KERNEL MATCHING PURSUITS PRIORITIZATION OF WAVELET COEFFICIENTS FOR SPIHT IMAGE CODING

Shaorong Chang and Lawrence Carin

Department of Electrical and Computer Engineering
Duke University
Durham, NC 27708-0291

ABSTRACT

The set partitioning in hierarchical trees (SPIHT), an efficient wavelet-based progressive image-compression scheme, is oriented to minimize the mean-squared error (MSE) between the original and decoded imagery. In this paper, we use the kernel matching pursuits (KMP) method to estimate the importance of each wavelet sub-band for distinguishing between different textures segmented by an HMT mixture model. Before the SPIHT coding, we weight the wavelet coefficients, with the goal of achieving improved image-classification results at low bit rates. A modified SPIHT algorithm is proposed to improve the coding efficiency. The performances of the original SPIHT and the modified SPIHT algorithms are compared.

1. INTRODUCTION

It is often useful to implement compression algorithms that account for the ultimate classification task associated with the decoded imagery, such as in detecting biological abnormalities in compressed medical images and in compressing aerial imagery for remote-sensing applications. The goal is to compress the image efficiently while accounting for the fact that the decompressed image will be employed in a classification task. The overall encoding scheme is shown in Fig. 1. We here focus on wavelet-based image compression algorithms, since such now represent the state of the art and are used in practical algorithms.

Hidden Markov trees (HMT) in the wavelet domain capture the statistical dependence of wavelet coefficients well, providing a reliable segmentation of image textures [3]. We propose an unsupervised image segmentation method using an HMT mixture model, the parameter estimation problem of which is solved by the EM algorithm, a widely applied technique in computational pattern recognition [4]. The hidden posterior probability distribution across the mixture components results in the image segmentation. The segmentation is performed autonomously at the encoder, and the goal is to prioritize for compression those wavelet coefficients that play important roles in this segmentation stage, despite the fact that these coefficients may be of small amplitude (and hence given low priority by conventional wavelet encoders).

Set partitioning in hierarchical trees (SPIHT) is an effective embedded wavelet-based image-coding algorithm [1]. It seeks to minimize the mean-squared error (MSE) at any bit rate by the progressive transmission of the partially ordered bit planes and the effective exploration of the self-similarity across the wavelet trees. However, the MSE-based measure is not in general well correlated with the image-recognition quality, especially at low bit rates (small wavelet coefficients, which may be important for classification, are given low priority and therefore a coarse representation by conventional MSE-based encoders). To optimize the image visual quality, perceptually weighted quantization demonstrates a significant improvement in visual quality [2]. Similarly, we estimate the recognition importance of each wavelet subband by the kernel matching pursuits (KMP) method. By weighting the wavelet coefficients properly, we order the transmission bits not only by the magnitudes of the wavelet coefficients but also by their contributions to image recognition. An efficient algorithm is developed especially for the weighted wavelet coefficients, to improve the coding efficiency.

The remainder of the paper is organized as follows. In Sec. 2 we give the definition of the HMT mixture model along with the expectation-maximization (EM) training algorithm. We consider in Sec. 3 an additive regression model to estimate the importance of the wavelet coefficients for texture recognition. In Sec. 4 the modified SPIHT coding scheme is discussed. Typical results of the algorithm are presented in Sec. 5, with conclusion in Sec. 6.

2. HMT MIXTURE MODEL

The wavelet transform is popular in signal processing and image compression, it effectively representing both the local features and global characteristics [7]. The persistence statistical property of the wavelet coefficients - that is, if the magnitude of the wavelet coefficient is large its children are likely to have large magnitudes - is captured by a hidden Markov tree (HMT) model. If the image contains more than one texture, we use HMT mixtures to model the statistical characteristics. The mixture-density parameter estimation is solved by the EM algorithm [4]. The probabilistic model is

0-7803-8484-9/04/$20.00 ©2004 IEEE

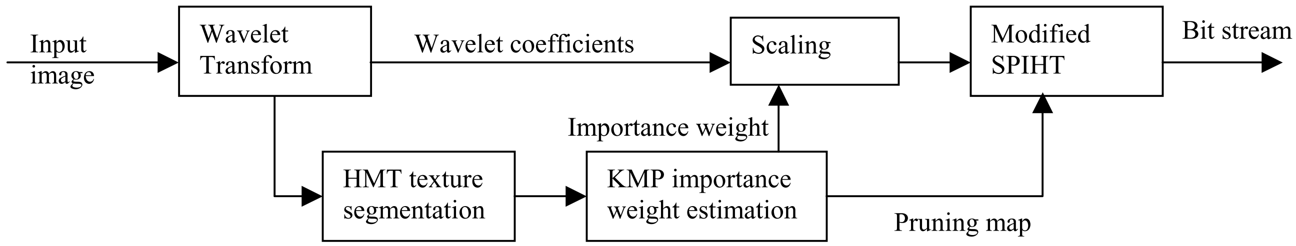

Figure 1. The Overall coding system

$$P(\mathbf{w} \mid \Theta) = \sum_{i=1}^{M} \alpha_i P(\mathbf{w} \mid HMT_i) \quad (2.1)$$

where $\alpha_i \geq 0, \sum_{i=1}^{M} \alpha_i = 1$ are the mixing coefficients of the M textures, which can be interpreted as the prior probabilities, and $\mathbf{w}$ are the wavelet coefficients of a wavelet tree. $P(\mathbf{w} \mid HMT_i)$ is a density function parameterized via the HMT for texture *i*.

We estimate the probability that the t^{th} wavelet tree is generated by texture *j* using Bayes's rule

$$p_j^{(k)}(t) = \frac{\alpha_j P(\mathbf{w}_t \mid HMT_j)}{\sum_{i=1}^{M} \alpha_i P(\mathbf{w}_t \mid HMT_i)} \quad (2.2)$$

and update the mixing coefficients as

$$\alpha_j^{(k+1)} = \frac{\sum_{t=1}^{N} p_j^{(k)}(t)}{N} \quad (2.3)$$

The parameters of each HMT are updated by the EM algorithm in [3] under the sample probability $p_j^{(k)}(t)$, t=1,2,…N. The samples that are associated with texture *j* with high likelihood make a greater contribution to the parameters of that texture component. We iteratively estimate the probabilities and update the model parameters until the model converges to a local optimal solution.

Based on the component probability, we segment the image by using maximum a posteriori (MAP) estimator, that is,

$$c(\mathbf{w}) = \arg\max_i P(\mathbf{w} \mid HMT_i)\alpha \quad (2.4)$$

The parameter M, the number of the textures in the image, is automatically selected via an information-theoretic model-selection method called the minimum description length (MDL) principle derived by Rissanen [6], considering both the model complexity and accuracy.

3. IMPORTANCE WEIGHT ESTIMATION

Transmitting wavelet coefficients prioritized based on their amplitude strength may not optimize the recognition rate (here characterized by distinguishing textures) based on the decoded imagery, so we introduce a rescaling process using estimated importance weights. The purpose of the importance weights is to help the encoder to order the output bit stream with consideration of the ultimate recognition task. The wavelet features that are highly correlated with the texture class labels (determined automatically, as discussed in Sec. 2) are important to distinguish the textures. These features should have higher priorities and be transmitted earlier. However, the wavelet features that are more important to segmentation may have small amplitudes and small variances compared to the wavelet features of less importance to segmentation. The rescaling process alleviates this phenomenon.

The kernel matching pursuits (KMP) [5,8] algorithm is employed to prioritize the importance of the wavelet coefficients for accurate segmentation. KMP is a generative learning algorithm and offers an efficient iterative method to select the kernel functions and estimate the weights. The sparseness of the algorithm not only implies good generalization but also allows us to prune a large number of wavelet coefficients, retaining those that are most important for the classification stage.

Assume we have a vector of wavelet coefficients $\mathbf{w}$ from a wavelet tree, and we wish to predict the texture label y (e.g., for two textures y is binary). The classifier is designed by using KMP [5], which is formulated as

$$f(\mathbf{w}) = \gamma_0 + \sum_{i=1}^{m} K(c_i, \mathbf{w})\gamma_i = \gamma_0 + \sum_{i=1}^{m} \left| w_{c_i} \right| \gamma_i \quad (3.1)$$

where $K(c_i, \mathbf{w})$ is a kernel function, selecting wavelet coefficient c_i from $\mathbf{w}$, γ_i is the corresponding kernel weight, reflecting the importance of c_i to the classifier and m is the number of wavelet coefficients selected by KMP, indicating the complexity of the classifier. In the generative learning progress, we select the kernel functions (wavelet coefficients) one by one to reduce the cost function defined as follows:

$$e = \tfrac{1}{N} \sum_{i=1}^{N} (\lambda(y_i - f(\hat{\mathbf{w}}_i))^2 + \left\| \mathbf{w}_i - \hat{\mathbf{w}}_i \right\|^2) \quad (3.2)$$

where $\hat{\mathbf{w}}_i$ is the quantized wavelet coefficients associated with the i^{th} wavelet tree, (m of these wavelet coefficients are chosen by KMP to be quantized to a specified precision, dictated by the weighting and the number of SPIHT iterations, and the other wavelet coefficients are quantized to zero). The scalar λ is a Lagrangian multiplier that constitutes a compromise between the quantization error and the regression error. The reason for using the quantization (MSE) error is that we not only want to select the features that facilitate the segmentation but also maintain a relative wavelet strength within the wavelet tree that is amenable to SPIHT. The MSE term in (3.2) penalizes choosing the wavelet features deep in the tree, which tend to have lower variances and thus smaller MSE importance.

The importance weight of the i^{th} selected wavelet coefficients is defined as:

$$\beta_i = \begin{cases} 1 + \lambda \bullet |\gamma_i| & i \le m \\ 0 & otherwise \end{cases} \qquad (3.3)$$

If λ is small, all selected wavelet coefficients are approximately equally important. If λ is large, the importance weight of the selected wavelet coefficient is approximately proportional to the kernel weight.

4. MODIFIED SPIHT

Between the wavelet transform and the SPIHT coding, we introduce a rescaling process, assigning finer quantization step sizes to the wavelet features with larger importance weights. After weighting the wavelet coefficients as designed in the previous section, the SPIHT algorithm may be less efficient for implementation of the spatial orientation zero tree structure (SPIHT was originally designed for a MSE cost function alone; the new weighting of importance to potentially small wavelet coefficients may undermine assumptions in designing SPIHT). We modify the SPIHT algorithm to improve the coding efficiency.

4.1. Review of the SPIHT algorithm

The effectiveness of the SPIHT algorithm originates from the efficient subset partitioning and the compact form of the significance information. The SPIHT algorithm defines spatial orientation trees, sets of coordinates and the recursive set partitioning rules [1]. The algorithm is composed of two passes: a sorting pass and a refinement pass. It is implemented by alternately scanning three ordered lists, called list of insignificant sets (LIS), list of insignificant pixels (LIP) and list of significant pixels (LSP), among which LIS and LIP represent the individual and sets of coordinates, respectively, whose wavelet coefficients are less than a threshold defined. During the sorting pass the significances of LIP and LIS are tested, followed by removal and set splitting operations to maintain the insignificance property of the lists. The LSP contains the coordinates of the significant pixels that are scanned in the refinement pass.

4.2. Modified SPIHT algorithm

The weighted wavelet coefficients, which have a larger dynamic range than the original ones, make the SPIHT algorithm inefficient, because this results in more scans of the wavelet coefficients. Consequently, a modified SPIHT is needed to overcome this problem.

First, we note that the wavelet coefficients with 0 importance weights are blocked out. As a prior knowledge, we need not waste the bit budget on those wavelet coefficients. Therefore, after each sorting pass, we delete the corresponding coordinates from LIP and the sets whose components all have 0 importance weights from LIS. This is done at both the encoder and the decoder.

Second, it is not necessary to encode the weighted coefficients using too many bits (for the small-amplitude coefficients, only enough bits are required to achieve the classification task, since they make a small contribution to the MSE). We can define an appropriate bit limit for each weight coefficient. When the wavelet coefficients reach the upper limits of the refinement bits, we can skip them in the refinement pass.

5. EXPERIMENTAL RESULTS

We select a high altitude optical aerial image of size 256x256 pixels from the USC-SIPI image database, to demonstrate the application of the algorithm proposed. This is an 8 bits/pixel black and white image with two distinct textures: a rural area and an urban area, as shown in Fig. 2.

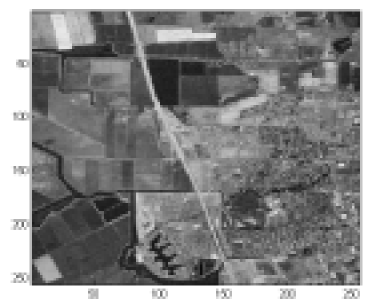

Figure 2. High altitude aerial image

After a three-level wavelet decomposition, we obtain wavelet trees **w** of size 64. We train the parameters of the 2-HMT-mixture model by the EM algorithm described in Sec. 2. In Fig. 3 we show the posterior probabilities of one texture component. We see that most of the city areas are brighter, which represents higher likelihood and the rural areas are darker indicating lower likelihood. The segmentation is consistent with human visual recognition. The task now is to prioritize those wavelet coefficients of importance for achieving this segmentation, even if they are of small amplitudes.

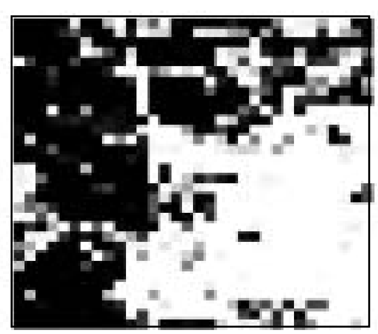

Figure 3. The posterior probability of the urban texture

The HMT mixture model provides a meaningful segmentation of a multi-texture image. However, it is too complicated to analyze the contribution of each wavelet sub-band to separate the textures. We introduce the KMP method to handle this problem. We iteratively select the wavelet coefficients that reduce the cost function defined in (3.2) and estimate the wavelet-coefficient weights γ_i jointly with all the selected coefficients. We choose the first 30 most important wavelet coefficients and calculate the importance weight of each wavelet feature as formulated in (3.3).

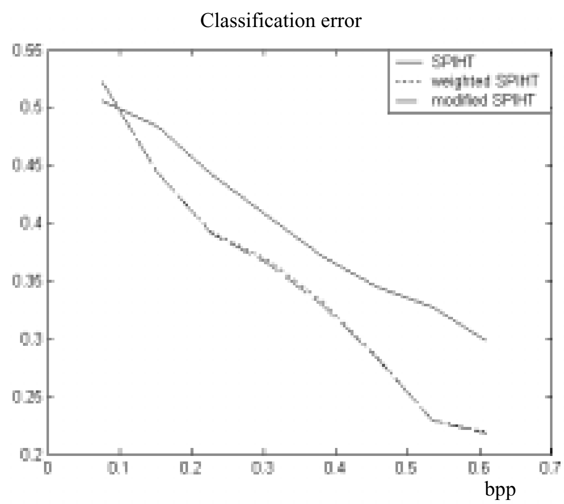

Figure 4. Classification performance in low bit rates

In Fig. 4 we compare the classification performances of the decoded image using different coding schemes (using HMT-based classification). The weighted SPIHT begins to code the wavelet coefficients important to segment the two textures in an earlier stage; therefore the classification rate is higher at low bit rates than the original SPIHT scheme. The modified SPIHT algorithm can save 2% of the bit budget to get the same performance in this case. It would be even more efficient in coding larger trees with more wavelet coefficients pruned.

Fig. 5 shows the decoded images using different coding schemes at the bit rate of 0.6 bit/pixel. The differences come from the particular subbands deep in the wavelet spatial orientation trees, which facilitate segmentation. The classification performance improves around 30% with little decrease in the image-reconstruction quality.

6. CONCLUSIONS

With the ultimate goal of image classification, we designed a scheme for properly pruning and weighing the wavelet coefficients before wavelet coding. We demonstrated improved segmentation performance of the decoded image, with little decrease in image quality at low bit rates. We also proposed a modified SPIHT algorithm, using the importance weight information, to save the bit budget. We tested the method on a high altitude two-texture aerial photographic image.

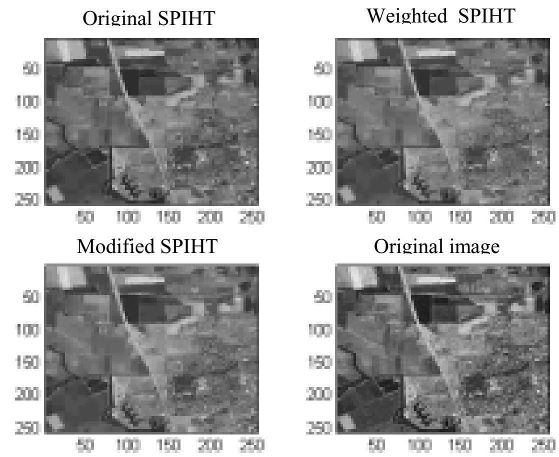

Figure 5. Decoded images at 0.6 bit/pixel bit rate

7. REFERENCES

[1] A. Said and W. A. Pearlman, "A New, Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees", *IEEE Trans. Circuits Syst. Video Technol.,* vol.6, pp. 243-250, June 1996

[2] Z. Wang and A. C. Bovik, " Embedded Foveation Image Coding", *IEEE Trans. Image Processing,* vol.10, pp. 1397-1410, Oct. 2001.

[3] M.S. Crouse, R.D. Nowak, and R.G. Baraniuk, "Wavelet-Based Statistical Signal Processing Using Hidden Markov Models", *IEEE Trans. Signal Processing,* vol. 46, pp. 886-902, April 1998.

[4] J. A. Bilmes, "A gentle tutorial on the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models", Technical Report ICSI-TR97-021, University of Berkeley, 1998

[5] X. Liao, H. Li, B. Krishnapuram, and L. Carin, " A Generative Learning Method for Kernel Machines with Application to Multi-aspect Target Identification", Technical Report, Feb. 2003, Department of Electrical and Computer Engineering, Duke University

[6] J. Rissanen, "Modeling by Shortest Data Description", *Automatic*, vol. 14, pp. 465-471, 1978

[7] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Trans. Inform. Theory*, vol. 38, pp. 719-746, March 1992

[8] P. Vincent and Y. Bengio. "Kernel matching pursuit", Machine Learning, 48, 165-187, 2002.