DETERMINATION OF THE VISIBILITY THRESHOLDS 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.
E-mail: scchan@hkueee.hku.hk

ABSTRACT

In this paper, a method to determine the visibility thresholds for a given subband system is introduced. Our approach identifies the important worse case combination of the quantization error in each basis function and determine the corresponding visibility thresholds using results previously measured in the DCT domain. Using the proposed method, it is relatively simple to determine the visibility thresholds for any subband system in the YCrCb domain. Simulation results on the Lapped Orthogonal Transform (LOT) show that the visibility thresholds are useful in visually lossless coding and perceptual weighted quantization.

1. INTRODUCTION

Compression of high-quality images is very important in digital multimedia and other applications. Due to the ease in incorporating the psychovisual rules for noise shaping, subband coding has emerged as a promising approach for coding images. Human visual system (HVS) is basically nonlinear. Properties that have been considered in image coding includes: perception of color, spatial frequency dependency, spatial orientation, luminance masking and spatial masking. Works in incorporating HVS can broadly be classified into codecs that achieve visually lossless encoded image [1-3] and codecs that use HVS to improve the visual quality of the encoded images [4-5].

In the first type of codecs, local visibility (or perceptual) thresholds in spatial or subband domain are calculated using various properties of the HVS. The image is quantized using these visibility thresholds hoping to produce visually lossless encoded images. Visibility thresholds in the DCT domain based on spatial frequency have been obtained by Lohscheller et al [1] in YCrCb space and by Peterson et al [2] in the RGB space. Later, a detection model for DCT domain in different viewing conditions is derived by Peterson et al [6]. More sophisticated models that take into account the spatial masking effects were proposed by Safranek et al [3]. In the second type of codecs, the frequency dependencies of the HVS are used to weight the distortion in each subband hoping to produce visually more pleasant images at lower bit rates. Ngan et al [7] and Chitpraet and Rao [4] had reported the use of the frequency sensitivity of the HVS in DCT domain. Perkins and Lookabaugh [5] used empirical formulas for the contrast and angular sensitivities and applied them to general subband coding. It should be noted that the visibility thresholds obtained in [1] can also be used as weighting factor in the DCT domain in the second type of codecs.

In this paper, we addressed the problem of determining the visibility thresholds (based on frequency sensitivities) and their application to both types of coder mentioned previously (i.e. visually lossless and visually lossy) in the subband domain. Unlike block transform such as DCT, the visibility thresholds in subband system cannot be determined by looking at each individual basis function alone. This is because subband coding relies on the overlapping of the basis functions to reconstruct the original image. In fact, the visibility of the error will also depend on the quantization error found in adjacent blocks. Here, we proposed a method so that these visibility thresholds can be determined from the visibility thresholds measured previously in the DCT domain [1,2]. Using the proposed method, it is relatively simple to determine the visibility thresholds for any subband system in the YCrCb domain.

II. THE PROPOSED METHOD

In [2], Peterson et al performed a series of subjective experiments to determine the visibility thresholds for the DCT domain in the RGB color space. Each visibility threshold is obtained by determining the amplitude necessary for the DCT basis function to be just noticeable to an average viewer. It might be expected that the same approach can be used to determine the visibility thresholds for all the basis functions in any given subband system. Unfortunately, due to the overlapping of the basis functions, the visibility of the error will depend on the quantization errors found in adjacent blocks. We must consider all the possible combinations in the overlapping basis functions. However, it is very difficult and time consuming to design such a subjective experiment.

Our approach is as follows: First, we identify the important worse case combination of the quantization error in each basis function. We then determine, using the DCT visibility threshold, whether this error image is visible or not. By increasing the quantization error until the error image becomes visible, the visibility threshold for that basis function is determined. The process is then repeated for all the basis functions to determine the visibility thresholds for the given analysis-synthesis system.

Suppose that we are given an error image, we want to determine whether this error image will be visible to the viewer using the visibility thresholds in the DCT domain. It is reasonable to assume that if any of the (8x8) blocks

within the error image is visible, then the entire error image will become visible. Therefore, we can transform all the (8x8) blocks to the DCT domain and compare the corresponding DCT coefficients with the visibility thresholds measured in [1]. Hence, we have the following assumption:

A1: *An error image is said to be invisible if each DCT coefficient in each (8x8) pixels block within the error image is below its corresponding visibility threshold in the DCT domain.*

Therefore, it is possible, in principle, to determine the visibility threshold for a given basis function by constructing all possible combination of the quantization errors. This is, however, computational intractable. A reasonable approach is to consider only the important worse case combination. Here, all the quantization errors in adjacent blocks are assumed to be their maximum value $t_{mn}^{(i)}$. Since the quantization error may be of different phase (in phase or 180 out of phase, i.e. positive and negative), thus there are nine different kinds of combinations as shown in Figure 1. For each combination, we obtain an error image. The amplitude $t_{mn}^{(i)}$ of the error image to be just noticeable is determined using (A1). The visibility threshold t_{mn} of this subband is the minimum value of $t_{mn}^{(i)}$ in the nine cases. Or equivalently, $t_{mn} = \min_i (t_{mn}^{(i)})$.

It should be noted that visibility threshold measured in [1] is dependent on the display conditions. Since we have made use of the results in [1], the display conditions is also the same. To apply these thresholds to different viewing conditions, the detection model for DCT developed by Peterson et al [7] can be used.

III. APPLICATION TO SUBBAND IMAGE CODING

The visibility thresholds determined in Section II can directly be used to quantize the subband coefficients to obtain a visually lossless subband image coder. They can also be used as weighting factors in perceptual weighted quantization. For simplicity and ease of comparison with the JPEG algorithm, we shall consider the separable subband coding system with (8x8) number of channels. 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. The problem can be stated as follows: Determine the quantizer step sizes $\{Q_k : k = 0, ..., 63\}$ to minimize the overall distortion:

$$D(\mathbf{Q}) = \sum_{n=1}^{N} \sum_{k=0}^{63} D_{n,k}(Q_k) \qquad (1)$$

subjected to the bit rate constraint:

$$R(\mathbf{Q}) = \sum_{n=1}^{N} R_n(Q_0, ..., Q_{63}) \le B \qquad (2)$$

where $D_{n,k}(Q_k)$ is the distortion in the k-th DCT coefficient of the n-th block if it is quantized with step size Q_k,

$R_n(Q_0, ..., Q_{63})$ is the number of bits generated in coding the n-th block with the quantization table $\{Q_0, ..., Q_{63}\}$, and $\mathbf{Q} = \{Q_0, ..., Q_{63}\}$ is the vector of quantization stepsizes. The effect of the frequency sensitivity can be taken into account by weighting the distortion in each subband as follows:

$$D_w(\mathbf{Q}) = \sum_{n=1}^{N} \sum_{k=0}^{63} D_{n,k}(Q_k) / w_k^2 \qquad (3)$$

where w_k are the visibility thresholds determined in Section II. The determination of the quantization table for the JPEG algorithm using the rate-distortion approach had been discussed in [8] and can be used here. Starting from an initial quantization table of large step sizes, the algorithm [8] decreases the step size in one entry of the quantization table at a time until a target bit rate is reached. In each iteration, the algorithm first computes the efficiency:

$$\lambda_k = \max_q \frac{-\Delta D(\mathbf{Q})|_{Q_k \to q_k}}{\Delta R(\mathbf{Q})|_{Q_k \to q_k}} \qquad (4)$$

for all values of k and then solves:

$$\lambda = \max_k \lambda_k \qquad (5)$$

where $\Delta D(\mathbf{Q})|_{Q_k \to q_k}$ and $\Delta R(\mathbf{Q})|_{Q_k \to k}$ are respectively the change in distortion and the change in overall bit rate when the k-th entry of the quantization table, Q_k, is replaced by q_k. The algorithm can be summarized as follows:

1. Initialize the quantization table by:

$$Q_k^{(0)} = \begin{cases} 16 & for & k = 0, \\ Q_{max} & for & k = 1, ..., 63 \end{cases}$$

2. Initialize tables of λ_k and $\tilde{Q}_k$ by searching q_k in $\{1, ..., Q_{max} - 1\}$ to solve (4) for k=1,...,63.

At iteration ℓ, we perform the following:

3. Search k in $\{1, ..., 63\}$ for p to solve (5).

4. Update the stepsizes by setting $Q_p^{(\ell)} = \tilde{Q}_p$.

5. Update λ_p and $\tilde{Q}_p$ by searching q_p in $\{1, ..., Q_p^{(\ell)} - 1\}$ to solve (4).

6. Repeat Steps 3 to 5 until $R(\mathbf{Q}^{(\ell)}) \le B$.

The maximum step size, Q_{Max}, is set to 128.

In this work, we shall use an improved version of this algorithm that we have recently proposed in [9] to perform the bit allocation.

IV. SIMULATION RESULTS

In order to investigate the performance of the proposed method, we applied the process described above to the Lapped Orthogonal Transform (LOT) in the YCrCb color space. LOT is a M-channel perfect reconstruction filter bank whose basis functions are linear-phase and of length equal to L=2M where M is the number of channels or subbands. It maps an LxL pixels block into MxM transform

coefficients. Thus, the LOT basis functions overlap adjacent blocks. In this work, we select M = 8. The visibility thresholds for the LOT that we have determined is shown in table 1. To test the effectiveness of the thresholds, the following image codecs were designed:

1. **LOT_VL**: A visually lossless LOT codec with the proposed visibility thresholds,

2. **DCT_VL** : A visually lossless JPEG codec with the visibility thresholds in [1],

3. **LOT_BA_P** : An LOT codec with perceptual weighted bit allocation.

4. **LOT_BA** : An LOT codec using the bit-allocation algorithm to generate the quantization table.

5. **JPEG** : A lossy JPEG codec with the default quantization matrix,

The structure of the LOT codecs are identical to the JPEG [9] algorithm except that the DCT are replaced by the LOT. Many grayscale and color images are encoded using the algorithms. For the visually lossless algorithms: **LOT_VL** and the **DCT_VL**, most of them are indistinguishable visually from the original images. Only in a small number of the test images like the boat image, small distortion appeared in the smooth and small luminance areas. This is reasonable because we haven't considered the local luminance contrast. Figure 2 shows the original and the encoded boat image using the two algorithms. Figure 3 shows that baboon image encoded to 0.35 bit per pixel using the **LOT_BA_P**, **LOT_BA** and the **JPEG** algorithms. It can be seen that the **LOT_BA_P** algorithm has the best visual quality followed by the **LOT_BA** and the **JPEG** algorithms.

V. CONCLUSION

In this paper, we present a method to determine the visibility thresholds for a given subband system. Our approach identifies the important worse case combination of the quantization error in each basis function and determine the corresponding visibility thresholds using results previously measured in the DCT domain. Using the proposed method, we determined the visibility thresholds of the Lapped Orthogonal Transform (LOT) in the YCrCb color space and demonstrated their usefulness with a visually lossless LOT coder and a lower bit rate image codec similar to the JPEG algorithm. Simulation results show that the visibility thresholds are useful in visually lossless coding and perceptual weighted quantization.

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)

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)

Table 1: a) Visibility Thresholds for Luminance (Y) component of the LOT, b) Visbility Thresholds for Chrominance (CrCb) components of the LOT.
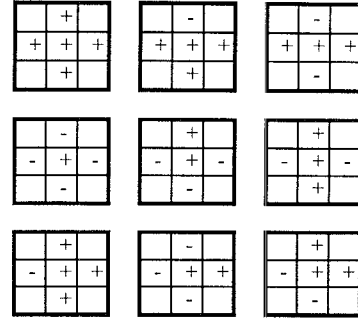

Figure 1. Worse case error combination patterns.

REFERENCES

[1] II. Lohscheller and U. Franke, "Color picture coding-algorithm optimization and technical realization," *Frequenz*, vol.41,pp.291-299,1987.

[2] 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.

[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] B.Chiprasert et al, "Human visual weighted progressive image transmission," *IEEE Trans. COM-38*, no.7, pp.1040-1044, July 1990.

[5] M. G. Perkins et al, "A Psychophysically justified bit allocation algorithm for subband image coding systems," *Proc. IEEE ICASSP-89*, pp.1815-1818.

[6] H. A. Peterson, et al "An Improved Detection Model for DCT Coefficient Quantization". *Proc. SPIE: Human vision, Visual Processing, and Digital Display IV*, pp.191-201, 1993.

[7] K. N. Ngan, et al ,"Cosine transform coding incorporating human visual system model," *IEEE Trans. ASSP-37*, No. 11, Nov. 1989, pp. 1743-1749.

[8] S. W. Wu and A. Gersho, "Rate-constrained picture-adaptive quantization for JPEG baseline coders," *IEEE ICASSP-93*, vol. V, pp. 390-392.

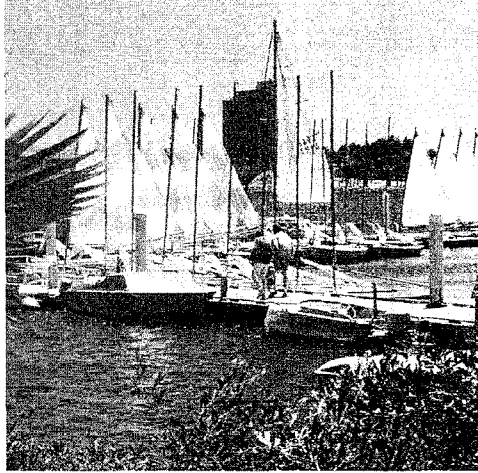[9] W. C. Fong, S. C. Chan and K. L. Ho, "Designing JPEG quantization matrix using rate-distortion approach and human visual system model," submitted for publication.
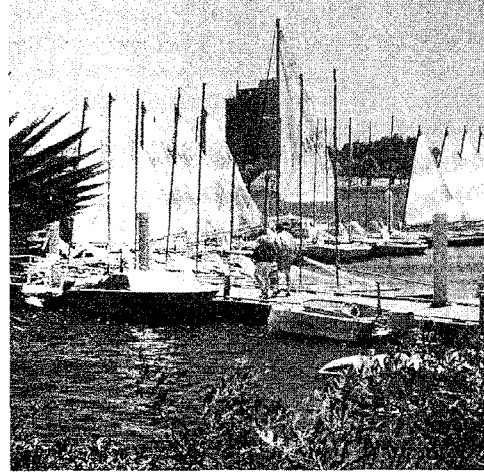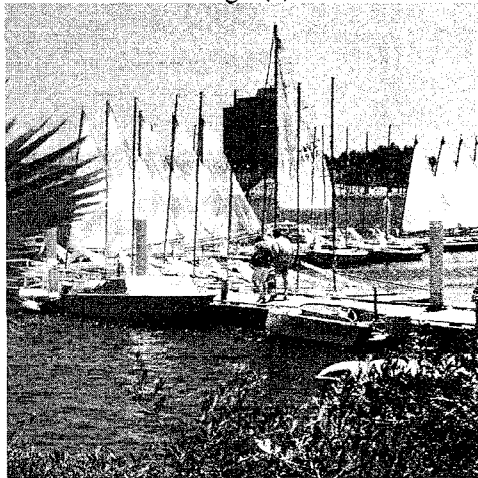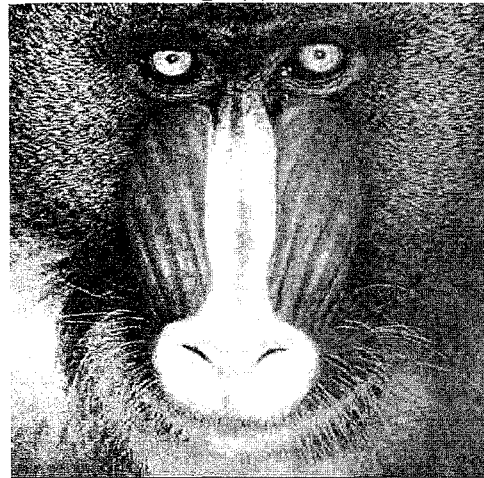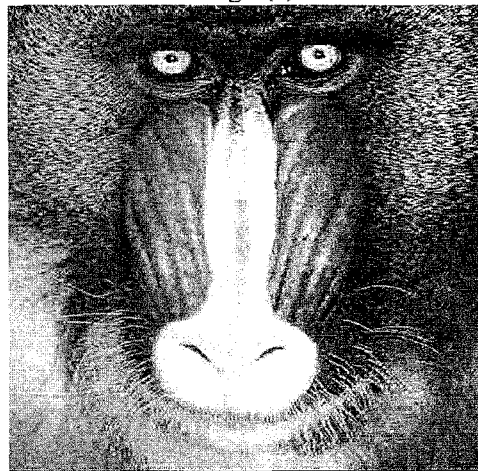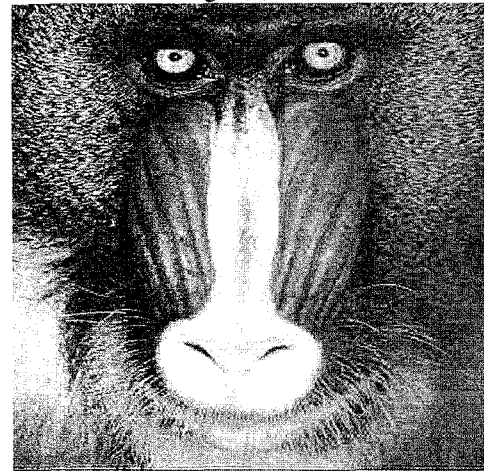
Fig.2(a)

Fig.2(b)

Fig.2(c)

Fig.3(a)

Fig. 3(b)

Fig. 3(c)

Figure 2. Boat images: (a) original, (b) compressed by DCT_VL, (c) compressed by LOT_VL algorithm.
Figure 3. Baboon images: (a) compressed by JPEG, (b) compressed by LOT_BA, (c) compressed by LOT_BA_P (0.35bpp).