ADAPTIVE VISIBLE WATERMARKING OF IMAGES

Mohan S Kankanhalli

School of Computing
National Univ. of Singapore
Singapore 119260
Email: mohan@comp.nus.edu.sg

*Rajmohan** & *K R Ramakrishnan*

Dept. of Electrical Engineering
Indian Institute of Science
Bangalore 560012
Email: krr@ee.iisc.ernet.in

ABSTRACT

Digital watermarks are emerging as a tool to provide copyright protection for high quality images and video. Though a lot of work has been done in the area of invisible watermarks, relatively less earlier work exists for visible watermarks. As a consensus on the various issues for invisible watermarks to be legally admissible eludes us, a visible watermark could serve as a deterrence to theft and also to provide instantaneous recognition of the owner or creator of an image. We propose a technique in which the location and strength of the watermark image to be embedded is varied in accordance with the underlying content of the image to be watermarked. We propose a new algorithm that classifies each block of 8×8 of pixels into one of 8 classes depending on the sensitivity of the block to distortion. We analyze the texture, edge and luminance information in the block for this purpose. The embedding process is automated and the bits are embedded in the DCT transform domain. Since the strength of the watermark in a block depends on the class to which the block belongs, the result is a pleasant and unobtrusively watermarked image irrespective of the type of image.

1. INTRODUCTION

The growth of the Internet as well as the availability of sophisticated software and fast hardware has made it possible to create and transfer very high quality images. Just as in the case of renowned paintings and artwork, the creation of such images entails a lot of hard work and talent coupled with the use of specially designed expensive equipment. It is therefore imperative to provide copyright protection or at-least a visible copyright notification for such images. One possible solution to the copyright protection problem is the embedding of a secondary signal into the primary image signal which is not perceivable. This secondary signal, which is called an *invisible watermark*, is bonded so well with the original data that it is inseparable and survives any kind of multimedia signal processing. This is currently an active area of research [11] and it is being investigated for all types of multimedia data – image [2, 10, 14], video [3, 9, 13] and audio [1]. It is however increasingly being felt that such watermarking techniques need to be utilized in conjunction with encryption, site security and a proper legal framework.

While these issues are being resolved, a logo or a line of text in the form of a secondary image (*visible watermark*), could be overlaid translucently onto a primary image to serve such a purpose. The main advantage of using a visible watermark is that they convey an immediate claim of ownership, providing credit to the owner. It also prevents or at least discourages, unauthorized use of copyrighted high quality images [12]. Such a deterrence is considered useful which is evidenced by the common practice of the use of a visible watermark in television broadcasting where the channel logo is translucently placed at the corner of the screen image. It also serves the dual purpose of providing a recognizable identity to the content.

Visible watermarks are used in much the same way as their bond paper counterparts, where the opacity of the paper is altered by physically stamping it with an identifying pattern. An ideal visible watermark, like its bond paper ancestor, should be therefore only noticed on careful examination of the image. It should also not be so delicate that it be lost by simple image processing operations that do not distort the original image appreciably. These contradicting requirements makes the creation of a robust and perceptually uniform watermark a hard and interesting problem.

Visual watermarking has been researched by the IBM group [12]. Here a visually unobtrusive watermark is embedded into a large area of the image by modifying pixel luminance values. A bright pixel is darkened or brightened or a dark pixel is darkened or

* Presently with the Corps of Signals, Indian Army

brightened by a perceptually equal amount. Randomization is added to the strength and the location of the watermark to make it less vulnerable to automated removal. We propose a new visible watermarking technique where we vary the intensity of the watermark in different regions of the image depending on the underlying content of the image. We achieve this by first analyzing the image content such as textures, edges and luminance and provide a just noticeable distortion level for every block. Using this measure for varying the strength of the overlaid secondary image ensures that it is perceptually uniform over different regions of the image and the process can be automated over a wide range of images.

The content based classification of the image will be discussed in section 2. In section 3 the classification will be utilized to adaptively embed a visible watermark. The experimental results will be given in section 4. The advantages of the proposed method and our areas of future work will be summarized in the concluding section.

1.1. Desired Characteristics of Visible Watermarks

Some of the desired characteristics of visible watermarks could be as follows :

- Visible in both color and monochrome images
- Visible only on careful examination of the image
- Should not significantly obscure the image details beneath it
- Should be difficult to remove automatically – should be *robust*
- The insertion process should be automated for all kinds of images

Watermarking techniques exploit the redundancies in the coded image to embed the information bits. Redundancies in images could be statistical or perceptual. We will briefly describe the discrete cosine transform (DCT) coded image due to its use in the JPEG standard. By transforming spatial data into another domain (such as spatial frequency), statistical independence between pixels as well as high-energy compaction can be obtained. The general method of DCT coding involves dividing the original spatial image into smaller $N \times N$ blocks of pixels, and then transforming the blocks to obtain equal-sized blocks of transform coefficients in the frequency domain. These coefficients are then thresholded and quantized in order to remove subjective redundancy in the image. To satisfy the desired characteristics, we embed in the DCT coded luminance channel of the image. The watermark is spread over a large area of the image so as to make it difficult to automatically remove it. The use of the content based classification ensures that the embedded image is visually pleasant and unobtrusive and the process can be automated.

2. CONTENT-BASED IMAGE SEGMENTATION

Watermarking an image is essentially the process of altering the pixel values of an image in a manner that ensures that the watermark is not obtrusive and there is only slight difference between the original and the watermarked image. Since any alteration of the image pixel values could be treated as a form of noise, we will interchangeably use the term distortion and alteration to imply the process of embedding or watermarking. Altering a large number of pixel values in an arbitrary manner will result in noticeable artifacts. We propose a new algorithm to classify the image regions based on their sensitivity to noise as shown in Figure 1 and Figure 2. We classify a block into one of the eight classes based on the sensitivity of the block to distortions. A highly textured block for example would be less sensitive to distortion and hence would be put into a higher class vis-a-vis a block with flat and featureless regions. The strength of the watermark is kept proportional to the class to make it uniformly visible. We have used a similar classification to obtain a new invisible watermarking algorithm [5].

The visibility of distortion in a region of the image depends on the underlying image content features as listed below

- **Edges.** Edge information of an image is the most important factor for our perception of the image. It is in fact the necessary and sufficient information that is to be transmitted if the final receiver is the HVS [7, 8]. It has the least noise sensitivity i.e. the lowest just noticeable distortion (*JND*) value and it is essential to maintain edge integrity so as to preserve the image quality.
- **Smooth Areas.** Psycho-visual studies [4, 7, 8] have shown that the HVS has a general bandpass characteristic. Smooth areas influence our perception together with the edge information. The JND perception thresholds are relatively low as compared to strongly textured regions.
- **Textures.** The distortion visibility is low when the background has strong texture [4, 7, 8]. A

strongly textured region has a very high noise-sensitivity level.

- **Brightness Sensitivity.** When the mean value of the square of the noise is the same as that of the background, the noise tends to be most visible against a mid-gray background [4, 7, 8]. The mid-gray regions have lower JND's as compared to the other regions.

2.1. Content based classification

2.1.1. Texture and Edge Analysis

In a block with abrupt changes between adjacent pixels, the signal energy tends to be concentrated in the AC coefficients. In a flat featureless region of an image the energy is concentrated in the low frequency components. Thus the energy in the AC coefficients can be used as a measure of roughness. Denoting the i, j DCT coefficient as X_{ij} and using the monotonic function log for range compression we arrive at the following expression for the energy in the AC coefficients.

$$E_{AC} = \log \sum_{i,j} (X_{ij})^2 - (X_{00})^2 \qquad (1)$$

where X_{00} is the DC DCT coefficient. The maximum energy is in the AC coefficient for a block with a checkerboard pattern with the adjacent pixels having the maximum and minimum permissible gray value. We denote this as E_{max} and is evaluated as

$$E_{max} = \log(G/2)^2 \qquad (2)$$

where G is the maximum permissible gray value.

Using E_{max} for the normalization factor we arrive at the measure for the roughness level R for the block b.

$$R_b = \frac{E_{AC}}{E_{max}} \qquad (3)$$

The range of R_b is uniformly split into 8 sub-groups and the block is given a block distortion index I_b where $I_b \in 1 \ldots 8$ and a high value of I_b implies lower distortion tolerance and vice versa.

2.1.2. Edge Separation and Reclassification

A block that has a large amount of energy in its AC coefficients could either contain prominent edges or be highly textured [6]. The spread of the fluctuations of energy in the block of pixel values would indicate whether it contains an edge or is highly textured. The measure of fluctuations at any location x in the pixel block is given by its gradient, ∇x. In a highly textured block, ∇x would be large at a large number of locations while in a block with prominent edges ∇x would have large values at much fewer locations. By comparing the magnitude of fluctuation to suitably selected threshold λ, it is possible to identify locations of large fluctuations within the image block. A count of these fluctuations can then used to decide if the block has an edge or is highly textured. The threshold count C, is arrived at statistically to make the decision. Therefore a block with large energy or in other words all the blocks satisfying the constraint $I_b \in \{6, 7, 8\}$ have to be subjected to the above explained test and distortion index is reallotted to such blocks with strong edges in the range 1 to 3 based on the edge strength. A block with large R_b has a strong edge if

$$||\{x | x \text{ is a pixel and } mag(\nabla x) > \lambda\}|| < C \qquad (4)$$

where $|| \cdot ||$ is the cardinality of the set.

Each block based on its corresponding I_b is mapped to two scaling factors α and β empirically.

$$\hat{\alpha}_b = 0.95 + 0.05 \times (I_b - 1) \qquad (5)$$
$$\hat{\beta}_b = 0.01 + 0.20 \times (I_b - 1) \qquad (6)$$

where I_b varies from $1 \ldots 8$ and $\hat{\alpha}_b$ and $\hat{\beta}_b$ are the scaling factors for the image and the watermark DCT coefficients respectively, in the block b.

2.1.3. Luminance Sensitivity

The effect of luminance is incorporated into the scaling factors of the block depending on the block DC coefficient. Since the distortion in an image is most noticeable in the mid gray region and sensitivity falls of parabolically as the gray value drifts on both sides we make a final correction to the block scaling factors. The scaling factors of α_b is decreased as the average luminance value of the block drifts away from the mid-gray level and that of β_b is increased.

$$\alpha_b = \hat{\alpha}_b - 0.2 \times \frac{(DC_{max}/2 - X_{b,00})^2}{0.25 \times {DC_{max}}^2} \qquad (7)$$
$$\beta_b = \hat{\beta}_b + 0.2 \times \frac{(DC_{max}/2 - X_{b,00})^2}{0.25 \times {DC_{max}}^2} \qquad (8)$$

DC_{max} is the maximum possible DC value and $X_{b,00}$ is the DC coefficient of the block b. The implementation of the above classification algorithm is illustrated in Figure 1 and Figure 2.

The results of the content based image segmentation are shown in Figure 3, Figure 4 and Figure 5. The fog image has been chosen to illustrate our classification as there are areas of fog that are very sensitive to distortion and areas of grass that are less sensitive. As is evident in Figure 4, the classification algorithm has segmented the image based on the content's noise sensitivity. The Lena image illustrates the effect of edges.

3. CONTENT-BASED ADAPTIVE VISIBLE WATERMARK

We now propose a new method of watermarking of images that exploits both the statistical redundancies and the subjective redundancies in DCT coded images. The aim here is to apply a watermark that is visually pleasant and unobtrusive in a large area of an image by modifying pixel luminance values in a "perceptually uniform manner". Thus a bright pixel is darkened or brightened and a dark pixel is darkened or brightened by a perceptually equal amount. The basic scheme of watermarking is as depicted in Figure 6. The RGB image is first converted to YCbCr domain. Only the Y component is then altered to embed the watermark. This ensures that the watermark is visible in both color and grayscale images. The image is then segmented into 8×8 blocks and the block DCT is evaluated. The DCT coefficients are then perceptually analyzed as described in the previous section. Once the blocks are classified into different classes then the scale factors α_b and β_b as in equation 7 and in equation 8 are calculated. The DCT coefficients of the corresponding blocks of the original image and that of the image to be embedded is then scaled by the scaling factors and then added as given below:

$$\hat{X_{ij}} = \alpha_b \times X_{ij} + \beta_b \times W_{ij} \quad (9)$$

where$\hat{X_{i,j}}$ is the i, j DCT coefficient of the watermarked image in the block b, X_{ij} is the corresponding DCT coefficient of the original image and W_{ij} the DCT coefficient of the appropriately rescaled and resized watermark image, α_b and β_b are the weight factors derived from the perceptual analysis of the image. For typical images α_b varies between 0.95 and 0.99 and β_b varies between 0.01 and 0.15.

4. IMPLEMENTATION AND RESULTS

We have implemented the content based watermarking technique as a C program interfaced with the Independent JPEG group's JPEG decoder. The watermark is obvious in both color and monochrome images as the watermark is applied to the luminance values. The watermark covers a large area of the image yet does not significantly obscure the image details beneath it. It is very difficult and time consuming to remove the watermark automatically without destroying the commercial value of the image. It is thus resistant to automatic and systematic removal. It requires very little to no human interaction and thus can be fully automated. An implementation is shown in Figure 7. A minor variation of the above method where the watermark is scaled and positioned in the corner of image has also been implemented. This could find application in digital TV transmission. An example is shown in Figure 8. More details of our experiments can be found in [6].

5. CONCLUSIONS

We have proposed a novel and robust method of adaptively embedding visible watermarks into images based on the human visual system (HVS). A new method of classifying a region of the image based on its sensitivity or tolerance to noise has been devised. The underlying content of the image is analyzed to evaluate the noise sensitivity of different regions of the image. The effects of image content features such as texture, edge and luminance have been investigated. We classify a block into one of 8 classes based on the block's content analysis. In the proposed technique the watermark image is resized and positioned roughly around the center of the image. The DCT coefficients of the corresponding blocks of the watermark and the original are added in varying proportions depending on the class to which the image block belongs. This ensures that the watermark least distorts the regions that are sensitive to changes and exploits perceptual redundancies in the areas of high detail to enhance the strength of the watermark. Thus the image that is overlaid is uniformly visible irrespective of the underlying content of the image. Our experiments have shown that the technique can be automated for a wide variety of images. Randomization of the location and strength further makes automatic and systematic removal of the watermark very difficult. We are currently working on extending these ideas to MPEG video and are also trying to incorporate other HVS factors, such as contrast masking, in the proposed JND model.

6. REFERENCES

[1] L Boney, A H Tewfik and K Hamdy. Digital Watermarks for Audio Signals. Proceedings of 1996 IEEE International Conference on Multimedia Computing and Systems ICMCS'96, Hiroshima Japan, pp. 473-480, July 1998.

[2] I J Cox and M L Miller. A Review of Watermarking and the Importance of Perceptual Modeling. *Proceedings of the SPIE International Conference on*

Human Vision and Electronic Imaging II, San Jose CA, pp. 92-99, Feb 1997.

[3] J Dittman, M Stabenau and R Steinmetz. Robust MPEG Video Watermarking Technologies. *Proceedings of the Sixth ACM International Multimedia Conference – ACM MM'98*, Bristol, UK, pp. 71-80, September 1998.

[4] J Johnston, N S Jayant and R Safranek. Signal Compression Based on Models of Human Perception. *Proceedings of the IEEE*, 81(10):1385–1422, Oct 1993.

[5] M S Kankanhalli, Rajmohan, and K R Ramakrishnan. Content-based Watermarking of Images. *Proceedings of the Sixth ACM International Multimedia Conference – ACM MM'98*, Bristol, UK, pp. 61-70, September 1998.

[6] Rajmohan. Watermarking of Digital Images. *ME Thesis report, Department of Electrical Engineering, Indian Institute of Science Bangalore 560012, India*, 1998.

[7] X Ran and N Farvardin. A Perceptually Motivated Three-component Image Model – Part I: Description of the Model. *IEEE Transactions on Image Processing*, Vol. 4, pp. 401-415, April 1995.

[8] X Ran and N Farvardin. A Perceptually Motivated Three-component Image Model – Part II: Applications to Image Compression. *IEEE Transactions on Image Processing*, Vol. 4, pp. 430-447, April 1995.

[9] F Hartung and B Girod. Copyright protection in video delivery networks by watermarking of pre-compressed video. *Proc. Multimedia Applications Services and Techniques - ECMAST*, 1997.

[10] C Y Lin and S F Chang. A Robust Image Authentication ALgorithm Surviving JPEG Compression. *Proceedings of the SPIE Conference on Starage and Retrieval of Image/Video Databases*, San Jose CA, Jan 1998.

[11] F A P Petitcolas, R J Anderson and M G Kuhn. Attacks on copyright marking systems *Second Workshop on Information Hiding*, Portland, Oregon, April 1998.

[12] A R Rao, F C Mintzer, G W Braudaway and M Yeung. Digital Watermarking for High-quality Imaging. *Proceedings of the IEEE First Workshop on Multimedia Signal Processing, Princeton NJ*, June 1997.

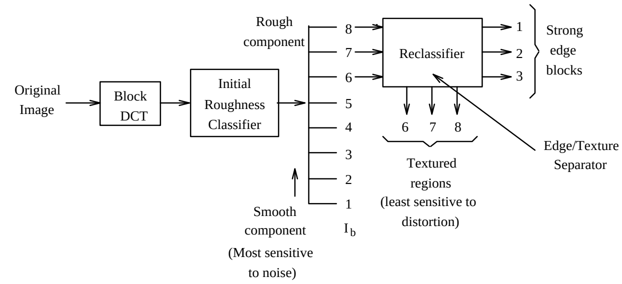

Figure 1: Block diagram for classifying the image sub-block based on noise-distortion sensitivity

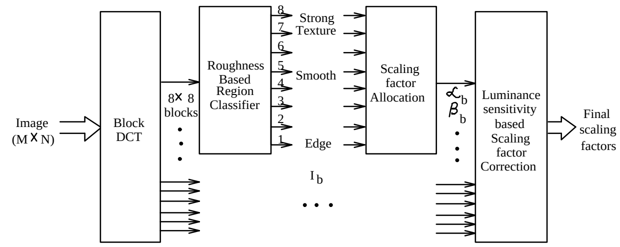

Figure 2: Block diagram for allocating scaling factors for every image block of 8×8 pixels

[13] L Qiao and K Nahrstadt. Watermarking Method for MPEG Encoded Video: Towards Resolving Rightful Ownership. Proceedings of 1998 IEEE International Conference on Multimedia Computing and Systems ICMCS'98, Austin TX, July 1998.

[14] A Piva, M Barni, F Bartolini and V Cappellini. DCT-based Watermark Recovering Without Resorting to Uncorrupted Original Image. *Proceedings of the 1997 IEEE International Conference on Image Processing*, Santa Barbara CA, Oct 1997.

Figure 3: The fog image

Figure 4: Content based classification of the fog image. Highly textured regions are shown in brighter shades of gray and the lesser textured areas are depicted in darker shades of gray. The brightness of the region is inversely proportional to the noise sensitivity of the region.

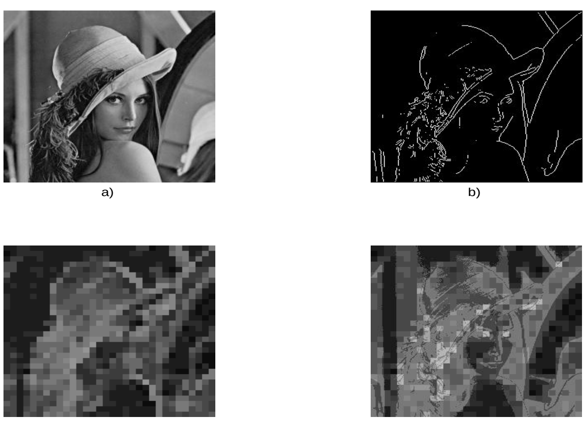

Figure 5: Perceptual region classification of the Lena image. a) Original image b) Prominent edges of the image. c) Classification based on the texture as given by the energy of the DCT AC coefficients. d) Classification based on texture, edge and gray level of the pixel. Bright areas are less sensitive to noise and vice-versa.

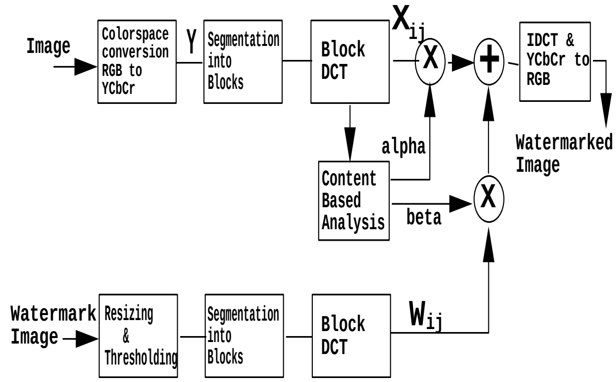

Figure 6: The block diagram of the proposed visible watermarking technique

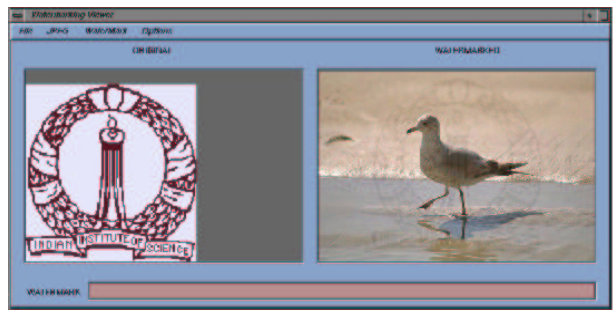

Figure 7: Visible watermark covering a large area of the image. Figures 7 and 8 can be seen properly in full resolution at http://www.comp.nus.edu.sg/~mohan

Figure 8: The logo embedded in the corner of the image