Image Compression with Adaptive Arithmetic Coding

S. Nigar Sulthana
ISRO Satellite Centre
HAL Airport Road
Bangalore -560017, INDIA

Mahesh Chandra
Department of Elec. & Commn Engg
Birla institute of Technology
Mesra, Ranchi, INIDIA

ABSTRACT

An image compression method using the wavelet transform, zero tree coding and adaptive arithmetic coding has been proposed. Here a novel static zeroth order adaptive arithmetic coder is being explored to improve the compression ratio. The proposed method decomposes an image into several subband images using the discrete wavelet transform, decorrelated coefficients quantized by Shapiro's embedded zerotree wavelet algorithm and encoded using static zeroth order adaptive arithmetic coder. The proposed static coder gives a better compression ratio while decreasing the coding time as compared to context based dynamic counterpart. The results obtained were comparable to those obtained by context modeling approach.

Categories and Subject Descriptors

I.4.2 [Computing Methodologies]: Compression (coding) – Approximate methods

General Terms

Algorithms, Performance, Design.

Keywords

Image compression, wavelet transform, arithmetic coding, embedded zero tree wavelets

1. INTRODUCTION

The explosive growth of remote sensing technology, internet and multimedia systems poses great challenge in handling huge amount of data .These huge amounts of data bring about the problems in transmission bandwidth, storage and processing, so efficient compression is indispensable. Generally, data compression algorithms [2, 7] can be classified into three separate stages: transform, quantization and entropy coding. The transform projects the original image onto a set of basis functions, thereby removing the redundancy due to the inter pixel dependencies in a given image. The 2-D wavelet transform has proved as an excellent candidate for the task. Quantization is done based on the embedded zero tree coding (EZW) proposed by Shapiro [4]. In addition to giving an excellent rate-distortion performance, the method has the property that bits in the bit stream are generated in order of significance so that it is ideal for progressive transmission. Entropy coding is by sophisticated arithmetic coder to achieve high compression. The performance of the arithmetic coder is directly related to coding gain. As entropy is always referred to a model, the fundamental problem in addressing arithmetic coder design is to build good models. Two major problems stand out as hurdles to efficient arithmetic coding are: 1) The modeling complexity grows exponentially with respect to the order of the Markov model, which exploits the redundancy between data symbols by context modeling, 2) Need of a practical context modeling technique. In this paper a histogram based static source model on the EZW's significant map output has been developed that addresses the above two practical problems.

The rest of the paper is organized as follows. Section 1 briefly introduces the concept of the wavelet transform and its implementation for 2-D signals. Section 2 gives brief introduction of wavelet transform to be used for image compression. Section 3 deals with the concept of zerotree and shows how zerotree coding can efficiently encode a significant map of wavelet coefficients. Section 4 explains adaptive arithmetic coding and the proposed algorithm. Experimental results are provided in section 5. The paper is concluded in section 6.

2. WAVELET TRANSFORM

The basic idea of wavelet transform is to represent any arbitrary function as a superposition of a wavelet basis. The wavelet basis is formed by dilation and translation of a special function, which is called the mother wavelet. The wavelet transform gives a spatial and frequency representation of signals. The wavelet transform can be implemented by a two channel perfect reconstruction (PR) filter bank [1]. A filter bank is a set of analysis and synthesis filters which are connected by sampling operators. The wavelet transform of a signal can be obtained by repeatedly applying a PR filter bank to the signal in a pyramidal scheme resulting in a dyadic decomposition. In a dyadic decomposition the lowest frequency band is decomposed in a recursive fashion and the number of these recursive steps is referred as levels. Two level, 1-D wavelet decomposition is shown in figure.1. The input is convoluted with low pass filter H_0 and high pass filter H_1 respectively and are sampled at every other pixel. The resulting signals are smooth signal f_{m+1} and detailed signal f'_{m+1} respectively. At each level, the

smooth and the detailed signal have all the information necessary to reconstruct the smooth signal at the next higher resolution level. The wavelet transform can be extended to multiple dimensions by the use of separable filters. Each dimension is filtered and sampled separately. Four subbands arise when wavelet transform is applied in two dimensions (vertical and horizontal) as shown in figure 2. The subbands labeled LH_1, HL_1 and HH_1 represent the
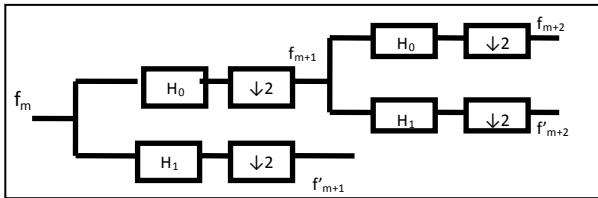

Figure 1. Two levels of 1-D wavelet decomposition

finest scale wavelet coefficients. At each coarser level, the coefficients represent a larger spatial area of the image, but a narrower band of frequencies. At each level, there are three subbands; the remaining lowest frequency subband is a representation of the information at all coarser levels. The filter used to compute discrete wavelet transform for the image compression described in this paper is 9/7 tap linear phase biorthogonal filter with the decomposition level five. These filters were chosen because in addition to their good localization properties, they produced good results empirically.

3. ZERO TREE CODING

In a dyadic wavelet decomposition explained in the previous section every coefficient is related to a set of coefficients at the next finer level that correspond to the same spatial location in the image. A coefficient at a coarse level is called parent, whereas its spatially related coefficients at the next finer level are referred as its children. The dependency is shown in figure 3. The coefficients at the coarsest level have only three children, whereas all the other
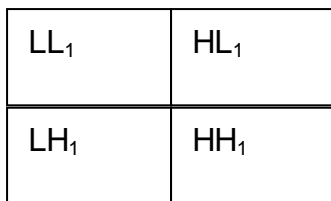

Figure 2. Sub-bands of single level two dimensional

coefficients except for the coefficients at the finest level have four. The coefficients at the finest level are childless. All the coefficients at the finer level that descend from a coefficient at a coarser level are called descendants. The EZW encoder [3] is an encoder specially designed to be used with wavelet transform. The zerotree is based on the hypothesis that if a wavelet
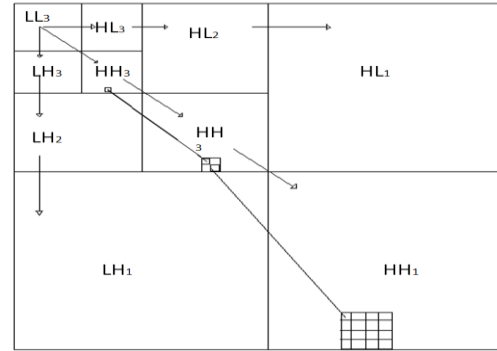

Figure 3. Three level zerotree dependency

coefficient at a coarser level is insignificant with respect to a given threshold, then all wavelet coefficients of the same orientation in

the same spatial location at finer scales are likely to be insignificant with respect to the same threshold. This approach can isolate interesting nonzero details by immediately eliminating large insignificant regions from consideration.

The procedure of zero tree coding is as follows. First the threshold T is determined by finding wavelet coefficient with the largest magnitude and setting the threshold to the integer power of 2 less than or equal to this value. The scanning of coefficients is performed in such away that no child node is scanned before its parent. For an N-scale transform, the scan begins at the lowest frequency subband LL_N and scans subbands HL_N, LH_N and HH_N at which point it moves on to scale N-1. Every coefficient is checked to determine whether its magnitude is greater than or equal to T, i.e.., whether it is significant with respect to the threshold. If a coefficient is found to be significant, it is coded as either positive significant (P) or negative significant (N) depending on its sign, and is placed on the list of significant coefficients. If the coefficient is not significant, then all its descendants are examined to see whether a significant descendant exists. If no significant descendants exist, it is coded as zerotree root (Z). If there exists a significant descendant, it is coded as isolated zero (T). The coefficients that descend from a zerotree root are not significant and need not be coded. This part of the algorithm is called dominant pass. The output of the dominant pass will be a string of the above said four letters. Next a subordinate pass is performed over the coefficients in the significant list. For every coefficient in this list the bit at the location $\log_2(T)-1$ in the coefficients binary representation i.e., the bit in the lower bit plane is coded. The output of the subordinate pass will be a string of ones and zeros. The zerotree encoder does progressive encoding to compress an image into a bit stream with increasing accuracy and that is achieved by successive approximation. Next the encoder halves the threshold and performs another dominant and subordinate pass. This process is iterated until the rate or distortion criteria are met. If a coefficient was determined to be significant at an earlier pass it will still be significant at the current

pass and need not be identified as significant again. The number of entries in the significant list grows monotonically as the threshold decreases

4. ADAPTIVE ARITHMETIC CODING

Arithmetic coding is a lossless data compression technique that encodes the input data by creating a string code that represents a fractional value between 0 and 1 on the real number line. Due to its high compression performance, arithmetic coding is used so ubiquitously in practice that it has become an indispensable part of any compression system. It encourages a clear separation between the model for representing the data and the encoding of information with respect to that model. It accommodates adaptive models easily and computationally efficient. The simplest model is a memory less one, in which the data symbols are considered as independent random variables. The maximum compression achievable by arithmetic coding is quantified in Shannon's noiseless source coding theorem, which states that the average code length is lower bounded by the entropy associated with the model. As entropy, the amount of compression is always referred to a mode; the fundamental problem in addressing arithmetic coder design is to build good models.

The simple memory less model is not efficient for compression. More sophisticated models are Markov models, in which redundancy between data symbols is exploited by context modeling [5, 8, 9]. The obvious goal of context modeling is for better compression which is guaranteed by another result in information theory which says that the higher the order of the Markov model, the lower the resulting conditional entropy. The better the model fits the source, the shorter the code length. Though better compression is achievable by increasing the order of the Markov model, in practice the modeling complexity grows exponentially with the order of the model and hence the speed of encoding suffers. If the number of model states is too large with respect to the size of the image, count statistics may not have enough samples to reach good estimation of conditional probabilities on the model states resulting in context dilution which leads to poor coding efficiency.

In the proposed approach as the bit streams resulting from the subordinate passes of the zerotree coding are hardly compressible even by higher order entropy coding, the bit streams of subordinate passes are outputted as they are. Arithmetic coding is performed only on the four letter strings resulting from dominant passes of the zerotree coding. Probability of each of the four letters occurring in the zerotree coding is statically calculated using histogram based zeroth order Markov model for the outputs of the dominant passes and this is done for each successive threshold separately. This source model is used for arithmetic encoding and this model is sent explicitly to the decoder. Arithmetic coding is done based on Elias scheme with scaling and rounding [6]. Sending this additional information of source model may result in increase in the total description length of the source. But the gain

achievable in the coding time is significant compared to the adaptive counterpart, where the source model is learnt on the fly from the data stream. In the case of remote sensing spacecrafts where huge image data has to be transmitted, the coding speed is more important than the progressive transmission.

5. RESULTS

The proposed adaptive arithmetic coder is used for wavelet based image compression. The PSNR numbers for the proposed arithmetic coder coupled with the EZW algorithm at various bit rates on the standard 512X512 'lena' test image are listed in Table 1. The proposed scheme achieves upto 10% bit rate saving than the memoryless model which is comparable with the results obtained by high order context modeling [10].

Table 1. Coding results for 'lena 512 X 512'

PSNR	30.16	32.37	33.34	33.42
bpp Huffman coding	0.17	0.38	0.75	1.46
bpp arithmetic coding with memoyless model	0.17	0.36	0.72	1.44
bpp arithmetic coding with the proposed adaptive model	0.15	0.33	0.68	1.35

6. CONCLUSION

The proposed adaptive arithmetic entropy coding technique for wavelet based image compression gives comparable results as that of systems using higher order context modeling. Studies are underway on obtaining an optimum context assigning method according to the local and global features of the image.

7. REFERENCES

[1] Jun Wang and Huang H.K, Medical Image Compression by Using Three Dimensional Wavelet Transformation, IEEE Transactions on Medical Imaging, Vol.15,no.4, August 1996.

[2] Lawson S. and Zhu J Image Compression Using Wavelets and JPEG2000", Comm, Electronics and Commn. Engg. Journal, June 2002.

[3] Li J, Cheng P.Y, and Kuo C.C.J., On the Improvement of Embedded Zerotree Wavelet (EZW) Coding, Proc. SPIE: Visual Communication and Image Processing, Vol. 2501,pp. 1490-1501, May 1995.

[4] Shapiro J.M, Embedded image coding using zerotrees of wavelet coefficients,IEEE Transactions on signal processing,vol.41,no.12,pp.3445-3462,December 1993.

[5] Seungjong Kim and Jechang Jeong, Image Compression Using the Wavelet Transform and Context-Based Arithmetic Coding, Proc. SPIE: Second International Conference on Image and Graphics Vol. 4875,pp. 133-140,2002.

[6] Tenkasi V. Ramabadran and Keshi Chen, "Efficient Compression of Medical Images Through Arithmetic Coding, Proc. SPIE: Medical Imaging, Vol. 1234,pp. 761-775, 1990.

[7] Witten I.H,.Nealand R,.cleary J.G, Arithmetic Coding for Data Compression, Comm, ACM, vol. 30, pp 520-540, June 1987.

[8] Wu X. and Chen J.H, Context Modelling and Entropy Coding of Wavelet Coefficients for Image Compression", Proceedings of the 1997 Inter.Conf.on ASSP, pp 3097-3100.

[9] Wu X and Memon N, CALIC-A Context Based Adaptive Image Codec, Proceedings of the 1996 IEEE International Conference on Acoustic, Speech and Signal Processing, pp. 1890-1893.

[10] Zixiang Xiong, Kannan Ramchandran and Michael T.Orchard, Efficient Arithmetic Coding For Wavelet Image Compression, Proc. SPIE: Vol. 3024, pp. 822-833, 1997.