A Transform Domain Hardcopy Watermarking Scheme

Avi Levy, Doron Shaked
HP Laboratories Israel
HPL-2001-309
November 27th , 2001*

E-mail: avi@hpli.hpl.hp.com, dorons@hpli.hpl.com

Digital image watermarking is the process of encoding information into a digital image without deteriorating the visual quality of the image. Digital watermarking techniques are designed to withstand various attacks such as standard image processing operations, compression and geometric transformations. In this report we focus on watermarking of printed images and argue that this application requires specialized methods. Previous work relating to hardcopy watermarking applied standard digital watermarking schemes or suggested techniques that exploit the specific printing (i.e. halftoning) process. This work proposes a novel transform domain hardcopy watermarking technique that is independent of the halftoning process and is rather based on a noise model for the printing and scanning operations. The proposed scheme allows for blind watermark decoding, i.e. the encoder does not know the original image. It is image adaptable, to enhance the visual quality, and is tuneable on the trade-off between information embedding, rate, and visual quality.

* Internal Accession Date Only
Technion city, Haifa 32000, Israel

A Transform Domain Hardcopy Watermarking Scheme

Avi Levy
HP Labs Israel
Technion city, Haifa 32000, Israel
Tel. 972-4-8231237
e-mail: avi@hpli.hpl.hp.com

Doron Shaked
HP Labs Israel
Technion city, Haifa 32000, Israel
Tel. 972-4-8231237
e-mail: dorons@hpli.hpl.hp.com

Abstract

Digital image watermarking is the process of encoding information into a digital image without deteriorating the visual quality of the image. Digital watermarking techniques are designed to withstand various attacks such as standard image processing operations, compression and geometric transformations. In this report we focus on watermarking of printed images and argue that this application requires specialized methods. Previous work relating to hardcopy watermarking applied standard digital watermarking schemes or suggested techniques that exploit the specific printing (i.e. halftoning) process.
This work proposes a novel transform domain hardcopy watermarking technique that is independent of the halftoning process and is rather based on a noise model for the printing and scanning operations. The proposed scheme allows for blind watermark decoding, i.e. the encoder does not know the original image. It is image adaptable, to enhance the visual quality, and is tuneable on the trade-off between information embedding, rate, and visual quality.

1. Introduction

A block diagram of a hardcopy watermarking system is depicted in Figure 1. The encoder embeds a message **m** in an original image **I,** thus producing a digital watermarked image **I**', which is subsequently being printed. The quality of the printed watermarked image should be similar to a printed version of the original image. The decoder receives the hardcopy version, scans it and produces a noisy digital version **I*** of the watermarked image. The decoder then estimates the encoded message from **I***, without resorting to the original image.

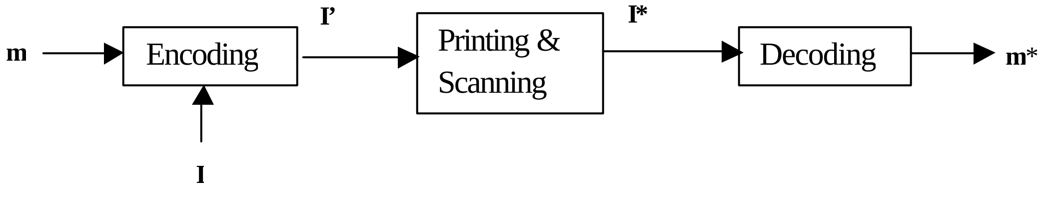


Figure 1: Hardcopy watermark block diagram

Hardcopy watermarking has been treated in certain works under the general context of digital watermarking [1,2]. In these works, the watermark signal is embedded in the middle frequencies range of a digital image using spread spectrum techniques. The watermarked image may be manipulated by image processing operations such as filtering or compression, and in some cases printing and scanning. The watermark signal is decoded from the modified version of the watermarked image using correlation-based methods. These approaches are based on the (some time implicit) assumption that the watermark-signal introduced by the encoder, and the noise signal introduced by the image manipulations, are relatively weak in comparison to the original image signal. We will show, in Section 2, that this assumption, which is true for most standard image processing operations including low quality compression, is not satisfied by the printing and scanning process. Therefore we suggest that hardcopy watermarking embedding rate can be increased if dedicated techniques are used.

The second approach to hardcopy watermarking exploits the halftoning method used by the printing process. This scheme has two variants - one for halftone screens [3,4,5] and the other for error-diffusion algorithms [6,7]. There are few disadvantages that this approach suffers from:
- It requires that the encoder have an access to the internals of the printing process.
It performs information embedding in the spatial domain and hence is sensitivity to geometric deformation such as image miss-alignment and to the presence of stains or scribbles on the printed media.
- In most schemes print resolution should be significantly lower than the scanning resolution in order to ensure errorless decoding. Therefore, the printed watermarked image has low visual quality in comparison to a high resolution printout of the original image.

In this report we propose a novel frequency domain watermarking technique, which is based on statistical features of the printing and scanning noise. Section 2 describes a printing-scanning noise model and provides the motivation for the new hardcopy watermarking scheme. Section 3 presents the details of the encoding scheme and Section 4 describes the decoding procedure. Section 5 includes discussion and description of the experimental results.

2. DFT domain statistical model for image processing noise

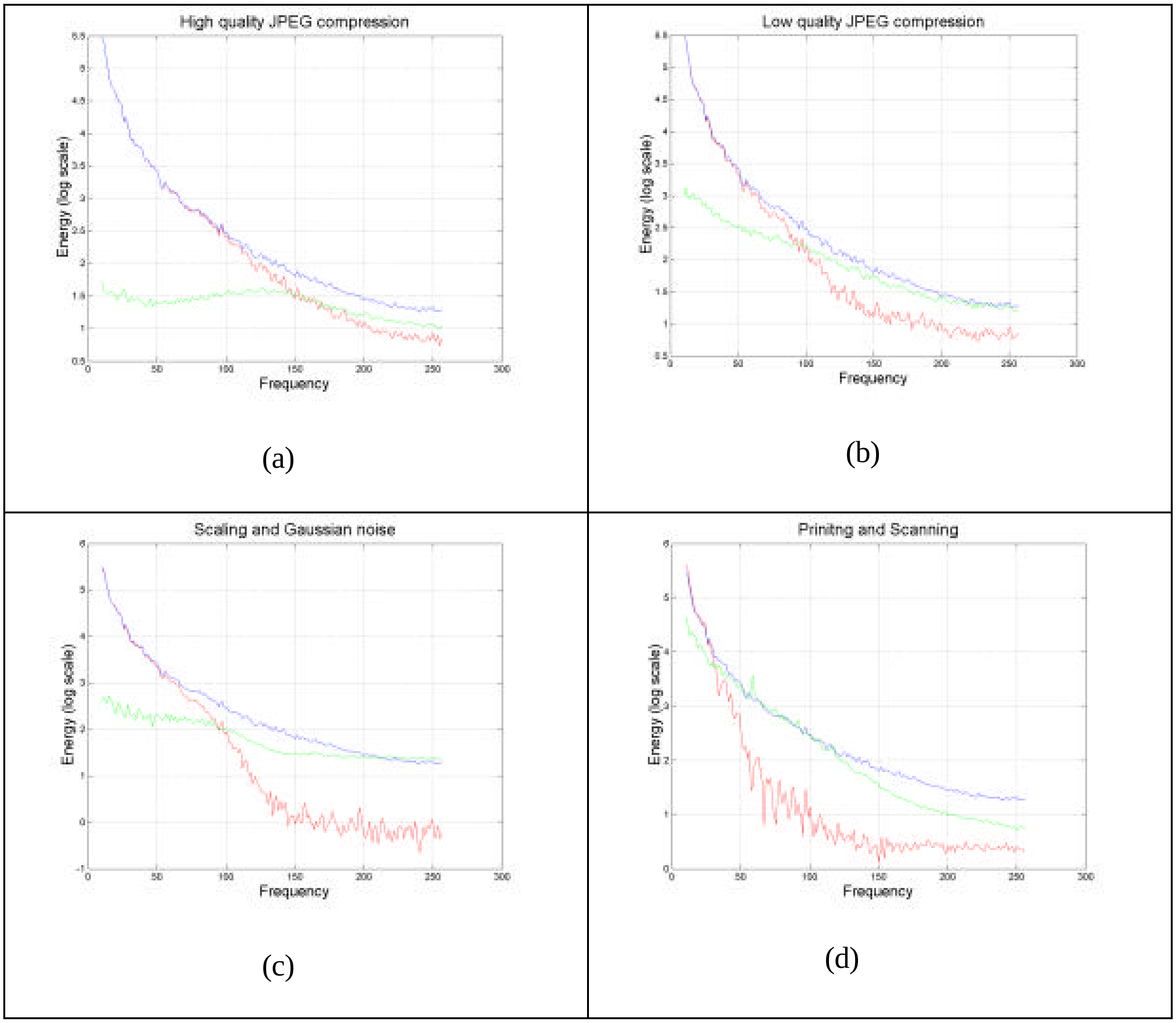
Image representations that are widely used in digital watermarking are the Discrete Cosine Transform (DCT) and the Discrete Wavelet Transform (DWT). In hardcopy watermarking there exists an additional pre-decoding task of aligning the scanned image. Therefore, the Discrete Fourier Transform (DFT) representation that has desirable invariance properties and is less sensitive to small misalignments is theoretically preferable. Due to this fact, which was verified also experimentally, we describe in this report the image watermarking noise model in the DFT domain. However, similar models can be developed in the DCT or DWT domain.

Given an MxM grey-level image, its DFT representation is an MxM array of complex coefficients. For the sake of this noise model, the DFT coefficients are grouped into square single-coefficient-wide rings, centered at the DC coefficient (assumed to be the central entry). These rings are indexed by their distance from the center and we relate to the index of a ring as the ring frequency.

The absolute values of the DFT coefficients in each ring of the original image are modelled as samples from a random variable (r.v.) denoted X_i , where i is the frequency of the ring. We assume that the absolute values of the corresponding DFT coefficients, after the image has been manipulated, are samples of a r.v. Y_i related to X_i by: $Y_i=A_iX_i+Z_i$. Where the A_i's are random coefficients and the Z_i's are zero mean noise r.v. independent (of X_i and A_i). Using standard regression technique we estimated the average value of the A_i's and the energy (variance) of the Z_i's for various image-processing operations. Figures 2 displays typical results of these estimations in four diagrams corresponding to: High quality JPEG, low quality JPEG, downscale + noise, and printing + scanning. In all diagrams the energy of the original image coefficients (X_i) is represented by the blue graph, the energy of the filtered coefficients (A_iX_i) is represented by the red graph and the green graph represents the energy of the noise signal (Z_i).

The middle frequency range, used by standard digital watermarking schemes, can be characterized by the following description: It starts after the low frequencies, which have overwhelming impact on the visually quality of the image and hence can't be modified by the encoder. It extends over those frequencies where the noise energy is small in comparison to the filtered signal energy, so that an embedded watermarked signal is not erased by a stronger noise signal. One can observe that while this range extends beyond the first 100 rings in diagrams (a), (b) and (c) of Figure 2 it reaches only the 30th (approximately) ring in the printing and scanning diagram (d). This observation is the foundation for the claim that embedding the message in the middle frequency coefficients is inefficient for hardcopy watermarking.

On the other hand, in hardcopy watermarking applications it is possible to embed information even in those coefficients that suffer from high levels of noise. This is due to the fact that the digital (pre-printed) version of the watermarked image is not required to have a high visual quality. It is only the printed version of the watermarked image that should be visually similar to the printed version of the original image. Therefore, a very strong watermark signal can be introduced into the DFT coefficients that are subject to a strong printing-scanning noise. Such a signal, though visible in the digital watermarked image, is invisible in the printed version, since the printing process has strong attenuation effect on these frequencies (This phenomena is expressed by the very small average values of the A_i's, corresponding to these frequencies). Still, the attenuated watermark signal can be picked up after the hardcopy image is scanned and hence the original message can be recovered. The next two sections describe in details an encoding and decoding schemes based on this principle.



(a)

(b)

(c)

(d)

Figure 2: Noise statistics for various image processing operations

(a) High quality JPEG compression (quality=50)
(b) Low quality JPEG compression (quality=10)
(c) Down Scaling (by half) and additive Gaussian noise with zero mean and std equal 5 (gray levels).
(d) 600 dpi printing (approximately one dot per image pixel) and scanning in 600 dpi resolution.

3. Encoding scheme

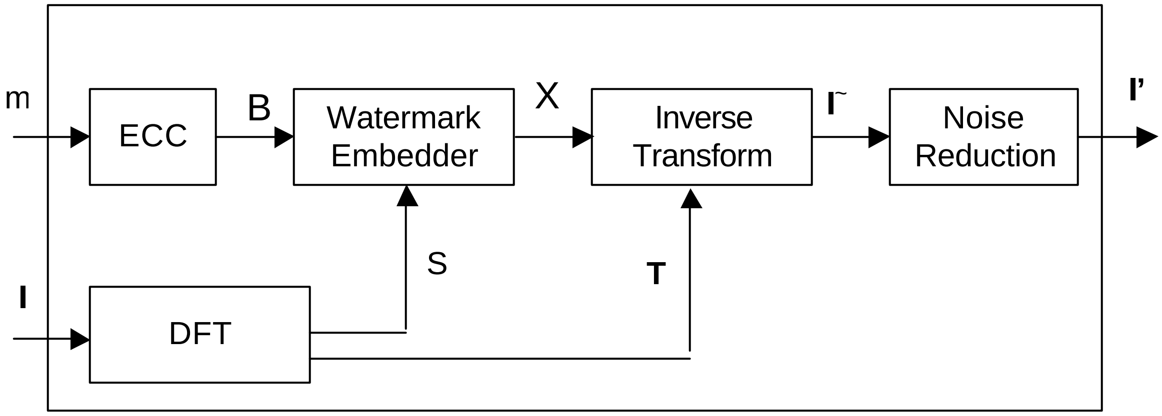


Figure 3: Encoding block diagram

The encoding scheme is based on changing the absolute values of selected pairs of DFT coefficients. Given a coefficient pair, a zero (one) bit is encoded by increasing the absolute value of the first (second) coefficient and decreasing the value of the second (first) one. We call a coefficient whose magnitude is increased a “u” coefficient and a coefficient whose value is decreased an “l” coefficient. As explained in Section 2, the absolute values of DFT coefficients in the middle frequency range can be significantly changed without deteriorating the visual quality of the printed image. It turns out that in spite of the significant attenuation effect and the strong additive noise introduced by the printing and scanning process, with very high probability, the relative order of the coefficient magnitudes in each pair is preserved. Using an Error Correction Code (ECC) and utilizing a ML decoding procedure it is possible to decode with negligible error rates. In order to avoid visible artefacts in the printed watermarked image the locations of DFT coefficients whose magnitudes are changed are pseudo-randomly distributed in the DFT domain. The optimal number of coefficient pairs and the degree to which their value can be changed is determined experimentally using a wide range of images. To achieve high visual quality the encoding stage is appended by a cleaning procedure that decreases the watermark signal strength in smooth areas of the image. The rest of this Section presents the details of the encoding scheme.

The encoder (see Figure 3) receives an original image **I** and a message **m** and outputs the watermarked image **I**’. The encoding procedure comprises the following stages:

1. Compute the DFT of the image **I**, denoted **T**, and rotate the coefficients in such a way that the DC coefficient is in the center. Choose N pairs of transform coefficients in the middle frequency range, where each pair is located in fixed positions relative to a small square. These non-overlapping squares are scattered in a pseudo random pattern in half of the middle frequency range - see Figure 4. The selected coefficients are

ordered in a sequence denoted $S=(s_1,\ldots,s_{2N})$ in such a way that the pairs are indexed s_{2i-1} and s_{2i} for $i=1,\ldots N$.

2. Encode the message **m** as a sequence of N bits denoted $B=(b_1,\ldots,b_N)$, using a linear ECC that maps each block of k message bits into an appropriate code word of length n. For the sake of simplicity we assume that N/n is an integer.
3. Encode the binary sequence B in the coefficient sequence S using the hardcopy watermark procedure described in Section 3.1 below, and obtain the watermarked coefficients sequence $X=(x_1,\ldots,x_{2N})$. The procedure encodes binary sequences of length n in a coefficient sequences of length 2n until the sequence $S=(s_1,\ldots,s_{2N})$ is exhausted.
4. Replace the coefficients corresponding to the sequence S in the DFT representation **T** with the watermarked coefficients of the sequence X. For each watermarked coefficients change the symmetric coefficient (in relation to the DC entry) in such a way that it becomes the complex conjugate of the watermarked entry. This ensures that the watermarked image pixels have real values.
5. Invert the resulting watermarked DFT representation, and obtain the watermarked image $\tilde{\mathbf{I}}$.
6. Reduce visual artefacts caused by the watermark signal by decreasing its amplitude in smooth regions of the image (the noise reduction procedure is described in Section 3.2 below).

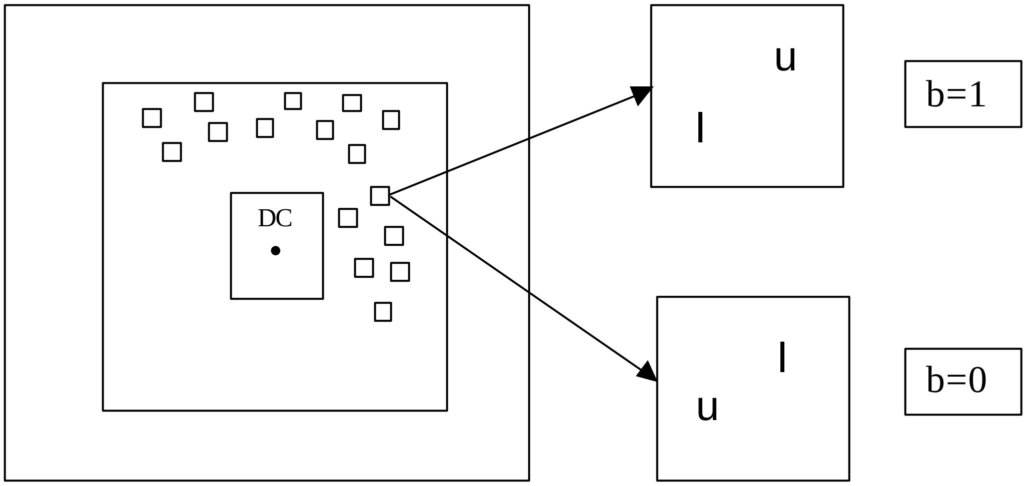


Figure 4: Encoding scheme: Square areas are scattered randomly in half of the middle frequency range of the DFT domain. A pair of coefficients are selected in each square and their magnitudes are increased (denoted u) or decrease (denoted l) according to the encoded bit - b.

3.1 Hardcopy watermarking procedure

The hardcopy watermark encoding procedure receives a sequence **b**=($b_1,\dots,b_n$) of n bits and encodes them in a sequence **s**=($s_1,\dots,s_{2n}$) of 2n transform coefficients. The procedure uses a parameter E that determines the total magnitude change of the coefficients whose absolute value is increased. The bits of the sequence **b** are embedded in the coefficient pairs of the sequence **s**, using the following formula:

$$\begin{cases} \mathbf{x}_{2i-1} = \left(\left| \mathbf{s}_{2i-1} \right| + \frac{\mathbf{E}}{\sqrt{\mathbf{N}}} \right) * \angle(\mathbf{s}_{2i-1}), & \mathbf{x}_{2i} = \mathbf{e} * \angle(\mathbf{s}_{2i}) \qquad \textbf{for } \mathbf{b}_i = \mathbf{0} \\ \mathbf{x}_{2i-1} = \mathbf{e} * \angle(\mathbf{s}_{2i-1}), & \mathbf{x}_{2i} = \left(\left| \mathbf{s}_{2i} \right| + \frac{\mathbf{E}}{\sqrt{\mathbf{N}}} \right) * \angle(\mathbf{s}_{2i}) \qquad \textbf{for } \mathbf{b}_i = \mathbf{1} \end{cases},$$

In the above formula $\angle(\cdot)$ is the phase of a complex number, ε is a negligible value and i=1,..,n.

3.2 Noise reduction procedure

The noise reduction procedure reduces the amplitude of the watermarked signal in smooth regions of the image thus improving visual quality. The procedure performs the following steps:

- Smooth the original image.
- For each pixel in the smoothed image calculate the standard deviation (std) of pixel values in a small neighbourhood, thus producing an “std image”.
- Smooth the std image.
- Normalize the entries of the smoothed std image, by linear scaling such that their values lie in the interval [0,1]. The resulting matrix is denoted A.
- Compute a visually enhanced watermarked image **I**’, by attenuating the watermark signal: $\mathbf{I}' = \mathbf{I} + \mathbf{A} \otimes (\mathbf{I}^{\sim} - \mathbf{I})$, where $\otimes$ denotes a term by term array multiplication.

4. Decoding scheme

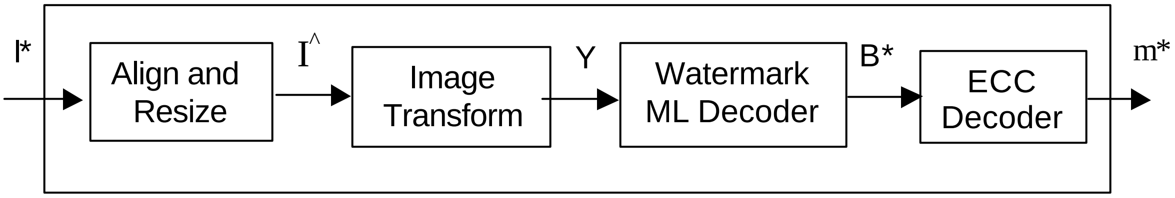


Figure 5: Decoding block diagram

Decoding is based on the fact that with high probability the magnitude order of the encoded DFT coefficients in each pair is preserved after the printing and scanning operations. If this order would always be preserved then decoding would amount to observing, for each pair, which coefficient has a bigger absolute value and determining the sign of the corresponding encoded bit accordingly. Since there are cases in which the order of the coefficient magnitudes is reversed, the decoder uses an ECC and a statistical model for estimating the encoded bit from the observed coefficient values. The details of the decoding procedure are listed below, where it is assumed that the decoder knows the dimensions of the original image and the locations of the encoded coefficient pairs in the DFT domain. Note however, the decoder has no information about the visual content of the original image.

The decoder (see Figure 5) receives a scanned version of the printed watermarked image denoted **I***. It firstly aligns and resizes it, using standard techniques, to ensure that its dimensions and orientation match those of the original image **I**. It then estimates the encoded message using the following procedure:

1. Compute the DFT transform of the aligned image $I^{\wedge}$ and choose the same N pairs of transform coefficients, as in stage 1 of Section 3.1. The selected coefficients are denoted $Y=(y_1,\ldots,y_{2N})$.
2. Using the estimation procedure described in Section 4.1 below, estimate the conditional probability distribution function (pdf) of the DFT coefficients' absolute value, given that they originate from a "u" type coefficients or an "l" type.
3. Extract the binary sequence B* from Y using the ML decoding procedure described in Section 4.2 below. The procedure extracts a binary sequence of length n from a coefficient sequence of length 2n. It is performed repeatedly for each block of 2n coefficients until the sequence Y is exhausted.
4. Decode the binary sequence B*, using the ECC decoding scheme, and obtain the estimated decoded message **m***.

4.1 Estimation of statistics parameters.

The statistics parameters estimation procedure is based on two experimental observations:

- The magnitude of a coefficient originated from a "u" coefficient is, with high probability, greater than the magnitude of its paired coefficient (originated from an "l" coefficient). This probability, denoted p_c, was estimated based on many measurements, and was set to 0.95 for low rate watermarking and 0.90 for high rate watermarking .
- The pdf of the absolute value of those coefficients having the high magnitude in their pair, can be approximated by a Gamma pdf - $G_{a,b}(t) \cong t^a e^{-bt}$ with parameters $a = a_x, b = b_x$. Similarly, the pdf of the absolute value of those coefficients having the low magnitude in their pair, can be approximated by a Gamma pdf with parameters $a = a_n, b = b_n$.

The parameters, a_x, b_x and a_n, b_n are estimated, based on the observed sequence $Y=(y_1,\ldots,y_{2N})$, using standard estimation procedure. Based on these estimations, the absolute value conditional pdf's - P_u and P_l (corresponding to the "u" and "l" type original coefficients) - are calculated by:

$$P_u(t) = p_c G_{a_x,b_x}(t) + (1-p_c) G_{a_n,b_n}(t)$$

$$P_l(t) = (1-p_c) G_{a_x,b_x}(t) + p_c G_{a_n,b_n}(t)$$

4.2 Maximum likelihood decoding procedure

The ML decoding procedure receives the coefficient sequence $\mathbf{y}=(y_1,\ldots,y_n)$ and outputs the decoded bit sequence $\mathbf{b^*}=(b^*_1,\ldots,b^*_n)$. It uses the estimated probability distribution functions P_u and P_l to compute the conditional probability of observing a pair of coefficients y_{2i-1},y_{2i} given that the corresponding encoded bit was 0 (respectively 1). These probabilities are given by:

$$P\{y_{2i-1}, y_{2i} \mid b_i\} = \begin{cases} P_u(|y_{2i-1}|)\cdot P_l(|y_{2i}|) & \text{for } b_i = 0 \\ P_l(|y_{2i-1}|)\cdot P_u(|y_{2i}|) & \text{for } b_i = 1 \end{cases}$$

The ML procedure comprises the following steps:

For all the ECC words, $\mathbf{b}=(b_1,\ldots,b_n)$, compute the following score:

$$Sc(\mathbf{b}) = P\{\mathbf{y} \mid \mathbf{b}\} = \prod_{i=1}^{n} P\{y_{2i-1}, y_{2i} \mid b_i\}$$

1. Having computed the set of scores for all the ECC words, choose the code word **b** that maximizes the score function and output it as the estimated sequence **b***.

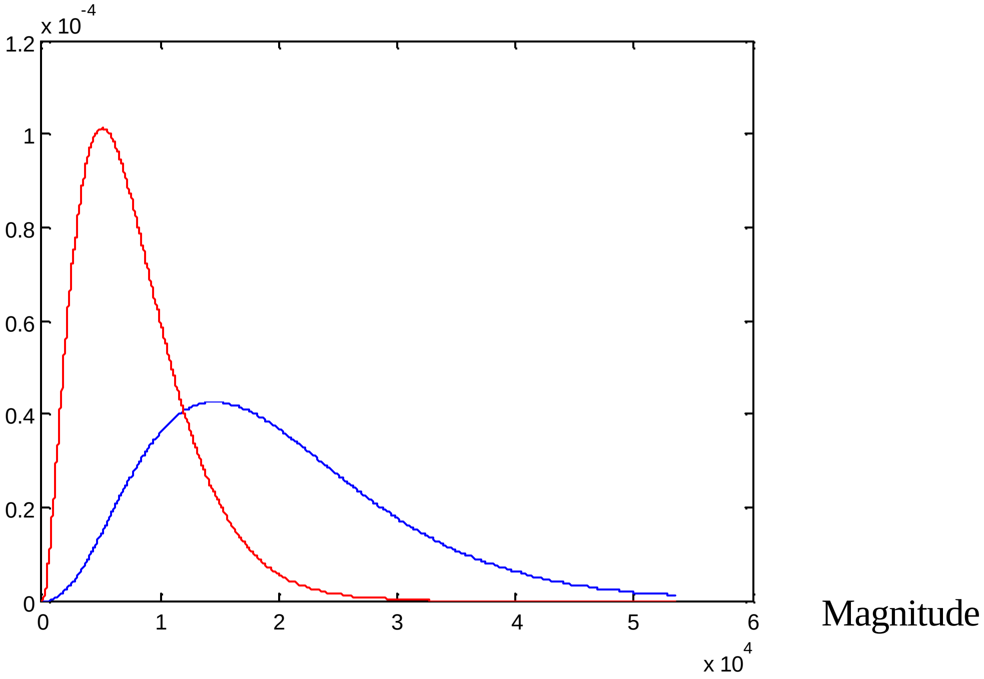


Figure 6 – Estimated energy pdf's of noisy "l" coefficients – red graph and noisy "u" coefficients – blue graph.

5. Discussion and experimental results

The watermarking scheme described in Sections 4 and 5 above was applied to a variety of images and was tested over a range of information embedding rates. The original images have sizes from 512x512 to 600x600 pixels. The watermarked images were printed with an HP DeskJet 970 cxi printer on plain paper in 600 dpi resolution and were consequently scanned with an HP ScanJet 5p with the same resolution. The size of the printouts was approximately one square inch for the 600x600 pixel images and proportionally less for the smaller images.

The information rate (the number of bits embedded in an image) the image quality, and the error rate are controlled as a trade-off by the number (N) of watermark coefficients pairs, the total energy (E) of the watermark signal and by the ECC expansion ratio (n/k). For reliable communication the error rate should be kept negligible and hence:

- High N value reduces visual quality and increases information rate.
- High n/k ratio reduces information rate but allows lower total energy value and hence improves the visual quality.
- In addition, the noise reduction procedure can be switched off, resulting in significantly higher information rate but poor image quality.

Figures 7, (b) and (c) depict the trade-off between the information rate and the visual quality of the printed watermarked image. Typical values of these parameters (for a 512x512 pixels image) range from N=270, E=8, n/k=15:5, noise reduction “on”, for an invisible hardcopy watermark application as in Figure 7 (b) to N=800, E=20, n/k=15:11, noise reduction “off” for high rate but obvious watermark signal as in Figure 7 (c).

In the invisible watermarking scenario, 10 messages of 100 bits, where embedded in each one of 5 different 600x600 images and were decoded without error. In the same experiments, for the high rate and low visual quality watermarking, the length of messages that were decoded without error was 800 bits.

In order to verify the robustness of the hardcopy watermarking scheme we have made few experiments where the printed watermarked images were significantly miss-aligned or where the printed media was defaced by scribbling. Even in the presence of these disturbances decoding was performed without error.

6. Summary

We have presented a novel DFT domain watermarking scheme that embeds a binary message in a printed image without interfering with the dithering algorithm. This scheme utilizes the noise features of the printing-scanning operations and hence is preferable over non-specialized digital watermarking scheme. For low embedding rate it produces watermarked prints that are perceptually indistinguishable from the original prints and it can trade visual quality for embedding rate. The scheme has been tested extensively and proved to be stable under a variety of input images and resilient to various types of image quality degradations.

(a) (b) (c)

Figure 7: Original hardcopy version of a 512x512 image (upper row) and corresponding scans (lower row).

(a) Original image (no watermark).
(b) Low rate hardcopy watermark – 91 bits - (~125 bits per square inch).
(c) High rate hardcopy watermark – 588 bits - (~807 bits per square inch).

Remark: The actual size of the printed images is approximately (5/6 inch)^2. However, in the figure they are enlarged, so that the differences can be clearly observed.

Bibliography

[1] I.J. Cox, J. Kiliant, T. Leighton, T. Shamoon - *Secured Spread Spectrum Watermarking for Multimedia*. NEC Research Institute, Technical report 95-10.
[2] A.M. Bruckstein, T.J. Richardson – *A Holographic Transform Domain Image Watermarking Method*. Circuits Systems Signal Processing, Vol. 17, No. 3, 1998, pp. 361-389.
[3] Zachi Baharav, and Doron Shaked - *Watermarking of Dither Halftoned Images*. SPIE Electronic Imaging, Vol. 3657, pp. 307-313, January 1999.
[4] J. Yen, D. Tretter and R. Kimmel - *Detecting Halftone Modulations Embedded in an Image*, HP Docket #10007849, Patent Application, April 2001.
[5] Hagit Z. Hel-Or – *Copyright Labelling of Printed Images*. ICIP 2000, September 2000, Vancouver Canada.

[6] D. Shaked, Z. Baharav – *Watermark rate, and distortion of error diffused halftones*. HP Laboratories Technical Report, HPL-98-89, April 1998.
[7] N. Demara Venkata – *Information Embedding in Error Diffused Halftones via Dot Shape Modulation*, HP Docket #10011356, Patent Application, November 2001.