Regional Classification of Texture in 3D Ultrasound of the Prostate

Jared Vicory[1], Mark Foskey[1], J.S. Marron[1], Aaron Fenster[3], Aaron Ward[2], Stephen Pizer[1]

[1] University of North Carolina at Chapel Hill
[2] Robarts Research Institute
[3] University of Western Ontario

Abstract. Locating an object, such as the prostate, in 3D ultrasound requires distinguishing the appearance of neighborhoods just outside the boundary from those just inside. That appearance should be based on both intensity and texture. Because these appearance differences vary strongly around the prostate, a regional approach is desired. The first of our methodological contributions is to have the appearance model be based on probabilities of each voxel being inside the target object (the prostate). The second is to compute these probabilities using regional classification of intensity and texture features. To represent the prostate we use deformable skeletal models that naturally generate these regions. The results of an experiment with 16 training cases and 13 target cases show our method can consistently identify significant regional differences in texture inside and outside the prostate boundary.

1 Introduction

In medical image analysis applications such as segmentation, an *image match* term measuring how well an object matches image data is used. This term is based on some appearance value for each voxel in an image. It can be desirable to have multiple appearance features per voxel. In our driving problem of prostate segmentation from 3D trans-rectal ultrasound, both intensity and texture information is important to distinguish the prostate from its surroundings.

Appearance of anatomy surrounding an object (and even of the object itself) can vary widely around the boundary of the object (see fig. 1), suggesting one build localized appearance models for many regions around the boundary of an object. Doing this requires a model that defines well corresponding local regions across many cases.

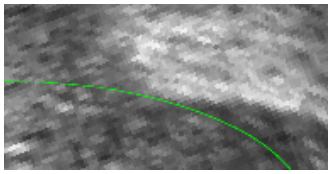

Fig. 1. Regions near the prostate boundary (green line) with significantly different appearance.

We combine these two ideas into an appearance model for the prostate in ultrasound images, where the overall image match value is the sum of the match values in all of the local regions. These

regional image match values must be computed from the vector-valued appearances. We accomplish this by using regionally-trained classifiers to compute, for each voxel, the probability that that voxel came from inside the prostate. Our regional image match is based on the mean probability of being inside the prostate for voxels inside and outside the boundary on each region.

We focus on the development of an appearance model which can be applied to a wide range of problems on different objects in different modalities. The details of our prostate segmentation method using this appearance scheme in a deformable-model segmentation approach is left to another paper; the success of the segmentation is used as one means of evaluation our appearance scheme.

2 Background

2.1 Previous Work

Previous work on prostate segmentation from ultrasound includes work done using intensity features [1], work done using texture [2], and work using texture classifiers [3–6], one of which uses large regions selected from a tiled boundary mesh [6] and another of which trains target-specific classifiers for prostate subvolumes [3].

2.2 Object Representation

To facilitate consistent local regions we use a skeletal representation, the discrete s-rep. S-reps produce shape spaces on an object population with few modes of variation [7]. Having so few modes of variation increases regional correspondence [8] among objects fit from this shape space.

A *discrete s-rep* is a sampling of the interior of an object. It consists of a grid of samples called *hubs* on the object's skeleton, which is approximately medial, and, at each of these samples, vectors called *spokes* which point from the skeleton to the object boundary and are nearly orthogonal to it. These spokes consist of both the direction from the skeleton to the boundary as well as the distance between skeleton and boundary. Hubs on the edge of the skeleton have three spokes: one which points to the top of the object, one which points to the bottom, and one which points to the crest of the object where the top sweeps around to the bottom. Hubs on the interior have only top and bottom spokes. An example s-rep is shown in fig. 2.

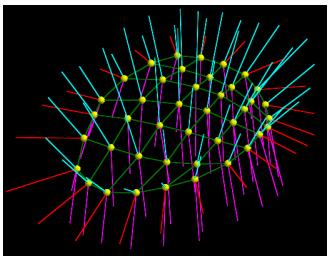

Fig. 2. S-rep model. The skeletal sheet positions are shown as yellow balls. The top spokes are shown in magenta, the bottom spokes are shown in cyan, and the crest spokes are shown in red.

We use a 7×6 sampled grid to represent the prostate. This gives 42 top spokes, 42 bottom spokes, and 22 crest spokes. We use these 106 spoke ends to define local regions on the boundary.

2.3 Classification

Classification of two classes of features is commonly done using a support vector machine (SVM), which computes a direction of highest separation between the two classes. Classification decisions are made by projecting values onto this direction. SVM separation directions have been shown to be less robust to noise than the method we use, distance-weighted discrimination (DWD)[9]. This is especially important in our driving problem, since ultrasound data is particularly noisy.

3 Method

To build a local appearance model, we train feature classifiers regionally. For each of these regions, voxels just inside the boundary form one class, while voxels just outside form the other.

For each region, a feature tuple $\mathbf{f}$ comprising intensity and texture (or other) derived features can be pooled from every member of the population. We then train classifiers on these pooled image features using DWD, giving a separation direction $\mathbf{v}$ in feature space which best separates voxels inside the object from those outside the object.

We then project each training point $\mathbf{f}$ onto the separation direction, giving a value d. If we consider the threshold on the separation direction to be at 0, then the sign of d gives which class the point is classified to, and the magnitude gives distance from the threshold. This distance gives a measure of how well classified the point is, as points further from the classification threshold will in general be classified correctly more often than those closer to it. We compute a d value for all n training points and form two histograms, one for d values from voxels inside the boundary and one for the outside voxels. These histograms represent the probability distributions $p(d|inside)$ and $p(d|outside)$, respectively.

If we are given a new feature tuple, such as one from a target image to be segmented, we wish to determine whether that tuple is from the inside or outside of the object. If we know which region of the boundary the tuple is from, we can project the tuple along that region's separation direction and get its d value. We wish to know the probability that this value comes from inside the object, $p(inside|d)$. Using Bayes Rule,

$$p(inside|d) = \frac{p(d|inside)p(inside)}{p(d|inside)p(inside) + p(d|outside)(1 - p(inside))}, \qquad (1)$$

where $p(inside)$ is the prior probability of being inside the object and must be chosen. We set $p(inside|d)$ to 0 (resp. 1) in the left (resp. right) tail of the distribution.

By computing this value for all voxels near the boundary of a target object, a *p-image* can be created, where each voxel's intensity represents the probability that that voxel is inside the object. We then want to use this image to compute a match value which measures how closely the target object fits the data. We

do this by computing the mean $p(inside|d)$ in the inside (μ^{in}) and the mean $p(inside|d)$ in the outside (μ^{out}) areas for each object region. Large values of $\mu^{in} - \mu^{out}$ correspond to better segmentations. A second term is added to favor probabilities which straddle $p = 0.5$. Without this term, cases with means 0.6 and 0.4 would not be differentiated from cases with 0.9 and 0.7 or 0.3 and 0.1. Our final match value E, for an object with k regions, is defined as

$$E = \sum_{\Omega_{i=1\ldots k}} \left[\left(\mu_i^{in} - \mu_i^{out} \right) - \left| \mu_i^{in} + \mu_i^{out} - 1 \right| \right] \tag{2}$$

4 Application

In this section, we describe the application of the methodology outlined in section 3 to our target problem of prostate segmentation from 3D TRUS.

4.1 Ultrasound Texture Features

For ultrasound images, using texture features is expected to produce better segmentation than using intensity alone.

In the problem of segmenting the prostate from ultrasound, intensity information alone does not yield an accurate basis for distinguishing areas just inside and just outside the prostate. We combat this problem by considering not only intensity information, but texture information as well.

Due to the nature of ultrasound imaging, much of the texture information is found along or orthogonal to the insonation direction. A 3D ultrasound image is typically made from a series of fans, with the insonation direction running from the inside to the outside of this fan. In order to better analyze the texture in an image, it must first be *unfanned* or *flattened*, as shown in Fig 3. This transformation is done so that the insonation direction now lies along the z-axis of the image, allowing for easier access to the directions most likely to contain meaningful texture. Also, the x-axis of the unfanned image goes across the fans.

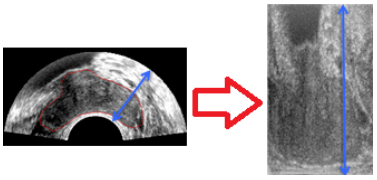

Fig. 3. An original, fanned ultrasound on the left and its defanning on the right. The blue line represents an insonation direction.

After the image has been flattened, the intensity and texture in the image must be separated. Speckle-reducing anisotropic diffusion [10] is used to reduce speckle and obtain more reliable intensities. These intensities can then be subtracted from the original image, leaving only texture information behind. We analyze this texture using oriented Gabor filters, commonly used in computer vision. These filters are computed in the two orthogonal planes containing the ultrasound insonation direction. We compute features in 6 orientations, 2 scales, and 2 phases. Together with intensity, we have a 49-dimensional feature tuple for each voxel in the image. Further improvements could be seen by using different texture features, including physically-based features.

4.2 Classifier Training

We define patches on an s-rep boundary as the portion of that boundary starting at the end of a spoke and extending halfway to its neighboring spokes. This region is computed by interpolating a denser grid of spokes given the initial discrete set of spokes. Our s-reps have 106 spokes; thus we have 106 boundary patches.

For each of these patches, we pool the feature tuples from multiple images to form texture profiles for the inside and outside areas. We then apply DWD to find a separation direction in the 49-dimensional feature space which best separates each patchs inside and outside voxels. We then project all of the training points onto this direction to construct the $p(d|inside)$ and $p(d|outside)$ histograms. These are then transformed into the $p(inside|d)$ function, which is used for segmentation.

5 Results

This section presents results on a data set consisting of 16 training and 13 target cases, each of which consists of a manually segmented ultrasound image of a prostate.

5.1 Significance of Regional Classifiers

First, the effectiveness of the regional classifiers must be analyzed. We first determine for each patch if there are meaningful differences between the texture features for the inside and outside classes in each region. To test this, each region was subjected to a Direction-Projection-Permutation (DiProPerm) hypothesis test [11].

In this test, all training features in the region are projected onto the separating direction. Then, the difference of the means of the d values for each class is computed. The training points are then randomly relabeled and a new classification direction is computed, followed by reprojection. This yields null inside and outside d distributions, and the difference of their means can be computed. By iterating this process, we can form a null distribution of these mean differences. By computing the z-score of the mean difference from the original labelings, we can get a sense of how different the two classes are.

Fig. 4 shows that for all 106 regions the difference of the means of the original labeling lies much greater than 3 standard deviations from the mean of the differences computed through the random relabelings. This indicates that the two classes are truly separate for all of our regions around the prostate boundary.

5.2 Misclassification Rates

While the z-score statistic is useful in determining if there are actually two groups present in a region, it does not directly indicate how well the trained classifiers do in determining whether a feature is inside or outside of the object. For this,

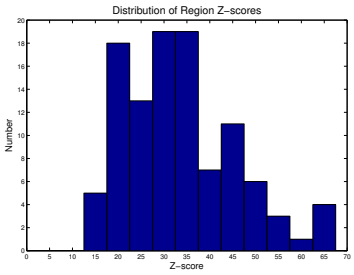

Fig. 4. Histogram of z-scores resulting from DiProPerm tests. These scores show strong significance for all regions.

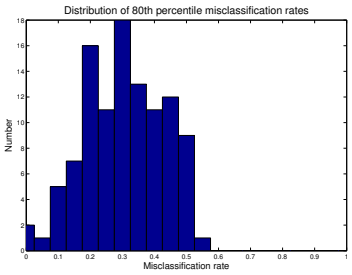

Fig. 5. Histogram of 80th percentile misclassification rates. Most of the regions perform well most of the time.

we must look directly at the misclassification rates of the regional classifiers on target images not used in the training.

For the thirteen target cases, we fit s-reps to the manual segmentations. We then compute inside and outside misclassification rates for the regions they imply. As fig. 5 shows, 80 of the 106 have lower than 35% misclassification at the 80th percentile level. All 106 patches have average misclassification rates lower than 35%. Even though parts of the boundary in a region may have poor classification, over the whole region the classification success can still be high. Fig. 6 shows a region where part of the outside is poorly classified but still performs well overall.

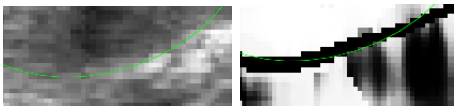

Fig. 6. Part of a region (left, in the ultrasound image) and its corresponding $p(inside|d)$ values (right). Brighter pixels represent higher probability. The prostate boundary is shown in each.

5.3 Probability Images

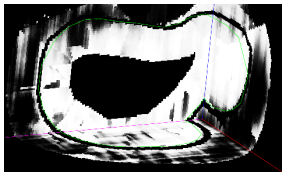

Fig. 7. Three slices of the prostate in target p-image. Intensities on the inside are higher than outside the green boundary.

Using the boundary regions defined by the fitted s-reps, we can classify all voxels near the boundary of the prostate in a target case and determine their p value. We can then compute an image where each voxel's intensity represents the probability that it came from inside the prostate. As fig. 7 shows, this image shows larger intensities inside the boundary compared to lower intensities on the outside in most regions.

To be useful for segmentation, these images must show good separation between the means of the inside and outside classes. We compute $p(inside|d)$ for all voxels both inside and outside each region. We create regional histograms of $p(inside)$ distributions for inside and outside classes and perform a 2-sample t-test to show that these distributions are significantly different. We find that, on average, 95 of the 106 distributions show significant difference between the inside and outside p distributions ignoring multiple comparison issues.

5.4 Segmentation Results

We apply this appearance model in a deformable-model-based segmentation method. Our s-rep shape space is learned from the deformations the prostate undergoes between MRI (planning) and TRUS (biopsy) images. The mean deformation is applied to an s-rep fit to a patient's MRI image to give a patient-specific initialization for TRUS segmentation. By deforming this model in the shape space and doing local refinements in regions with a low misclassification rate across all training cases, we can reliably find the prostate boundary. Figure 8 shows a result for one such case with a mean absolute difference between our result and the manual segmentation of 2.23 mm and a Dice similarity coefficient of 0.881.

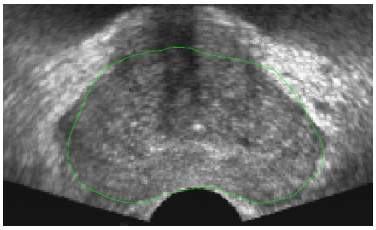

Fig. 8. 2D slice of preliminary 3D segmentation result of the prostate from ultrasound.

6 Discussion

The results presented in this paper show that region-specific statistical classification of intensity plus texture features to produce an image whose voxels give $p(inside)$ is an effective means of modeling the appearance of 3D ultrasound images. While ultrasound image segmentation has been seen as challenging, the derived probability images show properties consistent with strong performance in segmentation using difference of mean probabilities as an image match term.

The power of the regional approach within this method is demonstrated by the fact that, in all regions, it consistently identifies significant differences between areas just inside and just outside the boundary of an anatomical object. There are remaining questions on how exactly to define the regions and if their boundaries should be soft rather than hard that are worth exploring.

Future work should investigate which texture features and how many should be used for ultrasound segmentation. Physically-based features seem an attractive possibility. Also, the method should be tested with different subdivisions of our 29 ultrasound images into 16 training and 13 target images. As well, the method should be tested on other sets of segmented 3D ultrasound images.

While the results in this paper were obtained using specific texture features on ultrasound images, this method could be extended to work with any appearance model with many features per voxel, as long as the geometric model allows for easily creating regions which correspond over multiple instances of an object.

The use of the proposed appearance model in segmenting the prostate in ultrasound using our deformable model framework is left to another paper.

References

1. Fuxing Yang, Jasjit Suri, and Aaron Fenster. Segmentation of prostate from 3-d ultrasound volumes using shape and intensity priors in level set framework. In *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE*, pages 2341–2344. IEEE, 2006.
2. D. Shen, Y. Zhan, and C. Davatzikos. Segmentation of prostate boundaries from ultrasound images using statistical shape model. *IEEE Transactions on Medical Imaging*, 22:539–551, 2003.
3. Xiaofeng Yang, David Schuster, Viraj Master, Peter Nieh, Aaron Fenster, and Baowei Fei. Automatic 3d segmentation of ultrasound images using atlas registration and statistical texture prior. In *SPIE Medical Imaging*, pages 796432–796432. International Society for Optics and Photonics, 2011.
4. Amjad Zaim. Automatic segmentation of the prostate from ultrasound data using feature-based self organizing map. In *Image Analysis*, pages 1259–1265. Springer, 2005.
5. Yiqiang Zhan and Dinggang Shen. Automated segmentation of 3d us prostate images using statistical texture-based matching method. In *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2003*, pages 688–696. Springer, 2003.
6. Yiqiang Zhan and Dinggang Shen. Deformable Segmentation of 3-D Ultrasound Prostate Images Using Statistical Texture Matching Method. *IEEE Transactions on Medical Imaging*, 25(3):256–272, 2006.
7. Stephen M Pizer, Sungkyu Jung, Dibyendusekhar Goswami, Jared Vicory, Xiaojie Zhao, Ritwik Chaudhuri, James N. Damon, Stephan Huckemann, and J.S. Marron. Nested Sphere Statistics of Skeletal Models. In *Dagstuhl Workshop on Innovations for Shape Analysis: Models and Algorithms*, 2011.
8. Kaleem Siddiqi and Stephen M. Pizer. *Medial Representations: Mathematics, Algorithms, and Applications*. Springer, 2008.
9. J.S. Marron, M.J. Todd, and J. Ahn. Distance Weighted Discrimination. *Journal of the American Statistical Association*, 102:1267–1271, 2007.
10. Yongjian Yu and Scott T. Acton. Speckle Reducing Anisotropic Diffusion. *IEEE Transactions on Image Processing*, 11(11):1260–1271, 2002.
11. S. Wei, C. Lee, L. Wichers, G. Li, and J.S. Marron. Direction-Projection-Permutation for High Dimensional Hypothesis Tests. *submitted to Journal of Multivariate Analysis*, 2012.