IEICE TRANS. FUNDAMENTALS, VOL.E87–A, NO.6 JUNE 2004

PAPER

A Combined Approach to Integrity Protection and Verification of Palette Images Using Fragile Watermarks and Digital Signatures*

Chih-Hsuan TZENG[†a)], ***Member and*** **Wen-Hsiang TSAI**[†b)], ***Nonmember***

SUMMARY Conventional authentication methods, proposed mainly for gray-scale and color images, are not appropriate for palette images, which usually contain simple contents with a limited number of colors. In this paper, a new approach is proposed to verify the integrity of palette images and to locate tampered regions without re-quantization and re-indexing processes. The proposed approach is based on a combined use of both the fragile watermarking and the digital signature approaches, taking the advantages of both approaches and avoiding their drawbacks. To protect a block of an image, authentication signals are first generated according to a secret key. Based on an embeddability property defined in the study, the pixels of each block are classified as embeddable or non-embeddable. Only the former ones are used to embed the authentication signals. A corresponding digital signature is generated as well to compensate the possibly limited embedding capacity of the embeddable pixels that are insufficient in number. To authenticate a block, the recovered authentication signals, yielded from the extracted watermark and the received digital signature, are compared with the one generated according to the correct secret key, to prove the block's legitimacy. The effectiveness and the security of the proposed method are analyzed and tested with a variety of palette images. The results indicate that the proposed method can offer high authentication accuracy as well as maintain a good tradeoff between the authentication signal portability and the resulting image quality.

key words: *palette image, image authentication, fragile watermarking, digital signature, random palette mapping, embeddability*

1. Introduction

Palette images are ubiquitous in modern computer systems. Almost every web page contains palette images to depict icons, logos, maps, etc. Because palette images are digital in nature, they can be seamlessly manipulated by image processing tools. However, in medical, atmospheric, geographic, or military applications, it is required to keep the contents of palette images intact. Ease of seamless modifications of palette images becomes a security problem. The objective of this research is to devise a novel method to protect and verify the integrity of palette images.

The studies of image authentication can be categorized into two approaches: *use of digital signatures* and *fragile watermarking*. Digital signatures of image data are analogous to those of general data studied in traditional cryptography [1]. To protect an image, features are extracted from the image and encoded to form compact signatures. Changes in the original image will result in certain corresponding changes in the extracted features. Authentication of an image is accomplished by comparing two feature sets, one including those extracted from the image and the other from the corresponding digital signature. Certain features, such as color [4], color histogram [5], edge [6], [7], and moment invariant [7], have been proposed to generate digital signatures. Also, some features in the discrete cosine transform [8] and the discrete wavelet transform (DWT) domains [9] have been proposed. These features are based on the results of comparing randomly selected coefficient pairs.

On the other hand, the idea of the fragile watermarking approach is to imperceptibly embed certain coded information into an input image, called a *cover image*. The coded information hidden behind the cover image is called a *watermark*. Fragile watermarks are named to differentiate from robust ones for their low robustness against image manipulations. By examining the existence of the designated fragile watermark in a given image, the authenticity of the image can be proved. Many fragile watermarking methods have also been proposed [10]–[12]. An extensive survey of fragile watermarking techniques can be found in [2], [3].

Comparing the two authentication approaches, we see that fragile watermarking has better portability, because no extra data are generated. However, as the embedding capacity of an image is usually limited, it might be difficult to embed a watermark without introducing noticeable distortion. This is especially true for images with limited numbers of colors and smooth regions, like palette images. On the contrary, the digital signature approach leaves an image intact, so the image quality is preserved, although extra storage space is required to keep the signature. A combined use of both of the two approaches might result in a good solution for authenticating palette images, as is studied in this work.

More specifically, the *embeddability* of a pixel is first defined to classify pixels into *embeddable* and *non-embeddable* ones based on the characteristics of the human visual system. Since the values of embeddable pixels can be revised without causing noticeable distortion, they are used to embed and extract watermarks. Also, the use of a new type of digital signature is proposed to compensate the weakness of the limited embedding capacity of palette images. By gathering sufficient authentication information from both watermarks and digital signatures, it is proved

Manuscript received November 7, 2002.
Manuscript revised July 7, 2003.
Final manuscript received February 26, 2004.

[†]The authors are with the Department of Computer & Information Science at National Chiao Tung University, Taiwan 300, R.O.C.
a) E-mail: chtzeng@cis.nctu.edu.tw
b) E-mail: whtsai@cis.nctu.edu.tw

*This work was supported partially by the MOE Program for Promoting Academic Excellency of Universities under the grant number 89-1-FA04-1-4.

that the proposed method has good ability to locate tampered regions.

The remainder of this paper is organized as follows. In Sect. 2, a brief introduction to palette images is first given, followed by a description of the definition of pixel embeddability proposed in this study. The proposed scheme is presented in Sect. 3. The performance of the proposed method is discussed in Sect. 4. Section 5 shows some experimental results. At last, Sect. 6 includes the conclusion of this study.

2. Pixel Embeddability of Palette Images

2.1 Principle of Proposed Approach

A palette image, denoted as I, is composed of a color palette P_I and a set D_I of image data. Generated with a color reduction method, the palette P_I is a list of no more than 256 RGB colors, representative of those in D_I. A pixel X with color C is assigned an index c to the palette entry $P_I(c)$ whose RGB color value is closest to C. As the number of colors in P_I is limited and the colors are visually quite dissimilar, it is generally difficult to alter pixel colors in a given palette image without affecting the resulting image quality. This makes the design of watermark embedding methods for palette images a challenge.

We define a property of pixels in palette images, called, *embeddability*, based on certain human visual characteristics. The definition of this property will be given later in this paper. Accordingly, pixels are classified as *embeddable* or *non-embeddable*: embeddable pixels can be modified moderately to embed watermark data, but non-embeddable pixels must be left intact to preserve image quality. Then, given a pixel X with an index c of an image I, a binary watermark bit b, and a binary mapping function M of palette index values, the proposed watermark embedding process can be described briefly by the following steps:

1. Check the embeddability of X.
2. If X is non-embeddable, leave X intact; otherwise, perform the next step to embed b into X.
3. Check if $M(c) = b$. If so, regard b to be already existing at X; otherwise, replace c with an index c_k in P_I where $M(c_k) = b$.

Correspondingly, the proposed watermark extraction process can be described sketchily by the following steps:

1. Check the embeddability of X.
2. If X is non-embeddable, regard no watermark being embedded in X; otherwise, regard the bit value embedded in X as $M(c)$.

The concept of the above sketchy watermark embedding and extraction processes is to embed b into an embeddable pixel X with an index c by setting $M(c) = b$. As a result, the embedded watermark can be extracted easily by investigating the value $M(c)$ of an embeddable pixel X. Specifically, the value of an embeddable pixel X is changed only when $M(c) \neq b$. In this situation, a new palette index, whose mapping function value is equal to b, is selected to substitute the original one. In the following, we define some terms in advance. Based on these terms, the proposed embeddability of a pixel and the palette mapping function will be defined formally.

2.2 Color Distance and Precedent Neighbors of a Pixel

Let c be the index of a pixel p and let the corresponding triplet of the RGB color $P_I(c)$ be (r, g, b). Because the numerical distance in the L*a*b* color space is intuitively proportional to the perceived color difference, we define the color distance between two pixels in this study in terms of the L*a*b* color values. The transformation from an RGB color (r, g, b) to an L*a*b* one (l^*, a^*, b^*) can be found in [13].

Definition 1 (*Color distance between two pixel indices.*) Let X_1 and X_2 be two image pixels with palette indices c_1 and c_2, and L*a*b* colors (l_1^*, a_1^*, b_1^*) and (l_2^*, a_2^*, b_2^*), respectively. The color distance between X_1 and X_2 is defined as:

$$d(X_1, X_2) = d(c_1, c_2) = \sqrt{(l_1^* - l_2^*)^2 + (a_1^* - a_2^*)^2 + (b_1^* - b_2^*)^2}.$$

Definition 2 (*Precedent neighbors of a pixel.*) Given a pixel X in a 3×3 image block B, the *precedent neighbors* of X are those *four* pixels of the eight neighbors of X in B, which are visited first in a line-by-line raster scanning sequence. That is, if X is located at coordinates (h, k) in I, then its precedent neighbors are located at $(h-1, k-1)$, $(h-1, k)$, $(h-1, k+1)$, and $(h, k-1)$. We denote the set of the precedent neighbors of X by $Prec(X)$.

Definition 3 (*Maximum color difference between a pixel and its precedent neighbors, and that between a color and these neighbors.*) The maximum color difference between a pixel X with color index c and its precedent neighbors in $Prec(X)$ is defined as:

$$d_{max}(X, Prec(X)) = \max_{X_i \in Prec(X)} (d(X, X_i)),$$

and that between a given color c_k and $Prec(X)$ is defined as:

$$d_{max}(c_k, Prec(X)) = \max_{X_i \in Prec(X)} (d(c_k, X_i)),$$

where c_i is the color index of a pixel X_i in $Prec(X)$.

2.3 Embeddability of Pixels

Definition 4 (*Palette mapping function.*) A palette mapping function M takes a color palette P_I as its domain, and the output $M(c)$ of the function for each input palette index c in P_I is a binary value, 0 or 1.

Base on the characteristics of the human visual system, a visually seamless modification of a pixel X is possible when X is located in a color-abundant region where no sharp line or edge exists. Two constraints on X are devised in this study to determine whether X satisfies the criterion for seamless modification stated above: (1) The number α of distinct colors of pixels in $Prec(X)$ is larger than a pre-defined threshold T_c. This states the restriction that the change of a pixel's color is allowed only in a color-abundant region. (2) The maximum color distance $d_{max}(X, Prec(X))$ between X and its precedent neighbors is smaller than with a pre-defined threshold T_d. This states the restriction that modifications of colors are allowed only in regions with no sharp line or edge.

Despite of the seamless modification constraints, some other aspects need be taken into account before giving the definition of the embeddability of a pixel. From the above sketchy watermark embedding and extraction processes, it can be noted that only embeddable pixels are used to embed and extract watermark data. So, it is important to ensure that the embeddability of a pixel is *consistent* during both the embedding and the extraction processes. Next, the embeddability of a pixel will be changed only when a replacement of the original pixel index is conducted to embed a watermark bit during the embedding process. Therefore, if the pixel is assigned a new index value which fails to satisfy the embeddability constraints, the pixel will undesirably be regarded as non-embeddable.

Definition 5 (*Set of feasible replacement indices of a pixel.*) Given a pixel X with color index c, and a palette mapping function M, the set C_X of feasible replacement indices of X includes all of those palette indices satisfying the following properties: (1) $M(c) \neq M(c_k)$; and (2) $d_{max}(c_k, Prec(X)) < T_d$, where T_d is a pre-defined threshold.

Note that, when a replacement is needed, only indices with their palette mapping function values opposite to that of X's original index c can be considered as candidates to substitute c. This depicts the first property in the above definition. The second property guarantees that X, with the new index c_k, will not result in the creation of a sharp line or edge. That is, the second property satisfies the second constraint for seamless modification of pixel colors mentioned previously.

Definition 6 (*Embeddability of a pixel.*) Given a pixel X with color index c, let α be the number of distinct colors of the pixels in $Prec(X)$, and M be a palette mapping function. Then, X is said to be *embeddable* if the following three constraints are satisfied: (1) $\alpha > T_c$; (2) $d_{max}(X, Prec(X)) < T_d$; and (3) C_X is a non-empty set, where T_c and T_d are two pre-defined thresholds.

With the above definition, the embeddability of a pixel can be proved to be consistent during the embedding and the extraction processes, as done in the following.

Property 1 (*Preserving of a pixel's embeddability.*) Let X be a pixel with palette index c, and $c_k \in C_X$. If X is embeddable, then it is still embeddable after replacing c with c_k.

Proof. To prove that X is still embeddable, we prove that it satisfies the three constraints mentioned in Definition 6.

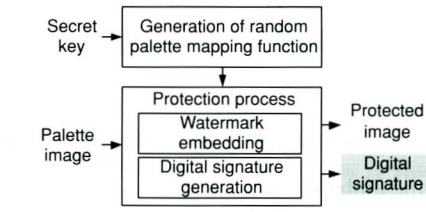
1. (Constraint 1) Because $Prec(X)$ is not changed after c is replaced with c_k, it is trivial to see that the constraint $\alpha > T_c$ is maintained.
2. (Constraint 2) Based on Definition 5, we have $d_{max}(c_k, Prec(X) < T_d$. This means that the second constraint is satisfied.
3. (Constraint 3) Because X is embeddable, we have $M(c) \neq M(c_k)$ and $d_{max}(c, Prec(X)) < T_d$. Hence, C_X is a non-empty set. So the third constraint is satisfied. This completes the proof.

□

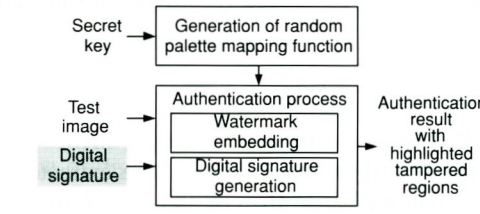
3. Proposed Image Protection and Authentication Process

An overview of the proposed method is illustrated in Fig. 1. As shown in Fig. 1(a), watermark embedding and digital signature generation are integrated into a single process, called a *protection process*. The inputs include a palette image and a generic secret key. The outputs include a protected image whose integrity can be verified, and a digital signature generated from the input palette image. Note that, in Fig. 1(a), the background of the digital signature output is darkened to emphasize that possibly no digital signature will be generated for a certain image by the protection process. This case arises when the embedding capacity of an image is large enough to convey sufficient authentication signals as a watermark.

Correspondingly, as illustrated in Fig. 1(b), watermark extraction and digital signature verification works are integrated into a single process, called an *authentication pro-*



(a) Protection process.



(b) Authentication process.

Fig. 1 Overview of the proposed method.

cess. The inputs include a palette image in question, a secret key, and the corresponding digital signature of the palette image. The output is the verification result of the image. If the image is verified as tampered, tampered regions will be located.

More specifically, before the protection or the authentication process starts, an input image I is divided into a set of non-overlapping blocks B_i's of the size $m \times n$. Each block B_i is processed individually subsequently. For simplicity, assume that the size of I is a multiple of $m \times n$. Furthermore, a palette mapping function M, as defined in Definition 4, is generated according to K for use in processing each B_i. The generation of M is described in the following.

3.1 Generation of Random Palette Mapping Function M

M can be taken as a vector of the size equal to that of the palette of the input image, and each entry of M contains a binary mapping value. Because palette colors can be permuted in different ways without affecting the image content, a color may be referenced by different indices under different permutations. For example, let the palette mapping function value of a color j in block B_i be $M(1)$. It will become $M(2)$ if a permutation switches the order of the first two colors in the palette. This will result in the generation of different M values from an identical palette. To ensure the uniqueness of M it is necessary to maintain a unique set of palette color indices. For this, we define a static color order for a palette in this study as follows.

Definition 7 (*Static color order of a palette index.*) Let $f(u) = 2^{16} \times r + 2^8 \times g + b$ be a function with RGB colors $u = (r, g, b)$ as inputs where $0 \leq r, g, b \leq 255$ and let c be an index of an image I with color palette P_I. If $f(P_I(c))$ is the t-th largest value of all 2^{24} possible outputs of f, then the static color order, denoted as $H(c)$, of c is defined to be t, i.e., $H(c) = t$.

Note that, no matter what the index value of a specific color in a palette is, the static color order of a palette index is unique according to Definition 7. Now, M can be generated as follows, assuming that P_I is composed of ℓ distinct colors so that M is also a vector of size ℓ:

1. generate a random bit stream $z_1 z_2 \ldots z_n$ with $n = 2^{24} = 16777216$ by a random number generator with K as the seed; and
2. for each palette index c_j where $j = 1$ through ℓ, set $M(c_j) = z_{H(c_j)}$.

3.2 The Detailed Protection Process

As illustrated in Fig. 2, let the number of pixels in a block B_i of an input image I be $m \times n$. The first step of the process is to generate a secret bit stream V of the length $m \times n$ as the authentication signals for B_i. To maintain a good trade-off between image quality and embedding capacity without affecting the authentication performance, those bit values of V, which cannot be embedded, are manipulated with certain pixel values of B_i to generate certain features which are then taken to form a digital signature. Consequently, a watermark w_i is hidden in the embeddable pixels of B_i, and the digital signature is proposed in this study to be stored in the header of the input palette image I for future references.

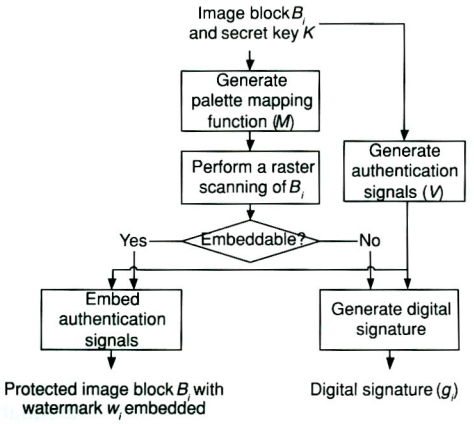


Fig. 2 The detail of the proposed protection process.

The idea to embed a bit into an embeddable pixel X in B_i is to enforce the value $M_i(X)$ of X equal to the bit to be embedded. If a modification of the color of X is needed, its index is replaced with an index c_{opt}, called the *optimal replacement index* for X, selected from C_X.

Algorithm 1. *The embedding process.*

***Input*:** An embeddable pixel X of B_i with color index c in an image I, a secret bit $V(j)$ to be embedded, and a random palette mapping function M.

***Output*:** A pixel X' with a new index c' implying the secret bit $V(j)$.

Steps:

1. Check if $M(X)$ is equal to $V(j)$. If so, regard $V(j)$ to be already existing at X, set c' as c and stop; otherwise, perform the next step.
2. Find the optimal replacement index c_{opt} in C_X.
3. Set c' as c_{opt} and stop.

The first step of the above algorithm checks whether X has "contained" the value $V(j)$ to be embedded. And the second step aims to select c_{opt}, which specifies the "most similar" color to that of X. Hence, the least distortion coming from pixel color changes is ensured. It is also ensured that c_{opt} always exists, and that the replacement will not affect the embeddability of X. Based on the above algorithm, we are now ready to describe the details of the proposed protection process in the following.

Algorithm 2. *The detailed protection process.*

***Input*:** A block B_i of a palette image I, and a secret key K.

***Output*:** A block B_i' with authentication signals embedded as a watermark w_i and/or stored as a digital signature g_i.

***Steps*:**

1. Let $g_i = \emptyset$ (the empty set).
2. Use K to generate a palette mapping function M.
3. Generate a secret bit stream $V = V(1)V(2)V(3)\ldots V(mn)$ as the authentication signals using a random number generator with K as the seed.
4. Set $j = 0$.
5. Perform a raster scanning of B_i, take a non-visited pixel X of B_i, and execute the following steps, either to embed a bit of V as a watermark signal of w_i or to generate a bit of the digital signature g_i, until all the pixels of B_i are visited.
6. Check the embeddability of X. If it is embeddable, then embed $V(j)$ into X by Algorithm 1, and go to Step 5; otherwise, compute a bit $e = V(j) \oplus M(X)$ where $\oplus$ is the exclusive-OR operation of two bits; and append e to the end of g_i.
7. Set $j = j + 1$, and go to Step 5 to continue the process to visit the next pixel in B_i, until V is processed to its end.

Note that V is first generated in the above algorithm based on K. Then, for each visited pixel in B_i, a bit of V is embedded if the pixel is *embeddable*. If not, a binary value is yielded as a new signature bit by "exclusive-ORing" the current processed bit of V and the palette mapping function value of the pixel. Also note that, to produce the entire digital signature of the input image I, the digital signatures of all image blocks need be concatenated. Because the length of a digital signature of a block is variable, we place $\lceil \log_2 mn \rceil$ bits before each g_i to indicate its length. Hence, the corresponding digital signature of a block can be correctly recovered from the digital signature of I in the authentication process described next.

3.3 The Detailed Authentication Process

The proposed authentication process, illustrated in Fig. 3, is based on sequential processing of blocks. A block B'_i of a test image I', the corresponding digital signature g_i of B'_i, and the secret key K are inputs to the process. Only persons, who have the secret key K identical to the one used in the protection process, can correctly verify the integrity of I'. The idea of the process is to recover the corresponding authentication signals V' from B'_i and compare it with the original authentication signals V. If no image pixel is tampered, the recovered V' will be identical to V, and this proves the authenticity of B'_i.

More specifically, to recover V' a raster scanning of B'_i is performed. For each pixel, its embeddability is evaluated. If it is embeddable, a bit of the embedded watermark is extracted; otherwise, a bit of the V' is generated from the digital signature. After all image pixels are visited, the original authentication signals in V are generated based on K. At last, V' and V are compared to verify the authenticity of B'_i. In the following, we describe the details of the proposed authentication process as an algorithm.

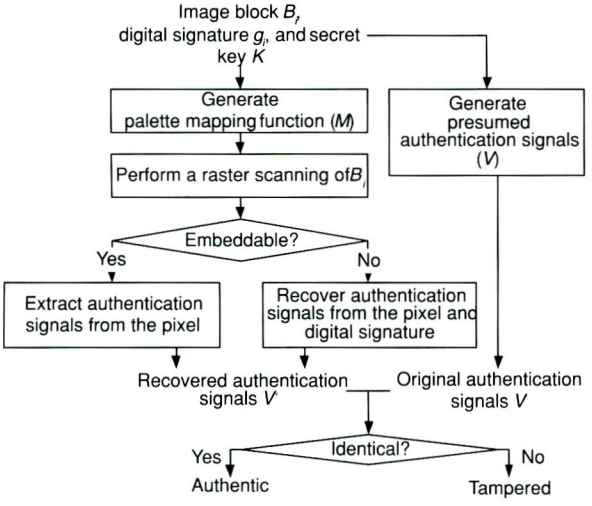


Fig. 3 The detail of the proposed authentication process.

Algorithm 3. *The detailed authentication process.*

***Input*:** A block B'_i of a test image I', the corresponding digital signature g_i of B'_i, and the secret key K.
***Output*:** The authentication result of B'_i.
***Steps*:**

1. Set $v = 0$.
2. Use K to generate a random palette mapping function M.
3. Perform a raster scanning of B'_i and process each pixel X in B'_i until all pixels of B'_i are visited.
4. Check the embeddability of X and perform the following steps, assuming that X is the j-th pixel visited during the raster scanning process:
   - 4.1 If X is embeddable, set $V'(j) = M(X)$.
   - 4.2 If X is non-embeddable, then set $V'(j) = g_i(v) \bigoplus M(X)$ and $v = v + 1$.
   - 4.3 Go to Step 3 to process the next pixel in B'_i.
5. Use K to generate a secret bit stream $V = V(1)V(2)V(3)\ldots V(mn)$.
6. If $V' = V$, then regard B'_i as authentic; otherwise, tampered.

Note that, in Step 4, the embeddability of X is first checked. Then, corresponding actions are taken to recover a bit $V'(j)$ of V'. In Step 4.1, if X is embeddable, then a bit of the original authentication signals, denoted as $V(j)$, is supposed to be embedded in it during the protection process. Hence, to extract $V'(j)$ from X, $M(X)$ is evaluated and assigned to $V'(j)$. Similarly, in Step 4.2, recall that in the protection process a bit e of g_i was generated by computing $e = V(j) \bigoplus M(X)$. Consequently, during the authentication process, the original $V(j)$, denoted as $V'(j)$, is obtained by evaluating $g_i(v) \bigoplus M(X)$.

4. Performance and Security Analysis

4.1 Performance Analysis

Two types of errors might occur in an authentication system: *miss* and *false alarm*. In this study, a miss refers to the situation that a block is tampered but is identified as authentic, while a false alarm refers to the situation that a block is authentic but is identified as tampered. We will analyze the probabilities of the occurrences of a miss and a false alarm in the following.

Because the proposed method ensures that the embeddability of each pixel of a block is preserved after applying the protection process as long as the block is not tampered, the hidden watermark and the digital signature of the block can be correctly extracted and verified by the authentication process. Hence, false alarm errors will *not* occur in the proposed system.

On the other hand, a miss error will occur when a received image is tampered but V and V' are checked to be identical. Therefore, the probability of the occurrence of a miss error equals to the probability of the generation of two identical random bit streams. Because the lengths of V and V' are both mn, the probability can be figured to be $2^{-m\times n}$. For example, when B_i is of the size 8×8, which requires the generation of a bit stream of 64 bits, the probability of the occurrence of a miss will be 5.4×10^{-20} which is small enough for security protection in practical applications.

4.2 Security Analysis

An effective attack is to tamper a protected image without triggering authentic alarms. If certain information must be presented during the authentication process to prove the authenticity of an image, they must also be required in conducting a successful attack. Four types of information are required in the authentication process: (1) the palette mapping function M; (2) the authentication signals V; (3) the secret key K which determines M and V; and (4) the embeddability of the pixels determined by the two threshold values T_c and T_d involved in Definition 6.

To find out the valid knowledge of the first two types of information mentioned above without the secret key K, the only way is to enumerate all possible combinations of V and M. Assume that the size of an image block is $m \times n$. The number of all possible V's is 2^{mn}. In addition, let r be the number of distinct colors in a palette. Then, the number of all possible M's is 2^r. Hence, the number of all possible combinations of V and M is $2^{nm}2^r$. This value can be increased to make it even difficult to discover V and M. For instance, M can be randomly generated to be distinct for each pixel of a block. Then, the number becomes $2^{nm}2^{rnm}$. If $n = m = 8$ and $r = 16$, then the number will become 21088 which is large enough to prevent a brute-force attack.

In addition, because the embeddability of a pixel cannot be known in advance without T_c and T_d, whether a pixel is used to embed a bit of a watermark or to generate a bit of a digital signature is not clear, leading to a certain degree of difficulty in guessing the contents of the watermark w_i and the digital signature g_i for each block. By randomizing T_c and T_d using K and a random number generator, it is harder to guess the possible values of T_c and T_d. For example, for each pixel, a random combination of (T_c, T_d) is generated by two random number generating functions f_{rg1} and f_{rg2} with the initial seeds of both functions set to K. In this study, we set $T_c = f_{rg1}(K)$ and $T_d = f_{rg2}(K)$, where $1 \le f_{rg1}(K) \le 3$ and $5 \le f_{rg2}(K) \le 50$. In this way, it will become harder to predict the values of T_c and T_d because these values are not globally set. In addition, a random number generator will always yield the same number sequence, as long as the input seed key is identical. Thus, the thresholds T_c and T_d will be identical in the protection and authentication processes.

In summary, the above discussions show that the proposed method has high immunity to security attacks.

5. Experimental Results

We have evaluated the performance of the proposed method using a large collection of palette images. These images were selected to simulate the use of palette images in real world applications. Some results are reported in this section.

In our experiments, a 8×8 image block is moderate for detection of localized manipulations. In addition, T_c and T_d are set as 2 and 15, respectively. Also, we compute the degrees of average distortion in the embeddable pixels of a block (ADEPB) based on Definition 1 using the following equation:

$$ADEPB = 1/N \sum_{i=1}^{N} d(c_i, c_i'),$$

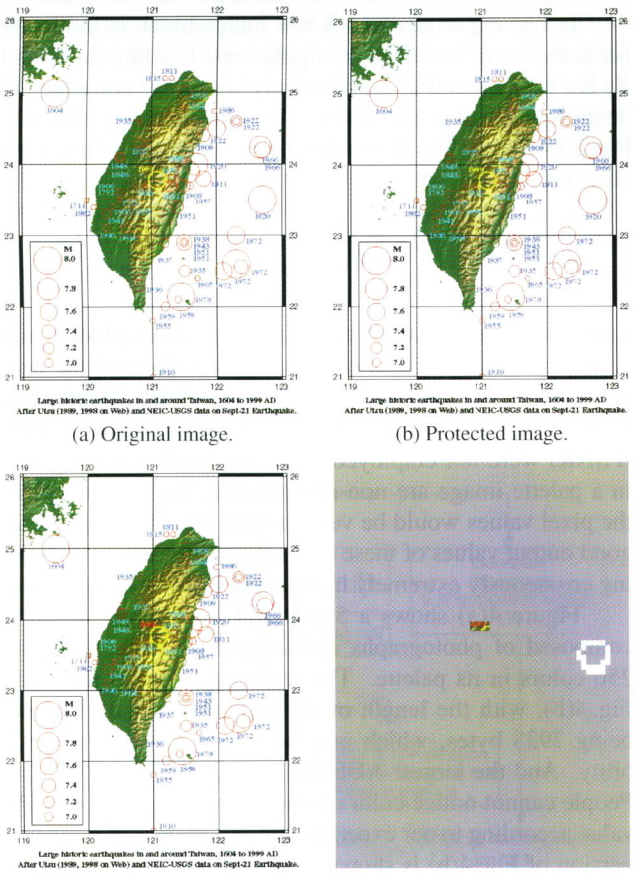
where N is the number of embeddable pixels in a block, and c_i and c_i' are the indices of the original color and the new one of the i-th embeddable pixel, respectively. Note that, conventional measures like the peak signal to noise ratio (PSNR) were not employed. A reason is that many pixels in a palette image are non-embeddable, and so changes in the pixel values would be very limited. This will yield very good output values of these conventional measures, indicating erroneously extremely high output image quality.

Figure 4(a) shows a 560×504 test image, which is composed of photographs and descriptive texts, and with 256 colors in its palette. The resulting image is shown in Fig. 4(b), with the length of the resulting digital signature being 7938 bytes, which may be said to have high portability. And the largest ADEPB value of the image is 8.25. People cannot notice color difference with such an ADEPB value according to our experimental experience. A tampered version of Fig. 4(b) is shown in Fig. 4(c) where the content was altered by switching the two scenic photographs in the image. The authentication output is given in Fig. 4(d), in which, as expected, only the regions of the switched photographs are considered as tampered. Another example is

(a) Original image. (b) Protected image.

(c) Tampered image. (d) Authentication result.

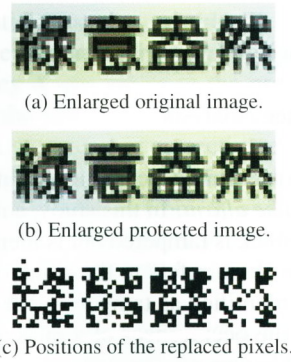
Fig. 4 The authentication result of a test image composed of text regions and photographs of the size 560 × 504 with 256 palette colors using the proposed method.



(a) Original image. (b) Protected image.

(c) Tampered image. (d) Authentication result.

Fig. 5 The authentication result of a test image of the size 488×672 with only 64 palette colors using the proposed method.



(a) Enlarged original image.

(b) Enlarged protected image.

(c) Positions of the replaced pixels.

Fig. 6 Enlarged sub-images of the original image and the protected images.

Table 1 Comparison of the proposed method with conventional fragile watermarking and digital signature approaches.

	Proposed method	Fragile watermarking	Digital signature
Operation in palette index domain	Yes	No	No
Need for re-quantization and re-indexing	No	Yes	No
Modification of image content	Depends	Yes	No
Need for extra storage space	Depends	No	Yes
Deal with smooth regions and limited colors	Yes	Unknown	Yes

shown in Fig. 5 in which Fig. 5(a) is a 488 × 672 web image with 64 colors used to show geological data. Figure 5(b) shows the image protected by the proposed method. The largest ADEPB is 9.76, and the length of the digital signature is 11907 bytes. The results again indicate a good trade-off between embedding capacity and image quality. Then, the image was tampered by cropping the circle and its caption "1920" at the right part of the image and replacing the color of the center caption "1999" with red as shown in Fig. 5(c). At last, the image is authenticated by the proposed method and the tampered regions are marked. As expected, only the tampered regions are pointed out by the proposed method. This shows the effectiveness of the proposed method.

Furthermore, Fig. 6(a) shows an enlarged sub-image of Fig. 4(a) which contains the text regions. Figure 6(b) is the corresponding enlarged sub-image of Fig. 4(b) where 280 pixels are replaced and their positions are shown in Fig. 6(c). It can be seen that the details of the characters in the original image are not affected. In this case, most authentication information is conveyed in the digital signature and the integrity of the regions still can be verified. At last, we compare the differences of three characteristics of the proposed method with those of conventional fragile watermarking and digital signature approaches in Table 1.

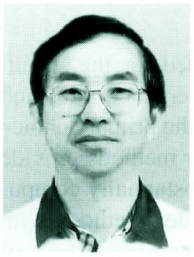
6. Conclusions

A novel method for integrity protection and verification of palette images has been proposed in this study. The proposed method is based on combining both the fragile watermarking and digital signature approaches to maintain good balance between output image quality and authentication data portability. We first defined the embeddability of a pixel based on certain human visual characteristics to classify pixels into embeddable and non-embeddable ones. Then watermark signals are embedded into embeddable pixels and a digital signature is generated from manipulating non-embeddable ones, so that sufficient authentication information can be acquired to achieve high accuracy of authentication results. All the processes are conducted in the palette index domain, and so no re-quantization and re-indexing is needed. Furthermore, we have shown that the proposed method is with low probabilities of yielding authentication errors, and with high security against intentional attacks. Good experimental results have been produced from extensive tests of the proposed method using a large collection of images. Both of the facts have proved the feasibility of the proposed method to real world application environments like WWW and digital libraries. Finally, applying the proposed method to pictures of graphic types like animation images, which are also palette-based, can be taken as a topic for future studies.

References

[1] B. Schneier, Applied Cryptography: Protocols, Algorithms, and Source Code in C, 2nd ed., John Wiley & Sons, New York, 1995.
[2] E.T. Lin and E.J. Delp, "A review of fragile image watermarks," Multimedia and Security Workshop in ACM Multimedia'99, pp.25–29, Orlando, FL, Oct. 1999.
[3] I. Cox, M. Miller, and J. Bloom, Digital watermarking, Morgan Kaufmann, San Francisco, CA, 2001.
[4] D.C. Lou and J.L. Liu, "Fault resilient and compression tolerant digital signature for image authentication," IEEE Trans. Consum. Electron., vol.46, no.1, pp.31–39, Feb. 2000.
[5] M. Schneider and S.F. Chang, "A robust content based digital signature for image authentication," 1996 International Conf. on Image Processing, pp.227–230, Lausanne, Switzerland, Sept. 1996.
[6] J. Dittmann, A. Steinmetz, and R. Steinmetz, "Content-based digital signature for motion pictures authentication and content-fragile watermarking," IEEE International Conf. on Multimedia Computing and Systems, pp.209–213, Florence, Italy 1999.
[7] M.P. Queluz, "Authentication of digital images and video: Generic models and a new contribution," Signal Process. Image Commun., vol.16, pp.461–475, 2001.
[8] C.Y. Lin and S.F. Chang, "A robust image authentication method distinguishing JPEG compression from malicious manipulation," IEEE Trans. Circuits Syst. Video Technol., vol.11, pp.153–168, Feb. 2001.
[9] C.S. Lu and H.Y.M. Liao, "Structural digital signature for image authentication," ACM International Conf. on Multimedia, pp.115–118, Los Angeles, CA, Nov. 2000.
[10] D. Kundur and D. Hatzinakos, "Digital watermarking for telltale tamper proofing and authentication," Proc. IEEE, vol.87, pp.1167–1180, July 1999.
[11] M. Wu and B. Liu, "Watermarking for image authentication," 1998 International Conf. on Image Processing, pp.437–441, Chicago, IL, 1998.
[12] M.M. Yeung and F. Mintzer, "An invisible watermarking technique for image verification," IEEE International Conf. on Image Processing, pp.680–683, Washington, D.C., 1997.
[13] R.S. Berns, Principles of Color Technology, 3rd ed., Wiley, New York, 2000.

Chih-Hsuan Tzeng was born in Hualien, Taiwan, R.O.C. in 1973. He received the B.S. and Ph.D. degrees both from the Department of Computer and Information Science of National Chiao Tung University, Taiwan in 1995 and 2003, respectively. He was an assistant research fellow in the MOE program for promoting academic excellency of universities. He was the recipient of the 2003 best thesis award of the Institute of Information and Computer Machinery (IICM). He is currently a programming officer in the navy. His research interest include information hiding, image compression, pattern recognition.



Wen-Hsiang Tsai was born in Tainan, Taiwan, R.O.C. in 1951. He received the B.S. degree in electrical engineering from National Taiwan University in 1973, the M.S. degree in electrical engineering from Brown University in 1977, and the Ph.D. degree in electrical engineering from Purdue University in 1979. He joined the faculty of National Chiao Tung University, Taiwan in 1979, is currently a Professor in the Department of Computer and Information Science and the Vice President of the University. He has served as the Head of the Department, the Dean of Academic Affairs of the University, the Chairman of the Chinese Image Processing and Pattern Recognition Society at Taiwan, the Editor of several international journals, and the Editor-in-Chief of *Journal of Information Science and Engineering*. He has published 293 academic papers, including 120 journal papers and 173 conference papers. Professor Tsai has received many awards, including four Outstanding Research Awards, two Special Researcher Awards, and one Distinguished Researcher Award, all of the National Science Council, R.O.C. He was also the recipient of the Academic Award of the Ministry of Education, the 13th Annual Best Paper Award of the Pattern Recognition Society of the U.S.A., and many academic paper awards made by several academic societies. Professor Tsai's major research interests include image processing, pattern recognition, computer vision, virtual reality, and information hiding. Dr. Tsai has supervised the thesis studies of 29 Ph.D. students and 113 master students. He is a senior *IEEE* member, a member of the Chinese Image Processing and Pattern Recognition Society, and the Chairman of the Computer Society of *IEEE* Taipei Section.