

# Lapped-orthogonal-transform-based adaptive image watermarking

Yuxin Liu\*

Bin Ni<sup>†</sup>

Xiaojun Feng

Edward J. Delp

Purdue University

School of Electrical and Computer Engineering

West Lafayette, Indiana 47907

E-mail: ace@ecn.purdue.edu

---

**Abstract.** A robust, invisible watermarking scheme is proposed for digital images, where the watermark is embedded using the block-based lapped orthogonal transform (LOT). The embedding process follows a spread spectrum watermarking approach. In contrast to the use of transforms such as discrete cosine transform, our LOT watermarking scheme allows larger watermark embedding energy while maintaining the same level of subjective invisibility. In particular, the use of LOT reduces block artifacts caused by the insertion of the watermark in a block-by-block manner, hence obtaining a better balance between invisibility and robustness. Moreover, we use a human visual system (HVS) model to adaptively adjust the energy of the watermark during embedding. In our HVS model, each block is categorized into one of four classes (texture, fine-texture, edge, and plain-area) by using a feature known as the texture masking energy. Blocks with edges are also classified according to the edge direction. The block classification is used to adjust the watermark embedding parameters for each block. © 2006 SPIE and IS&T. [DOI: 10.1117/1.2185658]

---

## 1 Introduction

In this paper we describe a robust, invisible watermark embedded in a host image. Such a watermark must be resistant to attacks and at the same time the visual quality of the image must be preserved. In order to improve the robustness of the watermark, one may increase the watermark embedding energy, which may increase the visibility of the watermark. Therefore, our goal is to achieve a balance between robustness and transparency in the embedding.

For invisible watermarking, the watermark must be inserted in such a way that the presence of the watermark is not noticeable under typical viewing conditions. Thus modeling the human vision system (HVS) plays a critical role in watermarking.<sup>1,2</sup> Algorithms for general watermark embedding/detection are described in Refs. 3–15. Digital watermarking is widely used for copyright protection.

---

\*Y. Liu is now with Sun Labs, Sun Microsystems Inc., Menlo Park, CA 94025 USA (E-mail: zoeliu@ieee.org).

<sup>†</sup>B. Ni is now with Marvell Semiconductor, Santa Clara, CA 95054 USA (E-mail: binn@marvell.com).

Paper 04019 received Feb. 27, 2004; revised manuscript received Jun. 9, 2005; accepted for publication Sep. 6, 2005; published online Mar. 24, 2006.

1017-9909/2006/15(1)/013009/\$22.00 © 2006 SPIE and IS&T.

Document and printing security using digital watermarking techniques are addressed in Refs. 16 and 17. The development of spread spectrum watermarking is a milestone in invisible robust watermarking.<sup>18,19</sup> The watermark is embedded using spread spectrum techniques in the most significant components in a host signal, thus guaranteeing its robustness. Approaches of watermarking evaluation are explored and benchmarking tools are developed.<sup>20–23</sup>

It is worth mentioning that with respect to image quality, watermarking is closely related to image compression. HVS research has been actively exploited in lossy image compression to improve the coding efficiency while maintaining a good subjective quality. Theoretically, every aspect of HVS research can be used in watermarking as well, particularly in trying to determine how the watermark will manifest itself in the watermarked image.

It has been recognized that watermark embedding in the frequency domain has many advantages in terms of robustness and transparency. Many watermarking methods have been based on this concept. Some of the methods obtain transforms of the entire image<sup>18</sup> while others use block-based transforms such as the discrete cosine transform (DCT).<sup>24</sup> The method we propose in this paper is based on the block-based lapped orthogonal transform (LOT).<sup>25</sup> The advantage of using LOT as opposed to conventional block-based transforms such as DCT is that it reduces the block artifacts caused by the watermark embedding. LOT was used in secure watermarking in Ref. 26 with the watermark embedded using a template with log-polar-mapping to combat the geometrical attacks. Our new method differs in the way the watermark is embedded using an HVS model. We use two types of characteristics in HVS to embed the watermark in an image-content-adaptive manner, thus improving our watermarking performance. Our technique does not currently use a synchronization template.

## 2 General Framework

In this paper, we combine the use of block-based LOT to combat the block artifacts in watermark embedding, and an HVS perceptual model for block classification to adaptively adjust the watermark embedding energy. Our proposed approach can achieve a better trade-off to balance the robustness and transparency of the embedded watermark. Also,

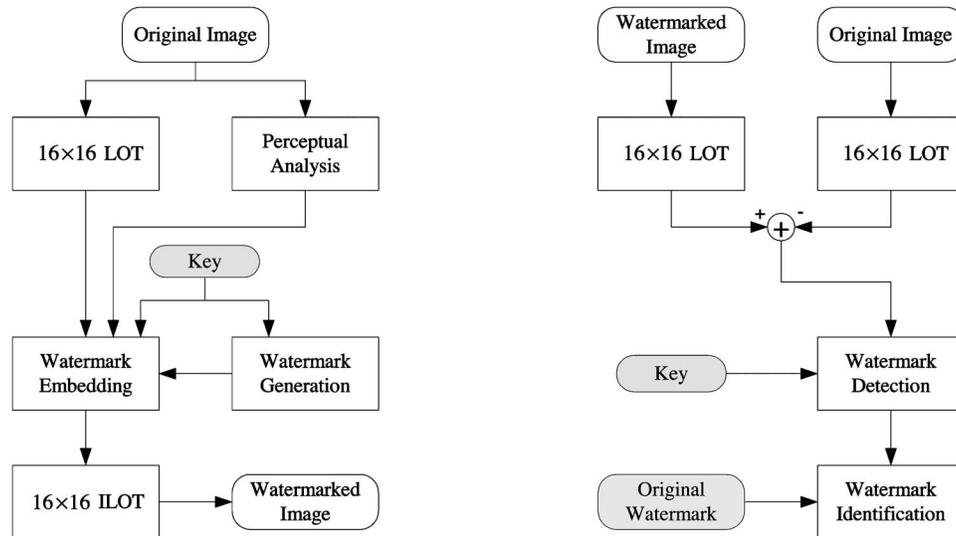


Fig. 1 Block diagram of nonblind LOT watermarking.

the use of LOT and our HVS model can be generalized and incorporated to any existing block-based watermarking schemes, either blind or nonblind.

Specifically, we describe the detailed design of a nonblind watermarking scheme as a proof of the effectiveness of our approach. As shown in Fig. 1, our watermarking scheme mainly follows the spread spectrum watermarking approach.<sup>18,27</sup> We propose a block-based robust, invisible watermarking scheme. We use block-based LOT and classify the blocks into specific categories using the HVS perceptual model.

As opposed to  $8 \times 8$  block-based DCT, LOT divides an image into overlapping  $16 \times 16$  blocks in the spatial domain and maps each block into an  $8 \times 8$  block in the frequency domain, hence obtaining an image in the transform domain with the same size as the original. The use of LOT helps reduce the block artifacts caused by block-based embedding. Details of LOT are discussed in Sec. 3.1.

To obtain a balance between robustness and invisibility of the embedded watermark, we introduce a perceptual analysis module that uses HVS properties to adaptively embed the watermark in each block. The perceptual analysis module first extracts a feature from each block, namely the texture masking energy (TME), described below. The blocks are then classified into four categories according to the TME. Blocks classified as edge by the use of the TME are further classified into four edge classes using an edge detector. The effect of this is that each block is classified into one of seven classes. Details of the perceptual analysis are presented in Sec. 3.2.

The watermark embedding energy in each block is adjusted accordingly to adapt to the HVS sensitivities to errors. We design a quantization matrix for each category of the blocks through incorporating the HVS analysis results. We take the quantization matrix as a just noticed difference (JND)-like matrix to adaptively adjust the watermark embedding energy to retain the transparency. We address the details of the design for the JND quantization matrix in Sec. 3.3.

An example of blind watermarking scheme that embeds

a binary watermark sequence is to quantize a transform coefficient of the host image and round it to either its nearest odd-numbered or nearest even-numbered quantized value based on the embedded watermark bin.<sup>28–30</sup> Using our approach, block-based LOT can be adopted and the quantization step for each hosting transform coefficient can be appropriately designed in a similar way as that for the nonblind watermarking scheme we describe in this paper.

### 3 Watermarking Using LOT and HVS Perceptual Analysis

#### 3.1 Lapped Orthogonal Transform

LOT is described in-depth in Ref. 25, which presents a DCT-based fast implementation for LOT and its inverse transform (ILOT). In the one-dimensional (1-D) case, for a discrete 1-D signal  $x_0$  with length  $MN$ , LOT divides the signal into  $M$  segments where each segment has length  $L$  that satisfies  $N \leq L \leq 2N$ . This means that every two adjacent segments have an overlapped portion of length  $(L - N)/2$ . LOT transforms each signal fraction into an  $N$ -entry coefficient vector, and thus the transformed signal has the same total length as the original signal. Like some other traditional transforms, it is straightforward to derive the two-dimensional (2-D) separable LOT from the 1-D version. As shown in Fig. 2, the 2D LOT maps every  $L \times L$  block in the spatial domain to an  $N \times N$  block in the frequency domain. We chose  $N=8$  and  $L=2N$  in our implementation. Since every block in the transform domain contains information of the adjacent blocks in the spatial domain, block artifacts caused by any processing in a block-by-block manner in the transform domain will not be as noticeable in the spatial domain. LOT usually requires 20–30% more computational complexity than DCT. LOT has been used in image and video compression.<sup>31</sup>

#### 3.2 Block Categorization Using HVS Perceptual Analysis

Characteristics of the HVS have been studied and exploited in the area of image and video compression.<sup>32</sup> The limits of

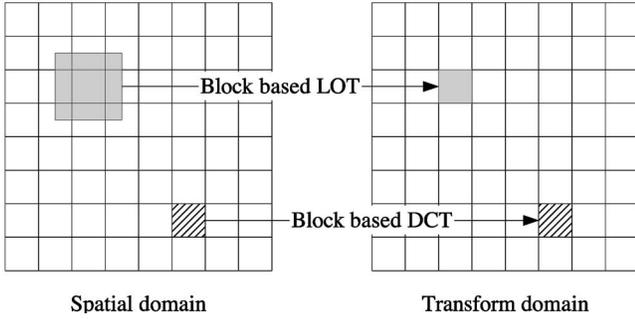


Fig. 2 Block-based LOT as opposed to block-based DCT.

the HVS in the intensity, spatial, and temple domain allows images to be represented with fewer bits but adequate subjective visual quality. For example, the HVS has different sensitivities to different image details. As a result, changes in a “flat” region are more perceptible than the same amount of changes in a complex texture region. The HVS perceptual analysis also provides methods in digital watermarking to increase the energy of the embedded watermark while keep it imperceptible.<sup>1,2</sup>

In our perceptual model, two types of HVS properties are used: the sensitivity to the complexity of texture and the sensitivity to the orientation of edges. We classify the image blocks into different categories based on these HVS properties. The block-by-block watermark embedding is then implemented accordingly. We design a JND-like matrix for each category of the blocks so that the capacity of each host block is maximized to retain the invisibility of the embedded watermark.

We first extract an HVS feature, known as the texture masking energy (TME), from each block as follows<sup>33</sup>:

$$T_E(k) = \left[ \sum_{\substack{i=0 \\ (i,j) \neq (0,0)}}^{N-1} \sum_{j=0}^{N-1} \hat{H}^{-1}(i,j)^2 X_k(i,j)^2 \right]^{1/2} \quad (1)$$

In the above equation,  $X_k(i,j)$  denotes the LOT transform coefficient located in position  $(i,j)$  of the  $k$ 'th block, and  $\hat{H}(f)$  is the HVS relative sensitivity function with respect to the spatial frequency  $f$ , where  $f$  is related with the spatial position  $(i,j)$  as follows:

$$f(\text{cycles/degree}) = \frac{\sqrt{i^2 + j^2}}{2N} (\text{cycles/pixel}) \times f_s (\text{pixels/degree}), \quad (2)$$

and we chose  $f_s = 32$ . The function  $\hat{H}(f)$  can be obtained as<sup>33</sup>

$$\hat{H}(f) = |A(f)|H(f),$$

$$H(f) = (0.31 + 0.69f)\exp(-0.29f),$$



Fig. 3 An enlarged block classification using the HVS perceptual analysis of crowd (Left: Block classification using the TME, where cross-hatched regions denote texture, hatched regions denote fine-texture, solid-line-bounded regions denote edge, and unmarked regions denote plain-area; Right: Edge classification, where line segments denote the edge orientation).

$$A(f) = \left\{ \frac{1}{4} + \frac{1}{\pi^2} \left[ \ln \left( \frac{2\pi f}{\delta} + \sqrt{\frac{4\pi^2 f^2}{\delta^2} + 1} \right) \right]^2 \right\}^{1/2},$$

where  $\delta = 11.636 \text{ deg}^{-1}$ .

Note that  $T_E^2(k)$  is a sum of the energy in the  $k$ 'th block weighted by the reciprocal of the square of the sensitivity function  $\hat{H}^2(f)$ . This implies that the TME provides a metric measuring the insensitivity, or equivalently, the resistance capability of a block to noise. We also divide each  $N \times N$  block into four  $N/2 \times N/2$  subblocks, and obtain the TME for each subblock, namely  $T_E(k,l)$ ,  $l=1,2,3,4$ , in a similar way.

Since the TME represents the insensitivities of a block and its subportions to noise, we use both  $T_E(k)$  and  $T_E(k,l)$  to classify the  $k$ 'th  $N \times N$  block into four major categories: texture, fine-texture, edge, and plain-area. This is done using a similar approach as in Ref. 33.

Studies of the HVS have shown that different edge orientations have different masking capabilities. The HVS is less sensitive to changes along the direction of an edge than changes perpendicular to the edge. Thus we further classify each edge block into four types according to the edge direction: vertical, horizontal, diagonal, and antidiagonal, using the Sobel operator.<sup>31</sup>

An example of the result of our perceptual analysis is shown in Fig. 3.

### 3.3 Design of JND Quantization Matrix

In this subsection, we design a JND quantization matrix for each category of the blocks. We take the standard quantization matrix used in JPEG<sup>34</sup> and create our JND quantization matrix through modulating the JPEG matrix with two specific matrices for each block category. Half of the quantization step for one LOT transform coefficient is treated as a measure of the just-noticed difference threshold. Hence the amount of watermark embedding energy can be determined without violating the invisibility.

We obtain our JND quantization matrix for the  $k$ 'th block as

$$Q_L(k) = Q_S \times M_{block}(k) \times M_{edge}(k), \quad (3)$$

where “ $\times$ ” denotes the element-wise multiplication, and  $Q_S$  is the standard quantization matrix used in JPEG,<sup>34</sup> which reflects the HVS sensitivity to spatial frequencies.<sup>24</sup> Any “noise” added to the transform coefficients shall be barely

perceptible, as long as the magnitude of the noise is within half of the quantization steps designated by  $Q_S$ . We design two additional matrices,  $M_{\text{block}}(k)$  and  $M_{\text{edge}}(k)$ , to modulate the quantization steps in  $Q_S$  for the  $k$ 'th block, based on our perceptual analysis,

$$M_{\text{block}}(k) = \begin{cases} 0.8 & \text{plain-area} \\ 1.0 + \min[T_E(k) \times 0.25/\omega_{\text{edge}}, 0.25] & \text{edge} \\ 1.0 + \min[T_E(k) \times 0.25/\omega_{\text{fine-texture}}, 0.25] & \text{fine-texture} \\ 1.25 + \min[T_E(k) \times 0.25/\omega_{\text{texture}}, 0.25] & \text{texture} \end{cases} \quad (4)$$

$$M_{\text{edge}} = \begin{cases} M_1 & \text{horizontal edge} \\ M_1^T & \text{vertical edge} \\ M_2 & \text{diagonal/anti-diagonal edge} \end{cases} \quad (5)$$

In Eq. (4), we chose  $\omega_{\text{edge}}=600$ ,  $\omega_{\text{fine-texture}}=300$ , and  $\omega_{\text{texture}}=400$ . In Eq. (5),  $M_1$  and  $M_2$  are designed as

$$M_1 = \begin{bmatrix} 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.1 & 1.3 & 1.3 \\ 1.1 & 1.0 & 1.0 & 1.0 & 1.0 & 1.1 & 1.3 & 1.3 \\ 1.2 & 1.2 & 1.0 & 1.0 & 1.1 & 1.1 & 1.3 & 1.4 \\ 1.2 & 1.2 & 1.2 & 1.0 & 1.1 & 1.3 & 1.3 & 1.4 \\ 1.2 & 1.2 & 1.2 & 1.2 & 1.3 & 1.3 & 1.4 & 1.5 \\ 1.2 & 1.2 & 1.2 & 1.3 & 1.3 & 1.4 & 1.5 & 1.5 \\ 1.2 & 1.2 & 1.3 & 1.3 & 1.4 & 1.5 & 1.5 & 1.5 \\ 1.2 & 1.3 & 1.4 & 1.4 & 1.5 & 1.5 & 1.5 & 1.5 \end{bmatrix}, \quad (6)$$

$$M_2 = \begin{bmatrix} 1.0 & 1.0 & 1.0 & 1.1 & 1.2 & 1.2 & 1.3 & 1.3 \\ 1.0 & 1.0 & 1.0 & 1.0 & 1.2 & 1.2 & 1.3 & 1.3 \\ 1.1 & 1.0 & 1.0 & 1.0 & 1.2 & 1.2 & 1.3 & 1.4 \\ 1.1 & 1.0 & 1.2 & 1.0 & 1.2 & 1.2 & 1.3 & 1.4 \\ 1.2 & 1.2 & 1.2 & 1.2 & 1.2 & 1.2 & 1.4 & 1.5 \\ 1.2 & 1.2 & 1.2 & 1.2 & 1.2 & 1.4 & 1.5 & 1.5 \\ 1.3 & 1.3 & 1.3 & 1.3 & 1.4 & 1.5 & 1.5 & 1.5 \\ 1.3 & 1.3 & 1.4 & 1.4 & 1.5 & 1.5 & 1.5 & 1.5 \end{bmatrix},$$

and  $M_1^T$  is the transpose of  $M_1$ .

The design of matrix  $M_{\text{block}}(k)$  in Eq. (4) used the HVS analysis result from the first step of our block classification approach. As we have discussed, HVS has the property that errors in high complex texture areas are less sensitive to those in low complex texture areas. The TME of each block is a measure of the texture complexity in the block, or equivalently, a measure of the resistance capability of the block to quantization errors. Hence the classification result using TME has provided the ability and extent of error insensitivity of each block category.

The four main block categories, sorted in an ascending order in terms of error insensitivity, are: plain-area, edge, fine-texture, and texture. The more insensitive a block is to errors, the larger amount of watermark embedding energy the block may allow to retain sufficient watermark invis-

ibility. Hence we generally scale down the quantization steps for blocks classified as plain-area to 0.8, but increase the quantization steps for edge and fine-texture blocks to 1.0~1.25, and texture blocks to 1.25~1.5. Moreover, we fine-tune the embedding energy for each block according to its own TME value. We empirically chose the thresholds for the TME value of each block category  $\omega_{\text{edge}}$ ,  $\omega_{\text{fine-texture}}$ , and  $\omega_{\text{texture}}$  through examining the average TME value of that specific category of blocks.

In addition to considering texture complexity of one block, we also mentioned that HVS has different sensitivities to errors present at different edge orientations. The design of matrix  $M_{\text{edge}}$  in Eqs. (5) and (6) aims to quantify this HVS property. It has been pointed out in Ref. 35 that the JND on both sides of an edge is raised up when approaching the edge, and the edge blocks concentrate their energy on the specific spatial frequency band at the particular orientation. Hence we may increase the quantization step size at those spatial frequencies in the vicinity of the particular edge band. We design different quantization modulation matrices  $M_{\text{edge}}$  for different edge blocks of particular edge orientations. The design of  $M_{\text{edge}}$  is empirically derived, which mainly depends on the edge structure in the edge block. The entries of  $M_{\text{edge}}$  only remain 1.0 in the specific edge frequency band, whereas the remaining entries have increased to a value larger than one. Entries at higher frequencies are larger than those at lower frequencies.

The entries in our JND matrix  $Q_L$  are an indication of the insensitivities of the LOT coefficients subject to quantization noise. That is, the smaller the entry in  $Q_L$ , the more perceptually significant the corresponding LOT coefficient is when the image block is distorted by quantization noise.

The modulated JND quantization matrix  $Q_L$  we designed here is an extended work of those described in Refs. 31 and 33. In our paper, we mainly use the concept of TME since we believe it is a better quantitative measure of HVS error insensitivity compared to other HVS analysis models based on our experience. We create two modulation matrices,  $M_{\text{block}}$  and  $M_{\text{edge}}$ , using the two-step block classification results. The design of our JND matrix could be used for both lossy image coding and invisible image watermarking. We will present the effectiveness of our HVS analysis for invisible image watermarking in the experimental results.

### 3.4 Watermark Generation, Embedding, Detection, and Identification

In this subsection, we use the spread spectrum approach to generate, embed, and extract the image watermark.<sup>18</sup> It is argued in Ref. 18 that a robust, invisible image watermark should be constructed as an independent, identically distributed (i.i.d.) Gaussian random sequence and embedded in the perceptually most significant spectral components of the host image following a spread-spectrum-like fashion. It has been proved that such watermarking schemes enable the robustness and security of the embedded watermark, since an attacker cannot remove the watermark without seriously degrading the fundamental structural components of the host image, and meanwhile, most image processing operations tend to leave the perceptually significant compo-



Fig. 4 Original images (Left: crowd, Right: umas).

nents intact. The choice of Gaussian distribution makes the embedded watermark capable of combating collusion attacks.

The invisibility of the watermark embedded in the perceptually most significant components is facilitated by the spread spectrum theory. An analogy exists between watermark embedding and secure signal communications over channels with noise. The frequency domain of the host image can be viewed as a communication channel, the watermark can be treated as a transmitted signal, and the malicious and unintentional lossy processing operations can be regarded as the channel noise. Hence, the same rationale used for secure signal communications can be applied in the field of invisible, robust watermarking. In spread spectrum communications, a narrow-band signal is transmitted over a much larger bandwidth such that the signal energy spread in any single frequency component is undetectable. Analogously, a watermark can be embedded over many frequency coefficients of the host image and the energy in any single coefficient is hence undetectable. The embedded watermark can only be destroyed through introducing noise of

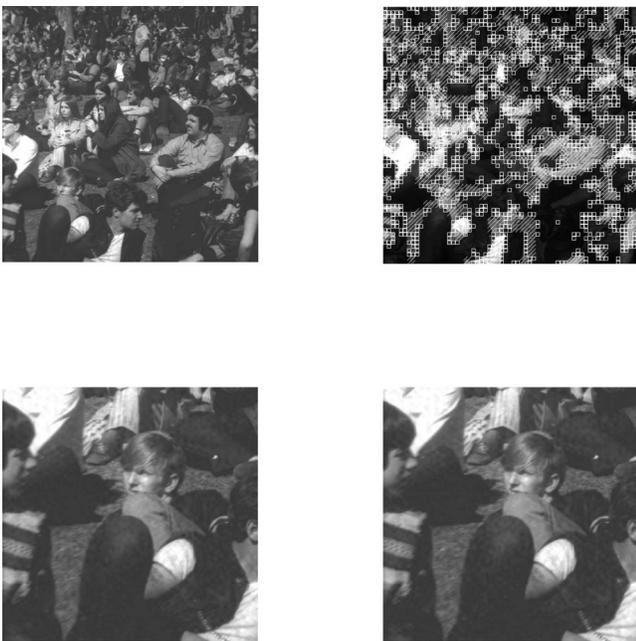


Fig. 5 LOT image watermarking for crowd,  $\alpha=1.0$  (Top-left: Original; Top-right: HVS analysis; Bottom-left; Watermarked using HVS, enlarged bottom-left part; Bottom-right: Watermarked without HVS, enlarged bottom-left part).

Table 1 Watermark identification after attacks using common filtering operations [ $\text{sim}(W, \hat{W})$ ].

Image	Scheme	$\alpha$	Median		Gaussian 3×3	Average 3×3	Sharpen
			2×2	3×3			
crowd	LOC	0.25	15.8	32.3	48.4	36.7	30.2
	DCT		15.7	28.9	44.1	34.5	33.2
	LOT	0.50	31.9	61.6	83.5	67.2	56.5
	DCT		30.8	57.5	76.9	62.9	62.2
	LOT	2.00	94.2	127.0	133.8	127.7	120.0
umas	DCT		90.0	124.7	130.1	122.7	120.5
	LOT	0.25	19.3	39.3	54.7	39.4	33.8
	DCT		15.7	35.7	42.1	33.2	40.5
	LOT	0.50	36.3	71.2	90.8	70.8	62.6
	DCT		31.1	66.2	75.7	62.7	71.5
	LOT	2.00	98.3	130.4	135.8	129.4	124.8
	DCT		91.2	129.0	130.7	123.8	125.8

large amplitude spread over many host coefficients, which will consequently cause serious image quality degradation. In contrast, at the stage of watermark detection and identification, using the knowledge of the location and content for the embedded watermark, it is possible to collect these many weak signals spread over a large range of host coefficients and provide a single output with sufficiently high signal-to-noise ratio (SNR).

We use the results of the perceptual analysis and the JND matrix described in Sec. 3.2 and 3.3 to adaptively embed the watermark in a block-by-block manner in the LOT transform domain. Motivated by the concept of spread spectrum watermarking,<sup>18</sup> we form our watermark using an independent, identically Gaussian-distributed pseudorandom sequence, with zero mean and unit variance. In our implementation, we fixed the number of LOT coefficients that bear the watermark in each block to be 5. Thus for a  $512 \times 512$  image, the total length of the embedded watermark is 20,480.

We choose the five visually most important AC coefficients in each block, whose corresponding entries in  $Q_L$  have the least values [except the one located in the (0,0) position that is excluded from watermark embedding], to bear the watermark. The watermark is embedded in the following way:

$$X'_k(i_{n,k}, j_{n,k}) = X_k(i_{n,k}, j_{n,k}) + \alpha Q_L(i_{n,k}, j_{n,k}) w(n), \quad (7)$$

where  $w(n)$  is the watermark element,  $(i_{n,k}, j_{n,k})$  denotes the specified position to bear the watermark in the  $k$ 'th block,  $X$  denotes the LOT coefficient of the original image,  $X'$  is the corresponding watermarked coefficient, and  $\alpha$  determines the overall intensity of the watermark embedding energy,

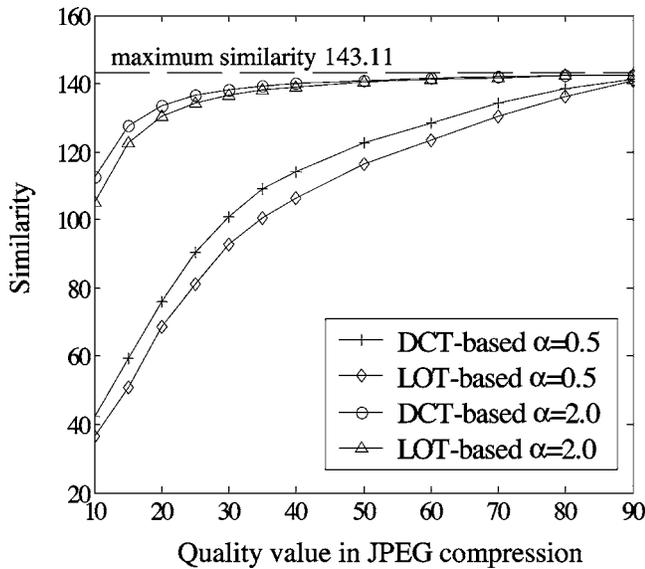


Fig. 6 LOT vs. DCT watermarking using a lossy JPEG compression attack to crowd.

which usually takes a value between 1/6 to 1.

To detect the watermark we need the original unwatermarked image, as well as the location and intensity information. We take LOT for both the original image and the possibly corrupted watermarked image. For each watermarked frequency component, we extract the watermark using

$$\hat{w}(n) = [\hat{X}'_k(i_{n,k}, j_{n,k}) - X_k(i_{n,k}, j_{n,k})] / [\alpha Q_L(i_{n,k}, j_{n,k})]. \quad (8)$$

Let  $\hat{W}$  denote the extracted watermark vector, and  $W$  a candidate watermark. We identify the watermark by obtaining the similarity (correlation) as follows:

$$sim(W, \hat{W}) = \frac{W \cdot \hat{W}}{\sqrt{\hat{W} \cdot \hat{W}}}, \quad (9)$$

where “ $\cdot$ ” denotes the inner production. It is easy to show that if  $W$  is a sequence of i.i.d. random variables with zero-mean unit variance, and also independent of  $\hat{W}$ ,  $sim(W, \hat{W})$  is a zero-mean unit-variance Gaussian random variable. If  $sim(W, \hat{W})$  is greater than a threshold, we claim that  $\hat{W}$  comes from the original watermark  $W$ , hence identifying the embedded watermark. In our implementation, we chose 6 as the threshold, which corresponds to a false alarm probability of  $10^{-9}$  [ $=1 - \Phi(6)$ , where  $\Phi(x)$  denotes the cumulative

Table 2 JPEG attack performance of LOT watermarking with increased  $\alpha$  crowd [ $sim(W, \hat{W})$ ].

Quality	10	20	30	40	50	70	90
LOT ( $\alpha=0.60$ )	42.8	78.0	101.6	114.8	123.0	133.8	141.4
DCT ( $\alpha=0.50$ )	42.1	76.1	100.7	114.0	122.8	134.0	141.4

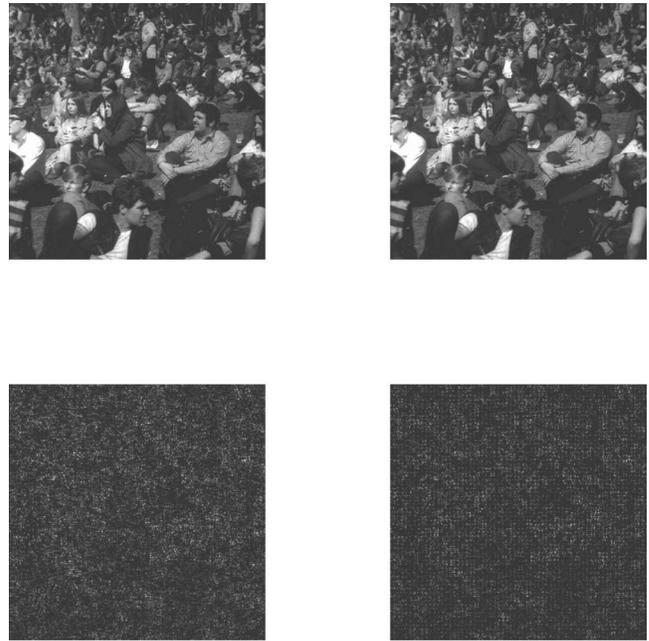


Fig. 7 LOT vs. DCT image watermarking for crowd (Top-left: Watermarked image using LOT,  $\alpha=0.6$ ; Top-right: Watermarked image using DCT,  $\alpha=0.5$ ; Bottom-left: Difference image using LOT,  $\alpha=0.6$ ; Bottom-right: Difference image using DCT,  $\alpha=0.5$ ).

distribution function (c.d.f.) of a Gaussian random variable with zero-mean unit variance]. Note that when  $\hat{W} = W$ ,  $sim(W, \hat{W})$  achieves its maximum value  $143.11 = \lceil [5 \times 512 \times 512 / (8 \times 8)]^{1/2} \rceil$ .

#### 4 Experimental Results

We used our LOT watermarking scheme to watermark several  $512 \times 512$  gray-scale images shown in Fig. 4. The Stirmark 3.1 benchmark<sup>21</sup> is used to attack the watermarked images and test the robustness of the embedded watermark.

To demonstrate the effectiveness of our HVS analysis and the use of our JND matrix, we compare the two watermarked images with the same watermarking embedding intensity, i.e., using the same  $\alpha$ . One image has the watermark embedded using the HVS result and one without the HVS result. For the one without using HVS analysis, we take the JPEG standard quantization matrix  $Q_S$  instead of our JND matrix  $Q_L$  for watermark generation, embedding, and extraction, similar to the approach described in Ref. 24. Due to the same  $\alpha$  that is used, both watermarked images have demonstrated almost the same robustness capability under various attacks and lossy image operations such as JPEG. However, as shown in Fig. 5, the watermarked image without using HVS demonstrates a larger quality degradation compared to the one using HVS, especially in those flat areas, for example in the area of the pants. This implies the use of HVS analysis improves the transparency of the embedded watermark at the same level of robustness.

To demonstrate the advantages of using LOT, we implemented a block-based DCT watermarking scheme for comparison. The only difference between the two schemes is the transform.

**Table 3** LOT watermark identification using a cropping attack to crowd ( $\alpha=0.60$ ).

Cropping [upper-left position (x,y)]	Area size: 171×170			Area size 51×51		
	(193, 169)	(292, 45)	(57, 290)	(150, 125)	(315, 53)	(15, 252)
[ $sim(W, \hat{W})$ ]	48.1	47.5	49.3	14.9	14.1	14.7

It can be observed from Table 1 that under the same embedding intensity  $\alpha$ , our LOT scheme is slightly more resistant to blurring attacks, but slightly more vulnerable to sharpening attacks.

Figure 6 suggests that using the same  $\alpha$ , the DCT watermarking is slightly more resistant to lossy JPEG compression. We have argued that one distinguished advantage of the block-based LOT, as opposed to the block-based DCT, is its capability of reducing block artifacts. Thus we can choose a slightly larger  $\alpha$  for our LOT method to increase its robustness without compromising the transparency. For example, if we chose  $\alpha=0.60$ , the LOT watermarking scheme will outperform the DCT scheme with  $\alpha=0.50$  as described in Table 2. The watermarked images and the difference images between the original image and the watermarked one are given in Fig. 7. Due to the larger intensity  $\alpha$  chosen for LOT, the LOT difference image presents more quantization noise compared with that using DCT. Nevertheless, the LOT watermarked image has the same or even better subjective quality compared with the DCT watermarked image, showing that LOT effectively reduces the block artifacts.

Our LOT watermark survives the cropping attack, as shown in Table 3. Intuitively, if we lose 90% of the image, we lose 90% of the watermark sequence. However, considering our goal is to verify the similarity between the extracted sequence and the original sequence, losing 90% of the extracted data only makes the similarity function to be

$1/\sqrt{10}$  of its original value. Usually the correlation is still well above 6, so we can still make a correct decision after a severe cropping. We chose two cropping sizes,  $171 \times 170$  and  $51 \times 51$ , which correspond to  $1/9$  and  $1/100$  of the original image size, respectively. The theoretical values for the maximum similarity between the extracted watermark and a candidate watermark [i.e.,  $sim(W, \hat{W})$ ] are  $47.70 [= (5 \times 512 \times 512 / 9)^{1/2}]$  and  $14.31 [= (5 \times 512 \times 512 / 100)^{1/2}]$  for the above two size choices. As discussed in Ref. 18, since we are using a nonblind approach in our watermarking, i.e., we have the original image available upon watermark detection, the watermark detector can easily obtain the position of the watermarked cropped area relative to the original image (such as using image correlation). For each cropping size, we tried three different cropping locations, as given in Table 3 and Fig. 8. In addition, we used the Stirmark 3.1 JPEG attack on the cropped area, and our LOT watermarking can still survive as long as the cropping size is large enough, as shown in Table 4.

One approach to improve our LOT watermarking scheme is to adaptively adjust the number of watermark elements embedded in each block rather than fix it at 5. A better transparency-robustness balance should be expected.

### 5 Conclusion

In this paper, we presented a LOT-based adaptive image watermarking scheme. LOT inherits all the best properties



(292, 45)



(462,214)

**Fig. 8** Image cropping attack (Left: Original image with the cropping area marked; Right: Enlarged cropped area,  $171 \times 170$ ).

**Table 4** LOT watermark identification under cropping and JPEG lossy compression attacks to crowd (cropping size of  $171 \times 170$  with upper-left corner (292, 45);  $\alpha=0.6$ ).

Quality	10	15	20	30	40	50	70	90
$\text{sim}(W, \hat{W})$	6.6	8.2	11.3	16.6	20.8	24.0	33.4	43.5

of the DCT, such as decorrelation and energy compaction, except that it requires 20–30% more computational complexity than DCT. Specifically, LOT greatly reduced block artifacts compared with DCT, making the embedded watermark less perceptible with same embedding energy. In addition, we used the HVS perceptual analysis to embed the watermark in an adaptive manner. Our experiments showed that our scheme obtains a good balance between robustness and transparency of the watermark. The use of LOT and our HVS model can be generalized and incorporated to any existing block-based watermarking schemes.

### Acknowledgments

The authors would like to thank the anonymous reviewers for their valuable suggestions and comments to help improve the quality of the paper. This work was partially supported by the Air Force Research Laboratory, Information Directorate, Rome Research Site, under research grant number F30602-02-2-0199.

### References

- R. B. Wolfgang, C. I. Podilchuk, and E. J. Delp, "Perceptual watermarks for digital images and video," *Proc. IEEE* **87**(7), 1108–1126 (July 1999).
- C. D. Vleeschouwer, J.-F. Delaigle, and B. Macq, "Invisibility and application functionalities in perceptual watermarking—an overview," *Proc. IEEE* **90**(1), 64–77 (Jan. 2002).
- C. I. Podilchuk and E. J. Delp, "Digital watermarking: algorithms and applications," *IEEE Signal Process. Mag.* **18**(4), 33–46 (July 2001).
- M. D. Swanson, M. Kobayashi, and A. H. Tewfik, "Multimedia data-embedding and watermarking technologies," *Proc. IEEE* **86**(6), 1064–1087 (June 1998).
- A. M. Eskicioglu and E. J. Delp, "An overview of multimedia content protection in consumer electronics devices," *Signal Process. Image Commun.* **16**(7), 681–699 (April 2001).
- F. Hartung and F. Ramme, "Digital rights management and watermarking of multimedia content for m-commerce applications," *IEEE Commun. Mag.* **38**(11), 78–84 (Nov. 2000).
- F. Hartung and M. Kutter, "Multimedia watermarking techniques," *Proc. IEEE* **87**(7), 1079–1107 (July 1999).
- J. A. Bloom, I. J. Cox, T. Kalker, J.-P. M. G. Linnartz, M. L. Miller, and C. B. S. Traw, "Copy protection for DVD video," *Proc. IEEE* **87**(7), 1267–1276 (July 1999).
- G. Voyatzis and I. Pitas, "The use of watermarks in the protection of digital multimedia products," *Proc. IEEE* **87**(7), 1197–1207 (July 1999).
- M. Barni, C. I. Podilchuk, F. Bartolini, and E. J. Delp, "Watermark embedding: Hiding a signal within a cover image," *IEEE Commun. Mag.* **39**(8), 102–108 (Aug. 2001).
- E. T. Lin, C. I. Podilchuk, T. Kalker, and E. J. Delp, "Streaming video and rate scalable compression: what are the challenges for watermarking?" *J. Electron. Imaging* **13**(1), 198–208 (2004).
- A. M. Eskicioglu, J. Town, and E. J. Delp, "Security of digital entertainment content from creation to consumption," *Signal Process. Image Commun.* **18**(4), 237–262 (April 2003).
- E. T. Lin, A. M. Eskicioglu, R. L. Lagendijk, and E. J. Delp, "Advances in digital video content protection," *Proc. IEEE* (submitted for publication).
- E. T. Lin and E. J. Delp, "Temporal synchronization in video watermarking," *IEEE Trans. Sig. Processing: Special Issue on Secure Media* (to be published).
- P. Moulin and M. K. Mihcak, "A framework for evaluating the data-hiding capacity of image sources," *IEEE Trans. Image Process.* **11**(9), 1029–1042 (Sep. 2002).

- E. J. Delp, "Is your document safe: an overview of document and print security," *Proc. Intl. Conf. Digital Printing Technologies*, San Diego, CA, September 29–October 4, 2002.
- A. K. Mikkilineni, G. N. Ali, P. Chiang, G. T. C. Chiu, J. P. Allebach, and E. J. Delp, "Signature-embedding in printed documents for security and forensic applications," *Proc. SPIE* **5306**, 455–466 (2004).
- I. J. Cox, J. Kilian, T. Leighton, and T. Shamoan, "Secure spread spectrum watermarking for multimedia," *IEEE Trans. Image Process.* **6**(12), 1673–1687 (Dec. 1997).
- I. J. Cox, M. L. Miller, and A. L. McKellips, "Watermarking as communications with side information," *Proc. IEEE* **87**(7), 1127–1141 (July 1999).
- F. A. P. Petitcolas, "Watermarking schemes evaluation," *IEEE Signal Process. Mag.* **17**(5), 58–64 (Sep. 2000).
- Stirmark 3.1, <http://www.petitcolas.net/fabien/watermarking/stirmark31/>
- B. Macq, J. Dittmann, and E. J. Delp, "Benchmarking of image watermarking algorithms for digital rights management," *Proc. IEEE* (to be published).
- H. C. Kim, H. Ogunleye, O. Guitart, and E. J. Delp, "The watermark evaluation tested (WET)," *Proc. SPIE* **5306**, 236–247 (2004).
- B. Tao and B. Dickinson, "Adaptive watermarking in the DCT domain," *IEEE Intl. Conf. Acoustics, Speech, and Signal Processing*, Vol. 4, pp. 2985–2988, Munich, Germany (April 1997).
- H. S. Malvar and D. H. Staelin, "The LOT: transform coding without blocking effects," *IEEE Trans. Acoust., Speech, Signal Process.* **37**(4), 553–559 (Apr. 1989).
- S. Pereira, J. J. K. O. Ruanaidh, and T. Pun, "Secure robust digital watermarking using the Lapped Orthogonal Transform," *Proc. SPIE* **3657**, 21–30 (1999).
- S. Pateux and G. L. Guelvout, "Practical watermarking scheme based on wide spread spectrum and game theory," *Signal Process. Image Commun.* **18**(4), 283–296 (Apr. 2003).
- D. Kundur and D. Hatzinakos, "Digital watermarking using multi-resolution wavelet decomposition," *Proc. IEEE Intl. Conf. Acoustics, Speech, and Signal Processing*, Vol. 5, pp. 2969–2972, Seattle, WA (May 1998).
- Q. Yuan, H. Yao, W. Gao, and S. Joo, "Blind watermarking method based on DWT middle frequency pair," *Proc. IEEE Intl. Conf. Multimedia Expo*, Vol. 2, pp. 473–476, Lausanne, Switzerland (August 2002).
- S.-H. Wang and Y.-P. Lin, "Wavelet tree quantization for copyright protection watermarking," *IEEE Trans. Image Process.* **13**(2), 154–165 (Feb. 2004).
- Y.-C. Li, T.-H. Wu, and Y.-C. Chen, "A scene adaptive hybrid video coding scheme based on the LOT," *IEEE Trans. Circuits Syst. Video Technol.* **8**(1), 92–103 (Feb. 1998).
- N. Jayant, J. Johnston, and R. Safranek, "Signal compression based on models of human perception," *Proc. IEEE* **81**(10), 1385–1422 (Oct. 1993).
- S. H. Tan, K. K. Pang, and K. N. Ngan, "Classified perceptual coding with adaptive quantization," *IEEE Trans. Circuits Syst. Video Technol.* **6**(4), 375–388 (Aug. 1996).
- G. K. Wallace, "The JPEG still picture compression standard," *IEEE Trans. Consum. Electron.* **38**(1), xviii–xxxiv (Feb. 1992).
- A. N. Netravali and B. G. Haskell, *Digital Pictures: Representation, Compression and Standards: Edition 2*, Plenum, New York (1995).



**Yuxin Liu** received her BS (with highest honors) and her MS from Tsinghua University, Beijing, China, and her PhD in 2004 from Purdue University, West Lafayette, Indiana, all in electrical engineering. She was with Multimedia Technologies Laboratory, Nokia Research Center (NRC), Nokia Inc., Irving, Texas as a research scientist from August 2004 to December 2005 and is currently with Sun Labs, Sun Microsystems Inc., Menlo Park, California, as a research staff engineer. During the summers of 2001 and 2002, she worked at Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey, as a summer intern. Her current research interests include image and video compression, information theory, and multimedia communications. Dr. Liu is a member of the IEEE and a member of Eta Kappa Nu.



**Bin Ni** received the BEng degree in electronic engineering from Tsinghua University, Beijing, China, in 2000 and the PhD in electrical engineering from Purdue University, West Lafayette, Indiana, in 2005. He has been a research assistant in the Spread Spectrum and Satellite Communications Research Laboratory (S3CRL) at Purdue University. He is now with Marvell Semiconductor as a senior design engineer. His research interests are in spread

spectrum communication, optical communication, storage systems, and digital signal processing.



**Xiaojun Feng** received the BS degree from Tsinghua University, China, in 2000, and the MS degree from Purdue University, West Lafayette, Indiana, in 2002. She is currently a PhD student in the Department of Electrical and Computer Engineering, Purdue University. Her research interest mainly lies in document image processing and image quality assessment.



**Edward J. Delp** received the BSEE (cum laude) and MS degrees from the University of Cincinnati, and the PhD degree from Purdue University. In May 2002 he received an Honorary Doctor of Technology from the Tampere University of Technology in Tampere, Finland. From 1980–1984, Dr. Delp was with the Department of Electrical and Computer Engineering, University of Michigan, Ann Arbor, Michigan. Since August 1984, he has been with the School of

Electrical and Computer Engineering and the Department of Bio-

medical Engineering at Purdue University, West Lafayette, Indiana. In 2002 he received a chaired professorship and currently is The Silicon Valley Professor of Electrical and Computer Engineering and Professor of Biomedical Engineering. His research interests include image and video compression, multimedia security, medical imaging, multimedia systems, and communication and information theory. Dr. Delp has also consulted for various companies and government agencies in the areas of signal, image, and video processing, pattern recognition, and secure communications. He has published and presented more than 280 papers. Dr. Delp is a Fellow of the IEEE, a Fellow of the SPIE, a Fellow of the Society for Imaging Science and Technology (IS&T), and a Fellow of the American Institute of Medical and Biological Engineering. In 2004 he received the Technical Achievement Award from the IEEE Signal Processing Society for his work in image and video compression and multimedia security. He is a member of Tau Beta Pi, Eta Kappa Nu, Phi Kappa Phi, Sigma Xi, and ACM. Dr. Delp is a registered professional engineer.