Multiscale texture segmentation of dip-cube slices using wavelet-domain hidden Markov trees

Ivan Magrin-Chagnolleau, Hyeokho Choi, Rutger van Spaendonck, Philippe Steeghs, and Richard G. Baraniuk
Rice University, Houston, Texas

Summary

Wavelet-domain hidden Markov models (HMMs) are powerful tools for modeling the statistical properties of wavelet transforms. By characterizing the joint statistics of wavelet coefficients, HMMs efficiently capture the characteristics of many real-world signals. When applied to images, the model can characterize the joint statistics between pixels, providing a very good classifier for textures. Utilizing the inherent tree structure of wavelet-domain HMMs, classification of textures at various scales is possible, furnishing a natural tool for multiscale texture segmentation. In this paper, we introduce a new multiscale texture segmentation algorithm based on wavelet-domain hidden Markov trees (HMTs). We apply this new technique to the segmentation of dip-cube time slices.

Introduction

The goal of an image segmentation algorithm is to assign a class label to each pixel of the image based on the properties of the pixel and its relation with neighborhood pixels. To capture the properties of each region of the image to be segmented, both the large and small scale behaviors should be utilized to properly segment both large, homogeneous regions and detailed boundary regions. In view of the utilization of properties in different scales, it is natural to approach the segmentation problem using multiscale analysis. There have been efforts to segment using multiscale autoregressive models (Basseville et al., 1992; Fosgate et al., 1997). In Bouman et al. (1994), a multiscale random field model was used to statistically analyze the multiscale structure of the segmentation labels. In this paper, we propose a multiscale texture segmentation algorithm based on the wavelet transform of the image.

For many classes of signals and images, wavelets provide a compact[1] and approximately decorrelated signal representation (Daubechies, 1992). For statistical applications such as detection and estimation, an accurate joint probability model for the wavelet coefficients of the signals plays an important role. In view of the compaction and decorrelation properties of wavelet transforms, modeling the wavelet coefficients instead of direct modeling of the signal may be more efficient and naturally provide a multiscale structure of the involved signal processing.

Recently, a wavelet-domain hidden Markov model (HMM) was proposed as a powerful tool for modeling the probability structure of wavelet transforms (Crouse et al., 1998; Crouse and Baraniuk, 1997). By modeling each wavelet coefficient as a Gaussian mixture and capturing the dependencies between the coefficients as hidden state transitions, HMMs provide a natural setting for exploiting the structure inherent in real-world signals and images for signal detection and classification.

In this paper, we employ the tree structure of wavelet-domain HMMs for multiscale signal classification (Choi and Baraniuk, 1999). By computing the likelihoods of each subblocks of the image in different scales, a "raw" segmentation is obtained. Segmentations in coarser scales are more reliable for large, homogeneous regions, while finer scale segmentations are more appropriate around boundaries between different textures. By combining these raw segmentations in different scales, we can obtain a robust and accurate segmentation result. We accomplish this interscale fusion by setting up a tree-structured model for segmentation results in different scales. By computing the likelihoods of each subblock in fine scale conditioned on the next coarser scale segmentation result, we can make a Bayesian multiscale classification. Unlike the approach in Bouman et al. (1994), we use the concept of "context" to capture the dependencies between segmentation results across different scales, resulting in an algorithm that is more flexible and easier to implement.

Wavelet transform

The discrete wavelet transform (DWT) represents a 1-D signal $z(t)$ in terms of shifted versions of a lowpass scaling function $\phi(t)$ and shifted and dilated versions of a prototype bandpass wavelet function $\psi(t)$ (Daubechies, 1992). For special choices of $\phi(t)$ and $\psi(t)$, the functions $\psi_{j,k}(t) = 2^{-j/2}\psi\left(2^{-j}t-k\right)$, $\phi_{j,k}(t) = 2^{-j}\ \phi\left(2^{-j}t-k\right)$, for j and $k \in \mathbb{Z}$, form an orthonormal basis, and we have the representation:

$$z(t) \ = \ \sum_k u_{j_0,k}\,\phi_{j_0,k}(t) \ + \ \sum_{j=j_0}^{\infty}\sum_k w_{j,k}\,\psi_{j,k}(t)$$

with $u_{j,k} = \int z(t)\,\phi^*_{j,k}(t)\,dt$ and $w_{j,k} = \int z(t)\,\psi^*_{j,k}(t)\,dt$.

To keep the notation manageable in the sequel, we will adopt an abstract index scheme for the DWT coefficients: $u_{j,k} \to u_i$, $w_{j,k} \to w_i$.

For multidimensional signals (e.g. images), the DWT can be easily generalized by alternately applying the 1-D DWT to each coordinate in order to get a 2-D wavelet transform.

Hidden Markov Models

The DWTs of many real-world signals tend to be sparse, with just a few non-zero coefficients containing most of the signal energy. Hence, the probability density function

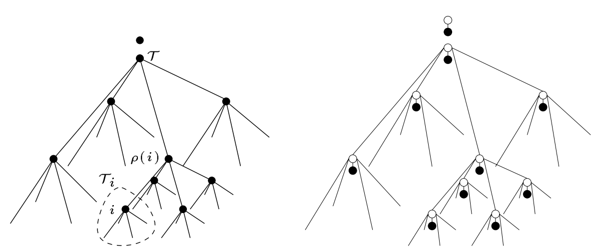
[1] lots of small coefficients and only a few large coefficients

(pdf) $f_{W_i}(w_i)$ of each wavelet coefficient is well approximated by a *Gaussian mixture* with two components, one corresponding to the low values of the energy, one to the high values. To each wavelet coefficient W_i, we associate a discrete hidden state S_i that takes on values $m = 1$ or 2 with probability mass function $p_{S_i}(m)$. Conditioned on $S_i = m$, W_i is Gaussian with mean $\mu_{i,m}$ and variance $\sigma^2_{i,m}$. Thus, its overall pdf is given by

$$f_{W_i}(w_i) \;=\; \sum_{m=1}^{2} p_{S_i}(m)\, f_{W_i|S_i}(w_i|S_i = m)$$

Wavelet-domain HMMs are multidimensional mixture models in which the hidden states have a Markov dependency structure (Crouse et al., 1998). The idea is to capture the dependencies in the wavelet coefficients through the transition probabilities between the states.

For example, the hidden Markov tree (HMT) model places a tree structure on the hidden states to capture wavelet dependencies across scale. The HMT model is specified via the mixture parameters $(\mu_{i,m}, \sigma^2_{i,m})$ and transition probabilities $p_{S_i|S_{\rho(i)}}(m|n)$, where $\rho(i)$ denotes the parent of node i. These parameters can be grouped into a model parameter vector $\mathbf{\Theta}$. The wavelet-domain HMT is trained to capture the wavelet-domain properties of the signals of interest using the iterative expectation maximization (EM) algorithm.



(a) (b)

Fig. 1: (a) Quadtree of 2-D wavelet coefficients for each subband, (b) 2-D wavelet-domain HMT model. Black nodes represent wavelet coefficients; white nodes represent hidden state variables.

For 2-D images, quadtrees are used for HMT models (see Fig. 1). In addition, because we have three different subbands in 2-D DWT, we have a distinct quadtree for each subband. The subbands are treated as independent although they are not, but it works well in practice.

Multiscale Classification

For simplicity, we assume that Haar wavelet is used for the underlying wavelet transform. Then, each dyadic subblock of pixels can be associated with a wavelet coefficient, i.e., each node of the quadtree (Fig. 1) corresponds to a dyadic subblock of the entire image. We define $\mathcal{T}_i$ to be the subtree of observed wavelet coefficients with root at node i, so that the subtree $\mathcal{T}_i$ contains coefficient w_i and all of its descendants (Fig. 1). Due to the quadtree structure, each $\mathcal{T}_i$ corresponds to a subblock of the entire image.

Given a HMT model $\mathbf{\Theta}$, for each subtree $\mathcal{T}_i$, we compute the conditional likelihood $\beta_i(m) = f(\mathcal{T}_i|S_i = m, \mathbf{\Theta})$ and the conditional probability $p(S_i = m|\mathbf{w}, \mathbf{\Theta})$, where $\mathbf{w} = \{w_i\}$ is the collection of all wavelet coefficients. Then, the conditional likelihood $f(\mathcal{T}_i|\mathbf{\Theta})$ can be computed as:

$$f(\mathcal{T}_i|\mathbf{\Theta}) = \sum_{m=1}^{2} \beta_i(m) p(S_i = m|\mathbf{w}, \mathbf{\Theta})$$

Suppose we have HMT models for two different textures, T_1 and T_2, whose parameters are respectively $\mathbf{\Theta}_{T_1}$ and $\mathbf{\Theta}_{T_2}$. The classification in each scale is accomplished by comparing the likelihood functions $f(\mathcal{T}_i|\mathbf{\Theta}_{T_1})$ and $f(\mathcal{T}_i|\mathbf{\Theta}_{T_2})$ for every node i. For each i, the subblock of image corresponding to the subtree $\mathcal{T}_i$ can be classified, producing a raw segmentation of the image up to 2×2 pixels image blocks. We call this segmentation a "raw" segmentation because we do not use any relationship between segmentations across different scales.

We expect this "raw" segmentation to be more reliable at coarse scales, because classification of larger image blocks is more reliable due to its richness in statistical information. However, at coarse scale, the boundaries between different textures cannot be captured accurately. At fine scales, the segmentation is less reliable, while the boundary is better captured.

When we compute the likelihoods $f(\mathcal{T}_i|\mathbf{\Theta})$, we ignore the scaling coefficients in the wavelet transform. This implies that we do not rely on the local brightness levels of the image for segmentation. Only the joint statistics between pixels are used as the discriminating feature. This is why only up to 2×2 blocks can be classified. The brightness of each pixel will be used to obtain the pixel level segmentation at the final stage of interscale fusion algorithm explained next.

Interscale Fusion

By combining the segmentation results from different scales, we can obtain better segmentation of the image. Because the "raw" segmentation is more reliable in coarse scales, the segmentation result in coarse scales can be used as extra information for segmentation in finer scales. When a subblock of an image at a given scale is classified as a certain texture, it is very likely that the corresponding subblocks in the next finer scale should be also classified as the same texture. This coarse-to-fine dependencies can be modeled as a 2-D tree (labeling tree), where each node corresponds to a subblock of image. The interscale dependencies are described as connecting branches between nodes. However, as the number of branches increases, analysis of this tree becomes extremely difficult.

Here, we propose a context-based method (Crouse and Baraniuk, 1997) to capture the interscale dependencies, which is easy to manipulate even when the number of incorporated neighborhood nodes is large. Let i be a node of the 2-D labeling tree. Denote the class label for the subblock of image corresponding to i as C_i, which takes values $1, \cdots, N$, where N is the number of different textures to be classified. Let v_i denote a context vector that describes the segmentation result in the next coarser scale. That is, v_i represents the segmentation results in the previous coarse scale in the neighborhood of the region centered at i.

The choice of v_i is not restricted at all, as long as it describes the results of the coarse scale segmentation well. For example, in the example of two textures T_1 and T_2, v_i can be chosen to be the number of neighborhood blocks in the previous coarse scale that were classified as T_1, with appropriately pre-specified neighborhood region for i.

In some sense, the context v_i indicates whether the part of image corresponding to node i falls in the interior of a homogeneous texture, or whether it is near a boundary between two different textures. Let $J(i)$ denote the scale of the node i, and denote the entire image (or equivalently the wavelet transform) as $\mathcal{T}$. Let k be the scale under consideration. And, let $\mathbf{v}^k$ denote the collection of all contexts at scale k. The contextual Bayes classification starts by estimating the probabilities $p_{C_i|V_i}(n|v_i)$ to maximize the likelihood of the entire image computed in scale k, given as

$$f(\mathcal{T}|\mathbf{v}^k) = \prod_{i \text{ s.t. } J(i)=k} \sum_{n=1}^{N} p_{C_i|V_i}(n|v_i) f(\mathcal{T}_i|C_i = n)$$

where $f(\mathcal{T}_i|C_i = n)$'s are given by the EM algorithm applied to HMTs for each candidate texture. The values of $p_{C_i|V_i}(n|v_i)$ are obtained by assuming that they are constant in each scale and averaging over all the nodes in that scale. In practice, we do not specify $p_{C_i|V_i}(n|v_i)$ directly, but rather specify $p_{V_i|C_i}(v_i|n)$ and apply Bayes rule. Thus, the actual probabilities that should be computed are $\epsilon_{i,n} = p_{C_i}(n)$ and $\alpha_{i,v,n} = p_{V_i|C_i}(v_i|n)$. Once $\epsilon_{i,n}$ and $\alpha_{i,v,n}$ are computed, the context-based Bayes classification is performed by computing

$$p_{C_i|V_i,\mathcal{T}}(n|v_i, \mathcal{T}) = \frac{\epsilon_{i,n}\alpha_{i,v,n} f(\mathcal{T}_i|C_i = n)}{\sum_{n=1}^{N} \epsilon_{i,n}\alpha_{i,v,n} f(\mathcal{T}_i|C_i = n)}$$

Then, we label node i so that $p_{C_i|V_i,\mathcal{T}}(n|v_i, \mathcal{T})$ is maximized. In practice, the computation of $\epsilon_{i,n}$, $\alpha_{i,v,n}$, and $p_{C_i|V_i,\mathcal{T}}$ is performed simultaneously by an iterative method similar to the EM algorithm. To obtain pixel level segmentation, the probability distribution of a single pixel for each of the textures is necessary. This distribution can be modeled from the histogram of pixel intensities for each texture. For textures, it can be well approximated as a Gaussian mixture model. The problem of pixel-level modeling and segmentation is discussed in more detail in Choi and Baraniuk (1999).

Application to dip-cube slices

Fig. 2 and Fig. 5 represent two time slices of a dip-cube obtained by using a 3-D local Radon power spectra calculation on a 3-D data set from the Ameland survey, located in the Dutch sector of the North Sea (Steeghs et al., 1998). We have selected two 64 × 64 blocks from each of the two original images to train the HMTs, each block corresponding to two different kinds of texture. The training was performed with intra-scale tying (Crouse et al., 1998) to avoid overfitting the models. Fig. 3 and Fig. 6 show the raw segmentation results without the interscale processing. Fig. 4 and Fig. 7 show the segmentation results after using the interscale fusion algorithm. As shown on the figures, we obtain excellent results for this two texture segmentation example, particularly on the second image, and we can see more clearly the channel on Fig. 4, and even more clearly the Salt structure on Fig. 7. A better choice of the training images may improve the results for the first image.

Conclusion

We have introduced a new framework for texture segmentation based on wavelet-domain hidden Markov models. By concisely modeling the statistical behavior of textures at multiple scales and by combining segmentations obtained at these scales, the algorithm produces a robust and accurate segmentation of texture images. We have applied this new segmentation technique to dip-cube time slices in order to help the interpretation of the slices. The next step will be to produce a 3-D texture segmentation of dip-cubes, including more than two classes.

Aknowledgement

This work has been sponsored by the Rice Consortium for Computational Seismic/Signal Interpretation (www.dsp.rice.edu/ccsi), by the National Science Foundation, grant no. MIP–9457438, and by DARPA/AFOSR, grant no. F49620-97-1-0513.

References

Basseville, M., et al., March 1992, Modeling and estimation of multiresolution stochastic processes: IEEE Transactions on Information Theory, **38**, no. 2, 766–784.

Bouman, C., and Shapiro, M., March 1994, A multiscale random field model for Bayesian image segmentation: IEEE Transactions on Image Processing, **3**, no. 2, 162–177.

Choi, H., and Baraniuk, R. G., 1999, Multiscale texture segmentation using wavelet-domain hidden Markov models: Submitted to IEEE Transactions on Image Processing.

Crouse, M., and Baraniuk, R., November 1997, Contextual hidden Markov models for wavelet-domain signal processing: Proceedings of the 31st Asilomar Conference.

Crouse, M., Nowak, R., and Baraniuk, R., April 1998, Wavelet-based statistical signal processing using hidden Markov models: IEEE Transactions on Signal Processing, **46**, no. 4, 886–902.

Daubechies, I., 1992, Ten lectures on wavelets: SIAM.

Fosgate, C., et al., January 1997, Multiscale segmentation and anomaly enhancement of SAR imagery: IEEE Transactions on Image Processing, **6**, no. 1, 7–20.

Steeghs, P., Fokkema, J. T., and Diephuis, G., 1998, 3-D local Radon power spectra for seismic attribute extraction: Proceedings of the 68th SEG Meeting.

Multiscale texture segmentation of dip-cube slices

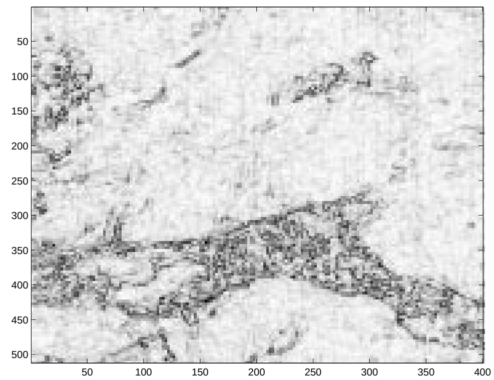


Fig. 2: *First time slice of the dip-cube.*

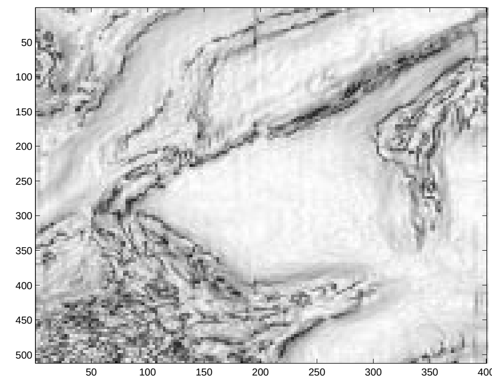


Fig. 5: *Second time slice of the dip-cube.*

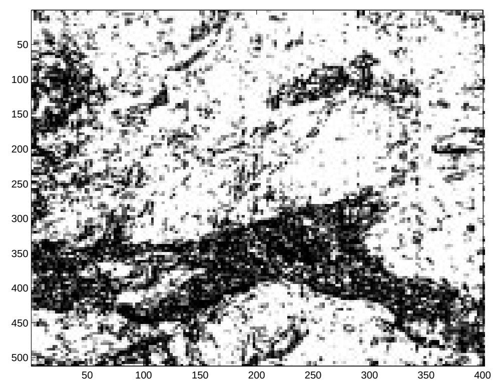


Fig. 3: *Raw segmentation without the interscale processing for the first time slice.*

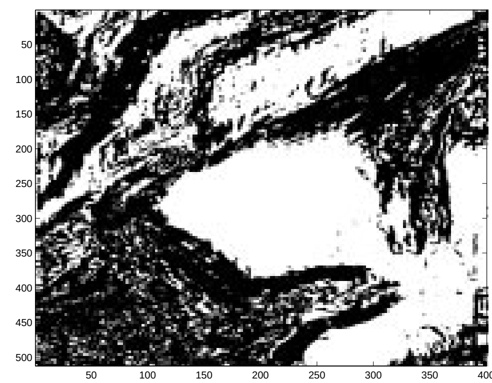


Fig. 6: *Raw segmentation without the interscale processing for the second time slice.*

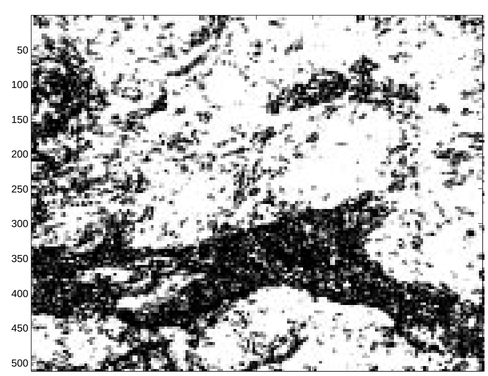


Fig. 4: *Segmentation after using the interscale fusion algorithm for the first time slice.*

Fig. 7: *Segmentation after using the interscale fusion algorithm for the second time slice.*