Document Protection Based on CHT Coefficients' Phase Modulation

Ivo R. Draganov[1] and Roumen K. Kountchev[2]

***Abstract* – In this paper we present a robust and efficient algorithm for document protection based on digital watermarking in phase domain using all complex coefficients of the two dimensional complex Hadamard transform (2D-CHT). High quality reconstruction and low visibility for the watermark is provided. Also broaden information capacity is guaranteed - up to ¾ of the number of pixels for the original image in bits. Detecting the watermark can be done only on the basis of a simple private/public key.**

***Keywords* – digital watermarking, document protection, complex Hadamard transform, phase modulation**

I. INTRODUCTION

Document protection is an important stage in many processing, transmitting and archival information systems – to detect unauthorized interference over electronic documents or usage without appropriate copyrights is of main importance here [1]. Digital watermarking provides the means for such protection – to prove copyright at one hand and to discover unwanted manipulation over the documents at the other.

Digital watermarking algorithms may be classified in three main groups as for the working domain in which some digital resource is being marked – space, frequency and phase [2]. Previous research suggest that watermarks introduced in space and frequency (amplitude alteration) often are robust enough but their visibility is high or the opposite – aiming low visibility leads to easy deletion of the watermark.

Our previous researches [3, 4] suggest that this contradiction may be suppressed by altering the phase spectrum of an image (or a sound signal). Most manipulations over images (attacks) affect least phases of the harmonics and as addition human eye (ear) is less sensitive to phase changes than to amplitude ones and the watermark is less visible.

To find appropriate phase spectrum a transform should be selected. We consider 2D-CHT as a proper choice since detailed examination of it had been done [5]. Initially we considered the possibility to mark only phases of a harmonics with amplitudes over a preliminary defined threshold but synchronism is easily lost when extracting the watermark so here enhanced form of the algorithm is presented using all 2D-CHT coefficients.

In part two we give detailed algorithm description along with in-depth analysis, in part three – some experimental results are given and then a conclusion is made in part four.

[1]Ivo R. Draganov is with the Faculty of Telecommunications, Technical University, Kliment Ohridski 8, 1000 Sofia, Bulgaria, E-mail: idraganov@abv.bg

[2]Roumen K. Kountchev is with the Faculty of Telecommunications, Technical University, Kliment Ohridski 8, 1000 Sofia, Bulgaria, E-mail: rkountch@tu-sofia.bg

II. ALGORITHM DESCRIPTION AND ANALYSIS

Inserting the watermark contains the following steps.

Step 1 – getting the input image, in color for the most general case in RGB color space, [$R(i,k)$, $G(i,k)$, $B(i,k)$], where $R,G,B = 0 \div 255$; $i = 0 \div M'-1$ and $k = 0 \div N'-1$. Conversion is made then to grayscale (or used by default):

$$I(i,k) = 0.30R(i,k) + 0.59G(i,k) + 0.11B(i,k), \qquad (1)$$

where $I(i,k) = 0 \div 255$.

Step 2 – extension the original image to size $M'' \times N''$ pixels ($M'' = w_M.2^n$, $M'' = w_N.2^n$, w_M, w_N – any positive whole number, $n \geq 2$, $M'' \geq M'$, $N'' \geq N'$, by symmetric repetition).

Step 3 – dividing the image into blocks with size NxN pixels where $N = 2^n$.

Step 4 – calculating the coefficients $s(u,v)$ for each block using 2D-CHT [5]:

$$s(u,v) = \sum_{i=0}^{N-1}\sum_{k=0}^{N-1} L(i,k) e^{-j\frac{\pi}{2}(ui+vk)} t(u,i)t(v,k), \qquad (2)$$

where $t(p,q) = \begin{cases} 1, & for\ n = 2; \\ (-1)^{\sum_{r=3}^{n}\left\lfloor \frac{p}{2^{r-1}} \right\rfloor \left\lfloor \frac{q}{2^{r-1}} \right\rfloor}, & for\ n = 3,4,... \end{cases}$ is a sign function, $\left\lfloor \frac{a}{b} \right\rfloor$ - operator for extracting the whole part of a quotient, u, v = 0, 1, …, N-1, $j = \sqrt{-1}$, p, q = 0, 1, …, 2^n-1, $n = \log_2 N$.

Step 5 – calculating the modules and phases of the spectral coefficients:

$$M(u,v) = \sqrt{[s_R(u,v)]^2 + [s_I(u,v)]^2}, \quad \varphi(u,v) = -arctg\left[\frac{s_I(u,v)}{s_R(u,v)}\right], \qquad (3)$$

where

$$s_R(u,v) = \sum_{i=0}^{N-1}\sum_{k=0}^{N-1} L(i,k)t(u,i)t(v,k)\cos[\tfrac{\pi}{2}(ui+vk)], \qquad (4)$$

$$s_I(u,v) = \sum_{i=0}^{N-1}\sum_{k=0}^{N-1} L(i,k)t(u,i)t(v,k)\sin[\tfrac{\pi}{2}(ui+vk)]. \qquad (5)$$

The conducted research on the distribution of φ (only for CHT complex coefficients) over the whole image reveals significant difference between natural scene and document images. This is confirmed by the presented averaged histogram $h(\varphi)$ in Fig.1.a for document images and in Fig.1.b – for natural scene images.

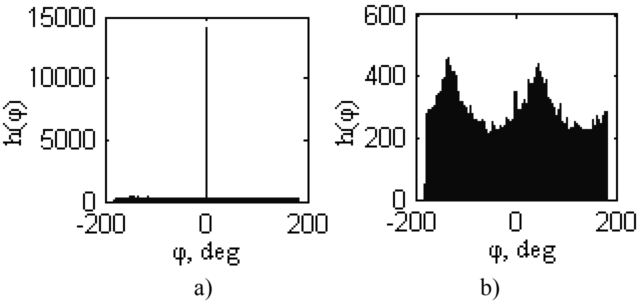


Fig.1. Phase histograms $h(\varphi)$ for different types of images a) document images b) natural images

Firstly all the distributions for document images appear very close – a huge peak for $\varphi = 0°$ persists and almost even spread for all other phases in the interval [-180°,+180°] with some small compared to the central peak maximums for $\varphi = -135°$ and 45°. As for the natural scene images nevertheless of their different content their distributions are also very close – the central peak here is absent, but clearly visible maximums for $\varphi = -135°$ and 45° are observed. These results can be explained by investigating (2) for the basic case of $n = 2$. The Hadamard coefficients (R_i for real – ¼$M''xN''$ in number and (C_i, C_i^*) for complex and conjugated complex ones – totally 3/4$M''xN''$) obtained from a block of 4x4 pixels are given in Fig.2.

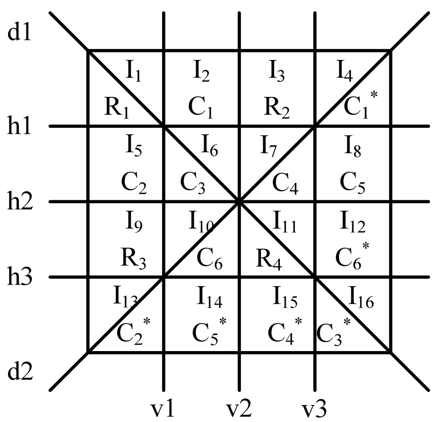


Fig.2. Complex Hadamard coefficients' location for block 4x4

The phase of the complex coefficient C_1 corresponding to $s(1,0)$ is calculated from (2)-(5):

$$\varphi(C_1)=\begin{cases} 0 & ,\ if \quad \sum\limits_{s=2,6,10,14} I_s = \sum\limits_{s=4,8,12,16} I_s \ and \ \sum\limits_{s=1,5,9,13} I_s = \sum\limits_{s=3,7,11,15} I_s \\ arctg\left(\sum\limits_{s=2,6,10,14} I_s - \sum\limits_{s=4,8,12,16} I_s \Big/ \sum\limits_{s=1,5,9,13} I_s - \sum\limits_{s=3,7,11,15} I_s \right) & -\ otherwise \end{cases} \quad (6)$$

It is obvious that for a homogenous block $\varphi(C_1) = 0°$, if a vertical transition $v1$ persists $\varphi(C_1) \approx -90°$ or 90° depending on the direction of the brightness gradient, for $v2$ – $\varphi(C_1) \approx -135°$ or 45°, for $v3$ – $\varphi(C_1) \approx 0°$ or 180°; for all horizontal transitions $h1$, $h2$, $h3$ $\varphi(C_1) = 0°$; for the diagonal ones $d1$ and $d2$ – $\varphi(C_1) \approx -135°$ or 45°. The analysis for the others $\varphi(C_i)$, for $i = 1 \div 6$, and for $n > 2$ is analogous.

Now we can explain the results from Fig.1. For document images vast homogenous areas appear and thus a strong peak for $\varphi = 0°$ is obtained. For natural scene images there is no preferable kind of transition (statistically) and because the number of $v1$, $h1$-$h3$, $d1$-$d2$ transitions is dominant over $v1$ and $v3$ maximums for $\varphi = -135°$ and 45° are formed. The latter exist in document images as well but considerably smaller than the central peak.

Step 6 – inserting the current bit w_r of the watermark p into the current phase for the coefficients $s(u,v)$ and $s^*(u,v)$ according to:

$$\varphi_w(u,v) = -\varphi_w^*(u,v) = \begin{cases} \varphi(u,v) + \Delta, & if \quad w_r(p) = 1; \\ \varphi(u,v) - \Delta, & if \quad w_r(p) = 0 \end{cases}, \quad (7)$$

for $r = 1, 2, \ldots, R$, where R is the number of the binary elements $w_r(p)$ of the watermark p and $\varphi_w(u,v)$ and $\varphi_w^*(u,v)$ are the phases of the marked coefficients $s_w(u,v)$ и $s_w^*(u,v)$. Parameter Δ is an angle defining the depth of the watermark and thus its robustness and visibility respectively.

The sequence w_r is obtained using XOR operation over each bit of the watermark and a bit from a randomly generated sequence describing private or public key depending on the cryptographic algorithm used. This sequence should have autocorrelation function close to delta impulse for higher reliability of the extracted watermark. If the current complex coefficient to be marked by phase has zero amplitude, the respective binary value of the watermark is omitted and only the value of the binary random sequence is used without applying XOR. Thus we reduce the probability for introducing errors in the recovered watermark later because in the opposite case the watermark value becomes inextricable – zero amplitude means zero phase even after change. This is highly important for document images where much more complex coefficients are zeros due to the presence of large homogenous areas. Also it is highly important to be no complex coefficient missed in this step because we may get the lack of synchronism at the stage of detecting or recovering the watermark – an improvement made to our previously developed algorithm [3].

Step 7 – calculating the pixel values $L_w(i,k)$ of the current marked block using the inverse two dimensional complex Hadamard transform (2D-ICHT):

$$L_w(i,k) = \frac{1}{N^2} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} s_w(u,v) e^{j\frac{\pi}{2}(ui+vk)} t(u,i)\, t(v,k), \quad (8)$$

for $i, k = 0, 1, \ldots, N$-1, where $s_w(u,v) = M(u,v)e^{j\varphi_w(u,v)}$.

Extracting the watermark contains the following steps.

Step 1-5 – repeat the steps 1-5 from the insertion of the watermark over the watermarked image for estimating the coefficients $s(u,v)$.

Step 6 – checking for the presence of p-th watermark from D possible ones calculating the coefficient $C_{m,p}$ which defines the correlation between the p-th and the m-th watermark. p-th is the watermarked originally used for all the blocks of the input image. The cross-correlation coefficient is given by:

$$C_{m,p} = \sum_{r=1}^{R} [\varphi_r + \Delta_r(m)]\,\Delta_r(p) = A_p + B_{m,p}\,, \qquad (9)$$

for $m, p = 1, 2, \ldots, D$, where $[\varphi_r + \Delta_r(m)] = \varphi_{rw}(m)$ is the phase of the r-th marked coefficient $s_{rw}(u,v)$,

$$A_p = \sum_{r=1}^{R} \varphi_r \Delta_r(p), \;\; B_{m,p} = \sum_{r=1}^{R} \Delta_r(m)\Delta_r(p)\,, \qquad (10)$$

where φ_r is the absolute value of the phase for the r-th coefficient $s_r(u,v)$ before watermarking and:

$$\Delta_r(p) = (-1)^{w_r(p)}\Delta = \begin{cases} +\Delta & if \;\; w_r(p) = 1; \\ -\Delta & if \;\; w_r(p) = 0. \end{cases} \qquad (11)$$

Here $w_r(p)$ are the binary elements of the random sequence consisting of R bits describing the p-th ciphered watermark.

When R tends to larger values and supposing that $\varphi_r \approx$ const. (in a block 4x4 there may be only one transition present or a homogenous area which leads to this assumption) according to (10) follows:

$$A_p = \varphi_0 \sum_{r=1}^{R} \Delta_r(p) \approx 0\,. \qquad (12)$$

If all coefficients are non-marked then $C_{0,p} \equiv A_p \approx 0$. For the phase marked spectrum coefficients it is true that:

$$C_{m,p} \approx \begin{cases} \sum_{r=1}^{R} [\Delta_r(m)]^2 = R\Delta^2 \;\; if \; m = p; \\ \sum_{r=1}^{R} \Delta_r(m)\Delta_r(p) \approx 0 \;\; if \;\; m \neq p. \end{cases} \qquad (13)$$

Step 7 – making a decision for detecting p_0-th watermark according to the following rule:

$$p_0 = \begin{cases} detected, \;\; if \;\; (C_{m,p_0} / R\Delta^2) \geq \theta; \\ absent \qquad - \;\; otherwise, \end{cases} \qquad (14)$$

for m, $p_0 = 1, 2, \ldots, D$, where θ is preliminary chosen threshold in the interval $0 < \theta < 1$. The number of errors introduces (false alarms and missed watermark altogether) depend on θ's value.

The extraction of the watermark itself is done by using the original image. It is assumed that the copyright holder is only authorized to do this and thus has access to it. After finding the phase spectrums for the original and the watermarked image a simple subtraction of the respective phases is sufficient to restore the random sequence which then is XOR-ed with the private key to get the watermark.

III. Experimental Results

As experimental dataset we use 10 natural scenes (8 bpp, 512x512 pixels, 72 dpi) and 10 document (containing text, equations, graphics and tables) grayscale images (8 bpp, 2550x3300 pixels, 300 dpi). The watermark is 100x100 pixels binary image.

First of all, the results for natural scene and document images appeared to be very close, so here averaged values are given. In Fig.3 peak signal-to-noise ratio (PSNR) is given between watermarked and original images. It is seen that for wide range of values for Δ the watermarked images preserve high quality not letting PSNR to drop below 35 dB. The size of the window used represented by n has no influence here.

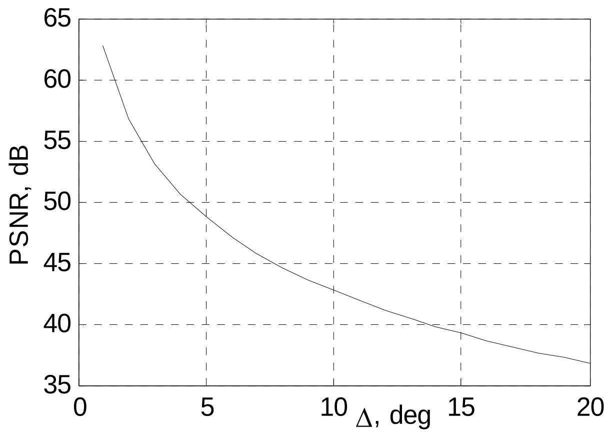


Fig.3. PSNR between watermarked and original images

In Fig.4 averaged results are given for the ratio between squared autocorrelation maximum when using the right private key and the average of the squared values when using other statistically independent random keys (200 in number) for the stage of watermark detection at different Δ. Even for small Δ around 5° this ratio is at least 2 which proves the robust detection for the algorithm proposed.

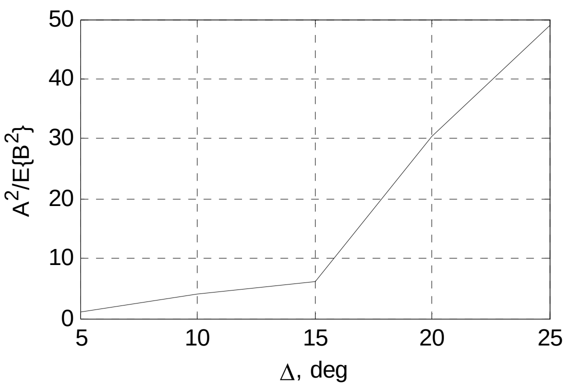


Fig.4. Autocorrelation maximum variation

In Fig.5 mean square error (MSE) is given between the original and extracted watermark when using typical value for Δ = 12° in the case of JPEG compressed (attacked, quality variation between 20 and 80 %) and non-compressed image (100% quality). Its almost constant value in the wide range of 20-80 % compression quality assures high resistance to high frequency attacks represented by the JPEG algorithm here.

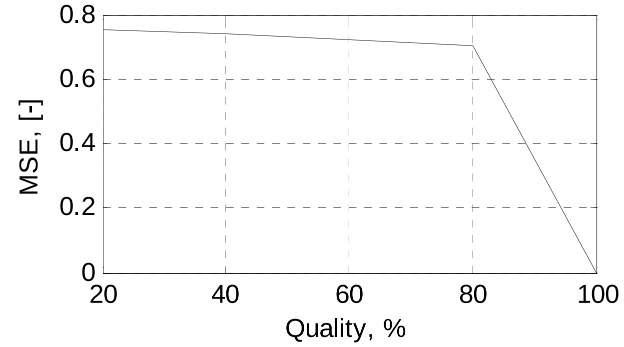
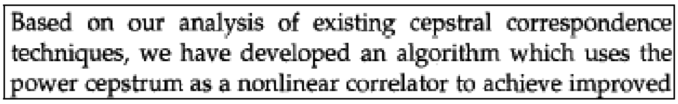
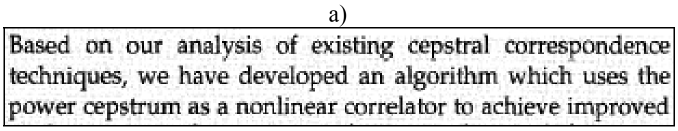


Fig.5 Mean square error for the extracted watermark at different quality levels for the JPEG compressed watermarked image

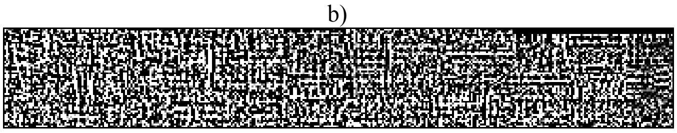
In Fig.6.a part of the original document containing printed text is given, and in Fig.6.b – the watermarked one for Δ = 12°. The amplified difference of 64 times between these two images is given in Fig.6.c. Slightly changed background is observed which is completely lost (invisible) if documents are stored in compressed form which is the most general case.


Based on our analysis of existing cepstral correspondence techniques, we have developed an algorithm which uses the power cepstrum as a nonlinear correlator to achieve improved

a)


Based on our analysis of existing cepstral correspondence techniques, we have developed an algorithm which uses the power cepstrum as a nonlinear correlator to achieve improved

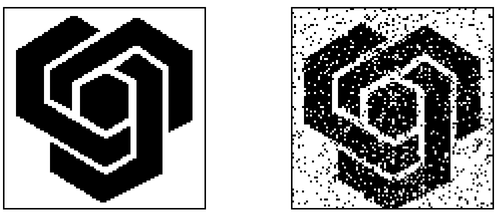
b)



c)

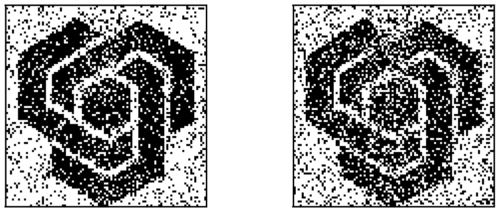
Fig.6. Original and watermarked printed text

In Fig.7.a the original watermark used is given, and in Fig.7.b - the extracted one from a JPEG compressed to 80 % quality image, Fig.7c – from 40 % and in Fig.7.d – from 20 %.



a) b)

Fig.7. Original and extracted watermark



c) d)

Fig.7. Original and extracted watermark (continuation)

IV. CONCLUSION

The proposed algorithm for document protection based on all 2D-CHT coefficients modulation is robust enough to high frequency filtration attack (JPEG compression) which is seen from the almost constant MSE for the extracted watermark. Because of the high PSNR for the watermarked image low visibility for the watermark is guaranteed. Also high capacity for the information inserted is present – ¾ of the total number of pixels in bits, and due to the use of all CHT complex coefficients synchronism is not lost when extracting the watermark. Detecting the presence of watermark is done without the original image. When applied to printed documents last preserve high quality.

ACKNOWLEDGEMENT

This paper was supported by the National Fund for Scientific Research of the Bulgarian Ministry of Education and Science (Contract - BY-И-305/2007).

REFERENCES

[1] S. Mao, A. Rosenfeld, and T. Kanungo, "Document Structure Analysis Algorithms: A Literature Survey", *In Proceedings of SPIE Electronic Imaging*, Vol. 5010, pp. 197-207, January 2003.

[2] J. Dugelay and S. Roche, "A Survey of Current Watermarking Techniques", *In Information Techniques for Steganography and Digital Watermarking*, S.C. Katzenbeisser et al., Eds. Northwood, MA Artec House, pp. 121-145, Dec. 1999.

[3] R. Kountchev and V. Mirchev, "Digital Watermarking Using Complex Hadamard Transform and Phase Modulation", *In ICEST'2005 Conference Proceedings*, Nis, Serbia and Montenegro, Vol. 1, pp. 330-333, June 29 – July 1, 2005.

[4] R. Kountchev, M. Milanova, C. Ford, and S. Rubin. "Multimedia Watermarking with Complex Hadamard Transform in the Inverse Pyramid Decomposition", *Proc. of the 2003 IEEE Intern. Conf. on Information Reuse and Integration*, Las Vegas, USA, , pp.305-310, October 2003.

[5] R. Mironov and R. Kountchev, "Analysis of Complex Hadamard Transform Properties", *In ICEST'2006 Conference Proceedings*, Sofia, Bulgaria, pp. 173-176, June 29 – July 1, 2006.