A BLIND WAVELET BASED DIGITAL SIGNATURE FOR IMAGE AUTHENTICATION[1]

Liehua Xie *Gonzalo R. Arce*
Department of Electrical and Computer Engineering
University of Delaware, Newark, DE 19711, U.S.A
e-mail: xie@ee.udel.edu, arce@ee.udel.edu

ABSTRACT

A content based digital image signature scheme for image authentication is proposed. The signature is implemented by image watermarking in the wavelet domain. We introduce interdependency between the watermark and the image data sequence, so that the detection and verification of the signature are independent of the original image and the knowledge of the original watermark sequence or location are not needed. The signature is embedded and tested within the Entropy Zerotree Wavelet Coding (EZW) model. The information capacity of the algorithm is determined.

1 INTRODUCTION

Digital signatures are electronic protocols used for the authentication of electronic documents whereby a receiver of a message can verify the identity of the sender and the integrity of the message [1]. With the advent of multimedia communications over the internet, it is natural, and critical in many applications, to provide security mechanisms in the transmission of imagery data. To this end, the need of digital image signatures emerges for applications where the security, integrity, and authenticity of images are important. Military and forensic imaging are two such areas where the security features of digital signatures are desirable.

In this paper, we develop a new approach to create digital signatures for imaging that adapts well to applications in multimedia communications. Rather than attaching the signature's bit-string as a header to the image file, we invisibly etch the digital signature into the image data using watermarking methods. Thus, the digital signature is embedded in the image data and cannot be removed by file conversions or image manipulations. Furthermore, the digital signature is "fragile" in that it tolerates small or negligible distortions caused by compression or other simple image manipulations. It does not, however, tolerate other malicious tampering that modifies the content of the image by the addition or removal of objects.

2 DIGITAL SIGNATURE SCHEME

In our previous work [2], we analyzed how the interdependency between the watermark and image data sequence affects the decoding of the watermark and argued that memory should be introduced during the coefficient's transformation in order to extract the watermark without previous knowledge of the original data. Here we propose a digital image signature scheme which introduces memory leading to a "blind" authentication algorithm. It exploits the rank order relationship in local areas throughout the entire image to encode the signature.

2.1 Binary Engraving Method

In essence, the proposed method engraves a digital signature by changing the median of a local area to a value set by its neighbors. Suppose we have the pixels of an image in a local window b_1, b_2, b_3 and the sorted elements $b_{(1)}$, $b_{(2)}$ $b_{(3)}$ in ascending order. We first split the range between $b_{(1)}$ and $b_{(3)}$ into intervals of length: $space \overset{\triangle}{=} \alpha\frac{b_{(1)}+b_{(3)}}{2}$, then we change the median according to the watermark as well as the region in which the median falls. α is a tuning parameter and the default value is 0.05.

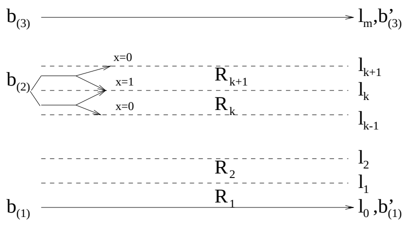

Figure 1: The median falls into different regions and adjusts the value of itself according to the watermark

Suppose the boundary of the regions are numbered from the bottom to the top by $l_0, l_1, \cdots, l_k, \cdots, l_M$. Let

[1] Prepared through collaborative participation in the Advanced Telecommunications and Information Distribution Research Program (ATIRP) Consortium sponsored by the U.S. Army Research Laboratory under the Federated Laboratory Program, Cooperative Agreement DAAL01-96-2-0002.

$l_k \triangleq$ the kth boundary, $R_i \triangleq$ the ith region. $l_0 = b_{(1)}$, $l_1 = b_{(1)} + space$, $\cdots$, $l_k = b_{(1)} + kspace$, $\cdots$, $l_M = b_{(3)}$. Let's denote the modified coefficient b_2 modified by the algorithm as $b'_{(2)}$. The median which falls into R_k that is bounded by l_{k-1} and l_k will be replaced by :

$$b'_{(2)} = \begin{cases} l_k & \text{case } A \\ l_{k-1} & \text{case } B, \end{cases}$$

where
case $A \triangleq k$ is odd and x is 1 or k is even and x is 0.
case $B \triangleq k$ is even and x is 1 or k is odd and x is 0.
and x is one bit of the watermark.

By sliding the window through the entire image, the watermark bits are etched in.

2.2 Digital Signature In A Wavelet Transformed Image

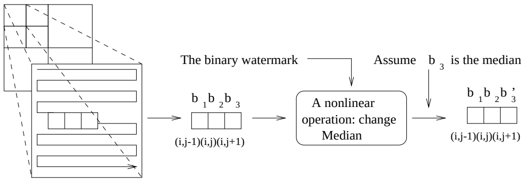

Figure 2: The watermark engraving structure

We implemented our signature scheme in a wavelet transformed image. First we generate the signature by combining a hashed image content information with public key encryption. The encrypted signature obtained is in the form of binary data and the wavelet decomposition coefficients will be changed according to this binary sequence. The signature is engraved in the top level of the pyramid (LL component) of the wavelet transformed image. The reason why we choose the LL band is due to a consideration of robustness. The watermark engraving structure is displayed in Fig. 2. A fixed size window runs without overlapping through the entire low frequency band of the wavelet transformed image. For ease of explanation, we assume a 3×1 window whose elements are denoted as b_1, b_2, b_3 which stand for the coefficients' value at locations with coordinate $(i-1, j)$, (i, j), $(i+1, j)$. The nonlinear transformation algorithm as we stated in last section is performed on the median of these three coefficients (See Fig. 3). Let's denote the modified coefficient b_2 by $b'_{(2)}$, then $b'_{(2)}$ is the function of $b_{(1)}$ and $b_{(3)}$:

$$b'_{(2)} = f(\alpha, b_{(1)}, b_{(3)}, x). \tag{1}$$

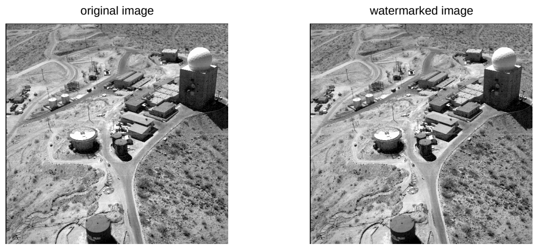

Figure 3: Left: the original image. Right: watermarked image with 2-level decomposition and $\alpha = 0.05$

At the receiver, watermark extraction is performed by inverting the engraving procedures. A rectangular window of the same size as the one that used in watermark engraving is applied. A sequence with elements: $B_{(1)}$, $B_{(2)}$ and $B_{(3)}$ can be obtained as the window is shifted. Since $B_{(1)}$ and $B_{(3)}$ are unchanged, we can get two possible values of $B'_{(2)}$:

$$B'_{(2)} = \begin{cases} f(\alpha, B_{(1)}, B_{(3)}, 0) & \text{if } x = 0 \\ f(\alpha, B_{(1)}, B_{(3)}, 1) & \text{if } x = 1 \end{cases}$$

where x is the possible value of watermark sample.

We compare the distance between $B_{(2)}$ and the outcome at $x = 0$ with the distance between $B_{(2)}$ and the outcome at $x = 1$, and we select the x which makes $B'_{(2)}$ closest to $B_{(2)}$. There is no need of the original image to retrieve the signature. The extraction is simpler than engraving, which is desirable for real time processing.

This procedure is repeated by shifting the window through the entire image. At the end, we have obtained the watermark from the unknown image. If encrypted when inserting the watermark, the binary sequence obtained must be subsequently decrypted. Furthermore, the image received for authentication is hashed using the same hash as was used to sign the original image. We compare it with the content information the watermark carries. If the two achieve high similarity, we can then declare the authenticity of the image. Otherwise, the image will be regarded as tampered.

2.3 Digital Signatures in a Wavelet Compression Scheme

We embedded the proposed scheme into the EZW coding algorithm [3]. The signature engraving has been modified to adapt to the compression scheme mainly because of the different level of quantization and coefficients truncation at different compression ratio. The watermark location is not limited in the top-level pyramid of wavelet transformed image. It's extended to any significant local area. The local area is significant if the image data in the area contains a margin larger than the size of quantization threshold. This extension is adaptive to quantization and it's now possible to embed more information bits.

2.4 Multi-bit Engraving

With a slight change, we can obtain "multi-bit engraving" within the compression watermarking framework. The principle of multiple bit engraving is the same as our binary engraving scheme. Information can be hidden in a local area where its variance overcomes quantization. Hence, a binary bit 0 or 1 can be hidden if the difference between two coefficients is larger than the quantization constant. The signature we engrave is represented by binary data so we engrave one bit each time when appropriate. Actually, the capacity can be greatly increased if we engrave as many bits as we can. When the difference is several times larger than the quantization constant, we could refine the engraving process and split the previous division into portions with the size of the quantization constant. Formerly, we inserted two values only: 0 and 1. Now if the distance is twice as large as that of the quantization constant, 0, 1 and 2, a 3-ary symbol can be inserted. This approach is called "multiple bit engraving". By using multi-bit engraving, we are able to insert more information.

3 WATERMARK BIT CAPACITY

The bit rate of a watermark refers to the amount of information that can be embedded within an image. The bit rate is measured in bits per pixel. Let's view a compressed image as an approximately continuous, band-limited channel. An image I is regarded as the signal and the quantization truncation becomes the source of noise.

We will show in this section that the information capacity for multi-bit engraving scheme is bounded by the rate given by Shannon's channel capacity theorem, although a different interpretation of noise is used. First, recall Shannon's capacity result [4]: Let P be the average transmitter power, and suppose the noise is white thermal noise of power N in the band W. By sufficiently complicated encoding systems it is possible to transmit binary digits at a rate

$$C \leq W \log_2 \ (1 + \frac{S}{N}) \tag{2}$$

where C is the channel bit rate, S is the signal power, N is the noise power and W is the bandwidth.

The inequality sets up the upper bound on the possible rate for an error-free communication channel subject to Gaussian noise. The principle Shannon used to derive the upper limit is that two symbols modulated at the sender in a communication channel must have a noise margin to get the correct decision at the receiver. With the assumption of white Gaussian noise, Shannon showed that the capacity problem can be converted to the following: Find the maximum rate of signals which satisfy:

1. Its average power is P and band-limited to W.
2. The distance between two signals in their geometrical representation is larger than the noise of power N.

In our model, we view the compressed image as an approximately continuous, band-limited channel. Let's assume the signal, i.e. the image I, has a power P and the quantization constant is q, the truncation threshold of the wavelet coefficients. How many bits of information can we embed? An equivalent conclusion can be reached by analyzing the watermark engraving. A watermark bit can be inserted if and only if the distance between the maximum and the minimum coefficients in the selected window is large enough to overcome the quantization. It is clear that whether or not a watermark signal can be engraved is determined by the distance of two signals. Thus, we build on a similar logical foundation. The information bit rate we can achieve by our engraving scheme is equal to the maximum rate of signals with noise power $N = q^2$. The following evolves directly from Shannon's theorem:

*Theorem:*The information bit rate we can engrave in an image is bounded by

$$B \leq W \ \log_2 \frac{P+N}{N}, \tag{3}$$

where B is the information bit rate, P is the image power, $N = q^2$, q is the quantization constant, and W is the image bandwidth.

An intuitive understanding of this result is important. When P is very small such that it goes to 0, $log_2 \frac{P+N}{N}$ goes to 0. The result fits for the type of images with constant background, in other words, with very small variance. On the contrary, when P is larger, the capacity is larger. So images with larger variance hide more information.

3.1 Information Capacity for Binary Engraving Scheme

In the coding algorithm, the most significant bit is coded first. So the quantization starts from the largest quantization constant and is refined gradually until the desired compression ratio is met. The quantization process is described by a descending sequence $q(i)$ $(i = 1 \ to \ n)$. $q(i)$ is the quantization constant used at the ith refining process. $q(n)$ is the last quantization constant when coding ends. Each $q(i)$ is the power of 2 and $q(i) = 2q(i+1)$, $q(0)$ is the largest quantization constant. We also define sequence $B(i)$ and $b(i)$. $B(i)$ and $b(i)$ are the bit rate at the ith refining process for multi-bit engraving and binary engraving respectively.

We derive the relation between the bit rate sequence of multi-bit and binary engraving. The proof of the next theorem is given in [5]:

Theorem: The bit rate of binary engraving is related to the bit rate of multiple bit engraving by:

$$\mathcal{O}(b(i+1)) = \mathcal{O}(B(i+1)) - \mathcal{O}(B(i)) \tag{4}$$

$$\mathcal{O}(B(i+1)) = \mathcal{O}(b(i+1)) + \mathcal{O}(b(i)). \tag{5}$$

This can be proved based on the assumption that $b(i) \approx B(i)$ at $i+1$, i.e. $B(i)$ and $b(i)$ are far less than

$B(i+1)$. The equations can be used to predict $B(i)$ and $b(i)$. We can predict $B(i)$ if we know $b(i)$, or we can predict $b(i)$, if we know $B(i)$. The latter is particularly useful since we have derived an upper bound for $B(i)$.

3.2 Experimental Results Towards Bit Rate

We studied the relationship between bit rate vs PSNR of the transformed image as compared to the original image and the observation on the relationship between bit rate vs quantization level. The experimental results also confirm our theoretical results. The assumptions we make for the estimation are reasonable as the experimental results show. Fig. 4 plots the bit rate versus $\log_2(1+\frac{P}{q^2})$. The higher the value of $\log_2(1+\frac{P}{q^2})$, the higher the amount of the information bits per image. The bit rate results include: experimental result using multi-bit engraving scheme, experimental result using binary engraving scheme and the upper bound we estimated using experimental data. The experimented image is the image in Figure 3. P and q are defined previously. The signal power is estimated by the image's variance that is measured in the pixel domain, treating the image as a data matrix. The sequence of quantization constants is 64, 32, 16, 8, 4, 2, 1. The upper bound is obtained by measuring the image's bandwidth in the FFT transformed domain to approximate W. We get the bandwidth by plotting the discrete power spectrum distribution of the image. The bandwidth is obtained at a point where the ratio of P_{max} and P_w is 30db. P_{max} is the maximum power, which is usually the image's power at the lowest frequency.

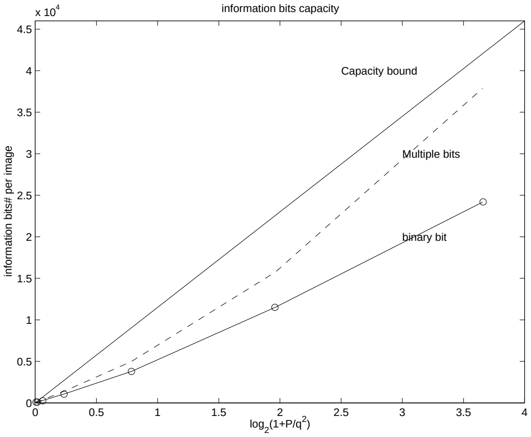

Figure 4: The relation of bit rate and quantization using binary and multi-bit engraving

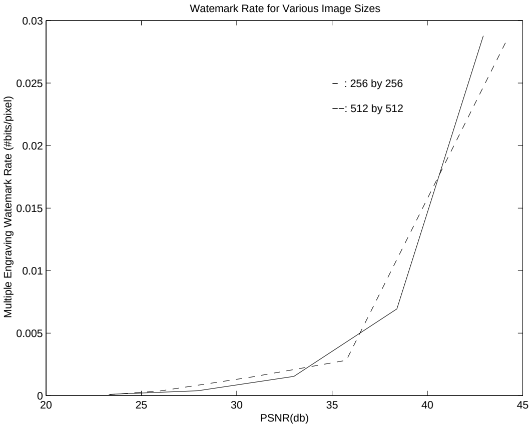

Figure 5: The relation between the information rate and PSNR

A second experiment which illustrate the information bit rate is depicted in Fig. 5. We measured the watermark bit rate in bits per pixel and PSNR for images with different sizes. We found that the higher the PSNR, the higher the watermark bit rate is. Thus, more information bits can be hidden when there is less compression. As we can see, the bit rates for the 512 by 512 image and the 256 by 256 image are comparable with the same PSNR. This result implies the relation between the SNR and the bit rate as well.

4 OTHER EXPERIMENTAL RESULTS

We tested the digital signature on its robustness by attacking the image in various ways. We tried to rescale the image and the signature can be recovered when we restore the manipulated image by resizing or padding. We repeatedly compressed the image with the signature and extracted the embedded information out each time. The compression ratio changed each time. As a result, the signature stayed in the image if the compression ratio used in later compression was lower than or equal to the compression ratio we applied when the signature was engraved. A comparison of the MSE and PSNR of the original image versus the compressed image with and without the signature was computed [2]. There is little difference brought by the signature engraving.

The edge information based signature system was developed and implemented by our signature scheme. The system can detect the deliberate tampering imposed on the image. Because of the multiresolution decomposition performed by wavelet transform, the system is able to roughly point out the place where the image has been tampered.

References

[1] A. Furche and G. Wrightson. *Computer Money.* Verlag fur Digitale Technologie, Heideberg:dpunkt, 1996.

[2] G. R. Arce L. Xie. A blind content based digital image signature. *Proceedings of the 2nd Annual Fedlab Symposium on Advanced Telecommunications/Information Distribution*, 1, College Park, MD, Feb. 1998.

[3] J. M. Shapiro. Embedded image coding using zerotrees of wavelet coefficients. *IEEE Trans. Signal Processing*, Vol. 41:3445–3462, Dec. 1993.

[4] C. E. Shannon. Communication in the presence of noise. *Proc. IRE*, 37, Jan. 1949.

[5] G. R. Arce L. Xie. A blind digital image signature in wavelet compression. *Proceedings of the IEEE International Conference on Image Processing*, Chicago, IL, Oct. 1998.