Probabilistic Sequential Segmentation and Simultaneous On-Line Shape Learning of Multi-Dimensional Medical Imaging Data

J. Chiverton[1], X. Xie[2], and M. Mirmehdi[3]

[1] School of Information Technology, Mae Fah Luang University, 57100, Thailand
jpchiverton@gmail.com
[2] University of Wales Swansea, Department of Computer Science, Swansea, SA2 8PP, UK
x.xie@swansea.ac.uk
[3] Department of Computer Science, University of Bristol, Bristol, BS8 1UB, UK
majid@cs.bris.ac.uk

Abstract. Accurate automatic segmentation of anatomical structures is usually considered a difficult problem to solve because of anatomical variability and varying imaging conditions. A prior description of the shape of the anatomical structure to be segmented can reduce the ambiguity associated with the segmentation task. However this prior information has to be prepared specifically for the structure of interest, usually supervised and under favorable imaging conditions. An alternative is to consider the shape of the object sequentially, along a particular dimension of the data. This is the approach taken here, i.e. on-line modeling of sequential shape information which is combined with sequential segmentation of the intensity distributions for the segmented structure and the surrounding region.

1 Introduction

A clinical Magnetic Resonance (MR) scan of a tumor in a human brain contains pixels that represent pathological and non-pathological tissues. The structure of these tissues can be considered on a slice by slice basis, i.e. in a sequential segmentation process. Potential benefits of sequential segmentation include adaptation to changes in the intensity distributions due to e.g. inhomogeneities in the magnetic field and adaptation to other unknown variations e.g. structural. Medical imaging data is often multi-dimensional (including 3 spatial dimensions and a time dimension). The medical analysis of such data is complex, requiring advanced medical and information technology knowledge and skills, see e.g. [1, 2]. Furthermore sequential segmentation approaches require considerably less memory in comparison to segmentation techniques that process large multi-dimensional data sets non-sequentially.

Sequential segmentation and summarization processes may ultimately be medically useful. A sequential segmentation process can be used to sequentially estimate pathological structures such as tumors or even non-pathological structures such as cardiac tissues that may or may not be defined over time. Shape is inherent to the sequential segmentation process, whether implicitly considered (i.e. for initialization purposes on subsequent image slices) or explicitly. We consider an explicit model of shape where

the shape information is not only estimated from the segmentation process but also used in subsequent estimation steps to assist with subsequent image slice segmentations. We consider shape to be an important factor, providing additional information for an intelligent segmentation framework.

An ideal medium for shape modeling are active contour models which have been extensively investigated in conjunction with prior shape knowledge since at least [3, 4]. These spline based approaches are limited by topological constraints unlike level set based active contour approaches that have been extensively investigated with the use of prior shape knowledge since [5]. PCA is often used in these techniques to compress and summarize the important components of a set of characteristic level sets [5, 6] or control points modeled using Active Shape Modeling (ASM) [7]. Many prior shape based segmentation methods have been demonstrated to be quite robust, providing accurate outlines of the shape of the object being segmented. However, preparation of extensive prior shape knowledge is not always convenient and even cumbersome. Furthermore, many methods can encounter difficulties if the structure of interest assumes a new shape that can not be easily approximated by the shape model. This is a realistic problem for medical images that are associated with an inherent variability associated with population differences and potential pathological structures. Sequential estimation techniques, requiring minimal user assistance are a valid alternative where shape and segmentation information is propagated slice to slice.

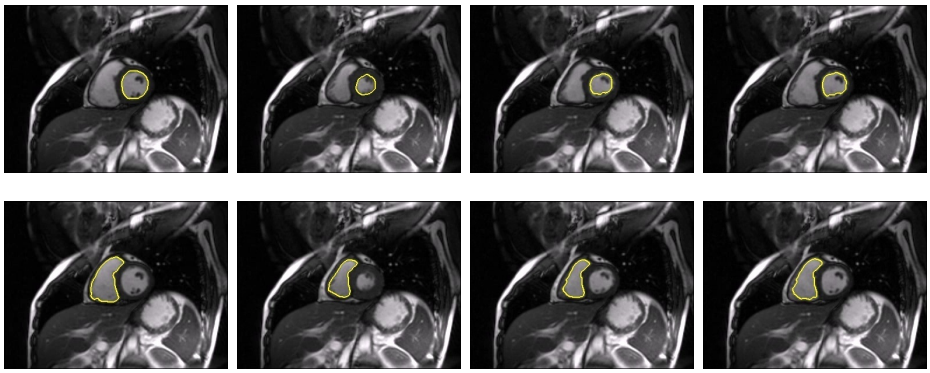

Fig. 1. Example of temporal sequential segmentation in the sagittal plane. Top row shows left atrium outline and bottom row, outline of the right atrium. Images correspond to a cardiac MRI scan acquired at Bristol Royal Infirmary. The algorithm was initialized in both cases by a manually defined region on an initial slice in the sequence for each result.

Sequential segmentation, results illustrated in figure 1 using the methodology described here, has previously been considered by a limited number of authors. An interesting paper by Cho et al. [8] described ways to augment the sequential segmentation process via physical measurements inherent in the MRI scanning process. However the additional physical information is often not available. Senegas et al. [9] utilized se-

quential segmentation techniques for cardiac sequences where shape information was propagated across the temporal dimension from a manually positioned mean location, however the authors utilized a statistical prior model for cardiac shape. Vaswani et al. [10] considered sequential segmentation techniques for both medical data and conventional video. Initialization was provided by manually placed geometric objects with manipulation of parameter settings. The included results appear to be rough estimates of the structures of interest.

We also consider shape in a sequential segmentation framework, where a structure of interest is manually defined for a single 2D image slice. Empirical parameter estimation is also undertaken for the model components. Section 2 describes our methodology, uniquely outlining a model for the sequential shape information in an on-line non-linear setting which is embedded in an image model that considers image information and the overall level set gradient descent based optimization process. Section 3 then presents exemplar results from spatial and temporal sequential estimation processes for pathological and non-pathological structures, respectively. Section 4 closes the work with a short discussion.

2 Methodology

The model of the shape information is described next, in Section 2.1. Section 2.2 then describes the important elements of an image model for n dimensional intensity information which also provides an ideal medium in which to embed the shape model defined earlier. Section 2.3 then summarizes the optimization process used to sequentially estimate the bounds of the structure of interest.

2.1 Shape model

The primary focus of the work here is the shape of a structure of interest S^j that evolves from one image slice j to the next $j+1$. This shape S^j is synonymous to a partition of the image space into two mutually exclusive regions, consisting of foreground $\mathfrak{F}^j = \{\mathbf{x}|\mathfrak{f}^j_{\mathbf{x}}\}$ and background $\mathfrak{B} = \{\mathbf{x}|\mathfrak{b}^j_{\mathbf{x}}\}$ pixels $\mathbf{x}$, where $\Omega^j = \mathfrak{F}^j \cup \mathfrak{B}^j$ and $\mathfrak{f}^j_{\mathbf{x}} \in \{0\ 1\}$ and $\mathfrak{b}^j_{\mathbf{x}} = 1 - \mathfrak{f}^j_{\mathbf{x}}$ are binary foreground and background labels, respectively. The partition of the image space is referred to here with $q^j = \{\mathfrak{F}^j\ \mathfrak{B}^j\}$ for image slice j. The foreground pixels $\mathfrak{F}^j$ define the structure of interest for image slice j.

We consider the level set $\phi^j \triangleq \mathsf{S}^j$ as the primary representation of shape information in our model. ϕ^j enables important geometric information to be conveniently incorporated into the modeling process and it can be considered synonymous to the image partition q^j where pixel level labeling information is fully encapsulated by the level set representation. This can be seen from the properties of the level set which include: $\phi^j_{\mathbf{x}_c} = 0$ on the coterminous foreground and background regions for contour points $\mathbf{x}_c$ and $\phi^j_{\mathbf{x}} = \pm \min |\mathbf{x} - \mathbf{x}_c|\ \forall \mathbf{x}_c | \phi^j_{\mathbf{x}_c} = 0$, i.e. the contour point with minimum Euclidean distance, see e.g. [11]. Also (here in this work) $\phi^j_{\mathbf{x}} \leq 0$ for $\mathbf{x} \in \mathfrak{F}^j$ and $\phi^j_{\mathbf{x}} > 0$ for $\mathbf{x} \in \mathfrak{B}^j$. The shape ϕ^i of the structure of interest for the current image slice can be controlled via comparisons with a set of shapes Φ^{i-1} from a dynamically built space of good shape hypotheses from previous image slices. The comparison of ϕ^i with Φ^{i-1}

should be invariant to translation $\mathbf{T}_{\mathrm{s}}^{i}$, scale $\mathbf{s}_{\mathrm{s}}^{i}$ and rotation $\mathcal{R}_{\mathrm{s}}^{i}$ to enable meaningful comparison resulting in a normalized shape space Ω_{s} representation, where the subscript s refers to the normalized shape space. Thus, Φ^{i-1} and ϕ^{i} have equivalent shape space forms given by $\Theta_{\mathbf{x}}^{i-1} = \Phi_{\mathcal{A}_{\mathrm{s}}(\mathbf{x})}^{i-1}$ and $\theta_{\mathbf{x}}^{i} = \phi_{\mathcal{A}_{\mathrm{s}}(\mathbf{x})}^{i}$ $\forall$ $\mathbf{x}$ where

$$\mathcal{A}_{\mathrm{s}}^{i}(\mathbf{x}) = \mathbf{s}_{\mathrm{s}}^{i}\mathcal{R}_{\mathrm{s}}^{i}\mathbf{x} + \mathbf{T}_{\mathrm{s}}^{i} \tag{1}$$

is the similarity transformation from image space $\mathbf{x} \in \Omega$ to shape space $\mathcal{A}_{\mathrm{s}}^{i}(\mathbf{x}) \in \Omega_{\mathrm{s}}$ for the object shape in image slice i. Shape comparisons also have to be in the current image space requiring the inverse transformation of (1), i.e. $\mathcal{A}^{i}(\mathbf{x}) = \mathbf{s}^{i}\mathcal{R}^{i}\mathbf{x} + \mathbf{T}^{i}$ where $\mathcal{A}^{i}(\mathbf{x}) \in \Omega$. As the shape information is learned on-line, without supervision, the resulting estimated shapes will not be perfect representations and hence can be considered to be inherently noisy. Thus we may define a probabilistic shape space with probability distribution $p_m(\theta|\mathcal{M}^{i-1})$ that represents the distribution of the learned noisy shapes over a normalized shape space $\mathcal{M}^{i-1} = \{\theta^{j}|0 \leq j \leq i-1\}$ consisting of shapes up to the current image slice. We can then define a locally weighted shape space expectation $\bar{\Theta}^{i-1}$ to provide a best estimate over the shape distribution $p_m(\theta|\mathcal{M}^{i-1})$ (which acts as a prior) and a local weighting distribution $p_w(\theta|\theta^{i-1})$. This best estimate can then be used to compare the currently evolving shape rather than a global mean or one based on assumptions on the linearity of the shape space or even one based on local integrity. The local weighting is given here by a Gaussian distribution $p_w(\theta|\theta^{i-1})$ centered on the previous image slice object shape θ^{i-1}, so that[4]

$$\bar{\Theta}^{i-1} = \int_{\theta} \theta \; p_m(\theta|\mathcal{M}^{i-1}) \; p_w(\theta|\theta^{i-1}) \; \mathrm{d}\theta. \tag{2}$$

θ^{j} for $j = 0...i-1$ are considered to be distributed according to $p_m(\theta|\mathcal{M}^{i-1})$, so that the expectation is approximated via

$$\bar{\Theta}^{i-1} = \frac{\sum_{j=0}^{i-1}\left[\theta^{j}\mathcal{W}^{i-1,j}\right]}{\sum_{j=0}^{i-1}\mathcal{W}^{i-1,j}} \tag{3}$$

where $\mathcal{W}^{i-1,j} = \exp(\frac{1}{|\Omega_{\mathrm{s}}|}\sum_{\mathbf{x}\in\Omega_{\mathrm{s}}}(\theta^{i-1}-\theta^{j})^{2})$ is the local weighting for θ^{i-1} and θ^{j} image slice shapes and previously identified object shapes are θ^{j} for image slice j. Each weight can then be considered to form an element in a weight matrix that encompasses the similarity of shapes at different image slices. A simple comparison between ϕ^{i} and Φ^{i-1} can then be the sum of squared differences in the current image space:

$$\mathcal{C}_{\mathfrak{s}}(\phi^{i}, \Phi^{i-1}) = \sum_{\mathbf{x}\in\Omega}\left(\phi_{\mathbf{x}}^{i} - \bar{\Phi}_{\mathcal{A}^{i}(\mathbf{x})}^{i-1}\right)^{2}, \tag{4}$$

where $\bar{\Phi}_{\mathcal{A}^{i}(\mathbf{x})}^{i-1} = \bar{\Theta}_{\mathbf{x}}^{i-1}$, and $\mathcal{A}^{i}(\mathbf{x})$, defined earlier, transforms the shape space estimate $\bar{\Theta}^{i-1}$ to a current image space estimate $\bar{\Phi}_{\mathcal{A}^{i}(\mathbf{x})}^{i-1}$. A sum of squared differences calculation implicitly assumes a Gaussian distribution. Therefore taking the exponential of

[4] The expected shape in a probabilistic region of the shape space (c.f. (2)) is $\bar{\Theta}^{i-1} = \int_{\theta\sim p_w} \theta \; p_m(\theta|\mathcal{M}^{i-1}) \; \mathrm{d}\theta = \int_{\theta} \theta \; p_m(\theta|\mathcal{M}^{i-1}) \; p_w(\theta|\theta^{i-1}) \; \mathrm{d}\theta = \int_{\theta\sim p_m} \theta \; p_w(\theta|\theta^{i-1}) \; \mathrm{d}\theta$.

(4) results in a Gaussian distribution and considering the partition representation then $p_{\mathfrak{s}}(q^i|\boldsymbol{q}^{i-1})$ can be used to symbolize the distributional form of (4), hence

$$p_{\mathfrak{s}}(q^i|\boldsymbol{q}^{i-1}) \propto \exp\left(-\mathcal{C}_{\mathfrak{s}}(\phi^i, \Phi^{i-1})\right). \tag{5}$$

This defines the shape distribution of the foreground region given past observations via a non-linear shape estimation process. A probabilistic image model is now defined, combined with this shape model, providing a medium for sequential segmentation of image structures in volumetric medical images that simultaneously takes account of both the shape and image information.

2.2 Image Model

Let $I^j_{\mathbf{x}} : \mathbb{R}^2 \times \mathbb{R}^+ \rightarrow \mathbb{R}^n$ be an n dimensional image intensity at pixel $\mathbf{x} \in \mathbb{R}^2$ and image slice $j \in \mathbb{R}^+$ where e.g. $n = 1$ for gray scale images. Considering all image slices j up to a current image slice i, i.e. $\forall j$, $j = 0...i$ then we can use Bayes theorem to calculate a conditional probability density $p(\boldsymbol{q}|\boldsymbol{I})$ for a set of image partitions $\boldsymbol{q}$ given a set of image intensities $\boldsymbol{I}$ up to slice i

$$p\left(\boldsymbol{q} = \{q^j|\forall\, j,\ j = 0...i\}|\boldsymbol{I} = \{I^j|\forall\, j,\ j = 0...i\}\right) = \frac{p(\boldsymbol{I}, \boldsymbol{q})}{p(\boldsymbol{I})}, \tag{6}$$

where $p(\boldsymbol{I})$ is the marginal data probability density which is not dependent on the image partition information and can therefore be ignored for the purposes of optimization. $p(\boldsymbol{I}, \boldsymbol{q})$ is a joint probability density which is expanded assuming Markov first order dependence of the image data, $p(\boldsymbol{q}|\boldsymbol{I}) \propto p(\boldsymbol{q}^{i-1}|q^i)p(q^i)p(I^0|q^0)\prod_{j=1}^{i} p(I^j|q^j, I^{j-1}, q^{j-1})$. $p(I^0|q^0)$ is the initial image's data likelihood term, $p(\boldsymbol{q}^{i-1}|q^i)$ is the probability density of all image partitions, except the current slice i.e. $\boldsymbol{q}^{i-1} = \{q^j|j = 0...i-1\}$ given the current slice partition q^i. $p(q^i)$ is the prior probability density of the current slice partition. $p(\boldsymbol{q}^{i-1}|q^i)$ and $p(q^i)$ are related to shape and contour labeling smoothness respectively.

The data likelihood term $p(I^j|q^j, I^{j-1}, q^{j-1})$ can then be divided into foreground and background terms (assuming conditional independent pixel intensities via the partition terms),

$$p(\boldsymbol{q}|\boldsymbol{I}) \propto p(\boldsymbol{q}^{i-1}|q^i)p(q^i)p(I^0|q^0)\prod_{j=1}^{i}\prod_{\forall \mathbf{x}\in\Omega}\left[p_{\mathfrak{F}}(I^j_{\mathbf{x}}|q^j, \mathfrak{m}^j_{\mathfrak{F}})^{\mathfrak{f}^j_{\mathbf{x}}} p_{\mathfrak{B}}(I^j_{\mathbf{x}}|q^j, \mathfrak{m}^j_{\mathfrak{B}})^{\mathfrak{b}^j_{\mathbf{x}}}\right] \tag{7}$$

where $\mathfrak{m}^j_{\mathfrak{F}} = \{I^{j-1}_{\mathfrak{F}}, q^{j-1}\}$ and $\mathfrak{m}^j_{\mathfrak{B}} = \{I^{j-1}_{\mathfrak{B}}, q^{j-1}\}$ and the powers $\mathfrak{f}^j_{\mathbf{x}}$ and $\mathfrak{b}^j_{\mathbf{x}}$ act as mutually exclusive switches between the foreground and background. $p_{\mathfrak{F}}(I^j_{\mathbf{x}}|q^j, \mathfrak{m}^j_{\mathfrak{F}})$ and $p_{\mathfrak{B}}(I^j_{\mathbf{x}}|q^j, \mathfrak{m}^j_{\mathfrak{B}})$ thus correspond to two different PDFs for the foreground $I^j_{\mathfrak{F}} = \{I_{\mathbf{x}}|\mathbf{x} \in \mathfrak{F}^j\}$ and background $I^j_{\mathfrak{B}} = \{I_{\mathbf{x}}|\mathbf{x} \in \mathfrak{B}^j\}$ image intensities respectively. The initial image's data likelihood can also be similarly expanded. However, to save space the non-expanded form will be retained.

A smooth labeling and a smooth boundary (both of which are defined as synonymous to each other here) separating the foreground and background regions are desirable properties of an image space partition for segmentation applications. These properties can be achieved by minimizing the length of the boundary of the partition q^i. Therefore $p(q^i) \triangleq p(\mathfrak{L}) \propto \exp(-\lambda_\kappa \mathfrak{L})$ where λ_κ is an exponential rate parameter and $\mathfrak{L}$ is the length of the contour defining the partition (c.f. [12]). Substituting this term into (7) gives

$$p(\boldsymbol{q}|\boldsymbol{I}) \propto p(\boldsymbol{q}^{i-1}|q^i)p(\mathfrak{L})p(I^0|q^0)\prod_{j=1}^{i}\prod_{\forall \mathbf{x}\in\Omega}\overbrace{\left[p_{\mathfrak{F}}(I^j_{\mathbf{x}}|q^j,\mathfrak{m}^j_{\mathfrak{F}})^{\mathfrak{f}^j_{\mathbf{x}}}p_{\mathfrak{B}}(I^j_{\mathbf{x}}|q^j,\mathfrak{m}^j_{\mathfrak{B}})^{\mathfrak{b}^j_{\mathbf{x}}}\right]}^{\textit{model based intensity competition}}. \tag{8}$$

The intensity part of this equation has been labeled as "*model based intensity competition*" indicating that the foreground and background terms work in competition with each other. Foreground/background competition is the basis of many active contour techniques e.g. as first proposed in [13, 14]. However, the intensity components used here are first order Markovian, i.e. they remember intensity information from the preceding slice via $\mathfrak{m}^j_{\mathfrak{F}}$ and $\mathfrak{m}^j_{\mathfrak{B}}$.

A Gaussian distribution possesses symmetry about the mean, so that we can define here $p_{\mathfrak{s}}(q^i|\boldsymbol{q}^{i-1}) \triangleq p(\boldsymbol{q}^{i-1}|q^i)$ (see (5)) which is of the form found in (8), so that

$$p(\boldsymbol{q}|\boldsymbol{I}) \propto p(\mathfrak{L})\prod_{\forall \mathbf{x}\in\Omega}\left[p_{\mathfrak{s}}(q^i_{\mathbf{x}}|\boldsymbol{q}^{i-1}_{\mathbf{x}})p(I^0_{\mathbf{x}}|q^0)\prod_{j=1}^{i}\left[p_{\mathfrak{F}}(I^j_{\mathbf{x}}|q^j,\mathfrak{m}^j_{\mathfrak{F}})^{\mathfrak{f}^j_{\mathbf{x}}}p_{\mathfrak{B}}(I^j_{\mathbf{x}}|q^j,\mathfrak{m}^j_{\mathfrak{B}})^{\mathfrak{b}^j_{\mathbf{x}}}\right]\right]. \tag{9}$$

This probabilistic model now incorporates memory based intensity competition terms, a spatial smoothness term and a shape based term. The current image slice partition q^i is estimated here using a gradient descent level set based optimization process which is now described.

2.3 Optimization process

A gradient descent level set based approach is used to optimize (9) and (1) in the experiments that follow, see e.g. [15, 16]. The optimization of (9) is made possible via the variational derivative (see Appendix) given by

$$\frac{\partial \phi^i_{\mathbf{x}}}{\partial t} = -2\lambda_{\mathfrak{s}}\left(\phi^i_{\mathbf{x}} - \bar{\Phi}^{i-1}_{\mathcal{A}(\mathbf{x})}\right) + \delta_0(\phi^i_{\mathbf{x}})\left(\lambda_\kappa \mathcal{K}_{\mathbf{x}} - \ln\left(\frac{p_{\mathfrak{F}}(I^j_{\mathbf{x}}|q^j,\mathfrak{m}^j_{\mathfrak{F}})}{p_{\mathfrak{B}}(I^j_{\mathbf{x}}|q^j,\mathfrak{m}^j_{\mathfrak{B}})}\right) + \lambda_{\mathfrak{a}}^{\ln\frac{\mathfrak{a}^i_{\mathfrak{b}}}{\mathfrak{a}^i_{\mathfrak{f}}}}\right) \tag{10}$$

where $\lambda_{\mathfrak{s}}$ is a shape term parameter which corresponds to the inverse variance of the shape density $p_{\mathfrak{s}}(q^i|\boldsymbol{q}^{i-1})$ and $\mathcal{K}_{\mathbf{x}} = -\nabla\cdot(\nabla\phi_{\mathbf{x}}/|\nabla\phi_{\mathbf{x}}|)$ is the curvature of the level set at point $\mathbf{x}$. The curvature result follows $p(\mathfrak{L}) \propto \exp(-\lambda_\kappa\mathfrak{L})$ (in (9)) and using the definition of length defined in [15], i.e. $\mathfrak{L} \triangleq \int_\Omega |\nabla H(\phi^i_{\mathbf{x}})|\mathrm{d}\mathbf{x}$ where $H(\cdot)$ is the

Heaviside function. An advection term has also been introduced to regularize the foreground area $\lambda_{\mathfrak{a}}^{\ln \mathfrak{a}_{\mathfrak{b}}^i}/\lambda_{\mathfrak{a}}^{\ln \mathfrak{a}_{\mathfrak{f}}^i}$ where $\lambda_{\mathfrak{a}}$ is an area term weight and $\mathfrak{a}_{\mathfrak{b}}^i = \sum_{\mathbf{x}\in\Omega} \Delta_{\mathbf{x}} \mathfrak{b}_{\mathbf{x}}^i$ and $\mathfrak{a}_{\mathfrak{f}}^i = \sum_{\mathbf{x}\in\Omega} \Delta_{\mathbf{x}} \mathfrak{f}_{\mathbf{x}}^i$ are the background and foreground areas respectively. This advection force penalizes large changes in the area of the segmented structure, to counteract the reduction in the object area that might occur due to the curvature based force and any mis-alignment of the template shape.

The evolving data PDFs $p_{\mathfrak{F}}(I_{\mathfrak{F}}^i|q^i, \mathfrak{m}^i)$ and $p_{\mathfrak{B}}(I_{\mathfrak{B}}^i|q^i, \mathfrak{m}^i)$ are approximated here with finite Gaussian mixtures with parameters estimated from the histograms of the image intensities from the currently estimated image regions $I_{\mathfrak{F}}^i$, $I_{\mathfrak{B}}^i$ and the image intensities from the previous slice, $I_{\mathfrak{F}}^{i-1}$ and $I_{\mathfrak{B}}^{i-1}$. The parameters of the finite Gaussian mixture models are estimated here using Expectation Maximization and the number of finite mixtures in the Gaussian mixture model was set to 6. This was empirically found to provide the best results for the test sequences.

In common with many active contour techniques manual parameter adjustment is also required to control the relative contribution of the individual components which have to be tuned for sequences with different properties. The similarity alignment transform (see [16]) also has weight parameters which can be kept constant once suitable values have been determined.

3 Experiments and Results

The sequential segmentation process described here can be applied to spatial and or temporal sequences. A single 2D manually defined segmentation is used here for initialization. Subsequent slices are segmented by automatically propagating the shape and intensity information from one slice to the next. Parameter estimation for the different components in (10) is performed using empirical methods.

3.1 Spatial sequential segmentation

A spatial sequential segmentation process result (3D) that identifies the bounds of a tumor in the transverse plane is illustrated in figure 2. This result uses a central transverse 2D slice for manual initialization. Intensity and shape information is then automatically propagated, via the model described here, to the remaining image slices in the 3D data volume. The bounds of the tumor are successfully located for each image slice in the 3D data volume.

The shape of the spatial pathological structure in figure 2 possesses very little variation other than changes in scale. However the shape of the temporal structure of interest that follows (cardiac MRI) possesses greater variation in shape.

3.2 Temporal sequential segmentation

A temporal sequential segmentation result (3D) was illustrated earlier in figure 1 for a Cardiac MRI sequence where the left and right atriums are successfully segmented. The algorithm was initialized using two 2D image slices, one for the right atrium and another for the left atrium.

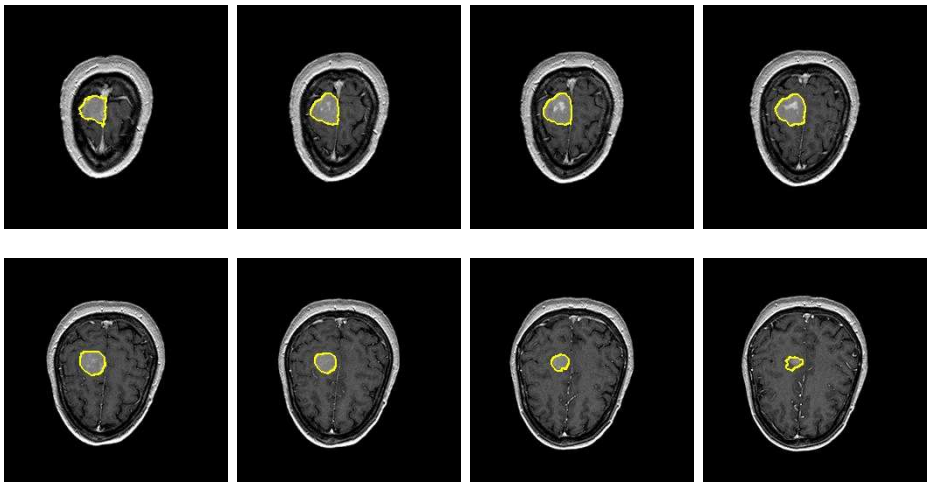

Fig. 2. Result from spatial sequential segmentation process in the transverse plane that identifies the bounds of a tumor for subsequent slices after manual initialization (bottom left). Images correspond to T1 + Gadolinium MRI scan, 59 Year old female at the NMR Center of the Massachusetts General Hospital with a 1.5 Tesla General Electric Sigma and provided by the Center for Morphometric Analysis (http://neuro-www.mgh.harvard.edu/cma/ibsr).

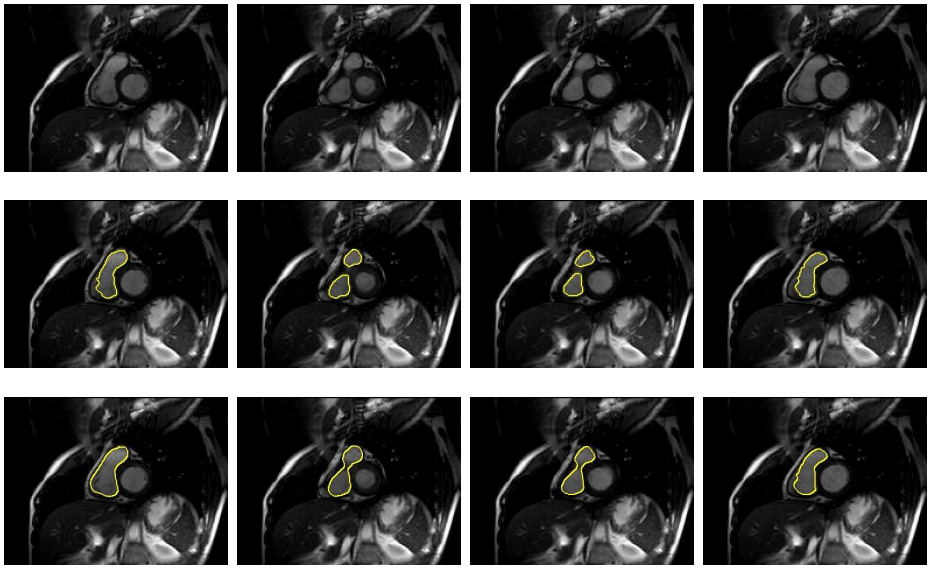

Fig. 3. Comparative temporal sequential segmentation of right atrium in the sagittal plane after single 2D image slice manual initialization. 1st row: original data; 2nd row: sequential segmentation using intensity and curvature information only; 3rd row: sequential segmentation using the proposed method. Images correspond to a cardiac MRI scan acquired at Bristol Royal Infirmary.

A further temporal sequential segmentation result (3D) of the right atrium after manual initialization is shown in figure 3 where the result is compared with a result using intensity information combined curvature but no shape model.

Figure 3 demonstrates an interesting result where the intensity information combined with a curvature based force is not able to identify realistic bounds of the right atrium. However the result corresponding to the work described here using learned online shape information enables prior information to be carried across from one image slice to the next, assisting with identifying the boundary of the right atrium.

3.3 4D: Combining temporal and spatial sequential segmentation

In the results above the temporal and spatial sequential segmentations are obtained in three dimensions. The same technique can be used to segment in higher dimensions. An example of 4D sequential segmentation in both the temporal and spatial domains is presented in figure 4.

The 4D sequential segmentation result in figure 4 was obtained by initializing the algorithm with a single 2D image slice, corresponding to a single sagittal section at a particular time instance. After all image slices in the temporal domain are segmented for a particular sagittal location, the algorithm propagates to another sagittally adjacent image slice for a particular time instance and then subsequent time instances are segmented for that sagittal location. The algorithm was able to successfully propagate the shape information to other time instances and other spatial locations in the sagittal plane. A comparison is also made with an algorithm using no shape model. The shape of the tracked atrium using only an intensity based model with curvature degenerates, particularly as new spatially located image slices are encountered. This is further illustrated by sensitivity calculations in comparison to partial volume calculated ground truths, in figure 5. The model described here has a much better sensitivity at the systolic stage of the cardiac cycle despite the significant changes in the shape of the heart.

4 Discussion and Conclusions

A new active contour based sequential segmentation framework has been presented. This utilizes high-level shape information that is learned on-line, adapting to new shape configurations whilst constraining the evolution of the active contour. Results have shown that the combined framework is able to segment structures of interest undergoing complex deformations of shape. The main shortcomings of the method are that it is computationally complex, requiring significant time to segment a single image slice and, similar to many active contour tracking frameworks, successful segmentation is highly dependent on an empirical selection of parameter values that control the relative contribution of the different model components.

Acknowledgments

This research was funded by the UK Leverhulme Trust and carried out at the Department of Computer Science, University of Bristol with some final experiments performed after the first author joined MFU, Thailand.

Using the model described in this paper:

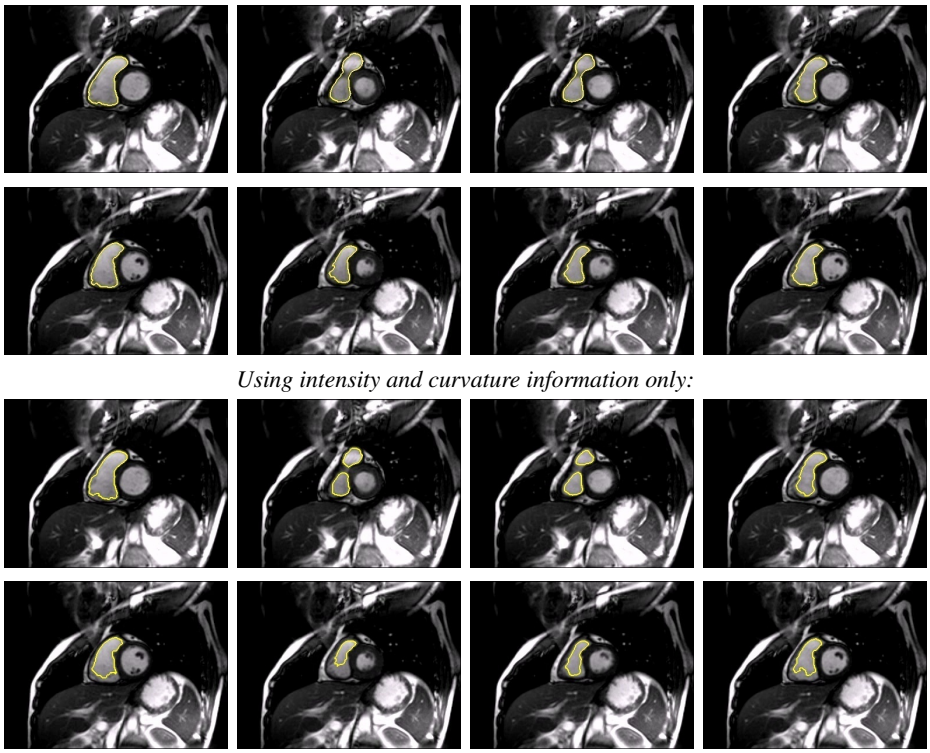

Fig. 4. Results of sequential segmentation in four dimensions in the temporal and spatial domains. Each row is a different sagittal slice through the subject and each column is a different time instance. The right atrium outline is delineated for all image slices for the model described here (top two rows), where the algorithm was initialized using a single 2D image slice in a single sagittal time instance. If only intensity information combined with curvature is used in the modeling process (bottom two rows) then the shape of the tracked region degenerates as more image slices are segmented. Images correspond to a cardiac MRI scan acquired at Bristol Royal Infirmary.

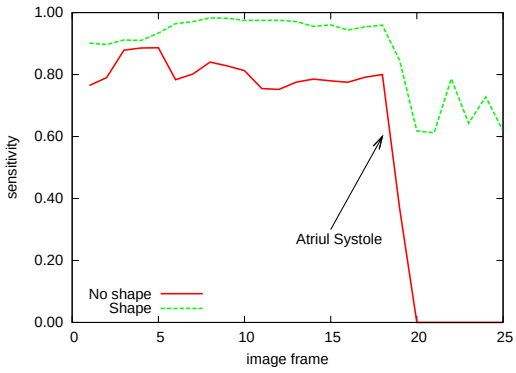

Fig. 5. Quantitative assessment for the 4D cardiac sequence results shown in figure 4.

Appendix

Maximization of (9) can be used to obtain a partition of the image data that most closely matches the given shape model and intensity information. This is also equivalent to minimization of the negative logarithm which is simpler, hence we minimize

$$\mathbb{E}(\boldsymbol{q},\boldsymbol{I}) = -\int_{\Omega}\Bigg[\ln p_s(q^i_{\mathbf{x}}|\boldsymbol{q}^{i-1}_{\mathbf{x}}) \tag{A-1}$$

$$+\ln p(q^i_{\mathbf{x}}|q^i_{\mathcal{N}_{\mathbf{x}}}) + \sum_{j=1}^{i}\Big[\mathfrak{f}^j_{\mathbf{x}}\ln p_{\mathfrak{F}}(I^j_{\mathbf{x}}|q^j,\mathfrak{m}^j_{\mathfrak{F}}) + \mathfrak{b}^j_{\mathbf{x}}\ln p_{\mathfrak{B}}(I^j_{\mathbf{x}}|q^j,\mathfrak{m}^j_{\mathfrak{B}})\Big]\Bigg]\,\mathrm{d}\mathbf{x}.$$

The foreground and background labellings $\mathfrak{f}^j$, $\mathfrak{b}^j$ are equivalent to functions of the Heaviside function on the level sets, i.e. $H(\phi^j_{\mathbf{x}}) = \mathfrak{b}^j_{\mathbf{x}} = (1-\mathfrak{f}^j_{\mathbf{x}})$ hence

$$\mathbb{E}(\boldsymbol{q},\boldsymbol{I}) = -\int_{\Omega}\Big[\ln p_s(q^i_{\mathbf{x}}|\boldsymbol{q}^{i-1}_{\mathbf{x}}) + \ln p(q^i_{\mathbf{x}}|q^i_{\mathcal{N}_{\mathbf{x}}}) + \tag{A-2}$$

$$\sum_{j=1}^{i}\Big[(1-H(\phi^j_{\mathbf{x}}))\ln p_{\mathfrak{F}}(I^j_{\mathbf{x}}|q^j,\mathfrak{m}^j_{\mathfrak{F}}) + H(\phi^j_{\mathbf{x}})\ln p_{\mathfrak{B}}(I^j_{\mathbf{x}}|q^j,\mathfrak{m}^j_{\mathfrak{B}})\Big]\Big]\mathrm{d}\mathbf{x},$$

where now $-\ln p(q^i_{\mathbf{x}}|q^i_{\mathcal{N}_{\mathbf{x}}}) = \lambda_\kappa|\nabla H(\phi^j_{\mathbf{x}})|$. Minimization of (A-2) can be performed via gradient descent on the variational derivative w.r.t. a gradient descent time parameter t, $\frac{\partial\mathbb{E}(\boldsymbol{q},\boldsymbol{I})}{\partial\phi^i} = -\frac{\partial\phi^i}{\partial t}$ where the level sets have been parametrized as a function of pixel location $\mathbf{x}$ and the gradient time parameter t, i.e. $\phi(\mathbf{x},t)$. The variational derivative can therefore be shown to be given by (10). An analytical solution to $\frac{\partial\phi^i_{\mathbf{x}}}{\partial t} = 0$ is not available, therefore finite difference approximations are used to iteratively find the solution, i.e.

$$\phi^i_{\mathbf{x}}(t+1) = \phi^i_{\mathbf{x}}(t) + \Delta_t\Bigg(-2\lambda_s\left(\phi^i_{\mathbf{x}} - \bar{\Phi}^{i-1}_{\mathcal{A}(\mathbf{x})}\right) \tag{A-3}$$

$$+\delta_0(\phi^i_{\mathbf{x}})\left(\lambda_\kappa\mathcal{K}_{\mathbf{x}} - \ln\left(\frac{p_{\mathfrak{F}}(I^j_{\mathbf{x}}|q^j,\mathfrak{m}^j_{\mathfrak{F}})}{p_{\mathfrak{B}}(I^j_{\mathbf{x}}|q^j,\mathfrak{m}^j_{\mathfrak{B}})}\right)\right) - \lambda_{\mathfrak{a}}^{\ln\frac{\mathfrak{a}^i_{\mathfrak{b}}}{\mathfrak{a}^i_{\mathfrak{f}}}}\Bigg),$$

where Δ_t corresponds to the size of the time step. The Dirac delta function restricts the computations to the contour rather than the entire level set. Therefore, following [17], the computations are extended to the entire set of pixels in the image space Ω by replacing $\delta_0(\phi^i_{\mathbf{x}})$ with $|\nabla\phi^i_{\mathbf{x}}|$ and a narrow band ($|\phi^i_{\mathbf{x}}| < \mathfrak{T}$ where $\mathfrak{T} \in \mathbb{R}^+$ is a real positive value) is also used to reduce the number of computations necessary to update the position of the contour. $|\nabla\phi^i_{\mathbf{x}}|$ is approximated here with a non-oscillatory upwind finite difference scheme, see e.g. [18].

The level sets are reinitialized here every t_{reinit} iterations to retain the smoothness and distance properties of the level set using a signed distance transform, [19]. The optimization of (A-3) continues until convergence which is assessed by comparing the sum of squared differences of the level set every t_{reinit} iterations or until a maximum number of iterations have been reached (t_{max}).

References

1. Paladini, D., Vassallo, M., Sglavo, G., Lapadula, C., Martinelli, P.: The role of spatio-temporal image correlation (STIC) with tomographic ultrasound imaging (TUI) in the sequential analysis of fetal congenital heart disease. Ultrasound in obstretics and gynecology **27**(5) (Apr. 2006) 555–561
2. Anderson, R., YenHo, S.: Sequential segmental analysis. In: Continuing Medical Education. Cambridge University Press (1997) 98–116
3. Blake, A., Isard, M., Reynard, D.: Learning to track curves in motion. In: 33rd IEEE Conf. Decision and Control. Volume 4. (1994) 4:3788–3793
4. Baumberg, A., Hogg, D.: An efficient method for contour tracking using active shape models. In: IEEE Workshop on motion of non-rigid and articulated objects. (1994) 194–199
5. Leventon, M., Grimson, W., Faugeras, O.: Statistical shape influence in geodesic active contours. In: Proc. IEEE CS Conf. Computer Vision and Pattern Recognition (CVPR'00). (2000) 316–323
6. Tsai, A., Yezzi, A., Wells, W., Tempany, C., Tucker, D., Fan, A., Grimson, W., Willsky, A.: A shape-based approach to the segmentation of medical imagery using level sets. IEEE Trans. Medical Imaging **22**(2) (Feb. 2003) 137–154
7. Cootes, T., Cooper, D., Taylor, C., Graham, J.: A trainable method of parametric shape description. Image Vision Comput. **10** (1992) 289–294
8. Cho, J., Benkeser, P.: Cardiac segmentation by a velocity-aided active contour model. Comp. Med. Imag. Graph. **30** (2006) 2006
9. Senegas, J., Netsch, T., Cocosco, C., Lund, G., Stork, A.: Segmentation of medical images with a shape and motion model: a Bayesian perspective. In: CVAMIA-MMBIA. Volume 3117 of LNCS., Springer (2004) 157–168
10. Vaswani, N., Yezzi, A., Rathi, Y., Tannenbaum, A.: Time-varying finite dimensional basis for tracking contour deformations. In: Proc. 45th IEEE Conf. Dec. Control, San Diego (Dec. 2006) 1665–1672
11. Osher, S., Sethian, J.: Fronts propagating with curvature-dependent speed. J. Comp. Phys. **79** (1988) 12–49
12. Cremers, D.: Nonlinear dynamical shape priors for level set segmentation. In: Proc. IEEE CS Conf. Computer Vision and Pattern Recognition (CVPR'07). (2007)
13. Paragios, N., Deriche, R.: Geodesic active regions: a new framework to deal with frame partition problems in computer vision. J. Vis. Comm. Image Rep. **13**(1-2) (2002) 249–268
14. Zhu, S., Yuille, A.: Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation. IEEE Trans. Pattern Analysis and Machine Intelligence **18**(9) (1996) 884–900
15. Chan, T., Vese, L.: Active contours without edges. IEEE Trans. Image Processing **10**(2) (2001) 266–277
16. Paragios, N., Taron, M., Huang, X., Rousson, M., Metaxas, D.: On the representation of shapes using implicit functions. In: Statistics and Analysis of Shapes. Springer (2006)
17. Zhao, H., Chan, T., Merriman, B., Osher, S.: A variational level set approach to multiphase motion. J. Comp. Phys. **127** (1996)
18. Aubert, G., Kornprobst, P.: Mathematical problems in image processing. Springer (2006)
19. Borgefors, G.: Distance transformations in digital images. Comp. Vis. Graph. Image Proc. **34**(3) (1986) 344–371