

영상 워터마킹을 위한 엔트로피 마스크 모델

김 성 환[†] · Shan Suthaharan^{††}

요 약

본 논문에서는 정지 영상과 동영상 콘텐츠에 적용하기 위한 새로운 워터마크 설계 가이드라인을 제공한다. 본 논문에서 제안하는 워터마크 설계 방법은 인간 시각 시스템 내의 시각 세포들의 기본적인 동작 특성인 상호간섭 및 배제 작용을 이용하여, 각각의 영상 콘텐츠의 특성을 반영하여 최대한의 워터마크 강도를 구하는데 있다. 영상 콘텐츠의 최대 워터마크를 결정할 수 있다면, 자연스럽게 워터마크의 강인성이 증가되어 일반적인 영상처리 공격이나 MPEG 공격에도 강인하게 동작할 수 있다. 정지영상에서는 여러 단계의 에지 성분이 나타나게 됨에 따라 인간 시각 시스템의 민감도가 급감하게 되고, 동영상에서는 움직임이 세밀해지고 또는 많아질수록 민감도가 급감한다는 점을 모델링하기 위해, 본 논문에서는 정지 영상에서의 엔트로피 마스크와 동영상에서의 움직임 엔트로피 마스크를 이용하여 민감도 저하 현상을 모델링 하였다. 또한, (움직임) 엔트로피 마스크 모델을 기존의 워터마크 기법에 적용한 결과를 통해, 일반적인 워터마크 기법에 엔트로피 마스크를 더함으로써, 워터마크의 투명성을 높이고, 자연스럽게 워터마크의 강건성을 높일 수 있음을 보였다.

An Entropy Masking Model for Image and Video Watermarking

Seong-Whan Kim[†] · Shan Suthaharan^{††}

ABSTRACT

We present a new watermark design tool for digital images and digital videos that are based on human visual system (HVS) characteristics. In this tool, basic mechanisms (inhibitory and excitatory behaviour of cells) of HVS are used to determine image dependent upper bound values on watermark insertion. This allows us to insert maximal allowable transparent watermark, which in turn is extremely hard to attack with common image processing, Motion Picture Experts Group (MPEG) compression. As the number of details (e.g. edges) increases in an image, the HVS decrease its sensitivity to the details. In the same manner, as the number of motion increases in a video signal, the HVS decrease its sensitivity to the motions. We model this decreased sensitivity to the details and motions as an (motion) entropy masking. Entropy masking model can be efficiently used to increase the robustness of image and video watermarks. We have shown that our entropy-masking model provides watermark scheme with increased transparency and henceforth increased robustness.

키워드 : 마스크 모델(Masking Model), 엔트로피(Entropy), 이미지 워터마킹, 비디오 워터마킹(Video Watermarking), 최대성, Multimedia, Real Time Steaming, Wireless LAN

1. Introduction

A digital watermark, or watermark in short, is an invisible mark inserted in digital media such as digital images, audio and video so that it can later be detected and used as evidence of copyright infringement. However, insertion of such invisible mark should not alter the perceived quality of the digital media (it is the transparency requirement) while being extremely robust to attack (it is a robust requirement) and being impossible to insert another watermarks for rightful ownership (it is a maximal capacity requirement) [1-4].

recent years, an overwhelmingly large amount of research work has been done in watermarking digital images for copyright protection. Cox [1] 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 embedded in a predetermined range of low frequency AC components. The embedded 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. It is a simple watermarking scheme with perceptual weighting consideration. Consequently, Podilchuk and Zeng (P&Z) [2] have extended this scheme using Watson visual model to adapt the watermark to each image and they believed that it could provide a maximum length and maximum power

* This work was supported by the Korea Research Foundation Grant (KRF 2002-003-D00295).

[†] 성 회 원 : 서울시립대학교 컴퓨터과학부 교수

^{††} 키 회 원 : University of North Carolina, Greensboro, NC, US

논문접수 : 2003년 5월 21일, 심사완료 : 2003년 8월 20일

watermark [2, 3].

We incorporated excitatory-inhibitory interaction between cells into our model in order to maximize the watermark. As the number of details (e.g. edges) increases in an image, the HVS decrease its sensitivity to the details. In the same manner, as the number of motion increases in a video signal, the HVS decrease its sensitivity to the motions. We model this decreased sensitivity to the details and motions as an (motion) entropy masking. In this paper, we have discussed only two of the most popular schemes proposed by Cox [1] and P&Z [2] and compared our proposed scheme with them, because our entropy masking model for watermark design is orthogonal to most of the watermarking schemes. Entropy masking model can be efficiently used to increase the robustness of image and video watermarks. We have shown that our entropy-masking model provides watermark scheme with increased transparency and henceforth increased robustness.

In section 2, we will introduce the basic idea and details of entropy masking model, and show experimental watermarking results using entropy-masking model with conclusion in section 4.

2. Entropy Masking Model for Maximal Watermark

Watermark insertion process uses human visual system properties to make the watermark imperceptible with maximal strength. We use Watson visual model as a baseline model to achieve this task. 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 as in the following equation [5].

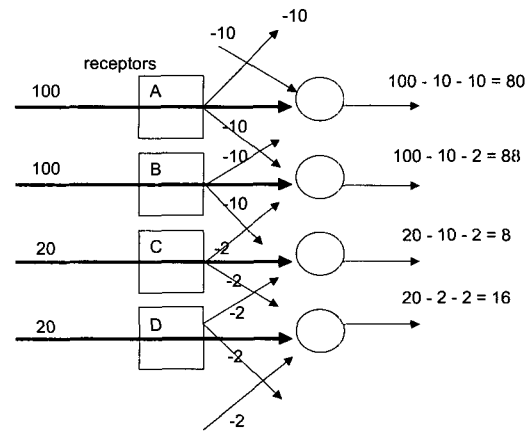
$$L_{u,v,b} = F_{u,v} \left(\frac{X_{0,0,b}}{X_{0,0}} \right)^\alpha \tag{1}$$

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

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.

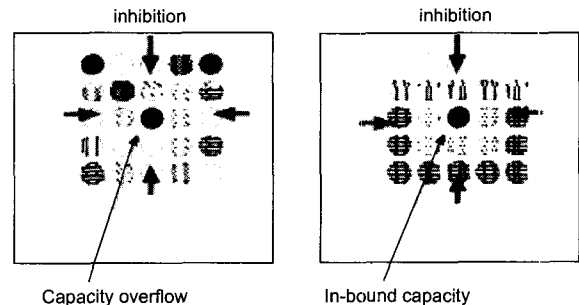
In this paper, we introduce a masking model, which exploits HVS cell behaviour. In HVS, lights on the eye ball will first meet retina cells (cones and rods), and they will

meet horizontal/amacrine cell and bipolar/ganglion cells. These cells show inhibitory and excitatory reactions depending on the input strength. As shown in the (Figure 1), horizontal and amacrine cells (denoted as A, B, C, and D in (Figure 1)) usually transmit signal to the neighbour bipolar and ganglion cells (denoted as circle in (Figure 1)), which inhibit their responses (i.e. lateral inhibition) [6]. In this scenario, we will not focus excitatory effect.



(Figure 1) Inhibitory-Excitatory Behaviour of HVS Cells

Watson suggested an entropy-masking concept in his paper [7], and we interpret the lateral inhibition in different way, in such a way that a group different strength cells will show complex strength patterns. As shown in (Figure 2), similarity or statistical non-uniformness of strength will show simple patterns, and naturally introduce an entropy concept. If the complexity of the pattern can be handled (in-bound capacity) in the HVS cells, they will be perceived.



(Figure 2) Lateral Inhibition and Cell Entropy

We made a lateral inhibition model using entropy masking as in the following equation

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

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

where $E_{u,v,b}$ is the entropy of $N(X_{u,v,b})$ 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 a variance of eight neighbours instead of the entropy of them. After identifying the threshold value for each DCT coefficients, we insert watermark into the coefficients as in $X_{u,v,b}^* = X_{u,v,b} + \text{JND } w_{u,v,b}$, where $X_{u,v,b}^*$ refers to the watermarked DCT coefficients, $w_{u,v,b}$ is the sequence of watermark values which takes zero when $X_{u,v,b} \geq \text{JND}$ [2]. We used $V_{u,v,b}$ for JND for our watermark sequence.

For video signals, we extend our model to quantify how much distortion we can put into each frame in a video sequence. As each moving blocks gets faster above some specified threshold, the spatial frequency sensitivity will decrease very rapidly. As shown in the (Figure 3), as the number of motion increases in a video signal, the HVS decrease its sensitivity to the motions.



(Figure 3) Increased motion decreases motion sensitivity

We can model this as motion entropy masking $W_{u,v,b}$ which is defined as follows.

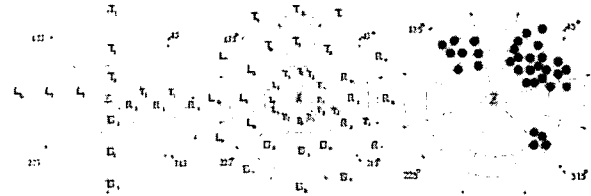
$$W_{u,v,b} = \max [V_{u,v,b} | V_{u,v,b} | \cdot M^{\mu Mb}] \quad (5)$$

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



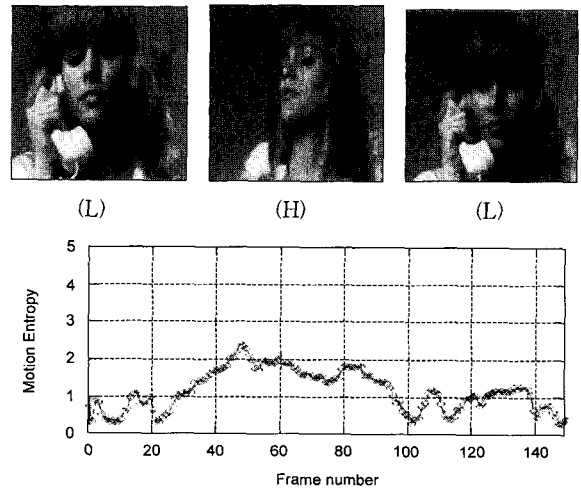
(Figure 6) Lateral Inhibition and Cell Entropy

where M_b is the motion information of a set $C(b)$ which the block belongs to, and there are 36 sets for C as shown in (Figure 4). Motion information for each set C is computed, and summed up to make the motion entropy for a frame : $M = \sum M_b$. η is set to 1.1 for our experimentation.

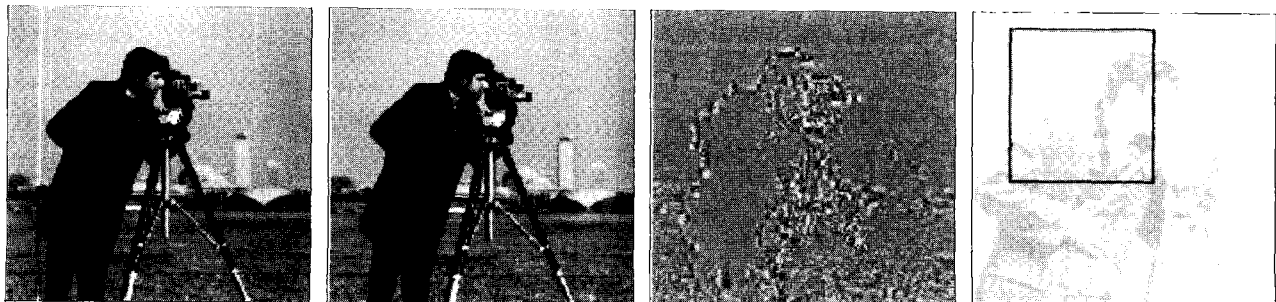


(Figure 4) Motion Class C(b)

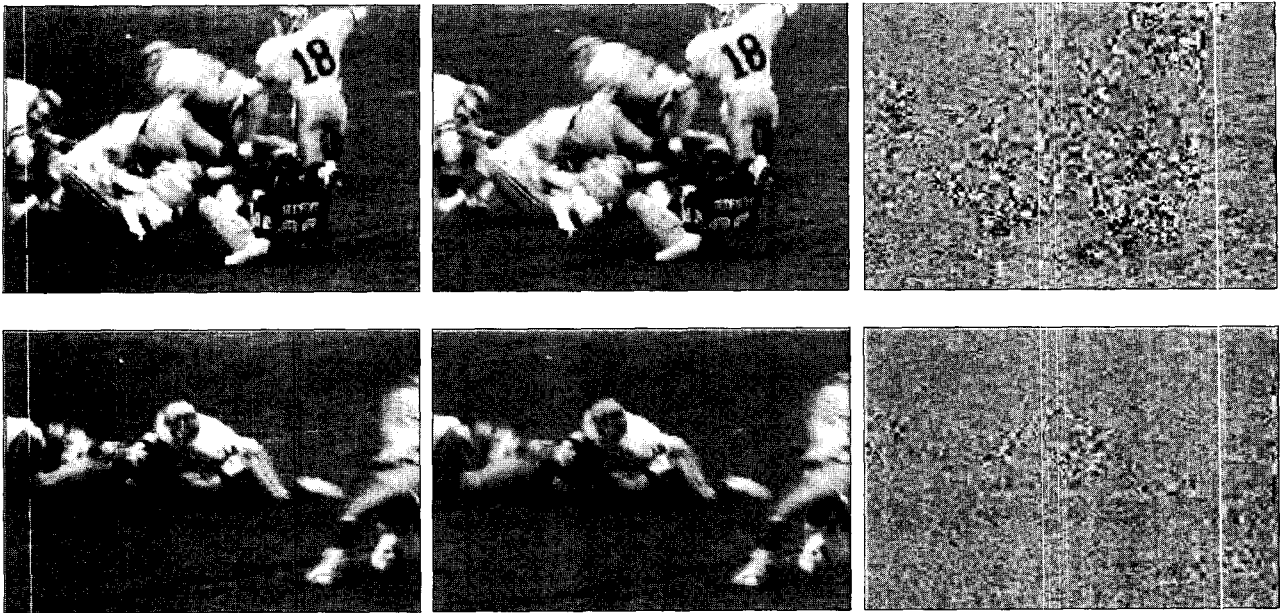
As shown in (Figure 5)(b), the motion entropy increases and decreases as the motion complexity does (Figure 5)(a) in the video sequences. The entropy can be low or high, and also it can be fluctuate or stable which reflects the image characteristics. We used the MPEG-1 motion estimation techniques to find motion vectors for each macro block. After finding motion vectors, we converted them to (magnitude, angle) notation for motion entropy computation.



(Figure 5) (a) Susie Test Sequence and (b) Motion Entropy Trace



(Figure 6) Lateral Inhibition and Cell Entropy



(Figure 7) Football Original Frame and Motion Entropy Identified Area

3. Simulation

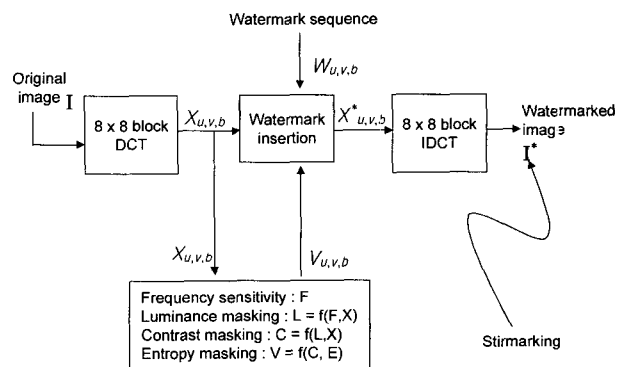
We applied entropy model to sample images as shown in (Figure 6). We can see that high entropy usually occurs in the edge regions because there are many lateral inhibition reactions and results in high entropy states.

As in the image case, we applied motion entropy model to sample video signals as shown in (Figure 7). We can see that high motion entropy occurs in the area where much motion occurs. Those areas are usually hard to perceive.

We compared our watermarking scheme with P&Z's and SS scheme in terms of imperceptibility, maximality, and robustness. We used a simple watermarking model as shown in (Figure 8). (Figure 9) shows the watermarked images for *Cameraman* after applying three watermarking schemes : SS scheme, P&Z's scheme, and ours. SS Scheme and ours are almost same image quality as original image, but P&Z's scheme shows some image impairment in some edge portions. However, the signal to noise ratio (SNR) is 29.73 dB for Cox, 39.06 dB for P&Z, and almost 41.89 dB for ours. In the figure, we can also see the difference image that is actually inserted watermark. The SS scheme has very little watermark, which implies that some malicious users can insert invalid watermark. In this aspect, P&Z's scheme and our proposed method show maximal watermark, and they satisfy the maximal capacity requirement.

To compare the robustness of our scheme, we used Stirmark that is known to defeat all the watermarking system

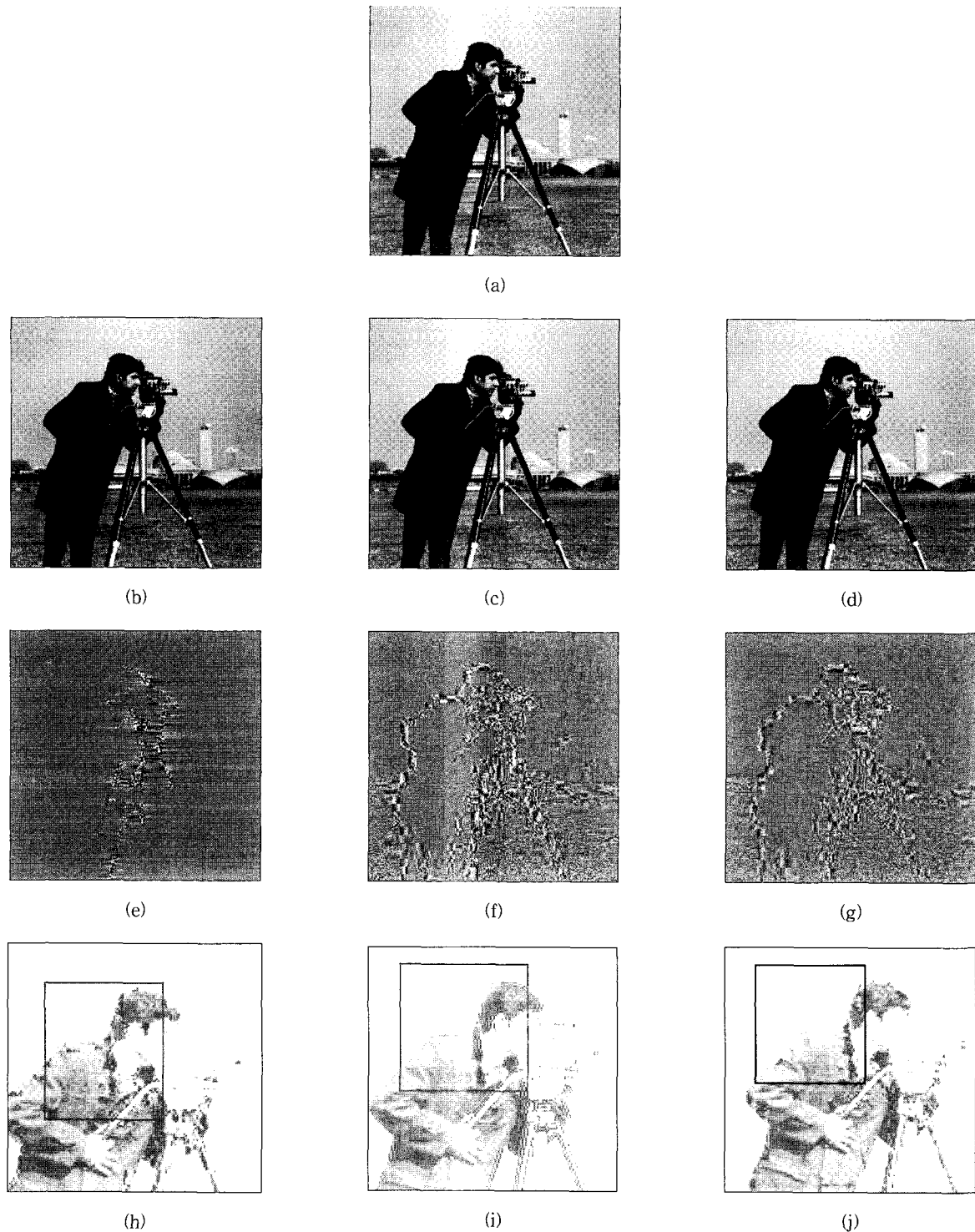
in the world. Stirmark basically simulates a re-sampling process : it introduces the same kind of errors into an image as we would expect if we printed an image on a high-quality printer and then scan the image again with a high-quality scanner [8]. Resulting images after Stirmarking usually do not show any degradation, but there is significant degradation of watermark in watermarked images. Our scheme shows relatively higher detection value (0.515) than Cox scheme (-0.566) and P&Z's scheme (-0.130) for *Cameraman* image.



(Figure 8) Simple watermarking test model

4. Summary

Inserting maximal allowable transparent watermark, which in turn is extremely hard to attack with common image processing, is important aspect of watermark design, We presented a new watermark design tool for digital images



(Figure 9) (a) Original Cameraman image, (b) Watermarked image using SS scheme, (c) Watermarked image using P&Z scheme, and (d) Watermarked image using the proposed scheme, (e) Difference image of SS scheme, (f) Difference image of P&Z scheme, (g) Difference image of the proposed scheme (h) Enlarged image which shows image impairment in SS scheme along edge portion (i) Enlarged image which shows image impairment in P&Z scheme along edge portion, (j) Enlarged image which shows no impairments in the proposed scheme.

and digital videos that are based on entropy concept, and it will be used to maximize the allowance of watermark depending on the image characteristics. As the number of details

(e.g. edges) increases in an image, the HVS decreases its sensitivity to the details. In the same manner, as the number of motion increases in a video signal, the HVS decreases its

sensitivity to the motions. We model this decreased sensitivity to the details and motions as an (motion) entropy masking. We have shown that our entropy-masking model provides watermark scheme with increased transparency and henceforth increased robustness. We have shown that the proposed design tool provides better results for SS and P&Z's schemes with adopting entropy-masking model, both in transparency and robustness.

References

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

[2] C. I. Podilchuk and W. Zeng, "Image adaptive watermarking using visual models," IEEE J. Selected Areas in Comm., 16, pp.525-539, 1998.

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

[4] S. W. Kim, S. Suthaharan, H. K. Lee and K. R. Rao, "A watermarking scheme using human visual model and BN distribution," Proc. of The 6th IEEE international workshop on intelligent signal processing and communication systems, pp.396-400, 1998.

[5] A. B. Watson, "DCT quantization matrices visually optimised for individual images," Proc. SPIE Conf. Human Vision, Visual Proc. and Digital Display VI, 1995.

[6] R. L. De Valois and K. K. De Valois, "Spatial Vision," Oxford University Press, 1990.

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

[8] F. A. P. Petitcolas, R. J. Anderson and M. G. Kuhn, "Attacks on copyright marking systems," the 2nd Workshop on Info. Hiding, pp.211-214, 1998.



김성환

e-mail : sigma@venus.uos.ac.kr

1991년 동국대학교 전산학과(공학사)

1993년 한국과학기술원 전자전산학과
(공학석사)

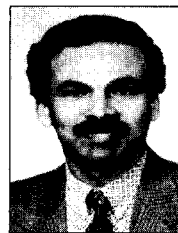
1999년 한국과학기술원 전자전산학과
(공학박사)

1997년~2000년 LG 전자/정보통신 IMT-2000 연구소

2000년~2001년 Cisco Systems Enterprise Wireless Business Unit

2002년~현재 서울시립대학교 컴퓨터과학부 조교수

관심분야 : 무선 패킷 비디오, 인지시각의 수학적 모델링, 워터마킹



Shan Suthaharan

e-mail : ssuthaharan@uncg.edu

1981년 Department of Statistics (BSc with Hons), University of Jaffna, Srilanka

1988년 Department of Computer Science (MSc), Dundee University, United Kingdom

1995년 Department of Computer Science (Ph.D.), Monash University, Australia

현재 Department of Computer Science' University of North Carolina, Greensboro (Head), Associated editor of the Journal of Real Time Imaging

관심분야 : image/video coding and compression, image restoration, multimedia computing and communication and computational intelligence