

Multi-object Deformable Templates Dedicated to the Segmentation of Brain Deep Structures

F. Poupon^{1,2,3}, J.-F. Mangin¹, D. Hasboun⁴, C. Poupon¹,
I. Magnin² and V. Frouin¹

¹ Service Hospitalier Frédéric Joliot, DSV, DRM, CEA, 91406 Orsay Cedex, France
e-mail : poupon@shfj.cea.fr

² CREATIS-INSA Lyon, 69621 Villeurbanne, France

³ IFSBM, 94805 Villejuif Cedex, France

⁴ Service de Neuroradiologie H. Fischgold, CHU Pitié Salpêtrière, 75013 Paris, France

Abstract. We propose a new way of embedding shape distributions in a topological deformable template. These distributions rely on global shape descriptors corresponding to the 3D moment invariants. In opposition to usual Fourier-like descriptors, they can be updated during deformations at a relatively low cost. The moment-based distributions are included in a framework allowing the management of several simultaneously deforming objects. This framework is dedicated to the segmentation of brain deep nuclei in 3D MR images. The paper focuses on the learning of the shape distributions, on the initialization of the topological model and on the multi-resolution energy minimization process. Results are presented showing the segmentation of twelve brain deep structures.

1 Introduction

The wide diffusion of magnetic resonance imaging has given rise to numerous potential applications for 3D segmentation algorithms, ranging from morphometric or functional studies to neuro-surgical assistance. Unfortunately most methods proposed by the image processing community are still far from reaching the robustness required by clinical applications. With regard to the brain for instance, statistical analysis of gray levels is insufficient to segment deep gray matter nuclei mainly because of partial volume effect [1]. A solution to overcome this kind of difficulties consists in embedding *a priori* knowledge in the segmentation algorithm, which is generally done using deformable models.

1.1 Deformable models

First active contours known as snakes [2] were introduced for regularization purpose. Since they include no *a priori* knowledge on global shape, they still required a precise initialization. The next generations of deformable models include global shape information. Hence they can be viewed as deformable templates. Two different directions have been investigated. The first one consists in using an elastic atlas which is mapped onto the image to be segmented, this atlas being generated by an expert manual delineation [3]. The second approach consists in restricting the class of admissible shapes to a family described by a few parameters using for instance superquadrics with global deformations [4]. This approach can be

improved by allowing local deformations of the template using splines [5]. Other families of models have been proposed including a continuum of parameters from global to local shape information using modal analysis [6] or wavelet analysis [7].

The more recent efforts aim at including some knowledge on anatomical variability in the deformable template. This can be achieved by endowing the model parameters with some probability distribution. This distribution generally relies on the parameter expected values and variances which are inferred from a training set [8]. The segmentation goal is then defined using for instance a Maximum a Posteriori criterion which leads to an energy minimization. An alternative consists in using principal component analysis in order to restrict the space of admissible shapes according to the main modes of deformation of a mean model [9]. The 3D extension of this approach is rather difficult because of parametrization problems. A solution recently proposed in [10] allows the computation of a homogeneous parametrization of an arbitrarily shaped simply connected object. From this parametrization, 3D Fourier snakes including a restricted deformation space relying on eigenmodes are proposed. Because of the shape space restriction, a final minimization can be required to recover shape local details. A fairly complete bibliography on deformable models can be found in [11].

1.2 Paper overview

In this paper we propose a new way of embedding shape distributions in a framework allowing simultaneous deformation of several objects. The underlying model is simply a label image initially endowed with the *a priori* known topology which is preserved during the deformation process [12]. Performing simultaneous deformation of several templates increases the robustness of the segmentation. Indeed, each object is confined to a domain defined by surrounding objects. The segmentation is achieved through the coarse-to-fine minimization of a global energy including data-driven terms, regularization and shape distributions.

This paper focuses on the construction of the initial topological model from the training set and the way of embedding shape distributions in an energy-based multi-object segmentation framework.

The paper is organized as follows. First we present the 3D multi-object deformable topological template endowed with the *a priori* known information on shapes. Then we present what is inferred from the learning step on the training set. Next we present the way of constructing the initial topological model from the training set followed by the presentation of the particular energy minimization scheme. Finally we propose different results obtained by this 3D multi-object segmentation process.

2 The deformable model

The deformable model is based on deformable regions initially endowed with the *a priori* known topology. Then, segmentation amounts to applying topology preserving deformations in order to reach an energetic goal [12]. The deformations are performed by applying an homotopic transformation that is a sequence of addition or deletion of simple points in the objects. Those simple points are efficiently characterized by two numbers of connected components in the local neighborhood [13]. To respect Jordan's theorem an object and its background

have to be taken with two different connectivity. Then the numbers are computed with the 26-connectivity for the objects of interest and with the 6-connectivity for the background defined by the surrounding of the set of objects. Knowledges on shapes are also embedded to constrain the deformations to an *a priori* known domain in the shape space during the segmentation process. These knowledges are shape distributions relying on descriptors corresponding to the 3D moment invariants [16, 14]. These descriptors are invariant relatively to location, orientation and scale.

The energy defining the segmentation goal has the form

$$E_g = \sum_M V_{l(M)}^{data}(i(M)) + \sum_{M_1, M_2} V^{Potts}(l(M_1), l(M_2)) + \sum_l V_l^{shape}, \quad (1)$$

where M is a point, $l(M)$ is a point label, $i(M)$ is the intensity of M , and M_1, M_2 are 6-neighbours. The $V_{l(M)}^{data}$ functions embed the data attraction on the deformable model. They are φ -functions roughly modeling for each region statistical knowledge on intensities

$$V_{l(M)}^{data}(i(M)) = \left(\frac{i(M) - m_{l(M)}}{\sigma_{l(M)}} \right)^4 / \left(1 + \left(\frac{i(M) - m_{l(M)}}{\sigma_{l(M)}} \right) \right)^4. \quad (2)$$

where $m_{l(M)}$ and $\sigma_{l(M)}$ are respectively the gray level mean and standard deviation of the object defined by the label $l(M)$.

V^{Potts} is a matrix of adjacency costs related to the Potts model used in Gibbs fields [15]. A set of potentials V_l^{shape} is added to the energy function to model the fact that the shape of each object is relatively stable between individuals. V_l^{shape} embeds the shape distributions related to the label l .

$$V_l^{shape} = V_l^v + V_l^g + V_l^o + \sum_{12} V_l^i(\tilde{I}_\beta^\alpha), \quad (3)$$

where V_l^v is related to the object volume, V_l^g is related to its gravity center position, V_l^o is related to its inertial frame orientation and $V_l^i(\tilde{I}_\beta^\alpha)$ is related to the moment invariants. The derivation of the \tilde{I}_β^α is described in [16, 14]. Those potentials are logarithm of Gaussians relying on statistical parameters inferred from a training set. Statistics on gravity center and orientation stem from an initial registration in the Talairach proportional reference frame like in [10].

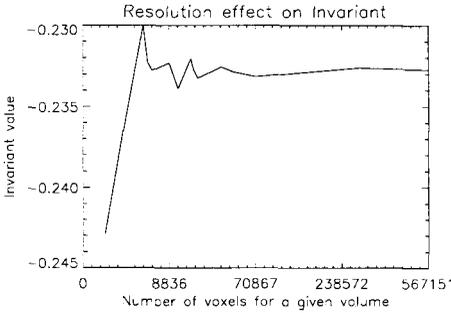
3 The learning step

The various *a priori* knowledge embedded in the deformable template is learnt from a training set. This training set was generated by manual delineation of the 12 desired structures on 6 different brains (Fig. 2). During this step statistics on volume, gravity center position, moment invariants and orientation are computed to form the shape distributions.

3.1 Moment invariants

The influence of resolution on the invariants has been studied from a training set composed of 8 different lateral ventricles by resampling each ventricles with

sixteen different resolutions. Fig. 1 shows the evolution curve of the invariant \tilde{I}_{22}^2 which is typical of other ones. It can be seen that the error on the values of the invariants becomes noticeable in low resolutions, which could be foreseeable since the object shape is perturbed. But by studying the inter-individual variability on this set of ventricles (Fig. 1) we have shown that the anatomical variability is greater than the variations induced by resolution. Therefore the coarse-to-fine implementation can be done with the same shape distribution at all resolution levels.



Invariants	R_LV		L_LV		P_LV
	m	σ	m	σ	
\tilde{I}_{00}^2	1.37	0.14	1.39	0.18	0.70
\tilde{I}_{11}^3	-0.79	0.13	-0.79	0.07	-0.35
\tilde{I}_{22}^2	0.97	0.10	0.99	0.14	0.41
\tilde{I}_{33}^3	-0.68	0.06	-0.68	0.06	-0.33
\tilde{I}_{222}^2	-1.07	0.11	-1.09	0.15	-0.44
\tilde{I}_{3111}^3	0.06	0.59	-0.15	0.57	0.23
\tilde{I}_{3131}^3	0.55	0.07	0.55	0.04	0.25
\tilde{I}_{3331}^3	-0.58	0.07	-0.58	0.05	-0.27
\tilde{I}_{3333}^3	0.66	0.06	0.66	0.07	0.32
$\tilde{I}_{112}^{2,3}$	-0.07	0.66	0.15	0.63	-0.25
$\tilde{I}_{312}^{2,3}$	-0.62	0.07	-0.63	0.05	-0.28
$\tilde{I}_{332}^{2,3}$	0.70	0.06	0.71	0.07	0.33

Fig. 1. Left: resolution curve of \tilde{I}_{22}^2 obtained for a lateral ventricle composed of 8836 voxels. Table: average (m) and standard deviation (σ) of the invariants evaluated for 8 right (R_LV) and left (L_LV) lateral ventricles and the invariants computed for a pathological one (P_LV).

It should be noted that most of the $\tilde{I}_{\beta}^{\alpha}$ show a relatively low variability between individuals. This result tends to prove that these $\tilde{I}_{\beta}^{\alpha}$ vary sufficiently slowly in the shape space to be efficiently used for constructing shape distributions. The results corresponding to a pathological ventricle show that it can be easily discriminated from the others due to its low values far from the distribution means. Nevertheless it means that in case of pathological structures some of the shape constraints have to be relaxed. This can be done either interactively or perhaps using a high level process driving the energy minimization.

3.2 Adjacencies

The adjacency cost matrix has also been constructed from the training set by looking for the different possible adjacencies in the manual segmentations. This can only be done if the manual delineations are performed in the same way to insure a reproductive topological arrangement of the brain structures. The twelve structure adjacency graph is given in Fig. 2.

3.3 Orientation

Since all the position and orientation parameters are computed in the Talairach reference frame we suppose that the orientations are relatively stable between individuals. Then each object average inertial frame is obtained by computing



Fig. 2. Left: different views of a manually segmented brain of the training set showing the 12 structures of interest. Right: the 12 structures adjacency graph. L. denotes Left and R. denotes Right. Abbreviations: Acc, Accumbens; Cd, Caudate nucleus; GP, Globus Pallidus; LV, Lateral Ventricle; Pu, Putamen; Th, Thalamus.

the average vector of each frame components \vec{m}_i , $i \in \{1, 2, 3\}$ from the training set. The standard deviations σ_i are computed with respect to angle of each frame components with its corresponding average vector. Then the orientation potential is obtained by a specific φ -function

$$\Phi(x(\vec{v})) = x(\vec{v})^4 / (1 + x(\vec{v})^4), \quad x(\vec{v}) = \cos^{-1}(|\vec{m}_i \cdot \vec{v}|) / \sigma_i. \quad (4)$$

Only the first orientation vector corresponding to the maximum eigenvalue of the inertial matrix is taken into account. It avoids the possible misplacement of the other vectors due to the relative symmetry of the objects. Another better approach could be designed from third order moments like in [16].

3.4 Data attraction

The next statistics to be determined is related to the data attraction term. It relies on the gray-level of the different structures in the volume. Since the gray-level distributions are different between acquisitions we propose to estimate the gray-level statistic parameters of each nucleus from the histogram of the brain we want to segment. First, the gray and white matter modes are automatically detected by a scale-space analysis of the histogram [17]. Then we suppose that each nucleus gray-level average value can be defined as a weighted function of the gray and white matter average values stable between acquisitions. For a brain nucleus n , we define its gray-level average value μ_n as

$$\mu_n = \alpha_n m_W + (1 - \alpha_n) m_G, \quad (5)$$

where m_W and m_G are respectively the white and gray matter gray-level average values and α_n the weighting coefficient for the nucleus n . The μ_n are inferred from the training set by computing the average gray-level value of the voxels belonging to a structure defined by a label. Then the α_n are estimated for each structure of each brain of the training set. These values are given in Tab. 1.

The α_n values are ranging from 0.0 for the gray matter to 1.0 for the white matter. Then the nuclei can be classified from the brightest to the darkest one as in Tab. 1.

Table 1. Average (m) and standard deviation (σ) of specific brain nuclei weighting coefficients α_n .

<i>Nuclei</i>	Globus Pallidus	Thalamus	Putamen	Caudate nucleus	Accumbens
$m(\alpha_n)$	0.911	0.618	0.458	0.287	0.244
$\sigma(\alpha_n)$	0.065	0.019	0.064	0.040	0.029

Anatomically, the nuclei we are interested in are made of neurons composing the gray matter. Each nucleus is connected to others or to the cortex by bright fibers composing the white matter. Some structures such as the thalamus are composed of a set of nuclei also connected to each other by white fibers. Then partial volume effects due to the MRI resolution results in a mixture of white and gray matter for some voxels where both neurons and fibers are present. That explains why some structures appear brighter than others. The definition proposed in eq. (5) was motivated by this assertion. Then α_n can be seen as an indicator that gives the proportion of gray and white matter gray-level that characterize a given nucleus. This α_n coefficients can not be computed for the Cerebro-Spinal Fluid (CSF) because its gray-level variations are too important between acquisitions. Anyway the CSF gray-level values are systematically lower than the gray matter ones. Therefore the ventricle attraction term is a simple linear potential function giving low values in the range $[0, m_G - 2\sigma_G[$, where σ_G is the standard deviation of the gray matter gray-level.

So the data attraction statistics of each nucleus of a new brain we want to segment except for the CSF is obtained by first extracting the gray and white matter gray-level average values from the brain histogram analysis [17] followed by the computation of their own average values according to eq. (5) using the α_n coefficients inferred from the training set in Tab. 1. Because nucleus standard deviations variability study is not yet achieved we chose them as the greatest values obtained during the learning step.

4 Energy minimization

The energy defined above stems from an extension of the usual Gibbs distribution form using additional global shape potentials. The global nature of these potentials and the topological constraints could lead to a failure of a simulated annealing algorithm. In fact, a good initialization can be provided to the minimization algorithm. Then, if the initial temperature is not too high, each object will be confined to a restricted area leading to local potentials. In our case where a complex energy is involved with many parameters, applying simulated annealing in normal resolution results in a prohibitive computation time. To overcome this problem, the whole deformation process is performed with a coarse-to-fine strategy. In this multi-resolution scheme, simulated annealing is only applied on the coarsest resolution level. Then, the resampled solution from coarser level is entered as initialization to the next finer level where the energy minimization is refined by an ICM like algorithm [12].

The whole homotopic deformation process is implemented using a multi-front strategy. At high temperature threads can appear around objects. They are difficult to delete during the following deformations because of the topological constraints. In order to prevent their apparition and to favour their deletion a

sophisticated front scanning is used. The idea is to make the number of time a point is treated depend on its neighborhood configuration. The goal is the decrease of the growth potential of threads while increasing their potential of being removed. This is achieved by attaching to each local configuration a probability of being treated. We chose this not normalized probability to be

$$P_N(n) = \begin{cases} 0.02 & \text{if } n \leq N, \\ 0.5 & \text{if } N < n < 26 - N, \\ 5.0 & \text{if } n \geq 26 - N. \end{cases} \quad (6)$$

where n is the number of identical labels in the 26-neighborhood and N is a parameter that specifies the two important transitions in the behaviour of the process (typically $N = 3$). The first condition forbids a point to be set to a specified label if there is not enough points of this label in its neighborhood so that the creation of threads is kept out. $P_N(n \leq N)$ is not set to 0.0 because if so, the relatively thin lines used for the initial topological model could not grow. $P_N(N \leq n \leq 26 - N)$ is set to 0.5 to randomize the treatment of points. Finally $P_N(n \geq 26 - N)$ is set to a high value to favour the deletion of threads. Practically this value is used to perform the Gibbs sampler probability test at most five times to have more chances to accept the transition of the label when we are in presence of threads.

5 The initial topological model

The initial topological model provides the first approximation of the organisation of the different structures to segment in the top-level pyramid volume. The major difficulty to generate this model is to assure the right topology with the good adjacencies.

Since all our objects are simply connected they can simply be defined as single points. This is the easier way to get the desired topology for the objects. The problem is the choice of such points because they must be relatively stable between individuals in the Talairach reference frame. Those points can not be the mean gravity center ones because since some of the structures are bended the gravity center can be located outside the object which could lead to energy minimization problems. So only points that are part of the objects should be used. We defined them as the one point reduction given by a homotopic skeletonization of the object. This skeleton is obtained from a front-based sequential deletion of simple points which assures a relatively stable localization of the final point between individuals. The statistics on the position of these points in the Talairach frame are computed during the learning step. Fig. 3.a shows the points obtained for the twelve structures of interest.

Next, we have to construct the adjacencies between these points. If the white matter is removed from the graph of Fig. 2, it simplifies to a simple chain of adjacent objects. So the skeleton points of the adjacent objects can simply be linked with lines to assure the right model adjacencies (Fig. 3.b).

Then we compute a Voronoï diagram associated to a simple chamfer distance conditionally to the connecting segments. This Voronoï diagram associates each line point to the closest seed (Fig. 3.c). Then this result is used as the initial

topological model of the arrangement of the different structures at the top level of the pyramid according to the multi-resolution scheme.

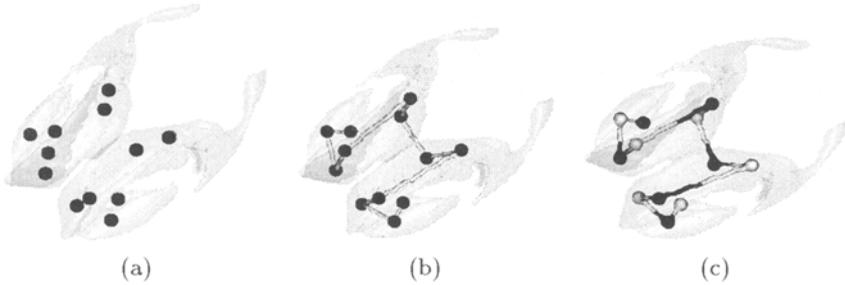


Fig. 3. Construction of the initial topological model. (a): the skeleton points in their surrounding structures. (b): the adjacent skeleton points linked by lines. (c): the initial topological model obtained after a Voronoi diagram.

6 Results

In the following the results are not presented using a triangulation-based rendering technique as in Fig. 2 but by representing each voxel by a cube to give a better insight of the topological nature of the process.

6.1 Without data attraction

The consistency of the shape information given by the moment invariants is studied at the top level of the multi-resolution segmentation scheme by minimizing the energy defined in eq. (1) without the data attraction term. Fig. 4 shows the results obtained for 4 of the structures with respect to their corresponding anatomical representations. They present a relatively good similarity with their anatomical corresponding structures. It should be noted that the purpose of the shape distributions is only to roughly characterize the shapes. Finer details have to be inferred from the data.

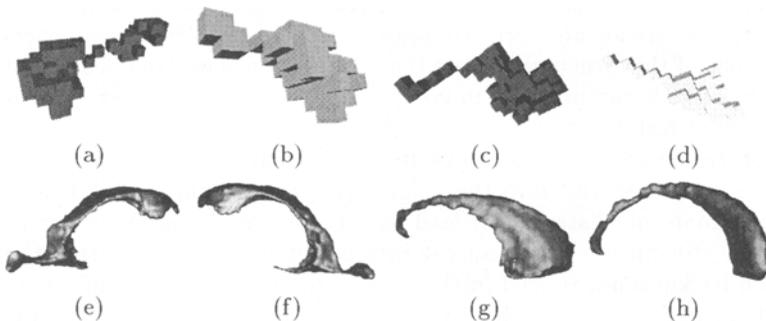


Fig. 4. Results obtained at the top level of the multi-resolution scheme only with shape information (top), with respect to their corresponding nuclei (bottom). (a), (e): right lateral ventricle. (b), (f): left lateral ventricle. (c), (g): right caudate nucleus. (d), (h): left caudate nucleus.

6.2 With data attraction

The whole multi-resolution scheme has been used to segment the twelve brain deep structures listed and represented in Fig. 2 in a given brain MR image.

The results obtained from the top level initial topological model to the normal resolution level of the pyramid are shown in Fig. 5.

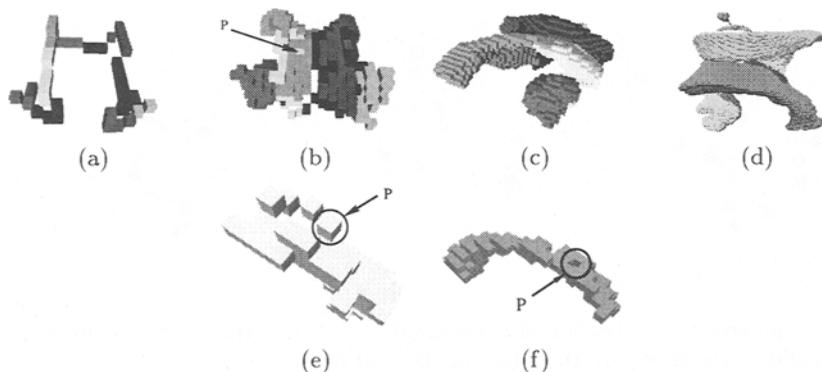


Fig. 5. Results obtained at the different level of the multi-resolution scheme. (a): initial topological model. (b): top level result. (c): intermediate level result. (d): normal resolution result. The arrow in (b) points out a particular configuration : in (e) P is 18-adjacent and the caudate nucleus is simply 26-connected and in (f) the ventricle is simply 26-connected. (e) and (f) have the expected topology, but P seems to create a hole in the ventricle (b).

The caudate nucleus point P in Fig. 5 can be seen as a hole for the ventricle but if we look at those structures separately (Fig. 5) both have the expected simple 26-connectivity. In fact, since each object topology is considered in 26-connectivity relatively to 6-connectivity for its complement, surprising configurations can arise when dealing with several objects. They are related to the usual paradoxes linked to Jordan theorem. Up to know, these strange configurations do not really perturb the segmentation process but further studies have to be done on the subject.

It appears that the structures well contrasted relatively to their surroundings such as the ventricles are correctly segmented. No problem is also encountered for the parts of the structures where the information is well defined in the model, i.e. where other structures are in competition with them. Their behavior leads to the desired result.

In return, areas where no competition occur can lead to errors. For instance the putamen can spread into the neighboring cortex; this result from the fact that both anatomical structures lead to similar gray levels. A solution to overcome these difficulties would consist in a better design of the attraction terms acting on background. Indeed in the current model the background is supposed to be made up from white matter which is far to be sufficient. We plan to devise a better attraction term for background by splitting it in three classes of tissues corresponding to CSF, gray and white matter. Each subclass would be regularized independently by the Potts model but the topological constraints would act on the union of the three sub-labels.

7 Conclusion

We have proposed a new method to embed global shape information in a framework allowing deformation of several objects. This information relies on shape

distributions constructed from 3D moment invariants. Indeed these invariants have been shown to vary sufficiently slowly in the shape space to be relatively stable between individuals. The whole segmentation process has been tested for 12 brain deep structures. It shows a good behavior for the well defined parts of the model, i.e. the regions where competition occur. In return, a better management of the background class seems to be required to achieve acceptable results. The segmentation of these brain structures with a slow simulated annealing temperature decreasing at the coarsest resolution level (.995 geometric reason) takes 20 minutes on a SPARC-ULTRA 1 station.

References

1. T. Géraud, J.-F. Mangin, I. Bloch, and H. Maître. Segmenting internal structures in 3D MR images of the brain by Markovian relaxation on a watershed based adjacency graph. In *IEEE ICIP*, 1995, pp. 548–551.
2. M. Kass, A. Witkin, and D. Terzopoulos. Snakes : active contour models. *International Journal of Computer Vision*, vol. 1, 1988, pp. 321–331.
3. R. Bajcsy and R. Kovacic. Multiresolution elastic matching. *Computer Vision, Graphics and Image Processing*, vol. 46, 1989, pp. 1–21.
4. F. Solina and R. Bajcsy. Recovery of parametric models from range images: the case for superquadrics with global deformations. *PAMI*, vol. 12, 1990, pp. 131–147.
5. D. Terzopoulos and D. Metaxas. Dynamic 3D models with local and global deformations: deformable superquadrics. *IEEE PAMI*, vol. 13, 1991, pp. 703–714.
6. J. Martin, A. Pentland, and R. Kikinis. Shape analysis of brain structures using physical and experimental modes. In *IEEE CVPR, Seattle*, 1994, pp. 752–755.
7. B. C. Vemuri, A. Radisavljevic, and C. M. Leonard. Multi-resolution stochastic 3D shape models for image segmentation. In *13th IPMI, Flagstaff*, 1993, pp. 62–76.
8. L. H. Staib and J. S. Duncan. Model-based deformable surface finding for medical images. *IEEE Transactions on Medical Imaging*, vol. 15, no. 5, 1996, pp. 720–731.
9. T. F. Cootes, A. Hill, C. J. Taylor, and J. Haslam. Use of active shape models for locating structures in medical images. *Image and Vision Computing*, vol. 12, no. 6, 1994, pp. 355–365.
10. G. Székely, A. Kelemen, C. Brechbühler, and G. Gerig. Segmentation of 2D and 3D objects from MRI volume data using constrained elastic deformations of flexible Fourier contour and surface models. *MIA*, vol. 1, no. 1, 1996, pp. 19–34.
11. T. McInerney and D. Terzopoulos. Deformable models in medical image analysis: a survey. *Medical Image Analysis*, vol. 1, no. 2, 1996, pp. 91–108.
12. J.-F. Mangin, F. Tupin, V. Frouin, I. Bloch, R. Rougetet, J. Régis, and J. López-Krahe. Deformable topological models for segmentation of 3D medical images. In *14th IPMI, Brest, France*. Kluwer Academic Publishers, 1995, pp. 153–164.
13. G. Bertrand and G. Malandain. A new characterisation of three-dimensional simple points. *Pattern Recognition Letters*, vol. 15, 1994, pp. 169–175.
14. F. Poupon, J.-F. Mangin, V. Frouin, and I. Magnin. 3D multi-object deformable templates based on moment invariants. In *10th SCIA*, vol. I, 1997, pp. 149–155.
15. S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the bayesian restoration of images. *IEEE PAMI*, vol. 6, no. 6, 1984, pp. 721–741.
16. C.-H. Lo and H.-S. Don. 3D moment forms: their construction and application to object identification and positioning. *IEEE PAMI*, vol. 11, 1989, pp. 1053–1064.
17. J.-F. Mangin, O. Coulon, and V. Frouin. Robust brain segmentation using histogram scale-space analysis and mathematical morphology. In *1st MICCAI*, 1998.