graph w est alors calculé sur toute l’image. Cependant, dans notre contexte de segmentation cérébrale, il ne semble pas judicieux de calculer w sur l’ensemble du domaine Ω . Ainsi, le graphe w n’est calculé que pour un proche voisinage : $w = \{w(\mathbf{x}, \mathbf{y}), \forall \mathbf{x} \in \Omega, \mathbf{y} \in \mathcal{N}(\mathbf{x})\}$, où $\mathcal{N}(\mathbf{x})$ est le voisinage du voxel \mathbf{x} . La taille de ce voisinage est directement reliée à la variabilité des structures à segmenter.

En nous appuyant sur le même principe que les moyennes non-locales, la carte de segmentation de l’image I peut être obtenue de la façon suivante :

$$\forall \mathbf{x} \in \Omega, L(\mathbf{x}) = \frac{\sum_{i=1}^n \sum_{\mathbf{y} \in \mathcal{N}(\mathbf{x})} w_i(\mathbf{x}, \mathbf{y}) \mathcal{L}_i(\mathbf{y})}{\sum_{i=1}^n \sum_{\mathbf{y} \in \mathcal{N}(\mathbf{x})} w_i(\mathbf{x}, \mathbf{y})}. \quad (2)$$

où $\mathcal{L}_i(\mathbf{y})$ est un vecteur de $[0, 1]^M$ (M est le nombre total de labels) représentant la proportion de chaque label pour le voxel \mathbf{y} de l’image \mathcal{L}_i . $L(\mathbf{x}) = (l_1(\mathbf{x}), l_2(\mathbf{x}), \dots, l_M(\mathbf{x}))$ représente la proportion de chaque label pour le voxel \mathbf{x} dans l’image I , tel que $\sum_k l_k(\mathbf{x}) = 1$ et $l_k(\mathbf{x}) \in [0, 1], \forall k \in [1, M]$. Ainsi, pour

chaque voxel \mathbf{x} de I , l’équation (2) fournit une segmentation floue de l’image I puisque $L(\mathbf{x}) \in [0, 1]^M$. Une segmentation « dure » H de I peut être obtenue en prenant la composante de $L(\mathbf{x})$ de plus grande valeur :

$$h_k(\mathbf{x}) = \begin{cases} 1 & \text{si } l_k(\mathbf{x}) = \|L(\mathbf{x})\|_\infty \\ 0 & \text{sinon.} \end{cases} \quad (3)$$

De part l’utilisation de blocs se recouvrant partiellement, il est possible de modifier l’équation 2 pour obtenir une estimation « multi-point » des labels :

$$\forall \mathbf{x} \in \Omega, P_L(\mathbf{x}) = \sum_{i=1}^n \sum_{\mathbf{y} \in \mathcal{N}(\mathbf{x})} w_i(\mathbf{x}, \mathbf{y}) P_{\mathcal{L}_i}(\mathbf{y}). \quad (4)$$

où $P_{\mathcal{L}_i}(\mathbf{y})$ est un bloc de la carte de labels \mathcal{L}_i centré au voxel \mathbf{y} , et $P_L(\mathbf{x})$ est une estimation pour le bloc centré au voxel \mathbf{x} . Ainsi, à cause du recouvrement des blocs de labels, pour chaque voxel $\mathbf{x} \in \Omega$, nous obtenons N estimées. Ces N estimées peuvent être agrégées en utilisant des règles de fusion, telle que la règle de la majorité par exemple.

3 Expériences

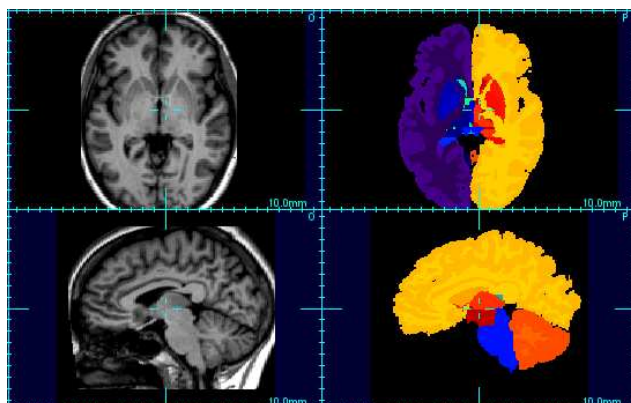


FIGURE 2 – Exemple d’image IRM pondérée en T1 provenant de la base IBSR (vues axiales et sagittales) et la segmentation manuelle en 32 structures d’intérêt correspondante.

Nous avons utilisée la base d’images IBSR, disponible sur Internet ¹, qui sert de référence pour l’évaluation d’algorithmes de segmentation de structures cérébrales en IRM. Les images anatomiques et les segmentations manuelles correspondantes sont fournies par le « Center for Morphometric Analysis at Massachusetts General Hospital ». L’ensemble contient 18 images de cerveaux sains et leur segmentation correspondante en 32 structures. Un exemple d’images est présenté sur la figure 2. Ces données ont été pré-traitées de la façon suivante : 1) correction du biais par l’algorithme N3 en utilisant MIPAV ², 2)

1. <http://www.cma.mgh.harvard.edu/ibsr>
2. <http://mipav.cit.nih.gov>

correction d’histogramme en utilisant 3DSlicer³, 3) recalage affine en utilisant ANTs⁴.

L’évaluation de la méthode s’effectue par validation croisée. Chaque image est sélectionnée comme image à segmenter et les 17 images restantes forment l’ensemble d’apprentissage. Pour chaque segmentation, afin de comparer avec d’autres méthodes de la littérature, nous avons utilisé l’index Dice (qui mesure le recouvrement entre deux segmentations) comme mesure de qualité :

$$DI = \frac{2.VP}{2.VP + FP + FN},$$

où VP est le nombre de vrais positifs, FP correspond aux faux positifs et FN aux faux négatifs. L’index de Dice est calculé pour les segmentations dures.

Pour ces expériences, nous avons utilisé des blocs de taille $3 \times 3 \times 3$ voxels et un voisinage local de $11 \times 11 \times 11$ voxels. Les résultats sont résumés dans le tableau 3, où sont également reportés les résultats d’autres méthodes de la littérature (utilisant des approches par recalage non-rigide ou des approches de segmentation par modélisation d’intensité des structures). Il est clairement montré que la méthode proposée est très compétitive par rapport aux méthodes récemment proposées.

Les figures 3 et 4 permettent une comparaison visuelle entre les résultats obtenus par la méthode proposée et ceux obtenus avec une propagation de labels par recalage non-rigide. La première ligne de la figure 3 correspond à l’image moyenne de l’ensemble des images d’apprentissage. Ce type d’image est fréquemment employé pour évaluer les méthodes de recalage non-rigide. Pour la méthode par patches, cette image est obtenue en calculant la moyenne pondérée des intensités au lieu d’utiliser les labels des images de la base d’apprentissage. La seconde ligne correspond aux résultats de labellisation. Enfin, la troisième ligne indique les différences entre les méthodes automatiques et la vérité terrain. Cette figure illustre le fait que la méthode proposée permet d’obtenir une meilleure segmentation des structures cérébrales, notamment au niveau du cortex.

Nous avons également comparé les temps de calcul de la méthode proposée avec une méthode de recalage (ANTs [3]) performante [8]. En utilisant un seul cœur d’un processeur Intel Xeon E5420 2.5GHz, l’approche présentée ici est de 30% à 50% plus lente que ANTs (60 minutes contre 40 minutes pour traiter une image de la base d’apprentissage). En revanche, en utilisant le fait que les patches se recouvrent partiellement, il est possible de réduire ce temps de 60 à 10 minutes. Enfin, puisque chaque voxel est traité séparément, l’algorithme est très simple à paralléliser. Ainsi, en utilisant un processeur à 8 cœurs, on arrive à un temps de calcul de 1,2 minutes par image de la base d’apprentissage, soit environ 20 minutes pour exploiter toutes les images de la base IBSR.

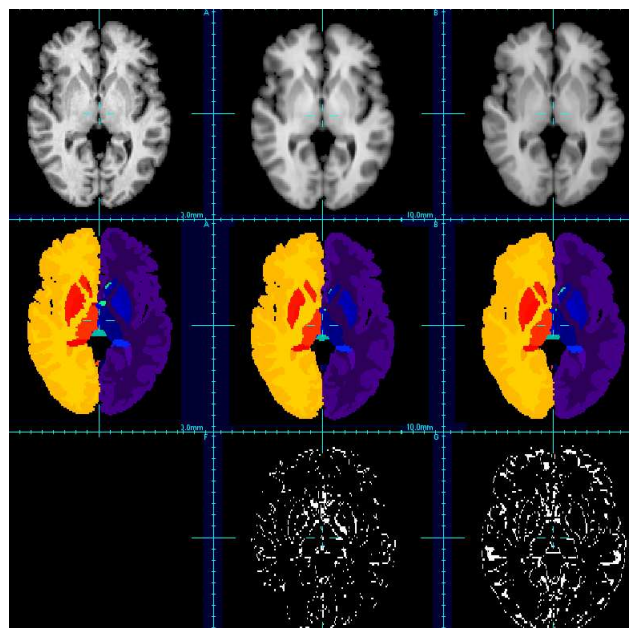


FIGURE 3 – Évaluation visuelle des résultats de segmentation (IBSR, image #7). Première ligne : images pondérées en T1, seconde ligne : résultat de segmentation, troisième ligne : voxels mal classés. Gauche : vérité terrain, milieu : méthode par patches, droite : approche par recalage non-rigide (en utilisant ANTs).

4 Discussion

La méthode proposée dans cet article permet de relâcher la contrainte de transformation bijective utilisée fréquemment par les méthodes de recalage non-rigide. Ainsi, si une des images de la base d’apprentissage ne fournit pas d’exemples pertinents pour l’image à segmenter, celle-ci ne sera pas prise en compte pour l’estimation de la carte de segmentation (en effet, les poids associés seront faibles). Les résultats obtenus sur la base IBSR montrent clairement que la méthode proposée dans cet article permet d’obtenir des segmentations de qualité, quelque soit la structure cérébrale à segmenter.

Reposant sur l’hypothèse de similarité (en intensité) entre les images de l’ensemble d’apprentissage et l’image à segmenter, la qualité de la segmentation peut dépendre du contraste

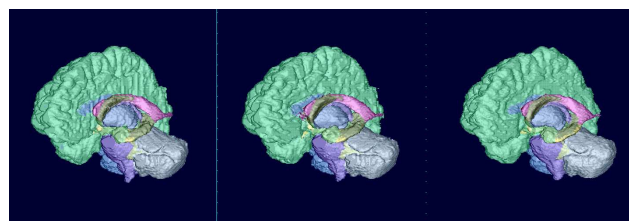


FIGURE 4 – Rendu de surface des résultats de segmentation (IBSR, image #7). Gauche : vérité terrain, milieu : méthode par patches, droite : approche par recalage non-rigide (en utilisant ANTs).

3. <http://www.slicer.org>

4. <http://www.picsl.upenn.edu/ANTS>

TABLE 1 – Comparaison (à partir de l’index Dice moyen) de l’approche proposée avec d’autres méthodes récentes de segmentation. L’index le plus élevé est indiqué en gras pour chaque structure.

structure cérébrale (gauche/droite)	Méthode proposée	LWV-MSD 2009 [2]	Fischl et al. 2002 [5]	Khan et al. 2009 [7]	MA-IDMIN+ 2010 [9]	Sdika 2010 [10]	Akselrod 2007 [1]
MB du cerveau	0.93 - 0.92	0.78 - 0.78	-	-	-	0.91	0.87
MG du cerveau	0.95 - 0.95	0.81 - 0.81	-	-	-	0.94	0.86
Ventricule Latéral	0.93 - 0.92	0.83 - 0.82	0.78	0.85	-	0.90	-
Vent. Lat. Inf.	0.60 - 0.56	0.22 - 0.22	-	-	-	0.44	-
MB du cervelet	0.88 - 0.88	0.79 - 0.79	-	-	-	0.86	-
MG du cervelet	0.95 - 0.95	0.86 - 0.86	-	-	-	0.95	-
Thalamus	0.89 - 0.89	0.87 - 0.88	0.86	0.89	0.89	0.89	0.84
Noyau caudé	0.88 - 0.89	0.83 - 0.83	0.82	0.83	0.85	0.86	0.80
Putamen	0.89 - 0.89	0.86 - 0.86	0.81	0.87	0.90	0.88	0.79
Globus pallidus	0.79 - 0.79	0.78 - 0.79	0.71	0.72	0.83	0.79	0.74
3ème ventricule	0.80	0.74	-	-	-	0.77	-
4ème ventricule	0.84	0.77	-	-	-	0.81	-
Tronc cérébral	0.93	0.91	-	-	-	0.93	0.84
Hippocampus	0.83 - 0.83	0.74 - 0.76	0.75	0.76	0.80	0.81	0.69
Amygdale	0.75 - 0.75	0.72 - 0.72	0.68	0.66	0.75	0.74	0.63
LCR	0.68	0.61	-	-	-	0.70	0.83
Noyau accumbens	0.68 - 0.66	0.67 - 0.67	0.58	0.61	-	0.72	-
Ventral DC	0.82 - 0.82	0.82 - 0.82	-	-	-	0.85	0.76

de la structure d’intérêt. Les variations d’intensité peuvent perturber la recherche de patchs similaires, ce qui conduirait à une labellisation finale obtenue à partir de patchs moins pertinents. Cette sensibilité peut notamment être modifiée à l’aide du paramètre de lissage β . Cependant, l’utilisation d’une valeur globale de β et d’une méthode de correction de biais associée à une correction d’histogrammes a permis d’obtenir des résultats très encourageants. Il peut être noté que la segmentation du pallidum n’est pas aussi satisfaisante que pour les autres structures. En effet, cette structure est très peu contrastée. Ainsi, contrairement aux méthodes reposant sur une étape de recalage non-rigide (qui permet de propager les relations spatiales entre structures d’intérêt), une méthode fondée uniquement sur l’intensité des images ne peut pas produire de résultats pleinement satisfaisants. Ceci fait clairement apparaître la possible complémentarité des deux approches, au prix d’un coût calculatoire fortement augmenté.

Remerciements Ces travaux de recherche ont été supportés par le conseil européen de la recherche (FP7/2007-2013 Grant Agreement no. 207667) et le NIH/NINDS Grant R01 NS 055064.

Références

- [1] A. Akselrod-Ballin, M. Galun, J. Gomori, A. Brandt, and R. Basri. Prior knowledge driven multiscale segmentation of brain MRI. *MICCAI*, 10(Pt 2) :118–126, 2007.
- [2] X. Artaechevarria, A. Munoz-Barrutia, and C. Ortiz-de-Solorzano. Combination strategies in multi-atlas image segmentation : application to brain MR data. *IEEE Transactions on Medical Imaging*, 28(8) :1266–1277, August 2009.
- [3] B.B. Avants, C.L. Epstein, M. Grossman, and J.C. Gee. Symmetric diffeomorphic image registration with cross-correlation : Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, 12(1) :26–41, February 2008.
- [4] A. Buades, B. Coll, and J. M. Morel. A review of image denoising algorithms, with a new one. *Multiscale Modeling & Simulation*, 4(2) :490–530, January 2005.
- [5] B. Fischl et coll. Whole brain segmentation : automated labeling of neuroanatomical structures in the human brain. *Neuron*, 33(3) :341–355, January 2002.
- [6] T. Gasser, L. Sroka, and C. Jennen-Steinmetz. Residual variance and residual pattern in nonlinear regression. *Biometrika*, 73(3) :625–633, December 1986.
- [7] A. R. Khan, M. K. Chung, and M. F. Beg. Robust atlas-based brain segmentation using multi-structure confidence-weighted registration. *MICCAI*, 12(Pt 2) :549–557, 2009.
- [8] A. Klein et coll. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *NeuroImage*, 46(3) :786–802, July 2009.
- [9] J. Lötjönen, R. Wolz, J. R. Koikkalainen, L. Thurfjell, G. Waldemar, H. Soininen, and D. Rueckert. Fast and robust multi-atlas segmentation of brain magnetic resonance images. *NeuroImage*, 49(3) :2352–2365, February 2010.
- [10] M. Sdika. Combining atlas based segmentation and intensity classification with nearest neighbor transform and accuracy weighted vote. *Medical Image Analysis*, 14(2) :219–226, April 2010.