Abstract

We introduce a segmentation framework which extends spatially varying classification to not only incorporate anatomical localization from shape estimation, but to also encode certainty of the localization by local shape variability. The method iterates between a classification step where a statistical classifier learned from feature selection is extended with anatomical localization features, and a shape estimation step where, given the class probability maps, shape is inferred by particle filtering using a level set shape model that accounts for local degrees of anatomical variability. The spatially varying classification is embedded in a geodesic active region framework which allows for local deviations from the inferred shape using an iteratively updated classification based region term. The method is evaluated on late gadolinium enhanced cardiac MRI and is to our knowledge the first automatic segmentation method demonstrated on this type of data.

1. Introduction

Automatic medical image segmentation methods face a number of challenges such as low contrast, noise, missing or ambiguous boundary information, and intensity variations occurring e.g. as a result of disease or of the image acquisition technique. Edge-based segmentation methods have been used in numerous applications (e.g. [2], [4]) but typically need to be initialized close to the desired solution. Region-based methods use statistics of entire regions hence are more global than edge-based [17]. In case of ambiguous boundary information the incorporation of prior shape information can be helpful [17] as well as global models for appearance [3], but for irregular appearance regional statistics need to account for local and non-linear intensity variations [6]. The image intensity might not be a sufficient or even the most useful feature for discrimination between regions.

Statistical classification can deal with local intensity variations, but lacks global boundary and shape information. Spatially varying statistical classification (SVC) was presented in [19], and improved classification by adding global information. Level set methods represent contours implicitly and are widely used in image segmentation [12]. They have been extended to an elegant geodesic active region formalism [13] which allows for local deviations from the inferred shape using an iteratively updated classification based region term. The method is evaluated on late gadolinium enhanced cardiac MRI and is to our knowledge the first automatic segmentation method demonstrated on this type of data.

We propose a method which exploits the benefits and compatibility of SVC and level set methods. From training data, a mean shape represented by a distance function with local shape variability is estimated, and feature selection is performed in order to find meaningful local descriptors for the classifier. The segmentation iterates between shape inference based on the classification, and classification...
tion extended with anatomical localization features where we define a weighting of the classification based on the local shape uncertainty. The inferred shape is a rigid transformation of the mean shape, therefore the SVC is embedded in a geodesic active region framework with a classification based region term, which allows for local deviations from the mean shape and for a segmentation that takes global shape, local shape variability, and local appearance into account. We demonstrate the method on late gadolinium enhanced cardiac MRI, which contains both irregular intensity patterns due to scar tissue at random locations inside the myocardium, and ambiguous boundary information due to low image contrast, scar tissue and the papillary muscles. The analysis of such data currently relies on manual segmentations of the myocardium.

2. Method

The main idea of this work is to incorporate localization certainty in spatially varying classification, which is described in sections 2.6 and 2.8. But first we present the statistical classifier in 2.1 followed by the shape model used described in sections 2.6 and 2.8. We use a Nearest Neighbor (kNN) classification, where class boundaries are determined from the labels of the k nearest neighboring labeled training points in a D dimensional feature space. Each pixel j is described by a feature vector \( u_j \), and the distance \( d \) squared to a training point \( v \) is \( d_j^2 = \sum_{f=1}^{D} (u_{j,f} - v_{f})^2 \). The posterior probability in j for class \( \omega_c \) is \( p(\omega_c | u_j) = \frac{k_{\omega_c}}{k} \), where \( k_{\omega_c} \) among the k nearest neighbors belongs to class \( \omega_c \). We use an approximate kNN classifier which allows for a tradeoff between computational efficiency and search error [1].

2.1. Statistical Classification

We use a \( k \) Nearest Neighbor (kNN) classification, where class boundaries are determined from the labels of the \( k \) nearest neighboring labeled training points in a \( D \) dimensional feature space. Each pixel \( j \) is described by a feature vector \( u_j \), and the distance \( d \) squared to a training point \( v \) is \( d_j^2 = \sum_{f=1}^{D} (u_{j,f} - v_{f})^2 \). The posterior probability in \( j \) for class \( \omega_c \) is \( p(\omega_c | u_j) = \frac{k_{\omega_c}}{k} \), where \( k_{\omega_c} \) among the \( k \) nearest neighbors belongs to class \( \omega_c \). We use an approximate kNN classifier which allows for a tradeoff between computational efficiency and search error [1].

2.2. Feature Selection

Feature selection can increase computational efficiency and improve classification performance. The feature space is initially empty and is expanded iteratively (sequential forward selection) until there is no performance improvement, then decreased (sequential backward selection) until there is clear performance degradation according to a criterion function [8], here we use the Dice similarity coefficient (DSC) [7]. The feature bank consists of all Gaussian derivatives from 0 up to \( 3^{rd} \) order, the intensity (original and Gaussian smoothed), the eigenvalues and eigenvectors of the structure tensor [9] and the Hessian, which describes the first and second order structure locally, on a range of scales. Gaussian derivatives are defined as \( I_{\sigma} = I * D_{\sigma} \),

where \( G \) is a Gaussian, \( D \) a differential operator and \( \sigma \) is the scale. The Hessian and structure tensor are described by:

\[
H(\sigma) = \begin{pmatrix}
I_{xx} & I_{xy} \\
I_{yx} & I_{yy}
\end{pmatrix},
\]

\[
T(\sigma, \sigma_T) = G_{\sigma_T} \ast \begin{pmatrix}
I_x I_y & I_x I_y \\
I_y I_x & I_y I_y
\end{pmatrix}.
\]

2.3. Shape Representation

We use a distance map shape representation which accounts for local variations by assuming a mean shape, \( \Phi_M \), with local degrees of shape variability, \( \sigma_M \), as described in [16]. This method assumes the distribution of distance function values in each pixel to be Gaussian,

\[
p_M(\Phi) = \frac{1}{\sqrt{2\pi\sigma_M}} e^{-\frac{(\Phi - \Phi_M)^2}{2\sigma_M^2}}.
\]

Given a set of \( N \) aligned (training) distance maps, the aim is to recover the distribution which has maximum support by minimizing (using the \(-\log\)) the energy \( E(\Phi_M, \sigma_M) = -\sum_{n=1}^{N} \int_{\Omega} \log[p_M(\Phi_n)] \) subject to the constraint that the mean shape remains a signed distance map \((|\nabla \Phi_M(x)|^2 = 1, x \in \Omega)\) and a smoothness constraint on the variability by minimizing \( E_v(\sigma_M) = \sum_{n=1}^{N} \int_{\Omega} \left[ \frac{1}{2\sigma^2} \right] \). Training shapes are aligned with respect to translation, rotation and scale, \( A = (T, \theta, s) \).

2.4. Shape Inference

Since SVC segmentation of an image relies on statistical classification we wish to infer shape based on the class probability maps. This can be realized in a variety of ways, for example a regional gradient descent approach as in [18] could be adjusted to incorporate classification. We use shape particle filtering as in [6] since that method describes shape inference from kNN classification. A set of shape hypotheses are sampled from transformations of the prior shape model \( \Phi_{M,i} \). Each hypothesis is associated with an image labeling which is compared to the probability maps, and weighted by the likelihood for each region:

\[
W = \exp(\gamma \sum_{j=1}^{n} \log p[u_j | \omega_c]),
\]

where \( r \) is a constant controlling the randomness and \( n \) is the number of pixels inside the template. New hypotheses sets are iteratively resampled proportional to the weights generated from the previous set, so successful shapes will multiply while the unlikely will vanish. The distribution converges to a \( \delta \)-peak at the maximum likelihood solution, which before convergence can be approximated by the strongest local mode.

2.5. Spatially Varying Classification

SVC [19] iterates between global shape matching and statistical classification, a combination that resolves ambiguities in feature space with anatomical context. The kNN
framework allows for a balance between anatomical localization and other features using Euclidean distance in a modified feature space. After initial classification and shape inference, the feature space is expanded with $F$ extra features (corresponding to the number of objects of interest) that provides anatomical localization from the global shape,

$$d^2_j = \sum_{f=1}^{D'} (u_{j,f} - v_f)^2,$$

with $D' = D + F$. With larger differences between $u$ and $v$ of the anatomical localization features, the distance in feature space increases, and this introduces a penalty to the classification. In [19], the object representation is converted to a distance map, giving a penalty that increases with the distance from the object boundary. This formulation fits well with the distance map representation in level set methods. In section 2.6 we describe how the shape model can be converted into features encoding not only anatomical localization but also localization certainty.

### 2.6. Anatomical Localization Features

We perform SVC with $D'$ features of which $F$ are anatomical localization features, one for each object in the image, and define them to be the current level set representations of the objects weighted by the shape model variability outside the object contour:

$$F_i = (1 - H_\alpha(\Phi_i)) \cdot \Phi_i / \sigma^2_{M,i}(A_i).$$

The interpretation of such a feature is that it introduces a classification penalty which is proportional to the distance from the contours weighted by the uncertainty of the boundary locations given the shape variability. Inside the contour the feature value is zero, so that the distance in Equation (1) is unchanged if both $u$ and $v$ are located inside the anatomical structure. Outside the structure, locations with large boundary variations in the aligned training set, thus high boundary uncertainty, leads to a low rate of change of $F_i$ in the normal direction to the contour at that location, and a small feature value relative to the distance from the boundary. Small local boundary variations, on the other hand, represent relatively high prior boundary certainty, and lead to a high rate of change in the direction normal to the boundary thus a larger feature value relative to the boundary. E.g., if $u$ is inside the structure and $v$ is outside, the distance in (1) will be larger if $v$ is at a location with low variability compared to one with high variability. The anatomical localization with local variability is demonstrated in Figure 1.

### 2.7. Geodesic Active Regions

The shape estimate itself is a rigid transformation of the mean shape, and we need to allow local deviations from the mean shape given the image information. Therefore the SVC with localization variability is embedded in a geodesic active regions framework described here. Contour- and region-based evolution terms can be implemented using approximations of the Dirac ($\delta_\alpha$) and Heaviside ($H_\alpha$) distributions [20]:

$$\delta_\alpha(\Phi) = \begin{cases} 0, & |\Phi| < \alpha; \\ \frac{1}{2\alpha}(1 + \cos(\frac{\pi |\Phi|}{\alpha})), & |\Phi| > \alpha \end{cases}$$

$$H_\alpha(\Phi) = \begin{cases} 0, & \Phi > \alpha; \\ 1, & \Phi < -\alpha \\ \frac{1}{2}(1 + \frac{\Phi}{\alpha} + \frac{1}{2}\sin(\frac{\pi |\Phi|}{\alpha})), & |\Phi| < \alpha \end{cases}$$

In these equations $\alpha$ is the region in which the distributions are approximated, and $\Phi$ is assumed to be negative inside the contour it represents. The classical active contour energy [2] can be reformulated as [16]

$$\frac{d}{dt} \Phi_i = \delta_\alpha [g \kappa_i + \nabla \phi_i \cdot \nabla g],$$

where $\Phi_i$ is the level set representation of the $i$th contour, $\kappa_i$ is the curvature and $g$ is a function of the image gradient. Evolution functions for regions are derived in [14], assuming no contour overlap:

$$\frac{d}{dt} \Phi_i = -\delta_\alpha(\Phi_i) \log \frac{p(\omega_i\mid u)}{p(\omega_i\mid u)} \nabla \Phi_i,$$

where $\omega_i$ and $\omega_o$ are the classes inside and outside the object of interest respectively. We use the class probability maps from the supervised classification to estimate $p(\omega_o\mid u)$. The shape term evolves towards a rigid transformation of the mean shape (annotated $\Phi_{M,i}(A_i)$) but is down-weighted where there is high variability,

$$\frac{d}{dt} \Phi_i = -H_\alpha(\Phi_i) \frac{\Phi_i - \Phi_{M,i}(A_i)}{\sigma^2_{M,i}(A_i)}.$$
active region terms which means the regional information in (4) is also iteratively updated from the SVC, hence indirectly benefiting from anatomical context. If the image position is one of the $D$ selected features, then the position in a test image can be shifted so that the object center of mass corresponds to the expected center of mass from training data. The SVC is iterated until there is insignificant change of class labels between iterations according to the classification, and when the region term is thus stabilized the geodesic active region terms evolve until low rate of change of the contours.

3. Experimental Results

3.1. Data Set

We evaluate the method on late gadolinium contrast enhanced MRI short axis data, with in-plane resolution $1.5 \times 2\text{mm}^2$ and thickness of $8\text{mm}$, acquired using a 1.5T system for myocardial scar as in [10]. Characterizing the extent of myocardial scar is essential for clinical treatment of patients with a history of infarction or chronic ischemia, and late contrast enhanced MRI is the leading technique for visualization of scar tissue. Segmentation of the myocardium is the initial and most crucial step in scar tissue analysis, which is important both in clinical studies of its impact on cardiac disorders and in computer aided interventions [10]. Example images and segmentations can be seen in Figure 2. The data set consists of 87 images from 11 scans, where 57 images (7 scans) are used for training and 30 for evaluation. Due to the large inter-slice distance we perform a 2D analysis on this data set, but the method can be extended to 3D applications.

Algorithm 1 SVC Framework

1: Classify image with $D$ features found from feature selection.
2: Infer shapes from class probability maps ($\Phi_{M,i}(A_i)$).
3: Initialize the contours $\Phi_i$ as $\Phi_{M,i}(A_i)$.
4: while Number of changes of class labels between iterations according to $\Phi_i$ are non-negligible do
5: Evolve curves $\Phi_i$ according to (3)-(5) and coupling force.
6: if Number of changes of class labels between iterations according to classification are non-negligible then
7: Calculate anatomical localization features $F$.
8: if Position is in the $D$ features then
9: Shift position feature towards center of mass of object in training data.
10: end if
11: Perform SVC with $D' = D + F$.
12: Infer shapes from the updated class probability maps ($\Phi_{M,i}(A_i)$).
13: end if
14: end while

3.2. Selected Features

Feature selection is performed with scales of 0.9, 1.5, 2.5, 3.5, 5 and 8 pixels, chosen to cover the myocardium thickness and structures of interest in the myocardium. After feature selection on the training data the feature set is the following, in decreasing significance: the position, the intensity smoothed on scales 5 and 8, $I_y$ on scales 5 and 8 (selected twice), $I_{yy}$ on scales 5 and 8,
enhanced MRI data, making these images difficult to segmentation, in particular the scar tissue, in the late gadolinium on our data set can be related to the ambiguous image information for 2 test images due to inferior image quality and the existing annotated data set our method performs well without the need of an advanced alignment method or point correspondences.

We attempted to substitute the region term with the intensity as the only feature (besides anatomical localization) in the supervised classification, so that the kNN classifier determines class boundaries from the prior intensity information only, but for this data set the intensity information could not guide the contours to the desired solution.

4. Discussion

We have presented a method for spatially varying classification with anatomical localization features from a level set shape model that incorporates boundary location uncertainty, which gives a classification penalty proportional to local boundary uncertainty. The SVC is embedded in a geodesic active region framework which allows for local deviations from the inferred mean shape with an iteratively updated, classification based, region term. Numerically, level set methods and spatially varying statistical classification fit well together since all calculations are made directly in the voxels, without need to switch back and forth between different representations. The method is automatic, does not require point correspondence during training, and makes use of prior information both in terms of global shape, local shape variability, and local descriptive features.

While there have been methods to automatically segment the myocardium from other imaging sequences, this is to our knowledge the first method that has demonstrated results of a fully automatic myocardial segmentation from late gadolinium enhanced MRI. Though the algorithm needs to be evaluated on a more extensive data set, the preliminary results presented here indicates that it may become useful for segmentation of this type of data and relieve cardiologists of time consuming labor in scar tissue analysis.

Acknowledgements

The authors would like to thank Marleen deBruijne at DIKU for sharing code and insights on shape particle filtering, and Ole Fogh Olsen at IT University of Copenhagen for discussions.
Figure 2. Three example segmentations from left to right. From top to bottom; input image where the rectangle signifies the region displayed in the following sequence: myocardium probability map from initial classification, same probabilities after convergence of the SVC, manual delineation by cardiologist, and automatic segmentation result. The endo- and epicardium are represented by red and blue contours respectively, and in between, inside the myocardium, scar tissue can be seen as relatively bright areas at varying locations in the images to the left and right.
References


