CLUSTERING ON IMAGE BOUNDARY REGIONS FOR DEFORMABLE MODEL SEGMENTATION

Joshua Stough, Stephen M. Pizer, Edward L. Chaney, Manjari Rao

Medical Image Display & Analysis Group (MIDAG)
The University of North Carolina at Chapel Hill

ABSTRACT

We present a novel approach, clustering on local image profiles, for statistically characterizing image intensity in object boundary regions. In deformable model segmentation, a driving consideration is the geometry to image match, the degree to which the target image conforms to some template within the object boundary regions. The template should account for variation over a training set and yet be specific enough to drive an optimization to a desirable result. Using clustering, a template can be built that is optimal over the training data in the metric used, such as normalized correlation. We present a method that first determines local cross-boundary image profile types in the space of training data and then builds a template of optimal types.

Also presented are the results of a study using this approach on the human kidney in the context of medial representation deformable model segmentation. The results show an improvement in the automatic segmentations using the cluster template, over a previously built template.

1. INTRODUCTION

Deformable model segmentation has been successfully applied in numerous medical image applications. Such segmentations proceed based on the optimization of an objective function that includes a term measuring geometry-to-image match (*image match*, for short) giving the likelihood of the model with respect to the image information. The likelihood function is often trained on a set of ground truth segmentations in order to drive the optimization to expert-like results.

Previous work on building the likelihood function have included intensity profiles associated with individual image points [1], full images referenced with a coordinate system defined by a collection of these points [2], and intensity profiles associated with a tiled boundary [3]. The methods typically make assumptions of a unimodal distribution of corresponding intensity profiles and two of them train each point separately. In our analysis the distribution is often widely variable and multimodal, leading to concerns in the ability to model each point by the family of training images at that point alone.

In this paper, we present a novel approach for training the image match, using clustering of image information in the boundary regions. The method depends upon object-based associations of positions in image space provided by a geometric model. Intensity profiles are generated along normal directions to the densely sampled model boundary from a training set of images to populate a high-dimensional space. As exemplified in our study of kidney CT images (Fig. 3), the profiles fall into a few types depending on the relative position of the anatomic regions surrounding the target object. Clustering determines profile types (cluster centers) prevalent in the data. At each boundary position on the model, a cluster center is then chosen that is the most representative of the profiles in the training set at that position. The likelihood function for new segmentations is then based on normalized correlation of the target image with a template consisting of these position dependent profile types.

The geometric model used in this method is the medial representation, or *m-rep* [4]. M-reps fitted to expert segmentations provide the spatial coordinates from which the corresponding intensity profiles are sampled.

Section 2 reviews the m-rep model, the context of profiles, and the image match computation. Section 3 describes the construction of the new image match template based on cluster centers in the profile space. An experimental study on human kidney segmentation and summary of results is presented in section 4, followed by conclusions and future directions.

2. M-REPS AND IMAGE MATCH

We begin with a brief description of the m-rep model and automatic segmentation using it, followed by details of the image match computation during optimization.

2.1. The M-Rep Model

The m-rep model for a single 3D figure such as a kidney or hippocampus has been previously described in [4]. It is an object representation made up of a manifold of medial atoms, each atom made from a hub and two equal-length spokes. The boundary implied by this representation passes orthogonally through the spoke ends (Fig. 1). A discrete grid samples the medial atoms, and properties of the manifold, like spoke length, are interpolated between grid vertices. A coordinate system on the manifold defines an explicit correspondence between surface points on deformations of the same m-rep model, as well as normal directions. This dense correspondence allows us to characterize image information locally and globally.

M-reps have proven to provide a powerful means of segmentation using deformable model techniques. In this approach a log posterior of the geometric parameters given the image data is optimized. Such optimization is equivalent to optimizing the sum of the log prior, which measures the geometric typicality, and the log likelihood, which measures image match. While geometric typicality is discussed in other papers [4], the image match is the topic of this paper.

2.2. Image Match

The match of a deformed model to an image is an approximation to log $p(\mathbf{I} \mid \mathbf{m})$, where $\mathbf{I}$ is the collection of profiles for the model defined by parameters $\mathbf{m}$. Assuming that $p(\mathbf{I} \mid \mathbf{m})$ is Gaussian, the log probability can be seen to be proportional to $-(\mathbf{I} - \mathbf{\mu})^T \mathbf{\Sigma}^{-1} (\mathbf{I} - \mathbf{\mu})$, where $\mathbf{\Sigma}$ is the covariance matrix of the image data, and $\mathbf{\mu}$ is the template image.

Consider the image and template, profile by profile and make the simplifying assumption that the profiles are uncorrelated. Then the log probability can be rewritten as the sum over all the profiles of the value of $-(\mathbf{I} - \mathbf{\mu})^T \mathbf{\Sigma}^{-1} (\mathbf{I} - \mathbf{\mu})$ for that profile. If each profile is furthermore normalized and assumed to have uncorrelated entries, then $\mathbf{\Sigma}$ becomes the identity matrix and (where $\mathbf{I}$ is again the collection of profiles) log $p(\mathbf{I} \mid \mathbf{m})$ becomes

$$-(\mathbf{I} - \mathbf{\mu})^T (\mathbf{I} - \mathbf{\mu}) = -\mathbf{I} \bullet \mathbf{I} + 2*\mathbf{I} \bullet \mathbf{\mu} - \mathbf{\mu} \bullet \mathbf{\mu} = 2*\mathbf{I} \bullet \mathbf{\mu} - 2 \qquad (1)$$

Thus, except for the addition and multiplication of a constant, optimizing the image match simplifies to maximizing a normalized correlation computation.

3. BUILDING THE TEMPLATE

This section details the construction of the template $\mathbf{\mu}$ from a set of training images, each equipped with a fitted boundary model, from which cross-boundary profiles can
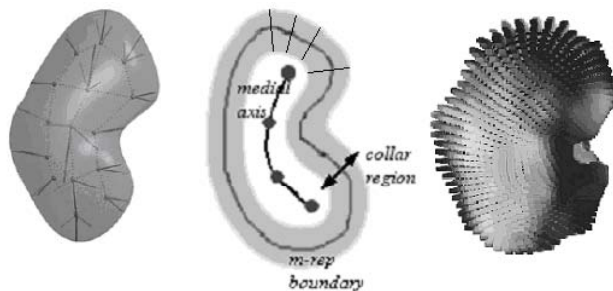

Figure 1 (left to right): A translucent surface rendering of the kidney m-rep model, with underlying grid structure. The middle image shows profiles traversing the collar region normal to the boundary, in the context of a 2D m-rep. Finally, profiles sampled from an abdominal CT scan, forming the target image.

be extracted. Our idea is that a template that is stable against image variability and effective in measuring image match can be made from profiles that are representative of the profiles observed in training images. We choose these representatives by clustering approaches on the collection of all training profiles. We then build the template according to which cluster center profile is most prevalent at the corresponding location. We first focus on building the cluster centers.

3.1. Formation of clusters and cluster centers

The following is done in the context of boundary models with corresponding profiles and correlation-based image match. In the image match, $\mathbf{\mu}$ is the "mean" image, where the match computes the target $\mathbf{I}$'s deviation from $\mathbf{\mu}$. The target is highly variable, both over different images and during the optimization itself, when a slight change of the parameters defining the surface model affects $\mathbf{I}$. Thus a robust estimate of the template is in order.

We begin by considering the space of observed profiles, over the entire surface and over all training cases. Each profile is normalized independently. When the data is clustered, the centers of the clusters represent prevalent profile types. Normalized correlation (which measures the cosine of the angle between two profiles) is used as the distance metric. The clustering algorithm begins with an initial seeding of a number of profile types. For example, for characterizing typical image boundaries, one might choose the positive and negative Gaussian derivatives. All observed profiles are binned according to their response to each type. This is not a fuzzy assignment—each profile is assigned to only one bin. The profile types are then recomputed as the mean of each bin. All profiles are compared against the new types, and the procedure iterates until convergence, when few profiles change bins between iterations. The results of the process are shown in Fig. 3.

The resultant profile types, the bin means computed in the above algorithm's final iteration, are most representative in the least squares and correlational sense:

$$\mu = \text{min over } \Re^m \text{ of } \Sigma \|\nu - \mu\|^2 = \Sigma_{\nu \in \beta} \ (\nu - \mu) \bullet (\nu - \mu) \quad (2)$$

for a bin β of profiles. By the same simplification shown in (1) for normalized vectors, μ gives the highest cumulative correlation with β: the minimizer of (2) is the maximizer of $\Sigma \nu \bullet \mu$. Thus, a segmentation that maximizes the target's correlation with μ will drive to a result that is optimal over all the training data.

3.2. Template Formation from cluster centers

In determining the local profile types, all of the profiles are tossed into a single space. However, the template used for segmentations consists of a profile at every surface position. The last step in building the template then is to use the explicit correspondence between training models to assign a profile type to each surface point. Each position is considered separately. The cumulative response of the observed profiles with each of the profile types—again, all normalized— is computed at that point over all cases. This results in three scores, the highest of which corresponds to that position's most representative profile type. The profile type for each position is computed in the same way. The result is that the template is the bundle of these maximally correlated profile types. This template will optimally drive future segmentations on images of the type represented in the training data.

4. EXAMPLES AND RESULTS

In order to test the efficacy of this template building approach on segmentation, we performed a study on the human kidney in CT images using m-rep deformable model segmentation. These results were compared to a previous study using a constant profile type per point, that of the Gaussian derivative. The right and left kidneys were segmented using different templates since the anatomical context is not the same on the two sides. Training of the templates was based on 52 images, with testing on 12 different CT scans. In this section we present details of this study.

Expert manual segmentations were performed on the training images, resulting in a set of binary image files, bright for kidney and dark for non-kidney voxels. An m-rep model of the kidney was then automatically deformed into each binary, leaving a set of fitted m-rep models. The template used for the image match in these binaries consisted at each point of a positive Gaussian derivative profile, with the inflection at the boundary. This is an effective edge detector for high-contrast images. By then
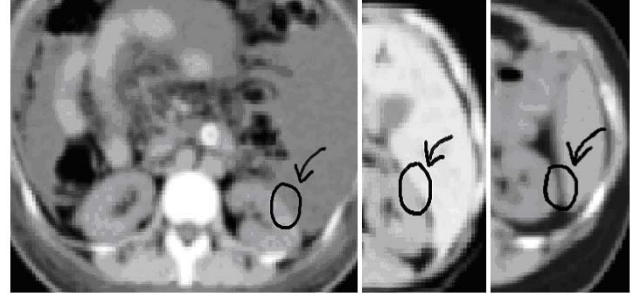

Figure 2: The right kidney is axially viewed in three separate CT scans at reasonably corresponding position. From left to right, the light-dark, dark-light, and notch type profiles are apparent in the marked region of interest.
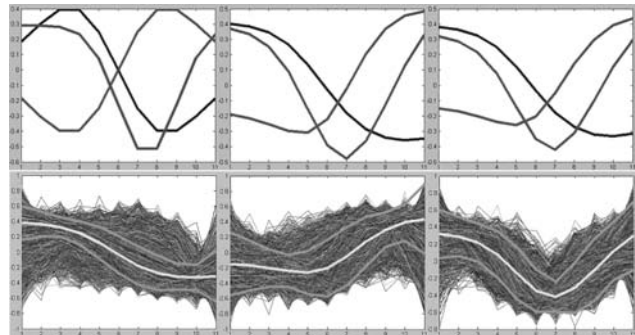

Figure 3: (top row) The three initial profile types converging over iterations of the clustering algorithm. (bottom row) The final binning, where the center lines are those rightmost in the top row. The outer lines show plus and minus two standard deviations in the data.

placing the fitted m-reps into the associated grayscales, the profiles were sampled at the same model boundary coordinates for each case.

The anatomical context of the kidney organ of the abdomen is somewhat consistent across cases. For example, the right kidney usually has the liver abutting or near the superior ventral lateral corner, and a rib supports the dorsal side. The context provides a wide variation in terms of image intensity regions across the surface of the kidney. In CT scans, bone and liver and spleen (the latter superior to the left kidney) all tend to have brightness similar to the kidney, which is in turn brighter than other surrounding tissue. As well, in that these organs do not always abut, there can be darker regions between bright tissues. Thus intuitively, one might expect there to be three profile types. From inside to outside the kidney, a light to dark (for most areas), a dark to light (for abutting liver and other bright tissue), or a light to dark and light again—notch (Fig. 2). As well, the three types can be found at similar position in the kidney, supporting our approach to find the best profile type to use over all cases, not for any particular case.

The three intuitive profile types were analytically modeled with positive and negative Gaussian derivatives
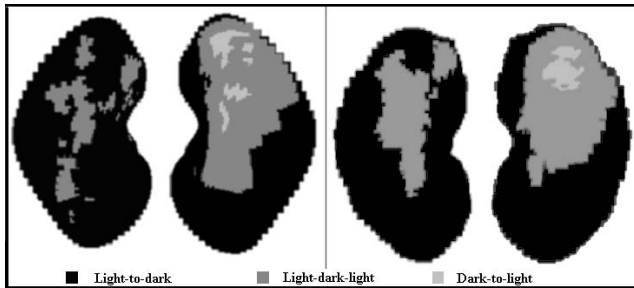
438

Figure 4: The left and right kidney (from left to right) with representative cluster center as shading on the surface, viewed from both sides. From darker to lighter shading the types are light-dark, notch, and dark-light along the profile.

and a negative Gaussian. These seeds quickly converged to prominent observed profile types (Fig. 3). The profile type per position on the surface was then chosen based on correlation of the training data with these cluster centers (see section 3.2). Figure 4 shows the choice as a shade on the surface of the left and right kidney models. The types chosen have intuitive appeal, considering general anatomy. For example, on the right kidney, the profile type chosen to represent the boundary intensity region where the liver often abuts is the dark to light type. Where bright tissue is generally nearby (within the domain of the profiles), there is the notch type.

The left and right templates were then used in m-rep automatic segmentations of the test images. The test images had two expert manual segmentations of the left and right kidneys each. To compare the continuous m-rep model to a binary ground truth image, one can either rasterize the m-rep model in order to compare binaries, or tessellate the boundary in the binaries and compare surface to surface. We chose the latter, using Marching Cubes [5] to find the boundary isosurface in each binary image. Volume overlap and average closest point distance between two surfaces were the metrics used for comparison.

Segmentation using the cluster template approach showed an improvement in 65% of the cases when considering average surface distance to the experts. The average increase in the volume overlap of automatic and expert was 1.3%. While these numbers are not striking, the major improvement in using the cluster template came qualitatively, in the degree of automation. In segmenting with a template of the Gaussian derivative profile type at every position, many cases required modified parameter settings in order to achieve the best results. Such parameters include the relative weights of geometric typicality and image match (see section 2.1). In contrast, all parameters were constant across the segmentations when the cluster template was used. The result is that slightly improved results were obtained with significantly less human input.

5. CONCLUSIONS AND DISCUSSION

We have presented a profile clustering approach for training an optimal template for image match in deformable model segmentation. Simplifying assumptions were made that assert directions for continuing research. To begin, intensity profiles at neighboring positions on the model surface are almost certainly not independent as was posited (section 2.2). One way to address this without hindering the simplicity of the image match computation as it stands could be using multiple scales. This would involve blurring along the surface (research that is ongoing). We are also looking to improve the correspondence between nominally associated boundary positions on the m-rep model. Specifically, the MDL [6] framework may be extended to include intensity statistics.

An additional simplification in the image match model was the choice of a single cluster center per position. The scores of the observed profiles at a point with respect to all the clusters can be turned into priors for a mixture model. While this will not be optimal in the correlation value overall, such an approach may improve the generality of the template, leading to better results in some cases.

More basically, we are working on characterizing the intensity inside an object or within nearby objects versus space interstitial to objects, all in an object relative coordinate frame.

11. REFERENCES

[1] Cootes, T.F., et. al., "Active Shape Models – Their Training and Application," *Computer Vision and Image Understanding*, Elsevier, Vol. 61 (1995), No. 1, pp.38-59.

[2] Cootes, T, et. al., "Unified Framework for Atlas Matching Using Active Appearance Models," *IPMI '99*, Springer LNCS, Vol. 1613, pp.322-333.

[3] Montagnat, J., H. Delingette, "Globally constrained deformable models for 3D object reconstruction," *Signal Processing*, Elsevier, Vol. 71 (1998), pp.173-186.

[4] Pizer, SM, et. al., "Deformable M-Reps for 3D Medical Image Segmentation," *Int. J. Comp. Vis.*, Vol. 55 (2003), Nos. 2/3, pp.85-106.

[5] Lorensen, W.E., H.E. Cline, "Marching Cubes: A High Resolution 3D Surface Construction Algorithm," Proc. SIGGRAPH '87, Vol. 21, No. 4, pp.163-169.

[6] Davies, Rh. H., et. al., "Shape Discrimination in the Hippocampus using an MDL Model," *IPMI '03*, Springer LNCS, Vol. 2732, pp. 38-50.