A Watermarking Scheme Based on The Characteristic of Addition Among DCT Coefficients

Minoru Kuribayashi[1] and Hatsukazu Tanaka[2]

[1] Division of Electrical and Electronics Engineering
Graduate School of Science and Technology, Kobe University
1-1 Rokkodai-cho, Nada-ku, Kobe, Japan 657-8501
minoru@es3.eedept.kobe-u.ac.jp
[2] Department of Electrical and Electronics Engineering
Faculty of Engineering, Kobe University
1-1 Rokkodai-cho, Nada-ku, Kobe, Japan 657-8501
tanaka@eedept.kobe-u.ac.jp

Abstract. Generally, low frequency domain may be useful to embed a watermark in an image. However, if a watermark is embedded into low frequency components, blocking effects may occur in the image. Then considering blocking effects, we study some characteristics among DCT coefficients and find some interesting mutual relations. Here, the robustness of many schemes based on the orthogonal transformations such as DCT may be doubtful against geometric transformations. For the distortions produced by some geometric transformations, we propose a searching protocol to find the watermarked block which is rotated and shifted. In the proposed scheme, a watermark can remain in the attacked image with very high probability in our scheme. Further, the watermark becomes more robust than the above scheme using error-correction.

1 Introduction

According to the spread of the internet, multi-media become to treat digital contents which can be copied easily without any degradation. It produces a very serious problem to protect the copyright of digital contents. Watermarking is one of the effective schemes to protect the copyright of digital contents. The watermarking is a technique to embed some information in the digital contents without being perceived. The embedded information can be extracted from the watermarked content by a tool. A watermark should convey information as much as possible, and it should be secret, which means that only authorized parties can access. Further more, it should not be removed from the original contents even if the original data were changed by signal processing such as data compression, hostile attack, etc. However, if an individual user knows how to embed and extract, unauthorized parties may forge the embedded information. Therefore, we establish a trusted center which knows the way of embedding and extracting in order to keep the copyright of every digital content. Every author who wants

to get the copyright must send his original contents to the trusted center and get the watermarked contents from it.

A watermark signal is sometimes designed in the spatial domain[1], but most scheme in the transformed domains such as DCT(Discreet Cosine Transform) domain[2][3][4] or wavelet domain[5][6] because the signal embedded into the transformed domain spreads all over the image. In transformed domain, a watermark is usually embedded into the high frequency components, as the changes of the components may be hard to perceive for man. Here, as mentioned in [7], it is difficult to perceive the changes of the very low frequency components as well as high. And the low frequency components may be useful to embed a watermark. However, many authors avoid to apply the low frequency components of DCT. Since blocking effects might be appeared in the image if the low frequency components of DCT would be changed. The blocking effects are usually noticeable because the boundaries of the blocks are appeared in the image.

In this paper, we propose a new watermarking scheme and apply the StirMark attack[8][9] to evaluate it, as the attack is used for evaluating the security of many watermarking schemes today. Using the characteristic of addition among DCT coefficients, a watermark is embedded into the very low frequency components of DCT without the blocking effects in the image. And considering the distortions caused by the geometric transformations, the distorted blocks where the watermark is embedded are found out in the searching protocol to synchronize the orthogonal axes and the positions. Further, encoding to the error correcting code, the watermark can become to withstand the effects caused by the attacks.

2 StirMark Attack

A watermark should retain the important feature such that it should not be deleted even if many kinds of signal processing, such as linear or nonlinear filtering, non-invertible compression, addition of noise, clipping, etc, were performed. In previous works, there are many watermarking schemes[10][11][12] to immunize such attacks. However, the specific distortions such as rotation, extension and reduction have often been neglected in spite of having the important feature such that they generate the serious degradation in PSNR(Peak Signal to Noise Ratio) being compared with only a little change of the visual characteristic of the image. Many watermarking schemes have been evaluated only by their own attacks and no standard tool to attack them has been proposed. Then StirMark attack, which performs some defaults attacks, has been proposed in order to standardize the attacks for watermarking. In the StirMark attack, an image is rotated, stretched, shifted and sheared by an unnoticeable amount. The main purpose is to give a short overview of the performances of watermarking systems and provide a standard tool for comparing them. Of course, some schemes withstanding StirMark attack have been proposed, but some problems are still remained in them. For example, the amount of embedded information is very small[13]. Then other attacks are added to a new version of StirMark attack

published in April, 1999. In this paper, we evaluate our scheme using the new tool of version 3.1[9], which includes low pass filtering , color quantization, JPEG compression, scaling, clipping, rotation, shearing, horizontal flip, removal of lines and columns, FMLR(Frequency Mode Laplacian Removal) attack, etc[8].

3 Proposed Scheme

In this section we consider how to embed watermark information and how to extract it. Taking into account of the distortions produced by embedding and attacks, we have proposed a new idea such that a watermark information bit embedded in a block is extracted from its inner sub-block.

3.1 Basic Idea

Generally, the set of basic orthogonal vectors has some interesting features. Then we try to analyze the characteristic of DCT. First, we begin with one dimensional DCT(1D-DCT). Fig.1(a) shows four low frequency basic vectors from $\boldsymbol{a_0}$ to $\boldsymbol{a_3}$. When two basic vectors are interacted ingeniously, an amplitude in the central region greatly increases and the waveform becomes similar to one of the basic vectors. Here we show the waveform of $\boldsymbol{a_0 - a_2}$ in the top of Fig.1(b). The waveform surrounded by a bold line is similar to a half scaled waveform of $\boldsymbol{a_0}$. Similarly the waveform of $\boldsymbol{a_1 - a_3}$ is similar to that of $\boldsymbol{a_1}$. This idea can be extended to

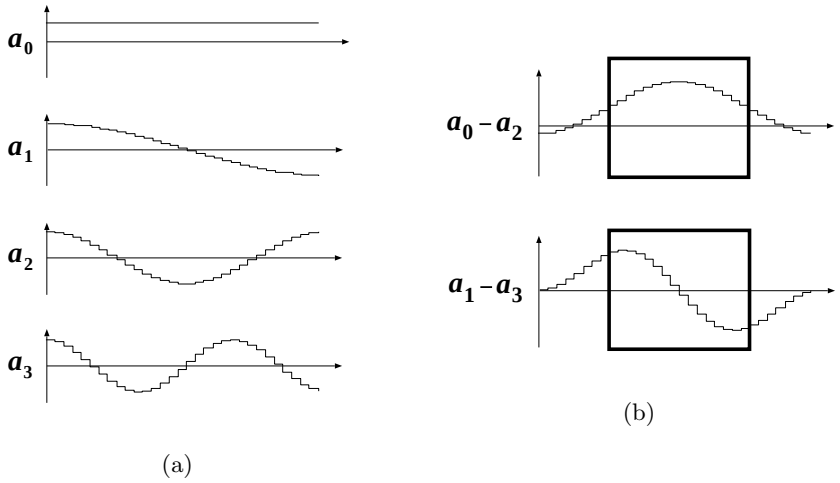


Fig. 1. Low frequency 1D-DCT basic vectors and those additive performance

two dimensional DCT(2D-DCT). Fig.2(a) shows 16 low frequency basic matrices from $\boldsymbol{A(0,0)}$ to $\boldsymbol{A(3,3)}$. When four basic matrices are interacted ingeniously, the similar phenomenon to the case of 1D-DCT is appeared. The calculated result

of $\{\boldsymbol{A(0,0)}-\boldsymbol{A(0,2)}-\boldsymbol{A(2,0)}+\boldsymbol{A(2,2)}\}$ has been shown in the top of Fig.2(b). Then the region in a oblique line is similar to a quarter scaled matrix of $\boldsymbol{A(0,0)}$. Similarly the region of $\{\boldsymbol{A(0,1)} - \boldsymbol{A(0,3)} - \boldsymbol{A(2,1)} + \boldsymbol{A(2,3)}\}$ is similar to that of $\boldsymbol{A(0,1)}$, the region of $\{\boldsymbol{A(1,0)}-\boldsymbol{A(1,2)}-\boldsymbol{A(3,0)}+\boldsymbol{A(3,2)}\}$ is similar to that of $\boldsymbol{A(1,0)}$ and the region of $\{\boldsymbol{A(1,1)} - \boldsymbol{A(1,3)} - \boldsymbol{A(3,1)} + \boldsymbol{A(3,3)}\}$ is similar to that of $\boldsymbol{A(1,1)}$. Further more, the amplitude outside of the oblique line becomes very small.

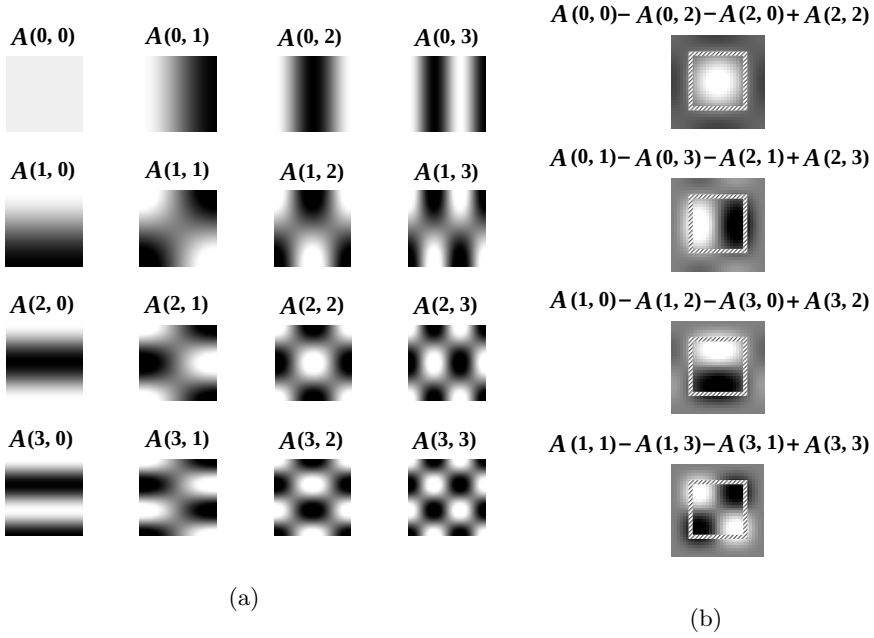


Fig. 2. Low frequency 2D-DCT basic matrices and those additive performance

Applying the feature for our scheme, we embed watermark information bit in a 32×32 block, and extract it from the inner 16×16 sub-block which exists in the middle of 32×32 block. When watermark information is embedded, the information bit is added to only four special DCT coefficients in a 32×32 block. Then the energy given to the four coefficients are spread over the block when IDCT is performed. However, when the inner 16×16 sub-block is transformed by DCT, the almost all spread energy is concentrated into only one DCT coefficient. Therefore, the embedded watermark information bit can be extracted from the sub-block. The four special DCT coefficients for embedding are specified as four columns p_1, p_2, p_3 and p_4 in Table.1, and then the coefficient for extracting is the special DCT coefficient given in the column P, where $F_{32}(*,*)$ means a DCT coefficient of 32×32 block, and $F_{16}(*,*)$ is defined similarly.

Table 1. Embedding and extracting coefficients

	embedding				extracting
set	p_1	p_2	p_3	p_4	P
i	$F_{32}(0,0)$	$F_{32}(0,2)$	$F_{32}(2,0)$	$F_{32}(2,2)$	$F_{16}(0,0)$
$\ddot{u}$	$F_{32}(1,0)$	$F_{32}(1,2)$	$F_{32}(3,0)$	$F_{32}(3,2)$	$F_{16}(1,0)$
$\ddot{\ddot{m}}$	$F_{32}(0,1)$	$F_{32}(0,3)$	$F_{32}(2,1)$	$F_{32}(2,3)$	$F_{16}(0,1)$
$\dot{w}$	$F_{32}(1,1)$	$F_{32}(1,3)$	$F_{32}(3,1)$	$F_{32}(3,3)$	$F_{16}(1,1)$

3.2 Embedding

Let I be an image, $\boldsymbol{w} = (w_0, w_1, \cdots, w_{n-1})$, $w_i = 1$ or 0, be a watermark information vector of which size is n, and m be an embedding intensity. Then watermark information $\boldsymbol{w}$ is embedded as follows.

1. I is divided into blocks of 32×32 pixel and each block is transformed by DCT.
2. The embedding coefficients in each block are first determined by selecting a set using a secret key from four sets given as raws in Table.1, and then the coefficients to be embedded are given by the elements of four columns p_1, p_2, p_3 and p_4.
3. The element w_t in $\boldsymbol{w}$ is embedded in the t-th block as follows.

   If $w_t = 0$ then

   $$p_1 = p_1 + m, \qquad p_2 = p_2 - m,$$

   $$p_3 = p_3 - m, \qquad p_4 = p_4 + m.$$

   else then

   $$p_1 = p_1 - m, \qquad p_2 = p_2 + m,$$

   $$p_3 = p_3 + m, \qquad p_4 = p_4 - m.$$

4. Each block is transformed by IDCT and the watermarked image I' which consists of the transformed blocks can be recovered.

3.3 Searching Protocol

A watermark should be extracted from the embedded image even if the embedded block were shifted and rotated. To search for the amount of shift and rotation, we calculate the MSE(Mean Square Error) in each shifted and rotated block, and estimate the most possible position where the MSE becomes minimum. Then, if the above procedure were performed with every considerable distortion caused by attacks such as shift and rotation, the computational complexity might become incredibly high. Therefore, we define the domain to search for and the candidates of distorted shapes. The block which may be shifted and rotated is assumed to

be in the searching domain block K of $(16 + 2d) \times (16 + 2d)$ shown in Fig.3, where d is called the searching distance, B is the block to be embedded and L is the sub-block. Then, the MSE is evaluated in each shape $T_s (0 \leq s \leq 12)$ given in Fig.4, where each shape is obtained by rotating the sub-block slightly. As an illegal image must be copied from the watermarked image, it is more efficient to use the watermarked image in order to search for the shifted and rotated block than to use the original one. But we need not to store both images, and it is enough to preserve only the original image because the watermarked image can be obtained by embedding the watermark using the secret key when the search is performed.

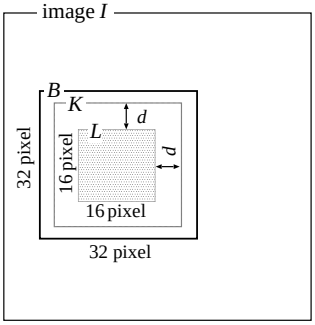


Fig. 3. Searching domain

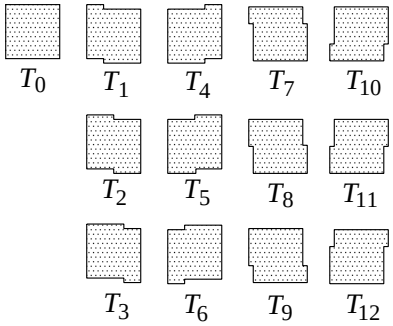


Fig. 4. The 13 candidates of rotated sub-block

Let I^* be the image which may be copied illegally from I', and each rotated and shifted block is searched as follows.

1. First, the following operations are performed.
   1.1. The block of the shape T_0 in Fig.4 is picked out from the left upper side of the searching domain block K^*. Then, the MSE in the block is calculated, and the MSE, shape and its position are preserved.
   1.2. Next, the block of the shape T_0 is picked out from the position shifted one pixel to the right or down and the MSE is calculated. Here, if the MSE in the block is less than that of the former, the MSE, shape and its position preserved before are changed to this new ones.
   1.3. The operation 1.2 is performed at all possible position in the searching domain block K^*.
2. The operation 1 is continued from the shapes T_1 to T_{12} repeatedly. Here in the operation 1.1, the MSE, shape and the position are changed when the MSE is less than the preserved one.
3. Finally, the position and shape where the MSE becomes minimum are selected and the rotated and shifted block is reformed to the square shape $\hat{L^*}$ of 16×16 in each block.

3.4 Extracting

The original image I is necessary to extract the watermark from the image I^*, because the embedded information can be extracted by subtracting the specified DCT coefficient in Table.1 of $\hat{L^*}$ from that of L. The procedure is given as follows.

1. The searching protocol is applied for each sub-block and $\hat{L_t^*}(0 \leq t \leq n-1)$ is obtained.
2. Each $\hat{L_t^*}$ and L_t are transformed by DCT.
3. A coefficient in the set of four elements P given in Table.1 is specified by the secret key.
4. The coefficient of $\hat{L_t^*}$ specified above is subtracted from that of L_t.
5. If the result is positive, then an extracted information $w_t^* = 0$, else, $w_t^* = 1$.

3.5 Improvement of Robustness

Generally a watermark must provide enough robustness even if any attacks might be performed. Our watermark has a strong tolerance for attacks, but it can not immunize a few attacks such as rotation, shift, etc. One of the reasons is that the shapes $T_s(0 \leq s \leq 12)$ do not meet every distortion caused by rotation. Of course, it is desirable to consider all possible rotated and shifted patterns. However, if we would evaluate the MSE with more patterns, it might take more computation time and need more memory in order to estimate all possible rotated and shifted blocks. To certify it, we have already tested the other patterns, but the improvement is only slight in spite of the great increase of the computational complexity. Further more, as many kinds of attacks are included in StirMark attack, the watermark information bits may be changed by some of them. Therefore, we must encode the watermark information bits by an error correcting code in order to improve the robustness against their attacks.

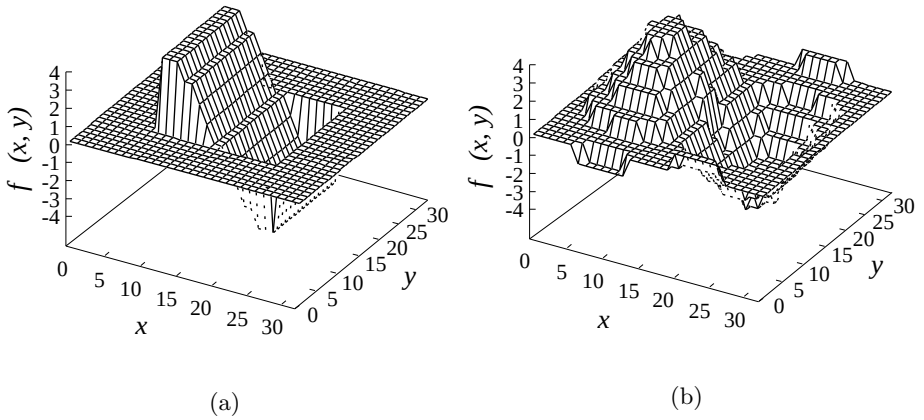
4 Consideration

DCT coefficients have the general property such that the change of the low frequency components causes the blocking effects to the image. For example, Fig.5(a) shows the blocking effects when the value of $F_{16}(1, 0)$ is changed, where $f(x, y)$ means the pixel intensity. However, considering the effects caused by the change of some low frequency components, the shape of blocking effects can be changed by the attentive modification of the special coefficients.

The suitable selection and modification of the coefficients can cause special effects that are hardly deleted by the attacks such as filterings. The selection means to use the special four coefficients in the columns p_1, p_2, p_3 and p_4 in Table.1. And the modification means that a certain constant value is added to the coefficients of p_1 and p_4, and subtracted from the coefficients of p_2 and p_3. Then, when the procedure given in the basic idea is applied, the energy of the four coefficients is concentrated into the specified DCT coefficient P in the 16×16 sub-block. For example, if we select the set $\ddot{u}$ and a certain value is

added to $F_{32}(1,0)$ and $F_{32}(3,2)$, and subtracted from $F_{32}(1,2)$ and $F_{32}(3,0)$, the energy of their coefficients concentrates into the DCT coefficient $F_{16}(1,0)$ in the sub-block. The result in the pixel domain is shown in Fig.5(b), and the result in the transformed domain is shown in Fig.6.

Hence, the distortions produced by embedding are greatly different from the blocking effects, and it is less noticeable as no edge is formed. In addition, the embedded signal is emphasized in the DCT coefficient $F_{16}(1,0)$, and there is no correlation among adjoining 32×32 blocks. Further more, some attacks such as extension and reduction do not cause a serious problem to the watermark, as the embedded signal is spread over outside of the sub-block.



(a) (b)

Fig. 5. Distortions in a block

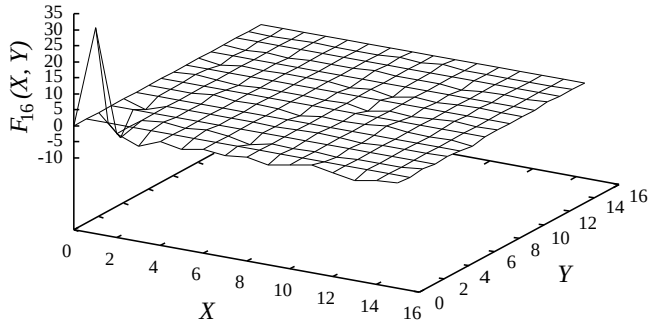


Fig. 6. Concentration of energy into the DCT coefficient $F_{16}(X,Y)$

5 Computer Simulated Results

In our simulation, we use a standard image "lenna" and standard images "girl", "baboon", "couple", "aerial", "moon" and "map" in SIDBA(Standard Image Data Base), each of which has 256 level gray scale with the size of 256×256. Then the maximum number of watermark information bits is 64 bits.

First, we show some results for "lenna" of which original image is shown in Fig.7. Fig.8 is produced by embedding a watermark into the original image, where embedding intensity $m = 30$, PSNR = 42.5 [dB], and Fig.9 is an image produced by applying the well-known StirMark attack to the watermarked image. It has no obvious visually difference from the original watermarked im-

Fig. 7. Original image

age though its PSNR decreases to 18.4[dB]. Here, we need not to say that the embedded watermark can be extracted also from Fig.9.

Fig.10 shows the average value of PSNR respect to embedding intensity m. Here, if m is set over 30, the distortions become to be remarkable in the watermarked image as shown in Fig.11. And if m is set under 15, the watermark is deleted easily by the StirMark attack. Therefore, the suitable range of m is between 15 and 30 for the image "lenna". For many images, our simulated results show that the suitable range of m is also from 15 to 30 though the distortions in the watermarked image can be blurred even if m is set greater than 30 for the

Fig. 8. Watermarked image ($m = 30$)

Fig. 9. StirMark attacked image

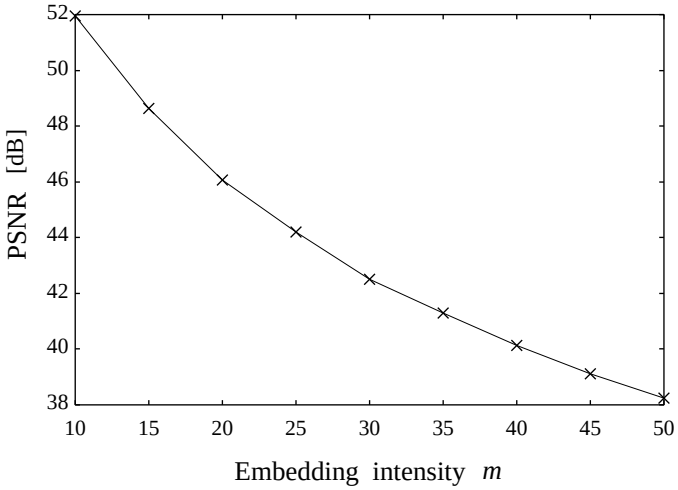


Fig. 10. PSNR versus embedding intensity m

Fig. 11. Watermarked image ($m = 50$)

image which includes many edges. Then we use the range from 15 to 30 in the following computer simulation.

If the StirMark attack which parameters are defaults is performed, a very few errors may be occurred in the extracted watermark information. Then, the number of errors may be less than or equal to 3. In such a case, the watermark information can be extracted correctly by use of an error correcting code. Then we will perform the computer simulation under the assumption that the watermark information has already been encoded by some error correcting codes , and hence, the extracted information has been corrected if the number of errors is less than or equal to the error correcting ability. Fig.12 shows the probability of the correct extraction for the searching distance $d(4 \leq d \leq 8)$ and the embedding intensity $m = 20$. The best results is obtained when the searching distance

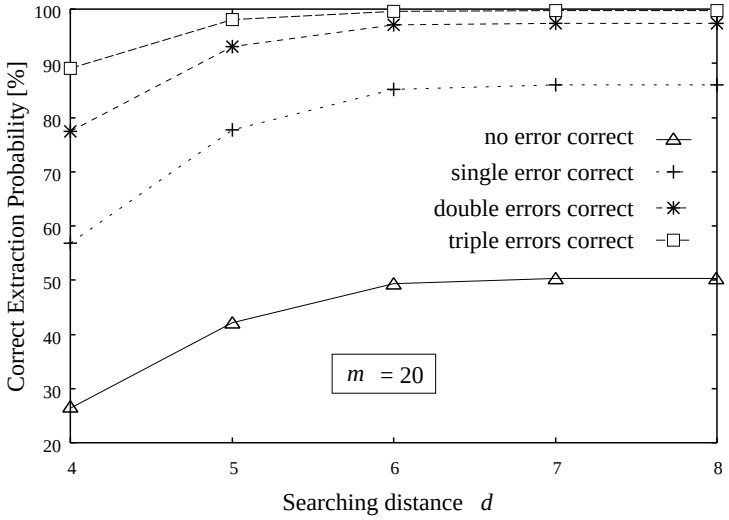


Fig. 12. Probability of correct extraction versus searching distance

$d = 8$. However, it is efficient to apply for $d = 6$ as the computational complexity is about half compared to the case of $d = 8$. Therefore we define the searching distance $d = 6$ in the following simulation.

Next, we evaluate the tolerance of each watermarked image against the StirMark attack. Table.2 to 5 show the computer simulated results of the correct extraction probability for each different error correcting abilities, when the number of simulated times is 10^5. The watermark is generated and embedded randomly in each time. The results mean that the watermark information can be extracted almost correctly from the image distorted by the StirMark attack when the embedding intensity is set to $m = 30$ and some triple errors correcting codes are applied. Here, if soft decision decoding is applied, the probability of correct

Table 2. Correct extraction probability[%] for the case of no error correction

m	lenna	girl	baboon	couple	aerial	moon	map
15	14.9	19.5	9.0	23.7	0.3	44.5	0.6
20	49.4	51.3	37.4	67.9	7.7	83.3	4.6
25	73.7	67.8	63.6	89.5	28.9	95.3	12.0
30	89.0	78.5	84.0	97.6	58.2	99.2	22.9

Table 3. Correct extraction probability[%] for single error correction

m	lenna	girl	baboon	couple	aerial	moon	map
15	44.5	53.7	31.2	57.7	2.0	80.3	4.2
20	85.2	88.1	74.9	94.4	28.1	98.6	20.3
25	96.4	95.8	92.6	99.4	65.4	99.9	39.6
30	99.4	98.3	98.5	100.0	90.0	100.0	59.2

Table 4. Correct extraction probability[%] for double errors correction.

m	lenna	girl	baboon	couple	aerial	moon	map
15	72.4	80.7	57.9	82.3	7.2	94.3	13.8
20	97.1	98.3	92.8	99.4	54.6	99.9	45.1
25	99.7	99.7	98.9	100.0	87.8	100.0	68.9
30	100.0	100.0	99.9	100.0	98.4	100.0	85.0

Table 5. Correct extraction probability[%] for triple errors correction

m	lenna	girl	baboon	couple	aerial	moon	map
15	89.3	93.9	79.0	94.1	17.5	98.4	30.1
20	99.6	99.8	98.4	100.0	76.4	100.0	69.5
25	100.0	100.0	99.9	100.0	96.7	100.0	87.9
30	100.0	100.0	100.0	100.0	99.8	100.0	96.1

extraction can be increased. For the images such as "aerial" and "map", the probability of the correct extraction becomes smaller than the others, because above two images include a lot of vulnerable edges by attacks. Then, for those two images, the embedding intensity can be set greater than 30, and the large distortions caused by the embedding will be blurred because of containing many edges. Therefore it is more preferable to set the suitable embedding intensity m for each image considering how many edges are included in the image.

Finally, our scheme can withstand not only StirMark attack, but also the well-known unZign attack[14], JPEG compression which quality parameter can be set under 25[%] and most of image processing.

6 Conclusion

We have proposed a new watermarking scheme in the tolerance of StirMark attack, and it has three important points. First, a watermark is embedded into the very low frequency components of DCT, then the blocking effects will not be appeared in the image. Next, the watermark embedded in a block is extracted from its inner sub-block. Finally, the amount of rotation and shift can be found in the searching protocol when the watermark is extracted.

In our scheme, the watermark can be extracted correctly with very high probability though the perceptual distortions caused by embedding are very few. Our simulation results mean that the strength of the watermark depends on the image and hence the suitable setting of parameters is inevitable.

References

1. I. Echizen, H. Yoshiura, T. Arai, H. Kimura and T. Takeuchi, "General quality maintenance module for motion picture watermarking" *IEEE Transaction on Consumer Electronics*, Vol.45, No.4, pp.1150-1158, (Nov, 1999).
2. C. F. Wu and W. S. Hsieh, "Digital watermarking using zerotree of DCT" *IEEE Transaction on Consumer Electronics*, Vol.46, No,1, pp.87-94, (Feb, 2000).
3. T. Mori and H. Imai, "The distribution of DCT coefficients to embed digital watermarks" *SCIS98*-3.2.E, (Jan, 1998).
4. S. Lee and H. Imai, "A new method of selecting blocks for the digital watermark using DCT" *SCIS98*-3.2.D, (Jan, 1998).
5. N. Nakanishi and H. Tanaka, "Watermark of images having tolerance for clipping attack" *SITA99*, pp.613-616(Dec, 1999).
6. M. Iwata and A. Shiozaki, "Index data embedding method for color images utilizing features of wavelet transform" *SITA99*, pp.181-184 (Dec, 1999).
7. H. Ishizuka, Y. Sakai and K. Sakurai, "A new digital watermarking method based on visual illusions" *SITA97*, pp.65-68, (Dec, 1997).
8. Fabien A. P. Petitcolas and Ross J. Anderson, "Evaluation of copyright marking systems" *Proc. IEEE Multimedia Systems'99*, Vol.1, pp.574-579, (Jun, 1999).
9. Fabien A. P. Petitcolas, R. J. Anderson and M. G. Kuhn, StirMark3.1 `http://www.cl.cam.ac.uk/~fapp2/watermarking/stirmark/`, (Apr. 1999).

10. I. J. Cox, J. Kilian, F. T. Leigthton and T. Shamoon, "Secure spread spectrum watermarking for multimedia" *Proc. IEEE Transactions on Image Processing*, Vol.6, No.12, pp.1673-1687, (Dec, 1997).
11. M. D. Swanson, M. Kobayashi and A. H. Tewlfik, "Multimedia data-embedding and watermarking technologies," *Proc. IEEE (Special Issue on Multimedia Signal Processing)*, Vol.86, No.6, pp.1064-1087, (Jun, 1998).
12. F. Hartung and M. Kutter, "Multimedia watermarking techniques" *Proc. IEEE*, Vol.87, No.7, pp.1079-1107, (Jul, 1999).
13. J. Tanimoto and A. Shiozaki, "A digital image watermarking scheme withstanding StirMark attack" *SCIS99*, pp.217-222, (Jan, 1999).
14. UnZign watermark removal software, `http://www.altern.org/watermark/`