A Robust Digital Image Watermarking Method using Wavelet-Based Fusion

Deepa Kundur and Dimitrios Hatzinakos
Department of Electrical and Computer Engineering
University of Toronto
Toronto, Ontario, Canada M5S 3G4

Abstract

We present an approach for still image watermarking in which the watermark embedding process employs multiresolution fusion techniques and incorporates a model of the human visual system (HVS). The original unmarked image is required to extract the watermark. Simulation results demonstrate the high robustness of the algorithm to such image degradations as JPEG compression, additive noise and linear filtering.

1 Introduction

With the increase in the availability of digital data such as multimedia services on the Internet, there is a pressing need to manage and protect the illegal duplication of data. One approach to address this problem involves adding an invisible structure to a *host image* to "mark" ownership of it. These structures are known as digital watermarks. To be effective, a watermark must be imperceptible within its host, discrete to prevent unauthorized removal, easily extracted by the owner, and robust to incidental and intentional distortions.

In this paper we address the watermarking of still image data. There are many benefits to embedding a watermark in an image. The use of digital watermarks can be used as an authentication tool, and as a method to discourage the unauthorized copying and distribution of electronic documents. Most of the recent work in watermarking can be grouped into two categories: spatial domain methods [1], [2], and frequency domain methods [3], [4],[5]. There is a current trend towards approaches that make use of information about the human visual system (HVS) [4], [5] to produce a more robust watermark. Such techniques use explicit information about the HVS to exploit the limited dynamic range of the human eye.

We introduce a novel approach of watermarking that also accounts for the characteristics of the HVS. Unlike [4] and [5] our approach involves embedding the watermark in the discrete wavelet domain. We make use of a multiresolution data fusion approach in which the image and watermark are both transformed into the discrete wavelet domain. The resulting image pyramids are then *fused* according to a series of combination rules that take into account the characteristics of the HVS.

The fundamental advantage of the data fusion approach lies in the method used to merge the watermark at the various resolution levels. This approach provides a simultaneous spatial localization and frequency spread of the watermark within the host image. In addition, the watermark merging process is adaptive as it depends on the local image characteristics at each resolution level, and is robust as it embeds the watermark more strongly into more salient components of the image. The combined result of these factors makes the proposed method attractive.

In the next section we introduce the proposed approach. In Section 3 we provide some results demonstrating the high robustness of the approach to JPEG compression, additive noise and linear filtering. Final remarks are provided in Section 4.

2 Proposed Watermarking Approach

2.1 General Description

The proposed method employs a multiresolution wavelet decomposition of both the host image and the watermark. When an image undergoes a wavelet decomposition, its components are separated into bands of approximately equal bandwidth on a logarithmic scale much as the retina of the eye splits an image into several components. It is, therefore, expected that use of the discrete wavelet transform will allow the independent processing of the resulting components much like the human eye.

For this reason, the use of wavelet decompositions for the *fusion* of images is popular. *Fusion*, or more specifically, *data fusion* refers to the processing and synergistic combination of information from various

0-8186-8183-7/97 $10.00

knowledge sources and sensors to provide a better understanding of the situation under consideration.

Existing literature has shown the usefulness of wavelets for data compression and data reconstruction. Since both image fusion and watermarking are essentially sensor-compressed information problems (i.e., they involve the merging of many images to a single fused result which contains the most important elements), it follows that wavelets are also useful for data merging. Existing literature on multiresolution wavelet-based fusion algorithms demonstrates that the approach is superior to other image merging techniques. Wavelet fusion methods can make use of information about the HVS to determine what information, from each image, is important to retain in the composite [7]. It is then expected that somewhat complementary HVS rules can be used to robustly embed a watermark imperceptibly inside the host.

2.2 Three Stage Method

We assume that the watermark to be embedded is a two-dimensional array of real or integer numbers. For robustness, it is desirable that the watermark have characteristics which are "noise-like". For simulations, we use binary watermarks comprised of $N \times N$ arrays of ones and negative ones. It is required that the size of the watermark in relation to the host image be "small". We assume, without loss of generality, that the watermark is smaller than the host by a factor of 2^M, where M is any integer greater or equal to one. We also assume that the dimensions of the watermark are $2N_{wx} \times 2N_{wy}$. We use $f(m,n)$ to denote the host image and $w(m,n)$ the watermark. The technique is comprised of the three main stages discussed below.

Stage I: The host image and the watermark are transformed into the wavelet domain. Specifically, we perform the Lth level discrete wavelet decomposition of the host image to produce a sequence of $3L$ *detail* images, corresponding to the horizontal, vertical and diagonal details at each of the L resolution levels, and a gross approximation of the image at the coarsest resolution level. L can be any positive integer less than or equal to M. We denote the kth detail image component at the lth resolution level of the host by $f_{k,l}(m,n)$ where $k = 1,2,3$ and $l = 1,\ldots,L$.

Only the *first* level discrete wavelet decomposition of the watermark is performed producing three detail images and an approximation. Similarly the resulting $N_{wx} \times N_{wy}$ detail coefficients are denoted by $w_{k,1}(m,n)$.

Stage II: The detail images of the host at each resolution level are segmented into non-overlapping $N_{wx} \times N_{wy}$ rectangles. We denote the segments by $f^i_{k,l}(m,n)$ where $i = 1,\ldots,2^{2(M-l)}$. The *salience* S (which is a numerical measure of perceptual importance) of each of these localized segments is computed using information about the HVS.

The watermark is embedded by a simple scaled addition of the watermark to the particular $N_{wx} \times N_{wy}$ detail component as described in Section 2.3. The scaling of the watermark is a function of the salience of the region. The greater the salience S, the stronger the presence of the watermark. The computation of S is described in Section 2.3.

Stage III: The corresponding Lth level inverse wavelet reconstruction of the fused image components is performed to form the watermarked image.

A general overview of the method is provided in Figure 1.

The watermark is extracted from the possibly corrupted watermarked image using the host image, by applying the inverse procedure at each resolution level to obtain an estimate of the watermark. The estimates for each resolution level are averaged to produce an overall estimate of the watermark.

2.3 The Merging Process

In this section we discuss the details of the watermark merging process which is performed in the second stage of the proposed method. Mathematically, contrast sensitivity is defined as the reciprocal of the contrast necessary for a given spatial frequency to be perceived. For this paper, we assume the well-known model given by Dooley [6]. We extend the model to two dimensions using the same approach as [7]. The resulting contrast sensitivity for a particular pair of spatial frequencies is given by:

$$C(u,v) = 5.05e^{-0.178(u+v)}(e^{0.1(u+v)} - 1), \qquad (1)$$

where $C(u,v)$ is the contrast sensitivity matrix and u and v are the spatial frequencies given in units of cycles per visual angle (in degrees). A conversion from cycles per visual angle to radians per pixel must be made prior to the use of C. For the simulations in this paper, we apply a conversion assuming a 256×256 host image and a viewing distance of 6 times the image size. C, therefore, provides an estimate the contrast sensitivity of the average human observer under these conditions. Once C is formed, we calculate the salience of the image components.

We define a mathematical quantity to measure the importance of an image component. This measure is called *saliency.* This quantity, similar to what is used in [7] to perform perceptual-based image fusion, is de-

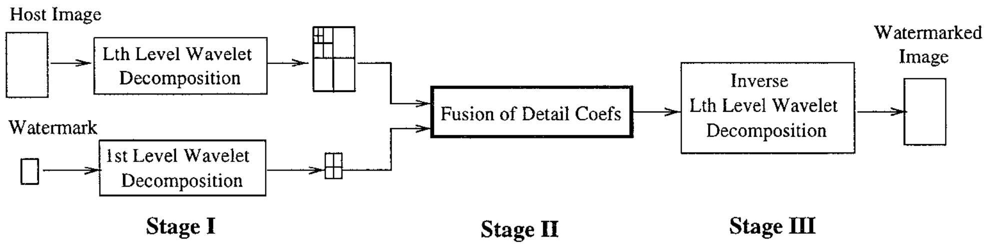

Figure 1: Proposed Watermarking Method.

fined as

$$S(f^i_{k,l}(m,n)) = \sum_{\forall(u,v)} C(u,v)|F^i_{k,l}(u,v)|^2, \quad (2)$$

where C is the contrast sensitivity matrix, and $F^i_{k,l}(u,v)$ is the discrete Fourier transform of the image component $f^i_{k,l}(m,n)$. After the salience is computed, the watermark is embedded using the following equation:

$$g^i_{k,l}(m,n) = f^i_{k,l}(m,n) + \gamma_{k,l}\sqrt{S(f^i_{k,l}(m,n))}w_{k,1}(m,n). \quad (3)$$

The user-defined parameters $\gamma_{k,l}$ for $l = 1, \ldots, L$, are positive real numbers which determine a trade-off between the visibility of the watermark and its robustness to signal distortion at each of the resolution levels. The following rule of thumb was determined to set these parameter values:

$$\gamma_{k,l} = \frac{\alpha}{\max\limits_{\text{over all }(m,n)} \sqrt{S(f^i_{k,l}(m,n))}}, \quad (4)$$

where α is 10% to 20% of the mean value of the host image. For the simulations in this paper, α was set to 10% of the mean value of the image. Equation 3 suggests that the watermark is embedded more strongly in the more salient image components, which should make the technique more robust to image distortions.

3 Simulation Results

Simulation results were conducted on the 256×256 host image shown in Figure 2(a). The 16×16 (32 bytes) binary watermark given in Figure 3 was embedding using the proposed technique for $L = 4$. The watermarked image is shown in Figure 2(b) and is perceptually identical to the original host. Simulation results were conducted to demonstrate the robustness of the technique to JPEG compression, additive noise and two-dimensional linear mean filtering. The robustness of the technique is evaluated by comparing the normalized correlation coefficient of the extracted watermark with the true one. The normalized correlation coefficient is defined as

$$\rho(w,\phi) = \frac{\sum_{\forall(m,n)} w(m,n)\phi(m,n)}{\sqrt{\sum_{\forall(m,n)} w^2(m,n)}\sqrt{\sum_{\forall(m,n)} \phi^2(m,n)}}, \quad (5)$$

where $\phi(m,n)$ is the extracted watermark and $w(m,n)$ is the watermark to detect for.

Figure 4(a) shows the effect of compression on the correlation coefficient. The correlation coefficient remains high for reasonable compression ratios. Visual image degradation occurs at compression ratios greater than 10. Severe image degradation in which the features of the face were not distinguishable occurred for compressions ratios of 34 and above. The results show that the watermark still remains present. Although the correlation coefficient reduces to 0.6, the correlation of the extracted watermark with 2000 other randomly generated binary watermarks produces correlation coefficients between -0.23 and 0.23 which still remain significantly lower than 0.6.

Figure 4(b) provides the results for degradation using additive white Gaussian noise. The proposed method performs well in the presence of additive noise. Severe visual image degradation occurred at signal-to-noise ratios of 15 dB and greater. Although the image appeared overwhelmed by noise, the watermark can be detected with a correlation of 0.85. The 2000 other randomly generated watermarks were also correlated with the extracted one to give correlation coefficients between -0.24 and 0.24.

The results for degradations from linear mean filtering are also presented in Figure 4(c). The watermarked image was filtered with a $K \times K$ linear mean filter. The results for filter dimension K values from 1 to 6 are shown. Highly noticeable image degradation began to occur for $K \geq 4$. The watermark can still be detected with a correlation of 0.42. The 2000 randomly generated watermarks produced corre-

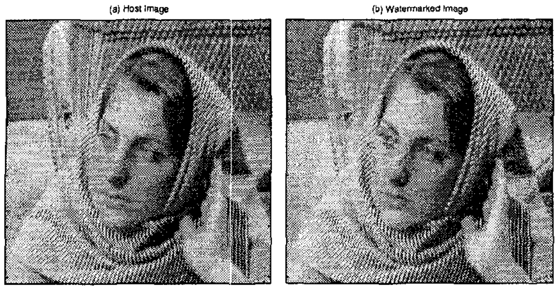

Figure 2: (a) Host Image (left), (b) Watermarked Image (right).

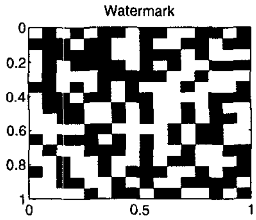

Figure 3: The 256 bit embedded watermark.

lations between -0.25 and 0.25. Because linear filtering reduces the details in the image, the watermarks were extracted and averaged only from the lowest and second lowest resolution levels (i.e., $l = 3, 4$).

4 Conclusions

In this paper we propose a robust method of still image watermarking based on concepts from wavelet-based data fusion. The proposed technique is highly robust to compression and additive noise. In fact, the images are almost completely destroyed, yet the watermark can be extracted fairly accurately. The approach is also quite resilient to moderate linear mean filtering.

Future work will concentrate on making the method more practical by modifying the technique such that the host image is not required to extract the watermark.

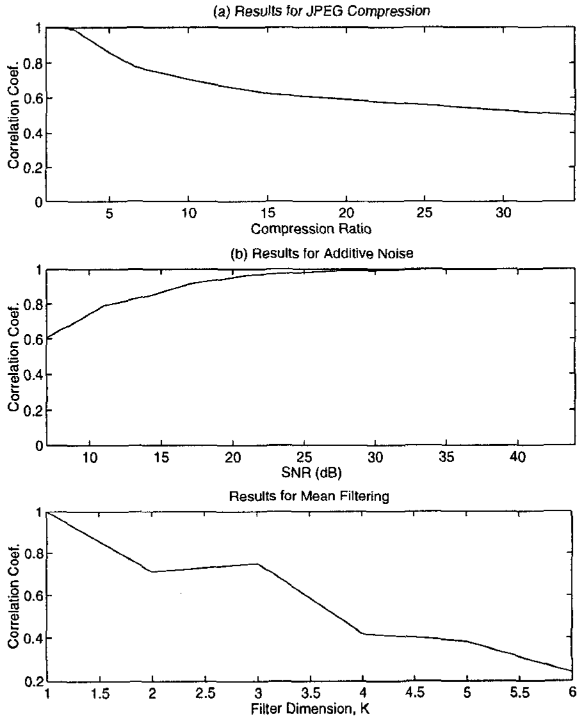

Figure 4: Correlation coefficient of the extracted watermark with the original for varying (a) Compression Ratios, (b) SNRs, and (c) Linear Mean Filtering Lengths.

References

[1] I. Pitas, "A method for signature casting on digital images," *Proc. IEEE Int. Conf. on Image Processing*, vol. 3, pp. 215-218, 1996.

[2] R. B. Wolfgang and E. J. Delp, "A watermark for digital images," *Proc. IEEE Int. Conf. on Image Processing*, vol. 3, pp. 219-222, 1996.

[3] I. J. Cox, J. Kilian, T. Leighton and T. Shamoon, *Secure spread spectrum watermarking for multimedia*, NEC Research Institute, Technical Report 95-10.

[4] M. D. Swanson, B. Zhu and A. H. Tewfik, "Transparent robust image watermarking," *Proc. IEEE Int. Conf. on Image Processing*, vol. 3, pp. 211-214, 1996.

[5] J.-F. Delaigle, C. De Vleeschouwer and B. Macq, "Digital watermarking", *Proc. SPIE, Optical Security and Counterfeit Deterrence Techniques*, vol. 2659, pp. 99-110, February, 1996.

[6] M. D. Levine, *Vision in Man and Machine*, McGraw-Hill, Toronto, 1985.

[7] T. A. Wilson, S. K. Rogers and L. R. Myers, "Perceptual-based hyperspectral image fusion using multiresolution analysis," *Optical Engineering*, vol. 34(11), pp. 3154-3164, November 1995.