A Stochastic Approach to 3-D Image Modeling

Dhiraj Joshi
Department of Computer Science and Engineering
Pennsylvania State University
University Park, PA 16802
djoshi@cse.psu.edu

Jia Li
Department of Statistics
Pennsylvania State University
University Park, PA 16802
jiali@stat.psu.edu

James Z. Wang
College of Information Sciences and Technology
Pennsylvania State University
University Park, PA 16802
jwang@stat.psu.edu

***Abstract*— Statistical modeling methods have been successfully used to segment, classify, and annotate digital images, over the years. In this paper, we present a 3-D hidden Markov model (HMM) for volume image modeling. The 3-D HMM is applied to volume image segmentation and tested using synthetic images with ground truth. Potential applications to 3-D biomedical image analysis are also discussed.**

I. Introduction

Computer based content analysis of digital images has become very important today. Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) scanners in hospitals can produce high-resolution 3-D images of the human brain or the human body. Conventional 2-D image modeling paradigms may not always be effective in volume image analysis as there is a third dimensional linkage that they cannot capture. Researchers have strived to extend existing algorithms for modeling and analyzing large-scale multi-dimensional data. 3-D Markov random fields based techniques have been proposed for medical image segmentation [5]. Among other modeling paradigms, hidden Markov models (HMM) have particularly demonstrated high effectiveness in modeling speech, image, and video. Pseudo 3-D HMMs have been proposed for face recognition [2]. 2-D multiresolution HMMs have been successfully used for supervised image segmentation and image annotation [3], [4]. Here we present a 3-D HMM, characterize its segmentation performance and discuss its computational complexity.

II. Model Assumptions of 3-D HMM

In 3-D modeling, a volume image is represented by feature vectors on a 3-D grid. An image may be divided into cubes which could be overlapping. In such a case, every cube corresponds to one position in the grid. The 3-D HMM model represents a 3-D image as a statistical process, the parameters of which have to be learned. More specifically, every point in a 3-D grid exists in a latent state, which is influenced by the states of its geometric neighbors. This geometric dependence is Markovian in nature as shown in Figure 1. Every state has an emission distribution associated with it which generates the feature vectors. The observed feature vectors are assumed to be conditionally independent given the states. In our model, the emission distribution is assumed to be multivariate Gaussian distribution, with the probability density function represented as

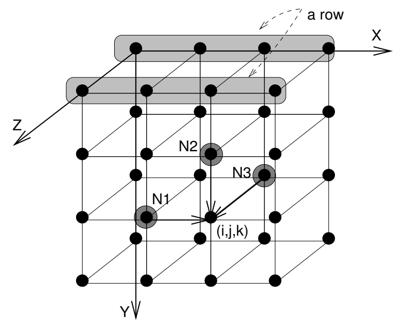

Fig. 1. A 3-D grid representing a 3-D image. Given the states of all the points that precede point (i,j,k), only the states of the three indicated neighboring points affect the distribution of the state at (i,j,k).

$$b_l(u) = \frac{1}{\sqrt{(2\pi)^d|\Sigma_l|}} e^{-\frac{1}{2}(u-\mu_l)'\Sigma_l^{-1}(u-\mu_l)}.$$

Here u represents the d dimensional feature vector of a point modeled as an instance of the multivariate Gaussian distribution with mean vector μ and covariance matrix Σ. The challenging issue here is to estimate the parameters of the 3-D HMM model for a given image. The parameters to be estimated consist of the state transition probabilities and the mean and covariances of the Gaussian distributions. Due to the large number of parameters, we regularize the transition probabilities by a partial 3-D dependence. In [1], we proposed a computationally efficient algorithm for doing so. Briefly, the estimation is performed in an iterative fashion using the Viterbi approach. Viterbi gives the optimal state sequence for a row of points under a fixed set of parameters. The parameters themselves get updated iteratively. The proposed approach is shown to be polynomial time with respect to the number of states of the model and the problem size.

*This work was supported in part by the US National Science Foundation under Grant Nos. IIS-0219272, IIS-0347148, and ANI-0202007, The Pennsylvania State University, the PNC Foundation, and SUN Microsystems under grants EDUD-7824-010456-US and EDUD-7824-030469-US. We acknowledge Stephen Wong for useful discussions during early stage of the project. The full version of this work will appear in *IEEE Transactions on Image Processing* in 2006.

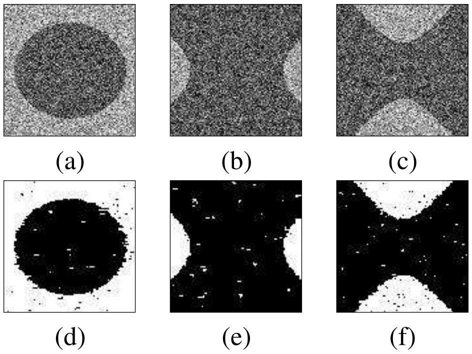

Fig. 2. Segmentation of a hyperboloid image into two classes using 3-D HMM. The figure shows 3 frames one each in the $X-Y$, $Y-Z$ and $X-Z$ planes. (a)-(c) is the original image; (d)-(f) is the segmented image. The parameters a = 35, b = 25, c = 30, and σ = 0.6. Error rate is 1.7%.

III. Unsupervised Segmentation of 3-D Images

The 3-D HMM has been applied to volume image segmentation. Experiments were performed using a large pool of synthetic volume images. Each image contained a standard mathematical 3-D shape, centered about the image. The initial parameter set was determined using a naive k-means clustering approach. Assuming that the clusters define the states of the model, the sample means and covariances determine the initial Gaussian distributions. The Viterbi approach is iteratively applied to refine the parameter set and state assignment to the points. The final segmentation depends upon the states of points at the end of the iterations. The method of image generation was as follows. Points in the interior of the 3-D shape were assigned black color while the rest were white. Each color voxel, black (μ = 0) and white (μ = 1), was perturbed by an additive Gaussian noise $\sim N(0,\sigma^2)$ and the voxel values were truncated to lie in the interval $[-2\sigma, 1+2\sigma]$. For the purpose of displaying images, voxel values in the interval $[-2\sigma, 1+2\sigma]$ were scaled to $[0,255]$. A unidimensional feature was used for each color voxel. The 3-D shape parameters (length of semi axes and radii, denoted by a, b, and c) and the noise parameter σ were varied to form a pool of 70 images.

IV. Discussion of Results

The regularization parameter α determines the extent of 3-D dependence. Segmentation performance as α is varied is shown in Table I. Note that the best performance usually occurs at an intermediate value of α. A trade-off between model complexity (complete 3-D model, $\alpha = 1$) and ease of estimation (2-D model, $\alpha = 0$), is preferred in most cases, and the results support this hypothesis. As the variance of Gaussian noise is increased the segmentation becomes harder. In Table II, we tabulate the best and median segmentation performances for different values of σ. As is evident from the results, 3-D HMM performs reasonably well segmentation even for large values of σ. The running times of 3-D HMM segmentation program for image sizes ($w \times w \times w$), where w takes values 50, 100, 150, and 200, were found out to be 32s, 280s, 798s, and 938s respectively. These numbers support the fact that the complexity of the algorithm is linear in problem size (w^3). Due to space limitations, we are unable to present all the results and comparisons here. Please see [1] for more details.

parameters		α					
a	c	0.0	0.2	0.4	0.6	0.8	1.0
20	20	0.0144	0.0116	0.0068	0.0072	**0.0064**	0.0816
20	30	**0.0104**	0.0114	0.0110	0.0106	0.0105	0.0276
20	40	0.0129	0.0124	0.0120	**0.0115**	0.0119	0.0372
30	30	0.0365	0.0342	**0.0314**	0.0338	0.0372	0.0373
30	40	0.0338	0.0598	0.0449	0.0511	**0.0233**	0.0353
40	40	**0.0091**	0.0115	0.0468	0.0184	0.0687	0.0456

TABLE I

Compare the segmentation performances of torii images (size $100\times100\times100$ and $\sigma = 0.5$) as the regularization parameter α is varied between 0 and 1. The best performance is indicated in bold.

σ	0.2	0.3	0.4	0.5	0.6	0.7
$\mathscr{P}_{best}$	0.0001	0.0005	0.0019	0.0037	0.0067	0.0058
$\mathscr{P}_{med}$	0.0004	0.0041	0.0157	0.0406	0.1207	0.1995

TABLE II

Compare the best ($\mathscr{P}_{best}$) and median ($\mathscr{P}_{med}$) segmentation performances over 70 images (size $100\times100\times100$) as the variance of the Gaussian noise increases from 0.2 to 0.7.

V. Conclusion

We presented 3-D HMM for volume image modeling. The model and estimation methodology were briefly discussed. We also demonstrated its performance on synthetic 3-D images. Model regularization parameter was varied to determine the best performance for a given problem and segmentation over large noise variances was found to be robust. The described framework can be useful for modeling 3-D medical image data. We are collaborating with certain experts in biology and are currently in the process of acquiring 3-D images. Through participation in this workshop, we wish to entail feedback from the audience, learn more about problems in biomedical imaging and foster potential collaboration with participants. The success of multidimensional hidden Markov models for 2-D image analysis gives us confidence in our 3-D HMM.

References

[1] D. Joshi, J. Li, and J. Z. Wang, "A Computationally Efficient Approach to the Estimation of Two- and Three-dimensional Hidden Markov Models," *IEEE Transactions on Image Processing*, 2006, to appear.

[2] F. Hulsken, F. Wallhoff, G. Rigoll, "Facial expression recognition with pseudo-3D hidden Markov models," *Proc. DAGM-Symposium, Lecture Notes in Computer Science,* vol. 2191, pp. 291-297, 2001.

[3] J. Li, R. M. Gray, R. A. Olshen, "Multiresolution image classification by hierarchical modeling with two dimensional hidden Markov models," *IEEE Trans. Information Theory,* vol. 46, no. 5, pp. 1826-41, August 2000.

[4] J. Li, J. Z. Wang, "Automatic linguistic indexing of pictures by a statistical modeling approach, " *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1075-1088, 2003.

[5] M. Li'evin, N. Hanssen, P. Zerfass, E. Keeve, "3D Markov random fields and region growing for interactive segmentation of MR data," *Proc. MICCAI*, 2001.