Implicit Active Shape Models for 3D Segmentation in MR Imaging

Mikaël Rousson[1], Nikos Paragios[2], and Rachid Deriche[1]

[1] I.N.R.I.A. Sophia Antipolis, France
{Mikael.Rousson,Rachid.Deriche}@sophia.inria.fr
[2] Ecole Nationale des Ponts et Chaussées, Paris, France
Nikos.Paragios@cermics.enpc.fr

Abstract. Extraction of structures of interest in medical images is often an arduous task because of noisy or incomplete data. However, hand-segmented data are often available and most of the structures to be extracted have a similar shape from one subject to an other. Then, the possibility of modeling a family of shapes and restricting the new structure to be extracted within this class is of particular interest. This approach is commonly implemented using active shape models [2] and the definition of the image term is the most challenging component of such an approach. In parallel, level set methods [8] define a powerful optimization framework, that can be used to recover objects of interest by the propagation of curves or surfaces. They can support complex topologies, considered in higher dimensions, are implicit, intrinsic and parameter free. In this paper we re-visit active shape models and introduce a level set variant of them. Such an approach can account for prior shape knowledge quite efficiently as well as use data/image terms of various form and complexity. Promising results on the extraction of brain ventricles in MR images demonstrate the potential of our approach.

1 Introduction

Object extraction is one of the first steps in medical imaging. Further analysis will highly depend on the quality of the segmented structures. However, medical images often suffer from noise, occlusions and incomplete data. Therefore, regularization constraints and prior knowledge are usually of good use. In this paper, we address this application with objective to recover a structure of particular geometric form.

B-splines deformable models as well as point distribution models are mathematical formulations introduced to the snake framework [4] to account for shape consistency. Active shape models [2] were a major breakthrough in object extraction and image segmentation. Such a framework consists of two stages; (i) the modeling and (ii) the segmentation phase.

During modeling the objective is to recover a compact representation for the geometric form of the structure of interest. Using a set of registered training examples, one can either represent prior knowledge using simple or more

complicated density functions. Gaussian distribution [2], mixture models [1] or non-parametric function [3] were considered in the past.

The segmentation/object extraction stage aims at recovering a geometric structure in the image plane that accounts for the desired image characteristics while being in the family of shapes generated by the model . To this end, a mechanism for recovering the most probable object location in the image was considered. Then, one can iterate and move closer to the target by updating the position of the model such that it gets closer to the desired image characteristics.

Level set representations [8] is an established technique for tracking moving interfaces in imaging, vision and graphics [7]. One can see numerous advantages for considering a level set variant of the active shape model. Such a formulation could account for various forms (boundary or regional) of data/image terms of various nature (edges, intensity properties, texture, motion, etc.), an important limitation of the active shape model. Furthermore, one can maintain the implicit and intrinsic property of the level set method as well as the ability to account for topological changes while being able to introduce prior shape knowledge, a task partially addressed up to now [6,12,10].

In this paper we propose a level set variant of active shape models that consists of various terms. Quite critical is the term that refers to the prior knowledge with objective to constrain the evolving surface to belong to a compact family of shapes - the one recovered through the training set. Such a term couples two unknown variables; (i) the evolving contour, (ii) the optimal projection parameters of this contour to the model space and imposes the active shape model behavior on the process. Furthermore, various image-driven terms - a major advantage/characteristics of the method - could be considered to guide the evolving contour towards the desired image characteristics.

The most closely related work with our approach, the active shape model can be found in [2]. In [6,12,10] substantial efforts to integrate prior knowledge within level set representations were considered. Worth mentioning is [6, 12] where modeling of prior knowledge is done in a consistent active shape model manner. Contrary to [6], where two optimization processes alternate, we propose a variational integration of data and prior terms. Moreover, the evolving surface is not restricted within the modeled space like in [12], but only attracted to this space, allowing more flexibility.

The reminder of the paper is organized as follows: in Section 2 we briefly introduce the level set representations, while in Section 3 we address the construction of the prior model in the space of level set functions. The main contribution of the paper, the level set variant of the active shape model is presented in Section 4, while in Section 5 we demonstrate the efficiency and the flexibility of our approach through the integration of a region-based data term for 3D segmentation in MR images.

2   Level Set Representations

Level set representations [8] are a useful mathematical formulation for implementing efficiently curve/surface propagation. One can also consider the level

set space as an optimization framework Let $\phi : \Omega \times \mathcal{R}^+ \rightarrow \mathcal{R}^+$ be a Lipschitz function with the following properties,

$$\phi((x,y);t) = \begin{cases} 0 & , (x,y) \in \mathcal{C}(t) \\ +\mathcal{D}((x,y), \mathcal{C}(t)) > 0 & , (x,y) \in \mathcal{C}_{in}(t) \\ -\mathcal{D}((x,y), \mathcal{C}(t)) < 0 & , (x,y) \in \mathcal{C}_{out}(t) = [\Omega - \mathcal{C}_{in}(t)] \end{cases}$$

where $(x,y) = p$, $\mathcal{C}_{in}(t)$ is the area enclosed by the curve $\mathcal{C}$, $\mathcal{D}((x,y), \mathcal{C}(t))$ the minimum Euclidean distance between the pixel (x,y) and $\mathcal{C}(t)$ at time t.

Let us also introduce the approximations of Dirac and Heaviside distributions as defined in [10]. Then one can define terms along $\mathcal{C}$ as well as interior and exterior to the curve using the Dirac and Heaviside functions:

$$(x,y) \in \Omega : \{\lim_{\alpha \rightarrow 0+} [\delta_\alpha(\phi(x,y)] = 1\} = \mathcal{C}$$
$$(x,y) \in \Omega : \{\lim_{\alpha \rightarrow 0+} [H_\alpha(\phi(x,y))] = 1\} = \mathcal{C}_{in}$$

Such terms will be used later to introduce the active shape prior term as well as data/image-driven terms that guides the contour ($\mathcal{C}$) towards the object of interest. The extension to higher dimensions is straightforward and in the following parts, we use this representation for an an evolving surface ($\mathcal{S}$) in $\mathcal{R}^3$.

3   Modeling Prior Knowledge in the Level Set Space

Learning the distribution of geometric/image structures is a common problem in computer vision with applications to segmentation, tracking, recognition, etc. It is clear that the selection of the representation is important. Given the selected optimization framework, level set functions is a natural selection to account for prior knowledge with numerous earlier described advantages. Let us consider a training set $\mathcal{C}_i$ of N registered curves or surfaces. Then, a distance transform can be used to represent $\mathcal{C}_i$ as a level set function ϕ_i.

The next step is the construction of the shape model, using the aligned contours. In order to create an invariant representation, one should first normalize the training set ϕ_i. Subtraction of the mean (that can be recovered by averaging ϕ_i's) is a common selection to this end. However, a simple averaging over the training will not give a istance function. To overcome this limitation, we consider a more rigorous approach [10], seeking to estimate the distance function (ϕ_M) that minimizes:

$$E(\phi_\mathcal{M}) = \sum_{i=1}^{n} \int_\Omega (\phi_i - \phi_\mathcal{M})^2 d\Omega, \quad \text{SUBJECT TO} : |\nabla\phi_\mathcal{M}|^2 = 1$$

One can optimize such a term though a gradient descent method:

$$\frac{d}{dt}\phi_\mathcal{M} = \sum_{i=1}^{n} (\phi_i - \phi_\mathcal{M})$$

while $\phi_\mathcal{M}$ is projected to the space of distance functions following [11]. The two steps alternate until the system reaches a steady-state solution. Then, we consider the modeling approach introduced in [6,12]. Once the samples ϕ_i centered
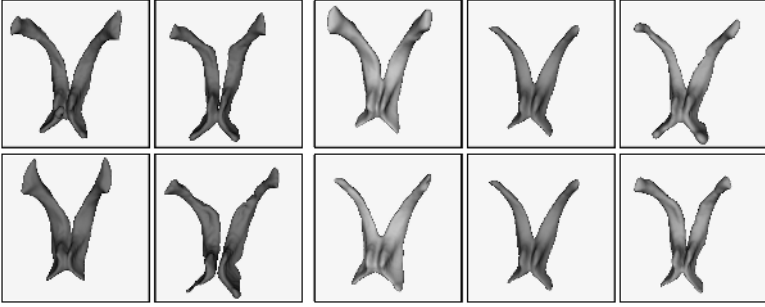
Fig. 1. LEFT: Some surfaces of the Training Set (segmented lateral brain ventricles of several patients), RIGHT: Model with the most important shape of variations [principal two modes after rigid alignment (blue:mean, red: $+\sigma$, green:$-\sigma$)].

with respect to $\phi_{\mathcal{M}}$, $[\psi_i = \phi_i - \phi_{\mathcal{M}}]$, the most important modes of variations can be recovered through Principal Component Analysis:

$$\phi = \phi_{\mathcal{M}} + \sum_{j=1}^{m} \lambda_j \, U_j$$

where m is the number of retained modes of variation, U_j are these modes (eigenvectors), and λ_j are linear weight factors within the allowable range defined by the eigenvalues.

An example of such an analysis is shown in [fig. (1)] for the 3D modeling of lateral brain ventricles. The model was built using 8 surfaces from different subjects. This example includes a difficult issue for classical parametric approaches because of different surface topologies within the training set. For example, the fourth surface in [fig. (1)] shows a separation between left and right ventricles. Our approach can deal naturally with this type of data. The obtained model gives a compact representation of the shape family: the first two modes of variation represent the major part of the class (80%), while the third one (9%) accounts for non-symetric properties of the ventricles that can be observed in some of the training samples. Moreover, the implicit representation of the surfaces make the modeling phase entirely automatic.

4   Introducing Prior Knowledge in the Level Set Space

Let us now consider an interface represented by a level-set function $\phi(\mathbf{x})$ as described in Section 2 (where $\mathbf{x}$ is in $\mathcal{R}^2$ or $\mathcal{R}^3$). We would like to evolve it while respecting some shape properties $\phi_{\mathcal{P}}(\mathbf{x})$ modulo a transformation $\mathcal{A}$ belonging to a predefined family. Assuming a rigid transformation $\mathcal{A}(\mathbf{x}) = \mathbf{R}\mathbf{x} + \mathbf{T}$, the evolving interface and the transformation should satisfy the conditions:

$$\begin{cases} \mathbf{x} \to \mathcal{A}(\mathbf{x}) \\ \phi(\mathbf{x}) \approx \phi_{\mathcal{P}}(\mathcal{A}(\mathbf{x})), \quad \forall \mathbf{x} \in \Omega \end{cases}$$
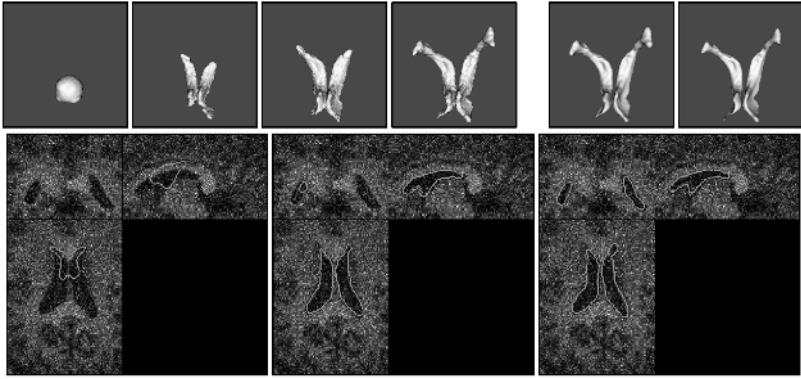
Fig. 2. Segmentation of lateral brain ventricles with Shape Prior ($b = 0.3$) of a noisy MR image. TOP LEFT: surface evolution, TOP RIGHT: projected surface in the learning space and ground-truth surface (from the training set), BOTTOM: surface cut (green) and its projection in the learning space (red) during surface evolution.

In that case, the optimal transformation $\mathcal{A}$ should minimize:

$$E(\phi, \mathcal{A}) = \int_{\Omega} \rho(\phi, \phi_{\mathcal{P}}(\mathcal{A})) d\Omega$$

where ρ is a dissimilarity measure. For the sake of simplicity, we will use the sum of squared differences. Scale variation can be added to the rigid transformation $\mathcal{A}$, leading to a similarity one $\mathcal{A}(\mathbf{x}) = \mathcal{S}\,\mathbf{R}\mathbf{x} + \mathbf{T}$ (for 3D images, we obtain 7 parameters: $\mathcal{S}$, $\mathbf{R}(\theta_1, \theta_2, \theta_3)$, $\mathcal{T} = (\mathcal{T}_x, \mathcal{T}_y, \mathcal{T}_z)^T$). In that case, the objective function should be slightly modified (refer to [10] for further details). Furthermore, one can assume that estimating and imposing the prior within the vicinity of the zero-crossing of the level set representation is more meaningful. Within distance transforms, shape information is better captured when close to the origin of the transformation. The prior can be thus rewritten:
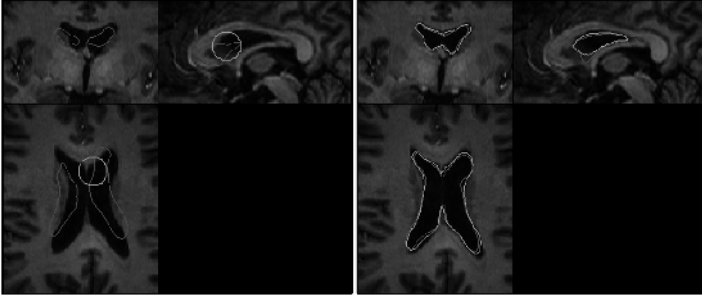
$$E(\phi, \mathcal{A}) = \int_{\Omega} \delta_\epsilon(\phi) \left(\mathcal{S}\phi - \phi_{\mathcal{P}}(\mathcal{A})\right)^2 d\Omega \quad \text{where } \epsilon \gg \alpha$$

During the model construction, we have analyzed the principal modes of variation within the training set. Including this information, the ideal transformation will map each value of current representation to the "best" level set representation belonging to the class of the training shapes. If a shape representation $\phi_{\mathcal{P}}$ belongs to this class, then it can be derived from the principal modes:

$$\phi_{\mathcal{P}} = \phi_{\mathcal{M}} + \sum_{j=1}^{m} \lambda_j\, U_j$$

Hence, we define a new objective function by introducing the modes weights $\lambda = (\lambda_1, \ldots, \lambda_m)$ as additional free parameters:

$$E(\phi, \mathcal{A}, \lambda) = \int_{\Omega} \delta_\epsilon(\phi) \left(\mathcal{S}\phi - \left(\phi_{\mathcal{M}}(\mathcal{A}) + \sum_{j=1}^{m} \lambda_j\, U_j(\mathcal{A})\right)\right)^2 d\Omega$$

Fig. 3. Segmentation of lateral brain ventricles with Shape Prior ($b = 0.3$) in an MR image which was not used during the learning phase: surface cut (green) and its projection in the learning space (red) at initialization (LEFT) and after convergence (RIGHT).

Fig. 4. Segmentation of lateral brain ventricles varying the influence of the shape prior term. From left to right, the shape prior weight b is 0, 0.3, 0.4, 0.5.

In order to minimize the above functional with respect to the evolving level set representation, the global linear transformation and the modes weights λ_j, we use the calculus of variations. The equation of evolution for ϕ is given by the calculus of its variations:

$$\frac{d}{dt}\phi = -2\delta_\epsilon(\phi)\mathcal{S}\big(\mathcal{S}\phi - \phi_\mathcal{M}(\mathcal{A})\big) - \frac{d}{d\phi}\delta_\epsilon(\phi)\big(\mathcal{S}\phi - \phi_\mathcal{M}(\mathcal{A})\big)^2$$

The differentiation with respect to the modes weights gives us a close form of the optimal parameters by solving the linear system $\bar{U}\lambda = b$ with:

$$\begin{cases} \bar{U}(i,j) = \int_\Omega \delta_\epsilon(\phi)U_i(\mathcal{A})U_j(\mathcal{A}) \\ b(i) = \int_\Omega \delta_\epsilon(\phi)(\mathcal{S}\phi - \phi_\mathcal{M}(\mathcal{A}))U_i(\mathcal{A}) \end{cases}$$

where $\bar{U}$ is a $m \times m$ positive definite matrix. Finally, the minimization of the energy with respect to the pose parameters is done by considering the gradient of each parameter:

$$\begin{cases} \frac{d}{dt}\mathcal{S} = 2\int_\Omega \delta_\epsilon(\phi)(\mathcal{S}\phi - \phi_\mathcal{P}(\mathcal{A}))(-\phi + \nabla\phi_\mathcal{P}(\mathcal{A}) \cdot \frac{\partial}{\partial\mathcal{S}}\mathcal{A})d\Omega \\ \frac{d}{dt}a_i = 2\int_\Omega \delta_\epsilon(\phi)(\mathcal{S}\phi - \phi_\mathcal{P}(\mathcal{A}))(\nabla\phi_\mathcal{P}(\mathcal{A}) \cdot \frac{\partial}{\partial a_i}\mathcal{A})d\Omega \quad \text{with } a_i \in \{\theta_1, \theta_2, \theta_3, \mathcal{T}_x, \mathcal{T}_y, \mathcal{T}_z\} \end{cases}$$

5   Active Shapes, Level Sets, and Object Extraction

In this section, we integrate the proposed level set variant of the active shape model to the Geodesic Active Region model [9], that on top of salient features uses global region statistics.

5.1   Geodesic Active Region

Introducing global region properties is a common technique to improve segmentation performance. To this end, one can assume a two-class partition problem where the object and the background follow different intensity distribution. Let $p_{\mathcal{C}_{in}}$ and $p_{\Omega - \mathcal{C}_{in}}$ be the densities of $I(\mathbf{x})$ in $\mathcal{C}_{in}$ and $\Omega - \mathcal{C}_{in}$. Then according to the Geodesic Active Region model [9] one can recover the object through the optimization of the following function:

$$E(\phi, p_{\mathcal{C}_{in}}, p_{\Omega - \mathcal{C}_{in}}) = (1 - a) \int_{\Omega} \delta_{\alpha}(\phi) g(|\nabla I|) |\nabla \phi| d\Omega$$
$$-a \int_{\Omega} \left[H_{\alpha}(\phi) \log(p_{\mathcal{C}_{in}}(I)) + (1 - H_{\alpha}(\phi)) \log(p_{\Omega - \mathcal{C}_{in}}(I)) \right] d\Omega$$

One can consider either parametric approximation [9] or a non-parametric density [5] functions to describe intensity properties. In both cases the new term will result in a local balloon force that moves the contour in the direction that maximizes the posterior segmentation probability as shown in [9].

5.2   Object Extraction

The Geodesic Active Region module is used jointly with the shape prior constraint. This data-specific information make the contour evolve toward the object of interest while keeping a global shape consistant with the prior shape family. For this purpose a variational formulation incorporating two terms is used:

$$E(\phi, \mathcal{A}, \lambda) = b \, E_{shape}(\phi, \mathcal{A}, \lambda) + (1 - b) \, E_{data}(\phi)$$

where E_{shape} is the shape prior and E_{data} is the Geodesic Active Region module.

This framework has been tested on the extraction of the lateral brain ventricles. [Fig. (2)] show the robustness to noise brought by the prior shape knowledge (the image is one of the training images but with additional Riccian noise). In [fig. (3)], we show the ability of our approach to extract objects from new images (not used for building the model). The active shape model is able to approximate the surface with a similar one from the modeled class while the object extraction allows small local variations with respect to the model. Finally, in [fig. (4)], we show the influence of the shape prior term by changing its weight. While prior knowledge improves the quality of the object extraction, overweighting shape prior will make object details to be missed. The possibility of tuning this parameter is an important advantage of our approach compared to [12].

6   Conclusion

We have proposed a level set variant of active shape models to deal with object extraction in medical MR images. Our approach exhibits numerous advantages. It can deal with noisy, incomplete and occluded data because of its active shape nature. It is intrinsic, implicit parameter and topology free, a natural property of the level set space. Examples on the brain ventricles extraction demonstrate the potential of our method. The nature of the sub-space of plausible solutions is a limitation of the proposed framework. Quite often the projection to this space does not correspond to a level set distance function. To account for this limitation, we currently explore prior modeling directly on the Euclidean space, and then conversion to the implicit space during the object extraction.

References

1. T. Cootes and C. Taylor. Mixture model for representing shape variation. *Image and Vision Computing*, 17:567–574, 1999.
2. T. Cootes, C. Taylor, D. Cooper, and J. Graham. Active shape models-their training and application. *Computer Vision and Image Understanding*, 61(1):38–59, 1995.
3. D. Cremers, T. Kohlberger, and C. Schn *o*rr. Nonlinear Shape Statistics in Mumford-Shah Based Segmentation. In *ECCV*, volume 2, pages 93–108, 2002.
4. M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. In *First International Conference on Computer Vision*, pages 259–268, London, June 1987.
5. J. Kim, J. Fisher, A. Yezzi, M. Cetin, and A. Willsky. Nonparametric methods for image segmentation using information theory and curve evolution. In *IEEE Internation Conference on Image Processing*, pages 797–800, Sept. 2002.
6. M. Leventon, E. Grimson, and O. Faugeras. Statistical Shape Influence in Geodesic Active Controus. In *IEEE CVPR*, pages I:316–322, 2000.
7. S. Osher and N. Paragios, editors. *Geometric Level Set Methods in Imaging, Vision and Graphics*. Springer Verlag, 2003.
8. S. Osher and J. Sethian. Fronts propagating with curvature dependent speed: algorithms based on the Hamilton–Jacobi formulation. *Journal of Computational Physics*, 79:12–49, 1988.
9. N. Paragios and R. Deriche. Geodesic active regions and level set methods for supervised texture segmentation. *The International Journal of Computer Vision*, 46(3):223, 2002.
10. M. Rousson and N. Paragios. Shape priors for level set representations. In A. Heyden, G. Sparr, M. Nielsen, and P. Johansen, editors, *Proceedings of the 7th European Conference on Computer Vision*, volume 2, pages 78–92, Copenhagen, Denmark, May 2002. Springer–Verlag.
11. M. Sussman, P. Smereka, and S. Osher. A Level Set Method for Computing Solutions to Incomprenissible Two-Phase Flow. *Journal of Computational Physics*, 114:146–159, 1994.
12. A. Tsai, A. Yezzi, et al. Model-based curve evolution technique for image segmentation. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 463–468, Dec. 2001.