AN EMBEDDED IMAGE CODING SYSTEM BASED ON TARP FILTER WITH CLASSIFICATION

Chao Tian and Sheila S. Hemami

School of Electrical and Computer Engineering
Cornell University
Ithaca, New York 14853

ABSTRACT

Recently, image compression systems based on the Tarp filter have attracted much attention in the image processing community. While providing very good performance when used in a non-embedded manner, the original Tarp-filter-based algorithm performs less competitively when used in an embedded manner (in spite of its operation on bitplanes), because of its raster scan encoding order. In this paper, we propose a Tarp-filter-based system which utilizes Classification of coefficients to achieve Embedding (TCE). The algorithm classifies the coefficients according to their statistical properties, and the Tarp filter only runs on the single class on which it tends to generate accurate probability estimates. TCE can achieve much better rate-distortion embedding performance than the original Tarp-filter-based system when used in an embedded manner; it achieves slightly better performance than SPIHT with arithmetic coding, and is comparable with JPEG-2000 performance on average.

1. INTRODUCTION

In a typical wavelet transform image coding system as shown in Fig. 1, an adaptive entropy coder is applied after the transform and quantization are performed. The performance of the system is significantly influenced by the efficiency of the entropy coder, which mainly depends on the accuracy of the probability estimates it uses.

It is widely believed that in the wavelet domain the neighboring coefficients capture essential context information, such as edges and patterns, and this information can help to achieve good compression. Many image codecs take this approach in the adaptive entropy coder [1–4], including JPEG2000. Recently, a new approach called a Tarp filter has been proposed [5], whose probability estimates are provided through an IIR filtering technique on the bitplanes of the wavelet coefficients. An image compression algorithm based on it can achieve performance comparable to JPEG2000.

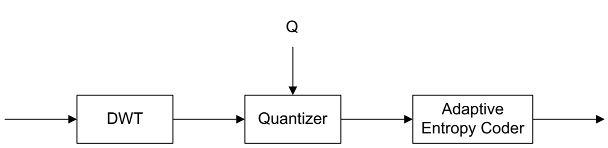

Fig. 1. Wavelet-based image coding diagram.

While providing very good performance when used in a non-embedded manner, the original Tarp-filter-based system [5] performs less competitively when used in an embedded manner, in spite of its operation on bitplanes. The reason is that the Tarp filter is designed to follow the raster scan encoding order; however, to achieve a good rate-distortion embedding, fractional bitplane methods ([2,4,6]) should be used, which require a more flexible encoding order. In this paper, we propose a Tarp-filter-based system which utilizes Classification of coefficients to achieve Embedding (TCE). The algorithm classifies the coefficients according to their statistical properties, and the Tarp filter only runs on the single class on which it tends to generate accurate probability estimates.

The paper proceeds as follows. Section 2 reviews the concepts of rate-distortion embedding, fractional bitplane methods and the Tarp filter technique. The TCE system is proposed in Section 3. Section 4 compares the performance of TCE with other codecs and Section 5 concludes the paper.

2. BACKGROUND

2.1. Rate-distortion embedding and fractional bitplane methods

An *embedded* image coding system generates a bitstream that can be truncated at any point and is still decodable. A *rate-distortion embedded* codec attempts to achieve the following: the embedded bitstream is generated in such a way that the coder is not only optimized at the final rate, but also optimized at every truncation point [2].

To achieve rate-distortion embedding, "the optimal strategy is to first encode those symbols with the steepest rate-distortion slope" [2]. Li *et al.* [7] showed that in a progressive bitplane coding system, the R-D slopes of the significance identification and refinement coding are different, and by placing the significance identification before the refinement coding, better rate-distortion embedding can be achieved. Later, this was systematically investigated by Ordentlich *et al.* [6] and Li and Lei [2]. It was found that the significance identification symbols which have high probability to become significant have the steepest R-D slope, and refinement bits are usually less important from a R-D point of view. These results motived the fractional bitplane method in EBCOT [4], which later was adopted in JPEG2000.

Specifically, Ordentlich *et al.* [6] observed that by classifying the coefficients based on the information of previous bitplanes, and encoding the classes in the following order, good rate-distortion embedding can be achieved:

1. *Non-zero neighbor* coefficients: the non-significant coefficients which have significant neighbors in the previous bitplanes;

2. *Non-zero parent* coefficients: the non-significant coefficients which have significant parents, but no significant neighbors in the previous bitplanes;

3. *Run* coefficients: non-significant coefficients that are not in the above two categories;

4. *Refinement* coefficients: significant coefficients that need refinement.

Four passes are made for each bitplane in the wavelet domain (and thus each represents a fractional bitplane), with each pass encoding the bit information of a specific class.

2.2. The Tarp filter technique

The Tarp filter technique was introduced in [5]. Consider a sequence of Bernoulli random variables, whose probability of being one is slowly changing. A simple estimate of the probability of the next variable being one can be obtained via the first order recursive filter:

$$p(t+1) = \alpha p(t) + (1-\alpha)v(t) \tag{1}$$

where $p(t)$ is the estimate of the probability of getting a one for position t, $v(t)$ is the observed value at position t, and α is the recursive parameter. This probability estimate can be used to drive a non-adaptive arithmetic coder to compress the information.

Simard *et al.* [5] generalized this idea to 2-D, by using three 1-D filtering steps, which results in the Tarp filter. In these three filters, the first filter runs from left to right; the second filter runs from right to left and is done after each full row has been processed; the third filter goes from top to bottom for each column. The probability estimate is calculated based on causal information, and the processing strictly follows a raster scanning order. By quantizing the wavelet coefficients with the same uniform quantizer on all the subbands, and using the Tarp filter to drive a non-adaptive arithmetic coder on the bitplanes of wavelet coefficient quantization indices, the Tarp-filter-based system can achieve performance comparable to JPEG2000.

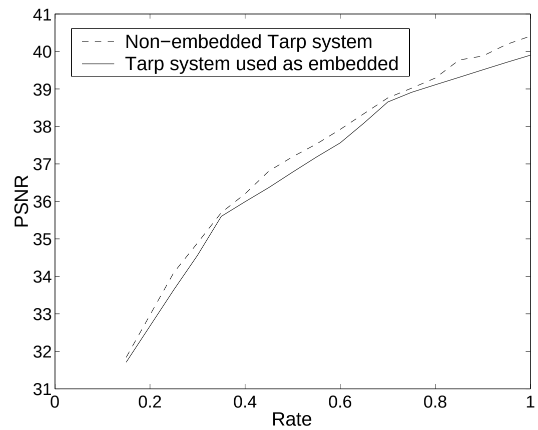

Fig. 2. Performance of the Tarp-filter-based system [5] and when it being used in an embedded manner on *Lenna*. The average difference is $0.31\ dB$, with the largest difference of $0.51\ dB$ at $1.0\ bpp$.

The Tarp-filter-based system in [5] does not have good rate-distortion embedding performance, in spite of its operation on bitplanes. The reason is that to achieve good rate-distortion embedding, the information to be placed into the bitstream has to be optimally ordered, but the raster scan encoding order of the Tarp filter prohibits this flexibility. In Fig. 2, the (non-embedded) Tarp system [5] (optimized for a final bit rate by adjusting quantization step-size) can achieve better performance, than when being used in an embedded manner (i.e., with a target rate ∞ but truncating the bitstream at different positions). For some images, this difference can be over $1\ dB$ at certain rates.

More precisely, the Tarp filter and the encoding order of fractional bitplane methods conflict as follows. Suppose the coefficients are classified as given in the previous section (i.e., as in [6]), and the raster scan encoding order is still to be used. During the scanning pass for *Non-zero neighbor* coefficients, if a coefficient belongs to other classes, it should not be encoded; but this implies this "observed" value can not be used in the filtering, which means the Tarp filter can't operate properly. A simple solution is to use a empirical value instead of the real "observed" value in the filtering, but the accuracy of the probability estimate will be jeopardized if the empirical value is not accurate.

An interesting fact we observed is that for the significance identification, the Tarp filter generates more accurate probability estimates in low activity areas; i.e., the Tarp filter is less suitable for areas that have a high density of ones. This disadvantage actually suggests an alternative encoding order. As aforementioned, fractional bitplane methods require that significance identification symbols with high probability of being one be coded before other information–these are the *non-zero neighbor* coefficients. Since the Tarp filter is less suitable for this class of coefficients, then instead of using a Tarp filter on them, a 1-D adaptive arithmetic coding can be used. By doing this, the overall performance of the system is not jeopardized, and in fact has the potential for improvement because of the better compression on the bits of the *non-zero neighbor* coefficients.

3. TARP WITH CLASSFICATION TO ACHIEVE EMBEDDING

3.1. The 3-pass structure of TCE

In the TCE algorithm, the *Non-zero parent* coefficients and *Run* coefficients are combined to form a single class, which we call *Zero-run* coefficients. For each bitplane, the TCE system performs 3 passes:

1. An adaptive arithmetic coder is used to encode bits of *Non-zero neighbor* coefficients;
2. The Tarp filter is used to produce the probability estimate for a non-adaptive arithmetic coder for *Zero-run* coefficients;
3. Bits of *Refinement* coefficients are encoded with an adaptive arithmetic coder.

The sign information bit of a coefficient is coded when needed with a probability estimate of 0.5. The encoding/decoding ends when the target rate is reached.

Note that in the second pass, the refinement bits are still not available, while the Tarp filter needs this information in the filtering. Empirical results suggest that the refinement bits are almost evenly distributed between zero and one. Thus, instead of using the true "observed" bits for the refinement coefficients in the filtering, $v(t) = 0.5$ is used.

3.2. Improving the accuracy of the probability estimate

More improvement can be achieved by refining the probability estimate in the second pass. Observe that before the second pass, some information is known for the busy areas. Specifically, the locations of the refinement coefficients and the bit values of the *Non-zero neighbor* coefficients are known. Thus a reversed Tarp filter can be applied before the second pass, and then combined with the Tarp filter running forward during the second pass to provide a better probability estimate. This benefit results from the 3-pass encoding, instead of only one-pass.

Cross-scale correlation in the wavelet domain can be utilized to refine the probability estimate even more. It is well-known that there are similarity between different scales at the same spatial location in the wavelet domain, which implies the probability estimates generated by the Tarp filter have this similarity too. However, because of the different energy distributions among the subbands, though the probability estimates in different subbands have similar variations, the magnitudes are very different. To solve this, the probability estimate of the parent coefficient can be properly scaled, which results in a reasonably good probability estimate for the children coefficients. Specifically, we calculate the scaling factor by the ratio of the number of the significant coefficients (normalized by the total number of coefficients in that subband) between the parent and the child subband. Suppose the probability estimate from the parent coefficient is $p_p(x, y)$ and the Tarp filter gives an estimate of $p_c(x, y)$ in the current subband, then a weighed summation is a more accurate probability estimate, which is given by:

$$p(x, y) = w_p p_p(x, y) + (1 - w_p) p_c(x, y) \tag{2}$$

where w_p is the weighting factor for the probability estimate from the parent subband. Empirically, we found $w_p = 0.3$ works well for all the subbands.

4. RESULTS

In this section, the performance of the TCE algorithm is compared with that of SPIHT with arithmetic coding and JPEG2000. The performance of the original Tarp filter system [5] and that when used in an embedded manner (under the name Tarp$_{eb}$) are also included.

In all the tests, the TCE's recursive parameter is set to $\alpha = 0.4$ (the performance of TCE is rather robust with $\alpha \in [0.3, 0.5]$ for natural images). The test images are the popular gray scale test images *Lenna*, *Barbara* and *Goldhill*, and the image *Woman* from JPEG2000 test suite. A 5-level wavelet decomposition is used with 9-7 tap filters. All the subbands are quantized with the same uniform quantizer with a deadzone twice the stepsize. Table 1 shows the performance of different systems in terms of PSNR (in dB). The JPEG2000 software used in this test is Jasper, which is publicly available. The performances of TCE, SPIHT and Tarp$_{eb}$ are tested by generating a single bitstream, but truncating it to different rates during decoding.

TCE outperforms SPIHT in the comparison, and its performance is comparable with JPEG2000. TCE can achieve similar performance to the Tarp-filter-based system [5], and it is much better than the original Tarp-filter-based system directly used in an embedded manner (Tarp$_{eb}$). For some test images, the improvement over Tarp$_{eb}$ is significant at certain rates (for *Barbara*, the difference is 0.79 dB at 1.0 bpp). TCE achieves better performance than JPEG2000 on

Lenna (512×512)					
Rate	J2K	Tarp [5]	SPIHT	Tarp$_{eb}$	TCE
0.15	31.58	31.84	31.89	31.71	**31.97**
0.25	34.04	34.10	34.11	33.64	**34.19**
0.50	37.22	37.20	37.21	36.78	**37.28**
1.00	40.31	40.41	40.41	39.90	**40.46**
Goldhill (512×512)					
Rate	J2K	Tarp [5]	SPIHT	Tarp$_{eb}$	TCE
0.15	28.98	28.97	28.96	28.79	**29.04**
0.25	30.51	30.54	30.56	30.41	**30.64**
0.50	33.21	33.16	33.13	32.97	**33.23**
1.00	36.53	36.61	36.55	36.18	**36.64**
Barbara (512×512)					
Rate	J2k	Tarp [5]	SPIHT	Tarp$_{eb}$	TCE
0.15	**25.93**	25.91	25.67	25.85	25.90
0.25	**28.36**	28.10	27.58	27.23	27.88
0.50	**32.26**	31.85	31.40	31.08	31.82
1.00	**37.15**	36.76	36.41	35.97	36.76
Woman (2560×2048)					
Rate	J2K	Tarp [5]	SPIHT	Tarp$_{eb}$	TCE
0.15	27.91	**28.05**	27.91	27.66	27.98
0.25	29.98	**30.06**	29.95	29.79	29.98
0.50	33.63	**33.65**	33.59	33.07	33.61
1.00	**38.44**	38.42	38.28	37.96	38.38

Table 1. Performance comparison of different codecs.

both *Lenna* and *Goldhill*, while JPEG2000 does better on *Barbara*. This is because *Barbara* has more details with certain pattern on which JPEG2000's context-based modeling works well, while *Lenna* and *Goldhill* have more smooth features on which the Tarp filter works slightly better.

Fig. 3 shows the rate-distortion behaviors of different embedded codecs for *Barbara*. At certain rates the performance of the Tarp-filter-based system when being used as an embedded coder is quite close to TCE, and these rates correspond to the transition points between bitplanes.

5. CONCLUSION

The Tarp filter technique is a promising new approach to provide accurate probability estimates in wavelet-based image codecs. We extend this technique to use it in an embedded system. The proposed TCE system can achieve much better rate-distortion embedding than the Tarp-filter-based system [5] directly used in an embedded manner. It achieves slightly better performance than SPIHT with arithmetic coding, and is comparable with JPEG2000 performance on average.

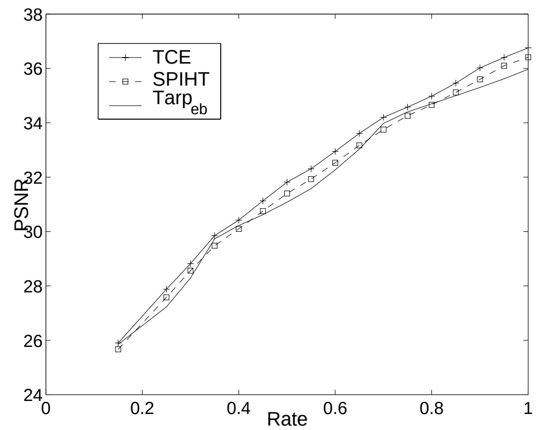

Fig. 3. Performance of embedded image codecs for *Barbara*. The average difference between TCE and SPIHT is $0.35\ dB$, with the largest difference of $0.45\ dB$ at $0.7\ bpp$; the average difference between TCE and Tarp$_{eb}$ is $0.48\ dB$, with the largest difference of $0.79\ dB$ at $1.0\ bpp$.

6. 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 and System for Video Tech.*, vol. 6, no. 3, pp. 243–250, Jun 1996.

[2] J. Li and S. Lei, "An embedded still image coder with rate-distortion optimization," *IEEE Trans. Image Processing*, vol. 8, no. 7, pp. 913–923, Jul 1999.

[3] X. Wu, "High-order context modeling and embedded conditional entropy coding of wavelet coefficients for image compression," in *Conf. Rec. 33rd Asilomar Conf. Sig., Sys. and Computers, Pacific Grove, CA*, Oct 1997, vol. 2, pp. 1378–1782.

[4] D. Taubman, "High performance scalable image compression with ebcot," *IEEE Trans. Image Processing*, vol. 9, no. 7, pp. 1158–1170, Jul 2000.

[5] P. Simard, D. Steinkraus, and H. Malvar, "On-line adaptation in image coding with a 2-d tarp filter," in *Proc. IEEE Data Compression Conference*, Snowbird, UT, Mar 2002, vol. 1, pp. 23–32.

[6] E. Ordentlich, M. Weinberger, and G. Seroussi, "A low-complexity modeling approach for embedded coding of wavelet coefficients," in *Proc. IEEE Data Compression Conference*, Snowbird, UT, Mar 1998.

[7] J. Li, P. Cheng, and C.-C.J. Kuo, "On the improvements of embedded zerotree wavelet (ezw) coding," in *Proc. SPIE:Visual Communication and Image Processing*, May 1995, vol. 2601, pp. 1490–1501.