A STEGANOGRAPHIC FRAMEWORK FOR DUAL AUTHENTICATION AND COMPRESSION OF HIGH RESOLUTION IMAGERY

Deepa Kundur

Texas A&M University
EE Department
College Station, TX
USA 77843-3128

Yang Zhao

University of Toronto
ECE Department
Toronto, Ontario
Canada M5S 3G4

Patrizio Campisi

Universitá degli Studi
di Roma "Roma Tre'
Via della Vasca Navale
84, 00146 Roma, Italy

ABSTRACT

This paper proposes an approach for the combined image authentication and compression of color images by making use of a digital watermarking and data hiding framework. The digital watermark is comprised of two components: a soft-authenticator watermark for authentication and tamper assessment of the given image, and a chrominance watermark employed to improve the efficiency of compression. The multipurpose watermark is designed by exploiting the orthogonality of various domains used for authentication, color decomposition and watermark insertion. The approach is implemented as a DCT-DWT dual domain algorithm. Simulations and comparisons of the proposed approach with state-of-the-art existing work demonstrate the potential of the overall scheme.

1. INTRODUCTION

It has been recently shown that steganography (related to hiding the "existence" of messages) and digital watermarking can be used for a diverse set of applications such as media authentication and compression. For authentication, the approach has the potential to provide the necessary "soft" integrity verification capabilities that traditional digital signatures cannot [1, 2] Furthermore, data hiding can be used to exploit the inefficiencies of certain lossy compression algorithms to provide even more compression reduction in size for color images [3]. In this work, we focus on providing both authentication capabilities and compression using semi-fragile digital watermarking and compressive data hiding. To the best of the authors' knowledge our approach is the first method that combines both processes using data hiding.

A review of existing digital watermarking approaches for image authentication shows that previous techniques can be classified as employing a *single domain host dependent watermark* [4, 5, 6] in which the image-dependent watermark generation and embedding are "mixed" in the same domain, or as involving a *host independent watermark* [1, 7], in which the watermark is a random sequence or logo independent of the image and embedded in a given domain. The former class of techniques suffers from high sensitivity or the inability to appropriately localize the degradations on the signal. The latter category requires the transmission the watermark W itself or an equivalent signal which makes the approach susceptible to eavesdropping and sophisticated attempts of fraud. In this work, we assert that the use of orthogonal subspace dual domains can keep the watermark embedding, which occurs in one image subspace, from interfering with watermark generation, which is applied to another orthogonal subspace, for more controlled soft authentication while overcoming the limitations of previous work.

The next section presents our framework. Section 3 describes our proposed dual domain authentication and compression scheme followed by simulations, and final remarks in Sections 4 and 5.

2. FRAMEWORK

Our framework is comprised of the following components:

1. The **generating function**, f_g, which produces the watermark signal W to embed as follows:
$$W = f_g(\iota, \kappa, Y) \tag{1}$$
where κ is the secret *generation key* known only to the sender and receiver, Y is the luminance of the host[1] image X, and ι is called the watermark "payload" which is comprised of a bit sequence independent of κ and Y. In our application, W has two parts: an *authenticator watermark* component W_a employed for security and a *chrominance watermark* component W_c to help with compression; we represent this relationship as a concatenation: $W = [W_a \| W_c]$ where $\|$ is the concatenation operator.
2. The **embedding function**, f_m, which inserts W into the luminance host data Y with the help of a secret *embedding key* K known only to the sender and receiver, yielding the watermarked data:
$$Y_w = f_m(Y, W, K) \tag{2}$$
such that Y_w is perceptually identical to Y.
3. The **lossy compression function**, f_l, which reduces the practical storage requirements of Y_w to form the compressed signal $\tilde{Y}_w$ as follows:
$$\tilde{Y}_w = f_l(Y_w). \tag{3}$$

[1]The *host* image by definition is the signal in which the watermark is embedded.

where $\tilde{Y}_w$ is the compressed secured version of Y.

4. The **extracting function**, f_x, which recovers the watermark information, $\hat{W}$, from the received watermarked data, $\hat{Y_w}$ (which may differ from $\tilde{Y}_w$ because of distortions in the image distribution chain), using the secret key K :

$$\hat{W} = f_x(\hat{Y_w}, K). \qquad (4)$$

5. The **recovery function**, f_r, which employs $\hat{W}$ for authentication and color recovery of the image:

$$[R_a, \hat{X}_w] = f_r(\hat{Y}_w, \hat{W}, \kappa') \qquad (5)$$

where κ' is a key available to the receiver that is different than (or the same as) κ if asymmetric (or symmetric) encryption is employed for authentication, R_a is a statistic that allows the authentication and tamper assessment of $\hat{Y}_w$, and $\hat{X}_w$ is the overall color-recovered version of $\hat{Y}_w$.

The authenticator watermark W_a should represent a secure content-based adaptive authenticator such that it is a function of image features invariant to predefined content-preserving image processing operations denoted Ω_R while fragile to specified content modification attacks denoted Ω_F. In addition, the component of the payload ι corresponding to the chrominance watermark W_c should be a compressed version of the color information to be later combined with the watermarked luminance image for color recovery.

For design simplicity, the host image should be partitioned into two distinct components – one in which to embed W_a and another for W_c – and employ different embedding approaches for each. This facilitates more straightforward control over achieving both tasks of authentication and compression. Furthermore, embedding should not affect authenticator watermark generation. Another consideration is that to achieve overall compression gains, chrominance embedding and lossy compression must work together. Specifically, given a coder structure f_l, the inefficiencies of compression should be exploited as unused bandwidth available for W_c embedding.

3. ALGORITHM

3.1. Orthogonality and Dual Domains

Our philosophy is to break an image into the following subspaces: V_c containing the chrominance information of the image to produce W_c, and V_l containing the luminance component. Furthermore, V_l is partitioned into subspaces V_{gen} for W_a generation, $V_{emb,a}$ for W_a embedding, and $V_{emb,c}$ for W_c embedding. Ideally, all subspaces should be orthogonal, so that any processing involved in these domains do not interfere with one another. Moreover, V_{gen} should allow access to "salient" image features that can be exploited by f_g to relate to the integrity of the image. Similarly, $V_{emb,a}$ should also contain features that are related to image credibility, but that can be used to characterize tampering, and $V_{emb,c}$ should be reasonably invariant to f_l so that the chrominance information can be robustly embedding. This can be achieved with the use of dual domains to produce V_{gen}, $V_{emg,a}$ and $V_{emb,c}$.

Given these basic principles, we next present an algorithm for joint authentication and compression of imagery.

3.2. Algorithmic Specifics

For our simulations, we make use of cultural heritage (CH) imagery and therefore reason that soft authentication must be forgiving of mild compression, low energy additive noise, and linear filtering to collectively form Ω_R as discussed in Section 4. In contrast, we would like the scheme to recognize forgery of the entire image, addition, removal or extreme changes in spatially localized visual features; these attacks collectively form Ω_F.

Our proposed method is summarized in Figure 1. The color image is first decomposed into the YIQ color space. The luminance component is passed through a soft authenticator generation algorithm to produce W_a. The chrominance components I and Q are subsampled using a 2-D discrete wavelet transform (DWT) to form W_c. Then W_a and W_c are embedded in turn to produce the watermarked image which is then compressed using an adaptive wavelet-based compression algorithm discussed in [3]. We next describe the soft authenticator generation, embedding and extraction. Details of the color information embedding and adaptive compression are described in [3]. The transforms used are all separable for computational simplicity.

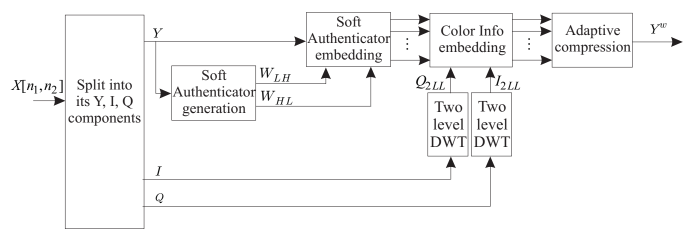

Fig. 1. *Dual domain compression and authentication watermark generation.*

The W_a generation, detailed below, aims to take the essential image features invariant to Ω_R and fragile to Ω_F and secure it cryptographically.

1. *DCT:* Take the 8×8 block DCT of the $M_x \times M_y$ luminance component Y to produce the coefficients $f_{D_{i,j}}(u,v)$ where (i,j) for $1 \leq i \leq \lceil \frac{M_x}{8} \rceil$, $1 \leq j \leq \lceil \frac{M_y}{8} \rceil$ denotes the particular block and (u,v) for $1 \leq u,v \leq 8$ is the frequency index where $(1,1,)$ represents the dc coefficient.

2. *Feature Extraction:* Compose a matrix of dc block coefficients D as follows: $D(i,j) = f_{D_{i,j}}(1,1)$ for $1 \leq i \leq \lceil \frac{M_x}{8} \rceil$, $1 \leq j \leq \lceil \frac{M_y}{8} \rceil$. Note: it is possible to determine $D(i,j)$ by taking the direct average of each 8×8 block. We consider both these stages for generality. Earlier implementations made use of non-dc DCT components.

3. *Binary Transform:* Initialize B as a matrix of zeros. Use the session key K_S to pair up every element $D(i,j)$ with another element $D(i',j')$ such that (i',j') is in a 3×3 neighborhood around (i,j). If $|D(i,j) - D(i',j')| < 16$ then find another pair member as follows. Consider a line connecting (i,j) to (i',j'). Rotate this line clockwise by

$\frac{\pi}{4}$ radians to find another possibility for $D(i', j')$, and repeat until a proper $D(i', j')$ can be found. Assuming a $D(i', j')$ is found such that $|D(i, j) - D(i', j')| \geq 16$, the following relations will be preserved under JPEG compression of 70% and moderate SPIHT compression: $D(i, j) = D(i', j')$, $D(i, j) > D(i', j')$ and $D(i, j) < D(i', j')$. Thus, these features (which we use to generate the watermark) are robust to reasonable levels of compression. If a proper $D(i', j')$ cannot be found even after scanning all eight directions, leave $B(i, j)$ as initially set to zero. Note: different coefficients may have the same pair member.

Create the $\lceil \frac{M_x}{8} \rceil \times \lceil \frac{M_y}{8} \rceil$ binary matrix B as follows:

$$B(i, j) = \begin{cases} 0 & \text{if } D(i, j) \geq D(i', j') \\ 1 & \text{otherwise} \end{cases} \tag{6}$$

Let the first part of the $\lceil \frac{M_x}{8} \rceil \times \lceil \frac{M_y}{8} \rceil$ authenticator watermark denoted W_{LH} be equal to B.

4. *Permutation:* For security, apply a element-level permutation on B making use of K_S to form the "random" $\lceil \frac{M_x}{8} \rceil \times \lceil \frac{M_y}{8} \rceil$ matrix $\tilde{B}$.
5. *Majority Function:* Reduce the size of B by taking its "raw" characteristics to produce W as follows:

$$W(k) = \begin{cases} 1 & \text{if } \sum_{j=1}^{\lceil \frac{M_y}{8} \rceil} \tilde{B}(k, j) > \lceil \frac{M_y}{8} \rceil / 2 \\ 0 & \text{otherwise} \end{cases} \quad \text{for } k = 1, \ldots, \lceil \tfrac{M_x}{8} \rceil \tag{7}$$

$$W(k) = \begin{cases} 1 & \text{if } \sum_{i=1}^{\lceil \frac{M_x}{8} \rceil} \tilde{B}(i, k - \frac{M_x}{8}) > \lceil \frac{M_x}{8} \rceil / 2 \\ 0 & \text{otherwise} \end{cases} \quad \text{for } k = \lceil \tfrac{M_x}{8} \rceil + 1, \ldots, \lceil \tfrac{M_x}{8} \rceil + \lceil \tfrac{M_y}{8} \rceil \tag{8}$$

6. *Map Function:* Map the $\lceil \frac{M_x}{8} \rceil + \lceil \frac{M_y}{8} \rceil$ length sequence, W to a $\lceil \frac{M_x}{8} \rceil \times \lceil \frac{M_y}{8} \rceil$ binary matrix, W_M using K_S suitable for encryption and watermark embedding.
7. *Encryption:* Use symmetric encryption to encrypt W_M using the secret key K_R (known only to the sender and receiver) to produce a binary matrix W_{HL} of the same dimension.

The W_a embedding strategy aims to be robust to Ω_R and fragile to Ω_F and works as follows:

1. *Two Level Haar DWT:* Take the two level Haar DWT of Y to obtain the $\lceil \frac{M_x}{4} \rceil \times \lceil \frac{M_y}{4} \rceil$ second level LH band and HL bands denoted Y_{2LH} and Y_{2HL}, respectively, as well as the $\lceil \frac{M_x}{2} \rceil \times \lceil \frac{M_y}{2} \rceil$ first level LH and HL bands denoted Y_{LH} and Y_{HL}, respectively.
2. *Group Embedding:* Embed the binary watermarks W_{LH} and W_{HL} in Y_{2LH} and Y_{2HL}, respectively, such that every 2×2 block contains one watermark bit; the sum of the absolute element values in each 2×2 block $S_g(i, j)$ is modified to produce Y_{2LH}^w and Y_{2HL}^w as the follows:

$$S_g(i, j) = \sum_{m=1}^{2} \sum_{n=1}^{2} |Y_{2LH/HL}(n + 2(i-1), m + 2(j-1))|$$

$$q(i, j) = \lfloor \frac{S_g(i, j)}{4\delta} \rfloor \tag{9}$$

$$Y_{2LH/HL}^w(n + 2(i-1), m + 2(j-1)) = \begin{cases} Y_{2LH/HL}(n + 2(i-1), m + 2(j-1)) \\ \text{if } \mathrm{mod}(q(i, j), 2) = W_{LH/HL}(i, j) \\ Y_{2LH/HL}(n + 2(i-1), m + 2(j-1)) \\ +\mathrm{sgn}(Y_{2LH/HL}(n + 2(i-1), m + 2(j-1)))\delta \\ \text{if } \mathrm{mod}(q(i, j), 2) \neq W_{LH/HL}(i, j) \end{cases} \tag{10}$$

where $n, m = 1, 2$, $\mathrm{sgn}(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ -1 & \text{if } x < 0 \end{cases}$ and δ is a user-specified quantization factor.

3. *Image Recomposition:* Recompose the image by taking the appropriate inverse discrete wavelet transforms (IDWTs) to produce the watermarked spatial domain luminance image Y^w.

For reasons of space the reader is referred to [3] for details on W_c generation, embedding and extraction. The overall W_a extraction process follows:

1. *Two level Haar DWT:* Take the two level Haar DWT transform on the received $M_x \times M_y$ luminance Y^r to obtain the second level bands Y_{2LH}^r and Y_{2HL}^r; and the first level bands Y_{LH}^r and Y_{HL}^r.
2. *Group Extraction:* Extract each watermark bit from every 2×2 block of Y_{2LH}^r and Y_{2HL}^r. W_{LH}^e and W_{HL}^e are extracted from Y_{2LH}^r and Y_{2HL}^r, respectively, as follows:

$$S_g(i, j) = \sum_{m=1}^{2} \sum_{n=1}^{2} |Y_{2LH/HL}^r(n + 2(i-1), m + 2(j-1))|$$

$$q(i, j) = \lfloor \frac{S_g(i, j)}{4\delta} \rfloor \tag{11}$$

$$W_{LH/HL}^e(i, j) = \begin{cases} 0 & \text{if } \mathrm{mod}(q(i, j), 2) = 0 \\ 1 & \text{if } \mathrm{mod}(q(i, j), 2) = 1 \end{cases} \tag{12}$$

For authentication, the extracted watermarks W_{LH}^e and W_{HL}^e are compared to a corresponding set generated from Y^r denoted $\hat{W}_{LH}$ and $\hat{W}_M$ (in the same fashion as W_{LH} and W_{HL}, respectively, were from Y). *Authentication matrices* are computed:

$$A_{LH}(i, j) = \hat{W}_{LH}(i, j) \oplus W_{LH}^e(i, j) \tag{13}$$

$$A_{HL}(i, j) = \hat{W}_M(i, j) \oplus W_{HL}^e(i, j) \tag{14}$$

where $\oplus$ is the exclusive OR binary operator and $1 \leq i \leq \lceil \frac{M_x}{8} \rceil$, $1 \leq j \leq \lceil \frac{M_y}{8} \rceil$. Visual inspection of the $A_{LH}(i, j)$ and $A_{HL}(i, j)$ can provide some information on the localization of tampering.

To further assess tampering, we introduce the notions of a credible, processed and fabricated image. A *credible* image is defined as one in which the essential content in is tact (e.g., through perceptual coding). An image is *processed* if the distortions result in extracted watermarks that do not exactly match the generated. An image is considered *fabricated* if the entire content of the image is not credible. Distinctions between R_{HL} and R_{HL} can be used to further characterize the tampering. It should be noted that a user without knowledge of the secret encryption key K_R cannot generate W_{HL} successfully, so A_{HL} ensures the source is legitimate, and A_{HL} and A_{LH} assess the integrity of the received image.

Quantitatively, we propose the use of an authentication statistic $R_a = [R_{LH} \| R_{HL}]$ where the error rates R_{LH} and R_{HL} are defined as the average values of $A_{LH}(i, j)$ and $A_{HL}(i, j)$, respectively over all (i, j). Using a user specified decision threshold $0 < \tau < 0.5$, a tamper categorization on the received image is made as follows: $R_{LH} = R_{HL} = 0$ means that the image content is credible and no modifications have been made; authentication of the sender is verified, $R_{LH}, R_{HL} < \tau$ means that the image content is credible, but the image has been processed, and otherwise we say that the image content is not credible.

Algorithms	Sub attack P_m %	Signal processing attacks P_f %					
		No attack	Hist.Equal.	salt.Pepper	Gaussian	JPEG 70%	Low-pass Filtering
[1]	3.2	0.0	23	1	2.5	3.4	15
[4]	1.0	1.1	31	19.5	1.1	6.7	58
[2]	0.0	0.0	45	80	75	7.2	58
[5]	0.8	0.0	47	0.3	12	45	53
dual domain	0.1	0.0	20	0.7	2.5	0.8	34

Table 1. Comparisons of the authentication capabilities of the proposed combined authentication-compression method.

4. SIMULATION RESULTS

We have tested our algorithm on CH artwork images acquired from an ancient book *"Le Livre des Mille Nuits et une Nuit."* Our algorithm is compared to the following four influential semi-fragile watermarking methods from the research literature [1, 4, 2, 5]. To assess the authentication performance, two figure of merit are used: probability of miss P_m, and probability of false alarm P_f; these standard measures are used to assess baseline performance of authentication watermarking schemes [8].

The error rates are computed over ten different test images each watermarked ten times. The quantization factor is set to $\delta = 12$ which results in a PSNR of 38 dB.

The attacks for which the error rates are computed include those from Ω_R and Ω_F. All tests were conducted using MATLAB. The content-preserving manipulations are well known attacks and include mild compression which we define as JPEG compression at 70% quality factor (which corresponds to a bit rate of 0.5 bpp), additive white Gaussian noise (at 30 dB SNR), 3×3 Weiner filtering (using the function `weiner2` in MATLAB), additive salt and pepper noise (at 1%). In addition, we also tested the approach on histogram equalization (using the function `histeq` in MATLAB). The results for malicious modifications involving sophisticated content substitution are also presented [8].

The results, reported in Table 1 for $\tau = 0.45$ (this value has been experimentally found to be optimal in terms of reducing P_m and P_f), show the overall better performance of the proposed dual domain authentication approach. Our method ranks number one for three of the seven attacks, and number two for the remaining four of the seven attacks. The other methods are each appropriate for different attacks, but do not exhibit the attractive global behavior of the proposed scheme. Furthermore, if a content change occurs, the proposed test cases are correctly able to identify it and its location.

The color recovery results in the presence of lossy compression f_l are also visually promising for the proposed algorithm, but are not reported for reasons of space.

5. CONCLUSIONS

This paper discusses an approach to combine image authentication with compression for the security within a digital watermarking paradigm. The overall algorithm makes use of orthogonal dual domains and compressive data hiding for an integrated algorithm. Application of the approach to real CH imagery provides an indication of the potential of the approach and its improved performance over existing research.

Overall, we have observed that image subspace orthogonality can be exploited in a digital watermarking framework to provide a flexible multipurpose algorithm for both security and compression. Various components can be individually optimized for performance with little interference, but the partitioning of subspaces must be well-suited for the intended application.

6. REFERENCES

[1] C.-S. Lu, H.-Y.M. Liao and C.-J. Sze, "Combined Watermarking for Image Authentication and Protection," *Proc. IEEE Int. Conf. on Multimedia and Expo*, vol. 3, pp. 1415–1418, August 2000.

[2] C.-Y. Lin and S.F. Chang, "A Robust Image Authentication Method Distinguishing JPEG Compression from Malicious Manipulation," *IEEE Transactions on Circuits and Systems of Video Technology*, vol. 11, no. 2, pp. 153–168, February 2001.

[3] P. Campisi, D. Kundur, D. Hatzinakos and A. Neri, "Compressive Data Hiding: An Unconventional Approach for Improved Colour Image Coding," *EURASIP Journal on Applied Signal Processing, special issue on Emerging Applications of Multimedia Data Hiding*, vol. 2002, no. 2, pp. 152-163, February 2002.

[4] M. Goljan, and J. Fridrich, "Invertible Authentication Watermark for JPEG images," *Proc. IEEE Int. Conf. on Information Technology: Coding and Computing*, pp. 223-227, April 2001.

[5] L. Xie, and G.R. Arce, "A Class of Authentication Digital Watermarks for Secure Multimedia Communication," *IEEE Transactions on Image Processing*, vol. 10, no. 11, pp. 1754–1764, November 2001.

[6] J. Dittmann, "Content-Fragile Watermarking for Image Authentication," *Proc.SPIE, Security and Watermarking of Multimedia Content III*, vol. 4314, pp. 175 – 184, Jan. 2001.

[7] D. Kundur and D. Hatzinakos, "Digital Watermarking for Telltale Tamper-proofing and Authentication ," *Proceedings of IEEE Special Issue on Identification and Protection of Multimedia Information*, vol. 87, no. 7, pp. 1167–1180, July 1999.

[8] O. Ekici, B. Coskin, U. Naci and B. Sankur, "Comparative Assessment of Semi-Fragile Watermarking Techniques," *Proc. SPIE, Multimedia Systems and Applications IV*, vol. 4518, August 2001.