
ELSEVIER

Pattern Recognition Letters 20 (1999) 1511–1517

Pattern Recognition Letters

www.elsevier.nl/locate/patrec

Embedding of any type of data in images based on a human visual model and multiple-based number conversion ☆

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

Department of Computer and Information Science, National Chiao Tung University, 1001 Ta Hsueh Road, Hsinchu 300, Taiwan, ROC

Received 5 February 1999; received in revised form 4 August 1999

Abstract

A novel approach to embedding any type of digital data into a cover image is proposed, which utilizes a human visual model to guarantee that the modification of the cover image is imperceptible. A quantized contrast function based on the model is constructed first. The gray values of each 3×3 sub-image of the cover image are used to compute, using the function, a range of gray levels with the same contrast as that of the central pixel of the sub-image. The embedding process proceeds by replacing the gray value of the central pixel by one of the values in the range. This guarantees that the changes be imperceptible. The data to be embedded is treated as a binary stream and is partitioned into a number of sub-streams. A multiple-base number conversion mechanism is used to convert each sub-stream of data into values which are then embedded in the central pixels of sub-images. The embedded data can be extracted out from the resulting stego-image without referencing the original image. Experimental results show that the proposed method is feasible. © 1999 Published by Elsevier Science B.V. All rights reserved.

Keywords: Data embedding; Cover image; Stego-image; Quantizer; Human visual system; Multiple-based number conversion

1. Introduction

Digitizing information such as text, image, audio and video is popular today. Digital media can be modified and reproduced easily. Several interesting applications about digital media have been explored, such as digital watermarking, tamper-proofing, security message delivering, annotation embedding, etc. These applications can be achieved by embedding information into digital media imperceptibly. The amount of information which need be embedded and the requirement of robustness vary with applications.

Data embedding in images is an application that hides secret messages in digital images. The message may be any type of bit stream data, such as text, image, voice, etc. Some terms about information hiding from Pfitzmann (1996) are followed in this study: an image that holds a secret message is called a cover image, and the output of the hiding process, which includes the secret message, is called a stego-image.

☆ This work is supported partially by National Science Council, R.O.C. under Grant No. NSC88-2213-E009-057.

* Corresponding author. Tel.: +886-3-572-0631; fax: +886-3-572-7382.

E-mail addresses: dcwu@mcu.edu.tw (D.-C. Wu), whtsai@cis.nctu.edu.tw (W.-H. Tsai)

[1] Also at: Department of Information Management, Ming Chuan University, Taipei, Taiwan 111, ROC.

PII: S0167-8655(99)00118-X

Many techniques about data embedding in images have been proposed. Some of the techniques are based on modifying the least significant bits (LSBs) of the pixels of the cover image (Turner, 1989; Walton, 1995). Bender et al. (1997) proposed a patchwork method for embedding information by adding or subtracting a value to or from the gray values of pixels. The texture block coding method (Bender et al., 1996) and the fractal-based steganography method (Davern and Scott, 1996) embed information by modifying small regions of images. Image hiding can also be applied in frequency or other transform domains, such as the method which uses randomly sequenced pulse position modulated codes (Koch and Zhao, 1995), the wavelet-based embedding methods (Ohnishi and Matsui, 1996; Chae and Manjunath, 1998), the secure spread spectrum method (Cox et al., 1997), and the DCT-based embedding methods (Hsu and Wu, 1998; Barni et al., 1998). Recently, a number of methods exploit the human visual system (HVS) to guarantee that the modification made to the cover image is imperceptible. These methods are based on spatial or frequency masking (Girod, 1989; Legge and Foley, 1990).

Three recent HVS-based researches have been conducted by Podilchuk and Zeng (1998), Swanson et al. (1998) and Delaigle et al. (1998). These methods are robust but need more complex computations than those of LSB methods. In our previous work (Wu and Tsai, 1998), we proposed an embedding method, which exploits the gray value differences between a cover image and a lossily compressed version of the cover image. The method needs the information of the original image to extract out the hidden data from the stego-image.

In this paper, we propose a novel, easy and efficient method to embed data of any type in gray-valued images. The method is designed in such a way that it can recover embedded data from the stego-image without referencing the original image. The method is designed to exploit a large amount of space in an image for embedding information. The embedding method can be used to deliver secret messages, embed annotations, or prevent tampering. The method is not robust and is not suitable for the applications that need high robustness, such as watermarking. The method is based on a human visual model which is a modified version of that proposed by Kuo and Chen (1996). The original model was used to remove the non-sensitive information in the process of image compression (Kuo and Chen, 1998). In our method, the modified model is used to construct quantized contrast functions for 3×3 sub-images. The quantized functions provide a set of visual thresholds. We use these visual thresholds to embed secret data in the central pixels of the 3×3 sub-images of the cover image. Since the method can obtain visual thresholds directly from the gray-valued domain, the computation is simpler than those of HVS-based methods which are performed in other transform domains. A pseudo-random number generator may be used to achieve cryptography, and so can prevent tampering access to the embedded data from illicit users. Moreover, a multiple-based number conversion which was proposed by the authors (Wu and Tsai, 1998) is used to convert the secret bit stream into values to be embedded in the chosen pixels of the cover image.

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

2. Human visual model

The human visual model proposed in (Kuo and Chen, 1996) considers Weber's law (Stockham, 1972; Netravali and Haskell, 1998). It was improved in their subsequent research (Kuo and Chen, 1998) by adding some considerations about visual masking effect and Mach band effect (Netravali and Haskell, 1998). These studies model the contrast function by Weber's law in the gray-valued domain to consider the sensitivity of the human eye to a luminance under a background. Let x be a gray value of a stimulus area which ranges from 0 to 255, and μ be the mean of the background gray values under the stimulus. In (Kuo and Chen, 1996), the contrast function $C(x)$ was modeled from the

combination of the conditions of the bright background and the dark background as follows.

For the bright background ($\mu \geqslant 128$),

$$C(x) = \begin{cases} \ln[c_1(c_L - x)/(c_L(127.5 - (x - c_1)))] \\ \quad \text{when } 0 \leqslant x \leqslant 128, \\ \ln[(x - c_1)(x - c_H)/(c_1(255 - c_H))] \\ \quad \text{when } 128 \leqslant x \leqslant 255, \end{cases} \quad (1)$$

where $c_1 = 127.5/2$, $c_H = (128 - 255e^{-k})/(1 - e^{-k})$, $c_L = 128/(1 - e^{-k})$, and k is defined as $k = 2.5/(1 + e^{(255-\mu)/55})$. Similarly, for the dark background ($\mu < 128$),

$$C(x) = \begin{cases} \ln[c_1 c_L/(127.5 - (x - c_1))(c_L - x)] \\ \quad \text{when } 0 \leqslant x \leqslant 128, \\ \ln[(255 - c_H)(x - c_1)/(c_1(x - c_H))] \\ \quad \text{when } 128 \leqslant x \leqslant 255, \end{cases} \quad (2)$$

where $c_1 = 127.5/2$, $c_H = (255 - 128e^{k})/(1 - e^{k})$, $c_L = -128e^{k}/(1 - e^{k})$, and k is defined as $k = 2.5/(1 + e^{\mu/25})$.

In (Kuo and Chen, 1996), a set of visual thresholds was used to quantize the contrast function values into n levels uniformly to form a quantized contrast function. An illustration of a quantized contrast function is shown in Fig. 1. Each quantization level expresses a set of gray values with the same sensitivity to the human eye under a specific background.

Firstly, we build up 256 possible contrast functions for each background luminance. For every possible background gray value b, where $b = 0, 1, \ldots, 255$, a contrast function C_b is obtained by gathering all function values $C_b(x)$ from Eqs. (1) and (2), where x is in the range from 0 to 255, respectively. In our proposed method, we use a visual model which represents the visual sensitivity of a 3×3 image. The model categorizes by experiments 3×3 images into n classes according to the standard deviation of the gray values of all pixels in each 3×3 image except the central pixel. The classes are categorized from smooth to edged. We quantize C_b uniformly to get a quantized contrast function Q_b^c for each class c by using different numbers of quantization levels, say l_c, where $c = 1, 2, \ldots, n$, respectively. For example, in our experiments, the image is categorized into four classes, and 32, 24, 16 and 12 levels are used. 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. The above calculations are off-line, i.e., they are constructed before the embedding steps begin and can be used for embedding any cover image.

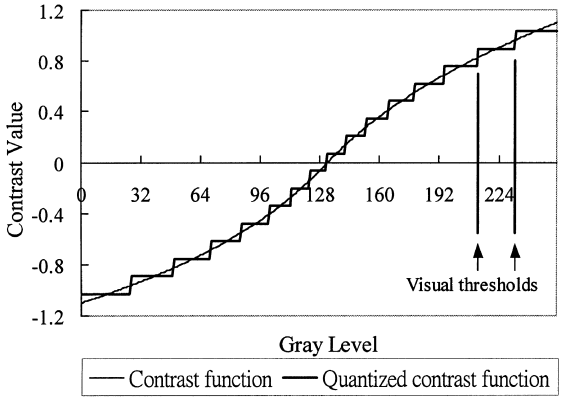

Fig. 1. Contrast function resulting from the condition that the gray value mean of the background is 60 and corresponding quantized contrast function for 16 quantization levels.

3. Proposed data embedding

In our proposed embedding method, the central pixel of every 3×3 sub-image in the cover image is used for embedding information. Fig. 2 shows the way to select the central pixels from a cover image. About one fourth of the pixels in a cover image can be used for embedding information by the selection strategy. The eight pixels which surround a central pixel are considered as the background of the central pixel in our proposed method. Each sub-image is categorized into one of the n classes, say into class c, by calculating the standard deviation of the gray values of the eight surrounding pixels. Then, we use the mean of the gray values of the eight pixels as the background luminance b to find the function value $Q_b^c(g)$ from the quantized contrast function Q_b^c, where g is the gray value of the central pixel. A range $(g^{\min}, g^{\max})$ defined by two extreme visual thresholds $g^{\min}$ and $g^{\max}$ thus can be obtained from Q_b^c where $g^{\min}$ and $g^{\max}$ are the smallest and largest values, respectively. The function value $Q_b^c(g')$ of any gray value g' in the range is the same as that of

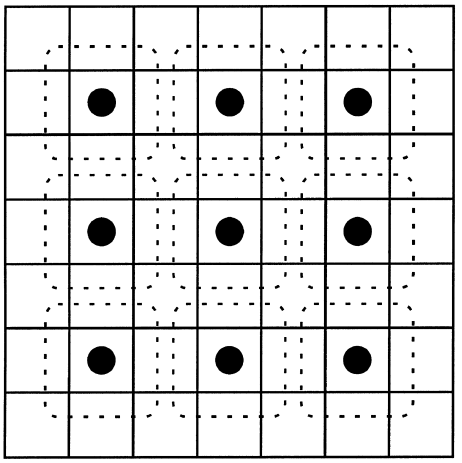

Fig. 2. Illustration of the central pixels in 3×3 sub-images, which are used for embedding information. Dotted boxes indicate the eight surrounding pixels of the central pixels.

$Q_b^c(g)$. That is, any gray value in the range has the same sensitivity as the gray value of the central pixel under the same background. So, if we substitute the gray value of the central pixel with any value in the range defined by the thresholds, the modification will be imperceptible. It means that we can embed a value which ranges from $g^{\min}$ to $g^{\max}$ into the central pixel. Since each central pixel of the 3×3 sub-image in the cover image has two distinct visual thresholds, the magnitude of information which can be hidden in each central pixel varies with the range defined by the two thresholds, i.e., varies with the value of $g^{\max} - g^{\min} + 1$. On the other hand, the secret data can be treated as a long bit stream. We randomly partition it into several variable-sized, say m-bit, sub-streams for easy embedding. Then, each sub-stream is embedded into the cover image sequentially. More specifically, suppose we want to embed an m-bit sub-stream of a secret bit stream. We collect n central pixels in the cover image first. For each central pixel p_i, let its gray value be g_i and visual thresholds be $g_i^{\min}$ and $g_i^{\max}$, where $i = 1, 2, \ldots, n$. The visual thresholds can be utilized to embed totally $\prod_{i=1}^{n}(g_i^{\max} - g_i^{\min} + 1)$ different information into the central pixels. So, we can embed the m-bit sub- stream into these n central pixels if the condition $2^m \leqslant \prod_{i=1}^{n}(g_i^{\max} - g_i^{\min} + 1)$ is met.

We use a multiple-based number conversion method which was proposed by the authors (Wu and Tsai, 1998) to convert the value of the m-bit sub-stream into a value to be embedded in the chosen central pixels. The m-bit bit stream is treated as an unsigned binary number, which is then converted into an n-digit multiple-based number with a form of $d_{n(b_n)}d_{n-1(b_{n-1})}\ldots d_{1(b_1)}$, where base $b_i = g_i^{\max} - g_i^{\min} + 1$, for $i = 1, 2, \ldots, n$. The coefficients $d_1, d_2, \ldots, d_n$ can be computed as the remainders of iterative integer divisions of the value of the m-bit bit stream by $b_1, b_2, \ldots, b_n$ progressively. Then we embed the digit d_i of the multiple-based number into each central pixel p_i by replacing its gray value g_i with the new value of $g_i^{\min} + d_i$ for $i = 1, 2, \ldots, n$. Since the new gray value is in the range defined by the visual thresholds, the modification is imperceptible.

To achieve cryptography, a pseudo-random mechanism is used also in some of the embedding steps, namely, in the selection of the sizes and bits which compose the m-bit sub-stream of the secret bit stream, and in the way of choosing 3×3 sub-images from the cover image where the secret sub-stream is embedded. Moreover, the mechanism is also used for encrypting the secret data. In our experiments, a sequence of random bit values of 0 and 1 is generated for each m-bit bit stream conduct a bit-wise exclusive-or operation to the m-bit bit stream before the multiple-based number conversion. Without using the seeds of the pseudo-random schemes, the embedded data cannot be reconstructed correctly. This can avoid easy illicit extraction of the embedded data from a stego-image.

4. Recovering embedded data from stego-images

The secret data extraction process is accomplished by using the stego-image only. The seeds which were used by the pseudo-random mechanism in the embedding steps must be provided for correct reconstruction of the secret data. We recover each m-bit bit stream of the secret data from the stego-image and combine the bit streams to reconstruct the entire embedded data. For each m-bit bit stream, we first collect n central pixels from the 3×3 sub-images in the stego-image with the same sequences as in the embedding process, and then obtain the visual thresholds of the central pixel of

each sub-image from the specific quantized contrast function. For central pixel p_i with gray value g_i, let its visual threshold be $g_i^{\min}$ and $g_i^{\max}$, where $i = 1, 2, \ldots, n$, and n is the minimum value that satisfies the condition $2^m \leqslant \prod_{i=1}^{n}(g_i^{\max} - g_i^{\min} + 1)$. Since the surrounding pixels remain unchanged during the process of embedding data, the visual thresholds of the central pixel under the background are the same as those obtained in the embedding step. After collecting n central pixels, the embedded value d_i in the central pixel p_i is computed by $d_i = g_i - g_i^{\min}$. We then group the resulting embedded values into an n-digit multiple-based number with a form of $d_{n(b_n)}d_{n-1(b_{n-1})}\ldots d_{1(b_1)}$, where base $b_i = g_i^{\max} - g_i^{\min} + 1$, for $i = 1, 2, \ldots, n$, and convert the number into an m-bit binary bit stream.

5. Experimental results

Four cover images as shown in Fig. 3, each with size 512×512, were used in our experiments, and a file (21 953 bytes) which consists of the text of this paper were used as the secret data. In the experiments, we classified the 3×3 images into four classes which are categorized further from smooth

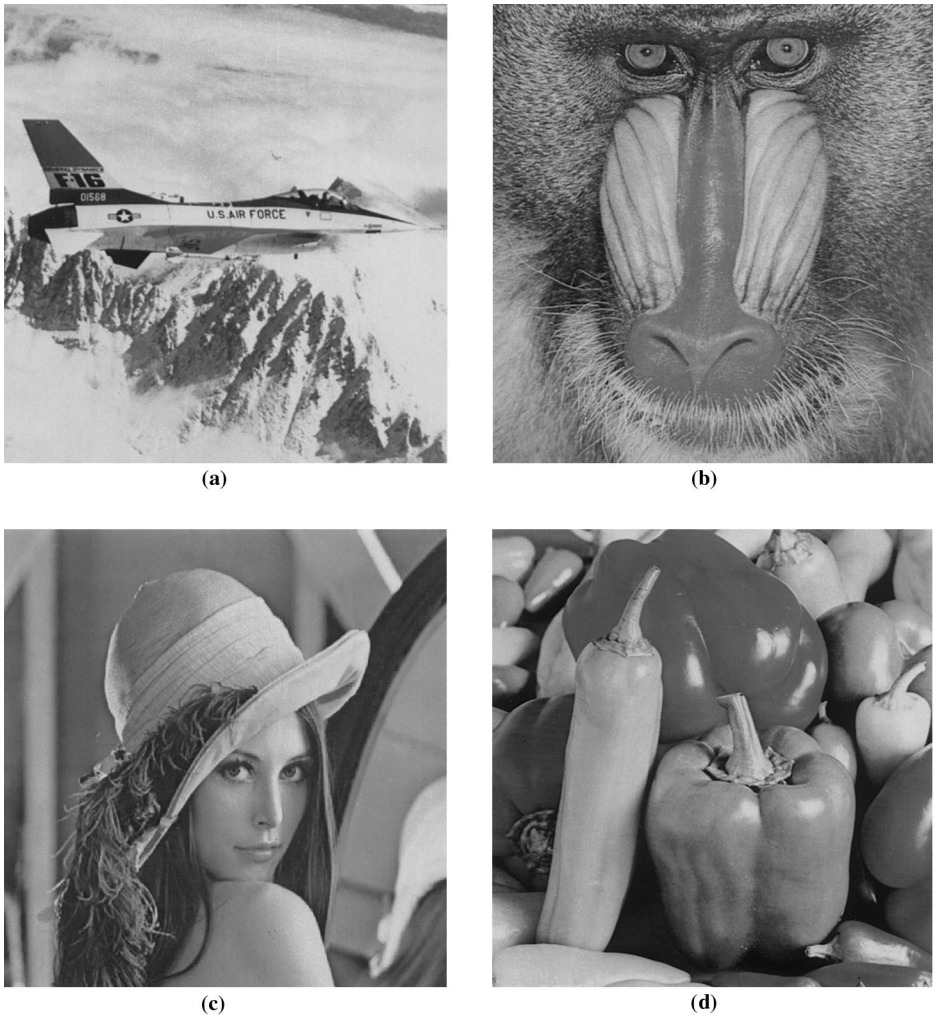

Fig. 3. Cover images used in experiments with size 512×512: (a) Jet, (b) Baboon, (c) Lena, (d) Peppers.

to edged. The criteria which were used to classify the classes and select the number of the quantization levels are shown in the following:

The number of quantized levels

$$= \begin{cases} 32 & \text{when} \quad \sigma \leqslant 2.4, \\ 24 & \text{when} \quad 2.4 < \sigma \leqslant 3.6, \\ 16 & \text{when} \quad 3.6 < \sigma \leqslant 4.8, \\ 12 & \text{when} \quad 4.8 < \sigma, \end{cases} \tag{3}$$

where σ is the standard deviation of the gray value of the surrounding pixels of a 3×3 sub-image. The stego-images resulting from embedding the secret data by randomly walking through the selected central pixels in the cover image are shown in Fig. 4. The results show that the proposed hiding method can embed data without noticeable images.

6. Concluding remarks

A novel, easy and efficient method for embedding any form of digital data into a gray-valued image by using a human visual model in the gray-valued domain has been proposed. Quantized con-

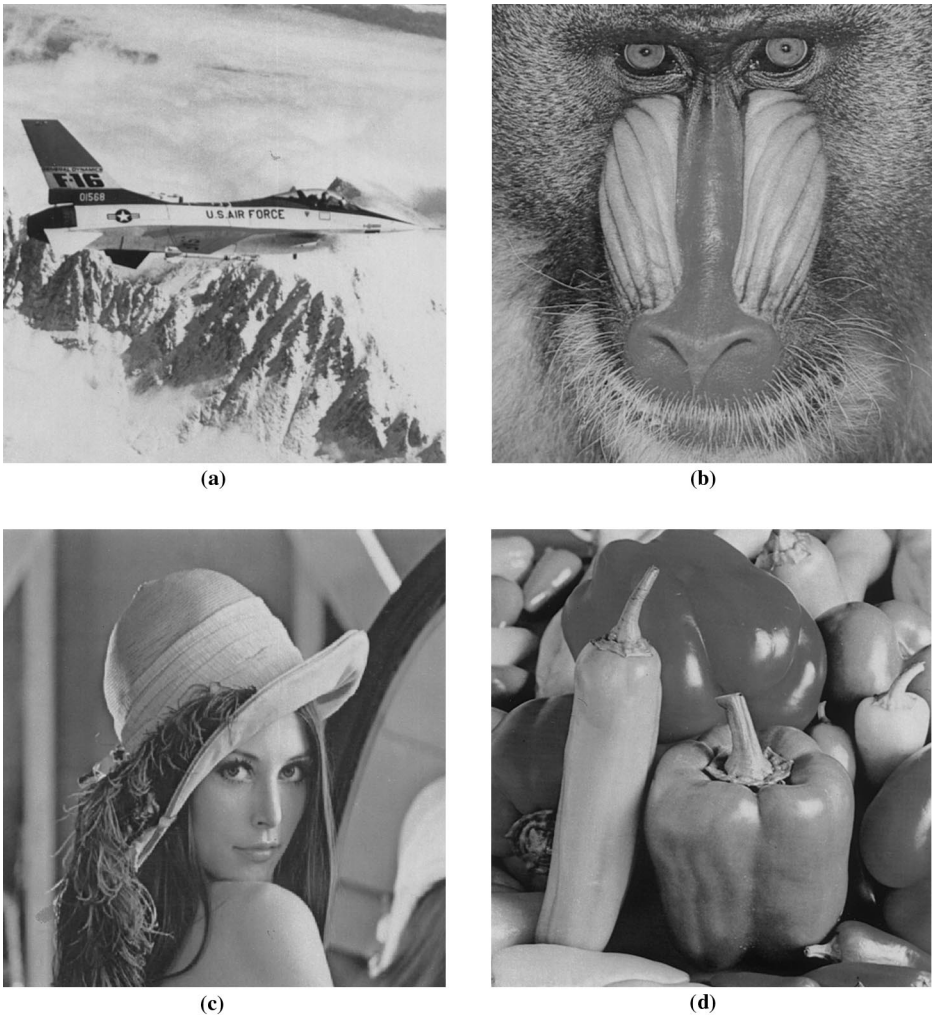

Fig. 4. Resulting stego-images after embedding a file consisting of the text of this paper.

trast functions which provide visual thresholds for the central pixels of 3×3 sub-images were constructed for use in the embedding process. The proposed method is computationally simpler than many HVS-based embedding methods in the steps of embedding and recovering data. A bit stream was embedded into a group of pixels by a multiple-based number conversion method and pseudo random mechanisms are applied in the embedding steps to achieve cryptography. The method provides a way for embedding any form of data in gray-valued images imperceptively. Good experimental results show the feasibility of the proposed method.

References

Barni, M., Bartolini, F., Cappellini, V., Piva, A., 1998. A DCT-domain system for robust image watermarking. Signal Processing 66, 357–372.

Bender, W., Gruhl, D., Morimoto, N., Lu, A., 1996. Techniques for data hiding. IBM System Journal 35 (3/4), 313–336.

Bender, W., Morimoto, N., Gruhl, D., 1997. Method and apparatus for data hiding in images. US Patent No. 5689587.

Cox, I.J., Kilian, J., Leighton, F.T., Shamoon, T., 1997. Secure spread spectrum watermarking for multimedia. IEEE Transactions on Image Processing 6, 1673–1687.

Chae, J.J., Manjunath, B.S., 1998. A robust embedded data from wavelet coefficients. In: Proc. SPIE EI'98, Storage and Retrieval for Image and Video Database VI, Vol. 3312, pp. 308–317.

Davern, P., Scott, M., 1996. Fractal based image steganography. In: First International Workshop Information Hiding. Lecture Notes in Computer Science, Vol. 1174, Springer, Berlin, pp. 279–294.

Delaigle, J.F., Vleeschouwer, C.D., Macq, B., 1998. Watermarking algorithm based on a human visual model. Signal Processing 66, 319–335.

Girod, B., 1989. The information theoretical significance of spatial and temporal masking in video signals. In: Proc. SPIE Human Vision, Visual Processing, and Digital Display, Vol. 1077, pp. 178–187.

Hsu, C.T., Wu, J.L., 1998. DCT-Based watermarking for video. IEEE Transaction on Consumer Electronics 44, 206–216.

Koch, E., Zhao, J., 1995. Towards robust and hidden image copyright labeling. In: Proc. IEEE Nonlinear Signal and Image Processing Workshop, Thessaloniki, Greece, pp. 452–455.

Kuo, C.H., Chen, C.H., 1996. A prequantizer with the human visual effect for the DPCM. Signal Processing: Image Communication 8, 433–442.

Kuo, C.H., Chen, C.F., 1998. A vector quantizer scheme using prequantizers of human visual effects. Signal Processing: Image Communication 12, 13–21.

Legge, G., Foley, J., 1990. Contrast masking in human vision. J. Opt. Soc. Amer. 70, 1458–1471.

Netravali, A.N., Haskell, B.G., 1998. Digital Picture Representation and Compression. Plenum Press, New York, pp. 245–299.

Ohnishi, J., Matsui, K., 1996. Embedding a seal into a picture under orthogonal wavelet transform. In: Proceedings of Multimedia'96, Piscataway, NJ. IEEE Press, pp. 514–521.

Pfitzmann, B., 1996. Information hiding terminology. In: Proceedings of the First International Workshop Information Hiding. Lecture Notes in Computer Science, Vol. 1174, Springer, Berlin, pp. 347–349.

Podilchuk, C.I., Zeng, W., 1998. Image-adaptive watermarking using visual models. IEEE Journal of Selected Areas Communication 16, 525–539.

Stockham Jr. T.G., 1972. Image processing in the context of a visual model. Proceedings of the IEEE 60, 828–842.

Swanson, M.D., Zin, B., Tewfik, A.H., 1998. Multiresolution scene-based video watermarking using perceptual models. IEEE J. Selected Areas Communication 16, 540–550.

Turner, L.F., 1989. Digital data security system. Patent IPN, WO 89/08915.

Walton, S., 1995. Image authentication for a slippery new age. Dr. Dobb's Journal: Software Tools for the Professional Programmer 20, 18–26.

Wu, D.C., Tsai, W.H., 1998. Data hiding in images via multiple-based number conversion and lossy compression. IEEE Transactions on Consumer Electronics 44, 1406–1412.