Imperceptible Image Indexing Using Digital Watermarking

Jobin Abraham[1] and Varghese Paul[2]

[1] M.G University, Kerala
[2] CUSAT, Kochin, Kerala, India
{jnabpc,vp.itcusat}@gmail.com

Abstract. Proposed image watermarking scheme embeds identification watermark in certain selected regions where modifications introduced during the process of watermarking is less sensitive to HVS (Human Visual System). Edge detectors are used to estimate regions in the image where intensity changes rapidly. Modifications to such pixel will not attract the attention of human eyes. Watermark is thus integrated imperceptibly into the digital images. The proposed is a scheme for embedding a unique index number as watermark for content tracking and identification.

Keywords: Digital watermarking, DCT based, edge detection, watermark embedding, extraction, psnr.

1 Introduction

Today, Internet is popular and widely used for communication and commercial purposes. Any digital resources such as image, music, video or multimedia data can be transferred to a buyer via Internet. Commercial vendors may market and sell their contents to potential buyers. Online business transactions are dependable only if powerful tools exist for controlling unauthorized replication and misuse of costly contents transferred. Watermarking is proposed as a powerful security mechanism that has immense applications as a tool for copyright protection, authentication, fingerprinting and many more [1]. Watermarking can also be used to restrict content misuse and illegal tampering.

Digital watermarking is the process of integrating identification information such as the owner name or his logo imperceptibly in the digital media. The identification watermark can be a text or logo image for uniquely identifying the content owner [2]. Two common approaches for watermark selection can be seen in the literature. One is the use of a pseudo random sequence [3] and the other is the use of an image as watermark [4]. The watermark image can be a logo or initials of the company or owner. In most methods, one dimensional array of binary digits is formed by preprocessing the watermark and these bits are then integrated with the original image during the process of watermarking.

Method proposed in [5], uses a binary image as watermark data. Non-overlapping 2x2 blocks from host image is taken and one pixel of watermark is embedded. Some

S.M. Thampi et al. (Eds.): SNDS 2012, CCIS 335, pp. 110–116, 2012.

proposed watermarking schemes embed a PN sequence in the LSB of host data. Though these spatial domain methods are simple and easy to implement, they are highly sensitive to signal processing operations that corrupt the embedded watermark.

Watermarking is still in evolutionary stage. Watermarking techniques has generated great interest as more and more vendors opt to sell their works through Internet. Watermarks should be able to withstand all intentional and unintentional attacks targeted to remove the identification mark. Creating robust watermarking methods is still a challenging problem for researchers as some methods withstand some attacks, but is broken by others. Current research is concentrated mainly on methods to suit to the specific application in hand and effectively address that issue.

A watermarking scheme that uses edge detection is proposed in [6]. The method use Gaussian noise of zero mean and unit variance as watermark. Though the method can address copyright issues, it is not pragmatic for sequentially ordering the digital resources using a numeric identifier.

[7] describes another method that can embed a watermark comprising eight symbols. Every element in the watermark is coded using a {+1, -1} to deduce a 64 bit ID. This is then embedded in host image.

The paper is organized as follows: section 2 introduces the aim and features of the proposed scheme. Section 3 outlines watermark embedding and extraction algorithm. Section 4 describes the experimental analysis and quality metrics used for assessing the proposed scheme. Finally, section 5 concludes the work.

2 Features of the Proposed Method

The proposed is a novel scheme for embedding a unique identification number as watermark for content tracking and identification. This may be used for sequentially numbering documents in a content archival system. The proposed method can also be used for fingerprinting digital contents with buyers ID. The seller can use this hidden information for identifying the owner and tracing the traitor in cases of illegal content replications.

Vidyasagar et al. [8] describe content archiving as a potential application for watermarking. The integrated watermark here acts as an object identifier. This has several advantages over conventional archiving that uses file names. Documents when archived using the file names, accidental changes in file names will create unpredictable issues in file access that may even lead to the non availability of contents.

Edges in images are areas with sharp variation in pixel intensity. There will be noticeable changes or abrupt discontinuities in pixel intensity in the vicinity of edges. These regions are detected using an edge detector and are effectively used in the proposed method for hiding an external watermark signal without distracting the eyes of the viewer.

Robustness is a measure of immunity of embedded watermark to attacks targeted to weaken them. Normally when images are cropped part of the embedded watermark is lost, making extraction of identification watermark cumbersome and unsuccessful. However, the method assures watermark retrieval even if the image is reduced to one-by-fourth of the actual size.

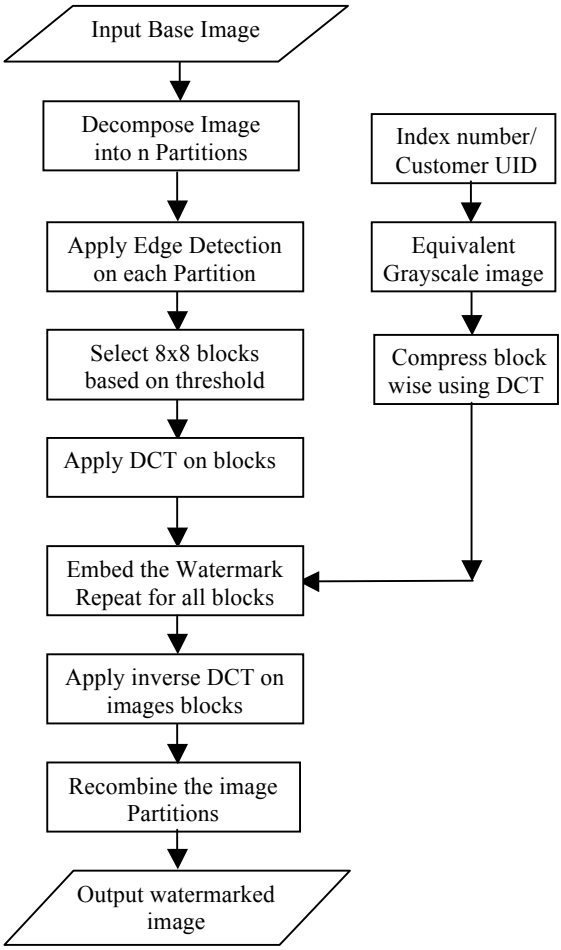


Fig. 1. Embedding process flowchart

3 The Algorithm

The algorithm has two phases. First is the watermark embedding phase and second the watermark extraction phase. General steps in watermarking are:

1) Preprocess the numeric watermark
2) Decompose the image I into n equal sized regions and estimate the edge detection threshold.
3) Construct non-overlapping 8x8 blocks and apply DCT.
4) Select blocks for watermark embedding based on the edge detection threshold.

5) Alter the coefficients in selected blocks to hide the watermark
6) Apply inverse DCT on blocks and recombine the sub-sections.

3.1 Embedding Algorithm

The watermark phase algorithm has two stages. First stage is the preprocessing and compressing of watermark image using DCT. Most significant coefficients are chosen for embedding in the base image. The second stage, adds the preprocessed watermark signal to selected mid-frequency locations in the regions selected using edge detector. Flowchart of the embedding process is shown in figure.1

Stage 1: Preprocess the watermark

Input: Unique ID, UID
Output: Equivalent grayscale watermark image, W

- Assign an equivalent grayscale values for each decimal digits in the number, d = {0, 1, 2, 3….9} from the set of grayscale values, egv = {25, 50, 75, ….225}.
- Construct a 4x4 block for every digit by repeating the selected equivalent grayscale value from the above list.
- Convert every digit in the UID in the same fashion above, to obtain a watermark image. For instance, a sixteen digit UID will result in a 16x16 grayscale image.

Stage 2: Watermark embedding

Input: Base image, I and watermark image, W
Output: Watermarked image, I_w

- Compress W by applying DCT on 4x4 non-overlapping blocks. Built an array, wmk(q) = $f_i(1,1)$. Here, i is the block number of a digit in the watermark.
- Split I into n regions. If n=4, results in four quadrants, each of size N/2 X N/2.
- Embed the watermark in each quadrant, considering one at a time.
- Use Sobel Edge detection for finding the threshold, t
- Consider 8x8 blocks and compute their threshold, t_b.
- Select the blocks with in a range of threshold. Say, (($t_b > 2t$) && ($t_b < 3t$))
- Embed the watermark in a mid-frequency coefficient, f(i,j) = f(i,j) + sf*wmk(q) where sf is the strength factor and f is the DCT of an 8x8 block from I.
- Repeat the above three steps to mark all eligible block in a region.
- Apply inverse DCT and recombine the sub-regions to output the watermarked copy, I_w.

3.2 The Extraction Algorithm

The proposed is a non-blind technique and requires the original image for extracting the hidden watermark signal. DCT of base image I and the watermarked copy I_w are found and the difference between the chosen mid-frequency coefficient is computed for decoding the embedded watermark information.

The watermark extraction algorithm is discussed below:

Input: Base image, I and Watermarked image, I_w
Output: Watermark, W

- First step is to decompose the image I and I_w into sub-regions, as done during watermark embedding.
- Determine the edge detection threshold t for the sub-region and t_b for each 8x8 block in consideration from I
- Extract embedded watermark information from I_w whenever threshold for the block is within the range employed during the watermarking process.

  $$wmk(k) = (f'w_i(p, q) - f_i(p, q))/sf.$$

  Here, i is the block number, sf is the strength factor, fw is DCT of an 8x8 block from I_w and f is DCT of the corresponding 8x8 block from I.
- Extract all watermark coefficients by considering the subsequent set of blocks from I_w and I.
- Reconstruct the watermark image by taking inverse DCT of each extracted element to generate the 4x4 grayscale blocks per digit.
- Use the proposed grayscale lookup table for finding the equivalent digit and repeat the above steps for extracting all the digits in the numeric watermark.
- Output the extracted watermark , *w*

4 Experimental Analysis

For evaluating the performance of the proposed scheme we watermarked various images. Figure.2 shows the results for two test cases, where images are watermarked using a sixteen digit number. Two important requirements to be met by any watermarking scheme is visual appeal ness and robustness to watermark removal attacks. PSNR (Peak Signal to Noise Ratio) ratio is computed to estimate the visual quality of the watermarked image. Higher the PSNR lesser the distortion and hence better the image quality. The measured PSNR values are tabulated in table.1. The robustness of the watermarked image is tested by subjecting to certain common signal processing attacks. The method is found to be dependable as the watermark is still decodable and recognizable. NCC (Non-correlation coefficient) is used to measure the similarity of the extracted the watermark with the original embedded watermark. NCC value ranges from 0 to 1. Any value close to 1 indicates that the extracted signal is closely similar to the original embedded watermark. Table.2 gives the NCC values for the extracted watermarks, supposing that there are four watermarked regions, I to IV, in a watermarked Lena image.

Table 1. Sinal to Noise Ratios

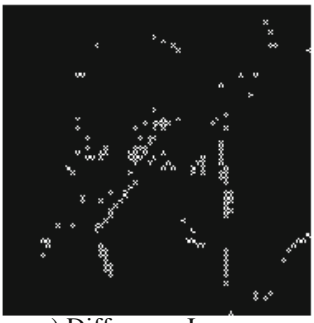
Image	**PSNR**	**MSE**
Lena	45.28	1.92
Deer	44.11	2.52
Boat	44.65	2.28

Fig. 2. Watermarking images. a) Base images b) Watermarked images

Figure.3 shows the pixels that were altered during the process of watermark insertion. The significant edges are alone carefully selected from the base image discarding the monotonous areas in the image. As there is a sharp change in intensity near the edges distortions are not observable to the human eyes. Hence it may be inferred that the proposed watermarking method do not significantly degrade the visual qualities of the output watermarked image. Figure 3.a shows the binary difference image of the base Lena image and its watermarked copy. And in figure 3.b, the watermark hidden areas in base image are highlighted.

Table 2. Correlation Measurements after attacks for watermarks from regions I to IV

Types of Attack	**NCC**			
	I	**II**	**III**	**IV**
No Attack	0.99	0.99	0.99	0.99
Salt & Pepper	0.95	0.99	0.97	0.94
Gaussian	0.92	0.93	0.87	0.82
Speckle	0.88	0.97	0.91	0.98



a) Difference Image

b) Watermarked Positions

Fig. 3. Watermark embedded positions in lena image

5 Conclusion

A method for imperceptibly indexing digital resources using watermarking is discussed here. The embedded index number could be used effectively for archiving multimedia digital documents in databases. It also helps in locating the resources in case of causalities like unintentional alterations. To ensure that watermark is embedded imperceptibly, the watermark is compressed using DCT and only significant coefficients are selected for insertion in base image. The algorithm is then implemented on various test image and the results for various quality metrics is found. It is observed that the embedding process does not degrade the quality of the base image and at the same is robust against various image processing attacks.

References

1. Cox, I.J., Miller, M.L.: The First 50 years of Electronic Watermarking. J. of Applied Signal Processing, 126–132 (2002)
2. Lian, S., Kanellopoulos, D., Ruffo, G.: Recent Advances in Multimedia Information System Security. Informatics 33, 3–24 (2009)
3. Al Haj, A.: Combined DWT-DCT Digital Image Watermarking. Journal of Computer Science 3(9), 740–746 (2007)
4. Sharkas, M., Elshafie, D., Hamdy, N.: A Dual Digital Image Watermarking Technique. World Academy of Science, Engineering and Technology (2005)
5. Dorairangaswamy, M.A.: A Novel Invisible and Blind Watermarking Scheme for Copyright Protection of Digital Images. IJCNS 9 (2009)
6. Ellinas, J.N.: A Robust Wavelet based Watermarking Algorithm Using Edge Detection. World Academy of Science, Engineering and Technology 34 (2007)
7. Kim, W.-G., Seo, Y.-S., Jung, H.-W., Lee, S.-H., Oh, W.-G.: Wavelet Based Multi-bit Fingerprinting Against Geometric Distortions. Key Engineering, 1301–1305 (2006)
8. Potdar, V.M., Han, S., Chang, E.: A Survey of Digital Image Watermarking Techniques. In: IEEE International Conference on Industrial Informatics (2005)