TRANSPARENT INFORMATION HIDING WITH AUTOMATIC EMBEDDING RANGE SELECTION FOR OWNERSHIP VERIFICATION

Farook Sattar, Dan Yu, Sirajudeen Gulam Razul, and Shankar M. Krishnan
School of Electrical and Electronic Engineering
Nanyang Technological University, Nanyang Avenue, Singapore 639798

ABSTRACT

This paper presents a transparent robust information hiding technique with automatic selection of embedding range. The discrete cosine transform (DCT) coefficients of grayscale watermark adaptively replace the DCT coefficients of original image within the selected embedding range. The embedding range is selected by investigating the perceptual quality of the watermarked image and the robustness of the embedded watermark. The proposed scheme successfully compromises the trade-off between the imperceptibility and robustness. Comparing to Tsai's random replacement embedding scheme, our method can achieve higher robustness against attacks, while retaining better perceptual quality of the marked image. Moreover, the use of an image-dependent dual-key containing embedding locations and scaling factors, can enhance the security level of this scheme in ownership verification.

1. INTRODUCTION

Information hiding [1] provides an efficient means for copyright protection of digital media. The information hiding technique is also known as *digital watermarking* in the context of hiding information for protection of the cover media. To verify the ownership of digital content, the robustness of a watermarking system is required to be high. However, watermarking systems must balance the requirements of three parties: *imperceptibility – robustness – capacity* [1]. *Imperceptibility* requires the marked data and the original data should be perceptually indistinguishable. *Robustness* requires that the embedded information should be reliably detectable or retrievable when the marked data is altered. *Capacity* refers to the amount of the information that is being embedded into the host cover data. If the amount of information to be embedded is decided, there always exists a trade-off between visual quality of the marked data and robustness of the embedded watermark. In general, the higher the embedding strength, the better robustness a watermarking system can achieve, however, at the same time it may result a poorer visual quality of the marked image. Therefore, the evaluation of a watermarking system should include the robustness test as well as the distortion measurement of the watermarked data introduced during embedding, for a fair performance evaluation.

This paper concentrates on a transparent robust information hiding scheme which can comprise the trade-off between the imperceptibility and the robustness of the embedded watermark. In particular, a grayscale watermark is hidden into the cover image in DCT domain transparently for verifying the ownership of the cover media. The automatic selection for an optimal embedding range is proposed.

2. TRANSPARENT ROBUST INFORMATION HIDING SCHEME

2.1. Adaptive Information Hiding Scheme

Suppose $\mathbf{I}$ be the original image of size $N1 \times N2$, and $\mathbf{W}$ be the grayscale watermark of size $M1 \times M2$. As shown in Fig. 1, whole DCT transform is performed for both $\mathbf{I}$ and $\mathbf{W}$, and the resulted DCT coefficients are denoted as $\mathbf{C_I}$ and $\mathbf{C_W}$, respectively. The 1-D sequence, $S_{\mathbf{I}}$, of $\mathbf{C_I}$ is generated using zig-zag scanning, whereas the 1-D sequence of $\mathbf{C_W}$, denoted as $S_{\mathbf{W}}$, is obtained by simple column-wise scanning.

The watermark embedding is performed through *adaptive replacement* of the selected DCT coefficients of original image ($S_{\mathbf{I}}$) by the DCT coefficients of watermark ($S_{\mathbf{W}}$), as shown in Fig. 2. Let the selected embedding range in $S_{\mathbf{I}}$ be [$Initial$,End], where $Initial$ denotes the starting index and End denotes the ending index. The embedding algorithm searches the nearest match of each watermark coefficient $S_{\mathbf{W}}(i)$ in the selected embedding range, in terms of the difference, d, between their absolute amplitudes,

$$d(i) = \left| |S_{\mathbf{I}}(location(i))| - |S_{\mathbf{W}}(i)| \right|, i = 1, 2, ..., M1 \times M2, \quad (1)$$

where $location(i)$ stores the location of the nearest match, and $S_{\mathbf{I}}(location(i))$ is the original image's DCT coefficient which is the nearest match for the i-th watermark coefficient. The nearest match $S_{\mathbf{I}}(location(i))$ for i-th watermark

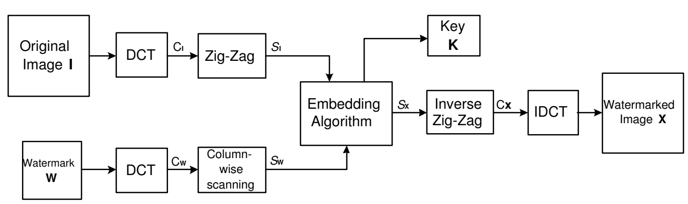


Fig. 1. Watermark embedding procedure.

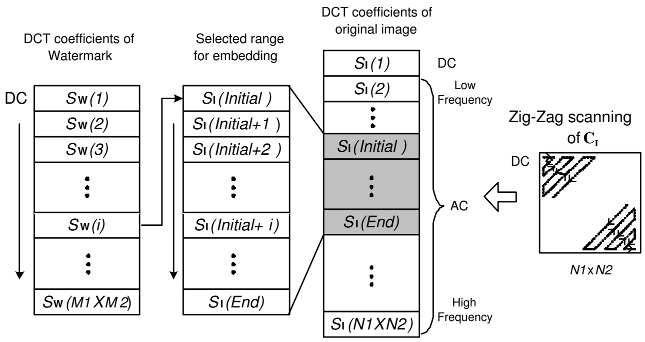


Fig. 2. The search engine for watermark embedding.

coefficient is not included again in the matching process for the following watermark coefficients.

A threshold T_0 is set to control the quality of the marked image. If the difference d is smaller than T_0, $S_{\mathbf{I}}(location(i))$ is replaced by the watermark coefficient $S_{\mathbf{W}}(i)$ directly. However, if the difference d is not less than T_0, the coefficient $S_{\mathbf{I}}(i)$ is kept unchanged, instead a scaling factor, $scale(i)$, between the $S_{\mathbf{I}}(location(i))$ and $S_{\mathbf{W}}(i)$ is generated for a perfect watermark recovery in the watermark extraction. To represent mathematically, the watermarked image's DCT coefficients' sequence, $S_{\mathbf{X}}(j)$ ($\mathbf{X}$ denotes the finial watermarked image, $j = 1, 2, ..., N1 \times N2$), is

$$S_{\mathbf{X}}(j) = \begin{cases} \text{sign}\{S_{\mathbf{I}}(j)\} \cdot |S_{\mathbf{W}}(j)| & \text{if } j \in location(i), \\ & i = 1, 2, ..., M1 \times M2 \\ & \text{and } d(j) < T_0 \\ S_{\mathbf{I}}(j) & \text{else} \end{cases}; \quad (2)$$

and the scaling factors – *scale* – are defined as follows:

$$scale(i) = \begin{cases} 1 & \text{if } d(i) < T_0 \text{ and} \\ & S_{\mathbf{I}}(location(i)) S_{\mathbf{W}}(i) \geq 0 \\ -1 & \text{if } d(i) < T_0 \text{ and} \\ & S_{\mathbf{I}}(location(i)) S_{\mathbf{W}}(i) < 0 \\ \frac{S_{\mathbf{I}}(location(i))}{S_{\mathbf{W}}(i)} & \text{if } d(i) \geq T_0 \end{cases}. \quad (3)$$

The scaling factors are either +1 or -1 if difference d between the two nearest matched candidates is smaller than the threshold T_0, depending on whether they are of the same sign. Otherwise, the scaling factor equals the ratio of these two candidates for the nearest match. Thus, T_0 is employed as a control parameter for watermarked image's quality, and *scale* is an important refinement parameter for good watermark retrieval.

The *Key* in the proposed watermarking system includes both the watermark embedding location – *location* and the scaling factors – *scale*. Note that the generated dual key of the algorithm is image-dependent, which implicates only the person, who knows the key, is able to do the ownership verification for the particular image content.

The final watermarked image $\mathbf{X}$ is obtained by an inverse zig-zag scanning of the embedded sequence $S_{\mathbf{X}}$ followed by a 2-D inverse DCT transform.

2.2. Ownership Verification Scheme

In an ownership verification system [2], the legal authority stores the watermark as well as one of the two keys. Only when the owner presents the other key, the final watermark detection can be done by the legal authority and the ownership is then verified. A dual-key watermarking system could be an advantage for ownership verification, especially when two owners sharing the ownership of the Work jointly. The key can be easily separated in two parts – one part contains embedding location and the other part is scaling information. Without requiring the original data, the rightful ownership can be verified when both owners present their keys.

Fig. 3 shows the blind watermark extraction procedure. The DCT coefficients, $C_{\mathbf{X}}{}'$, of the received image $\mathbf{X}'$ is obtained. The 1-D sequence, $S'_{\mathbf{X}}$, of $C'_{\mathbf{X}}$ is then generated by zig-zag scanning. Based on the key information of the watermark embedding locations (in vector *location*) and the scaling factors (in vector *scale*), the vector form for the DCT coefficients of the watermark, $S_W{}'$, can be retrieved by

$$S_W{}'(i) = \frac{S_X{}'(location(i))}{scale(i)}, i = 1, 2, ..., M1 \times M2. \quad (4)$$

The watermark's DCT coefficients, $C_W{}'$, in 2-D form, is obtained by rearrangement of $S_W{}'$ in column-wise sequencing. The final extracted watermark, $\mathbf{W}'$, is obtained by a 2-D inverse DCT transform of $S_W{}'$.

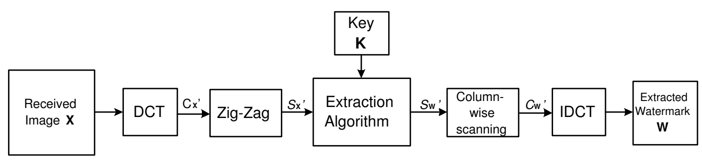


Fig. 3. Watermark extraction process.

3. PROPOSED AUTOMATIC SELECTION OF THE EMBEDDING RANGE

An automatic selection for the optimal embedding range, [$Initial$,End], is proposed by determining the End index followed by the estimation of the $Initial$ index. The process is demonstrated by using a Lena image of size 256×256 pixels (in Fig. 4(a)) as the original image and a grayscale 'NTU' image of size 38×111 pixels (in Fig. 4(b)) as watermark.

3.1. Selection of End Index

The End index of embedding range is determined by considering the robustness requirement of a watermark. One of the most important attacks, the effects of JPEG compression to the original image's DCT coefficients, is investigated. One would like to hide the watermark into the coefficients that are less sensitive to JPEG compressions. In particular, JPEG compression with quality factor 60% is used as a reference. The sensitivity of DCT coefficients

Fig. 4. (a) The original Lena image, (b) the grayscale watermark ('NTU' is the abbreviation of 'Nanyang Technological University'), (c) the watermarked Lena image and (d) the extracted watermark.

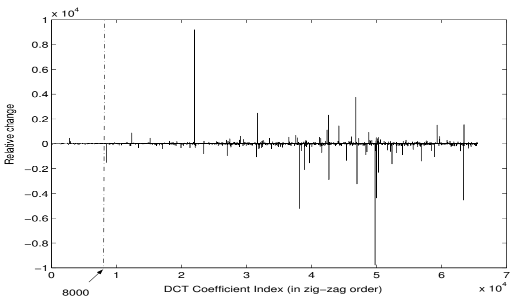


Fig. 5. The End index selection for embedding range.

against compression is measured by the relative change between the compressed coefficient $C_{compressed}$ and its original $C_{original}$, RC, defined by

$$\text{RC} = \frac{C_{compressed} - C_{original}}{C_{original}}. \tag{5}$$

As shown in Fig. 5, roughly the first 8,000 most significant coefficients are more resistive against compression because of the very small variance of RC. Hence, the End index is set as 8,000 from the robustness issue of the embedded watermark.

3.2. Selection of $Initial$ Index

The $Initial$ index of the embedding range is estimated from the perceptual quality of the watermarked data. One may not want to touch those most significant coefficients of the original image to avoid the easily visualized changes caused by varying those coefficients. A HVS-based quality measurement - weighted PSNR (wPSNR) [3] - has been defined by Voloshynovskiy *et al.* as an extension of the traditional quality metric – PSNR. Particularly, a perceptual model of human vision - Noise Visibility Function (NVF) [3] - is used to weight each term of the PSNR:

$$\text{wPSNR} \;=\; 10\log_{10}\frac{||255||^2}{||(\mathbf{I}-\mathbf{X})\,.\,\text{NVF}||^2}, \tag{6}$$

where $\mathbf{I}$ is the original image and $\mathbf{X}$ is the marked image. The level of NVF varies from 0 to 1. The magnitude of NVF approaches 1 in very flat region implying that the changes may be easily perceived for embedding in such regions.

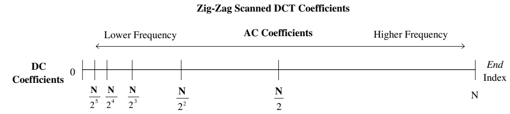


Fig. 6. $Initial$ index selection by the technique of octave-band division.

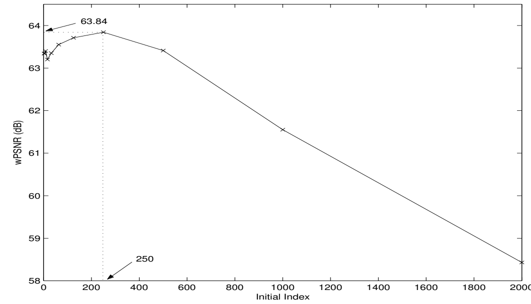


Fig. 7. The watermarked image quality (in wPSNR) with respect to the selection of the $Initial$ index.

The quality (in wPSNR) of the watermarked Lena image by selecting different $Initial$ index is examined for estimation of an optimal setting. Once the End index is fixed, the DCT coefficients within the range of [1, End] are divided into sub-bands with the octave-based division technique (see Fig. 6), which resembles the human visual system [4]. The coefficients of index N$\cdot 2^{-(s-1)}$ (N=End Index, $s \in \{1, 2, 3, ..., \lfloor \log_2 \text{N} \rfloor\}$), are selected as $Initial$ Index to test their corresponding qualities of the watermarked images. Fig. 7 shows the visual quality for the watermarked image measured in wPSNR, with respect to the selection of the $Initial$ index in the manner of the octave-band division. The optimal visual quality of the watermarked image is obtained when the $Initial$ index is 250. The embedding into the low frequency components below index of 250 results in degradations in the visual quality. Therefore, for this watermarking example, the $Initial$ index is set as 250 from the view point of the visual quality of the watermarked image.

For selection, the total number of DCT coefficients selected for the embedding range should be higher than the total number of DCT coefficients of the watermark image. For better nearest match, the embedding range needs to be chosen between 1.5 times to 3 times the length of the DCT sequence of the watermark. Because, if the embedding range is chosen too small as compared to the length of the watermark's DCT sequence, the number of selected coefficients would not be sufficient for good matching (see the lower wPSNR measures when $Initial$ $>$250). On the other hand, if the embedding range is too large, the robustness of the embedded watermark could be degraded due to the embedding in the less significant coefficients. In addition, the perceptual quality of the marked content could be degraded as well due to the embedding in those most significant coefficients (see the lower wPSNR measure when $Initial$ $<$250). Moreover, it is rather a time-consuming process if the searching range is very large.

4. PERFORMANCE EVALUATION

4.1. Embedding and Extraction Results

The threshold T_0 is set by considering the total distortions allowed in a watermarked image. For instance, the PSNR of the watermarked image $\mathbf{X}$ needs to be maintained above 40dB. The following relationships are obtained for watermarking the Lena image in Fig. 4(a):

$$\begin{aligned} \text{PSNR} &= 10\log\Big(\frac{255^2 \times Total\ image\ pixel\ no.}{\Delta E}\Big) \geq 40\text{dB}; \\ T_0 &\leq \sqrt{\frac{\Delta E}{Total\ no.\ of\ watermark\ samples}} \\ &= \sqrt{\frac{255^2 \times 256^2}{10^4 \times (38 \times 111)}} \doteq 10.0514, \end{aligned} \quad (7)$$

where ΔE is the total energy difference allowed between the watermarked image and the original image.

Fig. 4(c) shows a watermarked Lena image (wPSNR= 63.84dB and PSNR=61.20 dB) by setting T_0=10.0514, and embedding range of [250,8000]. The watermark extraction performance is evaluated by *normalized correlation coefficient*, *r*, of the extracted watermark $\mathbf{W}'$ and the original watermark $\mathbf{W}$ (of $M1 \times M2$ pixels) [1]

$$r = \frac{\sum_{i=1}^{M1}\sum_{j=1}^{M2}\widetilde{\mathbf{W}}(i,j) \times \widetilde{\mathbf{W}}'(i,j)}{\sqrt{\sum_{i=1}^{M1}\sum_{j=1}^{M2}\widetilde{\mathbf{W}}^2(i,j) \times \widetilde{\mathbf{W}}'^2(i,j)}}, \quad (8)$$

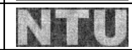
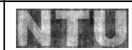
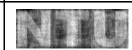
where $\widetilde{\mathbf{W}}$ and $\widetilde{\mathbf{W}}'$ are, respectively, the normalized original and extracted watermark. The magnitude range of r is [0, 1], and the unity holds if the extracted image perfectly matches the original one. Fig. 4(d) shows the extracted watermark ($r \doteq$1.0000) from the watermarked Lena image in Fig. 4(c) when there is no attack. The results for the other test images are not included here due to space limitation.

4.2. Performance Comparison

Our proposed scheme is compared with Tsai's method [5]. Tsai *et al.* [5] proposed a robust 8×8 block DCT based watermarking scheme, also by adopting gray-level watermarks. An adjusting quantization table is required for embedding and extraction. The middle frequency coefficients of the original image are *randomly* selected to be replaced by the quantized DCT coefficients of the watermark, therefore, the watermarked image could not achieve very high quality measure. The Tsai's random replacement embedding results in a watermarked image of PSNR 39.65dB [5], whereas our proposed adaptive replacement embedding achieves much better imperceptibility of the hidden watermark – a PSNR measure of 61.20dB for the marked image.

The watermark extraction results against various attacks such as JPEG compression, quantization, additive noise, geometric distortions and filtering, are shown in Table 1. Experimental results have shown its excellent resistance against a wide range of attacks. It has been demonstrated that the proposed scheme has better robustness performance than Tsai's schemes as well as other DCT-based watermarking schemes.

Table 1. Watermark extraction results against various attacks.

Attacks	PSNR (dB)	Extracted watermark	r
JPEG compression 50%	33.88		0.9996
JPEG compression 1%	22.79		0.9482
Color quantization to a binary image	5.77		0.8889
Additive Gaussian noise	13.61		0.9123
Rotation 45°, cropping and scale down by 40%	–		0.8299
Blurring	25.66		0.9858
Sharpening	9.59		0.9339

5. CONCLUSION

A transparent information hiding technique has been presented in this paper. Meaningful gray-level image is hidden transparently using adaptive replacement embedding method in DCT domain. The paper emphasizes a feasible solution for automatic embedding range selection from the view points of the robustness and the imperceptibility of the hidden watermark. It is also found that the replacement technique is more robust against conventional additive embedding due to the higher information hiding capacity (as in Cox's method [6]). Our proposed scheme successfully compromises the contradictory trade-off between the imperceptibility and the robustness of the embedded watermark. The generation of image-dependent dual-key provides a higher level of security of proposed scheme in the practical applications for ownership verification. It would be an efficient verification tool especially when the quality of the watermarked data is required to be high.

6. REFERENCES

[1] S. Katzenbeisser and F. A. P. Petitcolas, *Information Hiding Techniques for Steganography and Digital Watermarking*, Boston: Artech House, 2000.

[2] Y. Wang, J. F. Doherty, and R. E. V. Dyck, "A wavelet-based watermarking algorithm for ownership verification of digital images," *IEEE Trans. on Image Processing*, vol. 11, no. 2, pp. 77-88, Feb. 2002.

[3] S. Voloshynovskiy, A. Herrigel, N. Baumgaertner and T. Pun, "A stochastic approach to content adaptive digital image watermarking," *Proc. of International Workshop on Information hiding*, Dresden, Germany, Sep. 1999.

[4] Y. J. Jung, M. Hahn and Y. M. Ro, "Spatial frequency band division in human visual system based-watermarking," *Springer-Verlag Lecture Notes in Computer Science*, vol. 2613, pp. 224-234, May 2003.

[5] C.-S. Tsai and C.-C. Chang, "Embedding robust graylevel watermark in an image using discrete cosine transformation," *Distributed Multimedia Databases: Techniques and Applications*, Timothy Shih (Ed.), Idea Group Publishing, Jan. 2002.

[6] I. J. Cox, J. Kilian, T. Leighton and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Trans. Image Processing*, vol. 6, no. 12, pp. 1073-1087, 1997.