PERCEPTUALLY TUNED ROBUST WATERMARKING SCHEME FOR DIGITAL VIDEO USING MOTION ENTROPY MASKING

Seong-Whan Kim*, Shan Suthaharan**, Heung-Kyu Lee*, K.R. Rao***

Department of Computer Science*,
Korea Advanced Institute of Science and Technology (KAIST),
E-mail: {video,hklee}@casaturn.kaist.ac.kr

Department of Computer Science**,
Tennessee State University, Nashivile, TN 37221, USA.

Department of Electrical Engineering***,
University of Texas at Arlington, USA.

Abstract—
We present a watermarking scheme for digital videos that are based on human visual system characteristics. Our watermarking scheme inserts perceptually invisible watermark in discrete cosine transform (DCT) domain. We have shown that the proposed scheme provides better results than two other popular schemes both in transparency and robustness.

I. Introduction

A digital watermarking of electronic content is to insert some information into electronic contents such as images, audio and video for copyright protection. Cox proposed a watermarking scheme that is based on the spread spectrum (SS) communication [1]. In this scheme, DCT is performed on a whole image and then the watermark is inserted (added) in a predetermined range of low frequency AC components. The inserted watermark signal consists of a sequence of real numbers that are normally distributed and it is scaled according to the signal strength of the frequency components. Podilchuk and Zeng (P&Z) have recently extended this scheme using Watson visual model to adapt the watermark to each image and they believed that it can provide a maximum length and maximum power watermark [2], [4]. We improved their scheme in two aspects. Firstly we incorporated excitatory-inhibitory interaction between cells into our entropy masking model in order to maximize the watermark further. Secondly, in P&Z's scheme, watermarked image sometimes shows image impairments because normal distributed watermark values will make the pixel values exceed the just noticeable difference (JND). We used bounded normal (BN) distribution which yield only the values between -1 and 1 to overcome this problem (drawback).

II. Entropy Masking Model for Maximal Watermark

Our perceptual watermarking scheme inserts invisible watermark in the DCT domain and it can be used in the Moving Picture Experts Group (MPEG) compression scheme. As shown in Figure 1, the watermark

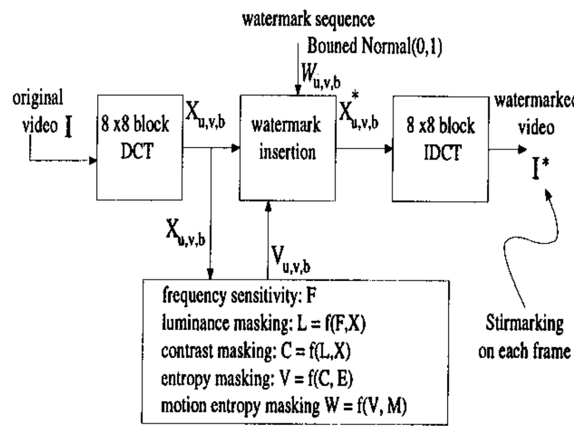

Fig. 1. Our perceptual watermarking scheme.

insertion uses JND to make the watermark which is imperceptible and also has maximal strength. Watson modelled three different properties of the human visual system: frequency sensitivity ($F_{u,v}$), luminance masking ($L_{u,v,b}$), and contrast masking ($C_{u,v,b}$) for each DCT coefficients in every block as in the following equation [4].

$$L_{u,v,b} = F_{u,v}(\frac{X_{0,0,b}}{X_{0,0}})^{\alpha} \quad (1)$$

$$C_{u,v,b} = max[L_{u,v,b}, |X_{u,v,b}|^{\beta_{u,v}}(L_{u,v,b})^{1-\beta_{u,v}}] \quad (2)$$

where $X_{0,0,b}$ is a DC coefficient of the block b, $X_{0,0}$ is a DC coefficient corresponding to the mean luminance of the display, $X_{u,v,b}$ is a (u,v)-th DCT coefficient of the block b, and α and $\beta_{u,v}$ are set to 0.649 and 0.7 to control the degree of luminance sensitivity and contrast sensitivity, respectively.

In human visual system, horizontal and amacrine cells transmit signal to the neighbour bipolar and ganglion cells, which inhibit their responses (i.e. lateral inhibition) [5]. Watson suggested an entropy masking model in his paper [6], and we made a lateral inhibition model using a entropy masking $V_{u,v,b}$ as in the

0-7803-5123-1/99/ $10.00 © 1999 IEEE

following equation

$$V_{u,v,b} = max[C_{u,v,b}, \ \ |C_{u,v,b}| \cdot (E_{u,v,b})^{\gamma}] \quad (3)$$

$$E_{u,v,b} = \sum_{x \in N(X_{u,v,b})} p(x) \cdot \log \frac{1}{p(x)} \quad (4)$$

where $E_{u,v,b}$ is the entropy of $\mathrm{N}(X_{u,v,b})$ which is a set of $X_{u,v,b}$'s eight neighbours. We used 0.5 for γ. If the entropy is large, the $X_{u,v,b}$ will get much inhibitory effect from its neighbours, so that it will increase its perceptual threshold $C_{u,v,b}$. For computational efficiency, we can use the variance of $\mathrm{N}(X_{u,v,b})$ instead of the entropy of $\mathrm{N}(X_{u,v,b})$.

As the number of motion increases in a video signal, the human visual system decrease its sensitivity to the motions. We can model this as a motion entropy masking $W_{u,v,b}$ which is defined as follows

$$W_{u,v,b} = max[V_{u,v,b}, \ \ |V_{u,v,b}| \cdot M^{\eta M_b}] \quad (5)$$

$$M_b = p(C(b)) \cdot \log \frac{1}{p(C(b))} \quad (6)$$

where M_b is the motion information (eg. motion vector) of a set $C(b)$ which the block b is belongs to, and we set 24 sets for C. Motion information for each set C is computed, and sumed up to make the motion entropy for a frame: $M = \sum M_b$. η is set to 1.1 for our experimentation.

After identifying the threshold value for each DCT coefficients, we insert source based watermark into the coefficients as in $X^{*}_{u,v,b} = X_{u,v,b} + JND_1 \cdot w_{u,v,b}$, where $X^{*}_{u,v,b}$ refers to the watermarked DCT coefficients, $w_{u,v,b}$ is the sequence of watermark values which take zero when $X_{u,v,b} \leq JND_1$ [2]. We used $V_{u,v,b}$ for JND_1 for our source watermark sequence. After that, we insert a destination watermark into the coefficients as in $X^{**}_{u,v,b} = X^{*}_{u,v,b} + JND_2 \cdot w_{u,v,b}$, where $X^{**}_{u,v,b}$ refers to the destination watermarked DCT coefficients, and JND_2 is set to $W_{u,v,b} - V_{u,v,b}$.

We use BN distribution to generate a watermark $w_{u,v,b}$. Watermark which is generated from a normal distribution N(0,1) sometimes results in values that exceed the JND, which in turn makes image impairment as will be shown in our experiment.

III. Simulation

We show that our watermarking is transparent and robust to video compression. We used MPEG-1 video coding with 44:1 compression ratio resulting in 22.1 dB (PSNR) on avearge. After compression, the watermarked image shows similar image quality to the non-watermarked image as shown in Figure 2. The detection response sharply drops its value after video compression, but still it shows positive response as shown in Figure 3. We compared our video watermarking schemes with Swanson's scheme [7], and our scheme shows more gap in the response of watermarked and non-watermarked images. Our scheme also shows considerable gain in the frame 7, 12, 18, and 23 because our scheme exploits motion information by using motion entropy.

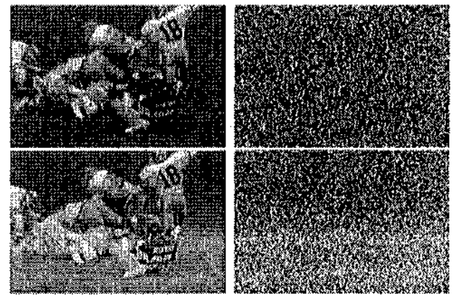

Fig. 2. MPEG (CR=44:1,PSNR=22.1 dB) compression tolerance: (a) watermarked Football after MPEG, and (b) difference, (c) no watermarked Football after MPEG, and (d) difference.

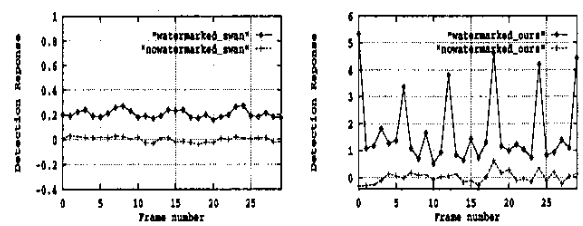

Fig. 3. Detection response for watermarked Football and no-watermarked Football after MPEG coding for (a) Swanson scheme and (b) Our scheme.

References

[1] I. J. Cox, J. Kilian, T. Leighton, and T. Shannon, "Secure spread spectrum watermarking for images, audio, and video," *in IEEE Proc. Int. Conf. Image Processing*, 1996, vol.3, pp. 243-246.

[2] C. I. Podilchuk and W. Zeng, "Image adaptive watermarking using visual models," *IEEE J. Selected Areas in Communication*, vol. 16, no. 4, pp. 525-539, 1998.

[3] J. Fridrich, "Combining low-frequency and spread spectrum watermarking," *in Proc. SPIE Int. Symp. Opt. Sci. Eng. and Inst.*, 1998.

[4] A. B. Watson, "DCT quantization matrices visually optimized for individual images," *in Proc. SPIE Conf. Human Vision, Visual Processing, and Digital Display VI*, 1995.

[5] R. L. De Valois and K. K. De Valois, Spatial vision, *Oxford University Press*, 1990.

[6] A. B. Watson, R. Borthwick, and M. Taylor, "Image quality and entropy masking," *in Proc. SPIE Conf. Human Vision, Visual Processing, and Digital Display VI*, 1997.

[7] M. D. Swanson, B. Zhu, and A. H. Tewfik, "Multiresolution scene-based video watermarking using perceptual models," *IEEE J. Selected Areas in Communication*, vol. 16, no. 4, pp. 540-550, 1998.

[8] F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn, "Attacks on copyright marking systems," *in the Second Workshop on Information Hiding*, 1998, pp. 211-214.