for the assessment. They are from CDS at 16 bits/sample, mono and with sampling rate = 44.1 kHz. They are coded with eight different bit rates ranging from 47.5 to 205 kbit/s. The perceptual audio quality measure (PAQM) developed in [5] was used to assess the effectiveness of the perceptual filter on reducing the noise disturbance in loudness due to quantisation. The segmental signal-to-noise ratio (SSNR) was also used to measure the objective performance. Listening tests and mean opinion scoring (MOS) were used to assess the subjective performance of the perceptual filter.

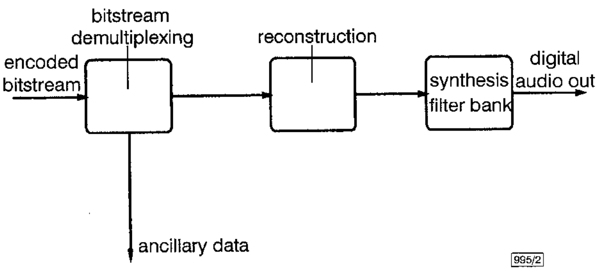

Fig. 2 *Sub-band audio decoder with perceptual filtering*

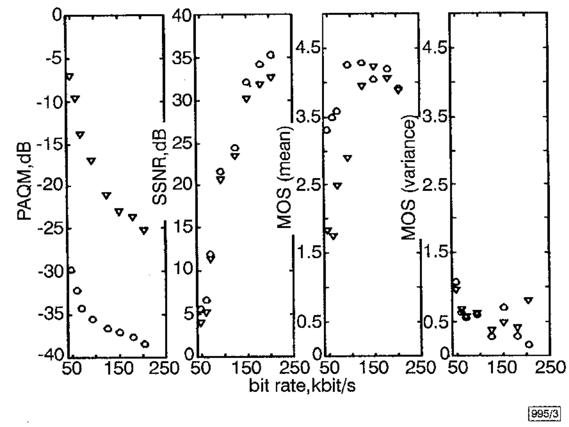

Fig. 3 *Test results of Haydn symphony*

○ PFC signal
▽ non-PFC signal

Fig. 3 shows the assessment results of the Haydn symphony. The amount of noise disturbance appears to depend strongly on the assigned bit rate. The lower the bit rate, the more serious is the noise disturbance. It is clear that the generalised perceptual filter has successfully reduced the noise disturbance by ~10–15 dB. The SSNR results reveal that, in all cases, the use of perceptual filtering improved the SSNR to a small degree (~0.5–3 dB). These results were expected, as the perceptual filter is not SSNR oriented. The MOS of PFC for the symphony is ~1.3–1.7 higher than that of non-PFC at bit rates < 100 kbit/s and the filtered signal achieved an average 3.5 MOS. However, for bit rates > 125 kbit/s, most listeners failed to distinguish the quality differences between different bit rate coders.

The performance of coders for the piano music is similar to that for the symphony, except that the SSNRs are higher, but only ~0.5 MOS improvement is observed for coders with a bit rate < 100 kbit/s. This may be due to the presence of strong attacks in the piano music, and the fact that the excitation pattern model from which the filter is derived is mainly for 'stationary' signals.

Conclusion: A generalised perceptual filter based on a psychoacoustic excitation pattern model has been developed to exploit the nonlinear property of human hearing. It is shown that this filter can be applied to a sub-band audio coder without increasing the decoder complexity and the bit rate requirement. Experimental results showed that the noise disturbance owing to quantisation was successfully reduced, particularly at low bit rate audio coding. This led to an improvement in subjective signal quality in mean opinion scoring tests.

23 February 1998

Electronics Letters Online No: 19980570

Y.H. Lam and R.W. Stewart (*Signal Processing Division, Department of Electrical and Electronic Engineering, University of Strathclyde, Glasgow, G1 1XW, United Kingdom*)

E-mail: (vicky,bob)@spd.eee.strath.ac.uk

References

1 JOHNSTON, J.: 'Transform coding of audio signal using perceptual noise criteria', *IEEE J. Sel. Areas Commun.*, 1988, **SAC-6**, (2), pp. 314–323

2 ZWICKER, E., and FASTL, H.: 'Psychoacoustics, facts and models' (Springer, Berlin, Heidelberg, 1990)

3 VELDHUIS, R.N.J.: 'Bit rates in audio source coding', *IEEE J. Sel. Areas Commun.*, 1992, **SAC-10**, (1), pp. 86–96

4 TSOUKALAS, D.E., MOURJOPOULOS, J., and KOKKINAKISH, G.: 'Perceptual filters for audio signal enhancement', *J. Audio Eng. Soc.*, 1997, **45**, (1/2), pp. 22–36

5 BEERENDS, J.G., and STEMERDINK, J.A.: 'A perceptual audio quality measure based on a psychoacoustic sound representation', *J. Audio Eng. Soc.*, 1992, **40**, (12), pp. 963–978

6 LAM, Y.H., and STEWART, R.W.: 'Perceptual suppression of quantisation noise in low bit rate audio coding'. Asilomar Conf. Signals, Systems and Computers, Monterey, CA, USA, 1997

Adaptive image watermarking scheme based on visual masking

Jiwu Huang and Yun Q. Shi

An image watermarking scheme in a DCT domain is proposed. It is adaptive in that a different strength of watermark is applied to different blocks according to their block classification. For the classification, visual masking (both luminance and texture masking) is taken into account. Consequently, the watermarks thus generated are robust and invisible.

Introduction: Providing copyright protection for digital media is becoming increasingly important. For this reason, digital watermarking, especially for images and video, has drawn extensive attention recently. Two of the most basic requirements of watermarking, i.e. invisibility and robustness, are in conflict with each other. For instance, some watermarks reported in literature are not robust since they are made to be invisible; it is vital to make a watermark both robust and invisible.

In this Letter, we propose an adaptive watermarking algorithm in a discrete cosine transform (DCT) domain based on visual masking. We embed watermarks in some significant DCT coefficients so that they cannot be easily damaged. To embed strong watermarks in images under the constraint of being invisible, classification based on visual masking is applied to the image blocks prior to watermark embedding. The watermarks are embedded adaptively according to the classification. Simulation results demonstrate that watermarks generated in this way are very robust against some typical image processing techniques such as compression, lowpass filtering and subsampling.

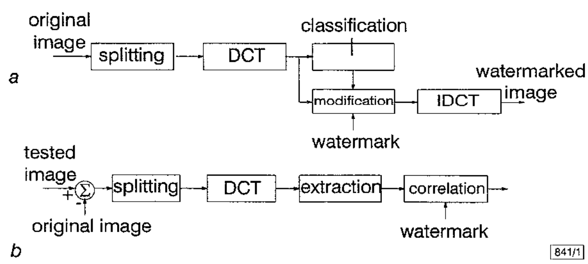

Fig. 1 *Proposed watermarking algorithm*

Overview of proposed algorithm: The proposed watermarking algorithm is shown in Fig. 1. The original image is first split into non-overlapped blocks of 8 × 8 and then each block is DCT-transformed. In the DCT domain, the average brightness and texture complexity for each block are estimated. Based on this estimation, all blocks are classified into three categories: dark and weak texture (class 1), bright and strong texture (class 3), and remains (class 2). The watermark, a random number sequence, is divided into groups. Each group, having three numbers, is inserted into three significant DCT coefficients of one of the blocks in the image, with different strengths: the strongest for class 3 and the weakest for class 1. After the modification of DCT coefficients, an inverse DCT is performed for each block in order to obtain the watermarked image.

The detection of watermark is quite straight forward. That is, with the original image, the possible watermark is extracted from the tested image. The correlation of the extracted watermark and the original watermark is computed and used to decide whether a watermark exists.

Watermark embedding: Watermark embedding is implemented via the following four steps: (i) image splitting and DCT; (ii) block classification in a DCT domain; (iii) watermark generation and casting; (iv) inverse DCT.

We first split the original image, $f(x,y)$, into non-overlapped blocks of 8 × 8, denoted as B_k, $k = 0, 1, \ldots$, for the sake of being compatible with the image and video compression standards. After splitting, the DCT of B_k, $F_k(u,v)$, is computed.

Embedding a watermark can be viewed as superimposing a weak signal onto a strong background. According to the visual features of the human vision system (HVS), there are two observations: (i) the brighter the background, the higher the embedded signal could be (luminance masking [2]), and (ii) the stronger the texture in the background, the lower the visibility of the embedded signal would be (texture masking [3]). Based on visual masking, we classify each B_k into one of three categories: class 1, class 2 and class 3, as defined above.

The average brightness information of the background is represented by the DC component in the DCT domain. To estimate the texture complexity, we quantise the DCT coefficients using the same method as that used in JPEG. Then the number of non-zero coefficients is computed. The larger the number, the stronger the texture in the block is.

If $F_k(0,0) < T_1$ and $number\{int(F_k(u,v)/Q(u,v)) \neq 0\} < T_2$, then B_k is classified into class 1. If $F_k(0,0) > T_1$ and $number\{int(F_k(u,v)/Q(u,v)) \neq 0\} > T_2$, then B_k is classified into class 3; otherwise, into class 2. Here, $int(\cdot)$ means taking integer and $Q(u,v)$ denotes the quantisation step at (u,v). T_1 and T_2 are two predefined thresholds.

As in [1], the watermark W is chosen to be a random number sequence of length n, i.e. $W = \{x_i, 0 \le i < n\}$, that obeys the normal distribution $N(0,1)$. To embed the watermark, the DCT coefficients are modified as follows:

$$F'_k(u,v) = \begin{cases} F_k(u,v) + \alpha_k \cdot x_i & 3k \le i < 3(k+1) \\ & (u,v) \in \{(0,1),(1,0),(1,1)\} \\ F_k(u,v) & \text{otherwise} \end{cases}$$

where α_k is a scaling factor that varies according to the classification mentioned above. Based on many experiments, α_k is set to be 2, 6 and 9 for class 1, class 2 and class 3, respectively.

The watermarked image is obtained by

$$f'(x,y) = \bigcup_k f'_k(x,y) = \bigcup_k \mathbf{IDCT}\{F'_k(u,v),\ 0 \le u,v < 8\}$$

We embed a part of the watermark in the first three low-frequency AC components in each block in order for the watermark to be robust. Experiments indicate that if more DCT coefficients are modified for watermark embedding, it may have a negative effect on either robustness or invisibility.

Watermark detection: The watermark can be detected by using a correlation technique. Let $F_k^*(u,v)$ denote the DCT coefficients of the corrupted watermarked image in block B_k. The corrupted watermark W^* is extracted by

$$W^* = \bigcup_k W_k^* \quad W_k^* = \{x_i^*, 3k \le i < 3(k+1)\}$$

$$W_k^* = F_k^*(u,v) - F_k(u,v) \quad (u,v) \in \{(0,1),(1,0),(1,1)\}$$

where x_i^* is the corrupted version of x_i, $0 \le i < n$.

To determine if a watermark exists, we compute the correlation between W^* and W:

$$corr(W^*,W) = \sum_{i=0}^{n-1} x_i^* \cdot x_i \Big/ \sum_{i=0}^{n-1} (x_i^*)^2$$

If $corr(W^*,W) > T_3$, it indicates that a watermark exists in the tested image. Experiments indicate that the response of the detector to false marks does not exceed four. So, T_3 is selected to be 6 here.

Simulation results and conclusions: To test the proposed algorithm, we generate 2000 random number sequences that have the normal distribution $N(0,1)$. The 1010th sequence is utilised as the true watermark. The experimental results on a 'Lena' image of 256 × 256 are shown in Figs. 2 and 3.

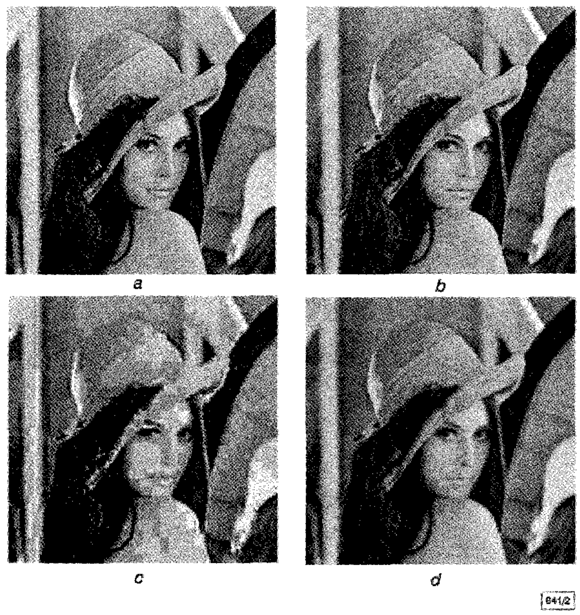

Fig. 2 *Demonstrations of invisibility and robustness*

Invisibility:
a Original 'Lena' image
b Watermarked image
Both are of 256 × 256
Robustness:
c Compressed watermarked image at 0.24 bit/pixel, $PSNR$ = 26.3dB
d Mean-filtered watermarked image

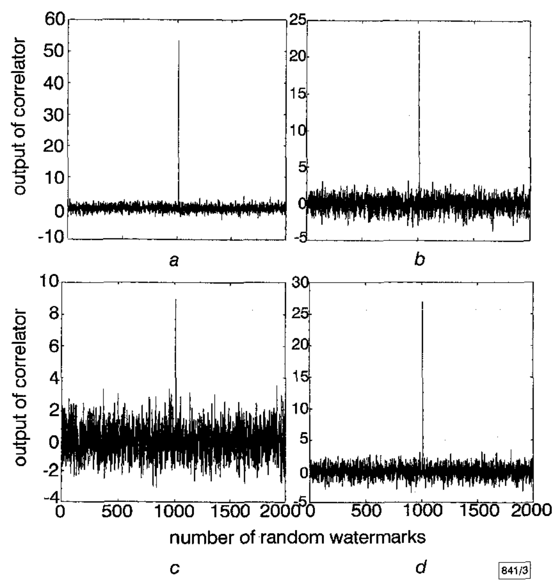

Fig. 3 *Demonstration of robustness*

a Response of detector to Fig. 2*b*
b Response of detector to subsampled watermarked image
c Response of detector to Fig. 2*c*
d Response of detector to Fig. 2*d*

Fig. 2 demonstrates that the watermark generated with the proposed algorithm is invisible, where *a* and *b* are the original and watermarked images, respectively. The response of the detector to the watermarked image is shown in Fig. 3*a*, where the response to the true watermark is much stronger than that to the false watermarks. Fig. 3*b* shows the response of the detector to a subsampled (2:1 in both horizontal and vertical directions) watermarked image. Fig. 2*c* shows the reconstructed watermarked image after JPEG compression (0.24bit/pixel, 26.3dB in PSNR) and Fig. 3*c* shows the corresponding response of the detector. Even though the watermarked image is seriously distorted, the watermark can still be detected robustly. Fig. 2*d* shows the watermarked image after 3×3 mean-filtering and Fig. 3*d* shows the corresponding response of the detector. Figs. 2 and 3 demonstrate the robustness of the watermark against subsampling, compression, and lowpass filtering, respectively.

In summary, we propose an adaptive watermarking algorithm in this Letter. The main contribution is to apply a classification procedure based on visual masking to watermarking embedding. The watermark generated by the proposed algorithm exhibits the following merits: (i) very robust and invisible, (ii) easy implementation, (iii) compatible with most compression standards. Simulation results support these observations.

16 February 1998

Electronics Letters Online No: 19980545

Jiwu Huang (*Department of Electronic Engineering, Shantou University, GD 515063, People's Republic of China*)

Yun Q. Shi (*Department of ECE, New Jersey Institute of Technology, University Heights, Newark, JN 07102, USA*)

Corresponding author: Yun Q. Shi

E-mail: shi@tesla.njit.edu

References

1 COX, I.J., KILIAN, J., LEIGHTON, T., and SHAMOON, T.: 'Secure spread spectrum watermarking for multimedia', *IEEE Trans. Image Process.*, 1997, **6**, (12), pp. 1673–1687

2 WATSON, A.B.: 'DCT quantization matrices visually optimized for individual images'. SPIE: Human vision, visual processing and digital display IV, 1993, Vol. 1913, pp. 202–216

3 JAYANT, N., JOHNSTON, J., and SAFRANEK, R.: 'Signal compression based on model of human perception', *Proc. IEEE*, 1993, **81**, (10), 1385–1422

Segmentation-based hybrid coding using luminance masking

Jiwu Huang and Yun Q. Shi

The authors propose a segmentation-based hybrid image coding algorithm taking the features of the human visual system into account. A novel segmentation criterion based on luminance masking is presented. The image is segmented into blocks of different sizes and different properties. Block mean, polynomial approximation and discrete cosine transform coding are then applied to the blocks accordingly.

Introduction: Segmentation-based coding has received extensive attention in MPEG-4 activities in recent years for achieving content based manipulation of video data and a high compression ratio. The quadtree based technique is one of the most commonly used segmentation techniques, owing to its simplicity and efficiency. To measure the smoothness of a block and then decide whether the block should be divided or not, some criteria are needed. Block variance, which measures mean square deviation (error) from block mean, is the most frequently used criterion.

It is well known that mean square error does not always closely match the features of the human visual system (HVS). This implies that block variance is not the best criterion for quadtree based segmentation. Motivated by this observation, we propose a luminance masking based criterion for segmentation. The effect of masking is to decrease the error in recognition during segmentation. Based on this criterion, a segmentation based hybrid coding algorithm is presented. The performance of the algorithm is verified experimentally.

Overview of proposed algorithm: The algorithm is implemented in three steps: segmentation, classification and encoding, as shown in Fig. 1.

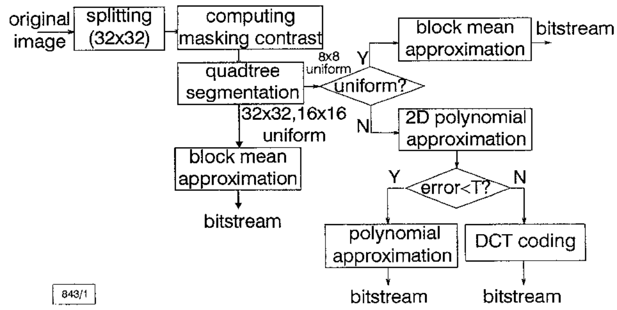

Fig. 1 *Proposed algorithm*

After the original image has been split into non-overlapping blocks of 32×32, the masking contrast of each block is computed, based on a visual model. This contrast is used as a metric to measure uniformity of the block. If the block is considered to be nonuniform, it will be split into four quadrant blocks. This procedure is repeated until either the block size reaches 8×8 or the block is classified as uniform. Therefore, once the segmentation is completed, the image is a union of many square blocks with three possible sizes. A two-level classifier is then applied to the 8×8 blocks. By measuring the contrast and the polynomial approximation error, we classify the 8×8 blocks into the following three categories: uniform, smooth and non-smooth blocks. To compress the image, block means, 2D polynomials and discrete cosine transform (DCT) coding are applied to the uniform blocks of variable sizes, smooth and non-smooth blocks of 8×8, respectively. A more detailed description of the algorithm is presented below.

Luminance masking based segmentation and classification: Using hierarchical segmentation and classification, an original image will be divided into the following five subsets:

(i) S_1: uniform blocks of 32×32
(ii) S_2: uniform blocks of 16×16
(iii) S_3: uniform blocks of 8×8
(iv) S_4: smooth blocks of 8×8
(v) S_5: non-smooth blocks of 8×8

A block with a variation in grey level can be viewed as a uniform background with a superimposed signal. If the amplitude of the signal is lower than the detection threshold of the HVS, the block can be considered to be uniform. Due to the amplitude nonlinearity of the HVS, the threshold depends on the background and the features of the signal. According to the features of the HVS, the brighter the background, the higher the luminance threshold (luminance masking [1]). Based on this idea, we discuss a luminance masking based segmentation criterion.

According to Weber's law [2], the luminance threshold ΔI is given by $\Delta I = 0.02I$, where I denotes the background luminance. Further research has indicated that the luminance threshold increases more slowly than is predicted by Weber's law [3]. Some more accurate contrast sensitivity functions (CSFs) have been presented in [3, 4]. According to [4], the luminance threshold can be expressed as

$$\Delta I = I_0 \cdot \max\{1, (I/I_0)^\alpha\} \qquad (1)$$

where I_0 is the detection threshold when $I = 0$ and α is a constant of ~ 0.6–0.7.

Taking account of the features of the HVS, for a block B_k of $n \times n$ in an image $f(x, y)$, we propose the following metric to measure uniformity, based on eqn. 1.

$$d(B_k) = \frac{1}{n^2} \sum_{(x,y)\in B_k} w(m_k) \cdot \frac{|f(x,y) - m_k|}{m_k} \qquad (2)$$

where m_k is the mean of B_k and $w(B_k)$ is the weight determined using eqn. 1:

$$w(m_k) = (1/m_k)^\alpha \qquad (3)$$

If $d(B_k) < T_1$, where T_1 is a threshold, B_k is considered to be uniform (S_1 or S_2). Experiments have demonstrated that using this criterion leads to better performance than is achieved by using block variance as the criterion.