Extraction of Multiresolution Watermark Images for Resolving Rightful Ownership

Wenjun Zeng*, Bede Liu+ and Shawmin Lei*

*Sharp Laboratories of America, 5750 NW Pacific Rim Blvd., Camas, WA 98607, USA

+Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA

ABSTRACT

Digital watermarking has been recently proposed as the mean for intellectual property right protection of multimedia data. We present some ways to "visualize" the invisible watermarks, both statistically and perceptually, for proving the ownership. A system which is capable of embedding a good resolution *meaningful* binary watermark image and later extracting different versions of that watermark image with varying resolutions is proposed. The system has the nice feature that the watermark detector (rather than encoder) is allowed to *adaptively* choose the trade-off between robustness degree and resolution of the extracted watermark image. It takes advantage of the high spatial correlation of the watermark image and the human visual system's super ability to recognize a correlated pattern to enhance the detection performance. While a statistical technique which can quantify the false alarm detection probability should be considered as a fundamental measure for a valid ownership claim, the ability to extract a *meaningful* watermark image will greatly facilitate the process of convincing the jury of an ownership claim.

Keywords: multiresolution watermark, scalable visual detection, rightful ownership, copyright protection, multimedia security

1. INTRODUCTION

The rapid growth of digital imagery, the increasingly easy access to digital media, and the increasingly powerful tools available for manipulating digital media have made media security a very important issue. Digital watermarks have been proposed recently as the means for intellectual property right protection of multimedia data. Digital watermarking is a process of embedding information (or signature) directly into the media data by making small modifications to them. With the detection/extraction of the signature from the watermarked media data, it has been shown that digital watermarks can be used to identify the rightful owner, the intended recipients, as well as the authenticity of a media data.[1–6]

In general, there are two basic but conflicting requirements of invisible watermarks. The watermarks should be perceptually invisible. They should also be robust to common signal processing and intentional attacks. Perceptual models have been incorporated to achieve the best tradeoff between imperceptibility and robustness to signal processing.[3,4,7,8] In,[5,6] we proposed a watermarking system which is able to detect the watermarks without directly involving the original image for the particular application of resolving rightful ownership. This system makes any counterfeit scheme impossible, and the watermark detector outputs a value which truly quantifies the false alarm detection probability.

This paper focuses on the particular application of resolving rightful ownership of digital images using invisible watermarks. We present some ways to "visualize" the invisible watermarks, both statistically and perceptually. We propose a system which is capable of embedding a good resolution meaningful binary watermark image and later extracting different versions of that watermark image with varying resolutions. While a large detector output value which quantifies the false alarm detection probability should be considered as a fundamental measure for a valid ownership claim,[5,6] the ability to extract a *meaningful* watermark image is very helpful in convincing the jury in the court for the claim of an ownership. Since the jury usually consist of non-technical people, the presentation of an extracted *meaningful* watermark image is much more convincing than a numerical value. An additional advantage is that it provides the opportunity to exploit the human visual system's super ability to recognize a correlated pattern. This advantage has been initially discussed in.[9] However, instead of directly extracting a

To appear in *IS&T/SPIE Electronic Imaging'99:Security and Watermarking of Multimedia Contents. Proc. of SPIE*, vol. 3657, 1999. Further author information: send correspondence to W. Zeng: E-mail: zengw@sharplabs.com.

watermark image, the super recognition ability of the human visual system is only used in[9] to visualize some detection results. In this paper, we allow the watermark decoder to directly extract a *meaningful* binary watermark image from the test image to prove the ownership. The human visual system's super recognition ability is exploited to enhance the possibility of convincing the jury. Therefore the system is virtually more robust to common signal processing. Furthermore, we propose a multiresolution watermark image extraction scheme. The observation here is that different images may tolerate different amount of watermarks without revealing visual artifacts, and more importantly, the watermarked image may undergo different types of processing before it is inputted as the test image to the watermark detector. From information theoretical point of view, the watermarks can be treated as the signal to be transmitted, and the original image can be considered as the transmission media. However the channel capacity is varying, depending on the original image characteristics, and more importantly on the amount of signal processing applied to the watermarked image. Since the channel capacity is unknown at the time of watermark embedding, it is difficult to determine how much information can be embedded and later reliably extracted at the watermark detector. To survive the most severe channel condition, the embedded data rate should be low, resulting in less embedded information. However, this "worst case scenario" consideration is not desirable because if the channel condition is good, you sacrifice the possibility of extracting more information to more effectively convince the jury. In our approach, a good resolution binary watermark image is always embedded. However, the watermark detector has the flexibility of extracting watermark images of different resolutions by exploiting the high spatial correlation nature of the watermark image. In other words, the detector has the capability to adapt to the channel conditions. When the channel condition is good, the detector will extract a good resolution watermark image which definitely will convince the jury for a valid ownership claim. On the other hand, when the channel condition is bad, the detector will still be able to extract a coarse resolution watermark image which hopefully will convince the jury too.

This paper is organized as follows. Section 2 summaries a previously proposed scheme which detects the watermark without directly resorting to the original image. This scheme provides a quantitative measure of "invisible" watermarks for the validity of an ownership claim. The multiresolution watermark image extraction scheme is presented in Section 3 to provide some enhanced performance, i.e., it provides a visual measure of "invisible" watermarks for the validity of an ownership claim. Experimental results are shown in Section 4.

2. DETECTING WATERMARKS WITH A QUANTITATIVE MEASURE

In this section, we summarize a previously proposed statistical technique[5,6] to detect watermarks without directly involving the original image for resolving rightful ownership. The system also serves as a basic structure for the multiresolution watermark image extraction scheme to be presented in Section 3.

Fig. 1 shows a general architecture of the watermarking system we proposed in.[5,6] To be more concrete, we focus on the so-called feature-based[10] watermarking schemes[2–4,7] in which an i.i.d. (independent identical distributed) pseudo random sequence $\{S_{1i}\}$ is embedded into a set of features $\{I_i\}$ derived from the original image I. For example, the feature set $\{I_i\}$ could be a subset of all the DCT coefficients of 8×8 blocks in the original image.[4,7] The signature S_{1i} is modulated by $G_i(I_i)$, where $G_i(\cdot)$ could potentially be a function of I_i. The resulting value is then added to I_i. Therefore, in the encoding process,

$$I_i' = I_i + G_i(I_i)S_{1i} \tag{1}$$

The watermarked image I' can be constructed based on the modified feature set $\{I_i'\}$ and other unmodified data.

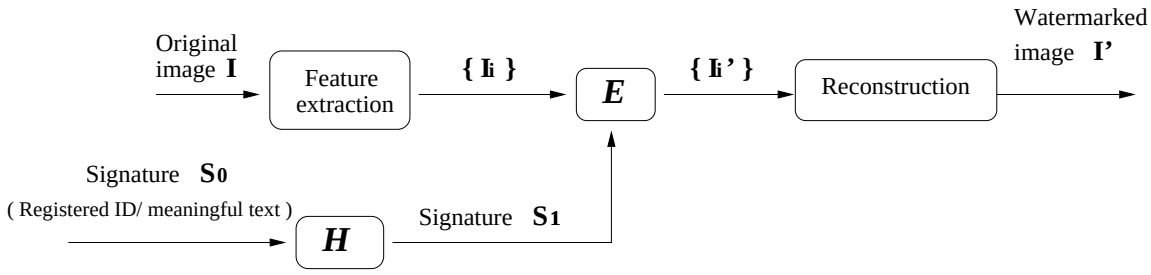
In the watermark detector, the test feature set $\{X_i\}$ are first derived from the test image X, and then correlated with a signature S_2 (also denoted as $\{S_{2i}\}$) which is usually a pseudo random sequence. S_2 should be highly correlated with S_1, but may not be exactly the same. Choice of S_2 can be optimized to improve the detector performance, as discussed in.[5,6] The correlator output q is compared to a threshold T to determine if the test image contains the claimed watermarks. Detection of the watermarks is accomplished via the hypothesis testing:

$$\begin{aligned} H_0 &: \quad X_i = I_i + N_i && \text{not contain the claimed watermark} \\ H_1 &: \quad X_i = I_i + G_i(I_i)S_{1i} + N_i && \text{contain the claimed watermark} \end{aligned} \tag{2}$$

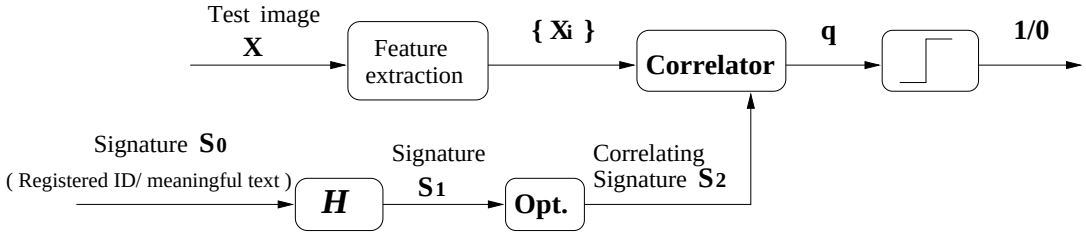
where N_i is noise, possibly resulted from some signal processing such as compression, lowpass filtering, etc..

The correlating detector outputs the test statistic q

$$q = \frac{\sum_{i=1}^{n} Y_i}{V_y\sqrt{n}} = \frac{M_y\sqrt{n}}{V_y} \tag{3}$$



(a) Watermark Encoder



(b) Watermark Detector

Figure 1. Block diagram for our previously proposed watermarking system

where $Y_i = X_i S_{2i}$, n is the size of the feature set $\{X_i\}$, M_y and V_y^2 are the sample mean and the sample variance of Y_i, given respectively by

$$M_y = \frac{\sum_{i=1}^{n} Y_i}{n}; V_y^2 = \frac{\sum_{i=1}^{n}(Y_i - M_y)^2}{n-1}; \tag{4}$$

Assume that the sequence $\{Y_i\}$ is stationary and at least l-dependent *[11] for a finite positive integer l, and that $\{S_{2i}\}$ is *zero mean* and uncorrelated with the original image I. Then under Hypothesis H_0, the test statistic q follows a zero mean student distribution with $n-1$ degrees of freedom.[12] For large n, q is approximately a normal distribution with zero mean and unit variance, i.e., $q \sim N(0,1)$. Note that it is not necessary to require $\{S_{2i}\}$ to be normal distributed in order to have this property hold, as opposed to the proof of a similar property discussed in.[2]

Let $E(\cdot)$ denote the expectation operator. Under Hypothesis H_1 and for large n, it is easy to show that q follows a normal distribution $N(m,1)$, where

$$m = \frac{(E(G_i(I_i)S_{1i}S_{2i}) + E(N_i S_{2i}))\sqrt{n}}{V_y} \simeq \frac{\sum_{i=1}^{n}(G_i(I_i)S_{1i}S_{2i} + N_i S_{2i})}{V_y\sqrt{n}} \tag{5}$$

Note that $E(N_i S_{2i})$ may not be equal to zero, since N_i could have some correlation with S_{2i}. For example, if the feature points are some DCT coefficients, then the noise introduced by zeroing out some of the coefficients (due to, e.g., coarse quantization) is highly correlated with S_{1i}, and thus may potentially have high correlation with S_{2i}.

To determine if a test image contains the claimed watermarks, the output q is compared to a threshold T. If $q > T$, the test image is declared to have been signed with the claimed signature $\{S_{1i}\}$. Otherwise it is not. The threshold T that minimizes the total detection errors (both Type 1 error - accept the existence of a signature under H_0, and Type 2 error - reject the existence of a signature under H_1) is $T = \frac{m}{2}$. In practice, however, nobody other than the true owner knows the value of m. People may only care about the false alarm detection probability, that is

*A sequence $\{Y_i\}$ is l-dependent if $(Y_1, \cdots, Y_k)$ and $(Y_{k+n}, \cdots, Y_{k+n+m})$ are independent for all k and m whenever $n > l$.

, the probability of error detection by accepting the existence of a signature when it does not exist. Refer to Table 1 for the false alarm detection probability as a function of the threshold T.

For a valid ownership claim, a *valid* correlating signature S_2 should satisfy some constraints.[5,6] First, it should be uncorrelated with the original image I. Second, it is necessary that the mean value of S_{2i} is zero. It can be proved[6] that, if $G_i(\cdot)$ is independent of I_i, then the choice of $S_{2i} = G_i S_{1i}$ is the optimal correlating signature which will results in the largest mean value m under H_1. On the other hand, if $G_i(\cdot)$ is a function of I_i, and assume that $G_i(\cdot)$ can be written as a product of two terms, i.e., $G_i(I_i) = U_i(I_i)W_i$ where W_i is independent of I_i, then a good choice of S_{2i} is $S_{1i}W_i$. An example of $G_i(\cdot)$ will be discussed in Section 4 where $G_i(\cdot)$ can be written as a product of two terms.[4,7] One is an *image independent* frequency threshold. The other one, which depends on I_i, accounts for the luminance sensitivity and contrast masking effect of the human visual system. It should be noted that setting S_{2i} to $G_i(I_i)S_{1i}$ is usually a very bad choice. Details of the analysis can be found in.[6]

In Fig. 1, it is shown that the signature S_1 is generated by mapping a *meaningful* text or *registered* owner ID using a potentially registered one-way deterministic function H. This step is to guarantee that the scheme meets the legitimate requirement that S_2 is generated independent of any image (thus uncorrelated with any image except potentially the watermarked image). It further precludes the possibility of doing exhaustive search (though sometimes computationally expensive) for a large detector output value by playing around with the seed of a random number generator. This step is believed to be a very important requirement for a valid ownership claim.[6] In fact, the false alarm detection probability discussed above makes sense only if only *one* sample random sequence S_2 that generated from certain seed is provided (**one trial**). If many trials are allowed, then one can always pick the sequence that results in the largest detector output value. Note that, in general, it is computationally expensive for a counterfeit attacker to analyze the test image to obtain a *correlating signature* S_2 (generated using a random number generator) which results in an output q significantly deviated from zero. However, if the presented detector output q is not too large, for example, $q = 5$ which corresponds to a large enough false alarm detection probability of 10^{-7}, then it is possible to computationally playing around with the seed of a random number generator to search for an output q which is greater than 5. In other words, without the proposed one-way deterministic mapping of a *meaningful* signature or a *registered* owner ID, a counterfeit attacker can simply try different values of the seed of a random number generator and find a seed that results in a large detector output q. For example, on the average, a counterfeit attacker can find a seed after about 10^7 trials (considering that $q \sim N(0,1)$, see Table 1) that results in an output q greater than 5 on *any* image, and therefore can claim that *any* image is his property. We have done experiments to verify this. For example, for the 256×256 "lenna" image, it takes about 117 ms to do one trial of detection using a Sun Ultra 2 workstation. We applied the detection scheme to the 256×256 original *unwatermarked* "lenna" image. The detector output a value of 4.93 after 617096 trials (or about 20 hours). This scenario suggests that, without constraints on the flexibility of choosing the seeds, given enough computing power, almost surely anyone can claim he has the ownership of any image. Many previously proposed watermarking schemes (whether the original image is involved in the detection process or not) are vulnerable to this counterfeit attack. Requiring the watermark to be dependent on the "original" image in the hope that with this constraint the "original" image can not be generated after the fake watermarks as suggested in[13,8] does not necessarily resolve this problem, because an attacker still has the flexibility to play around with the claimed "original" image to computationally search for an "original"-image-dependent fake watermark which has certain correlation with the extracted watermarks.[6]

The scheme described above makes any counterfeit scheme impossible,[5,6] and the watermark detector output value truly quantifies the false alarm detection probability. However, it is not straightforward for the jury to understand the physical meaning of the detector output value. We believe that, in addition to the detector output value which quantifies the false alarm detection probability, if one can extract a *meaningful* watermark image (e.g., a logo image) from the test image, it will greatly facilitate the process of convincing the jury of an ownership claim. In the next section, we present a scheme which can extract binary watermark images of different resolutions.

3. EXTRACTION OF MULTIRESOLUTION BINARY WATERMARK IMAGES

The scheme presented in last section essentially extracts one bit information from the entire test image X, i.e., whether the claimed signature is embedded in the test image or not. One can embed/extract more bits by segmenting the whole image into smaller segments and then embed/extract one bit for each segment. To be more concrete, once the random sequence S_1 is generated and segmented into smaller segments, each of which corresponding to one segment

of the original feature set $\{I_i\}$, we can modulate each segment of S_1 by either +1 or -1, then embed it into the corresponding segment of $\{I_i\}$. Detection of this one bit information is accomplished via the hypothesis testing:

$$\begin{aligned} H_0^{'}: &\quad X_i = I_i + G_i(I_i)S_{1i} + N_i \quad \text{a bit of +1 is embedded} \\ H_1^{'}: &\quad X_i = I_i - G_i(I_i)S_{1i} + N_i \quad \text{a bit of -1 is embedded} \end{aligned} \tag{6}$$

where N_i is noise, and index i corresponds to data in one particular segment. Using the test statistic q as shown in Eq. (3), and assuming that $\{S_{2i}\}$ is zero mean and uncorrelated with the original image I, we have that q follows normal distribution $N(m,1)$ and $N(-m,1)$ for $H_0^{'}$ and $H_1^{'}$ respectively, with m defined in Eq. (5). Therefore, the threshold that minimizes the total detection error is $T=0$. In other words, when q (or equivalently $\sum_{i=1}^{n} X_i S_{2i}$) is greater than 0, a bit +1 is extracted; otherwise, a bit -1 is extracted. To minimize the detection errors, we should choose S_2 to maximize m. As proved in,[6] if $G_i(\cdot)$ is independent of I_i, then the optimal choice of S_{2i} is $G_i S_{1i}$. On the other hand, if $G_i(\cdot)$ is a function of I_i, and assume that $G_i(\cdot)$ can be written as a product of two terms, i.e., $G_i(I_i) = U_i(I_i)W_i$ where W_i is independent of I_i, then a good choice of S_{2i} is $S_{1i}W_i$. It should be noted again that setting S_{2i} to $G_i(I_i)S_{1i}$ is usually a very bad choice.

We are interested in embedding a *meaningful* binary watermark image. While a large detector output value which quantifies the false alarm detection probability should be considered as the fundamental measure for a valid ownership claim, the ability to extract a meaningful watermark image is very useful in convincing the jury in the court for the claim of an ownership. The extracted watermark image serves as a visual measure of the "invisible" watermarks embedded in the test image. It has the additional advantage of providing the opportunity to exploit the human visual system's super ability to recognize a correlated pattern.[9] It is well known that, unlike traditional data, visual data can be lossy, and is more tolerative to detection errors. Human eyes usually can easily filter out some random noise and recognize a correlated pattern, in a way similar to how channel coding detects and corrects transmission errors. Another advantage is that visual data usually has high spatial correlation. This property can be used to enhance the detection performace as to be described in the following.

It can be seen from Eq. (5) that the larger the size n of each segment, the larger the value of m, and hence the smaller the detection error. However, increasing the size of each segment reduces the total number of bits that can be embedded. As a result, the prospective binary watermark image to be embedded has more constraints and less flexibility. Instead of enlarging the segment, we propose to embed one bit to each 8×8 image block. At the detector, we can usually correctly extract one bit from each 8×8 test image block, given that the watermarked image does not suffer from much image processing. However, if the watermarked image does undergo some image processing, the bit will not be extracted reliably and the detection error will increase. In this case, we propose to exploit the spatial correlation of the binary watermark image to improve the detection performance. In particular, we will increase the number of image blocks from which one bit will be extracted. For example, to extract the bit embedded in the current 8×8 image block, we can make use of the surrounding blocks (for example, a 3×3 window of blocks). In other words, we calculate $\sum_{i=1}^{n^{'}} X_i S_{2i}$, where $n^{'}$ is the total number of features within the detection window, and then compare the result to zero to determine the embedded bit of the *current* block. Note that if the test image X does contain the watermark image, then we have $\sum_{i=1}^{n^{'}} X_i S_{2i} = \sum_{i=1}^{n^{'}} (I_i + N_i)S_{2i} + \sum_{i=1}^{n^{'}} b_i G_i(I_i) S_{1i} S_{2i}$, where b_i is the corresponding bit (+1/-1) embedded in a particular 8×8 block. The second term accumulates and is the major factor to determine the embedded bit. A bit in a homogenous region of the binary watermark image usually has the same value as the surrounding bits. Thus by increasing the number of blocks involved in extracting one bit information embedded in the current block, we generally increase $|\sum_{i=1}^{n^{'}} b_i G_i(I_i) S_{1i} S_{2i}|$, thus reduce the detection error. However, for information bits around an edge in the binary watermark image, increasing the detection window size does not necessarily reduce the detection error, because b_i embedded in other blocks may not have the same sign as the b_i embedded in the current block. Therefore the extracted binary watermark image may lose its resolution around edges, though more robust to signal processing in the homogenous region. The trade-off here is robustness vs. resolution of extracted watermark image. The nice feature of our proposed system is that the watermark detector (rather than encoder) is allowed to *adaptively* choose the trade-off between robustness and resolution. When the test image does not suffer from signal processing, a small detection window is chosen and a good resolution watermark image will be extracted. On the other hand, when the test image suffers from severe image processing, robustness is a concern and should be increased. In this case, by increasing the detection window size, a coarse resolution watermark image can be extracted. It should be noted that increasing the detection window size to increase the robustness is different from applying some noise-reduction operations such as media filtering to the extracted watermark image

obtained by using only one block in the bit extraction process. In the latter case, each bit is detected independently first, thus is more vulnerable to channel noise. Once enough bits are in errors, rendering an unrecognizable extracted image, there is no noise-reduction operation that can recover a recognizable pattern, while increasing the detection window can still extract some meaningful pattern (see Fig. 7 for an example).

4. EXPERIMENTAL RESULTS

We start with the visual-model-based watermark encoding scheme described in,[4,7] and apply the proposed detection scheme to extract the binary watermark image. Two perceptually based watermarking schemes have been proposed in.[7] One is based on block-based DCT transform framework. The other is based on multiresolution wavelet framework which generally yields overall better performance. We will present our test results based on the DCT based perceptual watermark encoding scheme.[4] Better performance is expected with the wavelet based perceptual watermark encoding scheme.[7] The test images are 512×512 "baboon" and "lenna", and the watermark image is a 64×64 binary image shown in Fig. 2, with each bit to be embedded into the corresponding 8×8 image block. Note that, in general, it is not necessary that the coefficients carrying a particular watermark bit information belong to a single block. Other configurations are possible.[14]

In the DCT-based perceptual watermark encoding scheme,[4] the image is first divided into nonoverlapped 8×8 blocks. Then each block is DCT transformed. A frequency threshold value is derived based on measurements of specific viewing conditions for each DCT basis function, which results in an *image-independent* 8×8 matrix of threshold values, denoted as $T_f(u,v)$, $u,v = 1, \cdots, 8$. The feature set $\{I_i\}$ may consist of the AC coefficients which are larger than the corresponding $T_f(u,v)$, organized in the zigzag order within each block and from one block to another in the raster scan order. Denote the corresponding sequence of $T_f(u,v)$ to $\{I_i\}$ as $\{B_i\}$. One can use a more accurate perceptual model which also takes care of the luminance sensitivity and contrast masking effect of human eyes to find the just noticeable difference (JND) of each coefficient.[15,4] Luminance sentitivity is estimated as $T_l(u,v,b) = T_f(u,v)(X_{0,0,b}/\bar{X}_{0,0})^a$ where $X_{0,0,b}$ is the DC coefficient for block b, $\bar{X}_{0,0}$ is the DC coeffieient corresponding to the mean luminance of the display, and a is a parameter which controls the degree of luminance sensitivity. Then a contrast masking threshold, referred to as the JND, is derived as $T_c(u,v,b) = Max[T_l(u,v,b), T_l(u,v,b)(|X_{u,v,b}|/T_l(u,v,b))^{w_{u,v}}]$. where $w_{u,v}$ is a number between zero and one and can assume a different value for each DCT basis function. Note that, in general, the JND of a coefficient increases nonlinearly with the corresponding $T_f(u,v)$ and the magnitude of the coefficient. The contrast masking effect basically suggests that the larger the magnitude of the original coefficient, the larger amount of modification we can make to it without incurring visual artifacts. The feature set $\{I_i\}$ now consists of the AC coefficients which are larger than their corresponding JNDs (or equivalently, $T_l(u,v,b)$). Note that watermarks will not be embedded into those small coefficients (smaller than their corresponding $T_l(u,v,b)$) in order to avoid visual artifacts. This also avoids the potential negative effect on the compression performance if JPEG compression is to be applied to the watermarked image subsequenctly. Denote the corresponding sequence of JNDs to $\{I_i\}$ as $\{J_i\}$. Then J_i is used as $G_i(I_i)$ in Eq. (1).

Note that the locations of the feature set $\{X_i\}$ obtained from the test image may not exactly correspond to the locations of the feature set $\{I_i\}$ obtained from the original image, because they are determined by comparing the coefficients of the test/original image to some thresholds. In addition, for the detector, we used the *image-independent* $T_f(u,v)$ as the thresholds to determine the feature set $\{X_i\}$, instead of $T_l(u,v,b)$ which depends on the DC coeffieient of each 8×8 block. For synchronization between the encoder and the detector without reference to the original image, the same seed will be used at the encoder and the detector to generate a random sequence of the original image size, each element of which corresponds to one DCT coefficient. However, only those elements corresponding to the feature points $\{I_i\}$ or $\{X_i\}$ will be used as $\{S_{1i}\}$ for encoding or detection. An alternative is to use all the AC coefficients of the test image as the feature points at the detector. However, this will generally yield slightly worse performance, as shown in.[6]

A meaningful signature or registered owner ID S_0, for example "Sharp Laboratories", is mapped to an i.i.d. sequence S_1 with distribution of $N(0,1)$. S_1 is then modulated by the original watermark image and $\{J_i\}$, and then embedded into the original image according to Eq. (6). In the watermark detector, the signature S_0 is presented as the secrete key, and S_{2i} is chosen as $B_i S_{1i}$, which has been shown to be a near optimal choice in our experiments. Note that B_i is a factor of J_i which does not depend on I_i.

Fig. 3 shows both the original and the watermarked "baboon" images. They appear to be the same. No visual difference is observed. The visual model based watermark encoder is doing a good job. We extracted different resolutions of the binary watermark image from the test image under different channel conditions. Several detection window sizes are used in the extraction process. In Case 1, only the current image block is used. In Case 2, five image blocks are used (including the current block, the ones above, below, to the left, and to the right). In Case 3, a 3×3 window of blocks are used, and in Case 4, a 5×5 windows of blocks are used. When the watermarked "baboon" image does not suffer from any signal processing, the extracted versions with different detection windows are shown in Fig. 4. It is seen that Case 1 and Case 2 provide the best resolution of the binary watermark image. There are some random noise presented in the extracted images. They are, however, easily filtered out by the human eyes. It should be noted that if one increases the segment size for embedding one information bit in the encoding stage (in order to be robust to signal processing), no such detailed information can be embedded and later extracted. As the detection window size increases, lower resolution watermark images are extracted in which the edge parts become more and more jerky (see, e.g., the character "7"). Hence when the channel condition is good, the extracted good resolution watermark image of Case 1 or Case 2 can be presented to the jury to prove the ownership. Note that, since S_2 is generated independent of the test image, if the test image does not contain the claimed watermarks, the extracted binary image will look rather random as shown in Fig. 2. Fig. 5 shows the extracted watermark images from the watermarked 512×512 "lenna" image. They appear to be a little bit noisier than the corresponding results for "baboon" image. This suggests that different images may tolerate different amount of watermarks, and therefore the detection scheme should have some adaptability to account for the differing robustness.

When the watermarked image is subject to signal processing such as JPEG compression, the resolution of the extracted binary watermark image has to be traded for robustness. Fig. 6 shows the extracted watermark images with different detection windows from the watermarked image that suffers from JPEG compression with quality factor of 15% (compression ratio: 12:1). The extracted image in Case 1 is hardly recognizable. It is better to present to the jury the results of Case 2 or Case 3 in which both "PU" and "EE" are still recognizable. When the JPEG compression quality factor is 5% (compression ratio: 28:1), the extracted image is meaningless in Case 1 (See Fig. 7). The extracted watermark image in Case 2 is difficult to recognize too, while the extracted watermark images in Cases 3 and 4 are recognizable (at least for the bigger characters "P" and "U"). Note that it might be helpful to construct the binary watermark image in a way such that the content has some hierarchical structure and that higher level content consists of more redundant bits. It should also be noted that if the system presented in Section 2 is used to detect the watermark in the JPEG compressed watermarked image (with quality factor of 5%), then the detector outputs a value of 21 which virtually corresponds to zero false alarm detection probability (See Table 1). Fig. 8 shows the robustness of the detection scheme presented in Section 2 to JPEG compression. Note that in this case, the ownership claimer will present both the signature S_0 and the original binary watermark image, and the correlating signature $\{S_{2i}\}$ should be $\{B_i S_{1i} b_i\}$ where b_i is the corresponding bit (+1 or -1) of the original binary watermark image. This is a *valid* correlating signature because it is still generated independent of any image. It should be noted that there might be cases in which the extracted watermark images are hardly recognizable despite the detection window size used, while the detector output value of the system presented in Section 2 is still large enough to signify a low false alarm detection probability. This again suggests that the system presented in Section 2 provides, in the view of technical experts, a fundamental measure for a valid ownership claim. However, the proposed watermarking system in this paper makes more sense to ordinary people, thus will greatly facilitate the process of convincing the jury of an ownership claim. Only when the watermarked image has been subject to too much processing, resulting in unrecognizable extracted binary images despite the detection window size used, should it be necessary to call the technical expert to testify the physical meaning of the output value of the watermarking system presented in Section 2.

5. CONCLUSIONS

A watermarking system which is capable of embedding a good resolution binary watermark image and later extracting different versions of that watermark image with varying resolutions is proposed. The nice feature of the proposed system is that the watermark detector is allowed to adaptively, given the channel condition, determine the trade-off between robustness degree and resolution of extracted watermark image. We believe that this system, together with a statistical measure,[6] will greatly facilitate the process of convincing the jury of an ownership claim in the court.

Threshold T	$P_{err}(q > T)$
3	0.0013
5	2.86E-7
6	9.86E-10
8	6.22E-16
10	7.62E-24
12	1.77E-33

Table 1. False alarm detection probability P_{err} for the watermarking system presented in Section 2.

REFERENCES

1. E. Koch, J. Rindfrey, and J. Zhao, "Copyright protection for multimedia data," in *Proc. of the Int. Conf. on Digital Media and Electronic Publishing*, 1994.
2. I. Cox, J. Kilian, T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for images, audio and video," in *Proc. Inter. Conf. Image Proc.*, vol. 3, pp. 243–246, Sept. 1996.
3. M. Swanson, B. Zhu, and A. Tewfik, "Transparent robust image watermarking," in *Proc. Inter. Conf. Image Proc.*, vol. 3, pp. 211–214, Sept. 1996.
4. C. Podilchuk and W. Zeng, "Digital image watermarking using visual models," in *Proc. IS&T/SPIE Electronic Imaging: Human Vision and Electronic Imaging*, vol. 3016, pp. 100–111, Feb. 1997.
5. W. Zeng and B. Liu, "On resolving rightful ownerships of digital images by invisible watermarks," in *Proc. Inter. Conf. Image Proc.*, vol. 1, pp. 552–555, 1997.
6. W. Zeng and B. Liu, "An invisible watermark detection technique without using original images for resolving rightful ownerships of digital images." submitted to *IEEE Trans. Image Processing*, 1997, under revision.
7. C. Podilchuk and W. Zeng, "Image-adaptive watermarking using visual models," *IEEE Journal on Selected Areas in Comm., special issue on Copyright and Privacy Protection* **16**(4), pp. 525–539, May 1998.
8. M. Swanson, B. Zhu, and A. Tewfik, "Multiresolution scene-based video watermarking using perceptual models," *IEEE Journal on Selected Areas in Comm., special issue on Copyright and Privacy Protection* **16**(4), pp. 540–550, May 1998.
9. G. Braudaway, "Protecting publicly available images with an invisible image watermark," in *Proc. Inter. Conf. Image Proc.*, vol. 1, 1997.
10. S. Craver, N. Memon, B. Yeo, and M. Yeung, "Can invisible watermarks resolve rightful ownerships?," in *IS&T/SPIE Electronic Imaging: Storage and Retrieval of Image and Video Databases*, vol. 3022, pp. 310–321, Feb. 1997.
11. P. Billingsley, *Probability and Measure*, John Wiley & Sons, 1991, pp. 375-376.
12. A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, 1991, pp. 269-270.
13. S. Craver, N. Memon, B.-L. Yeo, and M. M. Yeung, "Resolving rightful ownerships with invisible watermarking techniques: limitations, attacks, and implications," *IEEE Journal on Selected Areas in Comm., special issue on Copyright and Privacy Protection* **16**(4), pp. 573–586, May 1998.
14. W. Zeng and S. Lei, "Transform domain perceptual watermarking with scalable visual detection - a proposal for JPEG2000," in *ISO/IEC JTC1/SC29/WG 1 N759*, Geneva, Switzerland, March 1998.
15. A. B. Watson, "DCT quantization matrices visually optimized for individual images," in *Proc. SPIE Conf. on Human Vision, Visual Processing, and Digital Display IV*, vol. 1913, pp. 202–216, 1993.

Figure 2. Left: 64×64 binary watermark image to be embedded. Right: extracted binary image if the test image does not contain the claimed signature.

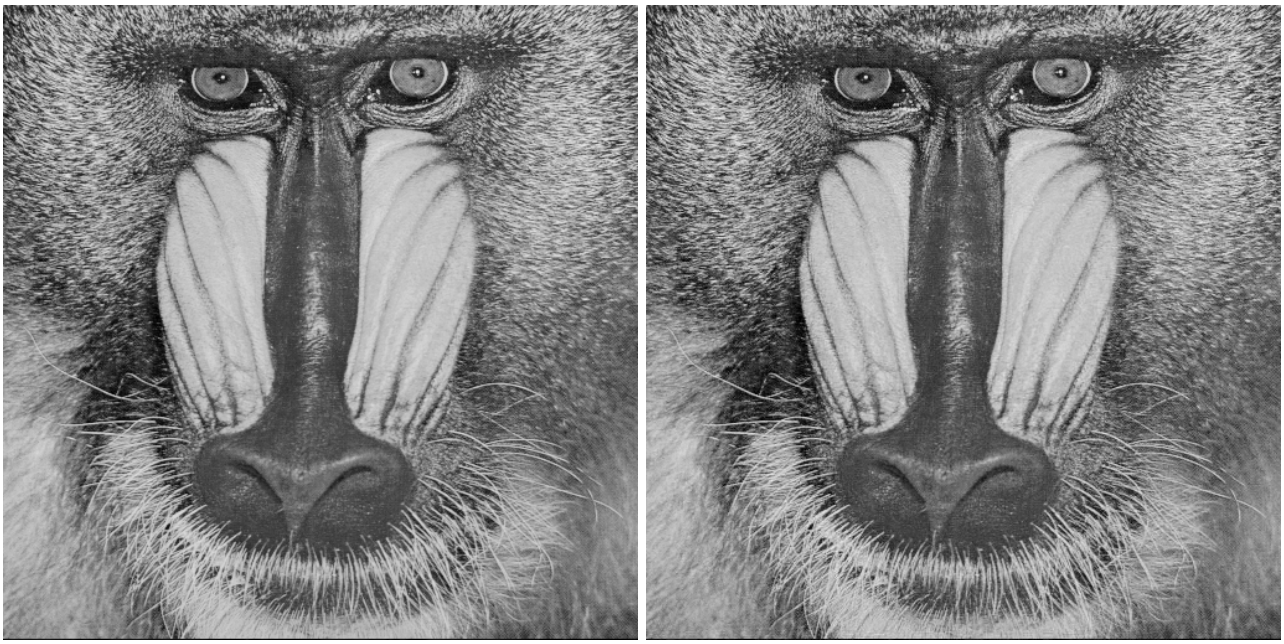


Figure 3. Left: original 512×512 "baboon" image. Right: watermarked "baboon" image.

Figure 4. Extracted watermark images from 512×512 watermarked "Baboon" for (from left to right) Case 1, Case 2, Case 3 and Case 4.

Figure 5. Extracted watermark images from 512×512 watermarked "lenna" for (from left to right) Case 1, Case 2, Case 3 and Case 4.

Figure 6. Extracted watermark images from 512×512 watermarked "Baboon" with JPEG compression with quality factor of 15% for (from left to right) Case 1, Case 2, Case 3 and Case 4.

Figure 7. Extracted watermark images from 512×512 watermarked "Baboon" with JPEG compression with quality factor of 5% for (from left to right) Case 1, Case 2, Case 3 and Case 4.

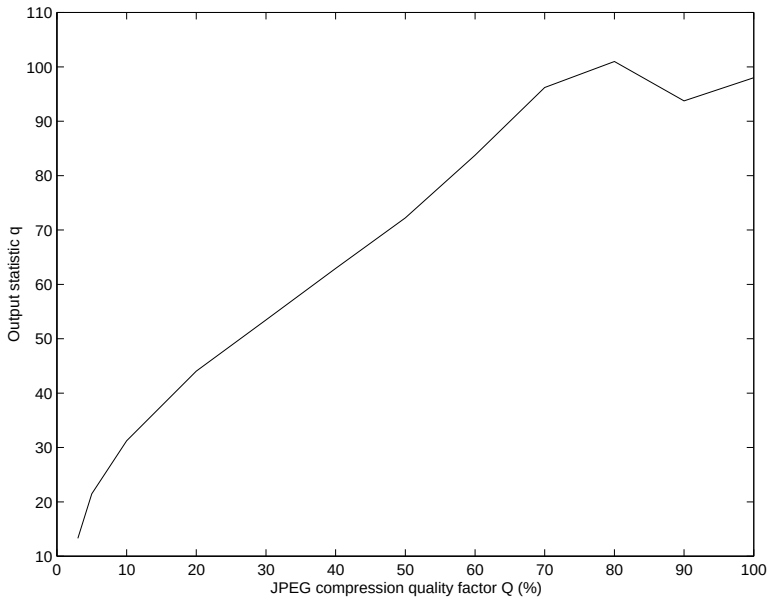


Figure 8. Output statistic q as a function of JPEG quality factor Q for 512×512 "baboon", for the scheme presented in Section 2.