A Content-Based Digital Image Watermarking Algorithm Robust Against JPEG Compression

Amir Najafi, Ali Siahkoohi, Mohammad B. Shamsollahi
Electrical Engineering Department
Sharif University of Technology

amir.najafi1990@gmail.com, ali.siahkoohi@gmail.com, mbshams@sharif.edu

Abstract—**watermarking (imperceptible insertion of data into the host image or soundtracks) has attracted many research interests in the recent decade. The aim of this correspondence is to introduce a none-blind, content-based method for image watermarking. It employs properties of Human Visual System (HVS) and is capable of embedding a high-energy watermark (consequently robust against lossy compressions and other kinds of attacks) while prevents perceptible degradation in the host image. To modulate the binary coded watermark on a spread-spectrum noise, the introduced algorithm is followed by spread-spectrum watermarking. The algorithm was simulated and experimental results represent magnificent improvement in watermarking capacity and accuracy of extraction.**

Keywords-non-blind watermarking; content-based; edge map; human visual system; JPEG compression

I. INTRODUCTION

Development of communication and multi-media systems magnify the necessity of means to protect copyright in data domain such as digital imagery and video recordings. Watermarking (embedding of information without perceptible degradation in fidelity of data) could be considered as a solution for this problem. Various methods have been introduced for this type of steganography [1][2][3], each consists of foibles and strengths of their own. In some applications, the main feature of an appropriate watermarking method is robustness against different kinds of unintentional attacks, such as JPEG compression, which could be applied to the image while transferring through the network. The above mentioned property could be considered as one of the main criteria for comparison of improvements introduced by recent researches [4][5][6][7][8]. Researchers have applied a number of highly different algorithms to obtain robustness against JPEG compression. Tao et al. [4] embeds a binary pattern in the form of a binary image in the LL and HH bands at the second level of Discrete Wavelet Transform (DWT) decomposition. Lee et al. [5] actively utilizes the JPEG quality level as a parameter in embedding a watermark into an image. Chang et al. [6] applied a two-level Discrete Wavelet Transform (DWT) and to ensure the robustness of the watermarks and protect the copyright of an image. Patra et al. [7] proposed a novel Chinese Remainder Theorem (CRT)-based technique for digital watermarking in the Discrete Cosine Transform (DCT).

Watermarking methods in terms of embedding and extraction of watermark can be classified into blind, semi-blind and non-blind algorithms. Non-blind watermarking, which is employed in this work, needs both the watermarked and the original version of the image for extraction procedure. In the proposed algorithm, based on the features of human visual system and content of the host, the approach is to insert the information in special regions of the image, which permits the user to embed high energy watermarks while prevents visible or perceptible changes for the viewer. Most of the previously introduced methods [4][5] seem to encounter critical uncertainties in the extraction procedure as the compression ratio keeps increasing. The proposed algorithm performs a brilliant immunity even for high orders of compression.

The rest of the paper is organized as follows: The proceeding section is devoted to a brief explanation about JPEG compression algorithm and its influence on the image. It focuses on the impact of changing the low or high frequency coefficients in DCT domain and quantization errors. In section III, development of the main idea based on experimental observations is discussed concisely, while in section IV the proposed method is fully explained and analyzed. Section V and VI will represent the simulation results and conclusions, respectively.

II. A BRIEF INTRODUCTION TO JPEG COMPRESSION AND ITS INFLUENCE ON IMAGE

In the first level of JPEG compression, the image is divided into non-overlapping, 8×8 blocks followed by transforming each block to DCT domain [9][10]. Second step corresponds to quantization of DCT coefficients through a set of division and rounding operations, using a specified quantization table, shown in Fig. 1. Each quantized component is specified as the integer part, obtained from dividing the original value by the corresponding entry in the matrix.

The Hoffman entropic compression is applied to the quantized coefficients at the final step [11].According to Energy compaction property of DCT, high frequency coefficients would have a relatively negligible role in reconstruction of the associated image block and consequently in the human visual system [15]. The stated property together with the potential capacity of the high frequency coefficients for data hiding, introduce them as appropriate candidates for watermarking. Unfortunately, based on experiments, these coefficients encounter a high degradation during a JPEG attack. According to the previous studies, low-frequency coefficients reveal more immunity in JPEG compressions, while manipulating those results to an undesirable impact on the quality of the image.

In selecting a suitable frequency band for signal in DCT domain, the explained features lead to a trade-off among the host fidelity and minimization of extraction errors.

III. ACONTENT-BASED APPROACH AS A REMEDY

The solution proposed in this work is to prioritize certain spatial regions of the host image, for the data insertion level, based on the content of the image. Regions, which include strong texture and edge patterns, would be considered as the primary candidates for data embedding, while smooth parts and low-frequency locations tend to be unchanged. The discussed idea is developed regarding the fact that high-frequency regions including textures and edge descriptors, seem to be invariant against small changes due to human perception. In other words, human visual system demonstrate more sensitivity to changes in smooth regions, while small details in textures patterns are usually ignored.

The experimental observations, which have justified the claimed properties, have been represented in Fig. 2 and Fig. 3. The latter, specifies the regions with maximum extraction error of the watermarked data (the black colored blocks), after a JPEG attack is applied to the watermarked image. Our heuristic investigation indicates the fact that rough sections of the image, show more immunity against the JPEG compressions.

Fig 1. The watermarked image, using an arbitrary watermarking algorithm proceeded by a JPEG compression.

Fig 2. Black blocks, indicate regions with maximum error during the extraction procedure. Rough regions, heuristically demonstrate more robustness.

IV. THE PRROPOSED ALGORITHM

A. Watermark Insertion

The proposed algorithm could be organized as follows:

*1) Block Selection:*The gray-scale host image is divided into non-overlapping 8×8 blocks and the edge map of each block is extracted.

Various methods can be employed to obtain the edge-map. Sobel algorithm has been utilized in this study [10][12].

Prioritizing the blocks for data insertion candidature is performed by comparing the average pixel intensities obtained from the edge-map associated with each region. Blocks with higher average intensity (consequently including more high-frequency regions) would be identified as primitive selections. The block selection procedure keeps progressing until the percentage of selected blocks reaches a user-defined parameter known as η (Selection ratio). Appropriate selection ratio for an arbitrary image, depends on both the size of the watermark and the content of the host data. Images including high amounts of detail and texture contain more potential capacity for watermarking.

2) Encryption of the Watermark: The so called watermark is considered as a binary image which can be represented as a bi-symbol array of {1,-1}. The idea of spread-spectrum watermarking [13] is to multiply the bi-symbol array by a pseudo random sequence, generated with a predetermined seed between the embedder and extractor. This in fact modulates the binary code on a spread-spectrum noise. This type of encryption not only enhances the security of data transmission, but also makes the watermark similar to a white-Gaussian noise and leads to more intangibility of the produced distortion by a third party.

*3) Embedding the Watermark:*The watermark should be amplified by a tunable parameter known as s (strength) before the insertion procedure, which enables the user to manually control the energy of embedding watermark.

The embedding procedure has been considered as an addition of the previously generated sequence to the low-pass corner of DCT coefficients obtained from the selected blocks (except of DC components). A schematic representation of the explained procedure indicated in Fig.4 shows the highlighted components which will be manipulated by adding the watermark sequence. The associated bandwidth$w$, for each

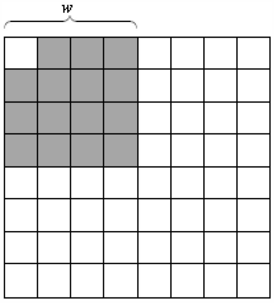

Fig 3. Selected squares (gray colored) in each selected 8×8 block for embedding the watermark

block, usually a global value for the entire blocks, is another user-defined parameter which performs a trade-off among the energy of distortion and capacity of data watermarking. The explained method employs $w^2 - 1$ binary symbols to be embedded in each block. Fig. 4 illustrates the procedure of embedding the watermark into a host image.

B. Watermark Extraction

After the embedding procedure, inverse Discrete Cosine Transform is applied to unchanged and manipulated blocks, in order to obtain the watermarked version of the image. Since the original host image is available, during the extraction process all of the previous steps are inverted and the embedded watermark could be uniquely extracted.

C. Capacity

If L indicates the dimension of a square-shaped gray-scale host image, then such data consists of $\left\lfloor \left(\frac{L}{8}\right)^2 \right\rfloor$blocks. During the block selection procedure, the number of blocks reduces to $\left\lfloor \eta\left(\frac{L}{8}\right)^2 \right\rfloor$ $(0 < \eta \le 1)$. As mentioned before, η is a tunable parameter and mostly depends on the content of the image and size of the watermark if selected optimally. As discussed in the preceding subsection, the capacity of data watermarking equals to $w^2 - 1$symbol per block. Thus, the total number of binary symbols which can be embedded in the host image would be:

$$N = \left\lfloor \eta\left(\frac{L}{8}\right)^2 \right\rfloor (w^2 - 1) \quad (1)$$

In the case of higher watermarking capability of the host image comparing to the length of the watermark sequence, the extra capacity could be considered for adding redundancy. This can be performed by Forward Error Correction (FEC) coding as an error reduction strategy in data extraction level [14].In this paper, in order to focus on the content-based aspect of the method, simple repetition of the watermark sequences is utilized. During the extraction, if the embedded symbol is optimally estimated, the majority voting method can be chosen as the symbol prediction process.

D. Coefficients Weight Allocation

As mentioned before, the JPEG compression distortion shows less influence in the low-frequency coefficients of DCT

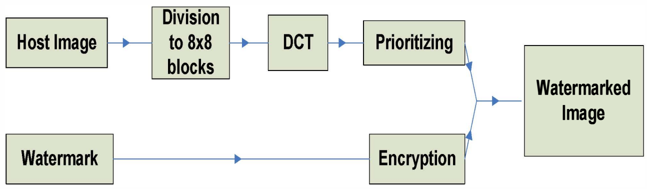

Fig 4. Stages of watermarking algorithm

at each block. In order to improve the decoding process one needs a minimum error in symbol prediction. To attain this we employed a more efficient version of the majority voting method. In the utilized voting method the weight of each coefficient is inversely proportional to the distance between the block of the coefficient and that of the DC component. In other words, the influence of an arbitrary component a_{mn}, is reduced proportional to $\left(\frac{1}{\sqrt{m^2+n^2}}\right)^p$:

$$v_{final} = \sum_{m=1}^{W} \sum_{n=1}^{W} v_{m,n} \left(\frac{1}{\sqrt{m^2+n^2}}\right)^p \quad (2)$$

where $v_{m,n}$ indicates the decision vote made by the (m, n) entry in the block and p is a tunable multiple. A defuzzification process should be applied to the v_{final} in order to make the final crisp decision. Note that the DC component would be omitted from the summation.

V. Experimental Results

The 512×512, gray-scale image of baboon is used as a host image and a 128×128 binary image as a watermark. Based on the content of the host and the size of the watermark, the following adjustments were applied:

$$\eta = 0.5$$
$$\text{Threshold of the Sobel filter} = 0.022$$

Fig 5. The block selected image with η=0.5 (un-selected regions are indicated as black blocks)

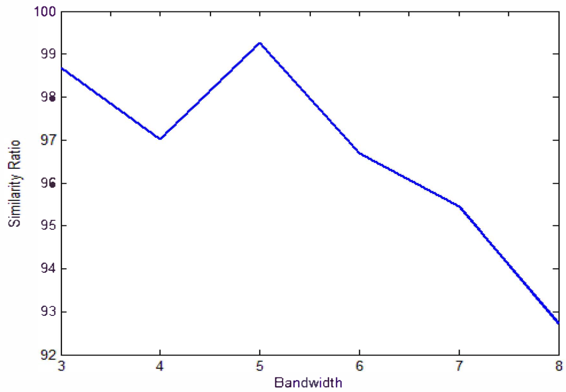

Fig 6. Similarity ratio as a function of bandwidth for a fixed JPEG quality of 30%.

A scheme representing the result of block selection procedure is shown in Fig.5.

Fig.6 depicts the similarity ratio among the original and the extracted watermark, proceeded by a JPEG compression at a constant quality of 30%, and different values of bandwidths. The mentioned diagram represents a maximum similarity at $W = 5$, which consequently will be chosen as the optimum value for further simulations.

The strength parameter s, is selected in such a way that distortions become imperceptible in the watermarked image. Figure 7 and 8 show the watermarked images with $\eta = 1$ and $\eta = 0.5$, respectively; magnifying the importance of block selection procedure in performance of the method. In other words, ignoring the content of the host image leads to a certain constraint in the energy of the watermark, while the proposed method is capable of inserting high energy watermarks without perceptible degradations. Fig. 9 represents the original host image.

Fig. 10 represents the improvement of similarity through ascending the quality of JPEG compression for various choices of s, demonstrating the significant performance of the method. It should be noted that the graphs are labeled by their corresponding $PSNR$ value (a directly controllable parameter by s) only for more clarity. However the proposed method is not guaranteed to be efficient in high $PSNR$ values, while it has been claimed that through a content-based procedure there would be intangible visual degradation even for high energy watermark insertions.

Some sample points of the diagram indicated by $PSNR$ of $32\ dB$ represented in Fig. 11 have been shown in Table 1, demonstrating the achievement of perfect extraction of the watermarked data for compression qualities equal or more than 40%, and a considerable robustness for lower qualities such as 25%. It has been observed for most of the previous methods [4][5], encountering many erroneous detections in compressed images with qualities lower than a method-dependent threshold, while the proposed method could reduce the mentioned threshold to a lower value.

Fig. 11 demonstrates the performance of the proposed method as a function of $PSNR$ for a fixed JPEG quality of 50%.

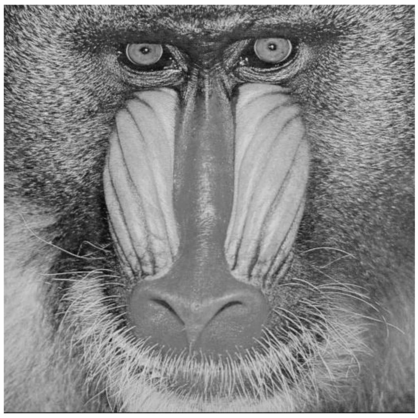

Fig 7. The JPEG version of the watermarked image with half-full selection ($\mathbf{\eta = 0.5}$) and quality of $\mathbf{30\%}$ ($PSNR = 28dB$).

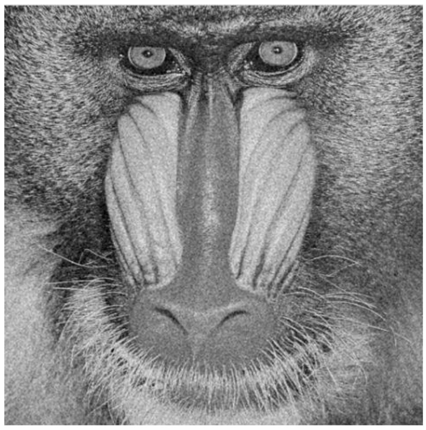

Fig 8. The JPEG version of the watermarked image with full selection ($\boldsymbol{\eta = 1}$) and quality of $\mathbf{30\%}$ ($PSNR = 25dB$).

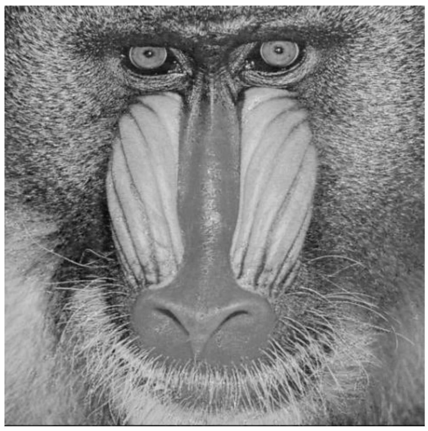

Fig 9. The original host image

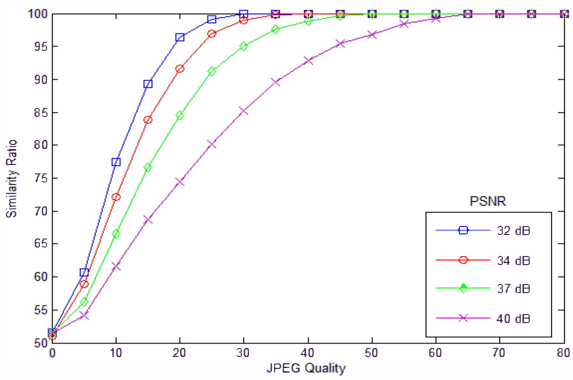

Fig 10. Similarity ratio among the original and extracted watermark as a function of JPEG quality for different values of PSNR.

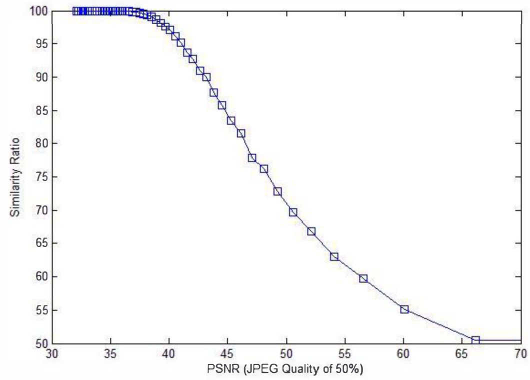

Fig 11. Similarity Ratio as a function of PSNR for a fixed JPEG quality of 50%.

TABLE I. Similarity ratio for some JPEG qualities ($PSNR = 32dB$)

JPEG quality	25%	35%	$\geq 40\%$
Similarity ratio	97%	98.8%	100%

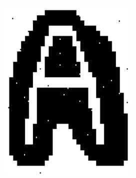

Fig 12. The original 128×128 binary image used as the watermark (left picture).

Fig 13. The extracted watermark after a JPEG compression of **35%** applied to the watermarked image (right picture).

VI. Conclusion

The aim of this correspondence was to introduce a method for watermarking digital images, with high robustness against lossy compressions such as JPEG. The proposed algorithm was based on the content of the host image for a heuristically optimal selection of regions in order to insert data, which in spite of embedding high energy watermarks into the host image, prevents perceptible degradation of fidelity. It has been observed that host images with strong texture patterns and high energy edge-maps demonstrate more robustness during the extraction process. Obvious improvements have been observed in the quality of extracted watermarks, even in high compression ratios for the mentioned categories. As an experimental analysis the performance of the proposed algorithm could be compared to a previous method [4] with achievement of 99.0% accuracy after a JPEG compression with %50 of compression ratio.Referring to table 1 the proposed method yields a significant robustness (approximately 2.98 times reduction in BER) during the mentioned attack. The unused parts of the image will have a considerable capacity for other types of watermarking.

Automation for optimal selection of the previously user controlled parameters could be considered as an extension of this work. From another aspect, the main criteria in this work for selecting the appropriate regions of data watermarking, was a simplified measurement of edge densities, which is capable of being optimized to a more complex and optimal formalism.

References

[1] V. M. Potdar, S. Han, E Chang, “A Survey of Digital Image Watermarking Techniques”, 2005 3rd IEEE International Conference on Industrial Informatics (INDIN).

[2] Abbas Cheddad, JoanCondell,KevinCurran, PaulMcKevitt, "Digital image steganography: Survey and analysis of current methods", Elsevier, 2010.

[3] Juan R.Hernández, Martín Amado, Fernando Pérez-González, "DCT-Domain Watermarking Techniques for Still Images: Detector Performance Analysis and a New Structure", IEEE Transactions On Image Processing, Vol. 9, No. 1, January 2000.

[4] Peining Tao, Ahmet M.Eskicioglu, “A Robust Multiple WatermarkingScheme in the Discrete Wavelet Transform Domain”, in Symposiumon Internet Multimedia Management Systems V, Philadelphia, PA, 2004.R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.

[5] Hye-Joo Lee, Ji-Hwan Park, and Yuliang Zheng, “Digital Watermarking Robust Against JPEG Compression”, In InformationSecurity, LNCS 1729. Springer-Verlag, Berlin, 1999,pp. 167-177.

[6] Chin-Chen Chang, Kuo-Nan Chen, Ming-Huang Hsieh, " A Robust Public Watermarking Scheme Based on DWT", Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2010.

[7] Jagdish C. Patra, Jiliang E. Phua, Cedric Bornand, "A novel DCT domain CRT-based watermarking scheme for image authentication surviving JPEG compression", Elsevier, 2010.

[8] Shinfeng D. Lin, Shih-Chieh Shie, J.Y. Guo, " Improving the robustness of DCT-based image watermarking against JPEG compression", Elsevier 2010.

[9] N. Ahmed, T. Natarajan, and K. R. Rao, "Discrete Cosine Transform", IEEE Trans. Computers, 90-93, Jan 1974.

[10] R.C.Gonzalez and R. E. Woods, "Digital Image Processing", 2nd Ed., 2002, Prentice Hall.

[11] D.A.Huffman, "A Method for the Construction of Minimum-Redundancy Codes", 1952.

[12] B.Jähne, H.Scharr, and S.Körkel. "Principles of filter design", In Handbook of Computer Vision and Applications, Academic Press, 1999.

[13] Ingemar J.Cox, Joe Kilian, F.Thomson Leighton, Talal Shamoon, "Secure Spread Spectrum Watermarking for Multimedia", IEEE Transactions On Image Processing, Vol. 6, No. 12, December 1997.

[14] Rohit Puri, Kannan Ramchandran, "Multiple Description Source Coding using Forward Error Correction Codes", in Conf. Rec. 33rd Asilomar Conf. Sig., Sys., & Computers, Pacific Grove, CA, Oct. 1999, vol. 1, pp. 342-346.

[15] K. R. Rao and P. Yip, "Discrete Cosine Transform: Algorithms, Advantages, Applications", Academic Press, Boston, 1990.