http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 10 (1): 120-126, 2011
ISSN 1812-5638

An Image Hashing Scheme based on Mean-removed Vector Quantization for Multiple Purposes

[1]Mei-Lei Lv and [2]Zhe-Ming Lu
[1]Department of Information and Electrical Engineering, Quzhou College, Quzhou 324000, China
[2]School of Aeronautics and Astronautics, Zhejiang University, Hangzhou, China

Abstract: Perceptual hashing has been proved to be an effective solution for multimedia indexing, authentication or watermarking. Traditional perceptual hashing schemes are typical designed only for one purpose. This study presents a multipurpose image-hashing scheme based on Mean-Removed Vector Quantization (MRVQ) for both copyright protection and content authentication. The main idea is to perform MRVQ on the original image to yield two index tables, one for copyright protection and the other for content authentication. The original gray-level image is first divided into non-overlapping small blocks. The mean value for each block is calculated and quantized by the scalar quantizer to get a mean index and the quantized mean is removed from the image block to obtain the residual vector that is further quantized by the vector quantizer to obtain the residual index. All obtained mean indices constructed the mean index table and all obtained residual indices construct the residual index table. The obtained two index tables are then transformed into two intermediate binary images based on two different mapping functions, respectively. One mapping function is based on the variance of indices in a 3×3 neighborhood and the other mapping function is based on the number of indices larger than the mean of indices in a 3×3 neighborhood. Finally, the authentication mark and permuted copyright logo are respectively XOR-ed with the two intermediate binary images to obtain final authentication and protection fingerprints. Experimental results demonstrate the effectiveness of the proposed scheme.

Key words: Perceptual hashing, perceptual image hash function, copyright protection, content authentication, multistage vector quantization

INTRODUCTION

With the rapid development of computer, multimedia and network technologies, the amount of audiovisual information available in the digital format has grown exponentially, resulting in information explosion and exceeding the limit of human's acceptability and thus several serious issues have emerged. On the one hand, as multimedia data are stored in digital formats, it is easy to modify and forge their content by widely available editing tools. The ability to detect changes in digital multimedia has been very important for many applications, especially for journalistic photography, medical or artwork image databases. Content authentication has therefore been one of the most important issues in the digital world (Lu *et al*., 2005). On the other hand, as businesses online have become ubiquitous, valuable digital artworks may be losslessly reproduced and arbitrarily distributed and thus copyright protection techniques are urgently required to protect the intellectual property rights. In addition, efficient search of desired multimedia content from the huge multimedia database is also a great challenge and therefore content-based retrieval has been an interesting and rapid developing research area since the 1990's. In general, content authentication and copyright protection techniques can be classified into three classes: digital signature based, watermark based (Fiaidhi and Mohammed, 2003; Khan *et al*., 2008; Lu and Li, 2006; Lu *et al*., 2000, 2003, 2005; Lu and Sun, 2000; Qureshi and Tao, 2006) and perceptual hash based (Dittmann *et al*., 1999; Lei *et al*., 2010; Lu and Liao, 2003; Monga *et al*., 2006; Monga and Evans, 2006; Monga and Mhcak, 2007; Venkatesan *et al*., 2000; Yu *et al*., 2010). A digital signature scheme is typically composed of three algorithms: (1) a key generation algorithm that selects a private key uniformly at random from a set of possible private keys. (2) a signing algorithm that produces a signature based on the given message and private key. (3) a signature verifying algorithm that either accepts or rejects the message's claim to authenticity. Digital watermarking is the technique to add some digital information to the multimedia data, such as images, voice

Corresponding Author: Zhe-Ming Lu, Teaching Building 5, Yuquan Campus, Zhejiang University, 38 ZheDa Road, Hangzhou 310027, People's Republic of China Tel: +86-571-87951589

and video signals and so on. It usually embeds visible or invisible watermarks in multimedia, without affecting the appearance and integrity of original document. Perceptual hash functions are designed for multimedia. Cryptographic hash functions generate a totally different hash value even if the input is changed by a single bit, while perceptual hash functions are expected to change the hash value only if the input is perceptually changed. Besides content authentication and copyright protection, perceptual hashing has also been a useful technique for content based retrieval.

The problem of mapping an image to a short binary string is known as image hashing. The image hash function should map perceptually identical images into the same hash value with high probability, while mapping perceptually different images into independent hash values. In addition, the hash function should be secure enough such that an attacker cannot predict the hash value from the image. An image hash function can be used as the robust feature of an image for image retrieval or copyright protection. An image hash function can be split into two stages. In the first step, a feature vector (or intermediate binary string) is extracted from the image to capture the important perceptual aspects of the image. In the second step, the feature vector (or the intermediate string) is securely transformed, compressed or quantized to obtain the final hash.

Existing image hashing schemes are typically designed only for one purpose, e.g., copyright protection or retrieval. They can be roughly classified into several categories such as statistics based (Venkatesan *et al*., 2000), relations based (Lu and Liao, 2003), low-level features based (Dittmann *et al*., 1999), feature points based (Monga and Evans, 2006), clustering based (Monga *et al*., 2006), non-negative matrix factorizations based (Monga and Mhcak, 2007) and DCT-domain statistics based (Lei *et al*., 2010; Yu *et al*., 2010). Venkatesan *et al*. (2000) adopted randomized signal processing strategies to irreversibly compress an image into random binary strings that are robust against image changes due to compression or geometric distortions. Lu and Liao presented a so-called structural digital signature for image authentication based on the phenomena that, in a subband wavelet decomposition, a parent node and its child nodes are uncorrelated, but they are statistically dependent. Dittmann *et al*. (1999) proposed the utility of feature points in perceptual hashing applications because they are sensitive to several perceptually insignificant modifications as well as content changing operations. Monga *et al*. adopted a wavelet based feature detector to extract significant image features based on the characteristics of the visual system. In Monga *et al*. (2006), they divided image hashing into two steps, i.e., feature extraction (intermediate hash) followed by data clustering (final hash). For any perceptually significant feature extractor, they proposed a polynomial-time heuristic clustering algorithm that automatically determines the final hash length needed to satisfy a specified distortion. In Monga and Mhcak (2007) utilized the non-negative matrix factorization (NMF) for image hashing, where they viewed images as matrices and the goal of hashing as a randomized dimensionality reduction that retains the essence of the original image matrix while preventing intentional attacks of guessing and forgery. Lei *et al*. (2010) presented a novel robust image hashing scheme for image authentication based on the Discrete Cosine Transform (DCT) and Least-Squares Line (LSL) fitting of Discrete Wavelet Transform (DWT) coefficients. Recently, Yu *et al*. (2010) proposed a novel image hashing scheme based on the statistical invariance of DCT coefficients, where they extracted the invariant parameters with the modified ML principle.

To achieve multiple purposes of copyright protection and content authentication simultaneously, in this study, we present a novel multipurpose image hashing scheme based on mean-removed vector quantization. The advantages lie in two aspects. One is that it is multipurpose. The other is that the copyright protection and authentication can be performed visually other than conventional image hashing schemes. The experimental results demonstrate the effectiveness of the proposed scheme.

MRVQ-BASED MULTIPURPOSE IMAGE HASHING SCHEME

Mean-removed vector quantization: Vector Quantization (VQ) is an attractive block-based image compression scheme. VQ can be defined as a mapping from k-dimensional Euclidean space R^k into a N-sized finite codebook $C = \{c_i | i = 0, 1, \ldots, N-1\}$, where c_i is called a codeword. The codebook C is generally generated from a training set using the well-known LBG algorithm (Linde *et al*., 1980). Before encoding, the image is first divided into non-overlapping blocks and then sequentially encoded block by block. In the encoding stage, for each input vector $x = (x_1, x_2, \ldots, x_k)^t$, we find the best matching codeword $c_i = (c_{i1}, c_{i2}, \ldots, c_{ik})^t$ in the codebook C, satisfying:

$$d(x,c_i) = \min_{0\le j\le N-1} d(x,c_j) \quad (1)$$

where, $d(x, c_j)$ is the error between x and c_j defined as follows:

$$d(x,c_j) = \sum_{l=1}^{k} (x_l - c_{jl})^2 \quad (2)$$

And then the codeword index i is transmitted over the channel to the decoder. The decoder possesses the same codebook as the encoder. In the decoding stage, for each index i, the decoder looks up the codeword c_i in Codebook C to reconstruct the input vector x.

In general, we are willing to deal with vectors that have zero statistical mean in the sense that the expected value of each component is zero. However, many vectors such as sampled image blocks have only nonnegative components and thus have nonzero means. The local means of image blocks can vary quite widely over an image. In fact, the mean of an image vector can be regarded as statistically independent of the variation of the vector. The mean of a vector refers to the sample mean, i.e., the average of all components in the vector and it is a scalar random variable given by:

$$m = \frac{1}{k}\sum_{i=1}^{k} x_i = \frac{1}{k} 1^t x \tag{3}$$

where, $1 = (1,1,\ldots,1)^t$ denotes the k-dimensional vector with all components equal to Eq. 1. The residual r of the random variable x is defined as:

$$r = x - \frac{1}{k}(1^t x)1 = x - m1 \tag{4}$$

Thus, x can be described as the sum of a mean vector m1 and the residual vector r as follows:

$$x = r + m1 \tag{5}$$

Therefore, we can naturally decompose the original image block into two separate features, a mean (representing the background gray-level) and a residual vector (representing the shape). Quantizing these two features based on separate VQ codebooks is called mean-removed VQ or mean-residual VQ (MRVQ).

In this study, we adopt the MRVQ structure as illustrated in Fig. 1. The mean of x is first computed and quantized by a mean codebook $C_m = \{c_{m0}, c_{m1}, \ldots, c_{m(Nm-1)}\}$ and the quantized mean $\hat{m}$ is then subtracted from each component of x to obtain the residual vector r. The residual vector r is then quantized with a residual codebook $C_r = \{c_{r0}, c_{r1}, \ldots, c_{r(Nr-1)}\}$. The MRVQ output consists of two indices for the mean and residual vector, respectively. The reconstructed vector after quantization of the mean and the residual vector is given by:

$$\hat{x} = \hat{r} + \hat{m}1 \tag{6}$$

Here, $\hat{m}$ is a codeword from the scalar codebook C_m of size N_m and $\hat{r}$ is a codeword chosen from the residual codebook C_r of size N_r. In fact, the equivalent codebook is the product codebook C that can be generated from the Cartesian product $C_m \times C_r$. Compared to the full search VQ with the product codebook C, the mean-removed VQ can reduce the complexity from:

$$N = N_m \times N_r \text{ to } N_m/k + N_r$$

Proposed image hashing scheme: Our MRVQ-based multipurpose image hashing scheme is depicted in Fig. 2.

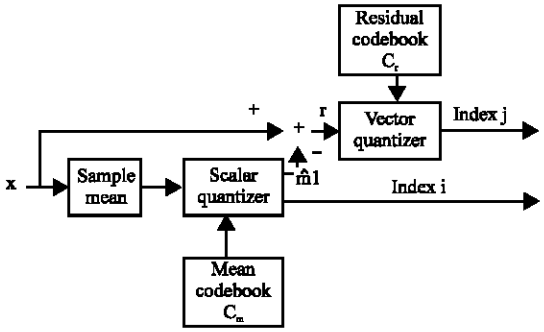

Fig. 1: Mean-removed vector quantization

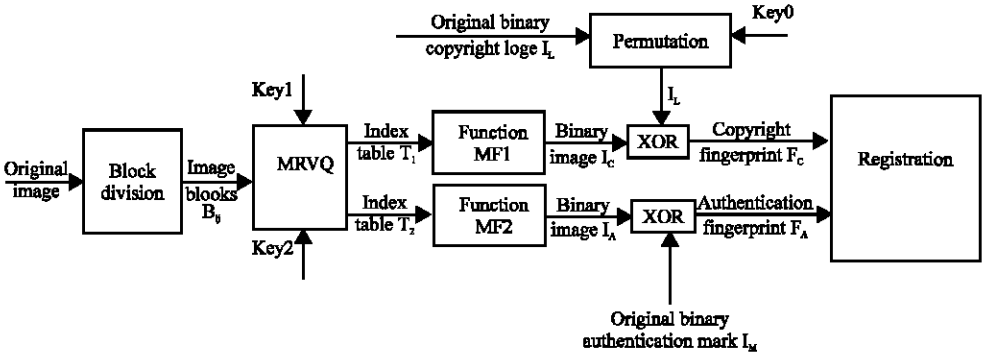

Fig. 2: MRVQ-based image hashing process

The main idea is to obtain two intermediate binary images from mean and residual index tables based on different specific mapping functions and then perform the first XOR operation between one intermediate binary image and the permuted copyright logo and the second XOR operation between the other intermediate binary image and the authentication mark, respectively. The resulting two binary images can be finally served as the authentication and copyright fingerprints of the original gray-level image. Assume that the input original image I_O is of size N×N, the input binary copyright logo I_L is of size N/4×N/4, the input binary authentication mark I_M is of size N/4×N/4, the output binary copyright fingerprint F_C is of size N/4×N/4 and the output binary authentication fingerprint F_A is of size N/4×N/4, then the whole multipurpose image hashing process can be illustrated as follows:

- **Step 1:** Segment I_O into non-overlapping blocks B_{ij} of size 4×4, where i, j = 1, 2,…, N/4 and permute I_L with the key, Key0, to obtain the permuted logo I_L
- **Step 2:** Calculate the mean value m_{ij} for each block B_{ij} and quantize m_{ij} by the scalar quantizer with Codebook $C_m = \{c_{m0}, c_{m1}, \ldots, c_{m(Nm-1)}\}$. All obtained indices construct the mean index table $T_1 = \{s_{ij}\}$, where s_{ij} is the mean codeword's index for Block B_{ij}, i, j=1, 2,…, N/4. Here we use the key, Key1, to permute the codewords in Codebook C_m before the scalar quantization for security
- **Step 3:** Calculate the residual vector $r_{ij} = B_{ij}$-$c_{ms_{ij}}1$ for block B_{ij} and quantize r_{ij} by the vector quantizer with Codebook $C_r = \{c_{r0}, c_{r1}, \ldots, c_{r(Nr-1)}\}$. All obtained indices construct the residual index table $T_2 = \{v_{ij}\}$, where v_{ij} is the residual codeword's index for Block B_{ij}, i, j = 1, 2,…, N/4. Here we use the key, Key2, to permute the codewords in Codebook C_r before the residual vector quantization for security
- **Step 4:** Map index tables T_1 and T_2 to binary images I_C and I_A based on mapping functions MF1 and MF2, respectively
- **Step 5:** Perform the XOR operation between I_C and I_L to obtain the final copyright fingerprint F_C. Similarly, perform the XOR operation between I_A and I_M to obtain the final authentication fingerprint F_A

Now we turn to introduce the two mapping functions used in step 4. The mapping function MF1 ($T_1 \rightarrow I_C$) can be described as follows: For each index s_{ij}, we get its 3×3 neighborhood centered by s_{ij} and calculate the average of these 9 indices:

$$a_{ij} = \frac{1}{9}\sum_{l=i-1}^{i+1}\sum_{k=j-1}^{j+1} s_{lk} \quad (7)$$

Then calculate the mean absolute error between s_{lk} (l = i-1,i,i+1; k = j-1,j,j+1) and a_{ij} as follows:

$$e_{ij} = \frac{1}{9}\sum_{l=i-1}^{i+1}\sum_{k=j-1}^{j+1} |s_{lk} - a_{ij}| \quad (8)$$

Based on e_{ij}, we finally binarize each block B_{ij} to obtain the binary image $I_C = \{I^C_{ij}\}$ as follows:

$$I^C_{ij} = \begin{cases} 1 & e_{ij} \geq N_m/4 \\ 0 & e_{ij} < N_m/4 \end{cases} \quad (9)$$

Here, N_m denotes the number of codewords in the mean codebook C_m.

Similarly, the mapping function MF2 ($T_2 \rightarrow I_A$) can be described as follows: For each index v_{ij}, we get its 3×3 neighborhood centered by v_{ij} and calculate the mean of these 9 indices:

$$a_{ij} = \frac{1}{9}\sum_{l=i-1}^{i+1}\sum_{k=j-1}^{j+1} v_{lk} \quad (10)$$

Then classify the Eq. 9 indices into two classes as follows:

$$\begin{cases} v_{lk} \in A & \text{if } v_{lk} \geq a_{ij} \\ v_{lk} \in B & \text{if } v_{lk} < a_{ij} \end{cases} \quad (l = i-1,i,i+1; k = j-1,j,j+1) \quad (11)$$

Count the number of indices belonging to class A and denote it as n_{ij}. Based on n_{ij}, we finally binarize each block B_{ij} to obtain the binary image $I_A = \{I^A_{ij}\}$ as follows:

$$I^A_{ij} = \begin{cases} 1 & n_{ij} \geq 5 \\ 0 & n_{ij} < 5 \end{cases} \quad (12)$$

The authentication process: The authentication process is shown in Fig. 3, which can be briefly expressed as follows:

- **Inputs:** The suspect image I_O', the registered copyright fingerprint F_C and the registered authentication fingerprint F_A, the original copyright logo I_L and the original authentication mark I_M
- **Outputs:** The binary decision result of the copyright existence and the binary decision result of the authenticity
- **Step 1:** Using the same steps 1-4 in the hashing process to obtain the binary images I_C and I_A from the suspect image I_O
- **Step 2:** Perform the XOR operation between I_C and F_C to obtain the suspect permuted logo I_{PL} and perform

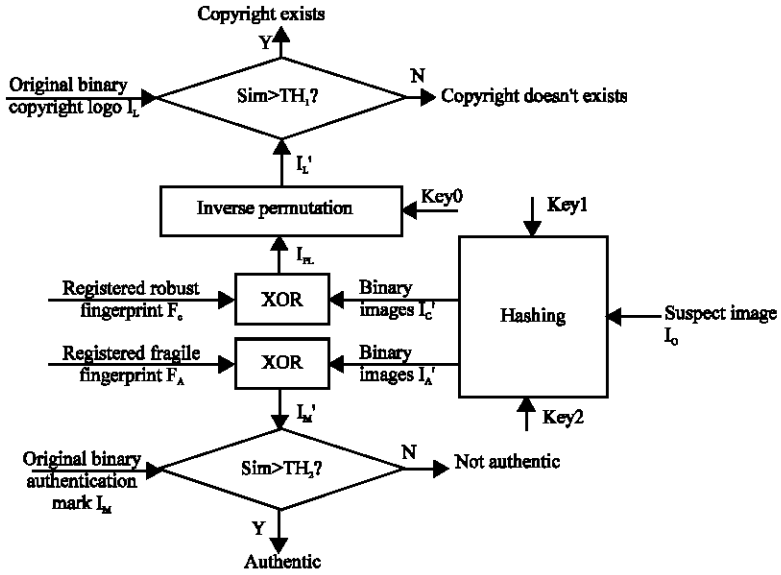

Fig. 3: The authentication process

Fig. 4: Four training images for MRVQ. (a) Lena image, (b) peppers image, (c) baboon image and (d) F16 image

the XOR operation between I_A and F_A to obtain the suspect mark I_M. Then, the suspect permuted logo I_{PL} is further inversely permuted with the key, Key0, to obtain the suspect logo I_L

- **Step 3:** Calculate the hamming similarity (i.e., the percentage of identical bits when comparing two binary strings) between the suspect logo I_L' and the original logo I_L. If the similarity is larger than the threshold TH_1, then the copyright exists; otherwise, the copyright doesn't exist. Similarly, compare the suspect mark I_M with the original mark I_M. If the similarity is larger than the threshold TH_2, then the suspect image I_O is authentic; otherwise, it is not authentic

EXPERIMENTAL RESULTS

To demonstrate the effectiveness of the proposed method, the 256 gray-level 512×512 Lena image is used for multipurpose hashing, as shown in Fig. 4a. The Lena image is divided into 16384 blocks of size 4×4 for MRVQ. The original copyright logo, the permuted copyright

Fig. 5: The attacked Lena images. (a) JPEG compression with QF = 50, (b) cropping in the middle part of the image, (c) rotation by angle 1°, (d)median filtering with the radius of 3 pixels, (e) contrast enhancement by 10% and (f) alteration of eyes

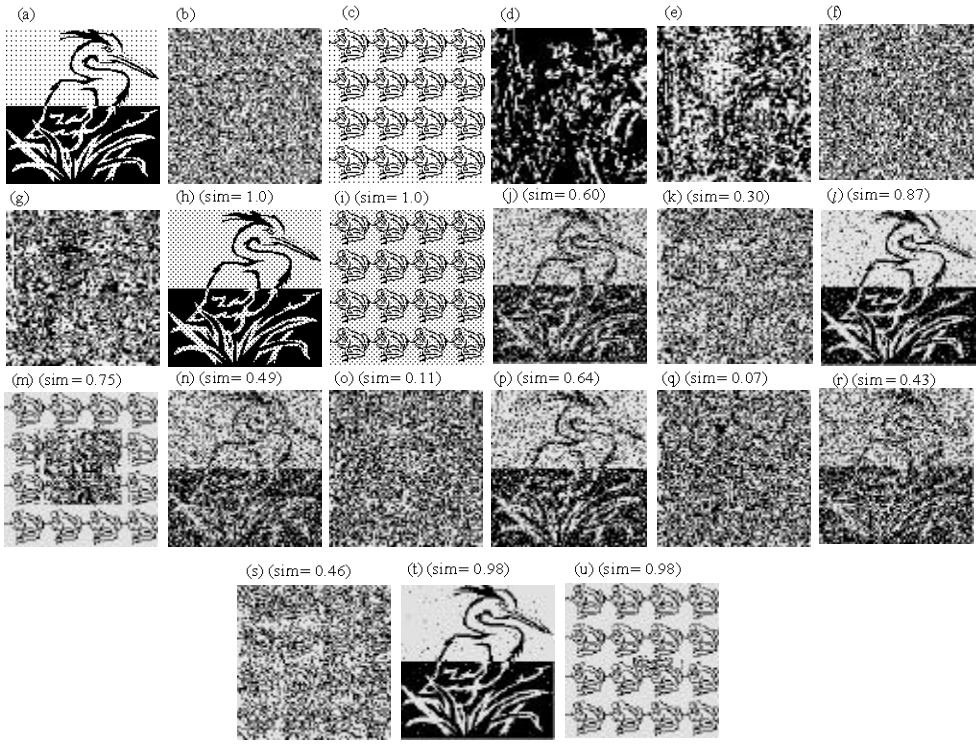

Fig. 6: Performance testing results. (a-c) Original copyright logo, permuted logo and original authentication mark, (d-g) obtained binary image 1, binary image 2, copyright fingerprint and authentication fingerprint, (h-i) obtained copyright logo and authentication mark without any attack, (j-u) obtained 6 copyright logos and 6 authentication marks under 6 attacks, i.e., JPEG compression with QF = 50, cropping in the middle part of the image, rotation by angle 1°, median filtering with the radius of 3 pixels and contrast enhancement by 10% and alteration of eyes, respectively

logo and the original authentication mark are shown in Fig. 6a-c, respectively. The mean codebook C_m of size 64 for copyright protection and the residual codebook C_r of size 128 for content authentication are obtained by the well-known LBG algorithm based on four training images, Lena, Peppers, Baboon and F16, as shown in Figs. 4a-d. The two binary images of size 128×128 obtained with the mapping functions MF1 and MF2 are shown in Fig. 6d, e and the final obtained two fingerprints are shown in Fig. 6f, g. The copyright logo and authentication mark extracted from the image without any attack are shown in Fig. 6h, i. To check the robustness and authentication ability of our algorithm, we perform several attacks on the original image, including JPEG compression with QF = 50, cropping in the middle part of the image, rotation by angle 1°, median filtering with the radius of 3 pixels, contrast enhancement by 10% and alteration of eyes. The attacked images are shown in Fig. 5a-f. The extracted logos and marks against these attacks are shown in Fig. 6j-u. From these results, we can easily see that the proposed method is effective.

CONCLUSIONS

In this study, we propose a novel multipurpose perceptual image hashing scheme based on mean-removed VQ by performing different mapping functions on mean and residual index tables respectively. In contrast with the traditional perceptual image hashing schemes, our scheme can be used to protect copyright and authenticate image content simultaneously and the authentication process can be visually recognized. Experimental results show that the hashing process for copyright protection is robust against most common image processes and the hashing process for content authentication is fragile.

REFERENCES

Dittmann, J., A. Steinmetz and R. Steinmetz, 1999. Content based digital signature for motion picture authentication and content-fragile watermarking. Proc. IEEE Int. Conf. Multimedia Comput. Syst., 2: 209-213.

Fiaidhi, J.A.W. and S.M.A. Mohammed, 2003. Towards developing watermarking standards for collaborative e-learning systems. Inform. Technol. J., 2: 30-34.

Khan, A., X. Niu and Z. Yong, 2008. A robust framework for protecting computation results of mobile agents. Inform. Technol. J., 7: 24-31.

Lei, Y.Q., K.Y. Chau, Z.M. Lu and W.H. Ip, 2010. DCT-domain global feature and DWT-domain least-squares line fitting based local feature for robust image hashing. Int. J. Innovative Comput. Inform. Control, 6: 2513-2521.

Linde, Y., A. Buzo and R.M. Gray, 1980. An algorithm for vector quantizer design. IEEE Trans. Commun., 28: 84-95.

Lu, Z.M. and S.H. Sun, 2000. Digital image watermarking technique based on vector quantisation. Electronics Lett., 36: 303-305.

Lu, Z.M., J.S. Pan and S.H. Sun, 2000. VQ-based digital image watermarking method. Electronics Lett., 36: 1201-1202.

Lu Z.M., C.H. Liu and D.G. Xu, 2003. Semi-fragile image watermarking method based on index constrained vector quantization. Electronics Lett., 39: 35-36.

Lu, C.S. and H.Y.M. Liao, 2003. Structural digital signature for image authentication: An incidental distortion resistant scheme. IEEE Trans. Multimedia, 5: 161-173.

Lu, Z.M., D.G. Xu and S.H. Sun, 2005. Multipurpose image watermarking algorithm based on multistage vector quantization. IEEE Trans. Image Process., 14: 822-831.

Lu, Z.M. and S.Z. Li, 2006. Multipurpose watermarking algorithm for secret communication. Chinese J. Electronics, 15: 79-84.

Monga, V. and B.L. Evans, 2006. Perceptual Image Hashing Via Feature Points: Performance Evaluation and Tradeoffs. Proc. IEEE Trans. Image, 15: 3452-3465.

Monga, V. and M.K. Mhcak, 2007. Robust and secure image hashing via non-negative matrix factorizations. IEEE Trans. Inform. Forensics Secur., 2: 376-390.

Monga, V., A. Banerjee and B.L. Evans, 2006. A clustering based approach to perceptual image hashing. IEEE Trans. Inform. Forensics Security, 1: 68-79.

Qureshi, M.A. and R. Tao, 2006. A comprehensive analysis of digital watermarking. Inform. Technol. J., 5: 471-475.

Venkatesan, R., S.M. Koon, M.H. Jakubowski and P. Moulin, 2000. Robust image hashing. Proc. IEEE Conf. Image Process., 3: 664-666.

Yu, F.X., Y.Q. Lei, Y.G. Wang and Z.M. Lu, 2010. Robust image hashing based on statistical invariance of DCT coefficients. J. Inform. Hiding Multimedia Signal Process., 1: 294-300.