

# A DCT-Based Secure JPEG Image Authentication Scheme

Mona F. M. Mursi, Ghazy M.R. Assassa, Hatim A. Aboalsamh, and Khaled Alghathbar

**Abstract**—The challenge in the case of image authentication is that in many cases images need to be subjected to non malicious operations like compression, so the authentication techniques need to be compression tolerant. In this paper we propose an image authentication system that is tolerant to JPEG lossy compression operations. A scheme for JPEG grey scale images is proposed based on a data embedding method that is based on a secret key and a secret mapping vector in the frequency domain. An encrypted feature vector extracted from the image DCT coefficients, is embedded redundantly, and invisibly in the marked image. On the receiver side, the feature vector from the received image is derived again and compared against the extracted watermark to verify the image authenticity. The proposed scheme is robust against JPEG compression up to a maximum compression of approximately 80%, but sensitive to malicious attacks such as cutting and pasting.

**Keywords**—Authentication, DCT, JPEG, Watermarking.

## I. INTRODUCTION

IMAGE authentication techniques have recently gained great attention due to their importance for a large number of multimedia applications. Digital images are increasingly transmitted over non-secure channels such as the Internet. Image authentication techniques have recently gained great attention due to their importance for a large number of multimedia applications. Digital images are increasingly transmitted over non-secure channels such as the Internet. Therefore, military, medical and quality control images must be protected against attempts to manipulate them; such manipulations could tamper the decisions based on these images.

Two approaches have been suggested for achieving the authenticity of digital images: 1) the digital signature-based method [1-4], and 2) the digital watermark-based method [5-10]. The first method uses an encrypted image hash (digital signature), which is generated in the capturing device.

Mona F.M. Mursi is with the Center of Excellence in Information Assurance (CoEIA) and the Department of Information Technology, King Saud University, Riyadh 11543, Saudi Arabia (corresponding author phone: +966-503288366; fax: +966-1-4781479; e-mail: e-mail: monmursi@coeia.edu.sa).

Ghazy M.R. Assassa is with the Center of Excellence in Information Assurance (CoEIA) and the Department of Computer Science, King Saud University, Riyadh 11543, Saudi Arabia, , on leave from the Faculty of Engineering at Shoubra, Benha University, Egypt (e-mail: ghazyassassa@gmail.com).

Hatim A. Aboalsamh is with the Center of Excellence in Information Assurance (CoEIA) and the Department of Computer Science, King Saud University, Riyadh 11543, Saudi Arabia (e-mail: hatim@ksu.edu.sa).

Khaled Alghathbar is with the Center of Excellence in Information Assurance (CoEIA) and the Department of *Information Systems*, King Saud University, Riyadh 11543, Saudi Arabia (e-mail: ghatbar@coeia.edu.sa).

A digital signature is based on the method of Public Key Encryption [11]. A private key is used to encrypt a hashed version (digest) of the image. This encrypted digest of the image is called the signature of the image; it provides a way to ensure that it cannot be forged. This signature then accompanies the image. The authentication process of the image needs an associated public key to decrypt the signature. The image received for authentication is hashed and the resulting digest is compared to the decrypted signature. If they match then the received image is authenticated. Authentication techniques based on digital signature are termed *fragile* since the least change to the image, even a single bit inversion, would deem the image tampered, despite the fact that such alterations to the image should be tolerated and the image should be authenticated.

The second method, the digital watermarking-based method, is a media authentication /protection technique that embeds invisible information into an image. For content authentication, the embedded watermark can be extracted and used for image verification purposes.

The underlying techniques used to implement either of these two approaches can be roughly classified into three main categories: robust, fragile, and semi-fragile. Robust techniques [12, 13] are primarily used in applications such as copyright protection and ownership verification of digital multimedia, because they can withstand nearly all attacks (such as lossy compression, spatial filtering, and geometric distortions). Fragile [14, 15] are mainly applied to content authentication and integrity attestation, because they are sensitive to almost all modifications. On the other hand, semi-fragile methods [1-9, 16, 17] are robust to incidental modification such as JPEG lossy compression, but sensitive to other modifications.

Lin and Chang [3] have shown a mathematical invariant relationship between two discrete cosine transform (DCT) coefficients in a block pair before and after JPEG compression and selected it as the image feature. Zhao et al. [10] proposed an approach for the combined image authentication and compression of images using a digital watermarking and data hiding technique. Lu et al. [7] presented a semi-fragile digital image watermarking method, based on index constrained vector quantization, which embeds the watermark information in the compressed bit streams. Zhou et al. [17] presented a semi-fragile method in which a watermark is extracted from the original image and then reinserted. The error correction coding ECC is adopted to encode the watermarks extracted from the image. To increase the security of this method, the user's secret key is used to encrypt and decrypt the watermark during the watermark extraction and insertion procedures. Most of these proposed authentication methods have the same merits and suffer from the same disadvantages. The common

advantage is that all methods can distinguish the JPEG lossy baseline compression from malicious manipulations, because the JPEG compression technique is typically used on the Internet. The common disadvantage is that there are no effective arguments and analysis to prove the validity of the methods [7,8,17, 18-20].

An effective authentication scheme should have the following desirable features: 1) to be able to detect malicious image tampering; 2) to be able to integrate authentication data with host image rather than as a separate data file; 3) to be oblivious, i.e. the authentication of the watermark should not need the original image or watermark; 4) to be invisible, i.e. the embedded authentication data be invisible under normal viewing conditions; 5) to allow the watermarked image be stored in lossy compression format; 6) to be tolerant to incidental image alterations due to noise, etc. An additional desirable feature is the ability to locate the tampering. Previously published methods for image authentication do not satisfy all the requirements.

The digital signature proposed in [21], as well as the content based signature reported in [22] and [23] do not satisfy requirement 3. The pixel-domain scheme [24] cannot be stored in lossy compression format. In addition, two frequency domain data hiding schemes [25, 26] may be used for authentication, but they cannot always locate the alteration and the distortion introduced by embedding introduces image artifacts.

In [27] the authors propose a novel approach to content-based watermarking for image authentication that is based on Independent Component Analysis (ICA). In the proposed scheme, 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. That authentication technique is claimed to be robust against incidental image processing operations, but detects malicious tampering and correctly locates the tampered regions.

In this paper, an image authentication scheme for verifying the authenticity of JPEG images is presented; it is based on a secret key and a secret mapping vector that is used in embedding the digital signature of a feature vector derived from the frequency domain of the image. The invariance of the relationship between the differences of the DCT coefficients is used to generate the image feature vector. The scheme presented is simple and yet solid, and can differentiate the practical JPEG lossy baseline compression from malicious image tampering. The image feature vector generation and embedding algorithms are simple and do not need extra memory space to store authentication signatures. Additionally the proposed method is secure and tolerant to noise. This scheme can be applied to compressed images using JPEG compression, and the marked image can be kept in the compressed format. This scheme is computationally efficient, and can be applied to video authentication.

The rest of the paper is organized as follows. In section II the proposed authentication scheme is described. In section III the feature vector generation is discussed. In section IV the data embedding scheme is presented. In section V the validation process is outlined. In section VI the experimental

results are presented. In section 7 the handling of special image cases is discussed, and finally section 8 is the conclusion.

## II. THE PROPOSED APPROACH

In this paper, we propose a watermarking technique for digital image authentication that does not require the receiver to maintain a database of watermarks. It is based on the observation of Ref. [3] who discovered a mathematical invariant relationship between two discrete cosine transform (DCT) coefficients in a block pair before and after JPEG compression. That invariant property for JPEG images is utilized in selecting the image's feature vector. The proposed technique authenticates the image by embedding an encrypted feature vector derived from the frequency domain of the image in the DCT coefficients of the image using a proprietary mapping vector.

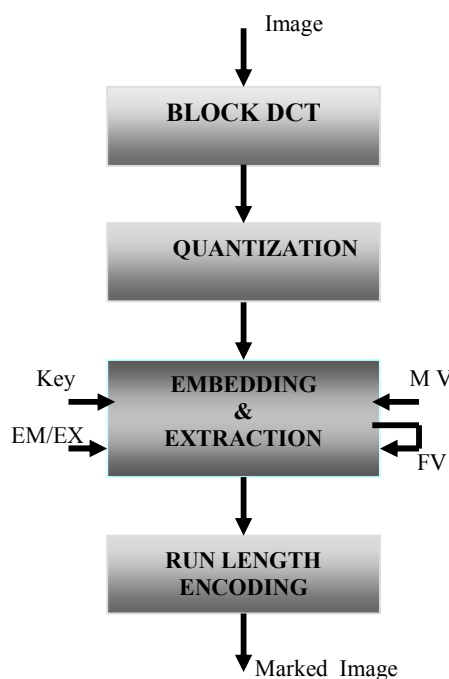


Fig. 1 Block Diagram of JPEG Embedding and Extraction Process

The mapping vector is generated at the source and destination via a secret *seed*. For added security, the derived feature vector is XOR-ed with a secret key known to sender and receiver. It could be possibly incorporated in the camera and downloaded in the authentication system upon system initialization.

In the proposed scheme we will consider only *grey scale* JPEG compressed images. The extension to color images is reasonably straightforward.

A block diagram for the embedding of data is given in Fig. 1. EM/EX is the embed and extract control. MV is the proprietary mapping vector, and the feature vector is derived from the DC components of the quantized image blocks. The process of image authentication consists of two steps: 1) the image watermarking step, and 2) the watermark extraction and

verification that it is identical to the one that was originally in the image.

### III. FEATURE VECTOR DERIVATION

The basic concept in Lin's and Chang's proposal in their paper [3] is based on the following theorem:

Assume  $F_p^{\wedge} = \text{IntegerRound}(F_p(v)/Q(v))$ , that means a DCT coefficient in the  $v^{\text{th}}$  position of a  $p^{\text{th}} 8 * 8$  non-overlapping block of the image quantized by quantization table  $Q(v)$ , and define :

$$\Delta F_{pq} = F_p - F_q^{\wedge}$$

$$\Delta F_{pq}^{\wedge} = F_p^{\wedge} - F_q^{\wedge}$$

Hence the following properties must be true:

1. If  $\Delta F_{pq} > 0$ , then  $\Delta F_{pq}^{\wedge} > 0$
2. Else if  $\Delta F_{pq} < 0$ , then  $\Delta F_{pq}^{\wedge} < 0$
3. Else  $\Delta F_{pq} = 0$ , then  $\Delta F_{pq}^{\wedge} = 0$

The only exception is that "greater than" and "less than" may become "equal" due to the rounding effect of quantization. The theorem only preserves the sign of coefficient differences.

Based on that theorem, a feature vector representing the relative magnitude of the consecutive image blocks, is derived, such that:

$$\text{If } DC_i > DC_j$$

$$\text{Then } V(i) = 1$$

$$\text{Else } V(i) = 0$$

Hence, as the relative values of DC coefficients before and after JPEG compression remain the same, we use this property to extract a feature vector from the image. The image's feature vector is used twice. First, as an invisible watermark embedded in the image; the second time, on the verification side, to compare it with the extracted watermark. The generation process is done by calculating the image's feature vector from the difference between two adjacent DCs of a  $16*16$  of the quantized blocks of the image as shown above. Hence for an image that contains  $M$  blocks, the feature vector will be of length  $M/4$ .

The sequence of bits thus derived, will form the image's feature vector. For added security, the obtained feature vector is XOR-ed with a secret key.

### IV. DATA EMBEDDING SCHEME

The next step is to embed each of encrypted feature vector ( $FV_e$ ) bits in consecutive quantized image blocks. This is done using the mapping vector.

Mapping vector (MV) is a proprietary mapping vector that contains a pseudo random string of ones and zeroes such that there are no more than two consecutive ones or zeroes. The MV bits are made to correspond to the range of the AC component values of the image via a look-up table:

TABLE I  
 MAPPING TABLE

AC	..	9	8	7	6	5	4	3	2	1	.
MV	.	1	0	0	1	0	1	1	0	1	.

The image's feature vector is embedded in such a way so as to avoid compromising it when the image is compressed. After the quantized DCT coefficients of the image are obtained, one bit of  $FV_e$  is inserted by modifying the low frequency AC values of a block using the mapping vector. After experimentation with a variety of images, better results were obtained when the change was performed only on the first five AC components. This choice was made so as not to degrade the watermarked image quality.

The embedding process is as follows: Assuming we need to embed a "1" into a block, then we examine the first AC component and its corresponding MV bit, from a table such as in table 1, if it is a "1" then we leave the AC unchanged, if it is a "0", then we find the nearest value in the table that has its MV bit "1", either to the left or right. In this case the AC value is changed to the value corresponding to the "1" in the table. The same process is carried out for all the first five AC components.

For example if the first five AC components were : 9, 4, 1, 5, 3 then after embedding they become : 9, 3, 1, 6, 3 (according to Table II).

The embedding is done by scanning through each bit of the watermark FV, and then altering a corresponding block in the original image using a mapping procedure as outlined above.

In this way, the host image is watermarked by the FV in the frequency-domain as it makes the watermark robust against attacks such as lossy compression. Fig. 2 shows the complete watermarking process.

TABLE II  
 QUANTIZED 8X8 BLOCK

15	9	3	5	3	-1	0	1
4	6	-5	-1	0	1	0	0
1	2	-2	3	-1	1	0	0
-3	-1	3	-2	0	1	0	0
-3	1	1	0	0	0	0	0
0	0	1	0	0	0	0	0
1	-1	1	0	0	0	0	0
1	0	0	0	0	0	0	0

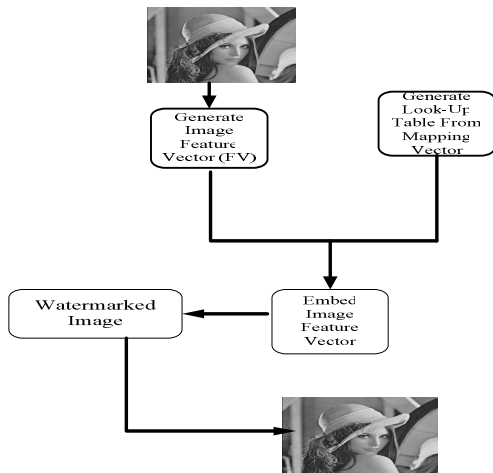


Fig. 2 The Image Watermarking Process

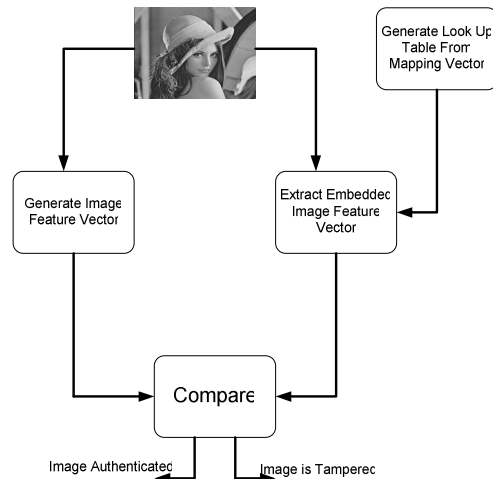


Fig. 3 The image authentication process

### V. VERIFICATION PROCESS

The main steps for image watermark verification i.e image authentication, are:

- Create a look-up table using the secret seed.
- Calculate the image's feature vector  $FV_e$  as was done in the watermarking side, and XOR it with the secret key to obtain  $FV_c$ .
- Extract the embedded authentication data (i.e.  $FV_e$ ) from the image by mapping the first five AC that hold a bit of the image's feature using the look up table.
- Check for a match with  $FV_c$ . If so, the image is authentic and there is no tampering, otherwise, if the extracted watermark is not the same as the calculated one, the image may be tampered or modified.

To extract the watermark  $FV_e$ , we extract the embedded bit sequence from the blocks of quantized DCT coefficients by mapping the first five AC's of a block that hold the embedded bit to its value 0 or 1 using the derived look up table. Due to noise, the value of some ACs may change. To minimize that effect, after extracting the values embedded in the first five AC components, the predominant value is selected as the value of the embedded bit in this block, i.e. voting between AC's values. The sequence of block derived bits forms the embedded image's feature vector. As an example, consider the block in figure 4. To derive the embedded bit using the look up table (LUT) as shown in table 1, we find that the values corresponding to the ACs : 9,4,1,6,3 from the LUT are 1,0,1,1,1; hence the embedded bit is the majority, i.e. 1. Final verification is done when  $FV_c = FV_e$ .

Fig. 3 shows the verification process.

### VI. EXPERIMENTAL RESULTS

Detailed experiments were conducted to validate the proposed scheme and it was found that the proposed scheme is able to distinguish the malicious and incidental attacks. An example of one of the images used is presented here.

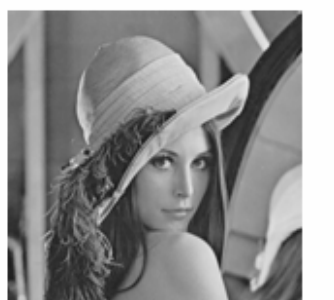
The Lena image of Fig. 4 is a 512x512 image (257 kb). In Fig. 4(a) the image is the original BMP image with no watermark. In Fig. 4(b) the image is the same image but watermarked, and is in the JPEG format. In figure 4(c) the image is the watermarked and tampered image. As can be seen there is hardly any perceptible difference in the first two images.

In Fig. 4 (b) the length of the feature vector is 1024, and the number of corrupted blocks is 23, i.e. an error of 2% which is within the allowable range. For the tampered image of Fig. 4 (c) the number of corrupted blocks is 184 which represents an error of 18% and hence is declared as tampered.

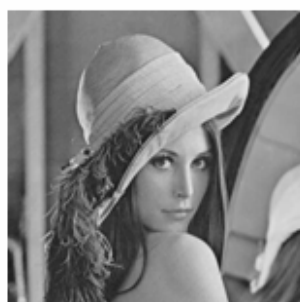
The Lena image was also subjected to a range of compressions, from 10% to 75%, and in all cases the image was deemed un-tampered. For further compression the image was declared as tampered. Thus the results obtained are within the acceptable range of performance.

### VII. CONCLUSION

Image authentication, requires that the verification method be able to resist incidental distortions while being sensitive to malicious manipulations. In this paper, an image authentication scheme for verifying the authenticity of JPEG images is presented; it is based on a secret key and a secret mapping vector that is used in embedding the digital signature of a feature vector derived from the frequency domain of the image.



(a)



(b)



(c)

Fig. 4 Lena Image (a) Before watermarking, (b) After watermarking, (c) After tampering the watermarked image

The invariance of the relationship between the differences of the DCT coefficients was used to generate the image feature vector. The encrypted image feature vector was then embedded into the original image by modifying the DCT coefficients where necessary according to the mapping vector.

A single bit is embedded in the low frequency AC components (specifically the first five) of the image DCT of an 8x8 block, and a voting technique is used in extracting the embedded bit for authentication purposes. The main contributions of our proposed solution are:

(1) This authentication system is watermark-based, not signature-based. Therefore, additional authentication signature is not required when verifying the authentication.

(2) The security goal of this system is implemented by a secret key and a pseudo randomly generated mapping vector during the procedures of image feature generation, embedding and authentication. If attackers want to counterfeit the image, they must know the secret key and the seed. Therefore, the security

of this system is protected based on the secret key and the mapping vector.

(3) The proposed scheme has redundancy in the embedding process, which renders the scheme more tolerant to noise than watermarking schemes which do not have this property.

(4) The results demonstrated that our proposed scheme has good visual quality of the watermarked image.

The detailed experiments were conducted and it was found that the proposed scheme is able to distinguish the malicious and incidental attacks and is also highly secure because the embedded feature vector is XOR-ed with a secret key. The experimental results also indicate that the proposed method is both feasible and effective.

#### REFERENCES

- [1] El-Din, S. N. and Moniri, M., "Fragile and Semi-Fragile Image Authentication Based on Image Self-Similarity", IEEE International Conference on Image Processing, Vol. 2, pp. 22-25, 2002.
- [2] Friedman, G. L., "The Trustworthy Digital Camera: Restoring Credibility to the Photographic Image", IEEE Trans. Consumer Electron., Vol. 39, pp. 905-910, 1993.
- [3] Lin, C. Y. and Chang, S. F., "A Robust Image Authentication Method Distinguishing JPEG Compression from Malicious Manipulation", IEEE Trans. on Circuits and Systems of Video Technology, Vol. 11, pp.153-168, 2001.
- [4] Sun, Q., Chang, S. F., Kurato, M. and Suto, M., "A New Semi-Fragile Image Authentication Framework Combining ECC and PKI Infrastructures", IEEE International Symposium on Circuits and Systems (ISCAS 2002), Vol. 2, pp. 440-443, 2002.
- [5] Li, C. T., "Digital Fragile Watermarking Scheme for Authentication of JPEG Images", Proc. of IEE on Vision, Image and Signal Processing, Vol. 151, pp. 460-466, 2004.
- [6] Lu, C. S. and Liao, H. Y. M. H. Y. Mark, "Multipurpose Watermarking for Image Authentication and Protection", IEEE Trans. Image Processing, Vol. 10, pp.1579 - 1592, 2001.
- [7] Lu, Z. M., Liu, C. H., Xu, D. G. and Sun, S. H., "Semi-Fragile Image Watermarking Method Based on Index Constrained Vector Quantization", IEE Electronics Letters, Vol. 39, pp. 36-37, 2003.
- [8] Lu, Z. M., Xu, D. G. and Sun, S. H., "Multipurpose Image Watermarking Algorithm Based on Multistage Vector Quantization", IEEE Trans. on Image Processing, Vol. 14, pp. 822 - 831, 2005.
- [9] Sun, R., Sun, H. and Yao, T., "A SVD- and Quantization Based Semi-Fragile Watermarking Technique for Image Authentication", IEEE International Conference on Signal Processing (ICSP 2002), Vol. 2, pp. 1592 - 1595, 2002.
- [10] Zhao, Y., Campisi, P. and Kundur, D., "Dual Domain Watermarking for Authentication and Compression of Cultural Heritage Images", IEEE Trans. on Image Processing, Vol. 13, pp.430-448, 2004.
- [11] Schneider, M. and Chang, S. F., "A Robust Content Based Digital Signature for Image Authentication", Proc. of IEEE Int. Conf. on Image Processing, Vol. 3, pp. 227-230, 1996.
- [12] Cox, I. J., Kilian, J., Leighton, F. T. and Shamoon, T., "Secure Spread Spectrum Watermarking for Multimedia", IEEE Trans. Image Process., Vol. 6, pp. 1673-1687, 1997.
- [13] Ramkumar, M. and Akansu, Ali N., "A Robust Protocol for Proving Ownership of Multimedia Content", IEEE Trans. on Multimedia, Vol. 6, pp. 469-478, 2004.
- [14] Celik, M. U., Sharma, G., Saber, E. and Tekalp, A. M., "Hierarchical Watermarking for Secure Image Authentication with Localization", IEEE Trans. on Image Processing, Vol. 11, pp. 585 - 595, 2002.
- [15] Lin, C. H. and Hsieh, W. S., "Applying Projection and B-spline to Image Authentication and Remedy", IEEE Trans. on Consumer Electronics, Vol. 49, pp. 1234-1239, 2003.
- [16] Bao, P. and Ma, X., "Image Adaptive Watermarking Using Wavelet Domain Singular Value Decomposition", IEEE Trans. on Circuits and Systems for Video Technology, Vol. 15, pp. 96-102, 2005.
- [17] Zhou, X., Duan, X. and Wang, D., "A Semi-Fragile Watermark Scheme for Image Authentication", IEEE International Conference on Multimedia Modeling, pp. 374-377, 2004.

- [18] Ho, C. K. and Li, C. T., "Semi-Fragile Watermarking Scheme for Authentication of JPEG Images", Proc. Of IEEE Int. Conf. on Information Technology: Coding and Computing, Vol. 1, pp. 7-11, 2004.
- [19] Wu, Y., "Detecting Tampered Image Blocks Using Error Correcting Code", Proc. of IEEE Int. Conf. on Multimedia and Expo, Vol. 3, pp. 2047-2050, 2004.
- [20] Li, C. T., "Digital Fragile Watermarking Scheme for Authentication of JPEG Images", Proc. of IEE on Vision, Image and Signal Processing, Vol. 151, pp. 460-466, 2004.
- [21] G. L. Friedman, "The Trustworthy Digital Camera: Restoring Credibility to the Photographic Image", IEEE Trans. on Consumer Electronics, Nov.1993.
- [22] M. Schneider, S-F. Chang, "A Robust Content Based Digital Signature for Image Authentication", ICIP, 1996.
- [23] D. Storck, "A New Approach to Integrity of Digital Images", IFIP Conf. on Mobile Communication, 1996.
- [24] M. M. Yeung, F. Mintzer, "An Invisible Watermarking Technique for Image Verification", ICIP, 1997
- [25] M. D. Swanson, B. Zhu, A. H. Tewfik, "Robust Data Hiding for Images", IEEE DSP Workshop, 1996.
- [26] E. Koch, J. Zhao, "Towards Robust and Hidden Image Copyright Labeling", IEEE Workshop on Nonlinear Signal and Image Processing, 1995.
- [27] L. Parameswaran, K. Anbumani, "Content-Based Watermarking for Image Authentication Using Independent Component Analysis", Informatica 32, 299-306, 2008.