A VISUAL MODEL FOR SUBBAND IMAGE CODING

W. C. Fong, S. C. Chan and K. L. Ho
Department of Electrical and Electronic Engineering
The University of Hong Kong, Pokfulam Road , Hong Kong.

ABSTRACT

In this paper, we propose a visual model for the general subband coding system. This model is a generalization of Watson's visual model for DCT. Based on this model, a new perceptual distortion is proposed which can be used together with an efficient bit allocation algorithm to determine the quantizer stepsizes of subband coders. Simulation results and subjective viewing tests demonstrate that the proposed approach is very effective in improving the subjective quality of the encoded image at low bit rate.

1. INTRODUCTION

The Human Visual System (HVS) is not an ideal receiver and does not response with equal sensitivity to the visual information. It is possible to take advantage of this fact in the encoding process [1]. The properties of HVS that have been incorporated in image coding include frequency sensitivity, luminance masking and texture masking. The most popular visual models that consider all these properties are the Watson [2] and Safranek-Jonhston [3] models. Watson developed perceptual masking threshold for DCT based coders. Safranek-Jonhston derived a similar nature model for a specific (4×4) subband coder. Watson's model is efficient and relatively simple to implement in which just the frequency sensitivity is required to be determined. In this paper, we generalize Watson's model to the general subband image coding and propose a perceptual distortion measure for subband coder to improve the subjective quality of the encoded image at low bit rate.

II. VISUAL MODEL

In the Watson and Safranek-Jonhston models, the masking thresholds are obtained by adjusting the base thresholds (frequency sensitivity) according to local brightness and textures of the picture. They both consist of three components: the base threshold, brightness correction and texture correction. The base threshold determines the frequency sensitivity of each subband for a flat mid-grey background. The brightness correction adjusts the base threshold for other background intensities. The texture correction compensates for texture masking effects. It should be noted that both models are frequency domain models which rely on the transform or subband coefficients to define the masking thresholds in the frequency domain. The differences are that they use different formula to estimate the texture component. For example, Watson models the texture masking as

$$m_{kn} = Max\left[t_{kn}, \left|c_{kn}\right|^{w_k} \cdot t_{kn}^{1-w_k}\right], \qquad (1)$$

$$t_{kn} = t_k \left(c_{0n} \big/ \bar{c}_0 \right)^{\alpha_\tau} . \qquad (2)$$

Here, the texture masking m_{kn} is estimated by the DCT coefficient c_{kn} alone, and is independent of other coefficients in the same block. The luminance masking t_{kn} is estimated by the DC coefficient c_{0n}. Whereas, Safranek's model uses the average value of the AC coefficients in adjacent blocks to estimate its texture component and use subjective experiments to adjust the base thresholds t_k to obtain the luminance masking.

Therefore, for the general subband image coding, it is reasonable to use the corresponding subband coefficients in place of the transform coefficients in equations (1) and (2) to calculate the masking thresholds. Here, c_{kn} becomes the coefficient at location n (block n) in the k subband, c_{0n} the DC coefficient at location n, $\bar{c}_0$ an average of the DC terms for the complete image, t_k is the base threshold for the *k-th* subband, and w_k is an exponent that lies between 0 and 1. The exponent may be different for different subbands. α_τ is a constant that controls the degree of luminance masking.

It can be seen from equations (1) and (2) that the parameters that we need to determine are the base threshold, t_k, and the constants, w_k and α_τ. From a series of subjective viewing of encoded images, we found that w_k and α_τ do not change significantly for different subband system. The values of w_k for the best subjective viewing of most images tested is tabulated in Table 1 for a block size of (8×8). It differs from the constant value of 0.7 used in [2] for DCT coefficients. The corresponding value for α_τ is found to be 0.649. Regarding to the most important basic component t_k, the base threshold, it is traditionally determined by a series of complicated and time-consuming psychophysical experiments as for the case DCT in [4] and for a 16-channel Generalized Quadrature Mirror Filterbank (GQMF) in [3]. Recently, a very simple and systematic method [5] is proposed to compute the base threshold t_k for any subband system in any color space. Using this method and the equations (1) and (2), masking thresholds for any subband system can easily be computed.

III. PERCEPTUAL DISTORTION MEASURE

The masking thresholds determined in Section II can be used to define a perceptual distortion measure. This measure, together with the bit allocation algorithm we propose in [6] can be used to determine the stepsizes of quantizers in subband coder. For simplicity and ease of comparison with JPEG algorithm, we assume that the subband coder is a separable (8×8) subband coding system. The quantization of the subband coefficients will also be identical to the JPEG algorithm. The remaining problem is to determine the quantization matrix of the coder.

Using the masking thresholds m_{kn}, we first define the perceptual distortion P_{kn} for the coefficient c_{kn} at location n (block n) in the k subband as follows:

$$P_{kn} = \left(\frac{d_{kn}}{m_{kn}}\right)^2 \qquad (3)$$

where d_{kn} is the quantization error of the subband coefficient c_{kn} and m_{kn} is the masking threshold for this coefficient. Then, we combine the local perceptual distortion P_{kn} into a single perceptual distortion measure as follows:

$$P = \sum_{n=1}^{N}\sum_{k=0}^{63} P_{kn} \qquad (4)$$

where P is the overall perceptual distortion of the encoded image which can be used as a subjective quality measure. It is then used together with the bit allocation algorithm [6] to determine the quantization matrix for subband coder. The problem can now be formulated as: Determine the quantizer step sizes $\{Q_k : k = 0,\ldots,63\}$ to minimize the overall perceptual distortion:

$$P(\mathrm{Q}) = \sum_{n=1}^{N}\sum_{k=0}^{63} P_{n,k}(Q_k) \qquad (5)$$

subjected to the bit rate constraint:

$$R(\mathrm{Q}) = \sum_{n=1}^{N} R_n(Q_0,\ldots,Q_{63}) \le B \qquad (6)$$

where $P_{k,n}(Q_k)$ is the perceptual distortion for the coefficient c_{kn} if it is quantized with step size Q_k, $R_n(Q_{0,}\ldots,Q_{63})$ is the number of bits generated in coding the n-th block with the quantization table $\{Q_{0,}\ldots,Q_{63}\}$, and $\mathrm{Q} = \{Q_0,\ldots,Q_{63}\}$ is the vector of quantization stepsizes.

The subjectively optimized quantization matrix for subband coders can now be obtained by using the procedure in [6] to solve this bit allocation problem with replacement of the distortion measure by the perceptual distortion proposed in this paper.

IV. SIMULATION RESULTS

In order to demonstrate the effectiveness of the proposed method for subband coder, we choose the Lapped Orthogonal Transform (LOT) [7] as an example. Here, the LOT used is a (8×8) channel perfect reconstruction filter bank whose basis function are of length L=16. The base threshold for this system is determined in [5] as shown in Table 2. Comparisons are made between the following algorithms:

1. **LOT_MSE**: A LOT codec with bit allocation algorithm in [6] and using the MSE as the distortion measure;
2. **LOT_PD**: A LOT codec with bit allocation algorithm in [6] and using the proposed perceptual distortion measure;
3. **JPEG**: JPEG baseline algorithm in [8].

Figure 1 shows the Lenna image (512×512) compressed to 0.30 bpp by the **JPEG**, **LOT_MSE** and **LOT_PD**. Note that there are far fewer blocking artifacts in the face and shoulder of the Lenna image encoded by the **LOT_PD** than the other two. Subjective viewing tests (using MOS) are also conducted to evaluate the subjective performance of these coders. The results of subjective tests also show that the perceptual **LOT_PD** coder has a better subjective quality than the **LOT_MSE** and **JPEG** coders at low bit rate.

V. CONCLUSION

In this paper, a visual model for the general subband image coding system is proposed. Using this model, a perceptual distortion measure is proposed which is used together with an efficient bit allocation algorithm to obtain the subjectively optimized quantization stepsizes for subband coders. Simulation results and subjective viewing tests show that this model is very effective in improving the subjective quality of the encoded image at low bit rate.

Table 1: The values for parameters w_k

0.2	0.2	0.3	0.6	0.6	0.6	0.6	0.6
0.2	0.3	0.3	0.6	0.6	0.6	0.6	0.6
0.3	0.3	0.6	0.6	0.6	0.6	0.6	0.6
0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6

Table 2: Visibility threshold for LOT (a) Luminance (b) Chrominance (CrCb)

Luminance (Y)							
10	9	9	12	19	23	32	34
11	11	11	14	23	27	39	37
13	12	15	21	28	39	45	43
13	16	19	26	38	51	59	56
16	20	25	37	49	70	75	72
22	26	33	51	76	97	92	93
28	34	48	63	87	100	99	100
37	37	45	58	83	101	100	100

(a)

Chrominance (CrCb)							
15	16	19	29	44	55	59	56
16	18	23	31	51	60	70	66
19	23	28	39	63	74	86	82
29	31	39	81	88	99	99	100
44	51	63	88	95	97	97	98
55	60	74	99	97	99	99	100
59	70	86	99	97	99	99	100
56	66	82	100	97	100	100	100

(b)

REFERENCES

[1] Visual Perception Academic Press, Inc. Orlando, FL, 1970.

[2] A. B. Watson, "DCT Quantization Matrices visually optimized for individual images," SPIE vol.1913 Human Vision, Visual Processing, and Digital Display IV (1993), pp. 202-216,1993.

[3] R. J. Safranek, et al, "A perceptually tuned sub-band image coder," *SPIE VOL. 1249 Human Vision and Electronic Imaging: Models, Methods, and Applications* (1990), pp.284-293.

[4] H.A.Peterson, et al, "Quantization of color image components in the DCT domain", *SPIE vol. 1453 Human vision, Visual processing, and Digital Display II* (1991), pp. 210-221,1991.

[5] W. C .Fong, S.C.Chan and K. L. Ho, "On the determination of frequency sensitivities for perceptual subband image coding," *IEE Electronics Letters, Vol. 33, No. 7, pp. 581-582, Mar. 1997.*

[6] W. C. Fong, S. C. Chan and K. L. Ho, "Designing JPEG quantization matrix using rate-distortion approach and human visual system model," in Proceedings IEEE ICC'97.

[7] H. S. Malvar and D. H. Ataelin, "The LOT: Transform coding without block effects," *IEEE Trans. ASSP-37*, pp.553-559, April 1989.

[8] ISO 10918: Digital compression and coding of continuous-tone still images.

Figure 1: Lenna image (a) is compressed to 0.30 bpp using (b) **JPEG** (c) **LOT_MSE** (d) **LOT_PD**