COMPENSATED SIGNATURE EMBEDDING BASED MULTIMEDIA CONTENT AUTHENTICATION SYSTEM

Sufyan Ababneh, Ashfaq Khokhar, Rashid Ansari

Dept. of Electrical and Computer Engineering, University of Illinois at Chicago, Illinois, USA

ABSTRACT

Digital content authentication and preservation is an extremely challenging task in realizing decentralized digital libraries. The concept of *compensated signature embedding* is proposed to develop an effective multimedia content authentication system. The proposed system does not require any third party reference or side information. Towards this end, a content-based fragile signature is derived and embedded into the media using a robust watermarking technique. Since the embedding process introduces distortion in the media, it may lead to authentication failure. We propose to adjust the media samples iteratively or using a closed form process to compensate for the embedding distortion. Using an example image authentication system, we show that the proposed scheme is highly effective in detecting even minor modifications to the media.

Index Terms— authentication, preservation, watermarking, compensated signature embedding

1. INTRODUCTION

The unprecedented accessibility of online multimedia contents in recent times has given rise to their vulnerability to corruption and tampering, particularly in an open, peer-to-peer, and possibly malicious environment. Several digital library initiatives have been launched to design and develop technologies for authentication and preservation of contents, format, presentation and functionality of online multimedia documents. In this regard, authentication and preservation of multimedia contents such as video, images and audio is extremely challenging particularly in an open decentralized environment where a central entity is not always available or may have been compromised. We argue that the amenability of non-textual multimedia data to imperceptible insertions offers feature-rich alternatives for authentication and preservation. Depending on the end user application, authentication may or may not tolerate modifications to the original data.

Both fragile and robust watermarking techniques have been used in conjunction with content-based signatures to authenticate a digital media. A disadvantage of these schemes arises from the fact that the very process of embedding a watermark alters the media, causing the subsequent authentication test to fail if the signatures are very fragile (highly sensitive to modifications). To prevent this, most of the schemes either do not use content-based signatures or tend to divide the media domain into two parts: verifiable part from which the signature is derived, and embedding part where the signature is embedded. The level of robustness or fragility is determined based on the application requirements.

Robust watermarking schemes combined with robust content-based signatures have been used for authentication and quality assessment of multimedia [1][2]. However, the target applications for these schemes -such as multimedia retrieval- require tolerance toward normal and incidental distortions while detecting tampering and significant distortions. On the other hand, fragile watermarking based authentication techniques can use high sensitive signatures (fragile) [3] [4]. In such schemes, the watermark is lost and signatures are altered as soon as any modification is applied to the media. It has also been used without content-based signatures. However, in such schemes the use of predefined patterns can not guarantee that no one has intentionally tampered with the media. This is because an adversary can forge a fragile watermark if the embedded pattern is not derived from media contents [5].

In this paper, a new multimedia authentication system is proposed which allows the combining of robust watermarking and fragile signatures to realize a highly effective authentication framework in the absence of a central entity or side information. The system allows for the flexibility to use any robust embedding technique in conjunction with the proposed compensated signature embedding (CSE) concept to provide feature-rich multimedia authentication. This scheme can provide an ideal mechanism for digital multimedia preservation in which verifying the authenticity of the multimedia is an essential step before further attempts for recovery are initiated. The proposed scheme can be applied on any type of digital media including still images, video, audio, text and any combination of them.

2. CSE AUTHENTICATION SYSTEM

The framework of the *compensated signature embedding* scheme (CSE) consists of two main functionalities: encoding and decoding. As shown in Fig.1, the encoder generates a fragile signature, embeds a robust watermark and compensates for signature embedding. The decoder extracts the embedded signature, generates a new signature and evaluates the results. All these functionalities operate on the media signal itself or its transformed version. For convenience, some notation is introduced.

The set of integers is denoted by $\mathcal{Z}$. For any positive integer K, let $I_K = \{k \in \mathcal{Z} : 0 \leq k \leq K-1\}$. The set D denote the domain of the signal. Let M and N be positive integers that are multiples of 2^L, where L is also a positive integer. Define $M_l = M/2^l$ and $N_l = N/2^l$, for $l \in \mathcal{Z}$, $1 \leq l \leq L$. For a raw $M \times N$ image, $D = I_M \times I_N$. We are interested in the subband/wavelet representation of the media using L-level decomposition. Although D defined for the raw media will work, we will define it differently in order to capture the structure of the wavelet decomposition at different levels and we will use image signals for illustration:

$$D = \{\mathbf{n} = (l, n_l, i, j) : l - 1 \in I_L, n_l \in I_{3+\lambda}, i \in I_{M_l}, j \in I_{N_l}\} \tag{1}$$

Where $\lambda = \delta_{lL}$ and δ_{kl} is the Kronecker's $delta$. The signal w (i.e. the image or its transformed version) is defined as a mapping $\{w : D \longrightarrow \mathbb{R}\}$, with some additional structure such as square summability or bounded real valued signal. The signature generation can be performed using some or all of the signal samples. Our choice is to use all samples to avoid leaving the unused samples open to tampering. We choose not to exclude from the signature generation process those samples that are used for embedding since the latter could be a substantial fraction of all samples.

The signature embedding is performed by modifying a subset of the samples. The embedding usually perturbs the signature, therefore, a different subset of the sample values is adjusted to compensate for the signature perturbation. Some more notation is introduced to explain this process.

Let $D_1 \subseteq D$ denote the subset of the signal domain used for generating the signature. As mentioned before, in our study we set $D_1 = D$. The domains of embedding and compensation are denoted by $D_2 \subset D$ and $D_3 \subset D$, respectively. The cardinality of the set D_k is denoted by N_k, $k = 1, 2, 3$. Preferably, D_2 and D_3 are disjoint, i.e. $D_2 \cap D_3 = \phi$, where ϕ denotes a null set. The signals w_1, w_2, and w_3, defined as the restrictions of S to D_1, D_2, and D_3, respectively, are used to conduct the system's operations for signature generation, embedding and embedding compensation.

The operations of signature generation, embedding, and compensation, are represented by σ_1, σ_2, and σ_3, respectively. The operation σ_j, $j = 1, 2, 3$, uses w_j with an optional key $\{k_j\}$ from a set K_j to support a secure operation, and an optional parameter p_j from a set P_j that could support user preferences, such as the level of robustness or other performance measures. An example of using a key for added security is to define options in using w_j for the operation σ_j based on a private key from a set K_j. In practice all signals are square summable over their domain:

$$S = \{w : \sum_{\mathbf{n} \in D} w^2(\mathbf{n}) < \infty\} \tag{2}$$

and

$$S_j = \{w_j : \sum_{\mathbf{n} \in D_j} w_j^2(\mathbf{n}) < \infty\} \tag{3}$$

The signature generation operation σ_1 creates a vector $\mathbf{F}$ of n_1 bits. These bits represent a fragile signature that is obtained from a suitable signal *feature space* such as signal energy or coefficients statistics:

$$\sigma_1 : S_1 \times K_1 \times P_1 \longrightarrow \{0,1\}^{n_1} \tag{4}$$

$$\mathbf{F} = \sigma_1(w_1, k_1, p_1) \tag{5}$$

The signature embedding operation σ_2 consists of modifying the signal samples over the embedding domain with a user-defined process such that the corresponding extraction process performed on the signal with the embedded signature vector reproduces the vector $\mathbf{F}$. The operation σ_2 represents the embedding:

$$\sigma_2 : S_2 \times K_2 \times P_2 \times \{0,1\}^{n_1} \longrightarrow S_2 \tag{6}$$

$$\hat{w}_2 = \sigma_2(w_2, k_2, p_2, \mathbf{F}) \tag{7}$$

Where the signal $\hat{w}_2$ which contains the embedded signature yields $\mathbf{F}$ upon signature extraction. The extracted signature $\bar{\mathbf{F}}$ is obtained through the signature extraction procedure ψ_2:

$$\bar{\mathbf{F}} = \psi_2(\hat{w}_2, k_2, p_2) \tag{8}$$

Where $\hat{w}_2$ is obtained as a result of introducing embedding noise to w_2. Since the embedding technique is robust, this means that $\bar{\mathbf{F}} = \mathbf{F}$. The signature generated after applying the compensation process is obtained by applying the signature generation on $\hat{w}_1$, where $\hat{w}_1$ is the modified w_1 as a result of adjusting w_3:

$$\hat{\mathbf{F}} = \sigma_1(\hat{w}_1, k_1, p_1) \tag{9}$$

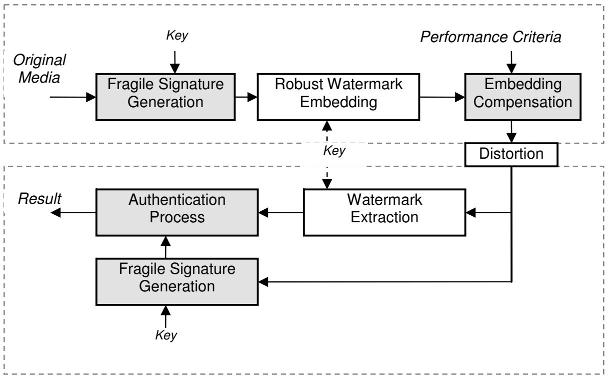


Fig. 1. encoder-decoder and authentication

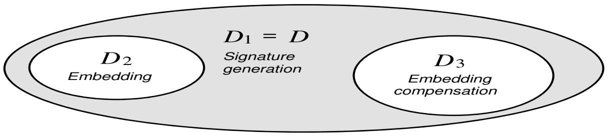


Fig. 2. partitioning the samples domain into three sets

The embedding compensation operation σ_3 consists of modifying the signal samples over D_3.

$$\sigma_3 : S_3 \times K_3 \times P_3 \times \{0,1\}^{n_1} \times \{0,1\}^{n_3} \longrightarrow S_3 \tag{10}$$

$$\hat{w}_3 = \sigma_3(w_3, k_3, p_3, \mathbf{F}, \hat{\mathbf{F}}) \tag{11}$$

Where the signal $\hat{w}_3$ represents the sample values after compensation. The embedded signature $\mathbf{F}$ is used for comparison with $\hat{\mathbf{F}}$ during the compensation process. Since $D_3 \subset D$ and $D_1 = D$, the signal changes that led to $\hat{w}_3$ will also lead to changing w_1 to $\hat{w}_1$. The goal of the compensation operation σ_3 is to generate $\hat{w}_3$ such that the newly generated signature $\hat{\mathbf{F}}$ is identical to the embedded signature $\mathbf{F}$ (i.e. $\hat{\mathbf{F}} = \mathbf{F}$). The fragility of the signature means that minor distortions introduced to the signal will lead to a different signature (i.e. $\hat{\mathbf{F}} \neq \mathbf{F}$).

The embedding compensation process is depicted in Fig. 3. It starts by obtaining the signal samples $w_3 \in S_3$, and then adjusting key-selected samples from this set. The adjustment process could be conducted using an iterative approach or using a closed-form approach [6]. The fact that the signature is fragile will cause $\hat{\mathbf{F}}$ to be different from $\mathbf{F}$ immediately after the embedding operation. The compensation operation adjusts the signal, and the operation is completed when the newly generated signature $\hat{\mathbf{F}}$ and the original media signature $\mathbf{F}$ become identical. The dotted line in the diagram represents an optional iterative adjustment. Performance criteria need to be applied to ensure system effectiveness and signal fidelity such as minimizing distortion based on the characteristics of the human visual system (HVS) in case of visual media, and human acoustic system (HAS) in case of audio.

3. IMAGE AUTHENTICATION USING THE CSE AUTHENTICATION SYSTEM

To illustrate the concepts proposed in the previous section, an image authentication system is developed. We choose the discrete wavelet transform (DWT) as the media sample domain D, where the DWT

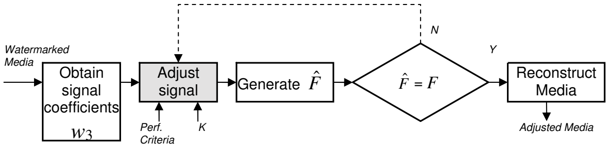


Fig. 3. embedding compensation process

coefficients represents the samples ($w \in S$) used throughout system operations. The energy of the selected coefficients is chosen as an example *feature space* for content-based signature generation. Fig. 4(a) shows a 5-level DWT tree, where HL_l, LH_l, HH_l and LL_l represent the 4 subband sample planes of level l. Any level l represents subbands with a frequency range lower than the frequency range of level $(l-1)$. The samples range S is partitioned into sets S_1, S_2 and S_3 as follow:

$S_1 = \{w_1 : w_1 \in S\}$, $S_2 = \{w_2 : w_2 \in \{HL_5 \cup LH_5 \cup HH_5 \cup LL_5\}\}$, $S_3 = \{w_3 : w_3 \in \{HL_4 \cup LH_4 \cup HH_4\}\}$, where the union $\cup$ of the subbands in this context denotes the union of their samples. S_2 and S_3 are independent of each other, so the compensation process does not impact the embedding process. The choice of subbands allocated to S_3 is not limited to $\{HL_4, LH_4, HH_4\}$ as long as image fidelity is maintained upon compensation completion.

3.1. Signature Generation

The signature generation is performed by operating on samples $w_1 \in S_1$. To provide signature security, a private key could be used to select the samples used for signature generation. Since the goal here is to validate the CSE concept, we choose all the samples for illustration purposes. The final signature is comprised of six concatenated subvectors:

$$\mathbf{F} = [\,\mathbf{f}_1 \,|\, \mathbf{f}_2 \,|\ldots|\, \mathbf{f_6}\,] \tag{12}$$

Where $\mathbf{F} \in \{0,1\}^{n_1}$, $n_1 = 144$. Each subvalue $\mathbf{f}_j$ is then rounded to a 24-bit integer value which forms a subvector $\mathbf{f}_j \in \{0,1\}^{24}$. A total signature size of 144 bits is consistent with the embedding capacity. The samples within the subbands are used for each signature subvalue $\mathbf{f}_j$ according to the following subbands partitions : $\mathbf{f}_1 \leftarrow \{HL_5, HL_4\}$, $\mathbf{f}_2 \leftarrow \{LH_5, LH_4\}$, $\mathbf{f}_3 \leftarrow \{HH_5, HH_4\}$, $\mathbf{f}_4 \leftarrow \{LL_5, HL_1, HL_2, HL_3\}$, $\mathbf{f}_5 \leftarrow \{LH_1, LH_2, LH_3\}$ and $\mathbf{f}_6 \leftarrow \{HH_1, HH_2, HH_3\}$. Each signature subvalue is obtained by calculating the average energy of the samples ${}^j w$, where ${}^j w$ represents the samples that belongs to each partition j of size N_j:

$$\mathbf{f}_j = \frac{1}{N_j}\sum_{i=1}^{N_j} {}^j w_i^2 \quad \longrightarrow i = 1,\ldots,N_j \tag{13}$$

3.2. Information Embedding and Detection

The signature embedding process is performed by employing an existing technique which uses dithered uniform scalar quantization watermarking method in DWT domain [1]. This technique is a special case of a more elaborate class of quantization index modulation (QIM) embedding technique [7]. It provides robustness to moderate distortions and uses two error detection/correction techniques, CRC and BCH, to improve robustness. It also provides security by using a private key. This technique is not the only choice, other embedding techniques ,with higher robustness, can be used as an alternative. To embed one bit of information $m \in \{0,1\}$ into a coefficient w, the coefficient is altered such that:

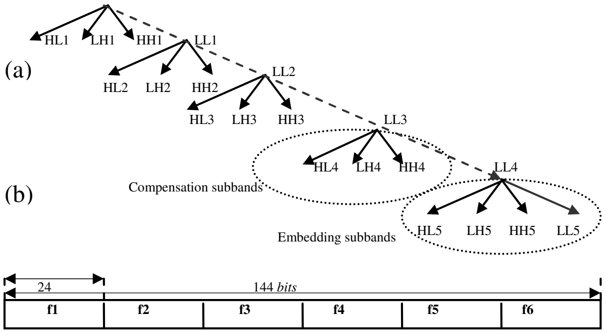


Fig. 4. wavelet decomposition and signature layout f1 to f6

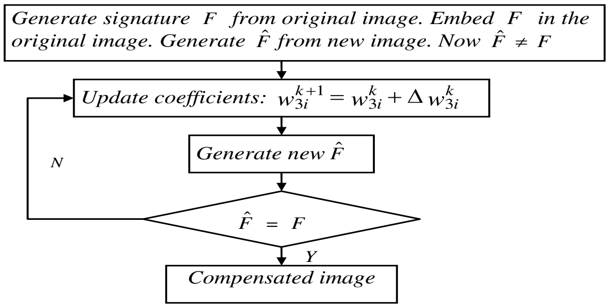

Generate signature F from original image. Embed F in the original image. Generate $\hat{F}$ from new image. Now $\hat{F} \neq F$
Update coefficients: $w_{3i}^{k+1} = w_{3i}^{k} + \Delta\, w_{3i}^{k}$
Generate new $\hat{F}$
N
$\hat{F} = F$
Y
Compensated image

Fig. 5. The flow of the embedding compensation process

$$w' = Q(w + d(m)) - d(m) \equiv Q^m(w) \tag{14}$$

$$d(m) = \begin{cases} -\tau/4 & \text{if } m = 0 \\ +\tau/4 & \text{if } m = 1 \end{cases} \tag{15}$$

Where w' represents the altered coefficient, $Q(.)$ is a base quantization operator with quantization step size τ, and $d(m)$ is a dithering operator. w in this subsection should not be confused with w_j notation which corresponds to a group of coefficients in S_j. At the decoder side, a distorted $\hat{w}$ coefficient is obtained and used to estimate the embedded bit based on the minimum distance criterion:

$$m'(\hat{w}) = arg_{m\in\{0,1\}} min \parallel \hat{w} - Q^m(\hat{w}) \parallel \tag{16}$$

3.3. Embedding compensation

The example embedding compensation operation ,presented here, is performed by adjusting selected wavelet samples $w_3 \in S_3$. An iterative approach is used to adjust the coefficients and generate a new signature $\hat{\mathbf{F}}$ at each iteration. As shown in Fig. 5, the iterative process continues until $\hat{\mathbf{F}}$ becomes identical to the original media signature $\mathbf{F}$ (i.e. $\hat{\mathbf{F}} = \mathbf{F}$). The main adjustment step is:

$$w_{3i}^{k+1} = w_{3i}^{k} + \triangle w_{3i}^{k} \tag{17}$$

Where w_{3i}^k is a coefficient value at iteration k (i.e. $\triangle w_{3i}^k = \alpha_k\ w_{3i}^k$), and α_k is a proportional scaling value (i.e. $\alpha_k = \varepsilon\ \mu_j^k$). ε is an adjustment factor and it is a small positive real number (in our simulation we use value 0.001). μ_j^k represents a scaling ratio calculated for each subvalue f_j at each iteration k, and its main objective is to

System’s operations	MSE	PSNR	CORR
embedding / watermarking	2.655	43.889	0.9995
embedding compensation	2.660	43.878	0.9995

Table 1. MSE, PSNR and *correlation* after each operation

Distortion Type	MSE	PSNR	CORR
localized gaussian blur	4.067	42.037	0.9993
localized gaussian noise	3.675	42.477	0.9994
global gaussian blur	409.755	22.005	0.9290
global gaussian noise	227.952	24.552	0.9632

Table 2. MSE, PSNR and *correlation* after applying distortions

help the compensation process converge smoothly; it starts at an initialized value of 1.0 and continues to decrease to adaptively reduce α_k:

$$\mu_j^k = |\hat{f}_j^k - {f_j}^k| \;/\; |\hat{f}_j^1 - {f_j}^1| \tag{18}$$

$$|\hat{f}_j^k - {f_j}^k| \;\le\; |\hat{f}_j^1 - {f_j}^1| \longrightarrow \mu_j^k \le 1.0 \tag{19}$$

Since the compensation operation is designed to be incremental and adaptive, this helps ensure low complexity and reasonable convergence speed as verified by the experimental results.

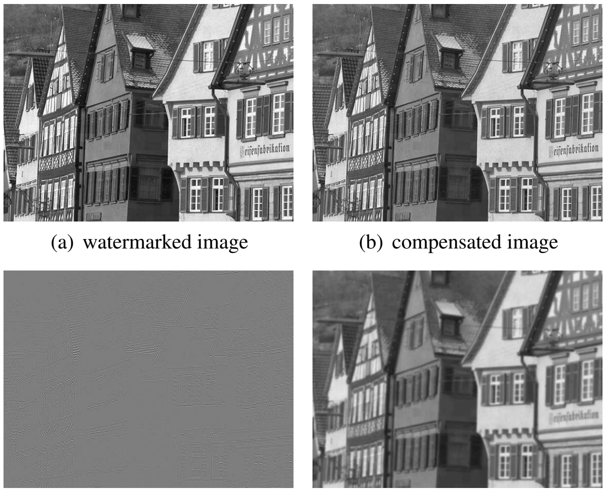
4. EXPERIMENTAL RESULTS

Extensive experiments are conducted to test the effectiveness of the image authentication system in terms of maintaining image fidelity after applying the system’s operations, and detecting tampering and distortions at different levels. Images of size (512x768) have been used. The distortions applied are: *white gaussian noise*, *gaussian blur*, *selective editing* by applying a localized white gaussian noise and localized gaussian blur in small areas (e.g. 20x20 pixels). Table 1 shows image quality measures obtained after embedding and after compensation with respect to the original image. The very slightly lower PSNR after compensation (difference of 0.011) indicates that the compensation impact is negligible, in addition to that, we define the distortion ratio Ω:

$$\Omega = \frac{PSNR_{after\,compensation}}{PSNR_{before\,compensation}} = \frac{43.878}{43.889} = 0.9997. \tag{20}$$

Fig. 6(a), 6(b) and 6(c) show that the image fidelity stays intact during the system’s operations. In addition, an informal visual test was collectively conducted to verify image fidelity and found that there is no visible change in the image.

Fig. 6(d) shows an image with gaussian blur distortion as an example. Table 2 shows different distortions introduced to the image for testing the authentication capabilities of the proposed system. As depicted in the column labeled as CORR (correlation with the original media), these distortions introduce 0.07% to 8% changes in the media. In *all* the cases, our authentication system was able to detect the changes successfully. The information embedding technique was successfully tested for robustness by applying the distortions mentioned above, JPEG and JPEG2000 compressions. The compensation operation converges to the desired solution in 74 iterations, which reflects a reasonable convergence speed for many applications.



(a) watermarked image

(b) compensated image

(c) difference (a)-(b)

(d) global gaussian blur

Fig. 6. Samples of images in different stages

5. CONCLUSION

In this paper, *compensated signature embedding* (CSE) based authentication system is proposed, which is founded on the new concept of using informed embedding compensation. The target media, watermarked with a content-based signature, is adjusted to compensate for the effect of the embedding process to achieve authentication. To illustrate the effectiveness of the proposed concept, an image authentication system is implemented by specifically using an energy content-based signature and embedding the signature in wavelet coefficients. The compensation process is demonstrated using an iterative and adaptive approach. Test results show the validity and effectiveness of the proposed scheme for image authentication while maintaining image fidelity.

6. REFERENCES

[1] Z. Wang, G. Wu, H. Sheikh, E. Simoncelli, E. Yang and A. Bovik ”Quality-Aware Images” IEEE Transaction on Image Processing, Vol. 15, No. 6, June 2006.

[2] Min-Jen Tsai and Hsiao-Ying Hung ”Wavelet Transform Based Digital Watermarking for Image Authentication” IEEE ICIS 2005

[3] Yiping Chu, Yin Zhang, Sanyuan Zhang and Xiuzi Ye ”Region of Interest Fragile Watermarking for Image Authentication” IEEE IMSCC 2006.

[4] Mehmet Celik, Gaurav Sharma, Eli Saber, A. Murat Tekalp ”A Hierarchical Image Authentication Watermark With Improved Localization And Security” Proceedings IEEE ICIP Oct. 2001.

[5] I. Cox, M. Miller and J. Bloom ”Digital Watermarking” by Academic Press, 2002.

[6] Hua Yuan and Xiao-Ping Zhang ”Multiscale Fragile Watermarking Based on the Gaussian Mixture Model” IEEE Transactions on Image Processing, Vol. 15, No. 10, October 2006.

[7] B. Chen and G. Wornell ”Quantization Index Modulation: A class of provably good methods for digital watermarking and information embedding” IEEE Transactions on Information Theory, Vol. 47, No. 3 May 2001.