Content-Based Watermarking for Image Authentication Using Independent Component Analysis

Dr. Latha Parameswaran
Professor, Department of Computer Science & Engineering,
AMRITA University, Coimbatore – 641 105, India
E-mail: lathapcp@yahoo.co.in

Dr. K. Anbumani
Former Director, Karunya School of Computer Science and Technology,
Karunya University, Coimbatore –641114, India
E-mail: anbumani_k@yahoo.co.uk

Keywords: digital watermarking, independent component analysis, discrete cosine transform

Received: September 19, 2007

This paper proposes a novel approach to content-based watermarking for image authentication that is based on Independent Component Analysis (ICA). In the scheme proposed here, ICA is applied to blocks of the host image and the resulting mixing matrix represents the features of the image blocks. Frobenius norm of the mixing matrix is adopted as the content-based feature. This is embedded as the watermark in a mid-frequency DCT coefficient of the block. This authentication technique is robust against incidental image processing operations, but detects malicious tampering and correctly locates the tampered regions.

Povzetek: Predlagana je nova metoda avtentikacije slik.

1 Introduction

A **digital watermark** is a piece of information that is hidden in a multimedia content, in such a way that it is imperceptible to a human observer, but easily detected by a computer. The principal advantage is that the watermark is inseparable from the content [1]. Digital watermarking is the process of hiding the watermark imperceptibly in the content. This technique was initially used in paper and currency as a measure of authenticity.

The primary tool available for data protection is encryption. Encryption protects content during the transmission of the data from the sender to receiver. However, after receipt and subsequent decryption, the data is no longer protected. Watermarking complements encryption [1].
Digital Watermarking involves two major phases:

(i) Watermark embedding, and
(ii) Watermark extraction.

Digital watermarks can be a pseudo random sequence or a logo of a company or an image. Watermark embedding is done in the watermark carriers such as Discrete Cosine Transform (DCT) or Discrete Wavelet Transform (DWT), etc of the original data resulting in watermarked data. The watermarked data may be compressed to reduce its size, corrupted by noise during its transmission through a noisy channel. It may be subjected to other normal image processing operations such as filtering, histogram modification etc. Also malicious intruders may tamper the data.

DCT is a widely used technique for watermarking [1]. Recently ICA is being adopted for watermarking [2] – [8]. In [2] ICA is applied to the blocks of the host image and the watermark image. The least-energy independent components of the host are replaced by the high-energy independent components of the watermark image. For watermark extraction the demixing matrices of both the watermark and the host images are required.

Dan Yu et al. [3] treat the host image, the key image, and the watermark image as the independent sources. Embedding is done by weighted addition of the key and the watermark to the host. For watermark extraction, two more mixtures are obtained by adding the key and the watermark using different weights. ICA is then applied to these mixtures to separate the host, the key, and the watermark. The host and the key are required for watermark extraction. In [4] the same procedure as in [3] has been used. The only difference is in the algorithm used for ICA.

Ju Liu et al. [5] use ICA for detection of the watermark which is a random sequence embedded in low-frequency DCT coefficients. Original DCT coefficients are required for watermark detection and for creating a second mixture needed for ICA.

Bounkong et al. [6] apply ICA to each block of the host image and obtain its independent components. The watermark is embedded in selected components using quantization and a modified image block is obtained

from these modified independent components. This is added to the host image block, obtaining the watermarked image. In the extraction phase, ICA is applied to each block obtaining the independent components. The watermark is then extracted from these through dequantization.

The technique of creating three mixtures is also employed in [7] and [8]. While [7] uses upsizing and downsizing, [8] uses the so-called redundant DWT (RDWT).

Many authors have worked on content-based watermarking for image authentication. Li et al. [14] discuss a content-based watermarking scheme that uses local features of the image such as edges and zero-crossings. Their scheme uses a look-up table to embed the watermark and the same table is required at the receiver end to extract the watermark.

Content-based watermark is generated based on salient features of the image either in spatial domain like edges, texture, fractal dimensions [15] etc. or in a transform domain such as singular values [16], eigenvalues [17], etc. Choices of image features vary with techniques and directly influence the robustness of the scheme. Some techniques generate a random binary sequence to embed the watermark based on the features of the images [18] and [19]. A content-based digital signature scheme for image authentication has been presented in [20].

In [21] - [24] a localization based method has been presented to verify the integrity of the received image. In these techniques the host image is divided into a number of disjoint blocks and watermark is embedded in each of these blocks. To verify the authenticity of the received image, blockwise authentication has been done.

In [18], [25], [26] image authentication has been done using content-based watermarks. But these schemes do not embed the watermark in the image content; instead embed them in the image header. These techniques distort the host image prior to watermark embedding.

In [27] a watermarking technique based on the quadtree is proposed. This scheme embeds a Gaussian sequence watermark into low-frequency band of the wavelet transform. In their technique, watermark is embedded into visually insensitive pixels in quadtrees.

Most of the above authors embed a specified watermark. Most have copyright protection as their goal and require a lot of information about the host image for watermark extraction. Additional image mixtures are artificially created and then ICA is used as a blind source separation technique to separate the host image, the watermark, and the key.

In the scheme proposed here, a different approach to the use of ICA is adopted. ICA is used to determine the mixing matrix which – specifically its Frobenius norm – represents the content of the host image. No information about the host is required for watermark extraction. Thus the proposed scheme is a novel, blind, content-based watermarking for content authentication that uses ICA and DCT.

2 Proposed content-based watermarking using ICA

2.1 Principles of ICA

Principles of ICA are discussed in [9]–[13]. Let x be a random vector of observations. ICA models the observed data as: x = As, where the vector s represents the independent sources that generate the observed data x. The matrix A is the mixing matrix. Each x_j is a linear combination of the independent source signals. Hence xi is called the mixed signal. ICA estimates A and s from given x.

In the context of images, each row is considered as one observation x. Thus the entire image is denoted as X, representing all the rows of the image. ICA models an image X as: X = AS.

In the general case, X is m x n and S is r x n, where $m \geq r$. In other words, the number of observed mixed signals must be greater than or equal to the number of independent components. Some significant points of ICA are:

- Each component s_j is independent of every other component.
- All independent components are non-Gaussian (with possible exception of one).
- The observation x is assumed centered.
- There is no order specified among the independent components.
- The independent components can be obtained from the observations as, s=Wx, where W is called the demixing matrix, estimated in ICA.
- Each column of the mixing matrix A represents a feature of the data x.

The last mentioned property is made use of in this proposed technique using ICA. Each column of the mixing matrix A represents a feature of the image. So, all the columns together represent all the features of the image. In order to get one single quantity to represent the image, the Frobenius norm of A has been chosen as the content-based feature that represents the image. The Frobenius norm of a matrix is the square root of the sum of the norms of all the columns of A.

The proposed content-based watermarking for blind authentication uses a hybrid of ICA and DCT. The host image is divided into small blocks. ICA is applied to each block and the mixing matrix of the block is determined. Frobenius norm of the mixing matrix is computed. This is considered as the content-based feature of the block. Such features are obtained for all the blocks of the host image. These constitute the content-based watermark used for authentication of the image.

In this technique, the watermark is embedded by replacing the chosen mid-frequency coefficient DCT(p, q) with a scaled value of the watermark but retaining the sign of the DCT coefficient, i.e.,

$$DCT(p,q) = sign(DCT(p,q)) * (\alpha * w)$$

where w is the content-based watermark. Here the watermark is the Frobenius norm of a block, which is always non-negative. The value for α is chosen based on the statistical details of the DCT coefficients and the watermark.

Watermark extraction is the reverse of the embedding process. The received watermarked image is divided into blocks and ICA is applied to each of them. The Frobenius norm of the mixing matrix is computed. DCT of each block is also performed. The watermark that was embedded is extracted from the chosen mid-frequency coefficient:

$$ExtractedWatermark = \frac{|DCT(p,q)|}{\alpha}$$

The percentage difference (Δ) between the extracted and the embedded watermarks is computed. If it is high, it indicates that the image has been tampered. Details of steps for the three phases of watermarking are given below

i. Watermark generation,
ii. Watermark embedding, and
iii. Watermark extraction and authentication.

Watermark generation

1. Segment the host image I of size n x n into blocks of size m x m resulting in K blocks.
2. Perform ICA of each block treating each row of the block as a vector.
3. Extract the mixing matrix A.
4. Compute the Frobenius norm of the mixing matrix; this is the content-based watermark w of the block.
5. Repeat steps 2 – 4 for computing the watermark for all the blocks. This set forms the watermark, $W = \{w_1, w_2, ..., w_k\}$

Watermark embedding

1. Perform DCT of each block.
2. Select the mid-frequency coefficient at the chosen location (p, q) in each block.
3. Replace the chosen coefficient with the watermark:

$$DCT(p,q) = sign(DCT(p,q)) * (\alpha * w)$$

4. Perform inverse DCT.
5. Repeat steps 1– 4 for all the blocks.

The resultant is the watermarked image I*.

Watermark extraction and authentication

1. Perform steps 1–5 of the watermark generation procedure on the received image I' and obtain the computed watermark, .
2. Perform DCT of each block.
3. Extract the embedded watermark from the chosen DCT coefficient:

$$w' = \frac{|DCT(p,q)|}{\alpha}$$

4. This set forms the extracted watermark,

$$W' = \{w'_1, w'_2, ..., w'_K\}$$

5. Calculate the blockwise percentage difference (Δ) between the watermark values w* and w':

$$\Delta = \frac{|w_i^* - w_i|}{\max\{w_i\}} * 100$$

In this scheme the percentage difference of the values corresponding to each block is used to detect any change in the block and thereby the authenticity of the image. If the difference is small – smaller than an experimentally chosen threshold value – the block and therefore the entire image is deemed authentic. If the difference of any block is greater than the threshold, that block is identified as the tampered block and hence the image is unauthentic. .

3 Experimental results

The proposed blind content-based watermarking scheme for image authentication has been tested using Matlab and Adobe Photoshop. The scheme has been evaluated on a set of three different categories of 512 x 512 gray scale images: (i) standard images, (ii) natural images, and (iii) images created using imaging tools.

Choice of parameters

In order to determine the block size for image segmentation various block sizes were tried. In [6] Bounkong has mentioned that choosing a block size is based on the processing time and relevant features. Blocks of small size leads to poor performance in watermarking process and larger blocks demand high computational time. Hence a trade off between these two is required to choose the block size. After experimentation, a block size of 16 x 16 was chosen as it resulted in better PSNR value, computational time and better feature representation.

In order to embed the watermark in a suitable location the proposed technique uses one of the mid-frequency coefficients. Embedding the watermark in low-frequency components, results in visual degradation of the host image. Similarly embedding the watermark data in high-frequency components is not advisable as they may be lost during compression. Hence embedding the watermark in mid-frequency components ensures robustness.

The mid-frequency coefficient (p, q) in which to embed the watermark is chosen as the mid-diagonal coefficient i.e. the location

$$\left(\frac{blocksize}{2}, \frac{blocksize}{2}\right)$$

For choosing a suitable value for the embedding strength α, statistics of the DCT coefficient values at that mid-diagonal location of all the blocks are obtained, specifically the standard deviation α_x. Similarly the standard deviation α_w is obtained for the watermark.

The value of embedding factor α is determined such that the watermark values are suitably scaled to have the same range of variation as that of the DCT coefficients: $\alpha = \frac{\alpha_x}{\alpha_w}$. In this experimentation after computation the value is α = 0.14.

Threshold for the percentage difference Δ between the watermarks has been experimentally determined as 15%. Lower thresholds resulted in false negatives; while higher thresholds made the technique to be fragile.

The ICA algorithm adopted in this proposed technique is the fastICA algorithm. This algorithm has been discussed in [9] and [13].

Quality of the watermarked image

The proposed content-based watermarking scheme has been implemented on a set of images of three categories. The metrics PSNR, Pearson Correlation Coefficient (PCC), Normalized Cross Correlation (NCC), and Image Fidelity (IF) are calculated between the host image and the watermarked image.

The test images after watermarking is shown in Fig. 1. It can be observed that there is no perceptually noticeable difference in the images due to watermarking.

Numerical values of the performance metrics for the test images given in Table 1 also corroborate this. PSNR values range from 96.09 to 102.45, with an average of 97.87 for all the test images**,** which is quite high. The other metrics, Pearson Correlation Coefficient (PCC) and Image Fidelity (IF) are also quite high. This shows that watermark embedding does not degrade the visual quality of the image.











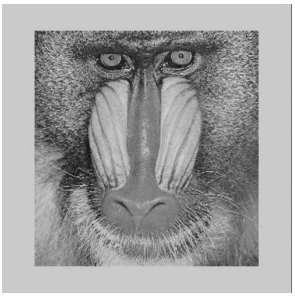


Figure 1: Images after watermarking using the ICA technique.

Extraction efficacy

The efficiency of the scheme in correctly extracting the watermark is given by the percentage difference between the computed and extracted Frobenius norm of the mixing matrix of the received image blocks. Table 2 gives the highest percentage difference Δ for some of the test images. The values are small, ranging from 1.87 to 9.60%, over all the test images**.** This indicates that the scheme extracts the embedded watermark accurately.

Table 1: Quality metrics after watermarking using ICA

	PSNR	IF	NCC	PCC
Bridge	98.8892	1.0000	1.0000	0.9999
Boy	98.6417	0.9999	1.0000	0.9999
Building	96.7310	0.9999	1.0000	0.9999
Cameraman	97.4249	1.0000	1.0000	0.9999
Clown	98.2414	0.9999	1.0000	0.9999
Couple	96.7900	0.9999	1.0000	0.9999
Jet Plane	98.6468	1.0000	1.0000	0.9999
Lena	98.0096	0.9999	1.0000	0.9999
Living Room	98.7276	1.0000	1.0000	0.9999
Mandrill	98.4030	1.0000	1.0000	0.9998
Peppers	96.0968	0.9999	1.0000	0.9998
Sail Boat	96.4620	1.0000	1.0000	0.9999
Bulb	102.4517	1.0000	1.0000	1.0000
Snow Tree	97.8179	1.0000	1.0000	0.9999
Specs	97.8651	1.0000	1.0000	0.9998
Trees	97.3633	0.9999	0.9999	0.9997
KeyClock	96.2998	1.0000	1.0000	0.9998
SunBark	96.8414	1.0000	1.0000	0.9999
Decor	97.2312	0.9999	0.9999	0.9989
Lamp	96.1245	0.9989	0.9999	0.9978
Average	97.7529	0.9999	1.0000	0.9997
Minimum	96.0968	0.9989	0.9999	0.9978
Maximum	102.4517	1.0000	1.0000	1.0000

Robustness against incidental image processing

Robustness of the proposed scheme against normal signal processing operations such as compression, noise and filtering has been experimentally evaluated on all the test images.

In this proposed watermarking technique the watermarked image is subjected to three types of distortions: compression, noise, and filter. Watermarked image has been compressed using JPEG compression with different quality factors. Additive white Gaussian noise (AWGN) and uniform noise has been added to the watermarked image. Also filtering such as low pass, sharpening, histogram equalization, and contrast stretching has been applied on the watermarked image. Results of the test image Lena is shown in Table 3.

For all these attacks, the values of highest percentage difference Δ, ranges from 1.25 to 6.45 with an average of 3.23 as given in Table 3. All the percentage differences all less than the determined threshold 15%, indicating that there is no tampering.

Table 2: Results after watermark extraction without attacks using ICA

Image	Highest percentage difference
Bridge	9.1100
Boy	8.3440
Building	8.5009
Cameraman	4.5248
Clown	8.6273
Couple	8.2430
Jet Plane	7.5966
Lena	8.9354
Living Room	9.2842
Mandrill	6.6092
Peppers	7.9915
Sail Boat	8.7662
Bulb	1.9682
Snow Tree	8.7526
Specs	8.2906
Trees	1.8791
KeyClock	9.6035
SunBark	8.2667
Decor	7.2345
Lamp	7.1124
Average	7.4820
Minimum	1.8791
Maximum	9.6035

Similar good performance of robustness of the proposed scheme has been obtained for other test images also. For example, Fig. 2 shows the robustness of the various test images against:

- JPEG compression medium quality (Quality factor = 5)
- AWGN with noise 5%
- Low pass filter with standard deviation 10.

Table 3: Results after incidental distortions on Lena using ICA

Attacks	Parameters	Highest percentage difference
JPEG Compression		
Maximum	Quality Factor =10	3.5728
High	Quality Factor =8	2.1245
Medium	Quality Factor =5	2.0000
Low	Quality Factor =3	1.9925
Noise		
AWGN	Percent = 5	6.4512
Uniform	Percent = 5	3.1928
Filter		
Low pass	Standard Deviation = 10	4.2686
Sharpening	-	3.2578
Histogram Equalization	-	4.2564
Gamma Correction	Gamma value = 3	3.2578
Contrast Stretching	Brightness = 15 Contrast = 15	1.2456
Average		3.2382
Maximum		6.4512
Minimum		1.2456

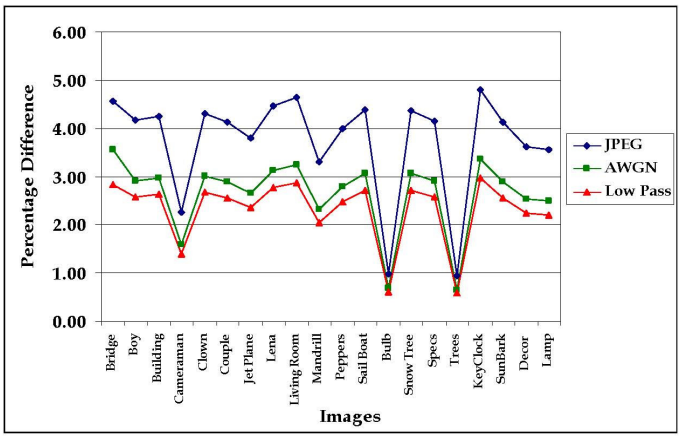

Figure 2: Robustness of ICA-based technique after incidental image processing

Detection of tampering

To demonstrate the ability of the proposed scheme in locating the blocks that have been intentionally tampered, the watermarked Lena image has been intentionally tampered by introducing small patches. In one case, a text has been placed at the top left corner of Lena, Fig. 3 (a), and the method correctly located the tampered blocks as shown in Fig. 3 (b). Patches were introduced at other points on Lena. In all the cases this technique correctly identified malicious tampering.

Figure 3: Lena with location of tampered blocks.

Table 4: Results after intentional tamper of Lena using ICA

Locations of tamper	Blocks identified as tampered	Percentage difference of tampered blocks
Top	(1,1) and (1, 2)	55 and36
Hat	(3, 18), (3, 19), and (3, 20)	27, 30, and 29
Shoulder	(30,51), (30, 52), (31,53), (31,51), and (31,52)	41, 61, 36, 35, and 31
Opposite to hat	(4,63), (4,64) (5,63), and (5,64)	31, 69, 32, and 65

4. Conclusion

This paper has discussed a new blind content-based watermarking scheme for image authentication using ICA and DCT. The watermark to be embedded is obtained from the host image itself in terms of the Frobenius norm of the mixing matrix obtained during ICA. These are embedded in the mid-frequency DCT coefficients. The proposed method correctly authenticates the image even under normal image processing operations and it correctly detects tampering and identifies the tampered regions of the image. The major quality of the watermarked image, this proposed technique is much superior to the techniques in the literature. The average PSNR value in the existing techniques is around 52.45; whereas the average PSNR value in this proposed ICA based technique is around 97. In terms of computation time, the ICA based technique takes a longer time. This is mainly due to the algorithms used to compute independent components. Exhaustive experimentation demonstrates the efficacy of the proposed scheme.

References

[1] Ingermar J. Cox, Matthew L. Miller, and Jeffrey A. Bloom, Digital Watermarking, Morgan Kaufmann Publishers, 2002.

[2] Francisco J. Gonzalez-Serrano, Harold. Y. Molina-Bulla, and Juan J. Murillo- Fuentes," Independent component analysis applied to digital image watermarking," International Conference on Acoustic, Speech and Signal Processing (ICASSP), vol. 3, pp. 1997-2000, May 2001.

[3] Dan Yu, Farook Sattar, and Kai-Kuang Ma, "Watermark detection and extraction using independent component analysis method," EURASIP Journal on Applied Signal Processing, vol. 1, pp. 92–104, 2002.

[4] Minfen Shen, Xinjung Zhang, and Lisha Sun, P. J. Beadle, F. H. Y. Chan, "A method for digital image watermarking using ICA," 4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA 2003), Nara, Japan, April 2003, pp. 209-214.

[5] Ju Liu , Xingang Zhang, Jiande Sun, and Miguel Angel Lagunas, "A digital watermarking scheme based on ICA detection," 4th International Symposium on Independent Component Analysis and Blind Signal Separation, (ICA 2003), Nara, Japan, April 2003, pp. 215-220.

[6] Stephane Bounkong, Boremi Toch, David Saad, and David Lowe, "ICA for watermarking digital images," Journal of Machine Learning Research 4, pp. 1471-1498, 2003.

[7] Viet Thang Nguyen and Jagdish Chandra Patra, "Digital image watermarking using independent component analysis," PCM 2004, Lecture Notes in Computer Science 3333, pp. 364-371, Springer-Verlag, 2004.

[8] Thai Duy Hien, Zensho Nakao, and Yen-Wei Chen, "Robust multi-logo watermarking by RDWT and ICA", Signal Processing, Elsevier, vol. 86, pp. 2981-2993, 2006.

[9] Aapo Hyvarinen, "Survey on Independent Component Analysis", Neural Computing Surveys, vol. 2, pp. 94-128, 1999.

[10] Hyvarinen, Karhunen, and Oja, "Introduction," Chapter 1 in Independent Component Analysis, John Wiley, pp. 1-12, 2001.

[11] Errki Oja, "Independent Component Analysis: Introduction," European Meeting on ICA, Vietri sul Mare, Feb. 21, 2002.

[12] Lucas Parra, "Tutorial on blind source separation and independent component analysis," Adaptive Image and Signal Processing Group, Sarnoff Corporation, Feb. 9, 2002.

[13] Bogdan Matei, "A review of independent component analysis technique," Tutorial, Electrical and Computer Engineering Department, Rutgers University, Piscataway, NJ, USA.

[14] Chang-Tsun Li, Der-Chyuan Lou, and Tsung-Hsu Chen, "Image authentication and integrity verification via content-based watermarks and a public key cryptosystem," Proceedings of International Conference on Image Processing, vol. 3, 2000 pp. 694 - 697

[15] Rongrong Ni, Quiqi Ruan, and H.D. Cheng, "Secure semi-blind watermarking based on iteration mapping and image features," Pattern Recognition, vol. 38, pp. 357-368, 2005.

[16] Herve Abdi, "Singular value decomposition and generalized singular value decomposition," in Neil Salkind (Ed.), Encyclopedia of Measurement and Statistics, 2007.

[17] Arto Kaarna, Pekka Toivnen, and Kimmo Mikkonen, "Watermarking spectral images through PCA transform," Proceedings of PICS, The Digital Prography Conference, May 2003, pp. 220-225.

[18] Chai Wah Wu, "On the design of content-based multimedia authentication systems," IEEE Transactions on Multimedia, vol. 4, no.3, pp. 385-393, September 2002.

[19] Eugene T. Lina, Christine I. Podilchuk, and Edward J. Delp, "Detection of image alterations using semi-fragile watermarks," Proceedings of SPIE International Conference on Security and watermarking of Multimedia contents, January 2000.

[20] Marc Schneider and Shih-Fu Chang, "A robust content based digital signature for image authentication," Proceedings of International Conference on Image Processing, vol. 3, September 1996, pp. 227-230.

[21] Roberto Caldelli, Franco Bartiloni, and Vito Cappellini, "Standard metadata embedding in a digital image," Proceedings of 14th International Workshop on Database and Expert Systems Applications, 2003.

[22] M.G. Albanesi, M. Ferretti, and F. Guerrini, "A taxonomy for image authentication techniques and its application to the current state of the art," Proceedings of the 11th IEEE International Conference on Image Analysis and Processing (ICIAP '01), 2001.

[23] Phen-Lan Lin, Po-Whei Huang, and An-Wei Peng, "A fragile watermarking scheme for image authentication with localization and recovery," Proceedings of the IEEE Sixth International Symposium on Multimedia Software Engineering (ISMSE'04), 2004.

[24] Huijuan Yang and Alex C. Kot, "Binary image authentication with tampering localization by embedding cryptographic signature and block identifier," IEEE Signal Processing Letters, vol. 13, no. 12, pp. 741-744, December 2006.

[25] Nasir Memon, Poorvi Vora, Boon-Lock Yeo, and Minerva Yeung, "Distortion bounded authentication techniques," Proceedings of International Conference on Watermarking and Multimedia Contents, February 2000.

[26] Chai Wah Wu, "Limitations and requirements of content-based multimedia authentication systems," Proceedings of International Conference of SPIE, vol. 4314, 2001, pp. 241-252.

[27] Kil-Sang Yoo, Mi-Ae Kim, and Won-Hyung Lee, "A robust image watermarking technique for JPEG images using quadtrees," Lecture Notes in Computer Science, vol. 3332, pp. 34-41, 2004.