A WAVELET DOMAIN HIERARCHICAL HIDDEN MARKOV MODEL

Zhen Ye, Cheng-Chang Lu

Department of Computer Science
Kent State University, Kent, OH, 44242

ABSTRACT

This paper proposes a wavelet-domain hierarchical hidden Markov model for an unsupervised texture segmentation. Based on a hybrid graph structure, the global dependencies can be captured by a quad-tree structure across all scales, and local dependencies at higher resolution scales can be captured by a pyramidal graph structure. A novel context model that includes different positions, orientations, and scales is introduced. Applications of an unsupervised texture segmentation are presented. Compared with other alternative approaches for several test images, this method can achieve a significant improvement in segmentation, especially at higher resolution scales.

1. INTRODUCTION

Textured image segmentation is one of the most challenging problems in image processing and computer vision. In order to obtain both reliable and accurate segmentation results for given textured images, multiscale approaches which can describe both large-scale and small-scale behaviors are widely applied and perform well. Recently, several multiresolution Bayesian segmentation approaches combined with Markov random fields (MRFs) have been proposed [1][2] [3][4][5]

Due to the significant relationships that exist among different scales and image blocks, how to represent globe and local dependencies are critical for multiscale statistical models. In [1], a multiscale random field (MSRF) was proposed by Bouman *et al*. Markovian dependencies across scales can be captured by a quad-tree model or a pyramidal graph model. In wavelet domain, Crouse *et al* proposed a hidden Markov tree (HMT) model [6] to capture the statistical properties of wavelet coefficients. This model has been applied to denoising and supervised segmentation [2]. The interscale dependencies between subband coefficients are characterized by a quad-tree structure.

One problem associated with the multiscale texture segmentation is that the segmentation results will be inaccurate at higher resolution scales. Texture is characterized not only by the gray value at a given pixel, but also by the gray value "pattern" in a neighborhood surrounding the pixel [7]. Therefore, due to the small block size at higher resolution scales, the statistical information is insufficient to describe each block's properties, which causes low segmentation accuracy.

In this paper, we propose a new wavelet-domain hierarchical hidden Markov model based on a hybrid graph structure for unsupervised texture segmentation. The global dependencies can be captured by a quad-tree structure across all scales. At higher resolution scales, a pyramidal graph structure is applied to capture both interscale and intrascale dependencies. Several novel contexts involving more local information are defined at higher resolution scales to provide additional information for small image blocks.

2. HIERARCHICAL HIDDEN MARKOV MODEL

Let the observed data W be the corresponding set of wavelet coefficients of an image I. In the HMT model [6], each individual wavelet coefficient w_i is associated with a hidden state s_i, and a two-state Gaussian mixture model associated with hidden states is applied to describe the non-Gaussian distribution property of w_i. The hidden states have a Markov dependency structure which is captured by the quad-tree structure. We can indicate the dependencies in wavelet coefficients by their hidden states. Fig. 1 shows the structure used by the HMT model to capture the interscale dependencies.

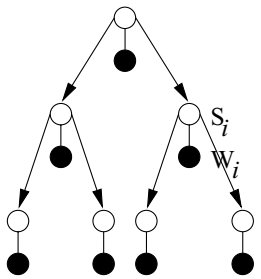


Fig. 1. One dimensional analog to quad-tree structure of the HMT model.

Work reported herein is supported in part by the CS Ph.D. enhancement program by the Ohio Board of Regents.

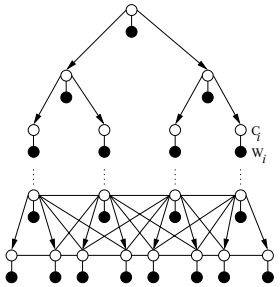


Fig. 2. One dimensional analog to a hybrid graph structure of the HHMM. Black dots represent the wavelet coefficients and white dots represent the hidden class labels.

However, the HMT models suffer from less statistical information at higher resolution scales. In order to consider more local information at higher resolution scales, in this paper, a new hierarchical hidden Markov model (HHMM) is proposed. A hybrid graph structure is used to capture both interscale and intrascale dependencies (see Fig. 2). Denote the index of the highest resolution scale by L, the HHMM has a quad-tree structure across all scales and a pyramidal graph structure only for scales at a higher resolution than l $(0 < l < L)$.

Similar to the HMT model, a hidden class label c_i^k is associated with each wavelet coefficient w_i^k in the HHMM, where i is the index of coefficients at scale k. Assuming the total number of classes in an image I is M[1], the overall marginal probability density function of w_i^k is an M-component Gaussian mixture

$$f(w_i^k) = \sum_{m=1}^{M} p(c_i^k = m) f(w_i^k | c_i^k = m),$$

where $p(c_i^k)$ is the probability mass function of hidden class labels c_i^k, and $f(w_i^k | c_i^k = m) = G(\mu_{i,k,m}, \delta_{i,k,m})$ is a Gaussian density function. To capture the interscale dependencies between scales, a first-order Markov chain is applied to hidden class labels by a quad-tree structure (see Fig. 1) that is the same as the HMT model. The transition probabilities can be described as

$$p(c_i^k = m | c_{\rho(i)}^{k-1} = n) = p(c_i^k = m | c_{\rho(i)}^{k-1} = n, c_{\rho(\rho(i))}^{k-2} \cdots),$$

where $\rho(i)$ is the parent of i.

Denote $\phi^t = \{\phi_m^t\}$ as the parameter vector of the quad-tree structure, we have $\phi_m^t = \{p_{rt}, p(c_i | c_{\rho(i)}), \mu_{i,m}, \delta_{i,m}\}$, where p_{rt} specifies the initial probability of a quad-tree's root. An iterative *Expectation Maximization* (EM) algorithm, which is the same as the training method employed by the HMT model, is used to estimate the parameters. The details of implementation can be found in [6].

[1] Here, we assume we have known the number of classes in an image.

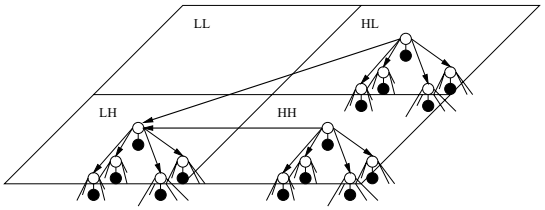


Fig. 3. An example of the inter-orientation context model. The context vector is a set of local correlated trees.

Suppose T_i is a quad-tree that originates from w_i. For any subtree T_j in T_i, we can define the α and β functions as: $\alpha_j(m) \equiv p(c_j = m, T_{i \setminus j} | \phi^t)$ and $\beta_j(m) \equiv f(T_j | c_j = m, \phi^t)$, where $T_{i \setminus j}$ is a tree obtained by removing the subtree T_j from T_i. According to the Bayes rule, we can have

$$p(c_j = m | T_i, \phi^t) = \frac{\alpha_j(m)\beta_j(m)}{\sum_{n=1}^{M} \alpha_j(n)\beta_j(n)}.$$

The calculations of the α and β functions were introduced in [6]. Thus, for any tree T_i^k at scale k, we can obtain the conditional class label probability $p(c_i^k | T_i^k, \phi^t)$. The probability $p(c_i^k | T_i^k, \phi^t)$ is useful for segmentation problems [1]. We can get the class label for each tree by the maximum value of $p(c_i^k = m | T_i^k, \phi_m^t), 1 \le m \le M$.

Due to small block size at higher resolution scales, only interscale dependencies are insufficient to indicate the property of T_i^k. Thus, at higher resolution scales, more dependencies, not only from the parent but also from local neighbors, will be included based on a pyramidal graph structure (see Fig. 2 and Fig. 3). A local context vector v_i^k is defined for a treeT_i^k, where k is greater than l. With the assumption that the T_i^k is independent, given its context v_i^k, we can have the condition probability

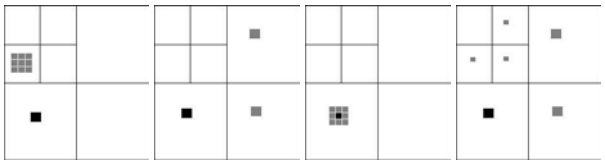
$$p(c_i^k | T_i^k, v_i^k) = \frac{p(c_i^k = m) f(v_i^k | c_i^k = m) \beta_i^k(m)}{\sum_{n=1}^{M} p(c_i^k = n) f(v_i^k | c_i^k = n) \beta_i^k(n)}.$$

Because of the high correlation between the v_i^k and T_i^k, additional information can be provided by the HHMM. Therefore, the properties of each wavelet coefficient can be described better by the HHMM, especially at higher resolution scales.

3. CONTEXT MODEL

For each tree T_i^k in wavelet domain, the context vector v_i^k can be defined as a set of trees which are locally correlated to T_i^k. An example of the inter-orientation context is shown at Fig. 3. At LH subband, the inter-orientation context $v_i^k(LH)$ of the tree $T_i^k(LH)$ is composed of two wavelet trees, $T_i^k(HH)$ and $T_i^k(HL)$.

According to different visual attributes, four types of local contexts based on different positions, orientations, and



(a) (b) (c) (d)

Fig. 4. Context vectors based on different visual attributes (black: current wavelet tree; gray: context trees). (a) Inter-frequency. (b) Inter-orientation. (c) Intra-subband. (d) Inter-frequency/Inter-orientation.

scales are applied: *inter-frequency* (v_{i1}), *inter-orientation* (v_{i2}), *intra-subband* (v_{i3}) and *inter-frequency/inter-orientation* (v_{i4}) (see Fig. 4)[8][9]. Thus, for a wavelet tree T_i^k (see the black blocks in Fig. 4), We have context vector $v_i^k = \{v_{i1}, v_{i2}, v_{i3}, v_{i4}\}$ and $v_{ij} = \{T_{q(i)}\}, j = 1, 2, 3, 4$, where $q(i)$ represents a set of indexes of context trees of T_i^k (see the gray blocks in Fig. 4). The following EM algorithm will be used to compute the $p(c_i^k|T_i^k, v_i^k)$ and $p(c_i^k)$,

E Step:

$$p(c_i^k|T_i^k, v_i^k) = \frac{p(c_i^k) f(v_i^k|c_i^k) f(T_i^k|c_i^k)}{\sum_{n=1}^{M} p(c_i^k = n) f(v_i^k|c_i^k = n) f(T_i^k|c_i^k = n)}.$$

M Step:

$$p(c_i^k) = \frac{1}{2^k} \sum p(c_i^k|T_i^k, v_i^k).$$

Assuming that T_i^k is independent, given its hidden class label c_i^k, a continuous-valued v_i^k can be modeled as an M-component Gaussian mixture,

$$f(v_i^k|c_i^k) = \sum_{j=1}^{4} \omega f(v_{ij}^k|c_i^k),$$

$$f(v_{ij}^k|c_i^k) = f(T_{q(i)}|c_i^k) = \prod_{r \in q(i), r \neq i} f(T_r|c_i^k),$$

where ω is the weight factor to indicate the significance of different context relationships.

4. UNSUPERVISED TEXTURE SEGMENTATION

In a spatial domain, a dyadic block d_i^k is associated with three wavelet coefficient trees $T_i^k(b), b \in \{HL, HH, LH\}$. Thus,

$$p(c_i^k|d_i^k) = p(c_i^k|T_i^k(HL), T_i^k(HH), T_i^k(LH)),$$

where $p(c_i^k|d_i^k)$ is the probability that class label c_i^k is present for a block d_i^k. The probability $p(c_i^k|d_i^k)$ can be computed by the Bayes rule. Assume that three subbands are independent, we have

$$\begin{aligned} p(c_i^k|d_i^k) &= p(c_i^k|T_i^k(HL), T_i^k(HH), T_i^k(LH)) \\ &= \kappa \frac{\prod_{b \in \{HL,HH,LH\}} p(c_i^k|T_i^k(b))}{p^2(c_i^k)}, \end{aligned}$$

where $\kappa = \frac{\prod_{b \in \{HL,HH,LH\}} p(T_i^k(b))}{p(T_i^k(HL), T_i^k(HH), T_i^k(LH))}$ is not dependent on c_i^k. The following method can be applied to calculate the probability $p(c_i^k|d_i^k)$,

E Step:

$$p(c_i^k|d_i^k) = \frac{\prod_{b \in \{HL,HH,LH\}} p(c_i^k|T_i^k(b))/p^2(c_i^k)}{\sum_{c_i^k=1}^{M} \prod_{b \in \{HL,HH,LH\}} p(c_i^k|T_i^k(b))/p^2(c_i^k)}.$$

M Step:

$$p(c_i^k) = \frac{1}{2^k} \sum p(c_i^k|d_i^k).$$

The final segmentation of each dyadic block d_i^k into class c_i^k can be described as:

$$c_i^k = \text{argmax}_{c_i^k} p(c_i^k = m|d_i^k).$$

5. EXPERIMENTAL RESULTS

Experiments are conducted for the HHMM using several synthetic and natural images. The results are also compared to the HMT model [6]. In all examples, the Haar wavelet is employed for simplicity. Moreover, the resolution threshold is set as $l = L - 3$, where L is the highest resolution scale.

Fig. 5(a) are two synthetic 256×256 images which the ground truth are known. A two Brodatz textures image and a three Brodatz textures image are used as examples. Fig. 5(b)(d) show results at higher resolution scales (4×4 and 2×2 scales) by the HMT model, and Fig. 5(c)(e) are segment results by the HHMM model. The segmentation results are obtained without any postprocessing. The result comparisons are also given by the Table 1, in which we show the average percentage segmentation error. The performance is improved by the HHMM model, especially at higher resolution scales.

Fig. 6(a) show two real images. One is a real aerial image downloaded from the USC Image Database, and another is a SAR image. The 2×2 scale results are provided by Fig. 6(b)(c) to show the performance of the HHMM. Compared to the HMT model, the HHMM model presents a more homogeneous segmentation result with fewer errors.

6. CONCLUSIONS

In this paper, a new hierarchical hidden Markov model was proposed to capture both interscale and intrascale depen-

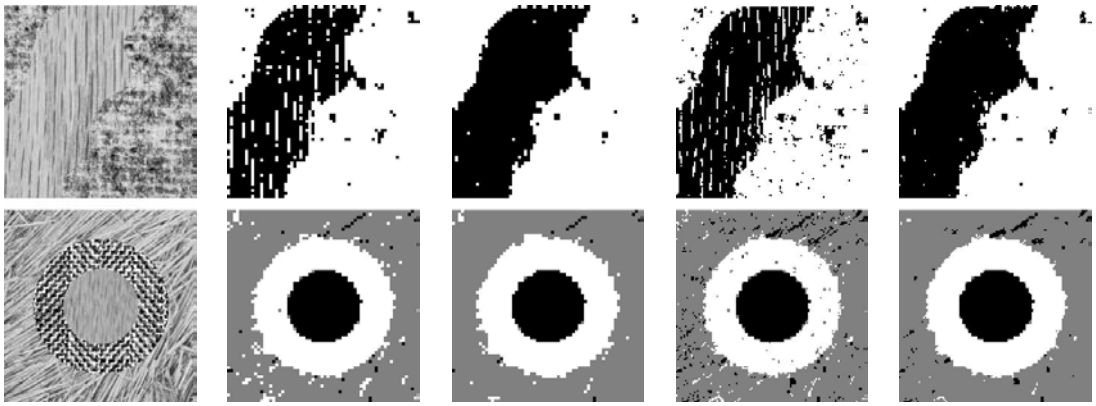


Fig. 5. (a) Original synthetic images. (b) and (d) Segmentation results at 4×4 and 2×2 scales by the HMT. (c) and (e) Segmentation results at 4×4 and 2×2 scales by the HHMM.

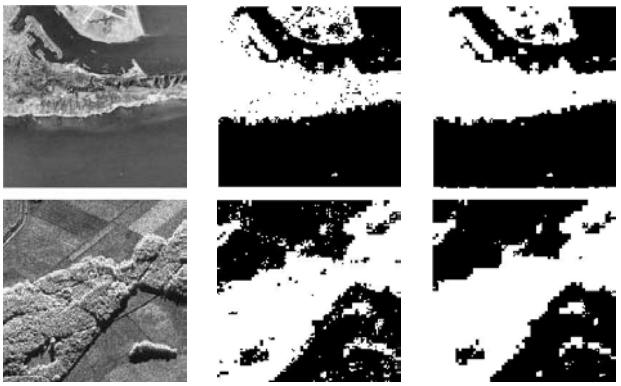


Fig. 6. (a) Original real-world images. (b) Segmentation results by the HMT. (c) Segmentation results by the HHMM.

dencies. New contexts that include different positions, orientations, and scales are introduced. The experimental results demonstrate the performance of the HHMM model. For texture segmentation, the HHMM model outperforms the HMT model, especially at higher resolution scales.

Table 1. Average error percentage in segmentation

	HMT		HHMM	
block	2-texture	3-texture	2-texture	3-texture
8×8	4.28%	5.65%	3.67%	5.63%
4×4	8.87%	6.58%	3.49%	5.04%
2×2	9.05%	7.23%	3.61%	4.84%

7. REFERENCES

[1] C. A. Bouman and M. Shapiro, "A multiscale random field model for bayesian image segmentation," *IEEE Trans. on Image Proc.*, vol. 3, no. 2, pp. 162–177, Mar. 1994.

[2] H. Choi and R. G. Baraniuk, "Multiscale texture segmentation using wavelet-domain hidden markov models," in *Proc. 32nd Asilomar Conference*, Nov. 1998.

[3] G. Fan and X. G. Xia, "Multiscale texture segmentation using hybrid contextual labeling tree," in *Proc. of the IEEE Int. Conf. on Image Proc.*, Vancouver, Canada, Sep. 2000.

[4] J. Li, R. M. Gray, and R. A. Olshen, "Multiresolution image classification by hierarchical modeling with two-dimensional hidden markov models," *IEEE Trans. on Info. Theory*, vol. 46, no. 5, pp. 1826–1841, 2000.

[5] J. Lu and L. Carin, "Hmm-based multiresolution image segmentation," in *Proc. ICASSP 2002*, Orlando, Florida, May. 2002.

[6] M. S. Crouse, R. D. Nowak, and R. G. Baraniuk, "Wavelet-based signal processing using hidden markov models," *IEEE Trans. on Signal Proc.*, vol. 46, no. 4, pp. 886–902, Apr. 1998.

[7] A. K. Jain and K. Karu, "Learning Texture Discrimination Masks," *IEEE Trans. on PAMI*, vol. 18, pp. 195–205, 1996.

[8] Henry Schneiderman, "A statistical approach to 3d object detection applied to faces and cars," *Ph.D Thesis, Carnegie Mellon University*, May. 2000.

[9] Z. Ye and C.-C. Lu, "A complex wavelet domain markov model for image denoising," in *Proc. of the IEEE Int. Conf. on Image Proc.*, Barcelona, Spain, Sep. 2003.