U. Uludag, B. Gunsel and A. M. Tekalp, "Robust watermarking of busy images," *Proceedings of SPIE Electronic Imaging 2001 Conference, Security and Watermarking of Multimedia Contents III*, San Jose, California, USA, January 2001, vol. 4314, pp. 18-25.
Correspondance: uludagum@cse.msu.edu

Robust watermarking of busy images

Umut Uludag[a], Bilge Gunsel[a], and A. Murat Tekalp[b]

[a]Multimedia Systems Group, Information Technologies Research Institute, TUBITAK Marmara Research Center, P.O. Box 21, Gebze, Kocaeli, 41470, Turkey
[b]Department of Electrical and Computer Engineering, University of Rochester, Hopeman 204, Rochester, New York, 14627-0126

ABSTRACT

This work extends the watermarking method proposed by Kutter *et al.* [1] to increase the watermark decoding performance for textured or busy images. The proposed algorithm modifies watermark embedding rule to utilize image characteristics, like local standard deviation and gradient magnitude, in order to increase the decoding accuracy for busy images. The method does not need original image for decoding and controls the watermark embedding process at encoder site, resulting in a more accurate decoding.

Keywords: Adaptive watermarking, data-hiding, digital security.

1. INTRODUCTION

Due to advantages like ease of production, distribution, editing and storing, multimedia data are mainly stored in digital format. Since digital data can be copied without any loss in fidelity, the protection of intellectual property rights poses a very big problem. Digital watermarks, which can be defined as codes imperceptibly embedded in the host multimedia data to carry information like origin, status or destination of the data, have gained considerable attention as a solution for the protection of ownership [2], [3]. A number of watermarking methods including space, frequency and transform domain techniques are developed in the last decade ([4], [5], [6], [7]).

This work extends the watermarking method proposed by Kutter *et al.* [1] in a way to increase the watermark decoding performance for textured or busy images. The original watermarking method of Kutter *et al.* [1] embeds watermarks to blue channels of images via amplitude modulation. A method describing a procedure for more robust watermark decoding after the watermarked image has undergone generalized geometrical transformations is given in [8]. The watermark embedding locations are determined by a secret key, and the blue channel pixel values at these locations are decreased or increased to produce watermark bits 0 and 1, respectively. The watermark decoding does not need the original image. The original pixel values at watermark embedding locations specified by the secret key are estimated as a linear combination of neighbor pixels. The sign of the difference between watermarked and estimated pixel values indicates the watermark bit. Namely, if this difference is negative, than the embedded watermark bit is 0; if it is positive, the watermark bit is 1. The higher the absolute value of this difference is, the more accurate will be the bit decoding. However, the estimation of original pixel values at watermark embedding locations does not give accurate results for textured or busy images and leads to incorrect decoding of digital data. This is mainly due to fact that in such image areas local variations in pixel values are high, and the estimation process, which can be thought of as a mean operation, is not adequate to obtain correct pixel values. This paper extends the original watermarking method introduced in [1] to overcome problems encountered in decoding.

Correspondance: U. Uludag. Other author information: B.G.: Email: gunsel@btae.mam.gov.tr; A.M.T.: Email: tekalp@ece.rochester.edu; U.U.: Email: umutu@btae.mam.gov.tr

The organization of the paper is as follows: Section 2 describes the modified watermarking method. Experimental results are given at Section 3. Section 4 concludes the paper.

2. ADAPTIVE WATERMARKING METHOD

In the proposed method, the watermark embedding rule is modified by adding two new terms which adapt the embedding process to the original data. The modified watermark bit embedding rule is

$$b_{WM}(i,j) = b(i,j) + (2s-1)L(i,j)q\left(1+\frac{SD(i,j)}{A}\right)\left(1+\frac{GM(i,j)}{B}\right) \quad (1)$$

where $b_{WM}(i,j)$ and $b(i,j)$ are gray values referring to watermarked and original blue channel pixels at watermark embedding location (i,j), respectively. The value of watermark bit is denoted as s. $L(i,j)$ is the luminance value at (i,j). Watermark embedding strength is denoted as q. The remaining two terms on the right hand side of the equation correspond to terms introduced by our method. $SD(i,j)$ denotes the standard deviation of pixel values around a local neighborhood of pixel at (i,j), and $GM(i,j)$ denotes the gradient magnitude at (i,j). A and B are normalization factors for the standard deviation and gradient magnitude, respectively.

$L(i,j)$ term is calculated as

$$L(i,j) = 0.299R(i,j) + 0.587G(i,j) + 0.114B(i,j) \quad (2)$$

where $R(i,j), \mathrm{G(i,j)}, \mathrm{B(i,j)}$ denote red, green and blue channel values at location (i,j).

Every watermark bit with value s in Eq.(1) is embedded multiple times to the blue channel pixels, whose locations are determined via the selected secret key. In addition to the real watermark data, two reference bits, 0 and 1, are embedded to the image. These reference data provides an adaptive threshold in determining the watermark bit value in decoding.

Decoding starts with finding the watermark embedding locations on the watermarked image, via the secret key used in watermark encoding stage. For every bit embedding location, the value of the original blue channel pixel, $\hat{b}(i,j)$, is estimated as the linear combination of pixels in a cross-shaped neighborhood of the watermarked pixel as follows:

$$\hat{b}(i,j) = \frac{1}{4c}\left(\sum_{k=-c}^{c} b_{WM}(i+k,j) + \sum_{k=-c}^{c} b_{WM}(i,j+k) - 2b_{WM}(i,j)\right) \quad (3)$$

where c is the neighborhood size. The difference between the estimated and current pixel values is calculated as

$$\delta = b_{WM}(i,j) - \hat{b}(i,j) \quad (4)$$

These differences are averaged over all the embedding locations associated with the same bit, to yield $\bar{\delta}$. For finding an adaptive threshold, these averages are calculated similarly for the reference bits, 0 and 1, as $\bar{\delta}_{R0}$ and $\bar{\delta}_{R1}$, respectively.

Then, the watermark bit value $\hat{s}$ is estimated as

$$\hat{s} = \begin{cases} 1 & if\ \bar{\delta} > \dfrac{\bar{\delta}_{R0} + \bar{\delta}_{R1}}{2} \\ 0 & otherwise \end{cases} \quad (5)$$

Standard deviation term $SD(i,j)$ can be computed as the standard deviation of the set containing the pixel values in a cross-shaped neighborhood of the watermark bit embedding location (i,j). Gradient magnitude term $GM(i,j)$ can be computed in a number of ways. Sobel operator is used in our work.

These additional terms in Eq. (1) adjust the strength of watermarking in an image adaptive way. For smooth image areas, they converge to zero. As can be seen from Eq. (1), this corresponds to the case where $SD(i, j)$ and $GM(i, j)$ terms do not have any effect on watermark embedding strength. On the other hand, if $SD(i, j)$ term is high, corresponding to an image region with high variance, $|b(i, j) - b_{WM}(i, j)|$, the absolute difference between the original and watermarked gray values will be higher, resulting in a higher accuracy in watermark decoding. Because accuracy of watermark decoding depends on this strength (Eq. (3), Eq. (4)). Same effect arises from $GM(i, j)$, namely for image regions with high gradient magnitude, corresponding to edge regions, the watermark embedding strength will be higher. Note that, although watermark decoding accuracy is increased as a result of the image adaptive increase in embedding strength, due to the fact that human visual system (HVS) is relatively less sensitive to pixel value changes in busy and edge image regions, the visibility of the watermark does not increase significantly. Also, embedding watermark data to blue channels of images decreases the visibility.

The watermarking method described does not need the original image for decoding. This fact increases the applicability of the method to some real world applications where the original image is only available to the copyright owner and the users can not (and must not) reach the original. On the other hand, watermarking methods using original data for watermark decoding can be more robust to attacks.

To further increase the watermark decoding accuracy, in addition to using image-adaptive new terms in calculating watermark embedding strength, controlling the watermark embedding process in the watermark encoder is utilized in the proposed method. The encoder uses a feedback loop for this control . The basic structure of the watermark encoder is given in Figure 1. As shown in this figure, the accuracy of the watermark decoding is checked by controller, during watermark embedding. If this analysis yields the result that the watermark decoding will be incorrect, the encoder adaptively adjusts the watermark embedding strength until correct decoding is guaranteed. Note that the algorithm also considers the invisibility criterion.

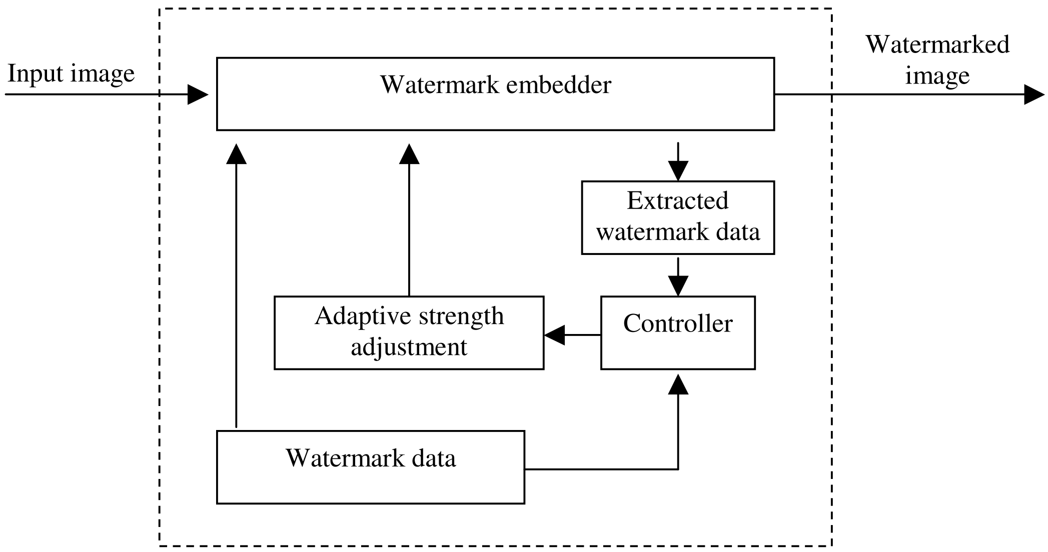

Figure 1. Watermark encoder structure.

In the proposed method, the feedback loop within the encoder changes the watermark strength q. For every bit embedding location, the term δ is calculated according to Eq. (3), with the initial q value. Than, if δ is calculated as a negative value for an embedded bit of "1", the value of q is increased until correct decoding is guaranteed, namely until δ is positive. Similarly, if the term δ is calculated as a positive value for an embedded bit of "0", the value of q is again increased until correct decoding is guaranteed, namely δ is negative. Furthermore, increasing the number of watermark embedding points, until the image capacity is reached, may improve the accuracy of watermark decoding, with the drawback of increased visibility.

3. EXPERIMENTAL RESULTS

In order to explore the performance of introduced adaptive watermarking method and the original method proposed in [1], a number of experiments are performed on color images. Watermark data are embedded into blue channels of images (512 by 512, 24 bit, color). The embedded bits correspond to the binary representation of the text data shown in Figure 2. The sample text consists of 463 characters.

> Due to advantages like ease of production, distribution, editing and storing, multimedia data are mainly in digital form. Since digital data can be copied without any loss in fidelity, the protection of intellectual property rights poses a very big problem. Digital watermarks, which can be defined as codes imperceptibly embedded in the host multimedia data to carry information like origin, status or destination of the data, have gained considerable attention.

Figure 2. Text representation of embedded watermark data.

The images[1] used in the experiments are shown in Figure 3. *New York* and *Baboon* images are examples for busy images; whereas *Lena* and *Sailboat* images are examples for relatively smooth images.

New York Baboon

Lena Sailboat

Figure 3. Images used in the experiments.

[1]New York image: Copyright photo courtesy of Patrick Loo, University of Cambridge; others are from USC-SIPI Database.

Watermark data decoding performance of the methods, measured as percentage of accurately decoded characters, is presented in Table 1. To better evaluate the efficiency of the proposed method, the tests are performed on images which are watermarked by a) the method in [1] (column 1), b) the introduced adaptive method, including only standard deviation (SD) term in Eq. (1) (column 2), c) the introduced adaptive method, including only gradient magnitude (GM) term in Eq. (1) (column 3), d) the introduced adaptive method, including both of the image adaptation terms (column 4) and e) the proposed method, including both of the image adaptation terms and controller at encoder (column 5).

The parameters used in the experiments are set to: initial $q = 0.1$, $A = 100$, $B = 1000$. The embedded watermark data size is 3243 bits. Totally, 40 % of the pixels in the image are modified in watermark embedding.

Table 1. The watermark data decoding performance (%).

Image	Method in [1]	Only SD	Only GM	SD+GM	SD+GM+ Controller
New York	73.87	90.5	83.59	94.82	99.78
Baboon	91.79	96.98	95.9	99.35	100
Lena	99.57	99.57	99.57	99.57	99.78
Sailboat	98.49	99.35	99.14	99.35	99.78

As can be seen from Table 1, especially the decoding performance for busy images, *New York* and *Baboon*, increases considerably by using image adaptive and controlled watermark embedding. For relatively smooth images, *Lena* and *Sailboat*, this increase is small and the performance of the method in [1] is nearly same as that of modified method. Using image adaptation terms individually, as shown in columns 2 and 3, also increases performance but using both of them and controller leads to highest performance.

The normalization factors, A and B , control the effect of standard deviation and gradient magnitude on watermarking, respectively. Decreasing A or B increases effects of standard deviation and gradient magnitude, respectively; while increasing A or B decreases effects of standard deviation and gradient magnitude, respectively.

For evaluating the visibility of the embedded watermarks, Figure 4 shows the luminance and blue channel images of i) original *New York* image, ii) image watermarked by the method in [1] and iii) image watermarked by the proposed method, including both of the image adaptation terms and controller. Similarly, Figure 5 shows the luminance and blue channel images of i) original *Lena* image, ii) image watermarked by the method in [1] and iii) image watermarked by the proposed method, including both of the image adaptation terms and controller. The figures indicate that the watermarks are invisible for the proposed method and the method in [1].

The method's resistance to blurring, JPEG encoding/decoding, rotation and image composition attacks is reported in [1]. Furthermore, we evaluated watermark data decoding performance of the adaptive watermarking method for cropped images. For this purpose, halves of the watermarked images are cropped. The performance results obtained from the remaining images are presented in Table 2.

Table 2. The watermark data decoding performance (%) after cropping attack.

Image	SD+GM+ Controller
New York	93.74
Baboon	95.03
Lena	93.52
Sailboat	84.02

Note that these performances will increase further if a smaller size data is embedded to the images. In that case, every bit will be embedded to more locations and the effect of cropping on the performance will reduce.

a)

b)

c)

Figure 4. Luminance (left column) and blue channel (right column) images of a) original *New York* image, b) image watermarked by the method in [1], c) image watermarked by the proposed method.

a)

b)

c)

Figure 5. Luminance (left column) and blue channel (right column) images of a) original *Lena* image, b) image watermarked by the method in [1], c) image watermarked by the proposed method.

4. CONCLUSIONS

An image adaptive watermarking method, which extends the method in [1], is presented. By utilizing standard deviation and gradient magnitude properties of the image regions in watermark embedding and by controlling watermark embedding process for correct decoding, the watermark data decoding performance has been increased, especially for textured or busy images. The embedded watermarks are invisible. In addition to blurring, JPEG coding/decoding, rotation and image composition attacks, it is shown that image cropping does not affect watermark data decoding performance considerably.

REFERENCES

1. M. Kutter, F. Jordan, and F. Bossen, "Digital signature of color images using amplitude modulation," *Proc. of SPIE-EI97*, Feb. 1997, pp. 518-526.
2. F. Hartung and M. Kutter, "Multimedia watermarking techniques," *Proceedings of the IEEE*, vol. 87, no. 7, July 1999, pp. 1079-1107.
3. M. D. Swanson, M. Kobayashi, and A. H. Tewfik, "Multimedia data embedding and watermarking technologies," *Proceedings of the IEEE*, vol. 86, no. 6, June 1998, pp. 1064-1087.
4. N. Nikolaidis and I. Pitas, "Robust image watermarking in the spatial domain," *Signal Processing*, vol. 66, 1998, pp. 385-403.
5. C. I. Podilchuk and W. Zeng, "Image-adaptive watermarking using visual models," *IEEE Journal on Selected Areas in Communications,* vol. 16, no. 4, May 1998, pp. 525-539.
6. J. F. Delaigle, C. De Vleeschouwer and B. Macq, "Watermarking algorithm based on a human visual model," *Signal Processing*, vol. 66, 1998, pp. 319-335.
7. S. W. Kim, S. Suthaharan, and K. R. Rao, "Image watermarking scheme using visual model and BN distribution," *Electronics Letters*, vol. 35, no. 3, February 1999, pp. 212-214.
8. M. Kutter, "Watermarking resisting to translation, rotation and scaling," *Proc. of SPIE: Multimedia Systems and Applications*, Nov. 1998, vol. 3528, pp. 423-431.