Optik - International Journal for Light and Electron Optics 186 (2019) 379–389

Contents lists available at ScienceDirect

Optik

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

Original research article

An image watermarking method based on visual saliency and contourlet transform

Yifeng Zhang[a,b,c], Yibo Sun[a,*]

[a] School of Information Science and Engineering, Southeast University, Nanjing, 210096, China
[b] Nanjing Institute of Communications Technologies, Nanjing, 211100, China
[c] State Key Lab. for Novel Software Technology, Nanjing University, Nanjing, 210093, China

ARTICLE INFO

Keywords:
Image watermarking
Visual saliency
Contourlet transform
Modified logarithmic quantization index modulation
Robustness
Imperceptibility

ABSTRACT

For the purpose of enhancing the imperceptibility and the robustness of watermark, a new image watermarking method based on visual saliency and contourlet transform is proposed. In this algorithm, salient object detection and contourlet transform are operated separately on host image. After dividing the transformed low-pass sub-band into nonoverlapping blocks, the saliency values and the energy distribution of the block determine the quantization step-size of each block together. The logarithmic quantization index modulation is modified so as to gain the robustness against the amplitude scaling attack. Then, in each low-pass sub-band block, the maximum singular value is chosen to embed watermark bits using the modified logarithmic quantization index modulation (MLQIM). The watermarked image is obtained after carrying out the inverse contourlet transform. The results of experiments demonstrate that the presented algorithm possesses satisfactory imperceptibility and is superior to other common algorithms in terms of the robustness against various attacks.

1. Introduction

Digital watermarking technology is used for protecting copyright, verifying the integrity, and proving reliability by embedding extra information into multimedia data [1,2]. Generally, robustness, imperceptibility, and capacity are three vital characteristics of digital watermark. They are opposite to each other and a satisfactory watermarking method should balance them well. Chen et al. [3] put forward the quantization index modulation (QIM) method to acquire good tradeoffs among the three characteristics, which has been broadly used and become one of the most important blind watermarking algorithms. QIM algorithm selects quantizer according to the watermarking information and can resist the Additive White Gaussian Noise (AWGN) attack well. The logarithmic quantization index modulation (LQIM) algorithm, which was presented in [4], applies a μ-Law logarithmic transformation to quantization watermarking algorithm and is an improvement of the QIM. The amplitude scaling attack is a common watermarking attack while both the QIM algorithm and the LQIM algorithm are sensitive to it, causing the bit-error rate (BER) of watermark increasing rapidly. In order to gain the robustness against the amplitude scaling attack, Ourique et al. [5] presented the angle quantization index modulation (AQIM) algorithm, in which the watermark information is embedded into the host signal vector's angle. Nevertheless, the AQIM algorithm shows insufficient performance in resisting the AWGN attack. Some modified AQIM algorithms are subsequently proposed in [6–8] for the sake of the robustness, while they still have some drawbacks, including high distortion or low capacity. For the purpose of enhancing the imperceptibility and the robustness of watermark, meanwhile, achieving the robustness against the

* Corresponding author at: Jiulong Lake Campus, Southeast University, Nanjing, 211189, Jiangsu Province, China.
E-mail addresses: yfz@seu.edu.cn (Y. Zhang), sunyb@seu.edu.cn (Y. Sun).

https://doi.org/10.1016/j.ijleo.2019.04.091
Received 15 February 2019; Received in revised form 18 April 2019; Accepted 19 April 2019
0030-4026/ © 2019 Elsevier GmbH. All rights reserved.

amplitude scaling attack, we propose a novel image watermarking method based on visual saliency and contourlet transform in this paper.

Visual saliency detection is widely used to localize the most significant and attractive regions to human beings in images. Visual saliency is a comprehensive attribute, which is determined by many factors, such as color, luminance, gradient, and boundaries [9]. It has been applied to digital watermarking [10–13] in order to achieve high transparency level. In [10], Bhowmik et al. employed all details coefficients on all wavelet scales in generating a visual attention model. Then the wavelet coefficients were embedded with watermark bits and the intensity of watermarking is consistent with the saliency of regions. Niu et al. [11] adjusted the just noticeable distortion (JND) by introducing visual attention to optimize image watermarking. The study in [12] chose the least salient pixels of the carrier image to implement watermark embedding for less visual distortion. Agarwal et al. [13] embedded watermark bits in prominent regions of the carrier image, which makes the method robust against various attacks.

The contourlet transform [14] is a directional multiscale transform, which solves the problem that the wavelet transform can not represent some types of singularities very well and has been applied extensively in all kinds of image processing algorithms. Meanwhile, image watermarking method based on contourlet transform has also become a new research direction [15–20]. The authors of [15] and [16] considered the low-pass sub-band coefficients in the contourlet transform domain is a good choice to embed watermark information. Ghannam et al. [17] embedded the watermark into different bands of contourlet transform to resist different kinds of attacks. In [18], watermark bits were multiplicatively embedded in the most energetic band-pass directional sub-band. Fazlali et al. [19] introduced an adaptive blind watermarking algorithm in multiple domains, that is embedding watermark bits redundantly in the DCT coefficients of directional sub-bands. Through making full use of the alpha-stable distributions, the study of [20] designed a blind watermarking scheme in the contourlet domain and achieved good performances.

Therefore, in this paper, we introduce visual saliency and contourlet transform to image quantization watermarking method. The logarithmic quantization index modulation is firstly modified so as to gain the robustness against the amplitude scaling attack. Salient object detection is operated on host image to obtain the corresponding saliency map. Then the contourlet transform is carried out on host image and the transformed low-pass sub-band is divided into nonoverlapping blocks. The saliency values and the energy distribution of the block determine the quantization step-size of each block together. After that, the modified logarithmic quantization index modulation is employed to embed the watermark information. Experimental results demonstrate that the presented algorithm possesses satisfactory imperceptibility and outperforms common watermarking algorithms in terms of the robustness.

We are the first to combine visual saliency and contourlet transform with the logarithmic quantization index modulation algorithm. Meanwhile, we solve the problem that the conventional logarithmic quantization index modulation algorithm can not resist the amplitude scaling attack. The imperceptibility of watermark is improved by using adaptive step-size and the robustness of watermark is enhanced through embedding information into the maximum singular value of low-pass sub-band block. Consequently, the proposed method can achieve a great tradeoff between the imperceptibility and the robustness.

The remaining part of this paper is arranged as shown below. A concise review of the contourlet transform and visual saliency detection is presented in Section 2. Section 3 provides the details of the proposed watermarking algorithm. Section 4 demonstrates the results of experiments and Section 5 draws a conclusion.

2. Preliminaries

2.1. The contourlet transform

The basic thought of contourlet transform is extracting directional information at multi-scale resolution. The contourlet transform can be regarded as a real sense of two-dimensional representation of images and can achieve decomposition in any direction at any scale. The contourlet transform possesses many excellent characteristics, such as multiresolution, directionality, localization, and anisotropy [12,14], and solves the problem of insufficient expression of wavelet transform in directional information.

A double filter bank structure is employed in the contourlet transform, including the Laplacian pyramid (LP) and directional filter bank (DFB), to get a sparse representation of images' smooth contour. The point discontinuities are firstly captured using the Laplacian pyramid and then are linked into linear structures by the directional filter bank. To realize the not shift invariant contourlet

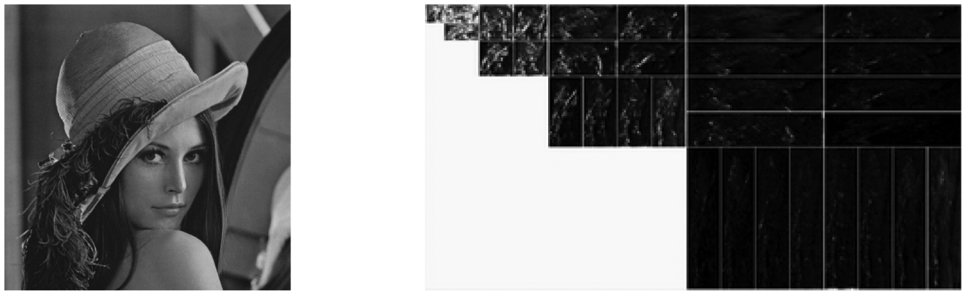

Fig. 1. Lena image(a) and sketch of the nonsubsampled contourlet transform on the Lena image(b).

transform, Cunha, Zhou and Do [21] proposed the nonsubsampled contourlet transform (NSCT), which has the advantages of easy design and implementation. An instance of the nonsubsampled contourlet transform is given in Fig. 1. The four-levels contourlet transform is carried out on the Lena grayscale image, taking two, four, eight, and sixteen directions, respectively. As Fig. 1(b) shows, the color white represents large coefficients while the color black represents small coefficients. The nonsubsampled contourlet transform decomposes the original image into multi-scale and multi-directional sub-bands.

2.2. Visual saliency detection

Visual saliency detection simulates human visual attention mechanism and localizes the most eye-catching regions in images. Prediction of visual saliency has become a significant goal in computer vision, which can be used as a pre-processing step. Some scalar values are consisted of the extracted saliency map, which represent the saliency at various locations [22]. Saliency value reflects the ability to attract human visual attention and higher saliency value indicates that the pixel is more noticeable to viewers.

Several bottom-up computation models have been presented to predict visual saliency using low-level hand-crafted features, including color contrast [9,23], boundary background [24], and center prior [25]. Several models employs high-level features, including contexts [26] and orientation contrast [27]. In recent years, convolutional neural network (CNN) has been widely used in visual saliency detection [28–33] and performed favorably against the conventional methods. The bi-directional message passing model proposed in [33] can be used to detect saliency commendably. Firstly, it designs the multi-scale context-aware feature extraction module to extract multi-scale contextual information. Then the gated bi-directional message passing module is introduced to integrate multi-level features. Making use of these incorporated multi-level features, saliency maps with different resolutions are generated. Finally, these forecasting results are fused to acquire the final saliency map. This model produces satisfactory results and outperforms other methods under different evaluation metrics.

To demonstrate the performance of the bi-directional message passing model in [33], we compare it with other common methods, including two conventional methods (FT [23] and CA [26]) and two deep learning based methods (RFCN [30] and Amulet [32]). Some saliency maps produced by these methods are shown in Fig. 2. It can be seen that deep learning based methods are superior to the conventional ones. Moreover, the saliency map obtained by the method in [33] has distinct outlines and is most similar to the Ground Truth (GT). Therefore, we employ the bi-directional message passing model in [33] to detect visual saliency.

3. The proposed method

3.1. Modified logarithmic quantization index modulation

Logarithmic quantization index modulation proposed in [4] introduces the μ-Law logarithmic transformation to quantization to achieve perceptual advantages.

The host signal x is transformed through the compression function as follows

$$c = \frac{\ln(1 + \mu\frac{|x|}{X_s})}{\ln(1 + \mu)}, \quad \mu > 0, X_s > 0 \tag{1}$$

where μ represents the level of compression and X_s represents the parameter scaling the host signal. The optimal X_s makes most of the host signals convert into the range [0,1]. After that, the transformed signal c is applied uniform quantization regarding the watermark bit. The quantized signal is converted back to the original domain following the inverse function

Fig. 2. Visual comparisons of saliency maps.

$$x_w = \text{sign}(x)\frac{X_s}{\mu}[(1+\mu)^z - 1] \tag{2}$$

where sign(·) represents the sign function, x_w and z represent the watermarked signal in the original domain and transformed domain, respectively.

Despite better performance of robustness compared with conventional QIM, LQIM is still extremely sensitive to amplitude scaling attack. Aiming at tackling this problem, we put forward a modified logarithmic quantization index modulation(MLQIM) algorithm.

In the proposed algorithm, a block-based embedding strategy is adopted and the X_s can be replaced with twice the mean value of each block's host signal. This is because after the watermarked image subjected to the amplitude scaling attack, the mean value of watermarked signals and the watermarked signal of each block are multiplied by the same scale factor. Then the scale factor can be divided out through division and keep the value in the transformed domain unchanged. Twice the mean value of host signals is a suitable choice which makes most of the host signals samples into the range [0,1]. The new compression and inverse function can be given by

$$c = \frac{\ln(1+\mu\frac{|x|}{2\overline{x}})}{\ln(1+\mu)}, \quad \mu > 0 \tag{3}$$

$$x_w = \text{sign}(x)\frac{2\overline{x}}{\mu}[(1+\mu)^z - 1] \tag{4}$$

where $\overline{x}$ represents the mean value of each block's host signal.

3.2. Selection of quantization step-size

In quantization index modulation, quantization step-size Δ is regarded as an important parameter. Large quantization step-size improves the robustness of watermark while infuse distortions and reduces the imperceptibility. Small quantization step-size yields high imperceptibility but poor robustness. For the purpose of achieving a satisfactory tradeoff between the imperceptibility and the robustness, we propose a novel algorithm, in which the quantization step-size of each block is determined by the saliency values and the energy distribution of the block so that the watermark is embedded adaptively.

In this paper, we detect visual saliency and get the saliency map of host image following [33]. The saliency map is consisted of saliency values in the range of [0,255] where salient parts have high values. Divide the saliency map into small nonoverlapping $B \times B$ blocks in accordance with the number of watermark bits to be embedded. In each block, the mean saliency value $\overline{S}$ is chosen to represent the degree of saliency. The larger $\overline{S}$ indicates the image block is more attractive to visual attention, therefore the quantization step-size of the image block should be smaller considering the imperceptibility of watermark. The mean saliency value of each block $\overline{S}$ and block's quantization step-size Δ_S have negative correlation as follows

$$\Delta_S = k_1 \times \frac{1}{\overline{S} + \delta} + b_1 \tag{5}$$

where k_1 is weight coefficient, b_1 is basic quantization step-size, and a constant $\delta = 1$ is added to avoid $\overline{S} = 0$.

In the following, we express the influence of image block's energy distribution on quantization step-size. The higher the energy of image block represents the stronger visual masking ability, hence a larger quantization step-size can be adopted. In our proposal, the host image carries out the contourlet transform and the transformed low-pass sub-band $\mathbf{I}_J$ is divided into nonoverlapping blocks. In each block, the mean value of square of the low-pass sub-band coefficients is calculated to represent the energy of image block E. The energy of each block E and block's quantization step-size Δ_E have positive correlation as follows

$$\Delta_E = k_2 \log_2 E + b_2 \tag{6}$$

where k_2 is weight coefficient, b_2 is basic quantization step-size, and logarithmic operation transforms the exponential growth of frequency coefficient to linear growth, which is consistent with human visual characteristics [34].

In our proposal, the block number of saliency map is assumed to be consistent with the block number of low-pass sub-band. Considering both the saliency and the energy distribution, the quantization step-size of i-th block Δ_i can be obtained by

$$\Delta_i = \Delta_{S_i} + \Delta_{E_i} = k_1 \times \frac{1}{\overline{S_i} + \delta} + k_2 \log_2 E_i + b \tag{7}$$

where $\overline{S_i}$ represents the mean saliency value of i-th block, E_i represents the energy of i-th block, k_1 and k_2 are proportional coefficients, and b is basic quantization step-size.

3.3. Proposed watermark embedding method

Fig. 3 gives the diagram of the watermark embedding process of proposed algorithm. The implementation details of the method are given as follows.

Step1: Operate salient object detection on host image $\mathbf{X}$ to obtain the corresponding saliency map $\mathbf{S}$. Then divide $\mathbf{S}$ into small nonoverlapping $B \times B$ blocks and calculate the mean saliency value of each block $\overline{S}$.

Step2: The nonsubsampled contourlet transform is carried out on host image $\mathbf{X}$. Thus low-pass sub-band $\mathbf{I}_J$ and the band-pass

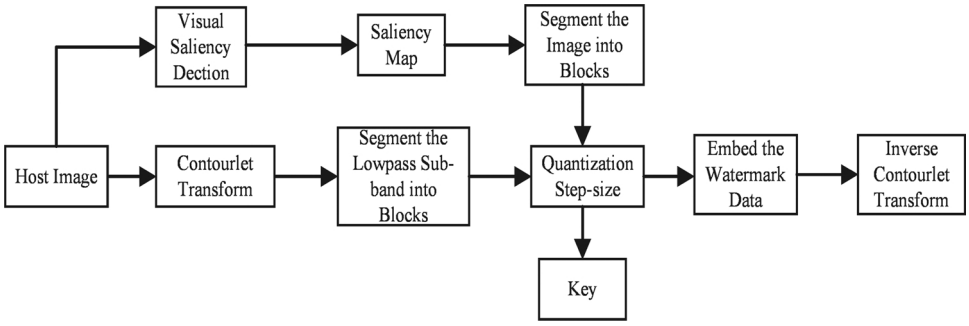

Fig. 3. Proposed watermarking scheme-embedding process.

directional sub-bands $d_{k,p}^{l_k}$ $(p = 0, 1, \cdots, 2^{l_k-1}, k = 1, 2, \cdots, K)$ are obtained, where k denotes the k-th level of LP decomposition, p denotes p-th band-pass directional sub-band decomposed by a l_k-th level DFB.

Step3: Divide low-pass sub-band $\mathbf{I}_J$ into M nonoverlapping $B' \times B'$ blocks and represent them as $\mathbf{L}_1$, $\mathbf{L}_2$, $\cdots\mathbf{L}_M$. Then calculate the energy of each block E.

Step4: Calculate the quantization step-size of i-th block Δ_i according to Eq. (7) and send it to the receiver as a key.

Step5: Perform singular value decomposition on the i-th block of low-pass sub-band following

$$\mathbf{L}_i = \mathbf{U}_i \sum_i \mathbf{V}_i^T \tag{8}$$

where $\sum_i$ denotes a diagonal matrix whose elements on the diagonal are singular values of $\mathbf{L}_i$ and represent the maximum singular value as σ_i.

Step6: According to MLQIM, calculate the mean value of each block's σ_i and denote it as $\overline{\sigma}$. Embed watermark bit into i-th block of low-pass sub-band as follows

$$c = \frac{\ln(1 + \mu\frac{|\sigma_i|}{2\overline{\sigma}})}{\ln(1 + \mu)} \tag{9}$$

$$c^q = Q_{b_i}(c, \Delta_i) = \left\lfloor \frac{c + b_i\Delta_i}{2\Delta_i} \right\rfloor \times 2\Delta_i + b_i\Delta_i \tag{10}$$

$$\sigma_i^q = \text{sign}(\sigma_i)\frac{2\overline{\sigma}}{\mu}[(1 + \mu)^{c_q} - 1] \tag{11}$$

where μ represents the level of compression, $b_i \in \{0, 1\}$ represents the information bit and σ_i^q represents the watermarked singular value.

It is important to note that the singular value reflects the intrinsic characteristics of the image. It has good stability and will not change greatly under small perturbation [35]. Therefore, we choose the singular values to embed watermark bits. Eq. (10) selects different quantizers according to watermarking bits. Then, the maximum singular values in the transformed domain c are quantified into different quantization points using the selected quantizer.

Step7: Replace the maximum singular value σ_i of $\sum_i$ with the watermarked singular value σ_i^q and obtain the watermarked diagonal matrix $\sum_i^q$. The i-th block of watermarked low-pass sub-band $\mathbf{L}_i^q$ can be obtained according to Eq. (8).

Step8: Combine the blocks of watermarked low-pass sub-band and utilize the inverse contourlet transform to get the watermarked image.

3.4. Proposed watermark decoding method

Fig. 4 gives the diagram of the watermark decoding process of proposed algorithm. The implementation details of the method are

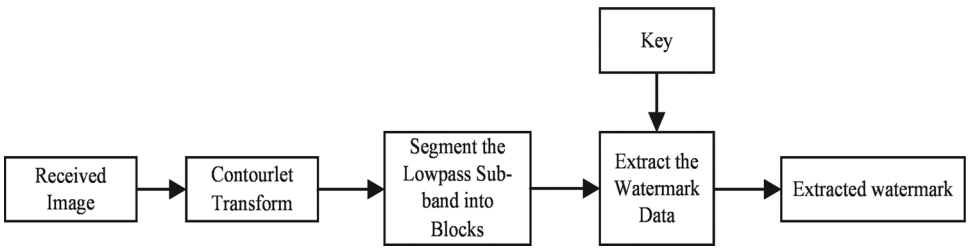

Fig. 4. Proposed watermarking scheme-decoding process.

given as follows.

Step1: The nonsubsampled contourlet transform is carried out on received image $\mathbf{X}'$. Divide low-pass sub-band $\mathbf{I}'_J$ into M non-overlapping $B' \times B'$ blocks and represent them as $\mathbf{L}'_1, \mathbf{L}'_2, \cdots \mathbf{L}'_M$.

Step2: Perform singular value decomposition on the i-th block of low-pass sub-band $\mathbf{L}'_i$ and obtain the maximum singular value as σ'_i.

Step3: Receive the key transmitted by the sender and obtain the quantization step-size of i-th block Δ_i.

Step4: Calculate the mean value of σ'_i and denote it as $\overline{\sigma'}$. Extract the embedded watermark bit in i-th block using the function as shown below

$$c' = \frac{\ln(1 + \mu \frac{|\sigma'_i|}{2\overline{\sigma'}})}{\ln(1 + \mu)} \tag{12}$$

$$b'_i = \arg \min_{b'_i \in \{0,1\}} |c' - Q_{b'_i}(c', \Delta_i)| \tag{13}$$

Eq. (13) employs a minimum distance decoder to extract the watermark bit from the set {0, 1} in a way that minimizes the distance between the extracted maximum singular value in the transformed domain c' and its respective quantization points calculated by Eq. (10)

Step5: Combine the watermark bits to obtain the watermarking image.

4. Experiment results

For the sake of convenience, we denote the presented algorithm as VSCT-MLQIM algorithm. For the purpose of demonstrating the performance of the VSCT-MLQIM algorithm, a series of experiments are conducted. The host images are of size 512 × 512 grayscale images with different texture features, including Lena, Peppers, Baboon and Barbara, as shown in Fig. 5(first row). Watermark information is a pseudorandom binary sequence. Four saliency maps obtained from saliency detection of host images are shown as Fig. 5(second row) and divided into nonoverlapping 16 × 16 blocks. Laplacian pyramid and directional filter bank employed in contourlet transform are all “pkva” filters. The image is applied to three-levels contourlet transform and the size of the low-pass sub-band block is set to 2 × 2.

4.1. Imperceptibility

1024-bit binary watermark is embedded in the four host images applying the presented VSCT-MLQIM algorithm. Each block is

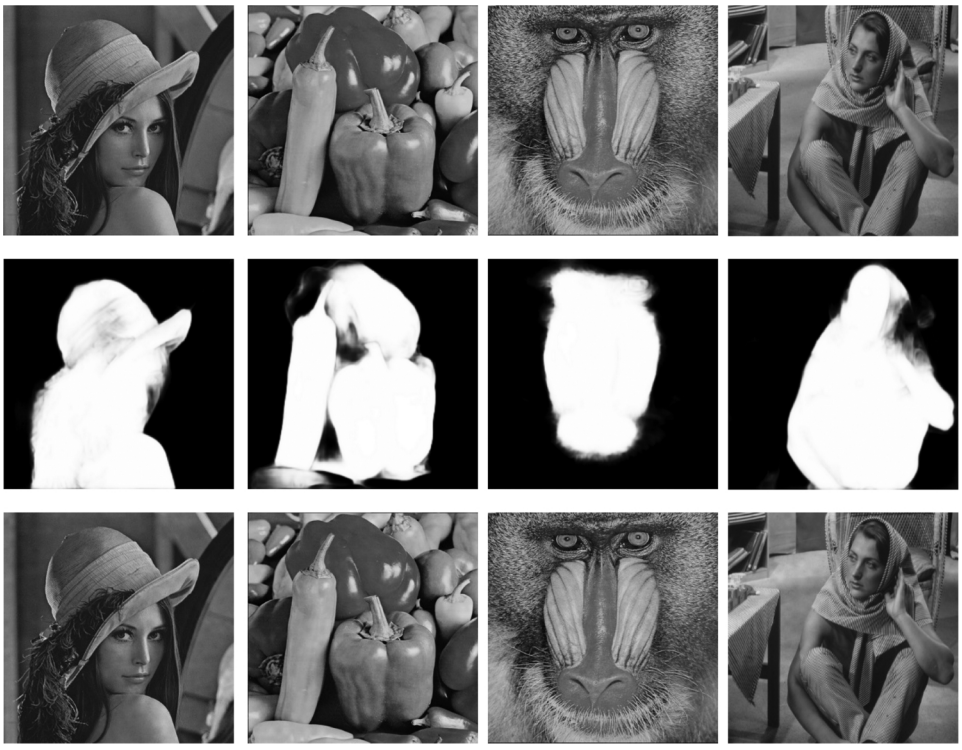

Fig. 5. Host images(first row), corresponding saliency maps(second row), and watermarked images(third row).

Table 1
SSIM indices of watermarked images.

Images	Lena	Peppers	Baboon	Barbara
SSIM	0.9989	0.9939	0.9988	0.9991

embedded with one bit watermark and the proportional coefficients $k_1 = 0.001$, $k_2 = 0.0001$, compression factor $\mu = 6$ are chosen. The basic quantization step-size b is adjusted such that the Peak Signal to Noise Ratio (PSNR) of four watermarked images are all equal to 42 dB. The structural similarity (SSIM) index [36], which is a perceptual metric quantifying the similarity between two images, is chosen to measure watermark embedding distortion. The SSIM index is in the range of [0,1] and higher SSIM index means higher similarity. The SSIM index of two images x and y is computed following:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{14}$$

where μ_x and μ_y represent the mean of the luminance of the host image and the watermarked image, σ_x and σ_y represent the variances of the luminance of the host image and the watermarked image, and σ_{xy} represents the covariance of the above both. The constants C_1 and C_2 are included to avoid instability.

The watermarked images are displayed in Fig. 5(third row) and their SSIM indices are shown in Table 1. The SSIM indices are all close to 1, indicating the watermarked images and the original host images achieve high similarity and little distortion. This is because the proposed VSCT-MLQIM selecting small quantization step-size in salient location is less visually detectable than selecting fixed step-size. Therefore, the watermarked images have satisfactory perceptual quality and it is difficult to distinguish whether images are watermarked or not. The proposed VSCT-MLQIM algorithm has good imperceptibility.

In this subsection, the visual saliency-induced index (VSI) [37] is also used to evaluate the imperceptibility of the VSCT-MLQIM algorithm. The VSI is an effective image quality assessment (IQA) index that makes full use of the visual saliency map. Same as the SSIM index, the VSI is also in the range of [0,1] and higher VSI means higher similarity. Different 1024-bit pseudorandom binary watermarks are embedded into the 4 test images, using the VSCT-MLQIM algorithm and the LQIM algorithm [4], respectively. We set the PSNR a constant as 38 dB. Table 2 shows the visual saliency-induced indices of watermarked images. It is obvious that the VSI of watermarked image using VSCT-MLQIM algorithm is larger than that using LQIM algorithm, which further validates the great imperceptibility of the proposed algorithm.

4.2. Effectiveness of visual saliency and contourlet transform

To evaluate the effectiveness of visual saliency and contourlet transform in VSCT-MLQIM algorithm, several contrastive experiments are conducted.

Algorithm A denotes the method that adopts MLQIM with fixed step-size to implement watermark embedding in the contourlet transform domain;

Algorithm B denotes the method that adopts MLQIM with adaptive step-size determined by saliency values and energy distribution to implement watermark embedding in the wavelet transform domain;

Algorithm C denotes the method that adopts MLQIM with fixed step-size to implement watermark embedding in the wavelet transform domain;

Algorithm D denotes the method that adopts conventional LQIM with fixed step-size to implement watermark embedding in the contourlet transform domain.

The basic quantization step-size b is adjusted such that the PSNR of four watermarked images using different algorithms are all equal to 42 dB. Fig. 6(a) illustrates the BER results of *Algorithm A*, *Algorithm B*, *Algorithm C*, *Algorithm D*, and proposed *VSCT-MLQIM algorithm* under AWGN attack with different standard deviations. The horizontal coordinate-axis is the standard deviation of noise and the vertical coordinate-axis is the BER of extracted watermark. Fig. 6(b) illustrates the BER results of *Algorithm A*, *Algorithm B*, *Algorithm C*, *Algorithm D*, and proposed *VSCT-MLQIM algorithm* after compression of JPEG with different quality factors (QF). The horizontal coordinate-axis is the value of QF and the vertical coordinate-axis is the BER of extracted watermark. It is obvious that the BERs of *Algorithm C* and *Algorithm D* are similar, which indicates that the proposed MLQIM does not weaken the robustness of conventional LQIM under AWGN attack and JPEG compression. It can be seen that the proposed *VSCT-MLQIM algorithm* acquires the

Table 2
Comparison between VSCT-MLQIM and LQIM based on the visual saliency-induced indices of watermarked images.

Methods	Image			
	Lena	Peppers	Baboon	Barbara
LQIM	0.9956	0.9956	0.9970	0.9962
VSCT-MLQIM	**0.9975**	**0.9978**	**0.9983**	**0.9979**

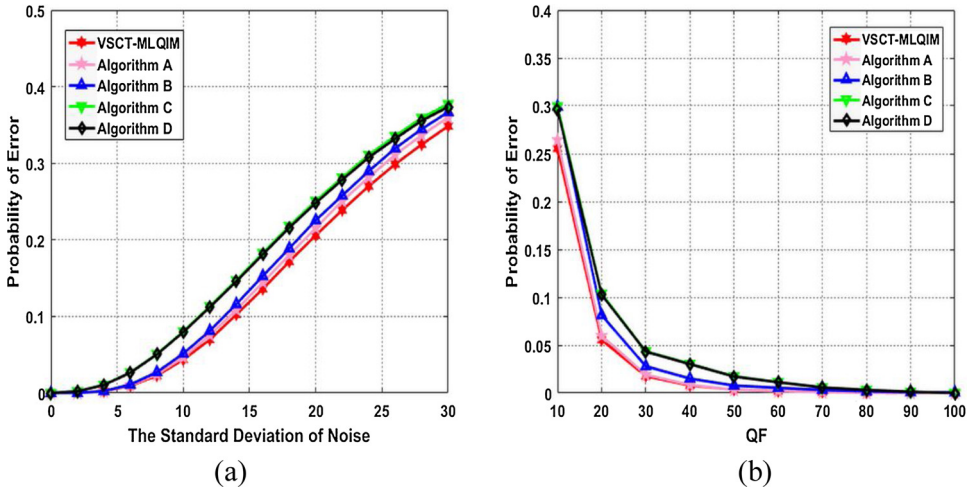

Fig. 6. Experimental comparisons of AWGN attack (a) and JPEG compression (b).

lowest BER under the two attacks, followed by *Algorithm A* and *Algorithm B*. The robustness of *Algorithm C* and *Algorithm D* is the worst, therefore, both the applications of visual saliency and contourlet transform can reduce the BER of extracted watermark and improve the robustness to AWGN attack and JPEG compression.

4.3. Robustness

The proposed VSCT-MLQIM algorithm replaces constant X_S in LQIM with twice the mean value of maximum singular value of low-pass sub-band block $\overline{\sigma}$, making the algorithm gain the robustness against amplitude scaling attack. This is because after the watermarked image subjected to the amplitude scaling attack, $\overline{\sigma}$ and σ_i are multiplied by the same scale factor, which can be divided out through division. Aiming at evaluating and demonstrating the robustness of the presented algorithm to different attacks, we make a comparison between the proposed VSCT-MLQIM algorithm and the GDWM [7], DAQIM [8] algorithms, which are also robust to amplitude scaling attack. In order to achieve fair comparisons, different 256-bit pseudorandom binary watermarks are embedded into the 4 grayscale test images while at the same time satisfying PSNR = 42 dB. VSCT-MLQIM algorithm selects 256 image blocks randomly and guarantees each block is embedded with one bit watermark. Table 3 displays the BER results of the VSCT-MLQIM algorithm and GDWM, DAQIM algorithms under different attacks for four test images, including AWGN attack (standard deviation = 10 or 20), salt and pepper noise (1%), amplitude scaling attack (scale factor = 2), rotational distortion (0.5°), JPEG compression (QF = 20 or 30), median filtering (the size of window = 3 × 3 or 5 × 5), and mean filtering (the size of window = 3 × 3 or 5 × 5).

It is easy to see that our presented VSCT-MLQIM algorithm can resist the amplitude scaling attack very well, which proves the validity of the proposed modified logarithmic quantization index modulation. The VSCT-MLQIM algorithm achieves the best performances against AWGN, rotational distortion and JPEG compression attacks, while the performances against mean filtering and salt and pepper noise are inadequate. This is because mean filtering and salt and pepper noise may make $\overline{\sigma'}$ of the received image change greatly, resulting in the deviation of the value of c' in Eq. (13) (12). Through investigating the BER results of the proposed method

Table 3
BER(%) result of the proposed algorithm, GDWM algorithm, and DAQIM algorithm under different attacks.

Images	Methods	AWGN (std)		S&P noise	Amplitude scaling	Rotational distortion	JPEG (QF)		Median filtering		Mean filtering	
		10	20	1%	scale = 2	0.5	20	30	3 × 3	5 × 5	3 × 3	5 × 5
Lena	Proposed	**0.04**	**3.60**	2.42	0.00	**15.78**	**0.19**	**0.00**	0.06	5.21	2.33	5.67
	DAQIM	1.79	12.66	**0.00**	0.00	35.86	1.71	0.39	0.01	**5.14**	**0.18**	**0.77**
	GDWM	1.85	13.52	0.03	0.00	36.37	1.65	0.41	**0.00**	6.10	0.21	1.46
Baboon	Proposed	**0.01**	**1.35**	0.52	0.00	**15.55**	**0.04**	**0.00**	7.38	17.90	2.65	6.17
	DAQIM	0.94	11.62	**0.31**	0.00	36.19	1.39	0.58	**2.79**	**16.35**	**0.13**	1.78
	GDWM	1.28	12.48	0.54	0.00	36.42	1.41	0.62	5.03	18.75	0.21	**1.24**
Peppers	Proposed	**0.22**	**3.33**	2.28	0.00	**24.03**	**0.16**	**0.02**	**0.00**	**1.39**	2.44	4.49
	DAQIM	1.01	11.29	**0.09**	0.00	35.01	1.30	0.19	1.00	5.37	**0.00**	**0.32**
	GDWM	1.32	13.47	0.11	0.00	35.89	1.33	0.18	1.17	6.64	0.11	1.56
Barbara	Proposed	**0.09**	**2.83**	1.86	0.00	**26.01**	**0.09**	**0.00**	**0.47**	5.60	3.62	5.63
	DAQIM	0.93	10.69	**0.09**	0.00	39.49	1.63	0.10	0.89	**3.59**	**0.06**	**0.97**
	GDWM	1.40	12.87	0.14	0.00	39.90	1.69	0.16	1.01	6.49	0.11	1.47

Table 4
BER(%) result of the proposed algorithm and the algorithm in [15] under different attacks.

Methods	S&P noise	Gaussian noise	Mean filtering	Median filtering
	1%	0.5%	5 × 5	5 × 5
[15]	11.7188	15.0391	16.1133	5.3711
Proposed	**6.8562**	**10.8682**	**7.5313**	**2.7471**

under different attacks, it should be noted that the VSCT-MLQIM algorithm is robust to most of attacks due to the stability of singular values. Altogether, the robustness of the proposed VSCT-MLQIM algorithm outperforms that of GDWM algorithm and DAQIM algorithm.

To further demonstrate the robustness of VSCT-MLQIM algorithm, we make several comparisons with other watermarking algorithms in [15],[16], and [19], which also implement watermark embedding in the contourlet transform domain.

Consistent with the algorithm in [15], the test image "Peppers" is embedded with different 1024-bit pseudorandom binary watermarks to make the PSNR = 40.2 dB. The BER results of the proposed algorithm and the algorithm in [15] under salt and pepper noise (1%), gaussian noise (0.5%), median filtering (5 × 5), and mean filtering (5 × 5) are shown in Table 4. It can be seen that the proposed VSCT-MLQIM algorithm has better robustness than the algorithm in [15].

Consistent with the algorithm in [16], test image "Lena" and "Peppers" are embedded with different 2048-bit pseudorandom binary watermarks to make the PSNR equal to 41.87 and 42.28 dB, respectively. In VSCT-MLQIM, each block is embedded with two information bits. The results in Table 5 distinctly shows that the algorithm in [16] is inferior to the proposed VSCT-MLQIM algorithm under amplitude scaling attack, JPEG compression, and gaussian noise.

Consistent with the algorithm in [19], test image "Barbara", "Couple" and "Bridge" are embedded with different 128-bit pseudorandom binary watermarks to make the PSNR equal to 42.89, 44.06, and 41.74 dB, respectively. In VSCT-MLQIM, we select 128 blocks randomly and guarantee each block is embedded with one bit watermark. The normalized correlation (NC) between the original watermark and the extracted watermark is calculated and used to compare the robustness of the algorithms. The value of NC is in the range of [0,1] and higher correlation represents higher robustness to attacks. The NC results of the two algorithms under gaussian filtering with different sigma and window size are shown in Table 6. As Table 6 demonstrates, the proposed VSCT-MLQIM is highly robust to Gaussian filter attack and outperforms the algorithm in [19].

In order to demonstrate the robustness of the proposed algorithm more intuitively, we embed the 32 × 32 "SEU" binary watermark image (as shown in Fig. 7(a)) into the test image "Lena" with the PSNR equal to 40 dB. Fig. 7(b)–(d) show the extracted watermark image under some combinatorial attacks, including JPEG compression (QF = 50) & salt and pepper noise (1%), gaussian filtering (sigma = 1, the size of window = 5 × 5) & gaussian noise (0.2%), and JPEG compression(QF = 50) & median filtering (3 × 3). The BER of the three extracted watermark images are equal to 7.91%, 2.25%, and 1.07%, respectively. The extracted watermark information is relatively accurate, which proves the proposed algorithm can resist some combinatorial attacks well.

5. Conclusion

This paper proposed a new image watermarking method based on visual saliency and contourlet transform. The logarithmic quantization index modulation was modified so that it could achieve the robustness to amplitude scaling attack. The step-size of each image block was determined by the saliency values and the energy distribution so as to improve the imperceptibility. The robustness was enhanced through embedding information into the maximum singular value of low-pass sub-band block. Several experiments are conducted to evaluate the performance of the proposed method. The SSIM index and VSI are used to verify the imperceptibility. Contrastive experiments proved the effectiveness of visual saliency and contourlet transform. The results also demonstrated that the presented method yields better robustness against various attacks compared with the GDWM, DAQIM, and the methods in [15,16,19]. The proposed method achieves a great tradeoff between the imperceptibility and the robustness, which has practical application value.

Table 5
BER(%) result of the proposed algorithm and the algorithm in [16] under different attacks.

Images	Methods	Amplitude scaling 1.5	JPEG QF = 50	Gaussian noise 0.03%
Lena	[16]	21.26	24.26	11.39
	Proposed	**0.00**	**2.48**	**4.01**
Peppers	[16]	21.92	25.00	11.60
	Proposed	**0.00**	**3.31**	**4.62**

Table 6
Normalized Correlation result of the proposed algorithm and the algorithm in [19] under gaussian filtering.

Images	Methods	Gaussian filtering								
		Sigma = 0.5			Sigma = 1			Sigma = 2		
		3 × 3	5 × 5	7 × 7	3 × 3	5 × 5	7 × 7	3 × 3	5 × 5	7 × 7
Barbara	[19]	1	1	1	0.99	0.99	0.99	0.99	0.85	0.63
	Proposed	1	1	1	**1**	**1**	**1**	**1**	**1**	**1**
Couple	[19]	1	1	1	0.98	0.98	0.98	0.98	0.82	0.62
	Proposed	1	1	1	**1**	**1**	**1**	**1**	**1**	**0.99**
Bridge	[19]	1	1	1	0.99	0.99	0.99	0.99	0.92	0.76
	Proposed	1	1	1	**1**	**1**	**1**	**1**	**1**	**0.99**

(a) (b) (c) (d)

Fig. 7. Original watermark image (a), the extracted watermark images under JPEG compression & salt and pepper noise (b), gaussian filtering & gaussian noise (c), and JPEG compression & median filtering (d).

CRediT authorship contribution statement

Yifeng Zhang: Conceptualization, Writing - review & editing. **Yibo Sun:** Validation, Writing - review & editing.

Acknowledgments

This work was supported in part by the Natural Science Foundation of Jiangsu Province under Grant BK20151102, in part by the Ministry of Education Key Laboratory of Machine Perception, Peking University under Grant K-2016-03, in part by the Open Project Program of the Ministry of Education Key Laboratory of Underwater Acoustic Signal Processing, Southeast University under Grant UASP1502, and in part by the Natural Science Foundation of China under Grant 61673108.

References

[1] S. Katzenbeisser, F.A.P. Petitcolas, Information Hiding Techniques for Steganography and Digital Watermarking, Artech House, Boston, 2000.
[2] S.W. Weng, Y. Zhao, J.S. Pan, R.R. Ni, Reversible watermarking based on invariability and adjustment on pixel pairs, IEEE Signal Process. Lett. 15 (2008) 721–724, https://doi.org/10.1109/LSP.2008.2001984.
[3] B. Chen, G.W. Wornell, Quantization index modulation: a class of provably good methods for digital watermarking and information embedding, IEEE Trans. Inf. Theory 47 (4) (2001) 1423–1443, https://doi.org/10.1109/18.923725.
[4] N.K. Kalantari, S.M. Ahadi, A logarithmic quantization index modulation for perceptually better data hiding, IEEE Trans. Image Process. 19 (6) (2010) 1504–1517, https://doi.org/10.1109/TIP.2010.2042646.
[5] F. Ourique, V. Licks, R. Jordan, F. Perez-Gonzalez, Angle QIM: a novel watermark embedding scheme robust against amplitude scaling distortions, 30th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) (2005) 797–800, https://doi.org/10.1109/ICASSP.2005.1415525.
[6] M.A. Akhaee, S.M.E. Sahraeian, C. Jin, Blind image watermarking using a sample projection approach, IEEE Trans. Inf. Forensic Secur. 6 (3) (2011) 883–893, https://doi.org/10.1109/TIFS.2011.2146250.
[7] E. Nezhadarya, Z.J. Wang, R.K. Ward, Robust image watermarking based on multiscale gradient direction quantization, IEEE Trans. Inf. Forensic Secur. 6 (4) (2011) 1200–1213, https://doi.org/10.1109/TIFS.2011.2163627.
[8] N. Cai, N.N. Zhu, S.W. Weng, B.W.K. Ling, Difference angle quantization index modulation scheme for image watermarking, Signal Process. Image Commun. 34 (2015) 52–60, https://doi.org/10.1016/j.image.2015.03.010.
[9] M.M. Cheng, G.X. Zhang, N.J. Mitra, X.L. Huang, S.M. Hu, Global contrast based salient region detection, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2011) 409–416, https://doi.org/10.1109/CVPR.2011.5995344.
[10] D. Bhowmik, M. Oakes, C. Abhayaratne, Visual attention-based image watermarking, IEEE Access 4 (2016) 8002–8018, https://doi.org/10.1109/ACCESS.2016.2627241.
[11] Y.Q. Niu, M. Kyan, L. Ma, A. Beghdadi, S. Krishnan, A visual saliency modulated just noticeable distortion profile for image watermarking, 19th European Signal Processing Conference (EUSIPCO) (2011) 2039–2043.
[12] A. Sur, S.S. Sagar, R. Pal, P. Mitra, J. Mukhopadhyay, A new image watermarking scheme using saliency based visual attention model, 2009 Annual IEEE India Conference (2009), https://doi.org/10.1109/INDCON.2009.5409402.
[13] C. Agarwal, A. Bose, S. Maiti, N. Islam, S.K. Sarkar, Enhanced data hiding method using DWT based on saliency model, 2013 IEEE International Conference on Signal Processing, Computing and Control (ISPCC) (2013), https://doi.org/10.1109/ISPCC.2013.6663414.
[14] M.N. Do, M. Vetterli, The Contourlet transform: an efficient directional multiresolution image representation, IEEE Trans. Image Process. 14 (12) (2005) 760–769, https://doi.org/10.1109/TIP.2005.859376.
[15] S.M. Zhu, J.M. Liu, Adaptive quantization index modulation digital image watermarking algorithm based on Contourlet transform domain, Acta Opt. Sin. 29 (6) (2009) 1523–1529.
[16] Zhou Chao, DC-QIM based image watermarking method via the contourlet transform, J. Comput. Methods Sci. Eng. 16 (3) (2016) 459–468, https://doi.org/10.

3233/JCM-160629.
[17] S. Ghannam, F.E.Z. Abou-Chadi, Enhancing robustness of digital image watermarks using contourlet transform, 16th IEEE International Conference on Image Processing (2009) 3645–3648, https://doi.org/10.1109/ICIP.2009.5414260.
[18] M.A. Akhaee, S.M.E. Sahraeian, F. Marvasti, Contourlet-based image watermarking using optimum detector in noisy environment, IEEE Trans. Image Process. 19 (4) (2010) 967–980, https://doi.org/10.1109/TIP.2009.2038774.
[19] H.R. Fazlali, S. Samavi, N. Karimi, S. Shirani, Adaptive blind image watermarking using edge pixel concentration, Multimed. Tools Appl. 76 (2) (2017) 3105–3120, https://doi.org/10.1007/s11042-015-3200-6.
[20] H. Sadreazami, M.O. Ahmad, M.N.S. Swamy, A study of multiplicative watermark detection in the contourlet domain using alpha-stable distributions, IEEE Trans. Image Process. 23 (10) (2014) 4348–4360, https://doi.org/10.1109/TIP.2014.2339633.
[21] A.L. Cunha, J. Zhou, M.N. Do, The nonsubsampled contourlet transform: theory, design, and applications, IEEE Trans. Image Process. 15 (10) (2006) 3089–3101, https://doi.org/10.1109/TIP.2006.877507.
[22] P.L. Callet, E. Niebur, Visual attention and applications in multimedia technologies, Proceedings of the IEEE Institute of Electrical and Electronics Engineers (2013) 2058–2067, https://doi.org/10.1109/JPROC.2013.2265801.
[23] R. Achanta, S. Hemami, F. Estrada, S. Susstrunk, Frequency-tuned salient region detection, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition(CVPR) (2009) 1597–1604, https://doi.org/10.1109/CVPR.2009.5206596.
[24] C. Yang, L.H. Zhang, H.C. Lu, X. Ruan, M.H. Yang, Saliency detection via graph-based manifold ranking, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition(CVPR) (2013) 3166–3173, https://doi.org/10.1109/CVPR.2013.407.
[25] Z.L. Jiang, L.S. Davis, Submodular salient region detection, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition(CVPR) (2013) 2043–2050, https://doi.org/10.1109/CVPR.2013.266.
[26] S. Goferman, L. Zelnik-Manor, A. Tal, Context-aware saliency detection, IEEE Trans. Pattern Anal. Mach. Intell. 34 (10) (2012) 1915–1926, https://doi.org/10.1109/TPAMI.2011.272.
[27] H.W. Tian, Y.M. Fang, Y. Zhao, W.S. Lin, R.R. Ni, Z.F. Zhu, Salient region detection by fusing bottom-up and top-down features extracted from a single image, IEEE Trans. Image Process. 23 (10) (2014) 4389–4398, https://doi.org/10.1109/TIP.2014.2350914.
[28] L.J. Wang, H.C. Lu, X. Ruan, M.H. Yang, Deep networks for saliency detection via local estimation and global search, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition(CVPR) (2015) 3183–3192, https://doi.org/10.1109/CVPR.2015.7298938.
[29] G.B. Li, Y.Z. Yu, Visual saliency based on multiscale deep features, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition(CVPR) (2015) 5455–5463, https://doi.org/10.1109/CVPR.2015.7299184.
[30] L.Z. Wang, L.J. Wang, H.C. Lu, P.P. Zhang, X. Ruan, Saliency detection with recurrent fully convolutional networks, Proceedings of European Conference on Computer Vision(ECCV) (2016) 825–841, https://doi.org/10.1007/978-3-319-46493-0_50.
[31] L.Z. Wang, H.C. Lu, Y.F. Wang, M.Y. Feng, D. Wang, B.C. Yin, et al., Learning to detect salient objects with image-level supervision, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition(CVPR) (2017) 3796–3805, https://doi.org/10.1109/CVPR.2017.404.
[32] P.P. Zhang, D. Wang, H.C. Lu, H.Y. Wang, X. Ruan, Amulet: aggregating multi-level convolutional features for salient object detection, Proceedings of the IEEE International Conference on Computer Vision(CVPR) (2017) 202–211, https://doi.org/10.1109/ICCV.2017.31.
[33] L. Zhang, J. Dai, H.C. Lu, Y. He, G. Wang, A bi-directional message passing model for salient object detetion, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), (2018), pp. 1741–1750.
[34] X.Y. Wang, H. Zhao, A novel synchronization invariant audio watermarking scheme based on DWT and DCT, IEEE Trans. Signal Process. 54 (12) (2006) 4835–4840, https://doi.org/10.1109/TSP.2006.881258.
[35] H.C. Xu, X.B. Kang, Y.H. Wang, Y.L. Wang, Exploring robust and blind watermarking approach of colour images in DWT-DCT-SVD domain for copyright protection, Int. J. Electron. Secur. Digit Forens 10 (1) (2018) 79–96, https://doi.org/10.1504/IJESDF.2018.10009833.
[36] 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, https://doi.org/10.1109/TIP.2003.819861.
[37] L. Zhang, Y. Shen, H.Y. Li, VSI: a visual saliency-induced index for perceptual image quality assessment, IEEE Trans. Image Process. 23 (10) (2014) 4270–4281, https://doi.org/10.1109/TIP.2014.2346028.

Yifeng Zhang received the B.E. degree in electrical engineering from Southeast University, Nanjing, China, in 1984, the M.E. degree in computer engineering from the Harbin Institute of Technology, Harbin, China, in 1989, and the pH.D. degree in electrical engineering from Southeast University in 1999, respectively. From 1999–2001, he was a Post-Doctoral Fellow with the Department of Radio Engineering, Southeast University. In 2009, he joined Southeast University, where he is currently an Associate Professor with the School of Information Science and Engineering. His current research interests include watermarking and information hiding, facial expression recognition, visual tracking.

Yibo Sun received the Bachelor's degree in the School of Information Science and Engineering from Southeast University, China in 2012. Currently, he is pursuing his Master degree in signal and information process at the School of Information Science and Engineering, Southeast University, China. His research interests include digital watermarking, information hiding, compressive sensing, etc.