Effect of the Number of Coupled Structures on the Segmentation of Brain Structures from MRI

Alireza Akhondi Asl[1,2] **Hamid Soltanian-Zadeh**[1,2,3]

[1]Control and Intelligent Processing Center of Excellence (CIPCE), School of Electrical and Computer Engineering , University of Tehran, Tehran, Iran

[2]School of Cognitive Sciences, Institute for Studies in Theoretical Physics and Mathematics (IPM), Tehran, Iran

[3]Image Analysis Lab. Radiology Dept., Henry Ford Health System, Detroit, Michigan, USA

a.akhundi@ece.ut.ac.ir hszadeh@ut.ac.ir

Abstract: We propose a three-dimensional, nonparametric, entropy-based, coupled, multi-shape approach to segment subcortical brain structures from magnetic resonance images (MRI). The proposed method uses PCA to develop shape models that capture structural variability. It integrates geometrical relationship between different structures into the algorithm by coupling them (limiting their independent deformations). On the other hand, to allow variations among coupled structures, it registers each structure separately when building the shape models. It defines an entropy-based energy function which is minimized using quasi-Newton algorithm. To this end, probability density functions (pdf) are estimated iteratively using nonparametric Parzen window method. In the optimization algorithm, analytical derivatives are used to improve speed and accuracy. Quantitative results show the improvement in the segmentation quality due to the integration of the coupling information into the segmentation process.

Keywords: Image segmentation, medical image processing, magnetic resonance imaging (MRI), shape modelling, entropy, nonparametric.

1. Introduction

Medical image segmentation is the most important step in visualization, surgical guidance and planning, diagnosis and quantitative measurement [1]. However, many important structures in medical images do not present a clear boundary for segmentation and have variations between different subjects. In addition, imaging methods have limitations such as low signal-to-noise ratio (SNR), partial volume effects, and field inhomogeneities [2]. These problems decrease image segmentation accuracy and make it a complicated problem. Several methods have been introduced in the literature to overcome these limitations. Yet, there is no general-purpose method for segmentation of all structures.

A major category of the methods proposed for the segmentation of brain structures from magnetic resonance images (MRI) optimizes an energy function with several parameters that represent to the underlying shapes. An exciting approach for the optimization of the energy function is based on the partial differential equations. These equations are defined from the derivatives of the energy function with respect to the model parameters. Kass et al [3] introduced the first work in this category which has been improved by others in recent years. For the shape representation, parametric active contours and geometric active contours have been used. In the definition of the energy function, earlier methods use the boundary information for the structures of interest [4]. Later methods use regional information such as intensity histogram (parametric and nonparametric, offline or online) or variance of an area [5]. Others combine boundary and regional information [6]. Recent methods benefit from a priori knowledge about the structures of interest. This makes the segmentation process robust to the imperfect

image conditions [7]-[8]. For the methods developed based on the a priori information, a registration process is essential to integrate the prior model into the segmentation process.

In addition, the anatomical structures in the brain are related to the neighboring structures through their location, size, orientation, and shape. An integration of these relations into the segmentation process improves accuracy and robustness as shown in [9]-[11]. The 3D deformable model introduced in [9] assumes multi-variate Gaussian statistics for the parameters of the prior coupled shape model of multiple structures. In [10], Litvin et al introduced shape distribution as a new concept for coupled object segmentation. Their prior shape model is constructed from a family of shape distributions of related features. Their method is 2D and its extension to 3D is not reported yet. Addition of new terms in their energy function leads to challenges in the calculation of the derivatives for curve evolution.

In this paper, we present a method based on entropy information of intensities. We add the flexibility of transformation to the coupled shapes, by enhancing the model construction process. We separate shape and pose variability. This is because inclusion of both poses and shape variability in a PCA based model generates huge variability and limits the model's benefits for segmentation. In other words, co-variations of shape and pose in the training datasets are too complex that may not be captured well by PCA. We align each of the shape classes in the training datasets individually to extract co-variations between shape classes without considering their spatial locations (poses). We use completely online pdf estimation, updated during iterations of the optimization algorithm. We use pre-register test datasets for segmentation. We use quasi-Newton method which is much faster and more robust for high dimensions than the steepest descent method.

In terms of gradient estimation for the optimization process, we calculate them analytically to benefit from their higher speed and accuracy.

The rest of the paper is organized as follows. Section II explains our shape model and its properties. In Section III, we explore our energy function and the proposed method for its optimization. Experimental results are presented and discussed in Section IV. Finally, in Section V, conclusions of our work are presented.

2. Shape Model

In many segmentation methods, a shape model is used where richer models generate more accurate results. Powerful methods for shape representation are based on distance function, implicit representation, and relationships among different shapes, including pose, orientation, and other geometrical relations [12]. Tsai, et al [11] use PCA to extract co-variations among different structures. Bijari et al [12] use symmetry properties of structures for segmentation. In this paper, we use shape relation and individual transformations for each structure for segmentation. To extract shape relation, we apply principal component analysis (PCA) on the 3D training datasets. To reach accurate results, we individually align each structure of interest in the training datasets.

2.1 Alignment

In knowledge based segmentation methods that use shape variation, alignment is a critical step. This procedure has four important constituents: a transformation matrix, a metric to compare similarity of the moving and fixed (reference) images, an interpolator to compute intensities at points with non-integer coordinates, and an optimizer to find the optimal transformation parameters. For 3D objects, several transformations such as similarity, affine, and Euler can be used for alignment. We use similarity transform with 7 parameters that include a scaling, 3 rotations, and 3 translation parameters. For transformation of points, we use:

$$\begin{bmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \\ 1 \end{bmatrix} = T(\mathbf{t}) R(\boldsymbol{\theta}) S(s) \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (1)$$

where *T, R,* and *S* are the translation, rotation, and scaling matrices, respectively, $\mathbf{t}$ is a vector of translation in 3D, $\boldsymbol{\theta}$ is the rotation vector with three parameters, s is the scaling parameter, $R(\boldsymbol{\theta}) = R_x(\theta_x) R_y(\theta_y) R_z(\theta_z)$. For the metric, several options are found in the literature. Since our goal is to align each structure in the datasets individually, we use a labeled 3D image that represents the desired structure with one (foreground) and other pixels with zero (background). The metric in our implementation is the cardinality metric that computes percent of pixels that are not matched between the moving and fixed images. This metric is equal to zero when all of the pixels are matched.

For interpolation of binary images, we use a nearest neighbor interpolator. For the optimizer, since the proposed metric is non-differentiable, we use an optimizer that does not need derivatives. The Nelder–Mead (amoeba) method is a popular direct search method for minimizing unconstrained real functions with our desired property [13]. This method is based on the comparison of function values at the $n+1$ vertices of a simplex where n is the number of parameters to be optimized. The algorithm changes simplex vertices through reflection, expansion, and contraction operations to find an improving point.

Using the above methods, we extract shape variability of the desired structures for model construction as explained in the following section.

2.2 Implicit Parametric Shape Representation

As stated in [7] and [11], implicit parametric shape representation has advantages such as computational efficiency, accuracy, capturing wide range of shape variability, and handling topological changes. We use a distance map for shape representation that is zero on the boundary of a shape and in other points is the Euclidian distance from the boundary (negative inside, positive outside). After extraction of the distance maps of m desired structures for n different training datasets (ψ_i^k shows the distance map of the kth structure of the ith dataset), we subtract the mean distance map of each structure, computed by averaging of the training datasets ($\bar{\Phi}^k$), from each of the n signed distance maps to remove similar parts in different shapes and show them with $\tilde{\psi}_i^k$. We use these $n \times m$ maps to show variability of different structures in the training dataset. We collect n column vectors of size $m \times N_x \times N_y \times N_z$ and use them to extract n eigenvectors for each of the m structures and show their variability (Φ_i^k). There will be up to n eigenvectors of size $m \times N_x \times N_y \times N_z$.

To allow limited, robust shape variability, we use $q \le n$ eigenvectors to represent each shape. In addition, to consider pose differences, we add 7 pose parameters (for local alignment of the structures) to the shape parameters of each structure. Our experiments show that using a single global transformation is not realistic and different shape models need independent local alignments. Finally, for each structure, we may write:

$$\Phi^k\left[\mathbf{w},\mathbf{p}^k\right](x,y,z) = \bar{\Phi}^k(\tilde{x}_k,\tilde{y}_k,\tilde{z}_k) + \sum_{i=1}^{q} w_i \Phi_i^k(\tilde{x}_k,\tilde{y}_k,\tilde{z}_k) \quad (2)$$

where $\mathbf{w}$ is the vector of eigenvectors multipliers and $\mathbf{p}^k$ is the vector containing 7 transformation parameters for the alignment of the kth structure according to the following equations.

$$\begin{bmatrix} \tilde{x}_k \\ \tilde{y}_k \\ \tilde{z}_k \\ 1 \end{bmatrix} = M(\mathbf{p}^k) \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (3)$$

$$M(\mathbf{p}^k) = T(\mathbf{t}_k + \mathbf{c}_k) R(\boldsymbol{\theta}_k) S(s_k) T(-\mathbf{c}_k)$$

where T, R, and S are the translation, rotation, and scaling matrices, respectively, $\mathbf{t}$ is the vector of translation parameters, and $\mathbf{c}$ is the center of rotation used for scaling and rotation. This center of rotation improves the local alignment accuracy, is fixed, and is computed using the mean of the centers of masses of the structures in the training datasets. In this manner, each structure's pose may change while shape classes co-variations are used for coupling.

In the next section, we present our proposed entropy-based segmentation method using the shape model described above.

3. Segmentation of Structures

After construction of the shape model and shape classes' co-variations, an energy function is defined for the segmentation process. In this section, we explain our energy model and optimization method.

3.1 Energy Model

The proposed entropy-based method classifies image voxels to distinguish regions by minimizing a weighted sum of the conditional differential entropies of different structures. To segment m coupled structures with closed boundaries, there are m regions for these structures. We set the area outside of the m structures as $m+1$ and use this notation throughout the paper. Based on the entropy of these regions $\{\Omega_k,\ k = 1 \cdots m+1\}$, the energy function is defined as

$$J(\Omega_1,\ldots,\Omega_{m+1}) = |\Omega| \sum_{j=1}^{m+1} P_{\Omega_j} \hat{H}(\Omega_j) \quad \text{where} \quad |\Omega|$$

represents the cardinality of a set Ω (number of pixels). When all of the regions are as uniform as possible, the energy function is at its minimum.

Nevertheless, there is the important problem of estimating conditional differential entropies.

We estimate the entropy of the *k*th structure using $\hat{H}\left(\Omega_k\right)=\frac{-1}{\left|\Omega_k\right|}\int_{\Omega_k}\ln\hat{p}\left(I\left(\mathbf{x}\right),\Omega_k\right)d\mathbf{x}$. In addition, we use $P_{\Omega_k}=\frac{\left|\Omega_k\right|}{\left|\Omega\right|}$ based on several previous publications [14]. In entropy estimation, $\hat{p}\left(I\left(\mathbf{x}\right),\Omega_k\right)$ is the approximate probability density function (pdf) in region *k* of the 3D image *I*. Many researchers estimated pdf's off-line. However, we observed dissimilar dynamic ranges of image intensities in different datasets and concluded that an off-line pdf is suboptimal; we use on-line estimation.

We estimate pdf's using the Parzen window method [15], one of the most powerful and robust methods in the literature, with the Gaussian kernel as $\hat{p}\left(I\left(\mathbf{x}\right),\Omega\right)=\frac{1}{\left|\Omega\right|}\int_{\Omega}K\left(I\left(\mathbf{x}\right)-I\left(\hat{\mathbf{x}}\right)\right)d\hat{\mathbf{x}}$. In this equation, *K* is the Gaussian kernel with a standard deviation (sigma) as its tuning parameter, which sets the resolution of the pdf estimation process. Choosing low values make pdf estimation sensitive to noise and high values remove useful details from the estimated pdf. In the literature, values between 1and 3 are used. Finally, we write the energy function as:

$$J\left(\Omega_1,\ldots,\Omega_{m+1}\right)=J\left(P\right)=\sum_{j=1}^{m+1}-\int_{\Omega_j}\ln\hat{p}\left(I\left(\mathbf{x}\right),\Omega_j\right)d\mathbf{x} \quad (4)$$

where *P* is vector of $m\times 7+q$ parameters (because each one of the local alignments have 7 parameters).

3.2 Energy Optimization

There are many ways to minimize the energy function. To minimize the energy function, we use Quasi-Newton algorithm with BFGS method for Hessian matrix estimation. Although, methods such as steepest descent are also popular in the literature [11], we obtained superior results using the Quasi-Newton algorithm. The gradients can be estimated using numerical methods but analytical computation is more robust and generates more accurate results. To compute the gradients analytically, we use Heaviside function $U\left(\Phi^k\right)=\begin{cases}1 & if\ \Phi^k\geq 0\\ 0 & if\ \Phi^k<0\end{cases}$. There are two types of parameters $\mathbf{w}$ and $\mathbf{p}^k$ and for the *i*th component of $\nabla_{\mathbf{w}}$ and $\nabla_{\mathbf{p}^k}$. We compute derivatives as follows:

$$\begin{aligned}\nabla w_i J=&\sum_{j=1}^{m}\left\{\iint_{\Gamma_j}\nabla w_i\Phi^j\left(\mathbf{x}\right)\left(\ln\hat{p}_j\left(\mathbf{x}\right)-\ln\hat{p}_{m+1}\left(\mathbf{x}\right)\right)d\mathbf{x}\right.\\&+\frac{1}{\left|\Omega_j\right|}\int_{\Omega_j}\left(\frac{1}{\hat{p}_j\left(\mathbf{x}\right)}\iint_{\Gamma_j}\nabla w_i\Phi^j\left(\hat{\mathbf{x}}\right)K\left(I\left(\mathbf{x}-\hat{\mathbf{x}}\right)\right)d\hat{\mathbf{x}}\right)d\mathbf{x}\\&\left.-\frac{1}{\left|\Omega_{m+1}\right|}\int_{\Omega_{m+1}}\left(\frac{1}{\hat{p}_{m+1}\left(\mathbf{x}\right)}\iint_{\Gamma_j}\nabla w_i\Phi^j\left(\hat{\mathbf{x}}\right)K\left(I\left(\mathbf{x}-\hat{\mathbf{x}}\right)\right)d\hat{\mathbf{x}}\right)d\mathbf{x}\right\}\end{aligned} \quad (5)$$

$$\begin{aligned}\nabla p_i^k J=&\iint_{\Gamma_j}\nabla p_i^k\Phi^k\left(\mathbf{x}\right)\left(\ln\hat{p}_k\left(\mathbf{x}\right)-\ln\hat{p}_{m+1}\left(\mathbf{x}\right)\right)d\mathbf{x}\\&+\frac{1}{\left|\Omega_k\right|}\int_{\Omega_k}\left(\frac{1}{\hat{p}_k\left(\mathbf{x}\right)}\iint_{\Gamma_k}\nabla p_i^k\Phi^k\left(\hat{\mathbf{x}}\right)K\left(I\left(\mathbf{x}-\hat{\mathbf{x}}\right)\right)d\hat{\mathbf{x}}\right)d\mathbf{x}\\&-\frac{1}{\left|\Omega_{m+1}\right|}\int_{\Omega_{m+1}}\left(\frac{1}{\hat{p}_{m+1}\left(\mathbf{x}\right)}\iint_{\Gamma_k}\nabla p_i^k\Phi^k\left(\hat{\mathbf{x}}\right)K\left(I\left(\mathbf{x}-\hat{\mathbf{x}}\right)\right)d\hat{\mathbf{x}}\right)d\mathbf{x}\end{aligned} \quad (6)$$

In the above equations, for simplicity, we have used $\hat{p}_k\left(\mathbf{x}\right)=\hat{p}\left(I\left(\mathbf{x}\right),\Omega_k\right)$. For $\nabla p_i^k\Phi^k\left(\mathbf{x}\right)$ and $\nabla w_i\Phi^k\left(\mathbf{x}\right)$ we have:

$$\nabla w_i\Phi^k\left(\mathbf{x}\right)=\Phi_i^k$$

$$\nabla p_i^k\Phi^k\left(\mathbf{x}\right)=\begin{bmatrix}\frac{\partial\Phi^k\left(\tilde{x}_k,\tilde{y}_k,\tilde{z}_k\right)}{\partial\tilde{x}_k}\\ \frac{\partial\Phi^k\left(\tilde{x}_k,\tilde{y}_k,\tilde{z}_k\right)}{\partial\tilde{y}_k}\\ \frac{\partial\Phi^k\left(\tilde{x}_k,\tilde{y}_k,\tilde{z}_k\right)}{\partial\tilde{z}_k}\\ 0\end{bmatrix}^T\times\frac{\partial M\left(\mathbf{p}^k\right)}{\partial p_i^k}\times\begin{bmatrix}x\\ y\\ z\\ 1\end{bmatrix} \quad (7)$$

where $\frac{\partial M\left(\mathbf{p}^k\right)}{\partial p_i^k}$ is computed using derivatives of each matrix element. Initial parameters are chosen for identity transformation of the mean shape.

4. Experimental Results

In this section, we show the results of applying the proposed method to real MRI data. This data is obtained from the Internet Brain Segmentation Repository (IBSR) and used for training and testing of the proposed method [16]. Datasets are T1-wighted volumetric images with different pixel sizes. There are 18 datasets for which expert physicians have segmented 43 structures. These datasets have non-cubical voxels. Each dataset consists of $256\times256\times128$ voxels. The smallest volume that covers each of datasets is

$256\times256\times192$ $(mm)^3$. Based on the structures of interest, a region of interest (ROI) is chosen to improve speed and robustness of the algorithm. This improves the execution speed and accuracy of the method since the pdf's of all regions inside and outside of each desired structure are needed. We estimate these pdf's using the Parzen window. We define the ROI as 1.05 times of the smallest cube that covers all of the desired structures in different training datasets. This additional 5% confidence interval has been sufficient for the test datasets (image segmentation) in our experiments. In this step, the initial parameters and shape models are set to start the segmentation process. After extraction of the ROI, we use binary images of the structure (k) in the training dataset (i) to construct distance maps (ψ_i^k). We set $\mathbf{W} = \mathbf{0}$ and $M\left(\mathbf{p}^k\right) = I$ as the initial parameters. Because of different scales of the parameters, a normalization step is needed. For example, we compute rotation in radian and translation in pixel. Therefore, we have to choose a smaller unit for rotation to have a uniform parameter space.

To evaluate the results, we use the Dice coefficient [17]. In addition, as an alternative evaluation measure, we use the segment Hausdorff distance [18]. For optimization and extraction of the principal shapes, we use MATLAB [19]. All programs are run on a 3.2 GHz (Dual Core) Windows XP workstation with 2 GB RAM.

Among different structures in the brain, we work with 12 (6 structures on the left and right sides of the brain): 1) caudate; 2) thalamus; 3) putamen; 4) pallidum; 5) hippocampus, and 6) amygdale. We use ten datasets (randomly chosen) for training and the remaining eight datasets to test the proposed segmentation algorithm. We segment each of the left and right structures individually using principal shapes extracted from the training datasets. We also segment the left and right structures or multiple structures by the proposed coupling method. Evaluation of the results using dice coefficient and segment Hausdorff distance for the amygdala and hippocampus for eight test datasets (different from the training datasets) are shown in Tables 1-2. Note that coupling of the left and the right structures generates superior segmentation results and are more accurate compared with the segmentation results obtained in the alone manner (i.e., without the coupling information) using either the Dice coefficient or the segment Hausdorff distance. In addition, note that the coupling of the amygdala with other structures generates superior results in all cases using both of the Dice coefficient and segment Hausdorff distance measures. For the hippocampus, adding other structures does not improve the segmentation quality compared to the segmentation results when using the left and the right hippocampuses in the coupled manner based on the Dice coefficient. On the other hand, coupling with pallidums, caudates, and thalamuses improves segmentation quality based on the segment Hausdorff distance criteria. In Fig. 1, sample segmentation results for the amygdala and hippocampus are shown, when segmentation process is done for the coupling of these 4 strcutres (left and right amygdala and hippocampus).

Table 1. Dice coefficients and Segment Hausdorff distances for the amygdala from 8 test datasets without and with the coupling information.

Coupled With	Dice Coefficient			Segment Hausdorff distance		
	L	R	LR	L	R	LR
None	0.55	0.59	0.57	5.91	6.17	6.04
Itself	0.57	0.59	0.58	5.83	6.17	6.00
2 Hippocampi	0.61	0.64	0.62	5.55	6.01	5.78
2 Pallidums	0.57	0.62	0.59	5.64	5.84	5.74
2 Putamens	0.61	0.66	0.64	5.54	5.92	5.73
2 Caudates	0.59	0.67	0.63	5.45	5.50	5.47
2 Thalamuses	0.59	0.66	0.63	5.63	5.57	5.60

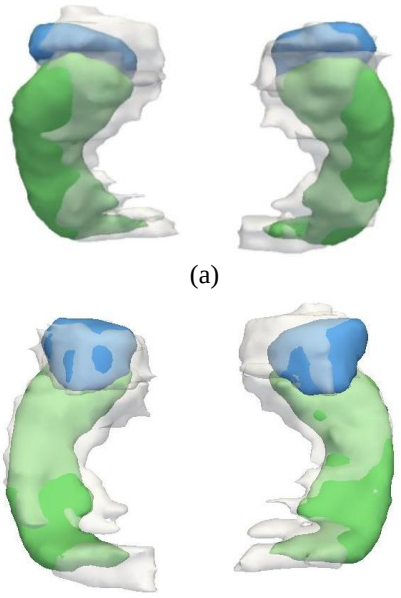
Table 2. Dice coefficients and segment Hausdorff distances for the hippocampus from 8 test datasets without and with the coupling information.

Coupled With	Dice Coefficient			Segment Hausdorff distance		
	L	R	LR	L	R	LR
None	0.68	0.63	0.66	6.3	7.15	6.73
Itself	0.71	0.68	0.70	6.48	6.73	6.61
2 Amygdalas	0.69	0.64	0.67	6.08	7.18	6.63
2 Pallidums	0.7	0.68	0.69	6.14	6.45	6.30
2 Putamens	0.69	0.64	0.67	6.62	7.51	7.07
2 Caudates	0.68	0.69	0.69	6.55	6.42	6.49
2 Thalamuses	0.7	0.67	0.69	6.16	5.17	5.67

5. Conclusions

We presented a new method for the segmentation of brain structures using their shapes relation extracted using PCA. We used separate shape class registration to extract each class variation with respect to the other shape classes. To do this, we removed variations due to shape poses by considering each shape representation locally. In addition, we added an independent transformation to each shape to allow more flexibility. This transform has seven parameters that cover rotation, scaling, and translation. Energy function

used for segmentation takes into account entropy of different shapes. With an automatic initialization of the structures and use of the quasi-Newton algorithm, a local minimum of the energy function is found. To achieve accurate results, probability density functions are calculated iteratively and gradients are computed analytically. Our method has low sensitivity to the parameters and is robust. Experimental results illustrate robustness and quality of the results generated by the proposed framework.



(a)

(b)

Fig. 1. Final left and right hippocampi (green) and amygdalas (blue) compared with the manual segmentation results (low opacity) extracted from the test datasets from anterior (a) and posterior (b) views.

References

[1] J.S. Suri, S. K. Setarehdan, and S. Singh, Advanced Algorithmic Approaches to Medical Image Segmentation, State of the Art Applications in Cardiology, Neurology, Mammography and Pathology, Springer-Verlag, February 2002.

[2] A. Macovski, "Noise in MRI," Magnetic Resonace in Medicine, Vol. 36, No. 3, pp. 494–497, 1996.

[3] M. Kass, A. Witkin, D. Terzopoulos, "Snakes: Active Contour Models," *First International Conference on Computer Vision*, pp. 259-268, 1987.

[4] E. D. Angelini, Y. Jin, and A. F. Laine, *State-of-the-Art of Levelset Methods in Segmentation and Registration of Medical Imaging Modalities*, The Handbook of Medical Image Analysis - Volume III: Registration Models, Kluwer Academic/Plenum Publishers, New York, 2005.

[5] A. Huang, G. M. Nielson, A Razdan, G. E. Farin, D. P. Baluch, and D. G. Capco, " Thin Structure Segmentation and Visualization in Three-Dimensional Biomedical Images: A Shape-Based Approach," IEEE Transaction On Visualization and Computer Graphics, Vol. 12, No. 1, pp. 93-102, January/February 2006.

[6] N. A. Paragios, "Geodesic Active Regions and Level Set Methods," *PhD Thesis, University of Nice*, January 2000.

[7] M. E. Leventon, W. E. L. Grimson, and O. Faugeras, "Statistical Shape Influence in Geodesic Active Contours," *IEEE Internatinal Conference on Computer Vision and Pattern Recognition*, Vol. 1, pp. 1316-1323, 2000.

[8] K. M. Pohl, J. Fisher, R. Kikinis, W. E. L. Grimson, and W. M. Wells, "Shape Based Segmentation of Anatomical Structures in Magnetic Resonance Images," *International Conference on Computer Vision*, Vol. 3765, pp. 489-498, 2005.

[9] J. Yang , L. H. Staib, and J. S. Duncan, "Neighbor-Constrained Segmentation with Level Set Based 3D Deformable Models," *IEEE Transactions on Medical Imaging*, Vol. 23, No. 8, pp. 940-948, August 2004.

[10] A. Litvin, and W.C. Karl, "Coupled Shape Distribution-Based Segmentation of Multiple Objects," *Boston University, Boston, USA, Tech. Rep. ECE-2005-01*, March 2005.

[11] A. Tsai, W. Wells, C. Tempany, E. Grimson, and A. Willsky, "Mutual Information in Coupled Multi-Shape Model for Medical Image Segmentation," *Medical Image Analysis*, Vol. 8, No. 4, pp. 429-445, December 2004.

[12] P. B. Bijari, A. R. Akhoundi-Asl, H. Soltanian-Zadeh, "Interactive Coupled Object Segmentation Using symmetry and Distance Constraint," *accepted and presented to third Cairo International Biomedical Engineering Conference*, Cairo, 2006.

[13] J. A. Nelder, and R. Mead, "A Simplex Method for Function Minimization". *The Computer Journal*, Vol. 7, pp. 306–313, 1965.

[14] J. Kim, J. Fisher, A. Yezzi, M. Cetin, and A. Willsky, "Nonparametric methods for image segmentation using information theory," *IEEE International Conference on Image Processing* Vol. 3, pp. 797–800, 2002.

[15] E. Parzen, "On the estimation of a probability density function and the mode," *Annals of Mathematical Statistics*, Vol.33, pp. 1065–1076, 1962.

[16] http://www.cma.mgh.harvard.edu/ibsr/

[17] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, Vol. 26, pp. 297–302, 1945.

[18] C. Guerra, and V. Pascucci, "3D segment matching using the Hausdorff distance," *Seventh International Conference on Image Processing And Its Applications*, Vol. 1, pp. 18-22, 1999.

[19] http://www.mathworks.com

[20] A. J. Worth, N. Makris, R. Patti, J. M. Goodman, E. A. Hoge, V. S. Caviness, and D.N. Kennedy," Precise Segmentation of the Lateral Ventricles and Caudate Nucleus in MR Brain Images Using Anatomically Driven Histograms", *IEEE Trans. on Medical Imaging*, Vol. 17, pp. 303-309, 1998.

[21] B. M. Dawant, S. L. Hartmann, J.-P. Thirion, F. Maes, D. Vandermeulen, and P. Demaerel," Automatic 3-D Segmentation of Internal Structures of the Head in MR Images Using a Combination of Similarity and Free-Form Transformations: Part I, Methodology and Validation on Normal Subjects," *IEEE Trans. on Medical Imaging*, Vol. 18, pp. 909-916, 1999.