Image Verification for Digital Rights Management Using Fragile Watermarks Based on a Human Visual Model

Da-Chun Wu[1,*] and Wen-Hsiang Tsai[2]

[1]Department of Computer and Communication Engineering
National Kaohsiung First University of Science and Technology
Kaohsiung, Taiwan 811, Republic of China
dcwu@ccms.nkfust.edu.tw

[2]Department of Computer Science
National Chiao Tung University
Hsinchu, Taiwan 300, Republic of China
whtsai@cis.nctu.edu.tw

Received 1 March 2009; Revised 1 April 2009; Accepted 6 April 2009

Abstract. A novel method for embedding perceptually based fragile watermarks in digital images is proposed. The method is designed for the purpose of image verification and tamper proofing for digital rights management applications. A human visual model is employed to guarantee that modifications in images are imperceptible. A set of quantized contrast functions based on a human visual model is constructed first. Then a given image is partitioned into 3×3 subimages. In each subimage, a range of gray levels with the same contrast as that of the central pixel of the subimage is obtained from the quantized contrast functions. One of the values in the range is chosen as a watermark value to replace the gray value of the central pixel of each subimage. This ensures that the change is imperceptible. To achieve higher security, a pseudo random mechanism is employed to generate random keys for use in generating watermark values. The verification process is proceeded with no reference to the original image. The proposed technique provides a visual inspection tool for detecting and localizing any image alteration. Experimental results have been conducted to show the feasibility of the proposed approach.

Keywords: watermarking, image verification, tamper proofing, data embedding, fragile watermark, human visual model.

1 Introduction

Digitizing information is popular today. Digital media can be reproduced and distributed easily. Slight modifications of digital data (images, video, audio, document images) may be accomplished without noticeable differences. To distinguish a genuine digital medium from a forgery is so an important technique in the applications of digital rights management [1-2]. In the following, we focus on images which are the type of data we deal with in this study. Many new applications about image data modification have been explored, such as copyright protection, tamper proofing, authentication, security message delivering, annotation embedding, etc.

Transparent digital image watermarking is a technique that embeds invisible signals into images. Recently, a number of watermarking methods [3-5] have been proposed, which exploit the sensitivity of the human visual system to guarantee watermark transparency. These methods can be categorized into spatial or frequency masking [6-9]. Such techniques are based on perceptual models and generate so-called perceptual watermarks. One of the detailed surveys of the use of perceptual watermarks for digital images can be found in [10].

The requirements and properties of digital image watermarking methods [11-13] vary with their applications. In copyright protection [14-15], the embedded watermarks must be resistant to common types of attacks, such as JPEG compression, geometric transformations, cropping, scaling, etc. To explore a large amount of space for efficiently embedding data into an image [16] is a prerequisite for the application of annotation embedding.

Tamper proofing applications require watermarking techniques to embed invisible signatures into images for integrity verification and authentication. Any tampering with a watermarked image can be detected by examining the embedded signature in the image. The verification process must be proceeded without referencing the original image and the non-tampered watermarked image. Such applications include trusted camera imaging [17], news imaging, commercial imaging, medical imaging [18], etc. In these applications, only when the integrity of an image is proved by a verification process can the value provided by the image itself be assured. Since the watermarks used in these applications are not robust to any slight modification, we call them fragile water-

*Correspondence Author

marks. There is another kind of fragile watermark with different properties, called semi-fragile watermarks [10], which are designed to be insensitive to simple image manipulations, such as lossy compression, but still sensitive to significant alterations of the context, such as replacing a portion of the image with another. In this study, we developed a novel method for creating fragile watermarks for digital image verification based on a human visual model.

Several watermarking techniques have been reported for image verification. In [19-20], least significant bit (LSB) planes are used for carrying the watermark information. The embedded watermark is not related to the contents of the original image. Another application [21] used check-sums information as the watermark message which is embedded into the LSB plane. In [22], a binary image was embedded into another for image verification. A pre-defined watermark extraction function is utilized to obtain a value for each pixel in the image according to the pixel value. The embedding process modifies the pixel value repeatedly until the function value calculated from the new pixel value matches the desired value in the binary image. Wong [23] partitioned images into blocks and used the LSB plane of each block for embedding watermark information. The information is generated by a cryptographic hash function which uses the pixel values of all pixels in the block and the dimension information of the image as the parameters. Fridrich [24] divided an image into 64×64 blocks, and inserted a watermark value into each block by modifying the middle third of the DCT coefficients of each block. The watermark is related to a secret key, the block number, and the content of the block.

In the proposed approach, we divide an image into 3×3 subimages, and use a set of pre-defined contrast functions based on a human visual model to obtain a range of gray levels with the same contrast as that of the central pixel of the subimage. A watermark generation function is used for generating a value in the range for replacing the gray value of the central pixel of each subimage. This preserves the contrast sensitivity and guarantees that the change is imperceptive. The function is related to a key and two visual thresholds calculated from the contrast functions.

The remainder of this paper is organized as follows. The adopted human visual model is discussed in Section 2. In Section 3, the proposed method for creating fragile watermarks is presented. In Section 4, the process for verifying images is described. And several experimental results are illustrated in Section 5. Finally, some concluding remarks are stated in Section 6.

2 Human Visual Model

The human visual system has been studied in the field of image coding and compression. It is possible to calculate a just-noticeable distortion (JND) threshold for each frequency band in an image. Any change below the JND threshold is considered to be imperceptive [25]. Another model [26] utilizes the Weber's law [27] and the spatial masking factor to form the JND gray value of each pixel. A perceptual image compression model [28] uses both frequency sensitivity and spatial masking to obtain a tolerable-error level (TEL) for each pixel. The JND threshold or the TEL can be directly obtained through the gray value of a pixel and its background intensity. Any modification of the pixel value which does not exceed the gray value ranges formed by the JND threshold or the TEL is regarded as imperceptive. In some watermarking techniques for copyright protection [10][14], the above approaches are used for embedding robust watermarks in images imperceptibly. But the original image need be referenced when extraction of the watermark out of a watermarked image is conducted. This does not meet the prerequisite of the watermarking application of image verification. In this study, we explore a human visual model which can be used in the application of image verification without referencing the original image.

In [29-30], an image compression technique utilizing the human visual model was proposed with consideration about the Weber's law, the visual masking effect, and the Mach band effect [31]. A group of contrast functions that considers the sensitivity of the human eye to individual luminance were constructed. And uniform quantization is applied to each contrast function to produce a corresponding quantized contrast function. Each quantization level in the output of the function expresses a set of input gray values with the same sensitivity to the human eye under a specific background. An illustration of a 12-level quantized contrast function with background luminance 180 is shown in Fig. 1.

In our proposed method, the visual model is modified to represent the visual sensitivity of a 3×3 image. We categorize 3×3 images into four classes according to the standard deviation of the gray values of all pixels except the central pixel in each 3×3 image. The classes are categorized from smooth to edged ones and each class is assigned with a specific number of quantization levels. The criteria used to classify the classes and select the number of the quantization levels are shown in the following:

$$\text{the number of the quantization levels} = \begin{cases} 32 & \text{when} \quad \sigma \le 2.4; \\ 24 & \text{when} \quad 2.4 < \sigma \le 3.6; \\ 16 & \text{when} \quad 3.6 < \sigma \le 4.8; \\ 12 & \text{when} \quad 4.8 < \sigma, \end{cases} \tag{1}$$

where σ is the standard deviation of the gray values of the eight pixels surrounding the central one in a 3×3 image. Larger numbers of quantization levels are used in the class of smooth images while smaller numbers of quantization levels are used in the class of edged images. In Figure 2, the image Lena is used for illustrating the regions of the four classes which are obtained by the above criteria.

Furthermore, for each background gray value b in a 3×3 image (computed as the average of the gray values of the eight pixels surrounding the central one in the image), we build up a contrast function, where the input gray values are in the range from 0 to 255. Next, each contrast function is uniformly quantized to get a quantized contrast functions for each class c according to the number of the quantization levels corresponding to the class. Finally, given the gray value g of the central pixel p of a 3×3 image, its background gray value b, and the image class c, we can find, from the quantized contrast function corresponding to b and c, the corresponding quantization level $L = [g_{\min}, g_{\max}]$ including g, where $g_{\min}$ and $g_{\max}$ are the two extreme values of L and are called the visual thresholds corresponding to g, b, and c. This process of finding the visual thresholds for each 3×3 image may be regarded as a function which we denote as Q in the sequel, and the range L may be taken as the function value $Q(g, b, c)$. The above calculations are image independent, i.e., they are constructed before the watermarking steps begin and can be used for watermarking any image.

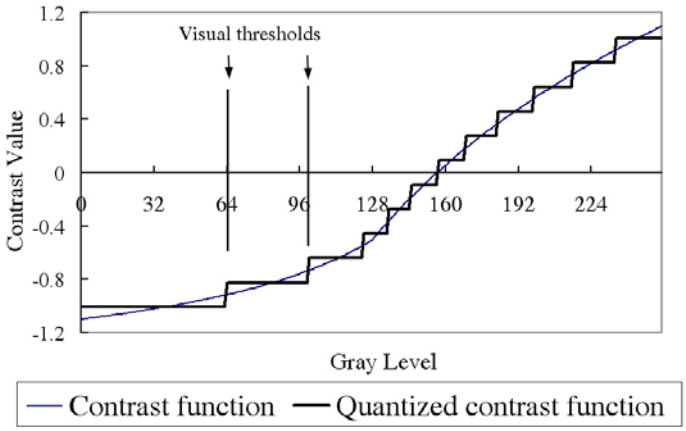


Fig. 1. Contrast function resulting from the condition that the gray value mean of the background is 180 and the corresponding quantized contrast function for 12 quantization levels. The visual thresholds 64 and 98 are picked out when the central pixel's gray value is within the range of [64, 98].

3 Proposed Watermarking Method

In the steps of creating the watermark, we first divide the given image into 3×3 subimages. The central pixel of every 3×3 subimage in the image is used for embedding the watermark. The pixels in each subimage which surround the central pixel are considered as the background of the central pixel in our proposed method. Let c be the class of the subimage, b be the background luminance which is computed as the mean of the gray values of the eight background pixels, and g be the gray value of the central pixel p. A range L=[$g_{\min}$, $g_{\max}$] defined by two visual thresholds $g_{\min}$ and $g_{\max}$ thus can be obtained as the function value $Q(g, b, c)$ where Q is as defined previously. The function value $Q(g', b, c)$ of any gray value g' in the range L is the same as that of $Q(g, b, c)$. From the viewpoint of the adopted human visual model, this means that any gray value in the range L has the same sensitivity as the gray value of the central pixel under the same background b and the same class c. If we substitute the gray value of the central pixel with a value in the range L, the modification will be imperceptible. We therefore embed a watermark value in the subimage by randomly choosing a gray value in the range of L=[$g_{\min}$, $g_{\max}$] to replace the gray value of the central pixel of the subimage. The detail is as described next.

Let W be a watermark generation function for the subimage. W is designed in this study to be related to a given key K, $g_{\min}$, and $g_{\max}$, as follows:

$$W(K,\ g_{\min},\ g_{\max}) = g_{\min} + K_{\bmod (g_{\max}-g_{\min}+1)}, \tag{2}$$

where mod is the modulus operator. Note that the function values of W range from $g_{\min}$ to $g_{\max}$. Then we take the value of W described by Eq. (2) as a watermark value and embed it into the central pixel p by replacing the gray value g of p with the watermark value. To achieve cryptography, a pseudo-random mechanism is used for generating a key for each subimage. As a result, without using the seed of the pseudo-random scheme, the watermark values generated by the function W of the subimages cannot be retrieved correctly. This can avoid easy illicit extraction of the watermark values from a watermarked image.

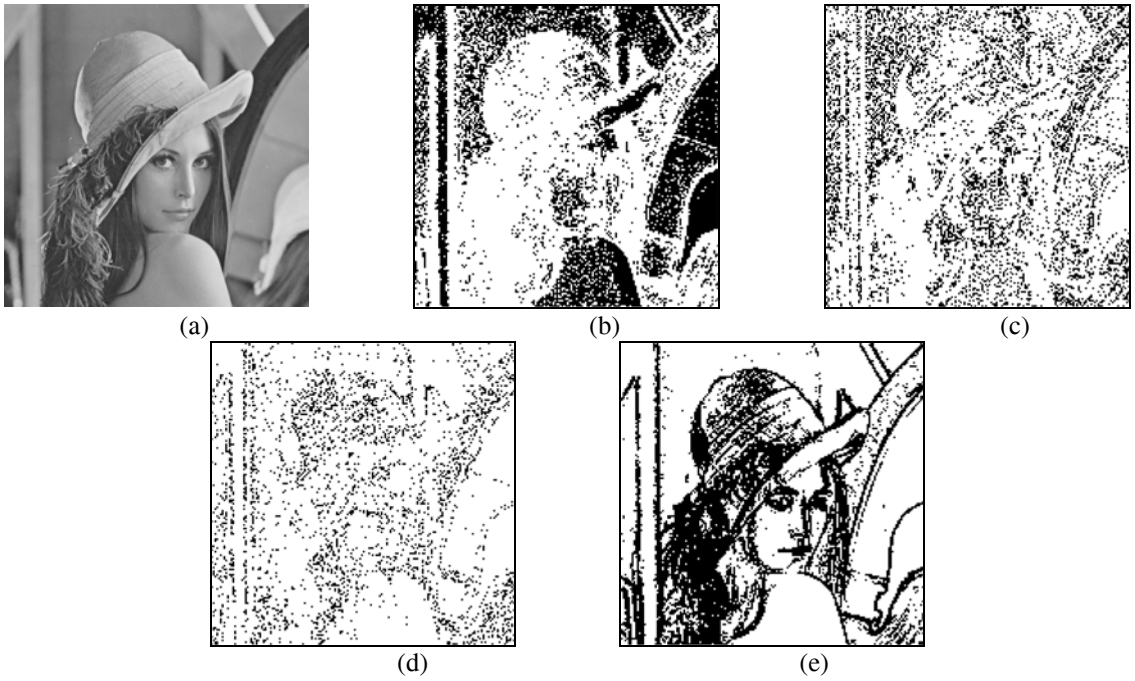


Fig. 2. (a) Illustration image "Lena". (b) The 3×3 regions of Lena with $\sigma \le 2.4$. (c) The 3×3 regions of Lena with $2.4 < \sigma \le 3.6$. (d) The 3×3 regions of Lena with $3.6 < \sigma \le 4.8$. (e) The 3×3 regions of Lena with $4.8 < \sigma$.

4 Image Verification

Now we describe the proposed image verification process which can detect any tampering with the watermarked image produced by the proposed watermarking method described previously. The verification process takes as input a watermarked image which is suspicious of being tampered. The seed for the pseudo-random mechanism which was used in the watermark embedding step must be provided. The key for each 3×3 subimage in the image can be generated correctly by using the seed. For each subimage, assume that c' is the class of the subimage, that g' is the gray value of the central pixel p and that b' is the background luminance which is the mean of the gray values of the eight surrounding pixels. Then, we obtain the visual thresholds $g'_{\min}$ and $g'_{\max}$ of the subimage from the function value of $Q(g',\ b',\ c')$, followed by the computation of the function value $w=W(K,\ g'_{\min},\ g'_{\max})$, where K is the key of the subimage. If the function value w is equal to g', then it is not difficult to figure out that the 3×3 subimage which contains p has not been tampered. On the other hand, if the value w does not match g', we decide that the subimage has been tampered.

For convenience, an image which provides a visual inspection tool for localizing any alteration in the watermarked image is constructed according to the matching results during the verification process. Dark regions in the resulting image indicate those 3×3 subimages in the watermarked image that have been altered.

Table 1. Values of RMSE's and PSNR's of images with embedded watermarks.

Image	RMSE	PSNR(dB)
Baboon	1.99	42.15
Lena	1.63	43.87
Jet	1.58	44.17
Cameraman	1.56	44.28

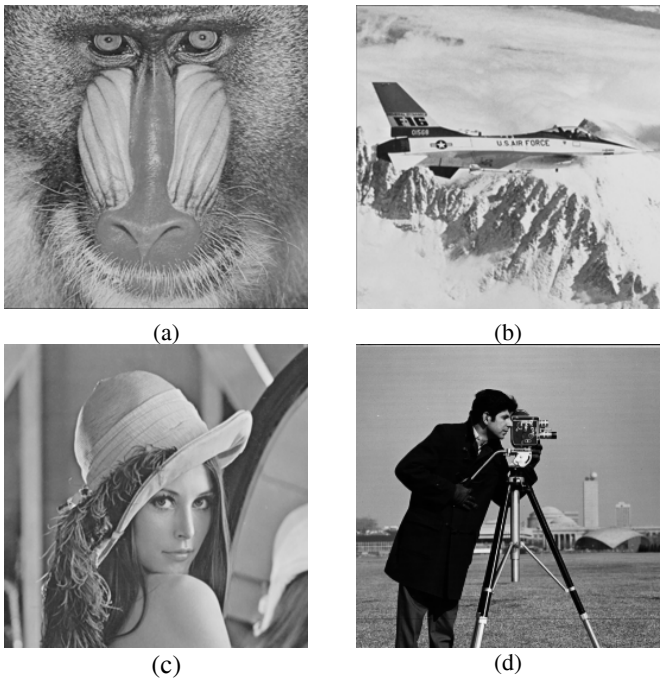


Fig. 3. Tested images. (a) "Baboon", (b) "Jet", (c) "Lena", and (d) "Cameraman".

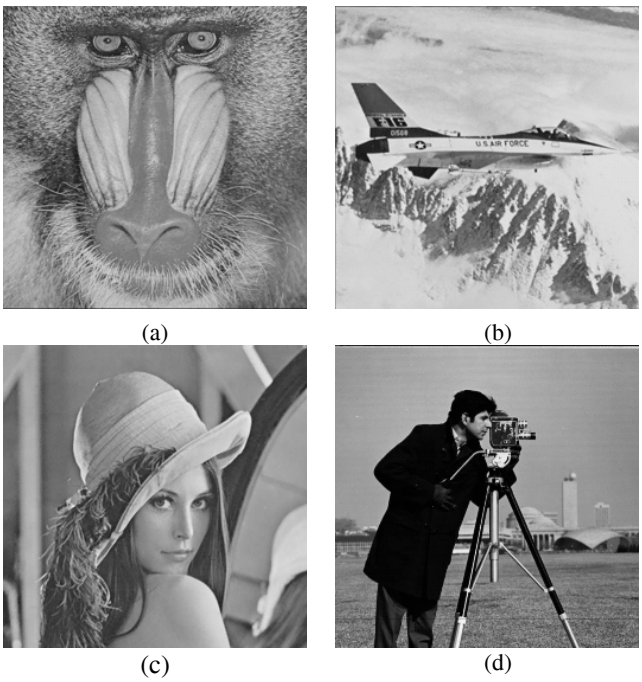
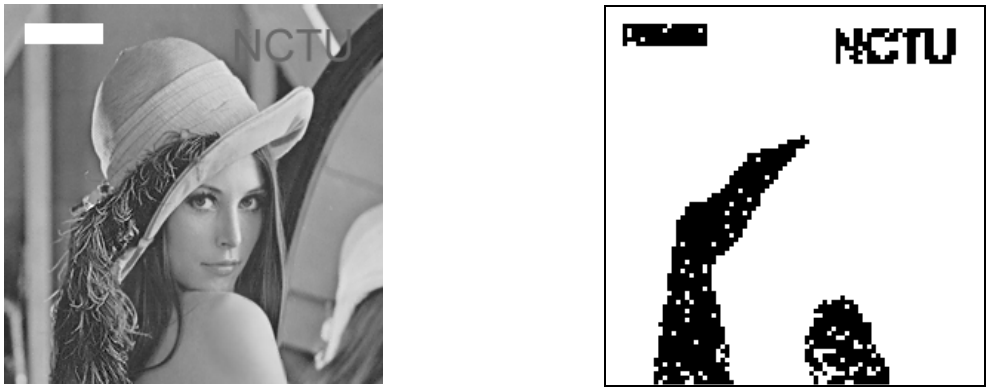


Fig. 4. Images with embedded watermarks. (a) "Baboon", (b) "Jet", (c) "Lena", and (d) "Cameraman".

5 Experimental Results

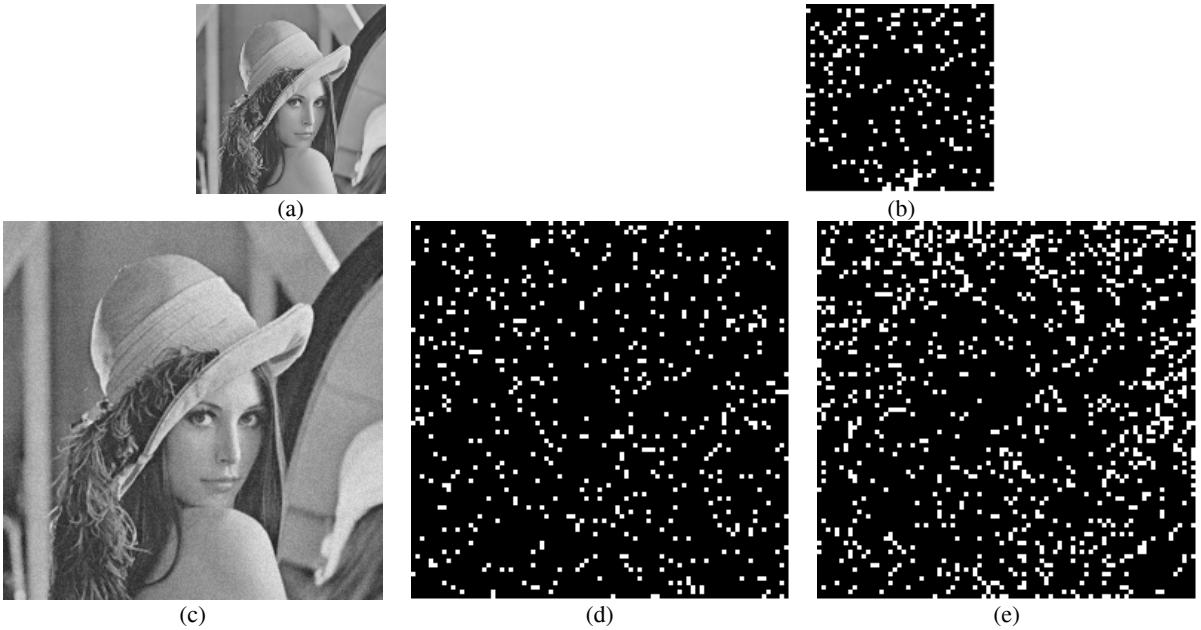
Four images as shown in Fig. 3, each with size 256×256, were used in our experiments. The images resulting from embedding the created watermark values in the subimage central pixels are shown in Fig. 4. The results show that the proposed watermarking method can embed watermarks without noticeable changes. In Table 1, the values of RMSE's and PSNR's of images with embedded watermarks are shown.

Shown in Fig. 5(a) is a tampered image of the watermarked Lena, which results from manipulations of sharpening the hair, smoothing the shoulder, inserting a text, and digging a rectangular region. The resulting image produced by the image verification steps using the tampered image as input is shown in Fig. 5(b). Dark regions indicate the altered regions in the watermarked Lena. The images resulting from tampering the watermarked Lena by resizing and adding Gaussian noise, and those produced by the verification steps (including one produced without using the correct seed) are shown in Fig. 6. In these cases, the resulting images include spread dark regions.



(a) (b)

Fig. 5. (a) The watermarked image manipulated by sharpening the hair, smoothing the shoulder, inserting a text, and digging a rectangular region. (b) The resulting image produced by the image verification steps using Fig. 5(a) as input.



(a) (b)

(c) (d) (e)

Fig. 6. (a) The Lena image after reducing its size to fourth using Fig. 4(a). (b) The image produced by the image verification steps using Fig. 6(a) as input. (c) The Lena image after adding Gaussian noise using Fig. 4(a). (d) The image produced by the image verification steps using Fig. 6(c) as input. (e) The image produced by the image verification steps with an incorrect key using Fig. 4(a) as input.

6 Concluding Remarks

A novel watermarking method for image verification using a human visual model in the gray-valued domain has been proposed. Quantized contrast functions which provide visual thresholds were constructed for use in the watermark embedding process. The proposed method provides a way for embedding transparent fragile watermarks in images. The verification steps detect image tampering in an efficient way without referencing the original image. Any alteration of the watermarked image can be detected and localized. Good experimental results show the feasibility of the proposed method.

References

[1] W. Lee and C. Hwang, "A Forensic Computing System Using a Digital Right Management Technique," in *Proceedings of Fourth International Conference on Fuzzy Systems and Knowledge Discovery*, Vol. 3, pp. 258-262, 2007.

[2] N. Nikolaidis and I. Pitas, "Image and Video Fingerprinting for Digital Rights Management of Multimedia Data," in *Proceedings of International Symposium on Intelligent Signal Processing and Communications*, pp. 801-807, 2006.

[3] A. Koz and A.A. Alatan, "Oblivious Spatio-temporal Watermarking of Digital Video by Exploiting the Human Visual System," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 18, No. 3, pp. 326-337, 2008.

[4] M. Swanson, B. Zhu, A. Tewfik, "Data Hiding for Video in Video," in *Proceedings of IEEE International Conference on Image Processing*, Vol. II, pp. 676-679, 1997.

[5] J. F. Delaigle, C. D. Vleeschouwer, B. Macq, "Watermarking Algorithm Based on a Human Visual Model," *Signal Processing*, Vol. 66, pp. 319-335, 1998.

[6] S.M. Zhu and J.M. Liu, "Adaptive Watermarking Scheme in Hybrid DWT-DCT Transform Based on Human Visual System," in *Proceedings of International Symposium on Knowledge Acquisition and Modeling*, pp. 668-671, 2008.

[7] J.P. Porter and P.K. Rajan, "Image Adaptive Watermarking Techniques Using Models of the Human Visual System," in *Proceedings of the Thirty-Eighth Southeastern Symposium on System Theory*, pp. 354-357, 2006.

[8] B. Girod, "The Information Theoretical Significance of Spatial and Temporal Masking in Video Signals," in *Proceedings of SPIE Human Vision, Visual Processing, and Digital Display*, Vol. 1077, pp. 178-187, 1989.

[9] G. Legge and J. Foley, "Contrast Masking in Human Vision," *Journal of the Optical Society of America A*, Vol. 70, No. 12, pp. 1458-1471, 1990.

[10] R. B. Wolfgang, C. I. Podilchuk, E. J. Delp, "Perceptual Watermarks for Digital Images and Video," *Proceedings of the IEEE*, Vol. 87, No. 7, pp. 1108-1126, 1999.

[11] W. Bender, D. Gruhl, N. Morimoto, A. Lu, "Techniques for Data Hiding," *IBM System Journal*, Vol. 35 No. 3/4, pp. 313-336, 1996.

[12] M. D. Swanson, M. Kobayashi, A. H. Tewfik, "Multimedia Data-embedding and Watermarking Technologies," *Proceedings of the IEEE*, Vol. 86, No. 6, pp. 1064-1087, 1998.

[13] G. Voyatzis and I. Pitas, "The Use of Watermarks in the Protection of Digital Multimedia Products," *Proceedings of the IEEE*,Vol. 87, No. 7, pp. 1197-1207, 1999.

[14] C. I. Podilchuk and W. Zeng, "Image-adaptive Watermarking Using Visual Models," *IEEE Journal on Selected Areas in Communication*, Vol. 16, No. 4, pp. 525-539, 1998.

[15] N. Nikolaidis and I. Pitas, "Robust Image Watermarking in the Spatial Domain," *Signal Processing*, Vol. 66, No. 3, pp. 385-403, 1998.

[16] D. C. Wu and W. H. Tsai, "Data Hiding in Images via Multiple-based Number Conversion and Lossy Compression," *IEEE Transactions on Consumer Electronics*, Vol. 44, No. 4, pp. 1406-1412, 1998.

[17] G. L. Friedman, "The Trustworthy Digital Camera: Restoring Credibility to the Photographic Iimage," *IEEE Transactions on Consumer Electronics*, Vol. 39, pp. 905-910, 1993.

[18] D. Anand and U. C. Niranjan, "Watermarking Medical Images with Patient Information," in *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Vol. 20, No. 2, pp. 703-706, 1998.

[19] J.C. Patra, K.K. Ang, E.L. Ang, "Hierarchical Multiple Image Watermarking for Image Authentication and Ownership Verification," in *Proceedings of International Conference on Image Processing*, Vol. 4, pp. 2661–2664, 2004.

[20] R.G.V Schyndel, A. Z. Tirkel, C. F. Australia, "A Digital Watermark," in *Proceedings of IEEE International Conference on Image Processing*, Vol. II, pp. 86-90, 1994.

[21] S. Walton, "Information Authentication for a Slippery New Age," *Dr. Dobb's Journal*, Vol. 20, No. 4, pp. 18-26 and 82-87, 1995.

[22] M. M. Yeung and F. Mintzer, "An Invisible Watermarking Technique for Image Verification," in *Proceedings of IEEE International Conference on Image Processing*, Vol. II, pp. 680-683, 1997.

[23] P.W. Wong, "A Public for Image Verification and Authentication," in *Proceedings of IEEE International Conference on Image Processing*, Vol. II, pp. 455-459. , 1998

[24] J. Fridrich, "Image Watermarking for Tamper Detection," in *Proceedings of IEEE International Conference on Image Processing*, Vol. II, pp. 404-408, 1998.

[25] C. R. Carlson and R. Cohen, "A Simple Psychophysical Model for Predicting the Visibility of Displayed Information," in *Proceedings of Society for Information Display*, Vol. 21, pp. 229-245, 1980.

[26] C.H. Chou and Y.C. Li, "A Perceptually Tuned Subband Image Coder Based on the Measure of Just-noticeable-distortion Profile," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 5, No. 6, pp. 467-476, 1995.

[27] T. G. Stockham Jr., "Image Processing in the Context of a Visual Model," *Proceedings of the IEEE*, Vol. 60, pp. 828-842, 1972.

[28] B. Zhu and A. H. Tewfik, "Low Bit Rate Near-transparent Image Coding," in *Proceedings of SPIE International Conference on Wavelet Applications for Dual Use*, Vol. 2491, pp. 173-184, 1995.

[29] C.H. Kuo and C.F. Chen, "A Prequantizer with the Human Visual Effect for the DPCM," *Signal Processing: Image Communication*, Vol. 8, pp. 433-442, 1996.

[30] C.H. Kuo and C.F. Chen, "A Vector Quantizer Scheme Using Prequantizers of Human Visual Effects," *Signal Processing: Image Communication*, Vol. 12, pp. 13-21, 1998.

[31] A. N. Netravali and B. G. Haskell, *Digital picture representation and compression*, Plenum Press, New York, pp. 245-299, 1989.