

Robust Brain Registration Using Adaptive Probabilistic Atlas

Jaime Ide¹, Rong Chen¹, Dinggang Shen^{1,2}, and Edward H Herskovits¹

¹ Department of Radiology, University of Pennsylvania, Philadelphia, PA
{jaime.ide,rong.chen}@uphs.upenn.edu

² Department of Radiology and Biomedical Research Imaging Center,
University of North Carolina, Chapel Hill, NC
dgshen@med.unc.edu

Abstract. Elastic image registration is widely used to adapt brain images to a common template space, and, in complementary fashion, to adapt an anatomical template to a subject’s anatomy. Although HAMMER is a very accurate image-registration algorithm, it requires a 3-class segmentation step prior to registration, and its performance is affected by segmentation quality. We here propose a new framework to improve this algorithm’s robustness to poor initial segmentation. Our new framework is based on *Adaptive Generalized Expectation Maximization* (AGEM) for unified segmentation and registration, in which we use an *adaptive* strategy to incorporate spatial information from a probabilistic atlas to improve segmentation and registration simultaneously. Our experiments using real MR brain images indicate that our integrated approach improves registration accuracy; we have also found that our iterative approach renders HAMMER robust to low tissue contrast, which hinders 3-class segmentation.

1 Introduction

Image registration [1,2,3,4] plays a central role in the field of medical image analysis [5,6,7]. In particular, it is used to adapt an anatomical atlas to an individual subject’s anatomy for performing atlas-based segmentation, which is especially important for difficult segmentation cases [8]. As one would expect, the performance of atlas-based segmentation is strongly related to the quality of atlas registration. HAMMER, a nonlinear warping algorithm developed by Shen and Davatzikos, has been proven to register brain MR images with high accuracy [9]. However, HAMMER requires a 3-class (GM/WM/CSF) segmentation as input, and HAMMER’s performance is affected by the quality of tissue segmentation.

In this paper, we propose an *Adaptive Generalized Expectation Maximization* (AGEM) framework for unified segmentation and registration. Probabilistic unified and iterative approaches¹ for segmentation and registration are considered to be an instance of the *Maximum a Posteriori* (MAP) estimation problem.

¹ In our approach, we use a probabilistic atlas rather than an atlas based on a single subject or standard.

The unification of segmentation and registration has been proposed independently by several research groups [10,11,12]. For example, Wyatt et al. [10], focused on rigid registration of synthetic validation images, and MRI mouse-heart images. More-complex registration problems were considered by Chen et al. and Pohl et al. [11,12]. Chen et al. extended the previous work of Wyatt et al. to include non-rigid B-spline based registration. In this paper, we present a novel solution for the segmentation-step (E-step) using an *entropy-based adaptive* atlas segmentation strategy, and using HAMMER-based registration technique [9] for the registration-step (M-step). Our MAP/EM framework differs fundamentally from previous work to solve the registration parameter estimation. Pohl et al. defined a *Kullback Leibler* based similarity measure [12, page 232], Chen et al. used a *mutual information* based similarity measure [11], whereas we define an *attribute vector* based similarity measure [9]. Pohl et al. employed a *structure-dependent* hierarchical registration, whereas we employ a *hierarchical attribute matching* registration. Moreover, similar to Wyatt et al. [10], we propose an adaptive strategy during the segmentation step. Wyatt et al. used the *class entropy convergence ratio* as their decision criterion, and we use the *probabilistic atlas class entropy* as our decision criterion to improve segmentation. Thus the main difference and contribution of this work is to incorporate an accurate elastic registration method, HAMMER, to solve the MAP segmentation and registration problems. Our experimental results demonstrate that our AGEM framework improves both the robustness and accuracy of registration, thus facilitating the construction of better probabilistic atlases for tissue segmentation.

This paper is organized as follows. In Section 2, we present the theory and implementation of our Adaptive Generalized Expectation Maximization framework. In Section 3, we present validation experiments for our approach. In particular, we demonstrate that our method manifests robust convergence (Section 3.1), that our approach outperforms existing methods such as HAMMER (Section 3.2), and that its performance is robust to low inherent tissue contrast (Section 3.3).

2 Method

In this section we formulate the MAP estimation framework to unify segmentation and registration, and introduce our Adaptive Generalized EM solution.

Let $\mathbf{S} = \{S_1, \dots, S_n\}$ represent the observed MR image intensity of a subject. Assume that we have a probabilistic atlas, $p(\mathbf{T})$, defined over the template image $\mathbf{T} = \{T_1, \dots, T_n\}$, where T_i represents a voxel i on a predefined label space Ω . Our goal is to obtain a labeled image $\mathbf{L} = \{L_1, \dots, L_n\}$ based on \mathbf{S} (intensity information) and \mathbf{T} (spatial information), where L_i assumes values from Ω . In order to use \mathbf{T} , we require a registration process to generate a deformation field \mathbf{D} . The deformation field \mathbf{D} takes each coordinate from the template image and maps it onto a coordinate in the subject's space. Once \mathbf{D} is known, the atlas can be warped into registration with the subject image, thereby generating a warped probabilistic atlas $p(\mathbf{L}|\mathbf{D})$.

Our problem domain contains three variables: \mathbf{S} , \mathbf{L} and \mathbf{D} . Since \mathbf{S} is observed, the maximum a posteriori (MAP) estimation of \mathbf{L} and \mathbf{D} is the mode of $p(\mathbf{L}, \mathbf{D}|\mathbf{S})$. This MAP estimation is intractable because of the sheer volume of the candidate-model space. Therefore, we compute the maximum a posteriori marginal (MPM) estimation of \mathbf{L} and \mathbf{D} , taking the mode of $p(\mathbf{L}|\mathbf{D}, \mathbf{S})$ and $p(\mathbf{D}|\mathbf{L}, \mathbf{S})$. We approximate MPM estimation by using the Generalized EM (GEM) algorithm [13].

Similar to the Bayesian framework presented in [10,12], we assume conditional independence between \mathbf{D} and \mathbf{S} given \mathbf{L} . Then the two steps of our EM algorithm are:

a) Segmentation step (E-step). In this step we compute the posterior distribution $p(\mathbf{L}|\mathbf{D}, \mathbf{S})$ given \mathbf{D} , based on

$$p(\mathbf{L}|\mathbf{S}, \mathbf{D}) = k \cdot p(\mathbf{S}|\mathbf{L}) \cdot p(\mathbf{L}|\mathbf{D}) \cdot p(\mathbf{D}), \quad (1)$$

where k is a normalization constant. The relationship between \mathbf{S} and \mathbf{L} is modeled by a Gaussian mixture model (GMM) $p(\mathbf{S}|\mathbf{L})$, where for a given voxel intensity S_i and class $k \in \Omega$, we have $p(S_i|\mathbf{L} = k) = \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp(-\frac{(S_i - \mu_k)^2}{2\sigma_k^2})$, where μ_k is the mean intensity of class k and σ_k^2 is its variance. We also use the EM approach to approximate the GMM solution. We compute $p(\mathbf{L}|\mathbf{D})$ by warping the probabilistic atlas $p(\mathbf{T})$ to the subject's image volume. For $p(\mathbf{D})$, we assume a uniform distribution².

b) Registration step (M-step). In this step we compute the mode of $p(\mathbf{D}|\mathbf{L}, \mathbf{S})$, given the updated \mathbf{L} computed in the previous step; we accomplish this goal by computing \mathbf{D} that maximizes the expected value of *complete-data log-likelihood* $\log p(\mathbf{L}, \mathbf{S}|\mathbf{D})$:

$$\mathbf{D} \leftarrow \operatorname{argmax}_{\mathbf{D}} E[\log p(\mathbf{L}, \mathbf{S}|\mathbf{D})] = \operatorname{argmax}_{\mathbf{D}} E[\log p(\mathbf{L}|\mathbf{D})]. \quad (2)$$

Generally, there is no analytical solution for Equation (2), so we formulate and solve this optimization problem in a more general way: instead of using a gradient-based optimization technique, such as Powell's method (used by [12]), we use a hierarchical attribute matching mechanism [9] (i.e., HAMMER) to optimize Equation (2). This solution is non-linear: the energy function is based on similarity of attribute vectors between the mode of $p(\mathbf{T})$ and $p(\mathbf{L}|\mathbf{S}, \mathbf{D})$. In fact, for a given expected value of \mathbf{L} , the deformation field \mathbf{D}^* that maximizes $E[\log p(\mathbf{L}|\mathbf{D})]$ is exactly the deformation field given by the registration method that matches image \mathbf{T} to label-field \mathbf{L} , according to some metric.

Probabilistic atlases encode brain-structure variability as discussed in [14]; consequently, when we take the mode of $p(\mathbf{T})$ and $p(\mathbf{L}|\mathbf{S}, \mathbf{D})$, register them and obtain the warped probabilistic atlas $p(\mathbf{L}|\mathbf{S}, \mathbf{D})$ to update segmentation (Equation(1)), in some cases this process may add misleading spatial information; we illustrate this situation in Figure 1. We assume three tissue classes

² We assume that there is not prior knowledge. It is possible to use a prior knowledge encoded in $p(\mathbf{D})$ if we construct it from a sample and consider anatomical variations of the brain.

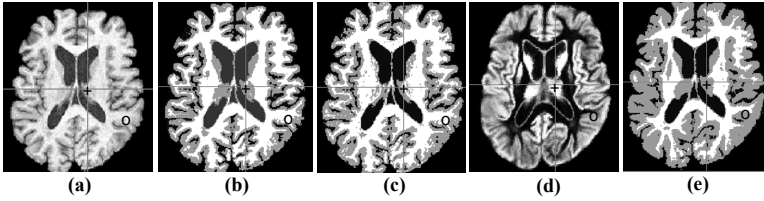


Fig. 1. Illustration of the case in which a warped probabilistic atlas is helpful (voxel marked by “+”) and a case in which a probabilistic atlas hinders segmentation (voxel marked by “o”). We show six images from the same subject at the same slice $z = 62$: (a) raw intensity image, (b) BLSA segmentation result (gray = GM, white = WM), (c) GMM segmentation, (d) the warped GM probabilistic atlas (white color = 1.0, black = 0.0), and (e) posterior segmentation using Equation(1).

$\Omega = \{WM, GM, CSF\}$. Figure 1(a) shows a raw-intensity T1-weighted brain image. Figure 1(b) demonstrates the result from the BLSA³ data set, which consisted of images carefully segmented into 3 classes before presentation to the registration algorithm. This algorithm classifies voxel “o” as WM and voxel “+” as GM. Using GMM and segmenting the subject based solely on intensity (Figure 1(c)), we classify “o” correctly as WM and “+” incorrectly as WM.

If we augment segmentation (Equation(1)) with spatial information (i.e., the warped probabilistic atlas), we obtain the segmentation shown in Figure 1(e) which classifies “o” incorrectly as GM and “+” correctly as GM. We obtain this result because GM in the probabilistic atlas (Figure 1(d)) asserts that “o” is GM with probability 0.62, which is enough to override the intensity information. Comparing GMM segmentation (Figure 1(c)) and MAP segmentation(Figure 1(e)), we find that the latter improves the initial noisy segmentation, particularly for subcortical regions of the brain, at the cost of structural details in cortical regions.

To improve previously described misclassifications, we propose an *adaptive* strategy to add the atlas’s spatial information based on its *entropy* $\mathbf{H}(p(\mathbf{T}))$, during the GEM iteration in Equation (1). The voxel-wise tissue class distribution entropy is defined by expression $\mathbf{H}(p(T_i)) = -\sum_{k \in \Omega} p(T_i = k) \cdot \log p(T_i = k)$, where Ω is the label space and i is the template voxel index. The proposed adaptive strategy has two components:

1) We adaptively compute posterior distributions of voxel classifications during segmentation updating (the E-step). For different regions of the brain, we use different updating expressions. In particular, for high-entropy regions of the warped atlas, we rely primarily on intensity information, whereas for low-entropy regions we rely primarily on atlas information.

2) We adaptively decrease the influence of atlas information across iterations. Initially, we rely primarily on the atlas’s spatial information, assuming that segmentation based solely on intensity information is noisy, but in successive

³ We will describe the BLSA data set [15] in Section 3.

iterations, since updated segmentations are less noisy, we rely more on intensity information to capture fine structural details.

Our **Adaptive GEM algorithm (AGEM)** is summarized as follows:

Input: subject image \mathbf{S} and a probabilistic atlas $p(\mathbf{T})$.

Output: labeled image \mathbf{L} of subject and a deformation field \mathbf{D} mapping template to subject space.

1. Load \mathbf{S} and $p(\mathbf{T})$.
2. Initialize label $\mathbf{L}^{(0)}$ using GMM, based only on intensity \mathbf{S} .
3. Start with $i = 0$ and repeat until the mode $\mathbf{L}^{(i)}$ of posterior segmented image distribution $p(\mathbf{L})^{(i)}$ converges:
 - 3.1. Run registration algorithm to compute the deformation field $\mathbf{D}^{(i)}$ between mode of $p(\mathbf{L})^{(i)}$ and $p(\mathbf{T})$;
 - 3.2. Apply $\mathbf{D}^{(i)}$ over $p(\mathbf{T})$ and get the warped probabilistic atlas $p(\mathbf{T})^* \leftarrow p(\mathbf{L}|\mathbf{D}^{(i)})$;
 - 3.3. if $(\mathbf{H}(p(\mathbf{T})^*) < h_t(i))$ update distribution $p(\mathbf{L})^{(i+1)} \leftarrow p(\mathbf{L}|\mathbf{S}, \mathbf{D}^{(i)}) = p(\mathbf{S}|\mathbf{L}).p(\mathbf{T})^*$, else update with $p(\mathbf{L})^{(i+1)} \leftarrow p(\mathbf{L}|\mathbf{S}, \mathbf{D}^{(i)}) = p(\mathbf{S}|\mathbf{L})$;
 - 3.4. update label $\mathbf{L}^{(i+1)}$ with mode of $p(\mathbf{L})^{(i+1)}$.
4. Return deformation field $\mathbf{D}^{(i+1)}$ and label $\mathbf{L}^{(i+1)}$.

At Line 2, we initialize the label distribution $p(\mathbf{L})$ based solely on the intensity information of image \mathbf{S} , assuming a GMM. At Line 3.1, we run HAMMER to register the atlas to the subject. After registration, we obtain an updated deformation field $\mathbf{D}^{(i)}$, which we apply over the probabilistic atlas $p(\mathbf{T})$, resulting in a warped probabilistic atlas $p(\mathbf{L}|\mathbf{D}^{(i)})$ (Line 3.2). Then, in Line 3.3, we compute the posterior label distribution $p(\mathbf{L})^{(i+1)}$, using the prior intensity information $p(\mathbf{S}|\mathbf{L}^{(i)})$ and the warped spatial atlas information $p(\mathbf{L}|\mathbf{D}^{(i)})$, if this has entropy lower than threshold $h_t(i)$. The idea is to decrease $h_t(i)$ as we increase the iteration number i . In Section 3.2, we refer to AGEM without the adaptive strategy (Line 3.3) as GEM algorithm.

3 Results

To validate AGEM, we performed three experiments: a) convergence analysis, in which we applied AGEM to 5 BLSA subjects [15] with two different segmentation initializations, to show that in both cases we obtain similar deformation fields upon convergence; b) we applied AGEM to 10 BLSA subjects, demonstrating that AGEM obtains atlas-based segmentation results as well as the current state of the art; and c) we applied AGEM to very low-contrast images, demonstrating that AGEM yields accurate and robust registration results. In these experiments, we set entropy thresholds $h_t(i)$, starting with $h_t(1) = 1.12$ and ending with $h_t(6) = 0.62$ in a linear way. This interval of values is based on the variation of voxel-wise entropy $\mathbf{H}(p(T_i))$ along the brain. This variation is a property of each constructed probabilistic atlas $p(\mathbf{T})$.

Table 1. Average deformation field difference between \mathbf{D}_1 and \mathbf{D}_2 (in mm) across iterations

		Iteration					
		1	2	3	4	5	6
Hippocampus region	Mean of average difference	3.014	2.255	1.369	1.010	0.824	0.758
	Deviation of average difference	1.177	0.875	0.529	0.343	0.300	0.308

We constructed our probabilistic atlas by using HAMMER to register 80 training subjects from the BLSA⁴ data set to the MNI template [16], measuring the overlap between corresponding segmentations (tissues: WM,GM and CSF), and computing the average percentage, as described in [17].

3.1 Convergence Experiment

The aim of this experiment was to show that AGEM provides deformation fields (obtained in the registration step) that converge to a near-optimal one, despite having started with several different initial segmentations. We selected 5 BLSA subjects; for each, we applied AGEM using each of two different GMM initializations⁵, and obtained two different deformation fields \mathbf{D}_1 and \mathbf{D}_2 . In Table 1, we present average deformation field differences between \mathbf{D}_1 and \mathbf{D}_2 , computed voxel-by-voxel in the hippocampus region across iterations. At iteration 6, for the hippocampus region⁶, average deformation field difference converged to 0.758 mm, which is smaller than 1.097 mm, the cube root of the volume of a voxel⁷; and the maximum deformation field error was 2.762 mm. In this experiment we also computed the rate of *label change* (percentage of brain voxels that have different classifications between consecutive iterations). At iteration 1 the average rate was $9.4 \pm 11.7\%$; it converged to $2.6 \pm 1.0\%$ on iteration 6.

3.2 Performance Experiment

To evaluate the performance of AGEM, we executed experiments similar to those ones described in [18,12], computing overlap ratios (ORs). In fact the segmentation quality measured by the OR reflects registration quality. We selected 10 BLSA subjects (different from the 80 subjects used to construct the probabilistic atlas) for which manual hippocampus labeling (from 2 raters) was available and compared them (using the DICE overlap ratio[19]) with labeling results obtained automatically by AGEM, HAMMER and GEM. The hippocampal OR is defined by the expression $OR = |A \cap B| / |A \cup B|$, where A is the manually labeled

⁴ The BLSA data set [15] studies approximately 100 healthy older-adult subjects, from whom annual 3-D T1-weighted SPGR MR examinations are collected.

⁵ Since the GMM is approximated using the EM method, we must initialize its centers[13].

⁶ The hippocampus plays an important role in neurological studies; in addition, it is a challenging segmentation region since it has relatively low contrast.

⁷ BLSA subjects' T1-weighted images have voxel dimension: $0.9375 \times 0.9375 \times 1.5$ mm.

Table 2. Comparison of overlap ratios (ORs) among manually and automatically labeled hippocampi, using HAMMER, GEM and AGEM to warp the probabilistic atlas, for 10 BLSA subjects. The OR between raters is also shown.

Subject	1	2	3	4	5	6	7	8	9	10
Raters	0.829	0.813	0.823	0.781	0.738	0.830	0.798	0.783	0.820	0.825
HAMMER	0.777	0.693	0.713	0.689	0.683	0.775	0.680	0.686	0.698	0.680
Adaptive GEM	0.754	0.719	0.681	0.665	0.702	0.774	0.694	0.694	0.702	0.681
GEM	0.717	0.739	0.678	0.642	0.667	0.723	0.656	0.670	0.681	0.665

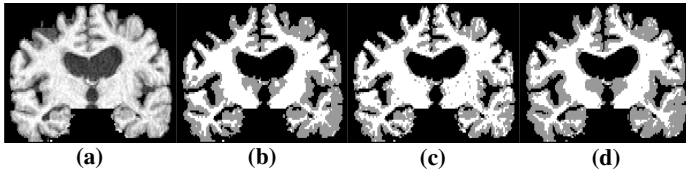


Fig. 2. Segmentation experiment for low-contrast images, year 1: (a) intensity image, (b) segmentation from BLSA, (c) GMM segmentation, (d) AGEM segmentation

hippocampal volume, and B is the automatically labeled volume. For each subject, we segmented the intensity images and registered them to the probabilistic atlas, obtaining a deformation field that mapped that atlas to that subject’s coordinate space. Then we applied this deformation field to warp the hippocampus in the atlas. We compared our ORs with those obtained using HAMMER (using carefully segmented BLSA subjects as input). The results are shown in Table 2; AGEM out-performs GEM in 9 subjects, and yields OR results better than HAMMER in 5 subjects⁸. Table 2 presents OR results obtained using AGEM, varying between 0.64 and 0.77, are comparable to automatic hippocampus labeling results described in [20] where an average of 0.7 ratio is obtained for the same OR measure.

3.3 Low-Contrast-Image Experiment

In this experiment we employed longitudinal images for a BLSA subject, at years 1 and 7. They have low tissue T1 contrast, and are therefore difficult to segment. Figure 2 shows the image for year 1, the segmentation result from BLSA [15], the GMM segmentation result using only intensity information, and the AGEM segmentation result. Note that AGEM (Figure 2(d)) improves the noisy GMM initial segmentation, (Figure 2(c)), and corrects the segmentation of the thalamus.

⁸ We found that AGEM is statistically similar to HAMMER (t-test $p = 0.96$). AGEM performs as well as HAMMER, even when HAMMER starts with a carefully segmented BLSA subject as input.



Fig. 3. Segmentation experiment for low-contrast images, year 07: a) intensity histogram of the whole brain, b) intensity image, c) FAST segmentation, d) segmentation from BLSA, and e) AGEM segmentation

Figure 3 demonstrates the year-7 image, and segmentation results obtained using different methods, for the thalamus. This figure demonstrates that AGEM provides robust segmentation of the thalamus, compared to the BLSA result and FAST [21] (FMRIB’s Automated Segmentation Tool version 3.53).

4 Conclusion

We have demonstrated a method for rendering an accurate elastic registration method, HAMMER, robust to poor segmentation input, using an adaptive generalized-EM algorithm to incorporate probabilistic atlas information. The utility of our proposed method is validated by experimental results. As future work we propose: a) to improve the current adaptive strategy and b) to add neighborhood intensity information during Bayesian segmentation updating, using a Markov Random Field model.

References

1. Bajcsy, R., Lieberman, R., Reivich, M.: A computerized system for the elastic matching of deformed radiographic images to idealized atlas images. *J. Comput. Assisted Tomography* 7(4), 618–625 (1983)
2. Bookstein, F.L.: *Morphometric Tools for Landmark Data*. Cambridge Univ. Press, Cambridge (1991)
3. Davatzikos, C.A., Prince, J.L., Bryan, R.N.: Image registration based on boundary mapping. *SPIE Proc. Image Processing* 15, 112–115 (1996)
4. Modersitzki, J.: *Numerical Methods for Image Registration*. Oxford University Press, Oxford (2004)
5. Gee, J.C., Reivich, M., Bajcsy, R.: Elastically deforming 3-d atlas to match anatomical brain images. *J. Comput. Assist. Tomogr.* 17, 225–236 (1993)
6. Hawkes, D.: Algorithms for radiological image registration and their clinical application. *Journal of Anatomy* 193(3), 347–362 (1998)
7. Scherzer, O.: *Mathematical Models for Registration and Applications to Medical Imaging*. Mathematics in industry: series, vol. 10. Springer, Berlin (2006)
8. Van Leemput, K., Maes, F., VanDermeulen, D., Suetens, P.: Automated model-based bias field correction of mr images of the brain. *IEEE Trans. Med. Imaging* 18(10), 885–895 (1999)

9. Shen, D., Davatzikos, C.: Hammer: Hierarchical attribute matching mechanism for elastic registration. *IEEE Trans. Med. Imaging* 21(8) (2002)
10. Wyatt, P., Noble, J.: Map mrf joint segmentation and registration. In: *MICCAI*, pp. 580–587. Springer, Heidelberg (2002)
11. Chen, X., Brady, M., Rueckert, D.: Simultaneous segmentation and registration for medical image. In: *MICCAI* (1), pp. 663–670 (2004)
12. Pohl, K., Fisher, J., Grimson, W., Kikinis, R., Wells, W.: A bayesian model for joint segmentation and registration. *NeuroImage* 31, 228–239 (2006)
13. Bishop, C.M.: *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer, Heidelberg (2006)
14. Thompson, P., Toga, A.: *Elastic Image Registration and Pathology Detection*. In: *Brain Warping*. Academic Press, London (1999)
15. Goldszal, A.F., Davatzikos, C., Pham, D., Yan, M., Bryan, R.N., Resnick, S.M.: An image processing protocol for the analysis of mr images from an elderly population. *Journal of Computer Assisted Tomography* 22, 827–837 (1998)
16. Talairach, J., Tournoux, P.: *Co-Planar Stereotaxic Atlas of the Human Brain*. Thieme Medical Publishers, Incorporated (1988)
17. Ashburner, J., Friston, K.: Multimodal image coregistration and partitioning - a unified framework. *NeuroImage* 6(3), 209–217 (1997)
18. Ashburner, J., Friston, K.: Unified segmentation. *NeuroImage* 26, 839–851 (2005)
19. Dice, L.R.: Measure of amount of ecological association between species. *Ecology*. 26, 297–302 (1945)
20. Chupin, M., et al.: Anatomically-constrained region deformation for the automated segmentation of the hippocampus. *NeuroImage* 34, 996–1019 (2007)
21. Zhang, Y., Brady, M., Smith, S.: Segmentation of brain mr images through a hidden markov random field model and the expectation maximization algorithm. *IEEE Trans. Med. Imaging* 20(1), 45–57 (2001)