CLUSTERING ON LOCAL APPEARANCE FOR DEFORMABLE MODEL SEGMENTATION

Joshua V. Stough, Robert E. Broadhurst, Stephen M. Pizer and Edward L. Chaney

Medical Image Display & Analysis Group (MIDAG)
University of North Carolina at Chapel Hill. 27599-3175, USA

ABSTRACT

We present a novel local region approach for statistically characterizing appearance in the context of medical image segmentation via deformable models. Our appearance model reflects the inhomogeneity of tissue mixtures around the exterior of the object of interest by determining mixture-consistent local region types relative to the object boundary. The region types are formed by clustering local regional image descriptors. We partition the object boundary according to region type and apply principal component analysis on the cluster populations to acquire a statistical model of object appearance that accounts for local variability in the object exterior.

We present results using this approach to segment bladders and prostates in CT in the context of day-to-day adaptive radiotherapy for prostate cancer. Results show improved fits versus those obtained with a previously developed method.

Index Terms— Local appearance, Clustering methods, Image match, Bayesian segmentation, Biomedical image processing

1. INTRODUCTION

Medical image segmentation is a challenging and high cost task. A goal common to many researchers around the world is automating this task. In Bayesian deformable model segmentation for medical images, a geometric model for an object of interest is deformed via its shape parameters to fit an image object that the model represents. Such segmentations proceed based on the optimization of an objective function that includes a term measuring geometry-to-image match–*image match*, for short. One image match that has proven successful is the likelihood, which up to a logarithm measures the likelihood of the model deformation parameters m given the image information I, written as $p(I|m)$. The likelihood function is often trained on a set of ground truth segmentations in order to drive the optimization to expert-like results.

Previous likelihood functions include those based on intensity profiles that are either associated with individual image points [1] or averaged over similar image points [2], and those based on tuples that are formed from intensity-derived features at an ordered collection of these points [3]. The

first two of these methods have difficulty capturing the interrelations among pixel intensities in a region. The third method makes the assumption of a unimodal distribution of each feature at each image point. However, we find that the distribution of intensities or other intensity-derived features at a point or within a region is often multimodal because within various image cases different tissue types fall at the same point.

Region-based methods address these concerns by modeling intensity distributions in object-relative regions. These approaches typically sample image intensities within the object interior or separately the interior and exterior. Some of the resulting intensity models are trained on summary statistics such as mean and variance [4], which themselves capture limited information. Other recently developed models use the entire intensity distribution, and either compute an image match with respect to a single reference distribution [5], or most recently, train the variation in the distribution and compute a real probability to use as the likelihood $p(I|m)$ [6].

In [6] we use two object-relative regions: interior and exterior to the object boundary. We sample from the regions and convert the resulting distributions to regional intensity quantile functions (RIQFs), which are amenable to linear statistical methods such as principal component analysis (PCA). One drawback of the method is its failure to account for the inhomogeneity of tissue mixture as one moves around the exterior. While it may be reasonable to model intensities in the interior of an organ as samples from a single distribution, the exterior may consist of neighboring organs, bones, and fat and connective tissue, the intensities of which should not be considered samples from a single source.

In this paper, we extend the work of [6] to account for the diverse local intensity distributions found on the exterior of an object of interest. We capture intensity quantile functions at many local regions outside the object boundary in training and cluster on the pooled data, deciding the number of sources, or region types. The region types robustly represent intensity in larger areas of relative homogeneity, such as the neighboring organs. We then partition the boundary on region type using the data only at corresponding regions. Finally, we apply PCA to the individual cluster populations to obtain a Gaussian model of the variability in the intensity distribution for each cluster. As a combination of the cluster probability per region, our likelihood thus reflects the local variability of
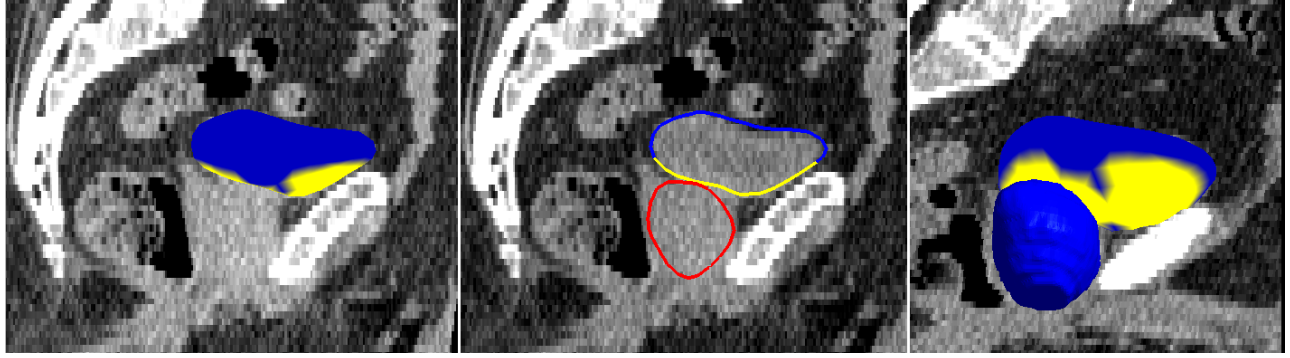
Fig. 1. Sagittal views of a male pelvis in CT, with bladder boundary colored by region type. Left: a 3D view. Middle: the in-plane contour also colored by region type, with the prostate shown for reference. Right: an off-sagittal 3D view of the same bladder, now with the prostate also in 3D and in the foreground. Note the brighter prostate and bone tissues inferior to the bladder correspond to a region type distinct from the darker bowel tissue regions superior to the bladder.

intensity on the exterior.

Section 2 reviews the intensity quantile methodology and describes our construction of image match using clustering and PCA in the quantile function space. Section 3 summarizes the segmentation framework and presents results on bladders and prostates. In section 4, we posit future directions.

2. QUANTILE FUNCTIONS AND IMAGE MATCH

We begin with a brief description of RIQFs, followed by details of our method in image match training and computation.

2.1. Regional Intensity Quantile Functions

In [6] we describe an approach to image match to probabilistically represent the appearance of an object in an image. Appearance is in the form of regional intensity quantile functions, derived from intensity histograms within object-relative regions, such as the interior near the object boundary. The space of RIQFs has the advantage that certain common changes in a distribution, such as mean shift and variance scaling, are represented linearly under an appropriate metric. Principal component analysis is then used to characterize the variability in a distribution.

In the following we review the construction of the RIQF in the context of the distance metric that provides linearity. Let q and r be the continuous, one-dimensional intensity distributions in two regions between which we wish to measure the similarity. The Mallows distance [7] between q and r, with cumulative distribution functions Q and R, respectively, is defined as

$$M_p(q,r) = \left(\int_0^1 |Q^{-1}(t) - R^{-1}(t)|^p dt \right)^{1/p}. \quad (1)$$

An n-dimensional RIQF is then the discretized inverse cumulative distribution on intensities in a region, i.e., $Q^{-1}(t)$ or

$R^{-1}(t)$ in the above equation. Let these discretized quantile functions be denoted $\mathbf{q}$ or $\mathbf{r}$. Coordinate j of $\mathbf{q}$ or $\mathbf{r}$ stores the average of the $[\frac{j-1}{n}, \frac{j}{n}]$ quantile of the intensity distribution for that region, i.e, $\mathbf{q}_j = \int_{(j-1)/n}^{j/n} Q^{-1}(t)dt$. Then the Mallows distance above corresponds (up to a scale factor) to the L_p vector norm between $\mathbf{q}$ and $\mathbf{r}$,

$$M_p(q,r) \approx \left(\frac{1}{n} \sum_{j=1}^n ||\mathbf{q}_j - \mathbf{r}_j||^p \right)^{(1/p)}. \quad (2)$$

Through RIQFs, distributions are understood as points in an n-dimensional Euclidean space in which distance corresponds to the M_2 metric, and mean and variance changes are linear.

2.2. Image Match

In this section, we describe our image match. In general, we want an image match that considers regions relative to the object of interest. The question is what constitutes a region. In the context of medical image segmentation, consider that an organ or other volume whose local intensity distributions are distinguishable from those of neighboring volumes constitutes such a region. That these regions exist is the cause of inhomogeneity in the object exterior. Short of the less tractable problem of modeling the 3D shapes of these regions, we consider them only as they affect the local distributions near the boundary of the object of interest. For this, we define our larger region types through clustering on the RIQFs of smaller regions. Our idea is that an image match that accounts for object-relative region type will provide superior fits to a model that considers only the global exterior.

We now detail the construction of our image match in training. The inputs to training are images of the object of interest and the geometric models fitted to those images. Our image match requires geometric correspondence on the boundary of the object model, which allows us to compute RIQFs
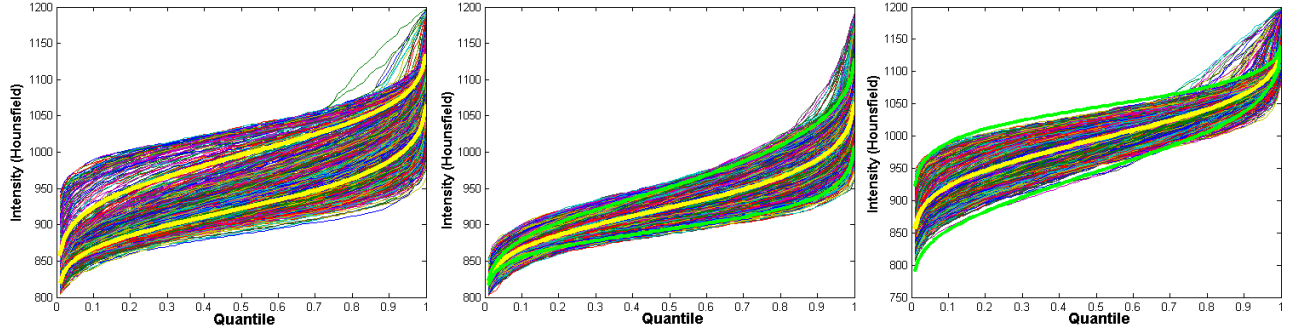
Fig. 2. Clustering on the local RIQFs for the bladder exterior. The left shows the pooled RIQFs with the two cluster centers overlaid. The middle and right images show the two cluster populations with mean and ± 2 standard deviations overlaid. The reasonable separation into lighter and darker distributions is evidenced in Fig. 1.

for regions centered at corresponding points densely sampled around the object boundary. We partition the boundary by RIQF cluster type, as measured in the space of local RIQFs. Then at target time we use appearance statistics on region type to compute an overall image match.

More precisely, over all training images I_p, we compute RIQFs $\mathbf{q}^{i,p}$ for many local regions anchored to individual points i on the object boundary (e.g., 306 for the bladder, see 3.1 for details). The contribution of a voxel to $\mathbf{q}^{i,p}$ is Gaussian weighted by its distance to the surface and is equal to zero if its closest point on the surface is not close enough to point i. Most voxels contribute to several local regions. This overlapping of regions means that the pooled samples may be correlated, potentially leading to a biased clustering. However, having overlapping regions smoothes the final image match and helps avoid local minima in the objective function.

We then cluster the pooled set of RIQFs for all boundary points and images, $\{\mathbf{q}^{i,p}, \forall i, p\}$, using Fuzzy C-means Clustering [8]. We choose the number of clusters according the target problem (see 3.1). The results are the set of region types, $M = \{\mu_k\}$ represented as RIQFs, and cluster membership scores $\{u_k^{i,p}, \forall i, p\}$ that minimize the sum of score-weighted distance in the M_2 metric used:

$$M = \min_{\{\mu_1, \ldots, \mu_K \in \Re^n\}} \sum_{\forall i, p} \sum_{k=1}^{K} u_k^{i,p} * ||\mathbf{q}^{i,p} - \mu_k||^2, \quad (3)$$

with $u_k^{i,p} \in [0, 1]$ and $\sum_k u_k^{i,p} = 1$ for given i and p (Fig. 2).

For the clustering procedure, all of the samples are considered together. For the next step, we use the explicit correspondence between training boundaries to consider each surface point separately and assign it to one of the μ_k from the clustering step. For a given point i, we compute the total response of the observed RIQFs to cluster k using the cluster membership scores, $s_k^i = \sum_p u_k^{i,p}$. The cluster for which s_k^i is maximum is most representative of the intensity distributions seen at that point in training. We partition the boundary by assigning the point i the corresponding region type (Fig. 1).

The last step in training is to perform PCA on the individual cluster populations in order to build Gaussian models of the probability of RIQFs per region type. We first segregate the pooled RIQFs according to $\max_k u_k^{i,p}$, the membership scores, to provide the populations. After PCA, each surface point is then assigned the Gaussian model of its previously assigned region type. Figure 2 shows the local region quantile functions in training, with the cluster statistics overlaid.

Some points will have quite variable RIQFs over training, so a particular cluster may not stand out as a clear choice. This may occur where neighboring organs have widely disparate configurations relative to the boundary of the object of interest. In this case we use PCA on all the RIQFs as its appearance rather than those of one of the cluster populations, to compensate for the greater variability at such points.

The final image match is best understood in the context of its computation during the segmentation of a target image. At the lowest iteration, a prospective boundary is supplied and its image match requested. The program samples from voxels near the boundary and populates the local RIQFs. For each boundary point and its observed quantile function, a Mahalanobis distance is computed with respect to the principal modes and variances of that point's region type. The Mahalanobis distances are summed over all points, weighted by the points' comparative importance (related to how many voxels contributed to the local RIQF).

3. RESULTS

We give segmentation results on bladder and prostate. Our data consists of five patient image sets, each of approximately 16 daily CT scans taken during radiotherapy courses. The images have an in-plane resolution of 512×512 with voxel dimensions of 0.98mm $\times$ 0.98mm and an inter-slice distance of 3mm. We consider the patients separately, segmenting the images from one patient in a leave-one-day-out study where the image match is trained on all daily images except the target. In section 3.1 we discuss our shape model and segmentation

framework. In section 3.2 we present segmentation results.

3.1. The Segmentation Framework

We use m-rep models to describe the shape of prostates and bladders [9]. The object representation is a discretely sampled grid of medial atoms, where each atom consists of a hub and two equal-length spokes. The boundary of the object model passes orthogonally through the spoke ends. Properties, such as spoke length and orientation, are interpolated between atoms in the grid. The model defines a coordinate system which provides an explicit correspondence between deformations of the same mrep model and the 3D volume in the object boundary region. We perform semi-automatic segmentation by starting with a mean model initialized in a target image using a similarity transform computed from two landmarks. Segmentation proceeds by a conjugate gradient optimization of the posterior of the geometric parameters given the image data.

We use a boundary point density that places 306 points on the bladder model surface and 290 on the prostate at fixed object relative coordinates. We set the number of clusters to be two for both object exteriors, with the following justification. The bladder is surrounded mostly by lower intensity bowel and fat, with much brighter tissue at the pubic bone area and prostate inferior to it. The prostate has brighter tissue exterior to it in the pubic bones areas and bladder, with darker tissue elsewhere (Fig. 1).

3.2. Segmentation Results

We are provided expert manual segmentations of the training images. M-reps are then automatically deformed into the manual segmentations, leaving a set of m-rep models whose boundaries are at the image object boundaries. By then placing the fitted m-reps into the associated greyscales, the local RIQFs are sampled in the training images and consequently the RIQF clusters (Fig. 2), region partitions (Fig. 1), and PCA results are computed.

We use volume overlap (intersection over average) with the expert manual segmentations to compare our boundaries to boundaries computed using the global exterior image match described in [6]. Segmentations using our local regions appearance scheme improve results in 57.5% (46 of 80) of bladders and 53.8% (43 of 80) of prostates over all patients. Considered separately (as they are trained and segmented), our scheme improved bladder segmentations in three of five patient image sets, with one wash and one failure, and improved prostates in three patients with two failures (where the failure is to improve results in a majority of the patient's images). In the successful patients, bladders were improved in 68.1% (30 of 44) of images while prostates were improved in 67.3% (33 of 49). These improvements are in the context of segmentations that already approach expert quality, with median overlap of 92.1%, exceeding the agreement we observe between experts.

4. FUTURE DIRECTIONS

Segmentation using our method based on RIQF clustering outperforms a global region approach on the given prostates and bladders. However, the geometric correspondence that the method depends on is in general problematic. Presently the image correspondence that we in essence capture with our region types is strongly tied to the geometric correspondence on the object of interest. However, for example, the region type representing the bladder may change position relative to the prostate across days, while the geometric correspondence of the prostates will not. Our goal is then to model the changes in the region type partition on the object surface. We look to improve results in cases where the geometric and image-implied correspondences disagree, when both the global region and our fixed region methods perform less well. We will use the target image to decide the region type locally, and then correct this initial partition to make it more representative of the corresponding regions as seen in training.

5. REFERENCES

[1] T.F. Cootes et al., "Active shape models - their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, January 1995.

[2] Joshua Stough et al., "Clustering on image boundary regions for deformable model segmentation," in *ISBI*, 2004.

[3] I.M. Scott et al., "Improving appearance model matching using local image structure," in *IPMI*. 2003, vol. 2732 of *LNCS*, pp. 258–269, Springer.

[4] Andy Tsai et al., "A shape-based approach to the segmentation of medical imagery using level sets," *IEEE Trans. Medical Imaging*, vol. 22, no. 2, pp. 137–154, 2003.

[5] Daniel Freedman et al., "Model-based segmentation of medical imagery by matching distributions," *IEEE Trans. Medical Imaging*, vol. 24, no. 3, pp. 281–292, 2005.

[6] Robert E. Broadhurst et al., "A statistical appearance model based on intensity quantile histograms," in *ISBI*, 2006.

[7] E. Levina, *Statistical Issues in Texture Analysis*, Ph.D. thesis, University of California at Berkley, 2002.

[8] Jim C. Bezdec, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York, 1981.

[9] S.M. Pizer et al., "Deformable m-reps for 3d medical image segmentation," *IJCV*, vol. 55, no. 2, pp. 85–106, November-December 2003.