IMAGE WATERMARKING USING DCT DOMAIN CONSTRAINTS

Adrian G. Borş *Ioannis Pitas*

Department of Informatics
University of Thessaloniki
Thessaloniki 540 06, Greece
E-mail: {adrian, pitas}@zeus.csd.auth.gr

ABSTRACT

Watermarking algorithms are used for image copyright protection. The algorithms proposed in this study select certain blocks in the image based on a Gaussian network classifier. The pixel values of the selected blocks are modified such that their Discrete Cosine Transform (DCT) coefficients fulfill a constraint imposed by the watermark code. Two different constraints are considered. The first approach consists of embedding a linear constraint among selected DCT coefficients and the second one defines circular detection regions in the DCT domain. A rule for generating the DCT parameters of distinct watermarks is provided. The watermarks embedded by the proposed algorithms are resistant to JPEG compression.

1. INTRODUCTION

Lately, multimedia and computer networking have known rapid development and expansion. This created an increasing need for systems that protect the copyright ownership for digital images [1]. By using watermarking algorithms, the unauthorized distribution of digital image copies can be tracked.

A watermarking algorithm embeds an invisible identification code that is permanently contained into the image. The watermark (digital signature) which is characteristic to a certain owner must be resistant to various image processing algorithms, especially to the standard compression algorithm, JPEG [2]. Various approaches were lately proposed for image watermarking. These algorithms embed the signal modifications either in the image intensity domain [3] or in the frequency domain [4, 5]. Using the 8×8 block DCT coefficients for embedding the watermark in image was proposed in [4]. In [5], the DCT is applied on the entire image and the watermarking method is similar to the spread spectrum techniques.

The proposed watermarking algorithm consists of two stages. In the first stage, 8×8 pixel blocks located at certain distances one from each other are selected by a Gaussian network classifier. In the second stage, constraints characterized by the watermark parameters are embedded in the DCT coefficients of the selected blocks. Two different approaches for embedding the DCT coefficient constraints are proposed. In the first approach, a linear constraint is embedded in the DCT coefficients based on the least squares algorithm. The second approach defines circular DCT detection regions. We provide a watermark similarity analysis by considering DCT detection regions for various watermarks. The same watermark code can be embedded in graylevel or color images and in image sequences.

2. EMBEDDING THE WATERMARK

The image is partitioned in blocks of 8×8 pixel size, similar to the block size considered by the JPEG compression algorithm. Blocks situated at certain distances from each other are selected based on a Gaussian network :

$$X = \sum_{i=1}^{L} \exp\left[- \sum_{j=1}^{N} \frac{(u_j - w_{i,j})^2}{r_{i,j}^2} \right] > \alpha \qquad (1)$$

where the inputs $\{u_j\}$, for $j = 1, \ldots, N$ are the distances between each two consecutive block locations, L is the number of Gaussian functions and $\{w_{i,j}\}$, $\{r_{i,j}\}$ are the network parameters provided by the watermark code. α is the activation level in order to validate the selection of a block location.

The Gaussian functions are activated in a certain order, only one at the time. The network coefficients must fulfill the following conditions :

$$w_{i,j} = w_{(i+1) \bmod L, j-1} \qquad (2)$$

$$r_{i,j} = r_{(i+1) \bmod L, j-1} \qquad (3)$$

where $i = 1, \ldots, L$ are the Gaussian functions, $j = 2, \ldots, N$ are the inputs, and mod L denotes the operation modulo L.

In the embedding stage, the geometrical distances between each two block locations are calculated based on the inverse of the expression (1). Let us assume as known the first $N-1$ distances between each two consecutive locations, $u_j = d_j$ for $j = 1, \ldots, N-1$. After applying the function ln in (1), we obtain the following range of distances, measured in the number of blocks :

$$w_{i,N} - T_{i,N} < u_N < w_{i,N} + T_{i,N}, \qquad (4)$$

$$T_{i,N} = r_{i,N} \sqrt{\ln\frac{1}{\alpha} - \sum_{j=1}^{N-1} \left(\frac{d_j - w_{i,j}}{r_{i,j}} \right)^2}. \qquad (5)$$

The distances between each two locations are recursively calculated, until the entire image is scanned.

After selecting the block locations according to (5), we compute the forward DCT transform for the respective blocks. The most suitable DCT coefficients to be used for embedding a given constraint are those corresponding to the middle frequency range [4]. In the first approach we consider a linear constraint among the DCT coefficients :

$$Y = \mathcal{F}\mathbf{Q} \tag{6}$$

where $\mathcal{F}$ is the modified DCT coefficient vector, $\mathbf{Q}$ is the weighting vector provided by the watermark code. For embedding the constraint, the DCT frequency coefficients are modified to fit the equation (6), based on the least squares algorithm.

The second proposed algorithm defines circular detection regions around certain DCT coefficients. Embedding the circular detection regions is similar to the vector quantization techniques :

$$\|\mathcal{F} - \mathbf{Q}_k\|^2 = \min_{i=1}^{K} \|\mathcal{F} - \mathbf{Q}_i\|^2 \text{ then } \mathcal{F} = \mathbf{Q}_k \tag{7}$$

where $\mathcal{F}$ is the DCT coefficient vector. After embedding the DCT coefficient constraint (6) or (7), the respective block is reconstructed based on the inverse DCT transform.

Only one block is chosen from the entire block range that fulfill the position condition given by (5). The block location is chosen that its average graylevel value is minimally distorted after embedding the DCT coefficient constraint :

$$k = \arg \min_{j=b+w_{j,N}-T_{j,N}}^{b+w_{j,N}+T_{j,N}} \sum_{x=0}^{7} \sum_{y=0}^{7} |g_j(x,y) - f_j(x,y)| \tag{8}$$

where $g_j(x,y)$ are the pixels of the jth block before embedding the watermark, f_j is the block after the image was reconstructed based on the modified DCT coefficients and b is the reference block position. The watermark can be identified in a part of the image depending by the watermark code and by the scanning order. The number of selected blocks and the size of the signed regions [6] are depending by the image characteristics, according to (8).

3. THE WATERMARK DETECTION

This stage implements a decision criterion in order to detect whether a certain watermark was or was not embedded in a given image. The watermark denoted as $\mathbf{S}$ can be considered as composed of two distinct parts $\mathbf{S} = [\mathbf{L}, \mathbf{C}]$. The first part $\mathbf{L}$, consists of the parameters used for finding the block locations, and the second part $\mathbf{C}$ is associated with the DCT coefficient constraint. Let us denote by x the random variable associated with block site position (1), by y the random variable associated with the DCT constraint given by either (6) or (7), and by $P(x,y|g,\mathbf{S}_k)$ the probability of jointly estimating these random variables when providing the watermark code $\mathbf{S}_k$ and the watermarked image g. The probability $P(x,y|g,\mathbf{S}_k)$ can be decomposed, based on the Bayes rule, in :

$$P(x,y|g,\mathbf{S}_k) \propto P(x|y,g,\mathbf{L}_k)P(y|g,\mathbf{C}_k) \tag{9}$$

where $P(y|g,\mathbf{L}_k)$ is the probability of detecting the DCT coefficient constraint (6) or (7), and $P(x|y,g,\mathbf{C}_k)$ is the probability of detecting the location constraint (1), provided that the respective blocks already fulfill the first constraint.

When evaluating the probability of the DCT coefficient constraint $P(y|g,\mathbf{C}_k)$, the algorithm should take into account the likelihood that the signal is distorted. The main JPEG distortions are caused by the DCT coefficient quantization and rounding [2]. These distortions produce changes in the DCT coefficients which may affect the constraints embedded in the image. In order to cope with the DCT coefficient distortions we define certain detection regions depending on the assumed distortion level. In the case when embedding (6), the detection region $\mathcal{D}_y^k$ is :

$$\mathcal{D}_{y,L}^k = \{y \in \mathbb{R} \mid |y - \mathcal{F}\mathbf{Q}| < D_L\} \tag{10}$$

and when embedding (7) :

$$\mathcal{D}_{y,C}^k = \{y \in \mathbb{R} \mid \|\mathcal{F} - \mathbf{Q}_k\|^2 < D_C^2\} \tag{11}$$

where the parameters D_L and D_C depend on the assumed level of watermarked image corruption. The blocks that are not chosen according to (1) and (8) should have their DCT coefficients modified such that they lie outside the decision regions.

The statistical decision for a certain watermark $\mathbf{S}_k$ is based on signal detection theory. The detection ratio is evaluated as the probability to have the constraints imposed by the watermark parameters embedded in the given image :

$$\frac{\int_{\mathcal{D}_x^k}\int_{\mathcal{D}_y^k} P(x,y|g,\mathbf{S}_k)dxdy}{\int_{\mathcal{D}_x^k}\int_{\mathcal{D}_y^k} q(x,y|g,\mathbf{S}_k)dxdy} \approx \frac{\sum_{i=1}^{p} s_i}{MR} > c \tag{12}$$

where $q(x,y|\mathbf{S}_k)$ is a normalizing probability, s_i is the region size, measured in pixels, where the watermark was successfully identified and $M \times R$ is the image size. Conventionally, the watermark is recognized when the given constraints are successfully identified in more than half of the image, i.e., for $c = 0.5$. The procedure of detecting the watermark in a certain image region s_i, as defined by (1) and (8) makes the watermark resistant to image region cropping. The original image is not necessary in the detection stage.

4. WATERMARK SIMILARITY ANALYSIS

Let us consider two watermarks with identical parameters, except the DCT constraint parameter $\mathbf{Q}$. We assume circular detection regions as provided by (11). Let us consider the image signed with a certain watermark $\mathbf{S}_k$ and that we try to detect another watermark $\mathbf{S}_j$ in this image. After compressing and uncompressing the watermarked image, we assume that the DCT coefficients are spread according to a uniform probability distribution function. The support region of this probability is defined in a circular area of a certain size having the center provided by the DCT parameter vector. Let us assume that in the detection stage we consider a radius D for the DCT detection region identical with that one considered in the embedding stage. However,

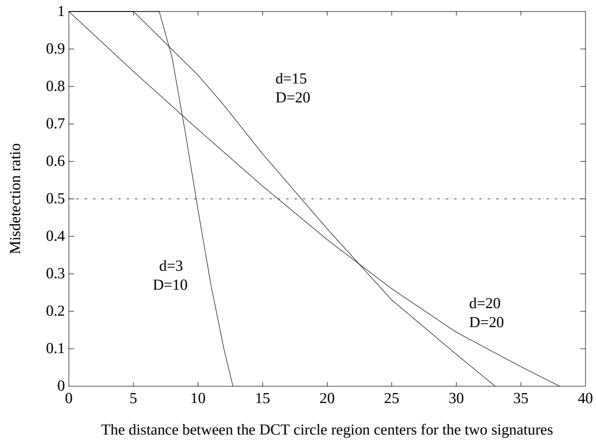


Figure 1: Theoretical watermark misdetection.

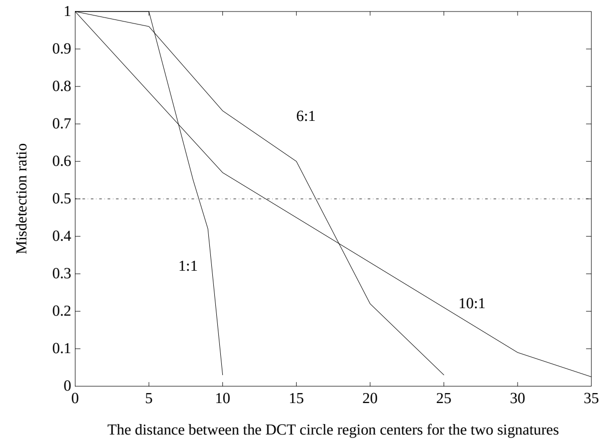


Figure 2: Experimental watermark misdetection.

the compression ratio really used can be smaller than that one assumed in the embedding stage. In this case, the radius d of the DCT coefficient spread is smaller than the detection region size, $d < D$. Under these assumptions, we estimate the probabilities from (12) by computing the overlapping areas of the detection regions in the DCT domain, corresponding to the two watermarks, normalized to the entire DCT detection region. We assume that each two circular DCT detection region centers assigned to the two watermarks are situated at a distance l from each other, measured in the DCT domain. The misdetection probability (12), evaluated with respect to the distance between each two DCT center regions is given by :

$$\frac{\int_{\mathcal{D}_y^j} p(y|g, \mathbf{C}_j)}{\int_{\mathcal{D}_y^k} p(y|g, \mathbf{C}_k)} = \frac{d^2 \arccos \frac{d^2+l^2-D^2}{2dl} + D^2 \arccos \frac{D^2+l^2-d^2}{2Dl}}{\pi d^2} - $$

$$-\frac{1}{2} \frac{\sqrt{2l^2d^2+2d^2D^2+2l^2D^2-d^4-l^4-D^4}}{\pi d^2} \qquad (13)$$

$p(y|g, \mathbf{C}_j)$ is the probability of having the DCT coefficient constraint when the watermark $\mathbf{S}_j$ is embedded in the image and D_y^j, D_y^k are the detection regions of the two watermarks. The variation of the misdetection probability for various distances l, calculated between the centers of DCT circular detection regions of two different watermark codes is represented in Figure 1. In order to test these results we cast a watermark on an image and after compressing it with various compression rates we try to detect a different watermark. The plot representing the percentage of the image where the watermark was misdetected is shown in Figure 2. The compression ratios are marked near each curve. As it can be seen from the two plots, the theoretical results under the given assumptions are very close to the simulation results.

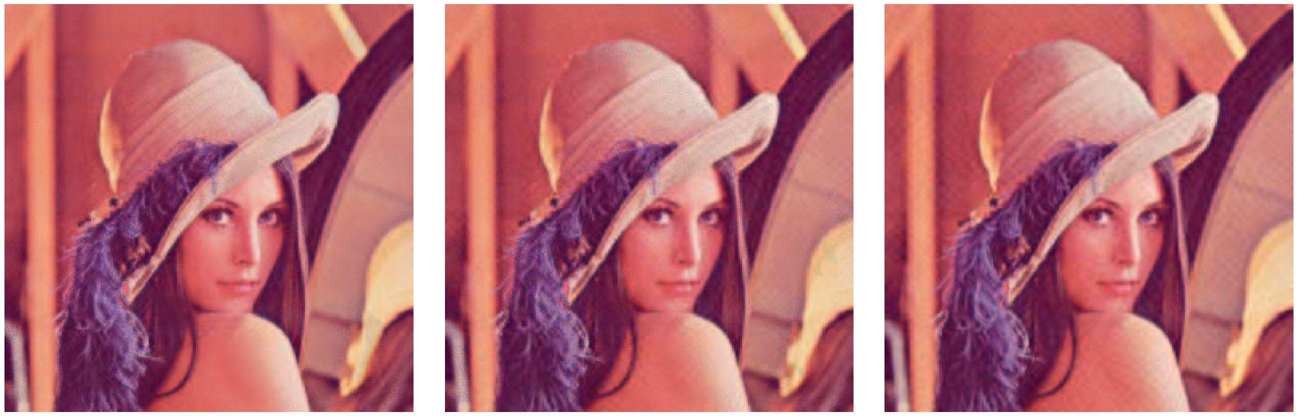
If the watermarked image is compressed with the compression ratio which was assumed in the embedding stage or if the watermark is detected considering the detection region size corresponding to the respective JPEG compression ratio, i.e., $d = D$ then, after denoting $\mu = l/D$ we obtain :

$$\frac{\int_{D_y^j} p(y|g, \mathbf{C}_j)}{\int_{D_y^k} p(y|g, \mathbf{C}_k)} = \frac{2}{\pi} \arccos\left(\frac{\mu}{2}\right) - \frac{1}{2}\sqrt{4\mu^2 - \mu^4}. \qquad (14)$$

The condition to fail identifying the wrong watermark j when assuming $c = 0.5$ is that $\mu > 0.8$. These results suggest that we can generate families of watermarks having the values of the DCT parameters situated at certain intervals from each other, where each watermark can be clearly distinguished from all others.

5. EXPERIMENTAL RESULTS

In [6] the proposed algorithms were applied on graylevel images. The watermark resistance to JPEG compression and to representative filtering algorithms was verified. In the case of color images we embed the watermark only in the luminance component (Y channel in the YUV color system). This procedure avoids the chrominance distortions and increases the efficiency under the JPEG color image compression [2]. The original "Lenna" image is shown in Figure 3 (a). The watermarked image based on (5) and (6), after it was compressed by JPEG at 1 bpp is shown in Figure 3 (b). In Figure 3 (c) the image watermarked by means of the algorithm provided in (5,7) and compressed at 1 bpp, is displayed. The image distortion produced by watermarking is measured by means of the Signal-to-Noise Ratio (SNR) between the watermarked image when assuming a certain embedding level D, and the original image. The plots representing the digital signature resistance to JPEG compression in the case of color images are shown in Figure 4 when embedding linear DCT constraint regions (10) and in Figure 5 when embedding circular DCT detection regions (11), respectively. The SNR, measured in dB, is written on each detection curve separately, and the watermark decision is marked by a dashed line. The watermarks applied in the color images are able to resist at even larger compression ratios than these reported if we subsample the U and V chrominance components. The watermarking detection results obtained when embedding circular DCT detection regions are better than those obtained when embedding linear DCT constraints. When the watermarks are designed to resist at higher compression ratios, the image distortion increases and becomes visible as a textured pattern. Because of the localization property, as provided by (4,5), the watermark can be detected from a certain part of the image, depending by the scanning order. This property



(a) original Lenna image (b) linear DCT constraint (c) circular DCT detection regions

Figure 3. Watermarked images after JPEG compression.

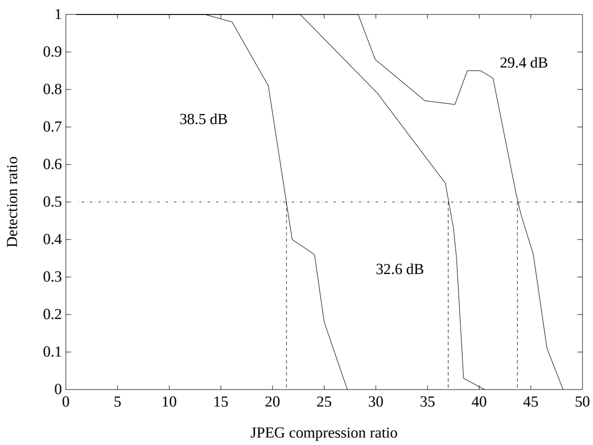


Figure 4. Watermark detection in color images when embedding linear DCT constraints.

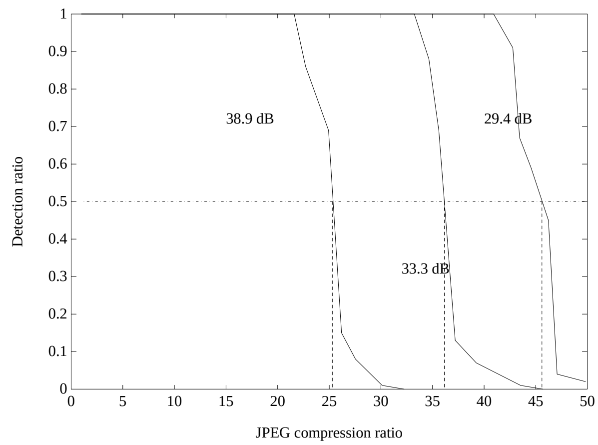


Figure 5. Watermark detection in color images when embedding circular DCT detection regions.

is useful in the case when watermarking image sequences. In this case, the respective watermark can be identified from each frame separately.

6. CONCLUSIONS

The increase in image transmission over the computer network determined the need to use watermarks (digital signatures) for image copyright protection. The proposed watermarking algorithms have two processing steps. In the first step, certain block locations are selected according to a Gaussian network classifier decision. In the second step, a DCT coefficient constraint is embedded in the selected blocks. Two different approaches are proposed for embedding the DCT constraint. The first one embeds a linear DCT constraint and the second defines circular DCT detection regions. The watermark detection is based on the probability detection theory. The simulation results suggest that the proposed algorithms generate watermarks which are able to resist at certain JPEG compression ratios. The extension of the proposed system to be applied in TV/Video systems is straightforward.

7. REFERENCES

[1] B. M. Macq, J.-J. Quisquater, "Cryptology for digital TV broadcasting," *Proc. of IEEE,* vol. 83, no. 6, pp. 944-957, June 1995.

[2] A. C. Hung, "PVRG-JPEG CODEC 1.1," *Stanford University,* Technical Report, 1993.

[3] I. Pitas, T. H. Kaskalis, "Applying signatures on digital images," *Proc. of IEEE Workshop on Nonlinear Signal and Image Processing,* Neos Marmaras, Greece, pp. 460-463, 20-22 June 1995.

[4] E. Koch, J. Zhao, "Towards robust and hidden image copyright labeling," *Proc. of IEEE Workshop on Nonlinear Signal and Image Processing,* Neos Marmaras, Greece, pp. 452-455, 20-22 June 1995.

[5] I. J. Cox, J. Kilian, T. Leighton, T. Shamoon, "Secure Spread Spectrum Watermarking for Multimedia," *NEC Research Institute,* Technical Report 95 - 10, 1995.

[6] A. G. Borş, I. Pitas, "Embedding parametric digital signatures in images," *Proc. of the European Signal Processing Conference (EUSIPCO'96),* Trieste, Italy, 10-13 Sep. 1996.