Information Sciences 426 (2018) 1–18

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Selective bit embedding scheme for robust blind color image watermarking[☆]

Thien Huynh-The[a], Cam-Hao Hua[a], Nguyen Anh Tu[a], Taeho Hur[a], Jaehun Bang[a], Dohyeong Kim[a], Muhammad Bilal Amin[a,d], Byeong Ho Kang[b], Hyonwoo Seung[c], Sungyoung Lee[a,*]

[a] Department of Computer Science & Engineering, Kyung Hee University (Global Campus), 1732 Deokyoungdae-ro, Giheung-gu, Yongin-si, Gyeonggi-do 446-701, Republic of Korea
[b] School of Computing and Information System, University of Tasmania, Hobart, TAS 7005, Australia
[c] Department of Computer Science, Seoul Women's University, 621 Hwarang-ro, Gongneung 2(i)-dong, Nowon-gu, Seoul, Republic of Korea
[d] National Research Foundation of Korea, 201 Gajeong-ro, Yuseong-gu, Daejeon 34113, Republic of Korea

A R T I C L E   I N F O

Article history:
Received 10 May 2017
Revised 20 September 2017
Accepted 6 October 2017
Available online 13 October 2017

Keywords:
Color image watermarking
Discrete wavelet transform
Quantization technique
Binary bit decomposition
2D Otsu thresholding

A B S T R A C T

In this paper, we propose a novel robust blind color image watermarking method, namely SMLE, that allows to embed a gray-scale image as watermark into a host color image in the wavelet domain. After decomposing the gray-scale watermark to component binary images in digits ordering from least significant bit (LSB) to most significant bit (MSB), the retrieved binary bits are then embedded into wavelet blocks of two optimal color channels by using an efficient quantization technique, where the wavelet coefficient difference in each block is quantized to either two pre-defined thresholds for corresponding 0-bits and 1-bits. To optimize the watermark imperceptibility, we equally split the coefficient modified quantity on two middle-frequency sub-bands instead of only one as in existing approaches. The improvement of embedding rule increases approximately 3 dB of watermarked image quality. An adequate trade-off between robustness and imperceptibility is controlled by a factor representing the embedding strength. As for extraction process, we exploit 2D Otsu algorithm for higher accuracy of watermark detection than that of 1D Otsu. Experimental results prove the robustness of our SMLE watermarking model against common image processing operations along with its efficient retention of the imperceptibility of the watermark in the host image. Compared to state-of-the-art methods, our approach outperforms in most of robustness tests at a same high payload capacity.

☆ This work was supported by the Industrial Core Technology Development Program (10049079, Develop of mining core technology exploiting personal big data) funded by the Ministry of Trade, Industry and Energy (MOTIE, Korea). This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (No.2017-0-00655). This research was supported by Korea Research Fellowship program funded by the Ministry of Science, ICT and Future Planning through the National Research Foundation of Korea (NRF-2016H1D3A1938039).

* Corresponding author.
E-mail addresses: thienht@oslab.khu.ac.kr (T. Huynh-The), hao.hua@oslab.khu.ac.kr (C.-H. Hua), tunguyen@khu.ac.kr (N.A. Tu), hth@oslab.khu.ac.kr (T. Hur), jhb@oslab.khu.ac.kr (J. Bang), dhkim@oslab.khu.ac.kr (D. Kim), m.b.amin@ieee.org (M.B. Amin), byeong.kang@utas.edu.au (B. Ho Kang), hwseung@swu.ac.kr (H. Seung), sylee@oslab.khu.ac.kr (S. Lee).

https://doi.org/10.1016/j.ins.2017.10.016

1. Introduction

Nowadays with the explosion of mobile devices and the Internet, digital images are easily and ubiquitously captured, stored and shared on common social networks such as Facebook, Twitter and Instagram. They are usually uploaded to the Internet directly using wireless communication channels without any preliminary protection schemes [40]. This makes several urgent issues relating to copyright protection and authentication in transmission, storage, and usage of images [12]. For example, a personal image shared on a social network can be accessed, downloaded, modified, and reused by the others illegally for commercial or other purposes. To prevent such kind of risks, digital image watermarking techniques are needed for owner authentication. Digital watermarking is further applied in video streaming services [7,28].

A general digital image watermarking model fundamentally includes two processes: (i) the embedding process encodes the watermark containing the private information of the owners into the host image in order to be invisible to the human eyes and (ii) the extraction process recovers the watermark from the embedded image to retrieve the hidden information for originality authentication. Most of existing digital image watermarking methods deal with the perceptibility in the embedment process and the robustness in the extraction process. At first, embedding a watermark requires a modification on the host image, which ordinarily degrades the quality of the host image and exposes the hidden information [1]. Many proposed models, therefore, concentrate to minimize the perceptibility of the watermark on gray-scale image instead of color images due to less visual recognition. Secondly, the watermark must be robust to withstand common image processing operations to precisely recover the hidden information from the embedded image [14]. In practice, the watermark can be modified, destroyed, and even removed from the host image by combining several operations. Additionally, the copyright privacy of an owner, an important impact whenever developing image watermarking models, is thoughtfully considered. As known as the most challenging model, the blind watermarking [23,39] requires neither the original image nor the watermark for the recovery process while the semi-blind watermarking [21] and non-blind watermarking [25] need the original host image and all of them, respectively. It is obvious to realize that the blind watermarking scheme is more convenient than the others. However, it should be noted that a secret key which contains variety of information such as wavelet block locations, embedded color profiles, permutation code, and etc., must be provided for extracting watermark. Another important aspect, which is usually ignored when evaluating a watermarking model, is the payload capacity. Most of existing works are developed for gray-scale images [13] and color images [10] with a binary image served as watermark [31,33]. However, the utilization of a binary image as the watermark has some limitations. At first, since binary images contain less information than gray-scale images, the content represented in binary images is more difficult for recognition than that of gray-scale images in case they are attacked by image processing operations. Secondly, binary images utilized as watermarks are sensitive and fragile to common image transformations due to only two states of a bit for decision. In the case of using gray-scale images as the watermark, as long as some most significant bits are correctly detected, we can precisely recover the primary content without paying much attention to remaining less significant bits. Currently, quite a few models use gray-scale image as watermark [25], in which, their performance in terms of imperceptibility and robustness is not impressive due to the constraint to payload capacity. Other approaches have trouble with high payload embedding, for example, the capacity of a gray-scale image is eight times the capacity of a binary image at the same resolution because the bit depth of gray-scale and binary image is 8 and 1, respectively. Hence, three major characteristics, i.e. imperceptibility, robustness, and payload capacity should be firstly addressed in any image watermarking models. The benefits of using gray-scale images as watermarks and also the challenges of developing a high payload embedding scheme for color images motivate us in this research.

In this work, we develop a robust color image watermarking method, namely selective MSB-LSB embedding (SMLE), where gray-scale image is aimed as watermark. A gray-scale watermark is decomposed to eight component binary images following digits ordering which are then embedded into middle-frequency sub-bands of the host color image in the wavelet domain. The embedding process is performed by a quantization technique where the coefficient difference of each wavelet block is quantized to two pre-defined thresholds by an improved embedding rule. The total error of coefficient modification caused on the host image by the embedding rule is minimized by equally spreading adjustment quantities of quantization on two sub-bands simultaneously. In order to precisely recover watermark in the extraction process, a classification threshold for detecting binary bits hidden in wavelet blocks of the watermarked image is calculated by 2D Otsu algorithm. Compared with existing approaches, the proposed watermark model has following advantages: (i) a selective MSB-LSB embedding scheme for efficiently hiding a gray-scale watermark image into a color image, (ii) an efficient embedding rule used for optimal coefficient quantization to improve image imperceptibility, (iii) a factor representing the embedding strength for a pleasant trade-off between robustness and imperceptibility, and (iv) an adaptive classification threshold defined by 2D Otsu algorithm for extracting watermark accurately. Although the model is designed with a very high embedding rate of 1/64 byte/pixel (or 1/8 bit/pixel), it achieves high robustness against typical image manipulations like average filtering, median filtering, linear motion blurring, scaling, additional noise, and JPEG compression.

The remaining of the paper is organized as follows. Related works in the field of digital image watermarking are reviewed in Section 2. Section 3 describes the proposed digital image watermarking model for color image. Experimental results and discussion are presented in Section 4. Finally, conclusion and future work are outlined in Section 5.

2. Related work

Due to natural limitations of watermarking in the spatial domain involving visualization and robustness, most image watermarking models are built on transformed domain, such as the cosine transform [8,18,23], Fourier transform [39], contourlet transform [22], wavelet transform [10,16], and gyrator transform [26]. Hu et al. [8] modulated a partly sign-altered (PSA) mean of selective discrete cosine transform (DCT) coefficients for binary embedding. The method is successful in enhancing imperceptibility of the watermark by handling the variation margin of each DCT coefficient based on either the JPEG quantization table or the just noticeable distortion (JND). Based on quantizing DCT coefficient difference of adjacent blocks at the same position to a particular pre-defined range, Parah et al. [18] modified coefficients by an amount which is depended upon the binary watermark bit and the median of certain zig-zag ordered AC-coefficients. Roy and Pal [23] developed a multiple watermarking approach which allows to embed two binary watermark images by adjusting the middle significant AC-coefficients of green and blue components of the host image. Although the multiple watermarking is so convenient for the multiple owners' authentication, it needs to be improved in terms of robustness, imperceptibility, and computational complexity. In [39], a good visual quality of embedded images was achieved by encoding binary bits into real quaternion Fourier coefficients, but the approach significantly consumed computational resources for least squares support vector machine (LS-SVM) training in the extraction process. Ranjbar et al. [22] developed a two-stage embedding scheme to preserve the visualization of host images by embedding a binary watermark bit for only one contourlet coefficient. The method is highly robust under several common attacks, however, its extracting speed is very slow because of the local and global watermark embedment on low and high frequency sub-bands. Compared to the cosine transform, the wavelet transform is much more efficient in multi-resolution representation and analysis of an image in the frequency domain, so it has been widely used in the field of image processing, for example, image decomposition for de-noising, quality assessment, feature extraction, image watermarking [9,10,36] and classification [27]. Nguyen et al. [16] and Huynh-The et al. [9,10] performed image watermarking in the wavelet domain for gray-scale and color image, respectively. Nguyen et al. proposed a reversible watermarking scheme that achieves high accuracy of watermark detection with less degradation of image quality. By formulating a color channel selection algorithm, a blind color image watermarking approaches [9,10] achieved high imperceptibility while maintaining great robustness of watermark under various image processing manipulations. A double random phase encoding (DRPE) scheme [26] was developed for image authentication with more security in the quaternion gyrator transform (GT) domain, an extension of the 2D fractional Fourier transform (FRFT). Moreover, some other transformation techniques have been explored to particularly address the geometric distortion such as nonsubsampled shearlet transform (NSST) [38], polar harmonic transform (PHT) [37], quaternion radial moments (QRMs) [31], radial harmonic fourier moments (RHFMs) [6], and quaternion exponent moments (QEMs) [20]. By optimizing NSST for image decomposition, Wang et al. [37] developed PHT to estimate the geometrical degradations parameters for precise watermark extraction. Besides expensively computational requirement, the approach is fragile to random bending, and column or line removal. Similarly, Wang et al. [6,20] calculated and selected the most robust fourier moments for moment magnitudes quantization, which was served for watermark embedment. However, rigorous design for the square image and poor robustness of the large scale cropping are two typical limitations of moments-based watermarking models compared to other approaches.

Recently, machine learning techniques have been utilized for digital image watermarking such as genetic algorithms (GAs) [15], artificial bee colony (ABC) algorithm [1,3], hidden Markov model (HMM) [4,33], support vector machines (SVMs) [34], artificial neural networks (ANNs) [2,29]. In details, Moghaddam and Nemati [15] presented a spatial watermarking model which applies imperialistic competition algorithm (ICA) to seek optimal regions for embedding. Due to processing on the spatial domain, the watermark extracted from the model is breakable under such frequency domain based operations as lossy JPEG compression. Ali et al. [3] employed ABC algorithm to determine the optimal threshold for singular value decomposition (SVD) coefficient quantization and the compensation parameter for visible distortion measurement in case of imperceptibility increment. In [1], the meta-heuristic technique developed on ABC algorithm was evaluated for several recent image watermarking models to prove its efficiency in the issue of multi-objectives optimization. Amini et al. [4] and Wang et al. [33] exploited HMM in the wavelet domain for locally optimal watermark detection. Although HMM-based watermarking models provide an efficient trade-off between robustness and imperceptibility, significantly increased complexity becomes a critical challenge for real-time processing. In [34], a synchronous correction algorithm was developed for watermark decoding using fuzzy least squares support vector machine (FLS-SVM). In training stage, the shape and scale parameters of Bassel K form (BKF) distribution of wavelet coefficients extracted from image dataset are fed to construct the FLS-SVM model. While Agarwal et al. [2] trained a hybrid back propagation neural network with genetic algorithm (GA-BPN) by inference rules which contains three human visual system (HVS) parameters standing for luminance, edge, and contrast sensitivity, Tsai and Liu [29] learned an ANN model with the JND profile to estimate watermarking without the original image as requirement. Almost machine learning based watermarking approaches deliver comparative performance in terms of robustness and imperceptibility, however, computational complexity is the major limitation compared to non-training methods.

Discrete wavelet transform (DWT), is one of the most popular time–frequency transformations, which has been widely used in digital image watermarking. In general, both embedding and extraction processes are employed in the wavelet domain by modifying DWT coefficients. Run et al. [24] encoded binary bits into wavelet-trees which are constructed by one 4-DWT coefficient and four 3-DWT coefficients of the low–high sub-band. Huynh-The et al. [10] grouped one coefficient of the low–high sub-band and one coefficient of the high–low sub-band to be a block for hiding binary information. Most

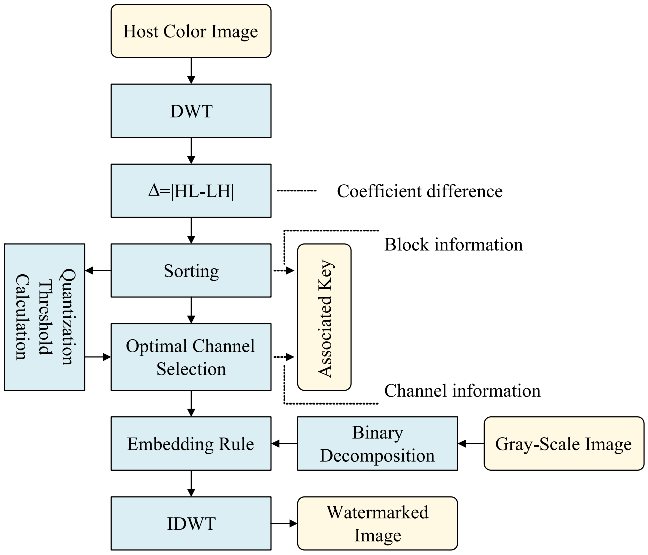


Fig. 1. The workflow of embedding process.

of the state-of-the-art watermarking approaches are proposed for encoding binary bits with a limited payload capacity, for instance, the embedding rate factor is 1/256 bit/pixel [24] and 1/64 bit/pixel [2,29,33] for host gray-scale image, 1/16 bit/pixel [10] for host color image. It can be realized that payload increment in wavelet-based watermarking models still remains as a touch challenge. Hence, controlling the performance balance of triangular imperceptibility-robustness-payload criteria is needed to build an efficient watermarking model. Obviously, the use of gray-scale image as watermark is rarely presented in color image watermarking systems, especially in the transformed domains. Recently, Saboori and Hosseine [25] proposed a novel color image watermarking method where a gray-scale image is served as watermark, however, its performance is not impressive and needs to be improved significantly.

3. The methodology

This section presents the proposed blind color image watermarking approach which consists of the watermark embedding and extraction processes.

3.1. Watermark embedding process

The embedding process performs the information encoding in the wavelet domain following the workflow in Fig. 1. At first, a gray-scale watermark image W is decomposed to eight binary images $w_{BI1:\,BI8}$, respectively representing order digits from LSB to MSB as shown in Fig. 2. The binary bits of these images are then embedded to a host color image I by a quantization technique in which the coefficient difference $\Delta_{i,k} = \left|HL_i^k - LH_i^k\right|$ of ith wavelet block in kth color channel is encoded to the pre-defined thresholds δ_0 and δ_1 for 0-bits and 1-bits, respectively, by varying wavelet coefficient values. HL (high–low) and LH (low–high), the coefficients of two middle-frequency sub-bands holding vertical and horizontal details, are obtained by applying DWT decomposition to color channels, e.g. red (R), green (G) and blue (B) of the host image. Following [10], wavelet blocks are sorted in the ascending order of difference, denoted $\Delta_{i,k}^S$, to code 0-bits for blocks of smallest coefficient differences and 1-bits for blocks of greatest coefficient differences. Two optimal color channels are chosen for binary embedment to minimize the differences between $\Delta_{i,k}^S$ and δ_0 for $w = 0$ and δ_1 for $w = 1$. Concretely, the watermark bits of four binary images $w_{k^{\#1}} = \{w_{BI2}, w_{BI4}, w_{BI6}, w_{BI8}\}$ are encoded to the first optimal channel $k^{\#1}$ and the watermark bits of remaining images $w_{k^{\#2}} = \{w_{BI1}, w_{BI3}, w_{BI5}, w_{BI7}\}$ are encoded to the second optimal channel $k^{\#2}$ as illustrated in Fig. 3(a). Channel selection is executed as follows:

$$
\begin{aligned}
k^{\#1} &= \begin{cases} \arg\min\limits_k \left(\left|\Delta_{i,k}^S - \delta_0\right|\right) & \forall bit_i \in w_{k^{\#1}} = 0 \\ \arg\min\limits_k \left(\left|\Delta_{i,k}^S - \delta_1\right|\right) & \forall bit_i \in w_{k^{\#1}} = 1 \end{cases} \\
k^{\#2} &= \begin{cases} \arg\min\limits_k \left(\left|\Delta_{i,-k^{\#1}}^S - \delta_0\right|\right) & \forall bit_i \in w_{k^{\#2}} = 0 \\ \arg\min\limits_k \left(\left|\Delta_{i,-k^{\#1}}^S - \delta_1\right|\right) & \forall bit_i \in w_{k^{\#2}} = 1 \end{cases}
\end{aligned} \tag{1}
$$

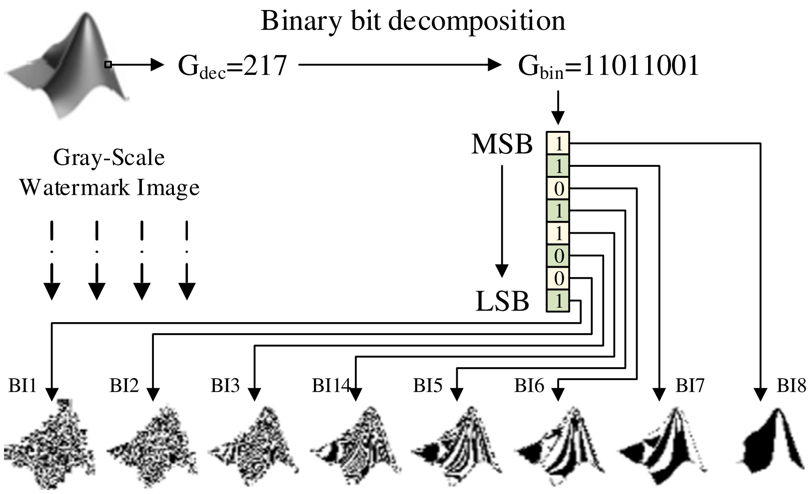


Fig. 2. The illustration of bit decomposition. The gray-scale watermark is decomposed to eight binary images respectively representing digits from LSB to MSB.

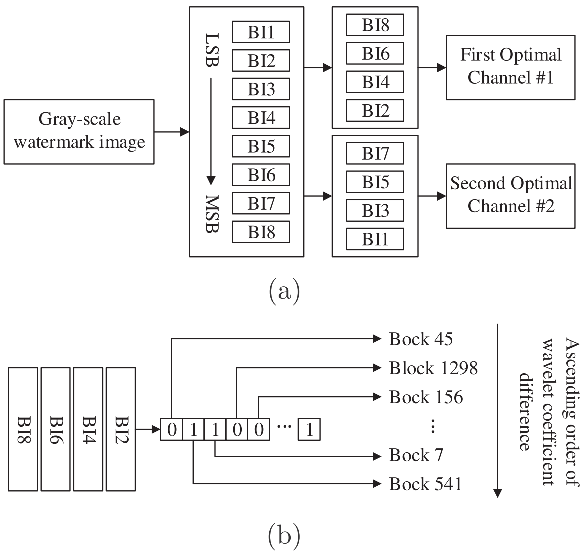


Fig. 3. The proposed selective scheme for watermark embedment: (a) splitting binary images into two classes to encode with two optimal channels, (b) embedding binary bits into ordered wavelet blocks in the case with the first optimal channel.

The notation $-k^{\#1}$ indicates the remaining color channels except $k^{\#1}$. In advance, two quantization thresholds δ_0 and δ_1 are calculated as follows:

$$(\delta_0 = \upsilon, \delta_1 = \upsilon + \lambda) = \arg\min_{\upsilon} (MSE_0 + MSE_1) \tag{2}$$

where λ is the robustness factor signifying the strength of watermark in the host image and υ is a value at which the total adjustment error in the embedment stage is minimized as small as possible. MSE_0 and MSE_1 are the corresponding mean square errors of 0-bits and 1-bits embedment, respectively, and defined as follows:

$$\begin{aligned} MSE_0 &= \tfrac{1}{(N_0-n)} \sum_{i=n \left| \Delta_n^S > \upsilon \right.}^{N_0} \left(\Delta_i^S - \upsilon \right)^2 \\ MSE_1 &= \tfrac{1}{(m-N_0-1)} \sum_{i=N_0+1}^{m \left| \Delta_m^S = \upsilon+\lambda \right.} \left(\upsilon + \lambda - \Delta_i^S \right)^2 \end{aligned} \tag{3}$$

where N_0 is the number of 0-bits of all binary images $w_{B1:\,B8}$. A larger λ presents the stronger embedment of watermark on the host image, that means, the watermark is more robust under several digital image transformations, yet its imperceptibility on the host image is worse and worse. Therefore, λ should be flexibly selected for a reasonable balance between watermark robustness and image imperceptibility. Encoding binary bits is executed by an embedding rule which is developed to correspondingly adapt the binary images in $w_{k^{\#1}}$ and $w_{k^{\#2}}$ with two already determined optimal channels. This process is implemented by replacing k^* by $k^{\#1}$ and $k^{\#2}$ as follows:

For 0-bits coding If $\Delta^S_{i,k^*} > \delta_0$

$$\begin{aligned} LH_i^{k^*} \geq HL_i^{k^*} &\rightarrow \begin{cases} LH_i^{k^*} = LH_i^{k^*} - \chi_i^0/2 \\ HL_i^{k^*} = HL_i^{k^*} + \chi_i^0/2 \end{cases} \\ LH_i^{k^*} < HL_i^{k^*} &\rightarrow \begin{cases} LH_i^{k^*} = LH_i^{k^*} + \chi_i^0/2 \\ HL_i^{k^*} = HL_i^{k^*} - \chi_i^0/2 \end{cases} \end{aligned} \tag{4}$$

If $\Delta^S_{i,k^*} \leq \delta_0$

$$\begin{aligned} LH_i^{k^*} &= LH_i^{k^*} \\ HL_i^{k^*} &= HL_i^{k^*} \end{aligned} \tag{5}$$

For 1-bits coding

If $\Delta^S_{i,k^*} < \delta_1$

$$\begin{aligned} LH_i^{k^*} \geq HL_i^{k^*} &\rightarrow \begin{cases} LH_i^{k^*} = LH_i^{k^*} + \chi_i^1/2 \\ HL_i^{k^*} = HL_i^{k^*} - \chi_i^1/2 \end{cases} \\ LH_i^{k^*} < HL_i^{k^*} &\rightarrow \begin{cases} LH_i^{k^*} = LH_i^{k^*} - \chi_i^1/2 \\ HL_i^{k^*} = HL_i^{k^*} + \chi_i^1/2 \end{cases} \end{aligned} \tag{6}$$

If $\Delta^S_{i,k^*} \geq \delta_1$

$$\begin{aligned} LH_i^{k^*} &= LH_i^{k^*} \\ HL_i^{k^*} &= HL_i^{k^*} \end{aligned} \tag{7}$$

where $\chi_i^0 = \Delta^S_{i,k^*} - \delta_0$ and $\chi_i^1 = \delta_1 - \Delta^S_{i,k^*}$ represent the modification quantities of original coefficients required to encode 0-bits and 1-bits, respectively. Compared to [10] where the modification is done on either LH or HL by χ_i^0 or χ_i^1, we split this adjustment quantity equally for two sub-bands to preserve the visual quality of the watermarked image. The image quality improvement is analyzed by Peak Signal-to-Noise Ratio as follows:

$$MSE_{[10]} = \frac{1}{N}\sum_{i=1}^{N}(\chi_i)^2 \tag{8}$$

$$\begin{aligned} MSE_{SMLE} &= \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\chi_i}{2}\right)^2 + \left(\frac{\chi_i}{2}\right)^2 \\ &= \frac{1}{2N}\sum_{i=1}^{N}(\chi_i)^2 \\ &= \frac{1}{2}MSE_{[10]} \end{aligned} \tag{9}$$

$$\begin{aligned} \Delta_{PSNR} &= PSNR_{SMLE} - PSNR_{[10]} \\ &= 10\log_{10}\left(\frac{255^2}{MSE_{SMLE}}\right) - 10\log_{10}\left(\frac{255^2}{MSE_{[10]}}\right) \\ &= 10\log_{10}2 \end{aligned} \tag{10}$$

Thanks to the improved embedding rule, the quality of watermarked image is increased approximately $10\log_{10}(2) \approx 3$ dB. According to the embedding scheme represented in Fig. 3(b), watermark bits are encoded following the ordering of significance, concretely, beginning with w_{B8} and ending with w_{B2} in group $w_{k^{\#1}}$ for the first optimal channel embedment, and beginning with w_{B7} and ending with w_{B1} in group $w_{k^{\#2}}$ for the second optimal channel embedment. As soon as the embedding process is done, the modified coefficient differences are either less than δ_0 or greater than δ_1. Additionally, the information of channel blocks is privately stored in an associated key, which is generated and maintained for original watermark recovery throughout the extraction process. Once the embedding process is completed, we rebuild the LH and HL sub-bands of all color channels by organizing the modified coefficients. As well as, the watermarked image is reconstructed from color-component images which are re-transformed by Inverse Discrete Wavelet Transform (IDWT). The particular notations used in our proposed watermarking method are summarized in Table 1. It is important to note that the security level of the gray-scale watermark can be increased by scrambling the binary bits using Arnold transformation. For the proposed method, this step can be done after the binary decomposition step in Fig. 1.

Table 1
Important notations used in the proposed watermarking method.

Notation	Description
$\Delta_{i,k}$	The coefficient difference of ith wavelet block in kth color channel
$\Delta^S_{i,k}$	The coefficient difference $\Delta_{i,k}$ after sorting
$k^{\#1}$	The first optimal channel
$k^{\#2}$	The second optimal channel
$w_{k^{\#1}}$	The MSB images assigned for embedding with $k^{\#1}$
$w_{k^{\#2}}$	The LSB images assigned for embedding with $k^{\#2}$
δ_0	The quantization threshold for 0-bit embedding
δ_1	The quantization threshold for 1-bit embedding
v	The value at which the total embedding error is minimized
λ	The robustness factor
χ_i^0	The modification quantity for 0-bit embedding
χ_i^1	The modification quantity for 1-bit embedding
δ_Δ	The extraction threshold defined by 2D Otsu method

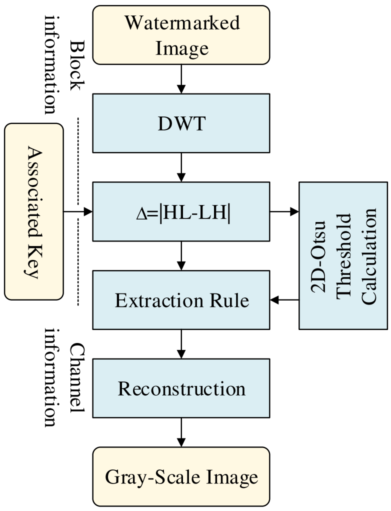


Fig. 4. The workflow of extraction process.

3.2. Watermark extraction process

As shown in Fig. 4, the first step of the extraction process is to calculate DWT coefficient differences of encoded blocks with the key that was created in the embedding stage. With a classification threshold denoted δ_Δ where $\delta_0 < \delta_\Delta < \delta_1$, watermark bits are basically recovered by the following comparison rule

$$w_i = \begin{cases} 1 & \forall \Delta^S_{i,k} \geq \delta_\Delta \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

It is important to note that δ_Δ must be determined in case of the unknown quantization thresholds, i.e. δ_0 and δ_1. Therefore, an adaptive two-dimensional (2D) Otsu threshold [11], regularly used in image segmentation, is computed for binary segmentation of DWT blocks. By maximizing the trace of the between-class variance matrix S_b, the threshold vector $(s = \delta_\Delta, t)$ is selected by the following equation:

$$(s, t) = \underset{\substack{0 \leq s \leq L \\ 0 \leq t \leq L}}{\arg\max} \; (Tr(S_b)) \tag{12}$$

where $L = \max\left(\Delta^S_{i,k*}\right)$ and the trace of discrete matrix is expressed as follows (more details in Appendix A)

$$Tr(S_b) = \frac{(\mu_{Ti}\omega_0 - \mu_i)^2 + \left(\mu_{Tj}\omega_1 - \mu_j\right)^2}{\omega_0(1 - \omega_0)} \tag{13}$$

In this paper, a fast recursive algorithm of 2D Otsu [41] is employed to obtain the optimal threshold s. Compared to 1D Otsu [17] exploited for watermark extraction [10], 2D Otsu algorithm adapts robustly to the noise issue in image segmentation

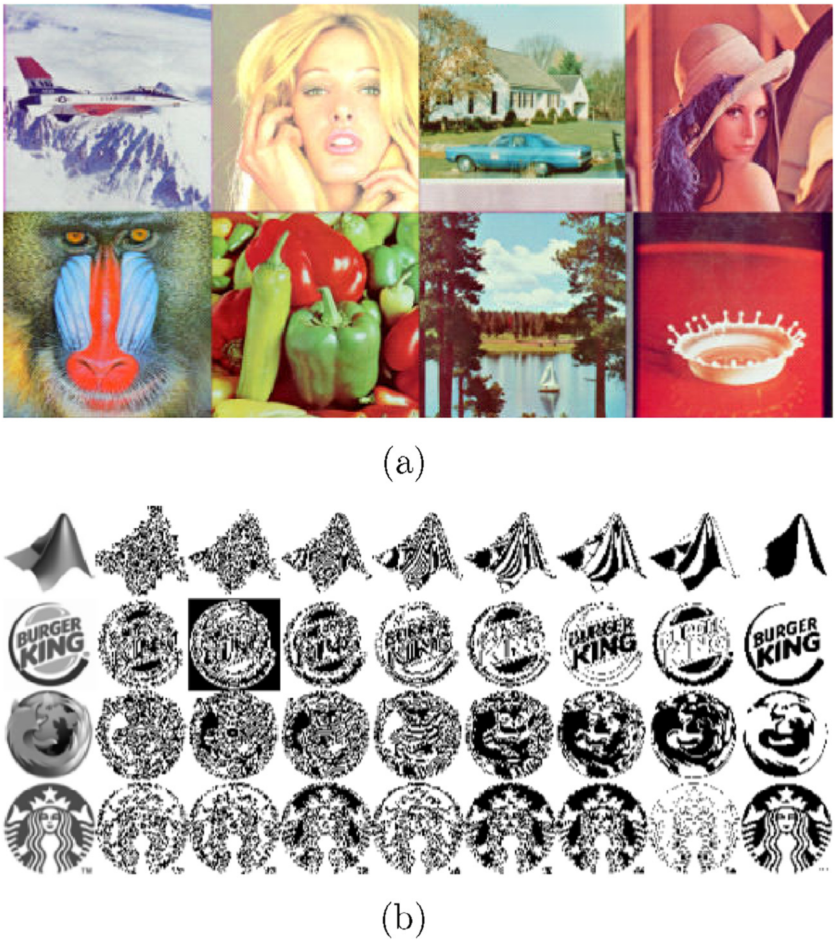


Fig. 5. Test images used for evaluation include (a) eight host color images (left to right): Airplane, Girl, House, Lena, Mandrill, Peppers, Sailboat, and Splash; (b) Four watermark samples (top to bottom): Matlab, Burger King, Firefox, and Starbucks are decomposed to eight corresponding component binary images.

due to its contents-independent characteristics. A gray-scale watermark image is finally reconstructed from eight binary images after decoding.

4. Simulation results and discussion

In this section, we benchmark the perceptibility of embedded images and the robustness of watermarks under various popular image transformations. Totally eight 512×512 color images served as host images [32] and four 64×64 gray-scale images used as watermark images are shown in Fig. 5. Color Peak Signal-To-Noise (CPSNR) and Structural Similarity Index (SSIM) [35] used to measure watermarked image quality (or the imperceptibility of a watermark in a host image) are calculated as follows:

$$CPSNR = 10\log_{10}\left(\frac{255^2}{\frac{\sum_{k=1}^{3}\sum_{i=1}^{p}\sum_{j=1}^{q}(I(i,j)-J(i,j))^2}{3\times p\times q}}\right) \tag{14}$$

$$SSIM = \frac{1}{T}\sum_{i=1}^{T} ssim_i\left(x_I, y_J\right) \tag{15}$$

where $p \times q$ is the resolution of the original I and watermarked J images, and T is the number of local windows which are divided for similarity verification. The *ssim* index, which is defined as the product of three similarity terms of luminance, contrast, and structural between two local windows x_I and x_J, is calculated as follows:

$$ssim(x,y) = [l(x,y)]^{\alpha} \cdot [c(x,y)]^{\beta} \cdot [s(x,y)]^{\gamma} \tag{16}$$

Table 2
Average PSNR and SSIM of embedded host images.

λ	CPSNR (dB)			
	Matlab	Burger King	Firefox	Starbucks
20	46.86 ± 2.91	48.44 ± 2.49	49.28 ± 1.78	48.33 ± 2.31
25	43.95 ± 2.67	45.62 ± 2.88	46.53 ± 1.96	45.55 ± 2.73
30	41.63 ± 2.43	43.29 ± 2.91	44.31 ± 2.31	43.25 ± 2.81
35	39.76 ± 2.19	41.32 ± 2.75	42.44 ± 2.52	41.30 ± 2.69
40	38.19 ± 1.98	39.64 ± 2.52	40.79 ± 2.58	39.63 ± 2.50
45	36.82 ± 1.83	38.19 ± 2.32	39.34 ± 2.51	38.19 ± 2.30
50	35.60 ± 1.71	36.92 ± 2.15	38.04 ± 2.40	36.92 ± 2.14
55	34.54 ± 1.59	35.80 ± 2.00	36.89 ± 2.27	35.79 ± 1.99
60	33.58 ± 1.48	34.78 ± 1.86	35.84 ± 2.13	34.78 ± 1.85
λ	**SSIM**			
	Matlab	Burger King	Firefox	Starbucks
20	0.997 ± 0.005	0.998 ± 0.003	0.999 ± 0.002	0.998 ± 0.003
25	0.994 ± 0.009	0.996 ± 0.006	0.997 ± 0.004	0.996 ± 0.006
30	0.991 ± 0.014	0.994 ± 0.009	0.996 ± 0.007	0.994 ± 0.009
35	0.987 ± 0.019	0.991 ± 0.013	0.994 ± 0.010	0.991 ± 0.013
40	0.983 ± 0.024	0.988 ± 0.017	0.991 ± 0.013	0.988 ± 0.017
45	0.978 ± 0.031	0.984 ± 0.022	0.988 ± 0.017	0.984 ± 0.022
50	0.972 ± 0.037	0.980 ± 0.026	0.985 ± 0.021	0.980 ± 0.027
55	0.966 ± 0.043	0.976 ± 0.032	0.981 ± 0.025	0.976 ± 0.032
60	0.960 ± 0.049	0.971 ± 0.037	0.977 ± 0.029	0.971 ± 0.037

where α, β, and γ are parameters used to adjust the relative importance of the three terms. The detail formulation and parameter set up are explained in [35]. Fundamentally, the higher the CPSNR and SSIM values are, the greater the imperceptibility of watermark in the host image is. The extraction performance is quantitatively rated by Normalized Cross-Correlation coefficient (NCC) and BER (Bit Error Rate) as follows:

$$NCC = \frac{\sum_{i=1}^{p'} \sum_{j=1}^{q'} (w(i,j) - \mu_w)\big(w'(i,j) - \mu_{w'}\big)}{(p' \times q' - 1)(\sigma_w \times \sigma_{w'})} \tag{17}$$

$$BER = \frac{B}{p' \times q'} \times 100 \tag{18}$$

where p' and q' are the height and width of the original w and extracted w' watermarks; μ and σ are the mean and standard deviation of gray values, respectively; and B is the number of erroneously detected bits. The value of NCC should converge to one for completely accurate watermark recovery. The method performance is benchmarked using Matlab R2014b on a notebook operating Windows 10 with 2.70 GHz i7-5700HQ processor and 16-GB RAM.

4.1. Watermark perceptibility

This sub-section examines the perceptibility of the watermark after being embedded into the host image by evaluating the quality of watermarked images. The effect of the robustness factor λ utilized to calculate quantization thresholds is further investigated. In particular, Table 2 reports the average CPSNR and SSIM with standard deviation of eight host images on various values of λ, and Fig. 6 shows the watermarked images of Lena sample using Starbucks as watermark. Obviously, the lower the λ is, the higher the CPSNR and SSIM are due to small modification of wavelet coefficients resulted by our improved embedding rule. However, the watermark is more fragile because the small distance between δ_0 and δ_1 does not guarantee a high-quality extraction, especially under intensive image transformations. As mentioned before, λ should be flexibly selected for a pleasant trade-off between watermark robustness and host image imperceptibility.

Compared to [10], the embedding rule in this paper is significantly improved to boost the quality of watermarked images. Based on the quantitative results plotted in Fig. 7, our improved rule is entirely better, approximately 0.002 of SSIM and 3 dB of CPSNR as formulated in (10). This proves the efficiency of equally splitting coefficient adjustment to two middle sub-bands instead of either one [10].

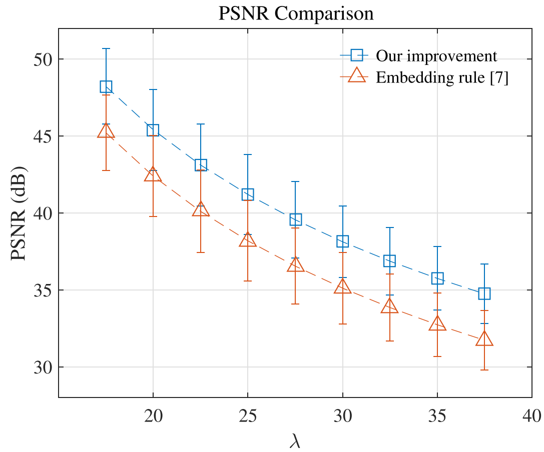
4.2. Watermark robustness

In this part, the proposed method is evaluated in aspect of hidden information recovery against attacks, i.e. digital image processing manipulations. For the latter, several popular transformations in the field are considered for robustness benchmark, such as median filtering, average filtering, linear motion blurring, size scaling, rotation, cropping, additional noise

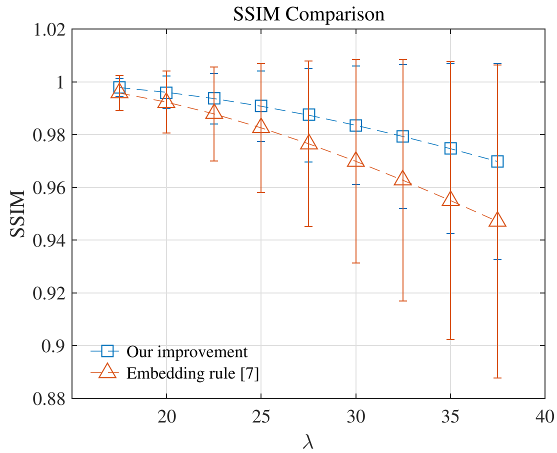
Fig. 6. Embedded images of Lena sample after embedding Starbucks watermark with various robustness values: (a) $\lambda = 20$, PSNR = 47.50 dB, (b) $\lambda = 25$, PSNR = 44.45 dB, (c) $\lambda = 30$, PSNR = 42.13 dB, (d) $\lambda = 35$, PSNR = 40.18 dB, (e) $\lambda = 40$, PSNR = 38.52 dB, (f) $\lambda = 45$, PSNR = 37.15 dB, (g) $\lambda = 50$, PSNR = 35.96 dB, (h) $\lambda = 55$, PSNR = 34.87 dB, (i) $\lambda = 60$, PSNR = 33.89 dB.

(Gaussian and Salt & Pepper), histogram equalization, and lossy JPEG compression. Fig. 8 illustrates the listed transformations and the corresponding watermarks extracted from attacked images, where Lena and Starbucks play the host and watermark role, respectively. The extraction accuracy under aforementioned attacks are presented in Table 3 with various values of λ. The quantitative results are presented as the average of eight host images associated with four watermarks. It is observed that the improvement of watermark robustness is proportional to the increment of embedding strength. However, the proposed watermarking model yields quite poor results in rotation and down-scaling attacks due to subsequent reasons: (i) DWT decomposes an image in the horizontal and vertical dimensions while rotation operates in the diagonal, (ii) scaling operation makes the loss of detail information by bi-cubic interpolation.

The robustness is further investigated under various attacking intensities of median filtering, average filtering, linear motion blurring, size scaling, rotation, Gaussian noise, Salt & Pepper noise, and lossy JPEG compression as graphically shown



(a)



(b)

Fig. 7. Watermark perceptibility comparison between the proposed embedding rule and Huynh-The et al. [10] scheme: (a) CPSNR, (b) SSIM.

in Fig. 9. As a result, the stronger the attack damages, the lower accuracy the watermark is generally extracted. We compare the watermark extraction accuracy between 2D Otsu and 1D Otsu thresholding. Based on the results in Fig. 9, 2D Otsu mostly extracts more precisely than 1D Otsu does, except Gaussian noise and lossy JPEG compression. To describe the relationship between imperceptibility and robustness in our proposed watermarking model when varying λ, we draw a performance trade-off graph as shown in Fig. 10. Our proposed SMLE method produces better imperceptibility but worse robustness when the robustness factor λ gets smaller and vice versa.

4.3. Complexity analysis

The complexity of digital image watermarking approaches is usually assessed in term of computation cost. This assessment is fundamentally important besides the performance evaluation of watermark embedding and extraction processes when developing a watermarking algorithm as an application for realistic systems. In this part, we analyze the processing time required for the embedding and extraction processes particularly. The results of processing time without optimization are plotted in Fig. 11 where the durations of major steps are indicated in details. Concretely, we measure the average processing time of embedding four gray-scale watermark images into eight host color images and also the average processing time of extracting watermarks from cover images under various attacks as digital image processing operations. It is observed that the steps of gray-scale to binary watermark conversion, DWT decomposition and IDWT construction of the host image take nearly 65% of the whole embedding processing time. Compared to 677 ms taken wholly during embedding stage, much more time is required for recovering in the extraction process, i.e. 1865 ms, in which most of duration is spent for calculating the 2D Otsu threshold. Therefore, reducing the processing time of this step needs to be done in future work.

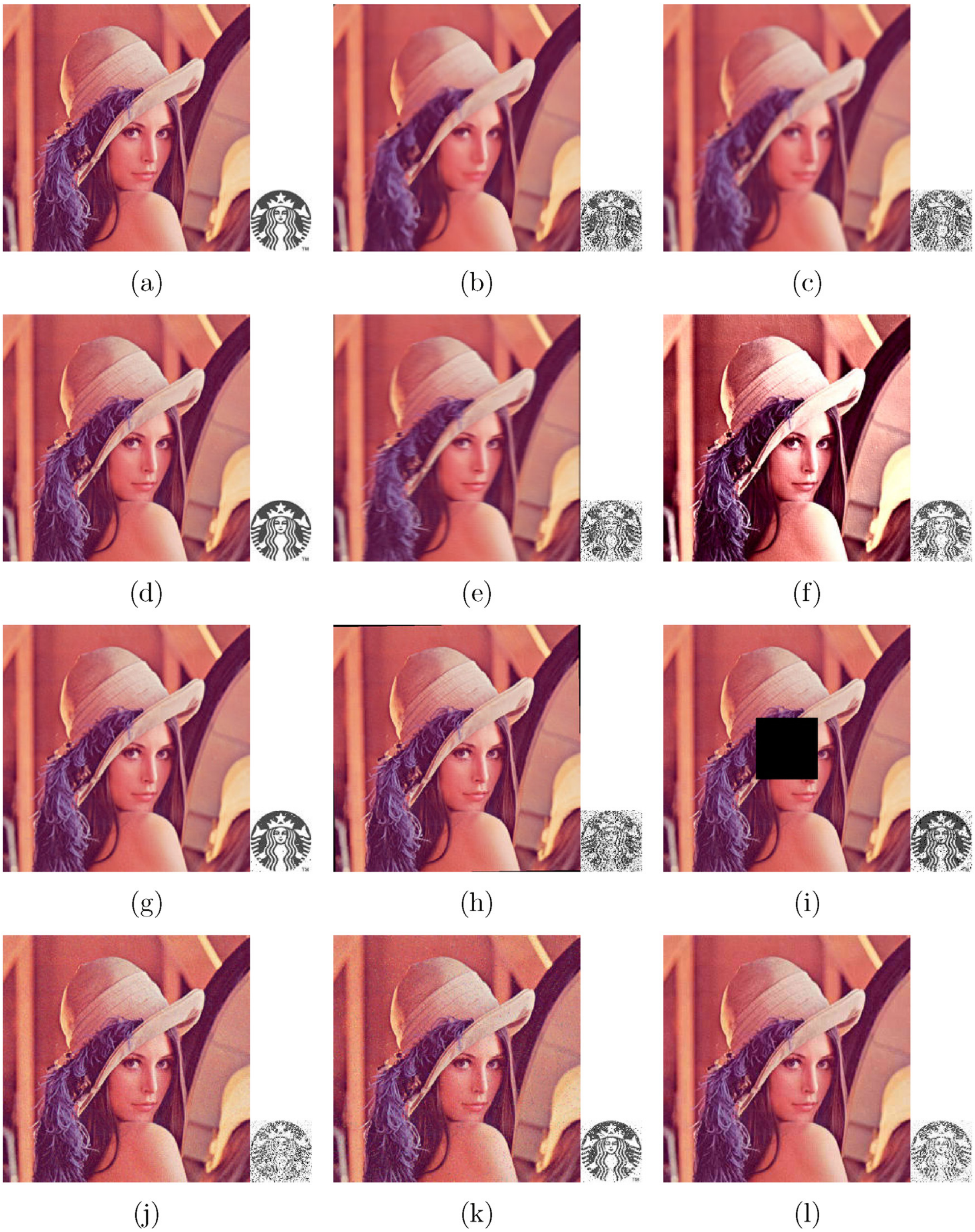


Fig. 8. Digital image processing operations on the Lena sample and corresponding extracted watermarks: (a) non-Attack, (b) median filtering 5×5, (c) average filtering 5×5, (d) Gaussian filtering 5×5, (e) motion blurring 5 pixels, (f) histogram equalization, (g) down-scaling 128×128, (h) rotation 0.5^0, (i) cropping 128×128 at the central, (j) Gaussian noise $(\mu = 0, \sigma = 0.05)$, (k) Salt & Pepper noise $(density = 5\%)$, and (l) lossy JPEG compression $(Quality = 50\%)$.

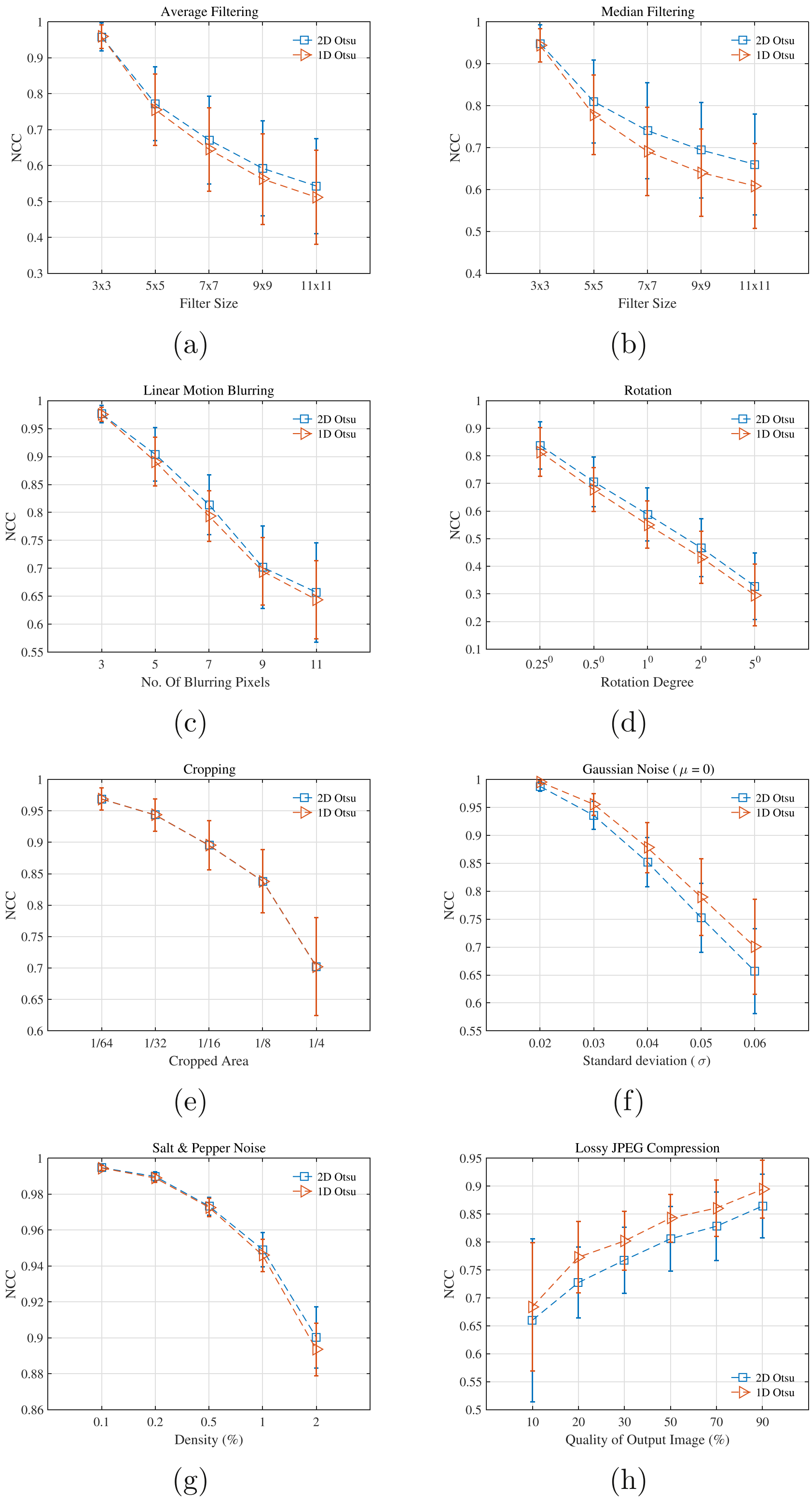


Fig. 9. Average watermark extraction accuracy measured by NCC using $\lambda = 40$ with four watermarks under various image transformations: (a) Average Filtering, (b) Median Filtering, (c) Motion Blurring, (d) Rotation, (e) Cropping, (f) Gaussian Noise, (g) Salt & Pepper Noise, and (h) Lossy JPEG Compression.

Table 3
NCC of extracted watermarks under different image transformations.

Image transformation	λ				
	20	30	40	50	60
Non-Attack	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000
Histogram equalization	0.755 ± 0.169	0.722 ± 0.180	0.699 ± 0.187	0.669 ± 0.199	0.670 ± 0.200
Median filter 3 × 3	0.924 ± 0.044	0.938 ± 0.044	0.948 ± 0.044	0.955 ± 0.043	0.960 ± 0.043
Median filter 5 × 5	0.783 ± 0.096	0.800 ± 0.097	0.810 ± 0.099	0.817 ± 0.100	0.819 ± 0.099
Median filter 7 × 7	0.716 ± 0.109	0.730 ± 0.113	0.741 ± 0.115	0.746 ± 0.115	0.751 ± 0.115
Average filter 3 × 3	0.938 ± 0.038	0.950 ± 0.038	0.958 ± 0.039	0.964 ± 0.038	0.969 ± 0.037
Average filter 5 × 5	0.757 ± 0.102	0.771 ± 0.102	0.772 ± 0.103	0.773 ± 0.103	0.771 ± 0.101
Average filter 7 × 7	0.652 ± 0.119	0.665 ± 0.120	0.671 ± 0.122	0.673 ± 0.124	0.671 ± 0.125
Gaussian filter 3 × 3	0.997 ± 0.002	0.998 ± 0.002	0.999 ± 0.001	1.000 ± 0.001	1.000 ± 0.000
Gaussian filter 5 × 5	0.997 ± 0.002	0.998 ± 0.002	0.999 ± 0.001	1.000 ± 0.001	1.000 ± 0.000
Gaussian filter 7 × 7	0.997 ± 0.002	0.998 ± 0.002	0.999 ± 0.001	1.000 ± 0.001	1.000 ± 0.000
Blurring 3 pixels	0.965 ± 0.019	0.971 ± 0.018	0.976 ± 0.016	0.980 ± 0.014	0.983 ± 0.012
Blurring 5 pixels	0.873 ± 0.048	0.892 ± 0.048	0.904 ± 0.048	0.913 ± 0.047	0.921 ± 0.045
Blurring 7 pixels	0.790 ± 0.057	0.804 ± 0.055	0.814 ± 0.053	0.821 ± 0.053	0.829 ± 0.051
Scaling up 200%	0.999 ± 0.001	0.999 ± 0.001	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000
Scaling down 50%	0.965 ± 0.023	0.973 ± 0.022	0.979 ± 0.021	0.984 ± 0.018	0.987 ± 0.015
Scaling down 25%	0.645 ± 0.121	0.658 ± 0.122	0.667 ± 0.125	0.673 ± 0.125	0.678 ± 0.125
Rotation 0.25°	0.812 ± 0.079	0.829 ± 0.082	0.838 ± 0.085	0.846 ± 0.087	0.850 ± 0.089
Rotation 0.5°	0.688 ± 0.095	0.700 ± 0.091	0.706 ± 0.090	0.705 ± 0.091	0.705 ± 0.089
Rotation 1°	0.567 ± 0.104	0.579 ± 0.099	0.589 ± 0.096	0.590 ± 0.097	0.593 ± 0.099
Rotation 2°	0.454 ± 0.105	0.462 ± 0.106	0.467 ± 0.105	0.471 ± 0.107	0.473 ± 0.107
Cropping 64 × 64	0.969 ± 0.017	0.969 ± 0.017	0.969 ± 0.018	0.969 ± 0.017	0.970 ± 0.017
Cropping 128 × 64	0.943 ± 0.025	0.944 ± 0.025	0.943 ± 0.026	0.943 ± 0.025	0.944 ± 0.024
Cropping 128 × 128	0.894 ± 0.040	0.895 ± 0.038	0.895 ± 0.039	0.896 ± 0.038	0.897 ± 0.037
Cropping 256 × 128	0.838 ± 0.053	0.837 ± 0.050	0.838 ± 0.050	0.838 ± 0.049	0.839 ± 0.049
Croping 256 × 256	0.702 ± 0.081	0.700 ± 0.078	0.702 ± 0.078	0.701 ± 0.076	0.699 ± 0.077
Gaussian noise ($\sigma = 0.02$)	0.915 ± 0.036	0.968 ± 0.015	0.988 ± 0.009	0.995 ± 0.004	0.998 ± 0.002
Gaussian noise ($\sigma = 0.03$)	0.782 ± 0.080	0.876 ± 0.043	0.936 ± 0.025	0.967 ± 0.017	0.981 ± 0.012
Gaussian noise ($\sigma = 0.04$)	0.657 ± 0.114	0.760 ± 0.068	0.852 ± 0.044	0.910 ± 0.029	0.946 ± 0.021
Gaussian noise ($\sigma = 0.05$)	0.560 ± 0.128	0.656 ± 0.091	0.753 ± 0.062	0.833 ± 0.041	0.890 ± 0.030
S&P noise ($den = 0.1\%$)	0.996 ± 0.002	0.995 ± 0.001	0.995 ± 0.002	0.995 ± 0.002	0.994 ± 0.002
S&P noise ($den = 0.2\%$)	0.991 ± 0.003	0.990 ± 0.002	0.990 ± 0.003	0.990 ± 0.003	0.990 ± 0.003
S&P noise ($den = 0.5\%$)	0.977 ± 0.005	0.976 ± 0.005	0.973 ± 0.005	0.974 ± 0.006	0.974 ± 0.005
S&P noise ($den = 1\%$)	0.954 ± 0.012	0.951 ± 0.009	0.949 ± 0.010	0.948 ± 0.011	0.948 ± 0.012
S&P noise ($den = 2\%$)	0.911 ± 0.019	0.904 ± 0.017	0.900 ± 0.017	0.898 ± 0.019	0.897 ± 0.022
Lossy JPEG ($QF = 10\%$)	0.626 ± 0.155	0.645 ± 0.152	0.660 ± 0.146	0.676 ± 0.125	0.683 ± 0.098
Lossy JPEG ($QF = 20\%$)	0.705 ± 0.104	0.721 ± 0.077	0.728 ± 0.064	0.740 ± 0.061	0.753 ± 0.060
Lossy JPEG ($QF = 30\%$)	0.729 ± 0.086	0.755 ± 0.069	0.768 ± 0.059	0.778 ± 0.060	0.785 ± 0.060
Lossy JPEG ($QF = 50\%$)	0.772 ± 0.075	0.793 ± 0.061	0.806 ± 0.058	0.812 ± 0.061	0.810 ± 0.064
Lossy JPEG ($QF = 70\%$)	0.803 ± 0.067	0.822 ± 0.060	0.828 ± 0.062	0.826 ± 0.064	0.823 ± 0.067
Lossy JPEG ($QF = 90\%$)	0.852 ± 0.061	0.858 ± 0.061	0.864 ± 0.057	0.871 ± 0.052	0.880 ± 0.049

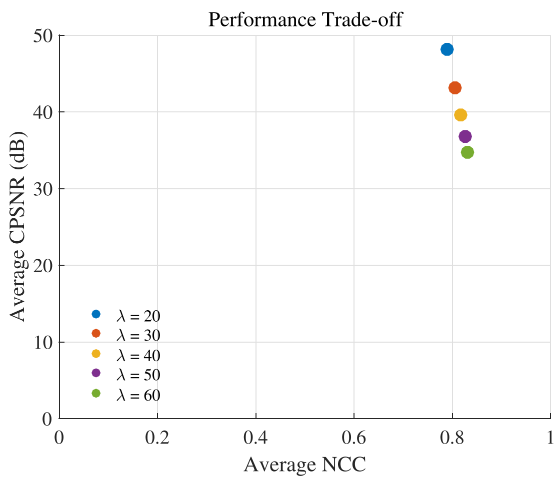


Fig. 10. The trade-off of watermarking performance between watermark robustness (NCC) and image imperceptibility (CPSNR).

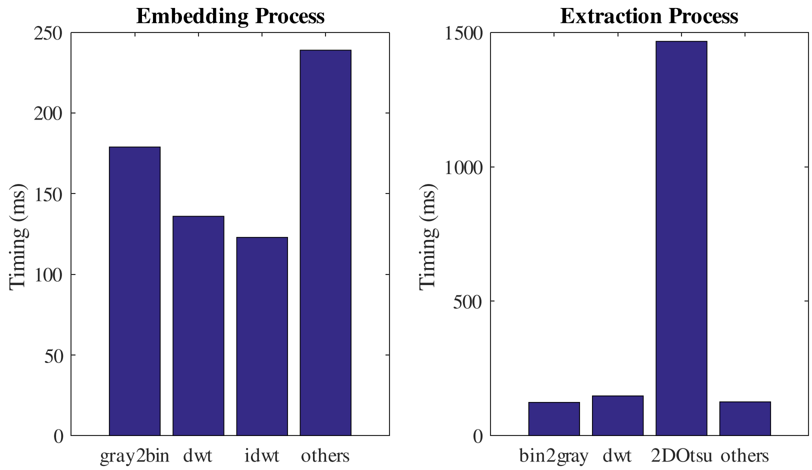


Fig. 11. Average computation time (in ms) of the proposed image watermarking method including the embedding process and the extraction process.

Table 4
Method comparison in the term of robustness for Lena sample.

Image transformation	Saboori et al. [25]		Proposed ($\lambda = 35$)
	RGB-B	YUV-Y	
Median filter 3×3	0.8510	0.8876	0.9560
Average filter 3×3	0.8454	0.8512	0.9593
Gaussian filter 3×3	0.9717	0.9788	0.9997
Rotation 0.25°	0.6234	0.6566	0.8534
Gaussian noise $\left(\sigma^2 = 0.001\right)$	0.8101	0.8838	0.8913
Salt & Pepper noise $(den = 0.1\%)$	0.9126	0.9488	0.9952
Lossy JPEG $(QF = 20\%)$	0.5543	0.6778	0.6890
Lossy JPEG $(QF = 30\%)$	0.5988	0.8123	0.7116
Lossy JPEG $(QF = 50\%)$	0.6356	0.9060	0.7409
Lossy JPEG $(QF = 70\%)$	0.7456	0.9820	0.7445
Lossy JPEG $(QF = 90\%)$	0.9188	0.9836	0.7521

4.4. Method comparison

For the last experiments, we compare the proposed watermarking method to Saboori's approach in the term of robustness. At the same embedding rate of 1/64 byte/pixel (or 1/8 bit/pixel), the quantitative results are listed in Table 4. Saboori et al. [25] embedded gray-scale watermark into color images in the spatial domain. The watermarking method, which is developed for two color-mode strategies, i.e., the blue channel of RGB and the luminance channel of YUV, encodes gray values into the first component of PCA. Although our proposed SMLE method is insignificantly worse in terms of watermark perceptibility, i.e., 40.12 dB vs 40.88 dB and 40.26 dB of RGB-B and YUV-Y, respectively, it remarkably outperforms Saboori's approach in the most of attacks in the watermark robustness benchmark, except the lossy JPEG compression. Two drawbacks of Saboori's approach are discussed: (i) the watermarking model is designed following the non-blind manner, that means, it requires the original host image in the extraction process, and (ii) the watermark is so sensitive to additional noise and geometric operations due to its spatial watermarking scheme. The results of our method in Table 4 are reported as the average NCC of Lena sample with various watermarks given $\lambda = 35$.

Although the color image watermarking method proposed in this paper is optimized for embedding a gray-scale watermark image, it is able to encode binary watermark with a minor change in the input setting. It is important to remind that our proposed SMLE method encodes binary bits which are decomposed from a gray-scale in the optimal manner with the partition of MSB and LSB binary component images. In this experiment, we do the comparisons in term of robustness, benchmarked by BER, at the same embedding rate (ER) for fairness. In particular, we compare the proposed watermarking method with other state-of-the-art approaches including Tsai and Sun [30], Peng et al. [19], Wang et al. [37], and Wang et al. [34] at $ER = 1/64$ bit per pixel (bpp); and with Chou and Liu [5] and Wang et al. [39] at $ER = 1/16$ bpp. The robustness results on several common image transformations are reported in Table 5 where one component binary image is randomly selected for encoding and Table 6 where four binary component images are randomly picked for encoding. In Table 5, our method significantly improves the watermark imperceptibility with 50.175 dB, higher than others 8.0–10.4 dB approximately, while outperforming in most of robustness tests such as lossy JPEG compression, filtering, scaling, and Salt & Pepper noise. Compared to the method [34] which is developed using the quaternion discrete Fourier transform (QDFT) and

Table 5
Method comparison in term of robustness (BER) at the same embedding rate $\frac{1}{64}$ bpp with Lena sample.

Method	Tsai and Sun [30]	Peng et al. [19]	Wang et al. [37]	Wang et al. [34]	Proposed
PSNR (dB)	41.527	42.179	39.783	40.546	50.175
Median filter 3×3	0.0300	0.4692	0.3892	0.0205	0.0071
Gaussian filter 3×3	0.0122	0.0640	0.1392	0.0027	0.0000
Average filter 3×3	0.0393	0.1084	0.1481	0.0296	0.0098
Gaussian noise $(\sigma^2 = 0.005)$	0.0109	0.0987	0.2166	0.0078	0.1011
Salt & Pepper noise $(den = 0.1\%)$	0.0075	0.0406	0.0849	0.0037	0.0054
Lossy JPEG $(QF = 90\%)$	0.0065	0.2066	0.0781	0.0061	0.0017
Lossy JPEG $(QF = 70\%)$	0.0149	0.2388	0.1485	0.0146	0.0027
Lossy JPEG $(QF = 50\%)$	0.0380	0.3399	0.2858	0.0276	0.0042
Lossy JPEG $(QF = 30\%)$	0.0927	0.4014	0.3931	0.0754	0.0146
Rotation 10°	0.0204	0.5068	0.5005	0.0066	0.3933
Scaling down 50%	0.0847	0.5114	0.5035	0.4370	0.0015

Table 6
Method comparison in term of robustness (BER) at the same embedding rate $\frac{1}{16}$ bpp with Lena sample.

Method	Chou and Liu [5]	Wang et al. [39]	Proposed
PSNR (dB)	39.92	36.10	40.89
Median filter 3×3	0.0514	0.0100	0.0270
Gaussian filter 3×3	0.0432	0.0000	0.0000
Average filter 3×3	0.0477	0.0237	0.0159
Gaussian noise $(\sigma^2 = 0.005)$	0.2122	0.0522	0.1738
Salt & Pepper noise $(den = 0.1\%)$	0.0758	0.0154	0.0029
Lossy JPEG $(QF = 90\%)$	0.0381	0.0000	0.0922
Lossy JPEG $(QF = 70\%)$	0.0555	0.0004	0.1106
Lossy JPEG $(QF = 50\%)$	0.0766	0.0208	0.1182
Lossy JPEG $(QF = 30\%)$	0.1011	0.0400	0.1367
Rotation 5°	N/A	0.0029	0.4056
Scaling down 50%	0.0402	0.0745	0.0043

fuzzy least squares support vector machine (FLS-SVM) to fight against to geometric attacks, our method is fragile under rotation. However, it should be noted that Wang's method [34] is time-consuming for training the FLS-SVM model. As a typical constraint in the field of image watermarking, either the watermark imperceptibility or the robustness will be downgraded if increasing the payload capacity. This explains for worse results of our method presented in Table 6 where the watermarking methods are compared at $ER = 1/16$ bpp. Although our method yields high quality of embedded images in PSNR metric and watermark robustness under filtering and scaling operations, it continuously gets a trouble with rotation. It shows the essential drawback of the wavelet transform compared with the real quaternion Fourier transform used in [39] for the task of domain transformation. The spatial shift, rotation, and scaling properties of quaternion Fourier transform (QFT) used in the image decomposition allow the embedded watermark to be robust against various geometric operations such as scaling and rotation.

5. Conclusions

This work proposes a robust blind digital color image watermarking method. By decomposing a gray-scale image to binary images from LSB to MSB for the embedment, a gray-scale watermark is completely encoded into a color host image using a quantization technique in the wavelet domain. Beside a color channel selection, the quality of watermarked images is significantly improved with an optimal embedding rule which minimizes the total wavelet coefficient modification. This improvement boosts approximately 3 dB of watermarked image quality. Moreover, 2D Otsu algorithm is exploited to define the classification threshold for accurate watermark extraction. The watermarking model is benchmarked on several standard color images and gray-scale images with various embedding strengths under different image processing operations. The experimental results prove that the proposed method reaches a high performance in imperceptibility of embedded host images and robustness of extracted watermarks. Compared to other similar watermarking approaches at same payload capacity, our proposed SMLE watermarking model outperforms in most of robustness tests, except lossy JPEG compression and rotation operation. For the future work, we will focus on handling the geometric attacks by exploiting quaternion discrete wavelet transform for image decomposition and will extend the proposed method for embedding color watermark image.

Appendix A. 2D Otsu algorithm

Suppose a gray-scale image with the size of $M \times N$ is presented by a gray level intensity $f(x, y)$ and its corresponding local average $g(x, y)$, ranging from 0 to $L-1$, where L is the number of gray levels. Let q_{ij} is the total number of occurrence (or frequency) of the pair (i, j) which is formed by $f(x, y) = i$ and $g(x, y) = j$. The joint probability mass function in 2-dimensional histogram is defined:

$$p_{ij} = \frac{q_{ij}}{M \times N} \tag{A.1}$$

The probability of two classes are given as

$$\omega_0 = \sum_{i=0}^{s-1} \sum_{j=0}^{t-1} p_{ij}; \quad \omega_1 = \sum_{i=s}^{L-1} \sum_{j=t}^{L-1} p_{ij} \tag{A.2}$$

The intensity means of two classes and the total mean of 2D histogram are expressed as follows:

$$\begin{aligned}
\mu_0 &= \left[\mu_{0i}, \mu_{0j}\right]^T = \left[\frac{\sum_{i=0}^{s-1} \sum_{j=0}^{t-1} i p_{ij}}{\omega_0}, \frac{\sum_{i=0}^{s-1} \sum_{j=0}^{t-1} j p_{ij}}{\omega_0}\right]^T \\
\mu_1 &= \left[\mu_{1i}, \mu_{1j}\right]^T = \left[\frac{\sum_{i=s}^{L-1} \sum_{j=t}^{L-1} i p_{ij}}{\omega_1}, \frac{\sum_{i=s}^{L-1} \sum_{j=t}^{L-1} j p_{ij}}{\omega_1}\right]^T \\
\mu_T &= \left[\mu_{Ti}, \mu_{Tj}\right]^T = \left[\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} i p_{ij}, \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} j p_{ij}\right]^T
\end{aligned} \tag{A.3}$$

The between-class variance matrix S_b is defined as

$$S_b = \sum_{k=0}^{1} \omega_k \left[(\mu_k - \mu_T)(\mu_k - \mu_T)^T\right] \tag{A.4}$$

By taking the trace of S_b following (13), where

$$\mu_i = \sum_{i=0}^{s} \sum_{j=0}^{t} i p_{ij}; \quad \mu_j = \sum_{i=0}^{s} \sum_{j=0}^{t} j p_{ij} \tag{A.5}$$

References

[1] A.M. Abdelhakim, H.I. Saleh, A.M. Nassar, A quality guaranteed robust image watermarking optimization with artificial bee colony, Expert Syst. Appl. 72 (2017) 317–326.
[2] C. Agarwal, A. Mishra, A. Sharma, Gray-scale image watermarking using GA-BPN hybrid network, J. Vis. Commun. Image Represent 24 (7) (2013) 1135–1146.
[3] M. Ali, C.W. Ahn, M. Pant, P. Siarry, An image watermarking scheme in wavelet domain with optimized compensation of singular value decomposition via artificial bee colony, Inf. Sci. (NY) 301 (2015) 44–60.
[4] M. Amini, M.O. Ahmad, M. Swamy, A new locally optimum watermark detection using vector-based hidden Markov model in wavelet domain, Signal Process. 137 (2017) 213–222.
[5] C.H. Chou, K.C. Liu, A perceptually tuned watermarking scheme for color images, IEEE Trans. Image Process. 19 (11) (2010) 2966–2982.
[6] C.-P. Wang, X.-Y. Wang, Z.-Q. Xia, Geometrically invariant image watermarking based on fast radial harmonic fourier moments, Signal Process. Image Commun. 45 (2016) 10–23.
[7] K.-L. Chung, C.-Y. Chiu, T.-Y. Yu, P.-L. Huang, Temporal and spatial correlation-based reversible data hiding for RGB CFA videos, Inf Sci (NY) 420 (Supplement C) (2017) 386–402.
[8] H.-T. Hu, J.-R. Chang, L.-Y. Hsu, Robust blind image watermarking by modulating the mean of partly sign-altered DCT coefficients guided by human visual perception, AEU Int. J. Electron. Commun. 70 (10) (2016) 1374–1381.
[9] T. Huynh-The, O. Banos, S. Lee, Y. Yoon, T. Le-Tien, A novel watermarking scheme for image authentication in social networks, in: Proceedings of the Ninth International Conference on Ubiquitous Information Management and Communication IMCOM'15, ACM, New York, NY, USA, 2015, pp. 48:1–48:8.
[10] T. Huynh-The, O. Banos, S. Lee, Y. Yoon, T. Le-Tien, Improving digital image watermarking by means of optimal channel selection, Expert Syst. Appl. 62 (2016) 177–189.
[11] L. Jianzhuang, L. Wenqing, T. Yupeng, Automatic thresholding of gray-level pictures using two-dimension Otsu method, in: Proceedings of the 1991 International Conference on Circuits and Systems, China, 1991, pp. 325–327 vol.1, doi:10.1109/CICCAS.1991.184351.
[12] A. Khan, A. Siddiqa, S. Munib, S.A. Malik, A recent survey of reversible watermarking techniques, Inf Sci (NY) 279 (2014) 251–272.
[13] S. Kumar, N. Jain, S.L. Fernandes, Rough set based effective technique of image watermarking, J. Comput. Sci. 19 (2017) 121–137.
[14] H. Liu, D. Xiao, R. Zhang, Y. Zhang, S. Bai, Robust and hierarchical watermarking of encrypted images based on compressive sensing, Signal Process. Image Commun. 45 (2016) 41–51.
[15] M.E. Moghaddam, N. Nemati, A robust color image watermarking technique using modified imperialist competitive algorithm, Forensic Sci. Int. 233 (1–3) (2013) 193–200.
[16] T.-S. Nguyen, C.-C. Chang, X.-Q. Yang, A reversible image authentication scheme based on fragile watermarking in discrete wavelet transform domain, AEU Int. J. Electron. Commun. 70 (8) (2016) 1055–1061.
[17] N. Otsu, A threshold selection method from gray-level histograms, IEEE Trans. Syst. Man Cybern. 9 (1) (1979) 62–66, doi:10.1109/TSMC.1979.4310076.
[18] S.A. Parah, J.A. Sheikh, N.A. Loan, G.M. Bhat, Robust and blind watermarking technique in DCT domain using inter-block coefficient differencing, Digit Signal Process 53 (2016) 11–24.
[19] H. Peng, J. Wang, W. Wang, Image watermarking method in multiwavelet domain based on support vector machines, J. Syst. Softw. 83 (8) (2010) 1470–1477. Performance Evaluation and Optimization of Ubiquitous Computing and Networked Systems.

[20] C.-P. Wang, X.-Y. Wang, Z.-Q. Xia, C. Zhang, X.-J. Chen, Geometrically resilient color image zero-watermarking algorithm based on quaternion exponent moments, J. Vis. Commun. Image Represent 41 (2016) 247–259.
[21] A. Rajpal, A. Mishra, R. Bala, Multiple scaling factors based semi-blind watermarking of grayscale images using OS-ELM neural network, in: Proceedings of the 2016 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), 2016, pp. 1–6.
[22] S. Ranjbar, F. Zargari, M. Ghanbari, A highly robust two-stage contourlet-based digital image watermarking method, Signal Process. Image Commun. 28 (10) (2013) 1526–1536.
[23] S. Roy, A.K. Pal, A blind DCT based color watermarking algorithm for embedding multiple watermarks, AEU Int. J. Electron. Commun. 72 (2017) 149–161.
[24] R.-S. Run, S.-J. Horng, W.-H. Lin, T.-W. Kao, P. Fan, M.K. Khan, An efficient wavelet-tree-based watermarking method, Expert Syst. Appl. 38 (12) (2011) 14357–14366.
[25] A. Saboori, S.A. Hosseini, A novel non-blind watermarking scheme for color image using pca transform and histogram matching technique, in: Proceedings of the 2016 Tenth International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP), 2016, pp. 1–5.
[26] Z. Shao, Y. Duan, G. Coatrieux, J. Wu, J. Meng, H. Shu, Combining double random phase encoding for color image watermarking in quaternion gyrator domain, Opt. Commun. 343 (2015) 56–65.
[27] C. Shi, C.-M. Pun, 3D multi-resolution wavelet convolutional neural networks for hyperspectral image classification, Inf. Sci. (NY) 420 (Supplement C) (2017) 49–65.
[28] P.-C. Su, T.-Y. Kuo, M.-H. Li, A practical design of digital watermarking for video streaming services, J. Vis. Commun. Image Represent 42 (2017) 161–172.
[29] H.-H. Tsai, C.-C. Liu, Wavelet-based image watermarking with visibility range estimation based on {HVS} and neural networks, Pattern Recogn. 44 (4) (2011) 751–763.
[30] H.-H. Tsai, D.-W. Sun, Color image watermark extraction based on support vector machines, Inf Sci (NY) 177 (2) (2007) 550–569.
[31] E. Tsougenis, G. Papakostas, D. Koulouriotis, E. Karakasis, Adaptive color image watermarking by the use of quaternion image moments, Expert Syst. Appl. 41 (14) (2014) 6408–6418.
[32] USC-SIPI-Database, Image database (1977). [online]. available: http://sipi.usc.edu/database/.
[33] C. Wang, J. Ni, J. Huang, An informed watermarking scheme using hidden markov model in the wavelet domain, IEEE Trans. Inf. Forensics Secur. 7 (3) (2012) 853–867.
[34] C. Wang, X. Wang, C. Zhang, Z. Xia, Geometric correction based color image watermarking using fuzzy least squares support vector machine and bessel k form distribution, Signal Process. 134 (2017) 197–208.
[35] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.
[36] R. Xie, C. Hong, S. Zhu, D. Tao, Anti-counterfeiting digital watermarking algorithm for printed QR barcode, Neurocomputing 167 (Supplement C) (2015) 625–635.
[37] X.-Y. Wang, Y.-N. Liu, S. Li, H.-Y. Yang, P.-P. Niu, Robust image watermarking approach using polar harmonic transforms based geometric correction, Neurocomputing 174 (2016) 627–642. Part B
[38] X.-Y. Wang, Y.-N. Liu, H. Xu, A.-L. Wang, H.-Y. Yang, Blind optimum detector for robust image watermarking in nonsubsampled Shearlet domain, Inf Sci (NY) 372 (2016) 634–654.
[39] X.-Y. Wang, C.-P. Wang, H.-Y. Yang, P.-P. Niu, A robust blind color image watermarking in quaternion Fourier transform domain, J. Syst. Softw. 86 (2) (2013) 255–277.
[40] J. Yu, B. Zhang, Z. Kuang, D. Lin, J. Fan, Iprivacy: image privacy protection by identifying sensitive objects via deep multi-task learning, IEEE Trans. Inf. Forensics Secur. 12 (5) (2017) 1005–1016.
[41] J. Zhang, J. Hu, Image segmentation based on 2d Otsu method with histogram analysis, in: Proceedings of the 2008 International Conference on Computer Science and Software Engineering, 6, 2008, pp. 105–108.