

## Research Article

# Blind Photograph Watermarking with Robust Defocus-Based JND Model

Chunxing Wang <sup>1</sup>, Xiaoxiao Li <sup>1</sup>, Meiling Xu <sup>2,3</sup>, Jun Wang <sup>2</sup>, and Wenbo Wan <sup>2</sup>

<sup>1</sup>School of Physics and Electronic Science, Shandong Normal University, Jinan 250014, China

<sup>2</sup>School of Information Science and Engineering, Shandong Normal University, Jinan 250014, China

<sup>3</sup>School of Information Science and Engineering, Shandong Huayu University of Technology, Dezhou 254300, China

Correspondence should be addressed to Wenbo Wan; wanwenbo@sdsu.edu.cn

Received 16 September 2020; Revised 22 October 2020; Accepted 5 November 2020; Published 1 December 2020

Academic Editor: Zhili Zhou

Copyright © 2020 Chunxing Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Just noticeable distortion (JND) is widely employed to describe the perception redundancy in the quantization-based watermarking framework. However, the existing JND models are generally constructed to treat every region of the photograph with an equal focus level, whereas the defocus effect has never been considered. In this paper, the defocus feature, which can portray the aesthetic emphasis in the photograph, is provided to improve the perceptual JND model. Firstly, two indicators which consider the block energy in the defocus measurement (DM) are proposed. Then, the defocus feature map (DFM) is obtained by integrating the influence of the circumambient blocks, and it is applied to the proposed JND contrast masking (CM) processing. In this way, a new blind photograph watermarking method, with emphasis on defocus-JND estimation combined with the proposed CM, is presented. Simulations show that the proposed JND is more suitable for watermarking framework than some existing JND models, and the proposed watermarking scheme with the improved defocus-based JND model has superior robustness compared with some watermarking schemes.

## 1. Introduction

In recent years, there has been a surge of interest in taking photos using a digital camera. The aesthetic effect can be reflected by high dynamic range images. Due to the advanced camera parameter settings, the photos we acquired often subject to different focus effects as shown as in Figure 1. We can see that the right one shows a nice inside focus and outside defocus photograph, which gives a better artistic photography. The difference between photographs and classic images can be briefly summarized in follows: The focus level will be more likely to vary from the object to its background in photographs. This phenomenon is related with the parameter setting of the camera nowadays. This preference will achieve a better aesthetic effect because a photograph that has visually dominant subjects usually induces stronger aesthetic interest. Professional photographers utilize contrast in sharpness and lighting to bring out the visual dominance of sub-

jects so that the viewer is directed to where the photographers intended. These techniques can change the original appearance which may bring some distortions. However, some aesthetic effect is achieved by these processing methods exactly right.

In our daily life, these aesthetic photographs are easily to be used in their own social software without the photograph owner's permission, which will cause a series of copyright issues [1]. This information security is a topic that is hardly out of date. In the past few years, some techniques have made some developments such as block cipher [2], encryption [3], and steganography [4]. Especially with the development of the internet, all kinds of multimedia can be searched easily [5–7]. It can even be achieved by cross-modal retrieval [8, 9]. Information hiding is an effective way to protect the meaningful host carrier [10–12]. Its principle is hiding the secret information into the original multimedia carriers without visual changes [13]. This feature related to the visual



FIGURE 1: Photos subject to different defocus effect: (a) original photographs; (b) photographs with different defocus level.

characteristics of human visual system (HVS) which can be measured by an objective parameter called structural similarity index (SSIM). It is significant to choose a suitable resolution mechanism [14–16]. Watermarking technology can achieve security under the premise of ensuring visual consistency. And exploring the embedding domains [17, 18] and adjusting the embedding ways [19] have made a series of progress in recent years. All these effects are made to improve the watermarking robustness in the premise of keeping an acceptable invisibility. That is to say, the balance between robustness and invisibility becomes a vital topic. And the commonly used just noticeable distortion (JND) model that related closely with HVS can get a better tradeoff between these two characteristics. Ma et al. [20] proposed an adaptive spread-transform dither modulation (STDM) approach by incorporating a new perceptual model into the conventional STDM framework. Wan et al. [21] presented a logarithmic STDM watermarking approach based on an improved JND model. This new JND model is got by revising the edge pixel density. Wang et al. [22] proposed a JND model based on the block types including orientation and regularity, and this model was used in watermarking framework.

In human's visual sense, the JND model can be a combination of positive and negative incentives for perception [23]. The positive effect is produced by the unexpected part that will cause visual attention. And the negative effect is brought by its own characteristics such as spatial frequency, luminance, texture, and surrounding stimuli. However, many exiting JND models only incorporated the negative effects together. In detail, JND models can be estimated in pixel domain or transform domain. In pixel domain, a JND model is often derived by investigating the luminance adaption (LA) and spatial contrast masking (CM) [24]. The JND models in transform domain usually utilize human vision sensitivities of different frequency components, and the contrast sensitivity function (CSF) is frequently applied in identifying the base thresholds [25]. Some JND models are calculated by fusing these multiple

masking factors, which will result in overestimation in the focus regions having strong attraction. How to combine these masking effects into a model linearly or nonlinearly has become a subject worthy of study. In general, defocus measurement is responsible for defining focus areas of a photograph that will attract more attention of the HVS. Distortion occurring in the focus area that attracts the viewer's attention is more annoying than that in other areas. Therefore, defocus measurement can be exploited to weight the JND model with different defocus levels in the watermarking framework. Because the local area in a photograph that is easy to be concerned, this application will satisfy this characteristic. What is more, this application can optimize the combination of masking effects getting a JND model that is more suitable for photograph watermarking framework. Until now, some defocus measurements have been proposed using spatial edge information such as edge width. It is exploited by combining the standard deviation of the edge gradient magnitude profile [26]. Then, as HVS is variable for different contrast levels, the conception of just noticeable blurriness (JNB) based on edge width is integrated into a probability summation model [27]. The above edge-based methods can achieve this goal by the theory that the edge will be affected by defocus in a big way. However, it is difficult to make a defocus judgement by detecting all edges firstly. Owing to the interference to edge detection caused by some common attacks in watermarking, the robustness will be uncertain. Besides edge information, the energy in different frequencies especially high frequency is another research field in defocus measurement, which is based on the theory that the area operated by defocus will lose many detail information. To eliminate the effect of noise, an adaptive weight is given to the coefficients at high frequency and middle frequency [28]. Though there is a natural masking effect in high frequency, the information in this subband is not stable, which cannot resist some common attacks in transmission. Thus, the robustness will not be satisfied when this measurement is applied in watermarking framework. Nowadays,

limited progress has been made in this research area, mainly due to the fact that the defocus measurement calculated in the original photo is inconsistent with that in the watermarked photograph. Therefore, the existing defocus measurements cannot provide an outstanding performance for a practical blind watermarking framework.

To tackle this challenge, in this paper, a novel blind watermarking method for the photograph is proposed, in which the defocus measurement in the discrete cosine transform (DCT) domain is introduced to measure the CM effect in the perceptual JND model. Experiments show that the proposed scheme has enhanced robustness against common attacks. The main contributions of this paper include the following:

- (1) In photograph copyright protection, we propose to apply the defocus measurement to JND model for the first time. In recent years, the unique photography technology divides a photograph into different levels of focus, which facilitate the emergence of aesthetic photograph. The existing JND models cannot satisfy the balance between invisibility and robustness better under this requirement
- (2) We propose a new defocus measurement and apply it into CM. In this measurement, as the influence of defocus differs in blocks with different energies, two types of blocks called strongly and weakly textured blocks are distinguished by some representative alternating current (AC) coefficients in DCT domain. And two coefficients at low and middle frequency are selected as the defocus indicator parameter. Then, for strongly textured blocks, the whole photograph's contrast is used as the threshold to measure these indicators. And for weakly textured blocks, it plays a key role that modulating the energy amplitude spectrum calculated by these two indicators and their corresponding spatial frequencies. Finally, the influence of circumbient blocks is modulated to a Gaussian function. The final results are used in the block types partitioning, and an enhanced CM factor is proposed. The experimental results show that the defocus measurement can take its role effectively
- (3) The proposed robust defocus-based JND model is implemented to estimate an optimum quantization step adaptively for the quantization-based watermarking framework. The experimental results show that this watermarking method performs better than some exiting watermarking methods. And the proposed JND model is more suitable to be used in watermarking framework than some existing JND models

The remainders of this paper are listed as follows. In Section 2, some related work is presented. In Section 3, we present a new JND model improved by the proposed defocus measurement. In Section 4, the proposed JND model is applied into photograph watermarking framework. In Section 5, some experiments are used to demonstrate our method's availability. Section 6 is a summary of this paper.

## 2. Related Work

*2.1. JND Models in Watermarking Technology.* A JND model needs to combine the reasonable masking factors that will affect HVS. The visual attention mechanism should be contained [29]. Watson model [25] is a classical JND model in DCT domain, which has put forward the DCT sensitivity coefficient table for the first time. In this model, the carrier coefficients are modified according to the luminance masking and contrast masking calculated by this DCT sensitivity coefficient table. Kim proposed a model based on DCT frequency and background brightness, where the contrast masking was not considered directly, but the texture complexity was used to construct LM-JND [30]. Psychological model is integrated into this JND. In [31], Wan proposed a new JND based on orientation regularity. By judging the directivity and texture characteristics of the image block and assigning different values to the CM, the proposed JND model was more in line with the visual characteristics of the human eyes. In the aspect of image compression, this JND can be well applied. Considering all these classic masking effect, Ngan has proposed a combined JND model which incorporates CSF, LA, and CM all together in DCT domain [32].

$$T_{\text{jnd}} = T_B(n, i, j) \cdot M_{\text{LA}}(n, i, j) \cdot M_{\text{CM}}(n, i, j), \quad (1)$$

where  $n$  is an index of a block and  $i$  and  $j$  are the DCT coefficients' indices in a block.  $T_B$  is used in depicting the threshold based on spatial CSF,  $M_{\text{LA}}$  is the index of luminance masking, and  $M_{\text{CM}}$  is the influence of CM which is a popular research point that has been developed deeply in recent years [33, 34]. This CM effect is strongly associated with the sensitivity of the HVS for different background.

Besides these JND models, some robust JND models applied in watermarking technology are proposed. In [21], the author used a perceptual JND model in his logarithmic STDM watermarking work. This model improved the CM by a new characteristic called edge pixel density and LA by another feature called average pixel intensity. Wang proposed a structural regularity-based watermark method employing an adaptive JND [22]. In this paper, three AC coefficients and one direct current (DC) coefficient were used to judge the regularity and directivity of image blocks, and then, different masking factors are given to different types of image blocks according to human visual characteristics.

*2.1.1. The CSF-Based Threshold.* CSF is the ability to characterize the spatial resolution and brightness resolution of human eyes. And its influence on the basic threshold can be defined as follows [32].

$$T_B(n, i, j) = \frac{1}{4\phi_i\phi_j} \frac{\exp(0.18\omega_{ij}) / (1.33 + 0.11\omega_{ij})}{0.6 + 0.4(\cos \varphi_{ij})^2}, \quad (2)$$

where  $\varphi_{ij}$  is the directional angle of DCT component; then, spatial frequency  $\omega_{ij}$  and DCT normalization factors  $\phi_{i(j)}$  can be defined by

$$\omega_{ij} = \frac{1}{2 \times 8} \sqrt{\left(\frac{i}{\theta_h}\right)^2 + \left(\frac{j}{\theta_v}\right)^2}, \quad (3)$$

$$\phi_{i(j)} = \begin{cases} \sqrt{\frac{1}{8}} & i(j) = 0, \\ \sqrt{\frac{2}{8}} & i(j) > 0, \end{cases} \quad (4)$$

where  $\theta_h$  and  $\theta_v$  represent the horizontal and vertical angles for each pixel, respectively. And they can be defined by Equation (5).

$$\theta = 2 \cdot \arctan \frac{1}{2 \times R_{vh} \times H}, \quad (5)$$

where  $R_{vh}$  stands for the ratio of viewing distance to screen height and  $H$  is the number of pixels in picture height. According to the international standard, the ratio of viewing distance to screen height should be a fixed number which is usually from 3 to 6 depending on the screen size. Here, its value is 4.5. Moreover, for most of the displays, PAR (pixel aspect ratio) is equal to 1. This means that the horizontal and vertical visual angles ( $\theta_h, \theta_v$ ) are identical.

**2.1.2. Luminance Adaption.** The response of human eyes to blocks with unequal brightness is different. When the brightness of a region is much greater or much less than the average brightness of the photograph, the masking property of the human eye is greater. But it is not difficult for our eyes to identify the normal luminance area in a photograph [32]. These features about human visual system can be described as

$$F_L = \begin{cases} \frac{60 - I_b}{150} + 1 & I_b < 60, \\ 1 & 60 \leq I_b \leq 170, \\ \frac{I_b - 170}{425} + 1 & I_b > 170, \end{cases} \quad (6)$$

where  $I_b$  is the average intensity of a photograph block.

**2.2. Spread Transform Dither Modulation.** STDM is an extension of classical quantization index modulation (QIM) [35]. QIM modulates the original carrier by using two quantizers corresponding to the watermark bits 0 or 1. When extracting the watermark bits, QIM uses the minimum distance detection method, which means that the quantization point closest to the watermarked carrier is used to determine the watermark bit. The original carrier is distorted in this quantization process. To improve the imperception quality, quantization dither modulation (QDM) is proposed by using a pseudorandom signal to reduce the quantization effect. It can be defined as follows.

$$\text{QDM}(\mathbf{x}, \Delta, d_m) = \Delta \cdot \text{round}\left(\frac{\mathbf{x} + d_m}{\Delta}\right) - d_m, m \in \{0, 1\}, \quad (7)$$

where  $\mathbf{x}$  is the original carrier,  $\Delta$  is the quantization step, and  $d_m$  is a dither signal corresponding to the watermarking bit  $m$ . And  $d_1$  is defined as

$$d_1 = d_0 - \text{sign}(d_0) \cdot \frac{\Delta}{2}, \quad (8)$$

where  $d_0$  is a pseudorandom number that is evenly distributed over  $[-\Delta/2, \Delta/2]$ .

STDM uses QDM to modulate the projection of carrier vector  $x$  along a given random direction vector  $u$ . The embedding process can be represented by Equation (9).

$$\mathbf{v} = \mathbf{x} + \frac{\text{QDM}(\text{proj}(\mathbf{x}, \mathbf{u}), \Delta, d_m) - \text{proj}(\mathbf{x}, \mathbf{u})}{\|\mathbf{u}\|_2} \mathbf{u}, \quad (9)$$

where  $\text{proj}(\bullet)$  represents the projection operation which can be defined by Equation (10).

$$\text{proj}(\mathbf{x}, \mathbf{u}) = \frac{\langle \mathbf{x}, \mathbf{u} \rangle}{\|\mathbf{u}\|_2} \mathbf{u}, \quad (10)$$

where  $\langle \mathbf{x}, \mathbf{u} \rangle$  represents the product of two vectors and  $\|\bullet\|_2$  represents the L-2 norm.

In the extraction process, the host carrier  $\mathbf{x}'$  that may be distorted is projected onto the same direction vector  $\mathbf{u}$  to get  $\mathbf{x}'_{\mathbf{u}}$ . Then, the watermarks 0 and 1 are embedded into  $\mathbf{x}'_{\mathbf{u}}$  according to Equation (9), respectively. And two new vectors  $\mathbf{v}'_{\mathbf{u}0}$  and  $\mathbf{v}'_{\mathbf{u}1}$  are obtained. The extracted watermark bit will be determined by the minimum distance between these two new vectors and  $\mathbf{v}'_{\mathbf{u}}$ . It can be represented by Equation (11).

$$m = \text{argmin} \text{dist}\left(\mathbf{x}'_{\mathbf{u}}, \text{QDM}\left(\mathbf{x}'_{\mathbf{u}}, \Delta', d_m\right)\right), m \in \{0, 1\}. \quad (11)$$

### 3. The Proposed Defocus-Based JND

In a photograph taken by professional photography, defocus regions can certainly optimize the quality of proliferating visual services, so the defocus measurement can bring a more realistic virtual environment as well as an enhanced visual experience, and it can avoid overestimation or underestimation in the redefinition of CM. With high defocus strength, the photograph has a strong contrast masking effect and high visibility threshold. In this paper, a defocus-based CM is proposed, and it is applied in a new JND. The details are discussed in the following.

**3.1. Defocus Measurement Based on Block Types.** In order to achieve the defocus measurement, the DCT is first applied. In DCT domain, a  $8 \times 8$  DCT block consists of one DC coefficient and 63 AC coefficients. The DC coefficient related closely with luminance that can reflect the block's average energy while the AC coefficients can represent the focus property [28]. However, the robustness of the watermarking method based on the proposed model should be ensured firstly. It means that the watermark extracted from the distorted watermarked photograph suffering from some common attacks should be acceptable. At the same time, the selected coefficients should make full expression of the characteristics in the defocus measurement. For saving computing cost, the study has proved that the coefficients centralized in the top-left corner of each subband can be a

representation of this subband [28]. Moreover, the signal energy usually concentrates on lower frequency bands. Therefore, to ensure the robustness and accuracy of the defocus feature, the representative AC coefficients  $AC_{0,1}$ ,  $AC_{1,0}$ ,  $AC_{1,1}$ ,  $AC_{0,2}$ ,  $AC_{2,0}$ , and  $AC_{2,2}$  (the subscripts refer to the row and column indices of the block) with local contrast features in  $n$ -th block are applied, and  $T_e$  is defined as Equation (12). The coefficients of a typical block in Equation (12) are presented in Figure 2. In this block, the coefficients with red are used in the defocus calculation.

$$T_e = \{AC_{0,1}, AC_{1,0}, AC_{1,1}, AC_{0,2}, AC_{2,0}, AC_{2,2}\}. \quad (12)$$

Defocus can be embodied in the relationship between low frequency coefficients and middle frequency coefficients. Generally, this relationship will vary by the low energy and high energy contained in these blocks. In detail, as the texture regions have more complex information than smooth region, defocus operation will cause a different visual effect, and the change regularity of AC coefficients is diverse. Therefore, to make the defocus measurement more precise, an energy parameter called  $E_T$  is defined to distinguish strongly textured photograph blocks and weakly textured photograph blocks firstly. Then, a defocus measurement (DM) will be obtained by analysing the focus or defocus characteristics of these blocks. One DC coefficient and  $T_e$  are used in this process, where DM is set as 1 if the characteristic is satisfied with the conditions of Equations (14) and (15); otherwise, it will be set as 0. The details will be described below.

$$E_T = \frac{\sum T_e^2}{DC^2}. \quad (13)$$

If  $E_T$  is greater than a threshold which is set as 0.02 here, this block will be set as a strongly textured block; otherwise, it will be classified into a weakly textured block.

For strongly textured blocks, when defocus happens, blocks will lose partial energy that will reflect in the reduction of DCT coefficients. It has been proved that the three AC coefficients on the low frequency can be used to extract the orientation feature of the block and further construct the direction feature map [36, 37]. Here, we use the max direction energy as an indicator of defocus. Uniquely, in Equation (14), not only the directional energy at low frequency  $E_L = \max \{AC_{0,1}, AC_{1,0}, AC_{1,1}\}$  but the directional energy at middle frequency  $E_M = \max \{AC_{0,2}, AC_{2,0}, AC_{2,2}\}$  are all applied to the defocus measurement in Equation 14. If any of the indices is less than their corresponding threshold, the corresponding DM's value will be set as 0; otherwise, it will be 1. Meanwhile, the thresholds will keep changing in different photographs. Here, an automatic method of setting threshold is established. A characteristic called  $C$  got by pooling  $E_T$  of all blocks is used, which can manifest the whole photograph contrast.

$$E_L < C_0 \cdot C^{0.4} \sqrt{E_M} < C_0 \cdot C^{0.25}, \quad (14)$$

where  $C_0$  is a parameter set at 16.

DC	$AC_{0,1}$	$AC_{0,2}$	$AC_{0,3}$	$AC_{0,4}$	$AC_{0,5}$	$AC_{0,6}$	$AC_{0,7}$
$AC_{1,0}$	$AC_{1,1}$	$AC_{1,2}$	$AC_{1,3}$	$AC_{1,4}$	$AC_{1,5}$	$AC_{1,6}$	$AC_{1,7}$
$AC_{2,0}$	$AC_{2,1}$	$AC_{2,2}$	$AC_{2,3}$	$AC_{2,4}$	$AC_{2,5}$	$AC_{2,6}$	$AC_{2,7}$
$AC_{3,0}$	$AC_{3,1}$	$AC_{3,2}$	$AC_{3,3}$	$AC_{3,4}$	$AC_{3,5}$	$AC_{3,6}$	$AC_{3,7}$
$AC_{4,0}$	$AC_{4,1}$	$AC_{4,2}$	$AC_{4,3}$	$AC_{4,4}$	$AC_{4,5}$	$AC_{4,6}$	$AC_{4,7}$
$AC_{5,0}$	$AC_{5,1}$	$AC_{5,2}$	$AC_{5,3}$	$AC_{5,4}$	$AC_{5,5}$	$AC_{5,6}$	$AC_{5,7}$
$AC_{6,0}$	$AC_{6,1}$	$AC_{6,2}$	$AC_{6,3}$	$AC_{6,4}$	$AC_{6,5}$	$AC_{6,6}$	$AC_{6,7}$
$AC_{7,0}$	$AC_{7,1}$	$AC_{7,2}$	$AC_{7,3}$	$AC_{7,4}$	$AC_{7,5}$	$AC_{7,6}$	$AC_{7,7}$

FIGURE 2: The coefficients for defocus feature calculation (the red blocks) in a photograph DCT block.

For weakly textured blocks, the coefficients are relatively smaller both in focus and defocus regions leading to unavailability of the above judgement. A study has pointed out that the magnitude spectrum will fall with frequency for natural scene inversely and regularly [38]. However, for regions with defocus operation, this drop rate will be faster. Consequently, in the calculation of the new slope  $S$ , the coefficients mentioned above and their spatial frequency are used, which can be denoted by

$$S = \frac{\log E_L - \log E_M}{\log F_L - \log F_M}, \quad (15)$$

where  $F_L$  and  $F_M$  are the corresponding low and middle spatial frequency [38]. If  $S$  is out of a certain range  $-1 < S < 1$ , the corresponding DM's value will be set as 0; otherwise, the corresponding DM's value is 1. Thus, DM can be calculated by Equation (16).

$$DM = \begin{cases} 0 & (E_L < C_0 \cdot C^{0.4} \sqrt{E_M} < C_0 \cdot C^{0.25}) \text{ or } (S < -1 \vee S > 1) \\ 1 & \text{otherwise.} \end{cases} \quad (16)$$

**3.2. Defocus Feature Map.** It is commonly accepted that the HVS is more sensitive to the center-surround differences from the blocks with nearer distance compared with those from the farther distance. Here, we use a Gaussian model to simulate this mechanism for weighting the center-surround differences among photograph blocks for defocus estimation. The final defocus feature map (DFM) of the photograph is calculated as follows.

$$\text{DFM}(n) = \frac{1}{\sqrt{2\pi}} e^{-d(n,\Omega)^2/2} \bullet \text{DM}(\Omega), \quad (17)$$

where  $\Omega$  is a local area centered at block  $n$ , and its size is  $5 \times 5$ . The parameter  $d$  is the Euclidean distance between the central block  $n$  and its circumjacent blocks in  $\Omega$ . If the result is less than the value set in the focus measurement above, that is 1, this will be considered to be a defocus block in DFM finally. Figure 3 shows examples of photographs' DFM using the proposed method. And Figure 3(b) is the result obtained by the sum of all AC coefficients. Because AC coefficients can be a representation of defocus, which comes from the theory that taking signal power of all coefficients in the corresponding subbands to measure focus is one possible way [28]. Figure 3(d) is obtained by the focus location of the human eyes. It is observed from the figure that the focus region is significantly located. Meantime, it is also noticed that the defocus photographs exhibit obvious artistic quality, particularly in the highlighted regions in Figure 3(d). Since photograph defocus measurement is a typical perceptual loss, which also sheds light on how the quality of defocus photographs should be measured. Thus, the proposed DFM in Figure 3(c) is effective in focus location whatever in subjective and objective results.

**3.3. The Defocus Feature Map-Based Contrast Masking.** Based on the above description, we can obtain the final DFM of a photograph. Existing studies have shown that the contrast masking effect and the visibility threshold are guided by the defocus strength. Therefore, the CM effect can be obtained with the guidance of the final DFM. Generally, the traditional CM overestimates the values in the focus regions because the higher the focus is, the lower the masking ability is. Thus, a smaller JND value is needed in the focus blocks, and the defocus factor should be larger than that in focus area. In addition, besides the effect of focus and defocus, the spatial frequency is considered in the calculation of CM, which comes from the theory that sensitivity of the HVS is different for low and high frequency. And our proposed CM factor deals with this problem as shown in Equation (18). The weights of defocus and focus are set as 1.75 and 1, and the weights of strongly textured and weakly textured blocks are set as 1.5 and 1, respectively.

$$M_{\text{CM}} = \begin{cases} 1 & \text{weakly textured focus block,} \\ \begin{cases} 2.625 & \text{strongly textured defocus block,} \\ 1.25 & \text{strongly textured defocus block,} \end{cases} \\ \begin{cases} 1.5 & \text{strongly textured focus block,} \\ 1.25 & \text{strongly textured focus block,} \end{cases} \\ \begin{cases} 1.75 & \text{weakly textured defocus block,} \\ 1.25 & \text{weakly textured defocus block.} \end{cases} \end{cases} \quad (18)$$

This  $M_{\text{CM}}$  will be applied in Equation (1) to obtain a new JND model which will be used in quantization-based watermarking framework.

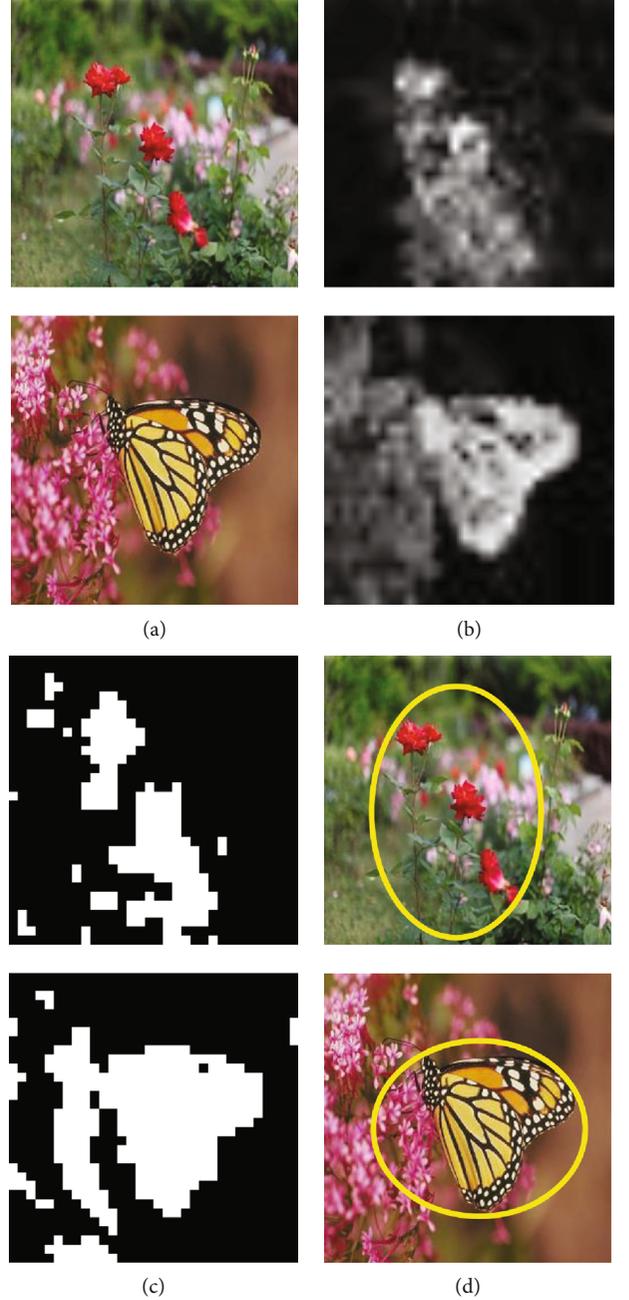


FIGURE 3: Illustration of the focus region in the original photographs and the defocus feature map calculated by Equation (6): (a) original photographs; (b) the comparison focus method in [28]; (c) the proposed defocus feature map (d) human eyes location in the yellow circle.

## 4. The Proposed JND-Based Photograph Watermarking Scheme

**4.1. Watermark Embedding.** The watermark is embedded into  $Y$  channel. And the characteristics of RGB channels are combined to get the final parameters, where their weights are 0.299, 0.587, and 0.114, respectively. The watermark embedding and extraction strategy are STDM, where the carrier vector  $\mathbf{x}$  is made up of  $AC_{0,3}$ ,  $AC_{0,4}$ ,  $AC_{1,2}$ ,  $AC_{2,1}$ ,  $AC_{2,3}$

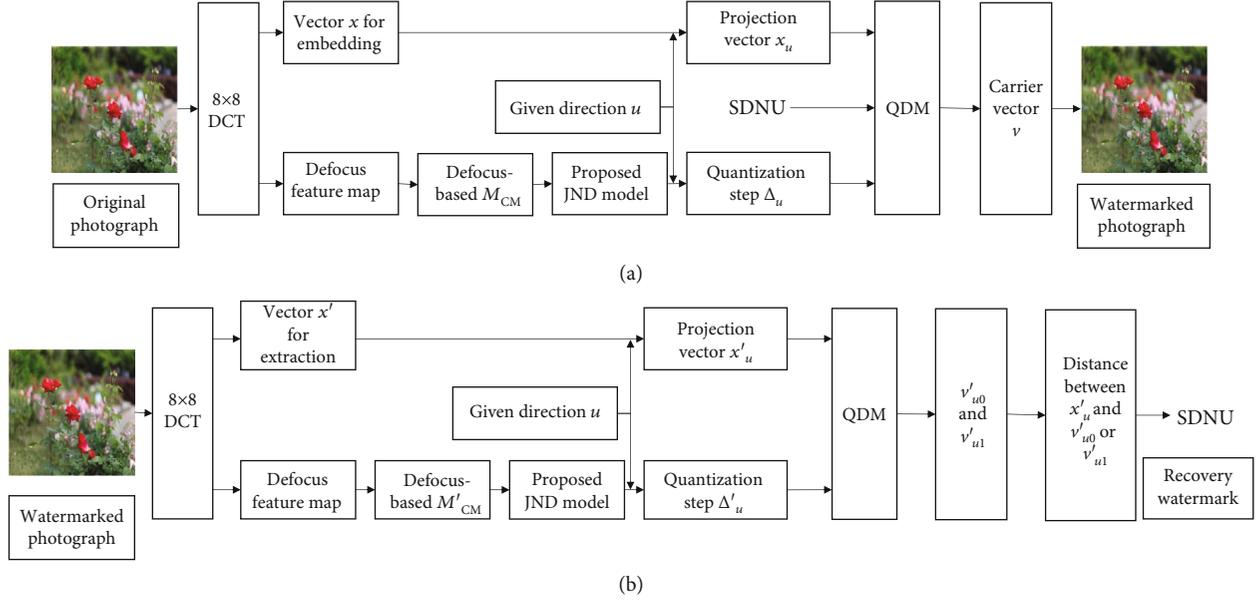


FIGURE 4: Block diagram of proposed watermark embedding process (a) and extraction process (b).

,  $AC_{3,0}$ , and  $AC_{4,0}$  (the subscripts refer to the row and column indices of the block) as shown in Figure 2. And the block diagram is shown in Figure 4.

In detail, the proposed JND model plays an important role here, which is used to calculate the STDN quantization step size. The value of the specific position of the JND for each DCT block is scanned to form a visual redundancy vector  $\Delta$ . Then, the vector is projected onto the given direction vector  $u$  to get the quantization step  $\Delta_u$  using Equation (10). At the same time, the carrier vector  $x$  is projected onto the same direction vector  $u$  to obtain the projection vector  $x_u$ . Then, QDM operation is implemented. And the watermark bits will be embedded in the carrier  $x$  by Equation (19). The details about this process are given in Related Work. The main steps of watermarking embedding can be described in Algorithm 1.

$$v = x + \frac{QDM(x_u, \Delta_u, d_m) - x_u}{\|u\|} u, \quad (19)$$

**4.2. Watermark Extraction.** And the watermarking extraction is the inverse process of watermarking embedding, where the watermarking bits can be extracted by Equation (20). Firstly, the watermarked photograph  $I'_w$  that may be distorted in RGB channels is converted into YCbCr channels. Then, the Y channel is divided into  $8 \times 8$  nonoverlapped blocks. For block  $n$ , the watermark bit can be extracted by Equation (20). The details are described in Related Work. And  $\Delta'$  is the original quantization step that is obtained by the calculated JND in the watermarked photograph  $I'_w$ . It is projected onto the same directional vector  $u$  getting  $\Delta'_u$ . The carrier vector  $x'$  is pro-

jected onto  $u$  getting  $x'_u$ . The main steps of watermarking embedding can be described in Algorithm 2.

$$m = \operatorname{argmin} \operatorname{dist}(x'_u, QDM(x'_u, \Delta'_u, d_m)), m \in \{0, 1\}. \quad (20)$$

## 5. Experimental Results

In this section, we give the experimental results and analysis. To prove the performance of the proposed watermarking scheme, we perform the experiments using the original code in MATLAB R2016b on a 64-bit Windows 10 operating system at 16 GB memory, 3.20 GHz frequency of Intel processors. Eight aesthetic RGB photographs ( $256 \times 256$ ) with local defocus effect from multifocus image dataset, CSIQ, and LIVE database are used in the experiment [39, 40]. They are named Flower1, Monarch, Toy, Couple, Parrots, Jug, Flower2, and Flower3. And the watermark is made up of 1024 binary bits. They can be shown in Figures 5 and 6.

The bit error rate (BER) is used in the robustness test with SSIM is 0.982 between the watermarked photograph and the original photograph. Because SSIM can reflect the human eyes' sensitivity better and HVS will not perceive the difference between the original host photograph and the watermarked host photograph. These two parameters are shown in Equations (21) and (22), respectively. The robustness of the proposed watermarking method is compared with some JND models [25, 30–32] and some watermarking methods [18, 20–22].

$$\operatorname{BER} = \frac{m \oplus m'}{M \times N}, \quad (21)$$

Watermarking embedding.

**Begin**

**Input:** Host photograph ( $I$ ); watermark ( $m$ );

**Output:** Watermarked photograph ( $I_w$ );

**Step 1:** Transform the RGB space of the photograph to YCbCr space. The Y channel is regarded as the watermark embedding channel, and the channels RGB are all used in feature extraction;

**Step 2:** Divide the Y channel image into  $8 \times 8$  non-overlapped blocks;

**Step 3:** The embedding strength can be set to control the photograph quality by modulating the SSIM of the original photograph and the watermarked photograph to 0.982;

**Step 4:** One bit of binary watermark  $m$  is embedded into a photograph block, according to the rules as follows:

**For each block do**

1. Estimate the perceptual JND factors including CSF, LA and CM effect by Equations (2), (6) and (18), respectively. And the block types in Equation (18) are distinguished by defocus measurement in Equations (13), (14), (15), (16) and (17);
2. Final JND value of each block combined with different factors is determined by Equation (1), which is the original quantization step  $\Delta$ ;
3. Obtain the DCT coefficients that comprise the embedding vector  $\mathbf{x}$ ;
4. Project  $\mathbf{x}$  and  $\Delta$  to the given direction  $\mathbf{u}$  and get two new vectors  $\mathbf{x}_u$  and  $\Delta_u$ ;
5. One bit of the watermark message  $m$  is embedded into this vector  $\mathbf{x}$  using STDM by Equation (19);
6. Generate the modified block;

**End**

**Step 5:** Generate the watermarked photograph Y by collecting all the modified blocks;

**Step 6:** Generate the watermarked color photograph by concatenating the modified Y channel with Cb and Cr photograph channel and then convert the color space from YCbCr to RGB;

**Return** Watermarked photograph  $I_w$ ;

ALGORITHM 1.

Watermarking extraction.

**Begin**

**Input:** Watermarked photograph  $I_w'$ ;

**Output:** Watermark Image  $m'$ ;

**Step 1:** Transform the photograph from RGB space to YCbCr space. Select the channel Y as the host photograph;

**Step 2:** Divide the host photograph into  $8 \times 8$  non-overlapped blocks;

**For each block do**

1. Obtain the vector  $\mathbf{x}'$  and the original quantization step  $\Delta'$  from the watermarked photograph  $I_w'$ ;
2. Project these two vectors to the given direction  $\mathbf{u}$  and get  $\mathbf{x}'_u$  and  $\Delta'_u$ ;
3. And one bit of the watermark message is extracted by Equation (20);

$$m = \operatorname{argmin} \operatorname{dist}(\mathbf{x}'_u, \operatorname{QDM}(\mathbf{x}'_u, \Delta'_u, d_m)), m \in \{0, 1\}$$

**End**

**Return** Watermark message  $m'$ ;

ALGORITHM 2.

where  $m$  and  $m'$  are the original watermark image and extracted watermark image whose size are  $M$  and  $N$ , respectively.

$$\text{SSIM} = \frac{(2u_X u_Y + C_1)(2\sigma_{XY} + C_2)}{(u_X^2 + u_Y^2 + C_1)(\sigma_X^2 + \sigma_Y^2 + C_2)}, \quad (22)$$

where  $u_X$  and  $u_Y$  represent the mean of the luminance of the reference photograph and the watermarked photographs,  $\sigma_X$  and  $\sigma_Y$ , represent the variances of two photographs, and  $\sigma_{XY}$  represents the covariance of  $X$  and  $Y$ .  $C_1$  and  $C_2$  are positive experience values.

### 5.1. Robustness Test for Single Attack

*5.1.1. BER Performance of Different JND-Based STDM Watermarking and Robust Watermarking Methods.* The robustness test is essential for watermarking, especially for the JND based on the defocus measurement, as the noise will give an interfere to this defocus judgement. Here, the performance of some JND models are used in the STDM watermarking method, which are all used to calculate the quantization step. A lower BER indicates a better balance between robustness and invisibility. The watermarked photograph Flower1 is added with some common attacks such like the Gaussian noise attack (GN) with mean zero and different variance, the Salt and Pepper noise attack (S&P) with different quantity, the JPEG compression attack where the JPEG quality factor varies from 40 to 80, the Scaling attack reduced

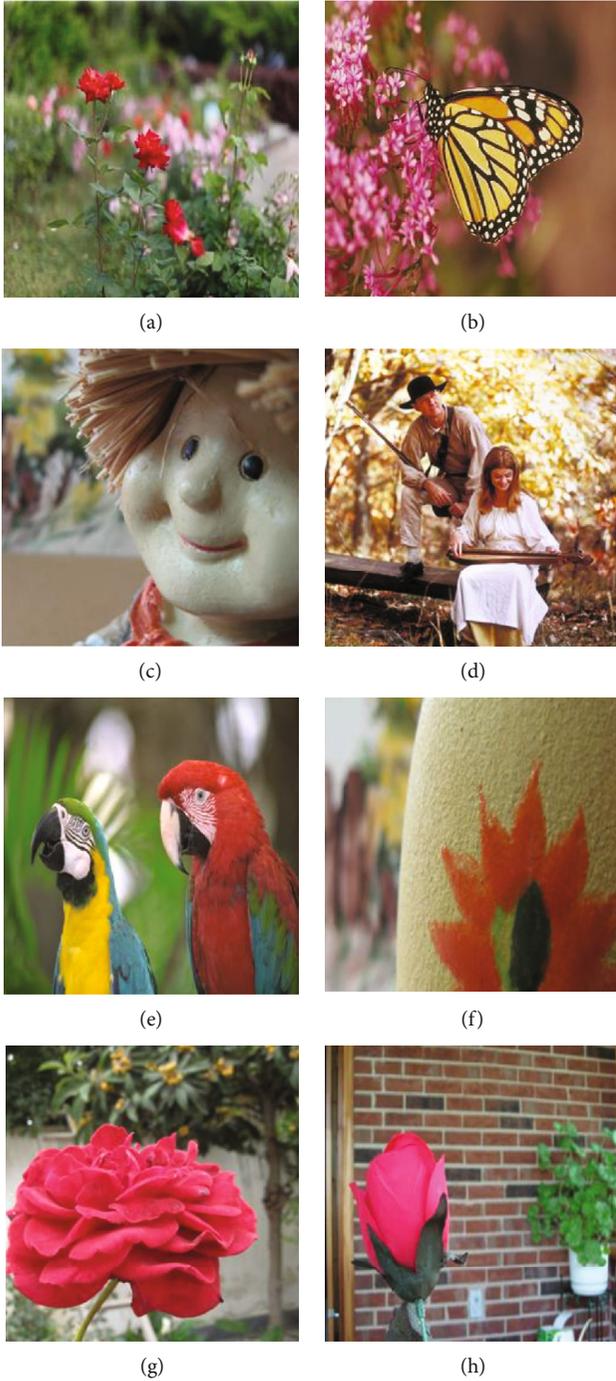


FIGURE 5: Typical photographs selected in the experiment.

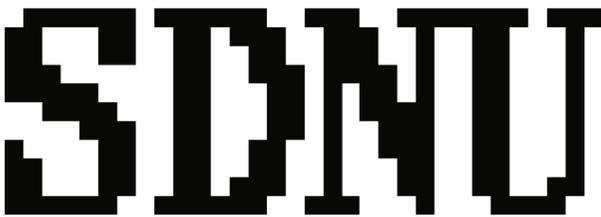


FIGURE 6: The binary watermark.

to 20% and increased to 140%, the Gaussian Filtering (GF) attack where its window size is  $3 \times 3$ , and Central Crop attack whose size is  $16 \times 16$  and  $32 \times 32$ . The ability of different JND models [25, 30–32] in resisting attacks is presented in Table 1.

From Table 1, we can see that the BER of the proposed JND model is lower than the presented JND models obviously. It can be seen that the robustness performance is almost improved by 15% for GN0.0015 attack in the tested photograph. For S&P0.0015 noise attack, the BER is about 2.5% lower than other methods. For JPEG attack, we can see that the watermarking framework using the proposed JND model have a lower BER in JPEG attack. Moreover, when the attack is Scaling 0.5, the performance of rest watermarking methods using JND as quantization even reaches to 28%. We can see that these JND models used as the quantization step in STDM may lack some robustness to Scaling attack. For Central Crop attack, these JND-based STDM watermarking methods show acceptable robustness, and the proposed method is the most robust. For GF, the proposed also shows a best performance. Thus, with the same visual quality, compared with these JND models, the proposed JND model shows a better robustness performance under some attacks, and we believe that it can adjust the quantization step better.

A JND model can be used as the quantization step in STDM to avoid unacceptable visual quality after embedding the watermark. From Table 1, we can see that the proposed JND model is more suitable for watermarking technology using STDM as its better robustness performance when keeping a same visual quality. However, it is not persuasive to show the robustness comparison of JND models because some existing JND models may involve features that cannot resist common attacks adaptively in general. That is to say, the coefficients selected in these JND models are not robust. Thus, it is essential to indicate the proposed JND-based watermarking methods have a better robustness than some existing robust watermarking methods. Besides JND models, we compare the robustness performance of other four watermarking methods in frequency domain, which combined encryption technique, perceptual models, or improved embedding way [18, 20–22]. These methods devote to improving the robustness of watermarking by typical ways. The comparison results are also shown in Table 1.

From Table 1, we can see that our proposed JND-based watermarking method shows a better robustness to noise and JPEG attacks. The robustness to GN0.0015 and S&P0.0015 is improved by about 8% and 2.5%, respectively. Compared with the BER of other method, this improvement can bring about a fresh effect. Because adding noise is the most common image processing technique to verify robustness of the watermarking framework. This robustness to noise attack achieves an important lift in watermarking framework. The robustness to JPEG40 and JPEG60 attack also presents a better result than other watermarking methods. For crop attack, the proposed method has the lowest BER among these robustness watermarking methods. For GF attack, the BER of method [18] is about 0.7% lower than the proposed method. For Scaling attack, the methods [18,

TABLE 1: The BERs of different watermarking methods for individual attacks in photograph Flower1.

Attack		[25]	[30]	[32]	[31]	[20]	[18]	[21]	[22]	Proposed
GN	0.0003	0.0459	0.0254	0.0107	0.1416	0.0283	0.0156	0.0068	0.0049	0.0010
	0.0009	0.1045	0.1006	0.0820	0.2754	0.0674	0.1016	0.1006	0.0771	0.0078
	0.0015	0.1553	0.1748	0.1680	0.2988	0.0977	0.1633	0.1836	0.1514	0.0166
S&P noise	0.0009	0.0293	0.0254	0.0264	0.0254	0.0352	0.0203	0.0244	0.0264	0.0029
	0.0012	0.0371	0.0313	0.0322	0.0332	0.0410	0.0322	0.0361	0.0332	0.0098
	0.0015	0.0498	0.0391	0.0410	0.0459	0.0449	0.0396	0.0410	0.0410	0.0156
JPEG	80	0.0234	0.0049	0.0029	0.0391	0.0215	0.0926	0	0.0020	0.0010
	60	0.0361	0.0166	0.0049	0.0713	0.0410	0.4226	0.0078	0.0068	0.0020
	40	0.0811	0.0586	0.0439	0.1299	0.0820	0.4819	0.0586	0.0449	0.0225
Scaling	0.5	0.0889	0.2168	0.2852	0.1035	0.0205	0.0005	0	0.0010	0.0166
	1.1	0.0098	0.0049	0.0020	0.0137	0.0186	0.0005	0	0	0
Central crop	16 × 16	0.0186	0.0068	0.0049	0.0107	0.0264	0.0240	0.0039	0.0049	0.0020
	32 × 32	0.0293	0.0156	0.0156	0.0186	0.0322	0.0283	0.0132	0.0127	0.0049
GF	3 × 3	0.0127	0.0195	0.0146	0.0322	0.0361	0.0032	0.0166	0.0088	0.0107

21, 22] also show an acceptable robustness performance. Especially for Scaling 0.5, the highest gap with our approach is at about 1.7%, which has little influence to the integrity of the watermark. But compared with the robustness performance to noise attacks, our method performs a better robustness as a whole.

*5.1.2. Average BER Performance within the above Watermarking Methods.* In the above subsection, we present the robustness performance of some JND models used in the STDM watermarking framework and some robust watermarking methods for photograph Flower. To make the result more reasonable, we further consider eight photographs as shown in Figure 5 to demonstrate the performance of the proposed method. Figures 7 and 8 present the average results of the above comparison methods under different attacks of varying intensity. Besides the attacks mentioned above, the BERs of GF attack and Central Crop attack are shown in Table 2.

In watermarking technology, adding noise is the most common attack to verify the robustness. In this paper, GN and S&P noise are chosen in this test. With the increase of the attack factors, all comparison methods' BER will increase. For the methods using JND as quantization step in STDM, the BERs of GN are more than 20% even 30%. The proposed JND-based watermarking method is about 8%. This improvement will ensure the availability of watermark. In the listed attacks with different noise quantity, the BERs are all less than 5%. In contrast, the BERs of the proposed method for S&P noise are the lowest in all methods.

JPEG is another common attack in photograph processing. The photographs may be attacked by JPEG during transmission. Figure 7(c) shows the partial experimental results of watermarked photographs with compression factors 20 to 80. When the JPEG factor is 40, the proposed method shows a better robustness with the BER at about 10%. When the JPEG factor is 60, the BER is about 2%. Compared with other

methods, the proposed method presents a higher robustness in JPEG compression.

For Scaling attack, the STDM watermarking methods using other JND models cannot satisfy the basic robustness. And the proposed method can resist this attack better.

From Figure 8, we can see that the proposed method shows a better performance in noise attacks and partial JPEG attack. In particular, the robustness to GN is enhanced to a better result obviously. As for Scaling, the BERs of these two methods [21, 22] are slightly lower.

In Figure 8, for other four watermarking methods, the change of BERs is great. Especially in GN0.0015, the BERs are within 14%-27%. The proposed JND-based watermarking method is about 8%. In the listed attacks with different S&P noise quantity, the BERs are all less than 5%. And the BERs are less than 2.5%. Thus, from the performance of the proposed method to noise attack, a high robustness is obtained than other robustness watermarking methods, which can indicate the importance of the defocus measurement in JND.

The JPEG format is common for photographs. Thus, the watermarking method needs to show an acceptable robustness performance to this attack. Figure 8 shows the partial experimental results of watermarked photographs with compression factors 20 to 80. We can see that the method [20] performs well in JPEG20 and JPEG30, but the BER declines slowly in other JPEG factors. For the common JPEG50 attack, the proposed method shows a lower BER. Therefore, the coefficients used in defocus feature calculation and watermark embedding are robust to this attack.

For Scaling attack, we give the average BERs when the watermarked photographs are scaled from 10% to 140%. We can see that the method [21, 22] performs better in Scaling attack.

From Table 2, it can be seen that the BER of [22] is the lowest in GF attack, which is 0.25% lower than the proposed scheme. They are all at about 4%. The BERs of other methods

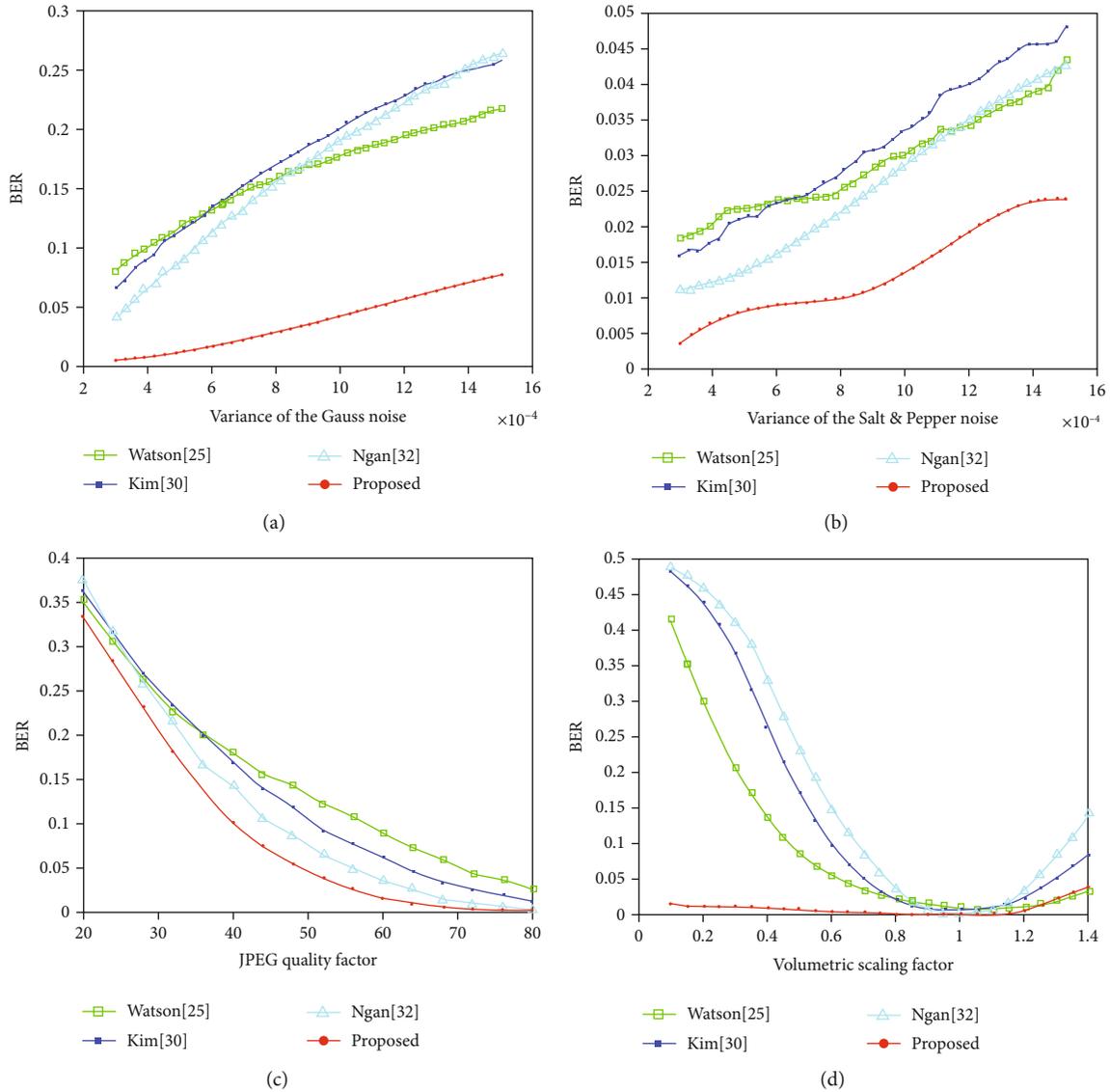


FIGURE 7: The robustness performance of different JND models in STDM watermarking.

are also almost 4% except [18, 31]. This difference will not bring a great change in the extracted watermark. For Central Crop attack, the BERs of these methods are all within 4%. But the BER of the proposed method is less than 0.4%.

Above all, the methods [21, 22] perform better with some attacks. But these two methods do not have enough robustness to noise attacks. Thus, the proposed method achieves a better robustness in the whole. The importance of the defocus measurement in JND is indicated.

**5.2. Robustness Test for Combined Attacks.** For photographs, resisting some common attacks under JPEG is important in transmission. Thus, the combined attack experiments are carried out, especially stressing on the influence of JPEG. And the experimental results are list in Table 3. It can be seen that the noise and filtering attacks are combined with JPEG50

and JPEG70. And the proposed method shows a better robustness performance under all these tested attacks.

For the combination of GN noise and JPEG, the BER of the proposed JND-based watermarking method achieves a better performance than other methods of comparison. When attacked by JPEG50 + GN0.0003, the BER is 7.56% while the lowest BER of the other methods is 12.96% in [20]. Similar to JPEG50 + GN0.0003, the BER of JPEG50 + GN0.0009 attack is more than 5% lower than other methods. In contrast with JPEG + GN, there is little difference in the robustness to JPEG + S&P attack of all methods. However, the BER of the proposed method is more than 2% lower than others under JPEG50. This superiority will decrease to about 1% under JPEG70. For the combination of GF and JPEG, the proposed method shows a superior result. In sum, the proposed JND-based watermarking method shows a satisfy performance in the combined attacks.

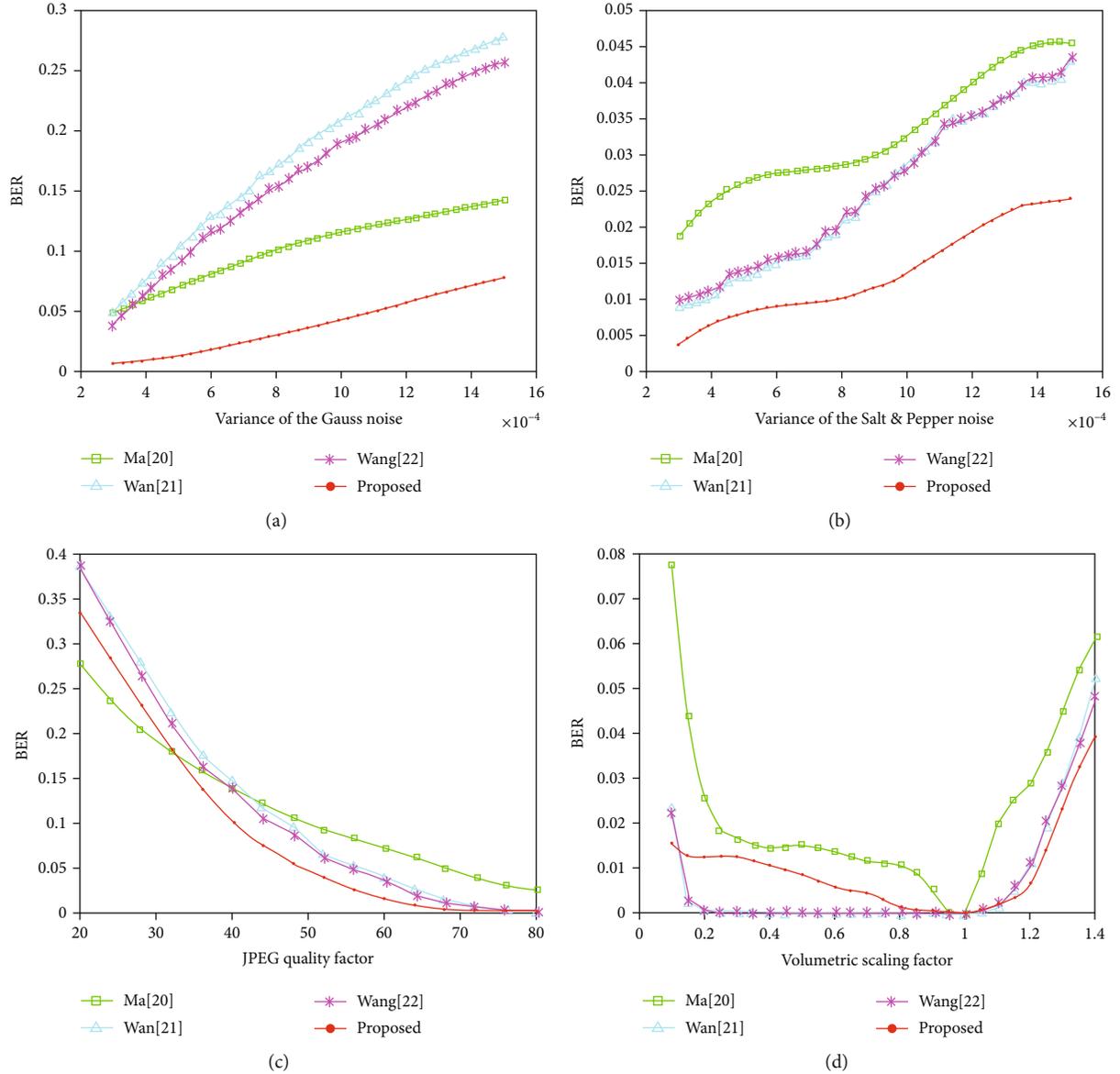


FIGURE 8: The robustness performance of different robust watermarking method.

TABLE 2: The BERs of different watermarking methods in individual attacks.

Attack		[25]	[30]	[32]	[31]	[20]	[18]	[21]	[22]	Proposed
Central crop	$16 \times 16$	0.0142	0.0137	0.0059	0.0112	0.0240	0.0273	0.0039	0.0044	0.0020
	$32 \times 32$	0.0240	0.0195	0.0151	0.0157	0.0283	0.0313	0.0132	0.0117	0.0039
GF	$3 \times 3$	0.0396	0.0412	0.0398	0.0735	0.0420	0.0830	0.0393	0.0363	0.0388

Then, to indicate the availability of watermark, Figure 9 shows the watermarks extracted from the distorted watermarked photographs. And from Figure 8 and Table 2, it can be seen that the methods [21, 22] perform better with Scaling and GF attacks. Thus, the extracted watermarks using these two methods are displayed in Figure 9. It displays the visual effects of two attacks that combine noise attack, JPEG, or filtering attack. In Figure 9, the photograph Couple, Flower2, and Flower3 in Figure 5 are selected to present the

visual effect of the extracted watermark. It can be seen that the difference of visual effect in the Scaling+JPEG70 attack is not obvious. But for noise attack combined with JPEG70, the visual effect of the extracted watermark makes a better achievement. Because for GN attack, the extracted watermark used method [21, 22] even cannot be recognized in the photographs, which will make it useless to embed a watermark. In addition, in the combination of GN and GF attack, the above two comparison methods cannot extract

TABLE 3: The BERs of different watermarking methods in combined attacks.

Attack		[25]	[30]	[32]	[31]	[20]	[18]	[21]	[22]	Proposed
JPEG50 + GN	0.0003	0.1766	0.1605	0.1342	0.2103	0.1296	0.4384	0.1479	0.1343	0.0756
	0.0009	0.2238	0.2373	0.2230	0.3215	0.1636	0.4560	0.2396	0.2217	0.1142
JPEG70 + GN	0.0003	0.1201	0.1007	0.0740	0.1556	0.0858	0.3847	0.0803	0.0709	0.0198
	0.0009	0.1840	0.2000	0.1815	0.2896	0.1274	0.4160	0.1964	0.1801	0.0596
JPEG50 + S&P	0.0003	0.1294	0.1121	0.0826	0.1781	0.1038	0.4257	0.0856	0.0797	0.0523
	0.0009	0.1360	0.1212	0.0930	0.1853	0.1106	0.4169	0.0956	0.0903	0.0641
JPEG70 + S&P	0.0003	0.0577	0.0396	0.0229	0.0951	0.1038	0.3295	0.0227	0.0201	0.0106
	0.0009	0.0652	0.0527	0.0371	0.1069	0.0612	0.3544	0.0376	0.0336	0.0195
JPEG50 + GF	3 × 3	0.1489	0.1396	0.1216	0.2059	0.1168	0.4287	0.1268	0.1130	0.0962
	5 × 5	0.1484	0.1394	0.1218	0.2057	0.1172	0.4277	0.1272	0.1133	0.0963
JPEG70 + GF	3 × 3	0.0851	0.0718	0.0601	0.1321	0.0736	0.3251	0.0627	0.0549	0.0540
	5 × 5	0.0858	0.0723	0.0602	0.1322	0.0741	0.3252	0.0645	0.0553	0.0541

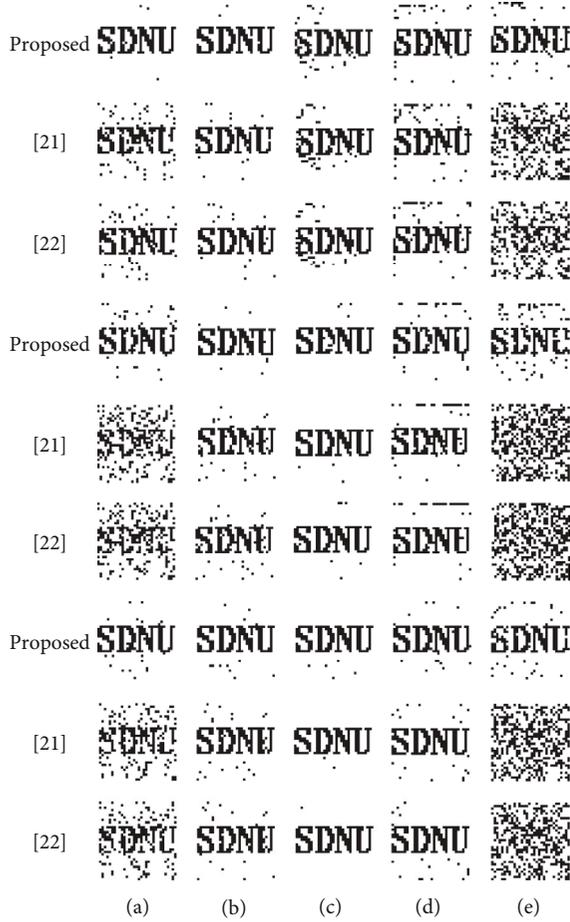


FIGURE 9: The photographs. Comparison of the extracted watermark image in [21, 22] and the proposed method: (a) JPEG 70 + GN0.0012; (b) JPEG70 + S&P 0.0012; (c) JPEG70 + Scaling1.1 ; (d) JPEG70 + GF5 × 5; (e) GN0.0009 + GF3 × 3.

the watermark in visual effect. In brief, the proposed scheme presents a better watermark visual quality than [21, 22].

In summary, our proposed watermarking method works well than other methods that treat every area of a photograph

with an equal focus level. In detail, the reason of our proposed JND model performs better than other JND models used in STDM watermarking framework is that the chosen AC coefficient is robust in defocus measurement when the visual effect is maintained at a reasonable and acceptable level. And compared with some watermarking methods, the performance of our proposed method indicates the significance of defocus measurement in the protection of aesthetic photographs.

*5.3. Average BER Performance within Different Watermarking Methods for Classic Images.* In order to show that our algorithm is not only effective for photographs with obvious defocus areas, we also use eight images of areas that are not visible to the human eyes to carry out the experiment. The purpose of this test is that although the core of our algorithm is the measurement of defocus, if there is no obvious defocusing effect in the image, the proposed model will be calculated according to the amount of information contained in the image block. Some common single attacks and combined attacks are tested in eight images ( $256 \times 256$ ) in CSIQ and LIVE database as shown as in Figure 10. And Figure 6 is used as the watermark.

From Table 4, it can be seen that for noise attacks, the proposed method is more robust than other methods. For GN0.0015 attack, the BER of the proposed method is 6.51% while the best performance in other methods reaches to 12.33%. For S&P0.0015 noise attack, the BER performance is improved by 1.32%. For JPEG attack, the method in [22] performs better for JPEG80, which is 0.53% lower than the proposed method. And for JPEG40 and JPEG60, the proposed method performs a better robustness. For Scaling attack, the method in [21, 22] performs better on the whole. The BER is decreased by 0.37% and 1.88%. For Crop attack, the proposed method achieved a better robustness performance in the crop block size of  $32 \times 32$ , which increased by 0.56%. In contrast, the BER in [22] is about 0.3% lower than the proposed method when the crop block size is  $32 \times 32$ . For GF attack, the BER is 0.08% higher than [22]. To sum up, the method in [21, 22] performs better than the proposed



FIGURE 10: Classic images selected in the experiment.

TABLE 4: The BERs of different watermarking methods in individual attacks.

Attack		[25]	[30]	[32]	[31]	[20]	[18]	[21]	[22]	Proposed
GN		0.0003	0.0710	0.0514	0.0291	0.1212	0.0499	0.0864	0.0316	0.0070
		0.0009	0.1444	0.1589	0.1327	0.2211	0.0901	0.1138	0.1429	0.1281
		0.0015	0.1873	0.2274	0.2081	0.2623	0.1233	0.1494	0.2235	0.2047
S&P noise		0.0009	0.0243	0.0314	0.0205	0.0316	0.0301	0.0796	0.0210	0.0220
		0.0012	0.0305	0.0400	0.0293	0.0399	0.0356	0.1035	0.0299	0.0313
		0.0015	0.0393	0.0486	0.0381	0.0491	0.0407	0.1309	0.0374	0.0397
JPEG	80	0.0267	0.0156	0.0040	0.0907	0.0299	0.1948	0.0013	0.0007	0.0060
	60	0.0713	0.0474	0.0240	0.1930	0.0650	0.3438	0.0254	0.0212	0.0187
	40	0.1492	0.1294	0.0990	0.3021	0.1190	0.3931	0.1077	0.1001	0.0891
Scaling	0.5	0.0984	0.2188	0.2887	0.1453	0.0195	0.0771	0	0.0002	0.0188
	1.1	0.0101	0.0200	0.0112	0.0370	0.0220	0.0791	0.0033	0.0022	0.0059
Central crop	16 × 16	0.0145	0.0198	0.0063	0.0148	0.0202	0.0771	0.0039	0.0035	0.0067
	32 × 32	0.0231	0.0205	0.0133	0.0215	0.0240	0.0776	0.0107	0.0106	0.0050
GF	3 × 3	0.0461	0.0571	0.0496	0.1422	0.0526	0.1136	0.0519	0.0420	0.0428

method for some attacks. However, these improvements are within a small range, while the proposed method's improvements of other attacks are obvious. Thus, the robustness performance of the proposed method is acceptable in classic images. Moreover, the watermarked images will also be attacked by the combined attacks. Table 5 lists some BER results under combined attacks. And to indicate the availability of the extracted watermark, Figure 11 shows the extracted watermark under some combined attacks.

From Table 5, we can see that the proposed method has a better performance than others except JPEG70 + S&P attack. For the combination of JPEG and GN0.0003, the BER is about 4% lower than [22]. When the intensity of GN is

0.0009, this optimization is up to about 7%. For JPEG50 + S&P0.0003 attack, there is about 1% decrease in BER of the proposed method. With the increase of S&P intensity, the BER difference between [22] and the proposed method is narrowing. When the JPEG factor is 70, the BER in [22] is 0.04% and 0.43% lower than the proposed. For JPEG + GF attack, the proposed method has a better performance though there are no big improvements. Figure 11 shows the extracted watermark of three watermarked images. In JPEG 70 + GN0.0012 and GN0.0009 + GF3 × 3 attacks, the method [21, 22] even cannot extract the distinguishable watermark in some images. Taken altogether, the proposed method has a better visual effect under the listed combined attacks.

TABLE 5: The BERs of different watermarking methods in combined attacks.

Attack		[25]	[30]	[32]	[31].	[20]	[18]	[21]	[22]	Proposed
JPEG50 + GN	0.0003	0.1449	0.1309	0.1084	0.1959	0.1120	0.4131	0.1182	0.0999	0.0616
	0.0009	0.1896	0.2051	0.1904	0.2655	0.1378	0.4224	0.2018	0.1831	0.1196
JPEG70 + GN	0.0003	0.0997	0.0853	0.0580	0.1422	0.0738	0.3232	0.0616	0.0542	0.0214
	0.0009	0.1548	0.1783	0.1541	0.2402	0.1090	0.3491	0.1669	0.1475	0.0719
JPEG50 + S&P	0.0003	0.1180	0.0836	0.0590	0.2367	0.0929	0.3809	0.0650	0.0563	0.0471
	0.0009	0.1243	0.0935	0.0691	0.2419	0.0963	0.3892	0.0750	0.0654	0.0625
JPEG70 + S&P	0.0003	0.0533	0.0314	0.0184	0.1438	0.0477	0.2935	0.0164	0.0139	0.0143
	0.0009	0.0601	0.0433	0.0298	0.1537	0.0537	0.3032	0.0277	0.0260	0.0303
JPEG50 + GF	3 × 3	0.1383	0.1224	0.1018	0.2737	0.1131	0.3813	0.1096	0.0929	0.0916
	5 × 5	0.1388	0.1218	0.1016	0.2738	0.1133	0.3828	0.1099	0.0931	0.0922
JPEG70 + GF	3 × 3	0.0822	0.0691	0.0625	0.1981	0.0773	0.3081	0.0684	0.0557	0.0544
	5 × 5	0.0825	0.0696	0.0632	0.1989	0.0775	0.3101	0.0693	0.0566	0.0551

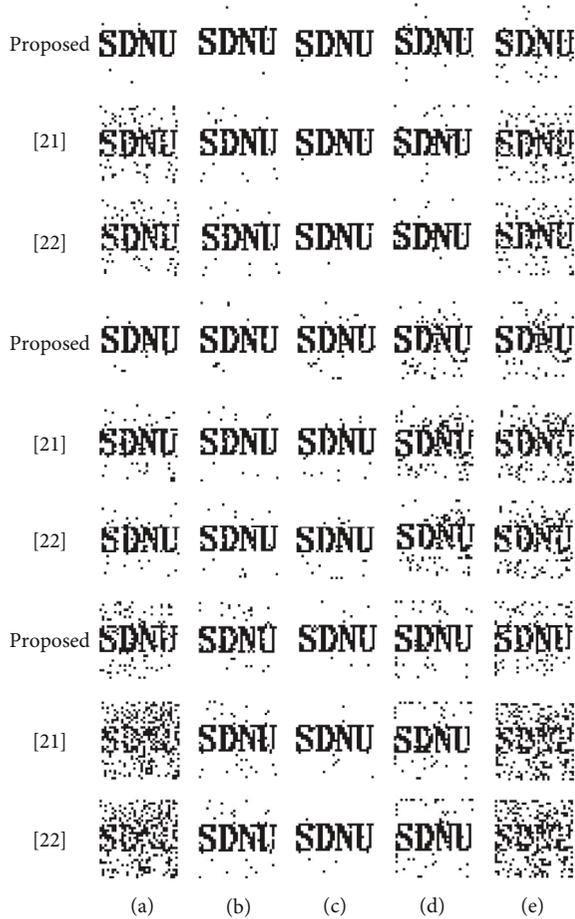


FIGURE 11: The classic images. Comparison of the extracted watermark image in [21, 22] and the proposed method: (a) JPEG70 + GN0.0012; (b) JPEG70 + S&P0.0012; (c) JPEG70 + Scaling1.1; (d) JPEG70 + GF5 × 5; (e) GN0.0009 + GF3 × 3.

## 6. Conclusion

In this paper, a blind watermarking method is proposed to protect aesthetic photographs based on a new JND model with defocus analysis. Firstly, we propose a new defocus-

based JND model which can reflect the characteristic of HVS that the eyes will take more time to gaze at focus areas. Then, the calculated JND is applied as the quantization step in STD. Moreover, the selected binary watermark image is embedded into the host vector that is made up of some AC coefficients in DCT domain. The proposed watermarking scheme is tested under some common attacks such like noise attack, compression attack, scaling attack, filtering attack, and crop attack. In robustness test, the watermarked photographs will be carried out individual attacks such like noise attack, compression attack, scaling attack, filtering attack, and crop attack. Besides, in view of the noise attack may be added under JPEG for photographs particularly, some combined attacks are performed to show the effectiveness of the proposed method. And the performance of the proposed method is also tested in classic images. The experimental results demonstrate that our proposed method has a better robustness performance in the same visual quality than other JND models and watermarking methods. For future work, we may use the proposed defocus measurement for JND model by edge detection, because defocus will increase the edge width where HVS is sensitive. And the aesthetic photograph database may be created in the future for further study.

## Data Availability

The image data used to support the findings of this study are included within the article.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This project is supported by the Natural Science Foundation of China (61601268 and U1736122), the Natural Science Foundation for Distinguished Young Scholars of Shandong Province (JQ201718), and the Shandong Provincial Key Research and Development Plan (2017CXGC1504).

## References

- [1] H. Wang, D. He, J. Shen, Z. Zheng, C. Zhao, and M. Zhao, "Verifiable outsourced ciphertext-policy attribute-based encryption in cloud computing," *Soft Computing*, vol. 21, no. 24, pp. 7325–7335, 2017.
- [2] W. Zhang and W. Zhang, "Algebraic techniques on searching linear diffusion layers in block cipher," *Security and Communication Networks*, vol. 9, no. 17, pp. 4285–4294, 2016.
- [3] Y. Song, H. Wang, X. Wei, and L. Wu, "Efficient attribute-based encryption with privacy-preserving key generation and its application in industrial cloud," *Security and Communication Networks*, vol. 2019, Article ID 3249726, 9 pages, 2019.
- [4] Z. Zhou, Y. Mu, and Q. M. J. Wu, "Coverless image steganography using partial-duplicate image retrieval," *Soft Computing*, vol. 23, no. 13, pp. 4927–4938, 2019.
- [5] X. Zheng, J. Tian, X. Xiao, X. Cui, and X. Yu, "A heuristic survivable virtual network mapping algorithm," *Soft Computing*, vol. 23, no. 5, pp. 1453–1463, 2019.
- [6] Z. Zhou, K. Lin, Y. Cao, C. N. Yang, and Y. Liu, "Near-duplicate image detection system using coarse-to-fine matching scheme based on global and local CNN features," *Mathematics*, vol. 8, no. 4, p. 644, 2020.
- [7] J. Sun, X. Liu, W. Wan, J. Li, D. Zhao, and H. Zhang, "Video hashing based on appearance and attention features fusion via DBN," *Neurocomputing*, vol. 213, pp. 84–94, 2016.
- [8] E. Yu, J. Li, L. Wang, J. Zhang, W. Wan, and J. Sun, "Multi-class joint subspace learning for cross-modal retrieval," *Pattern Recognition Letters*, vol. 130, pp. 165–173, 2020.
- [9] F. Shang, H. Zhang, L. Zhu, and J. Sun, "Adversarial cross-modal retrieval based on dictionary learning," *Neurocomputing*, vol. 355, pp. 93–104, 2019.
- [10] W. Zhang and V. Rijmen, "Division cryptanalysis of block ciphers with a binary diffusion layer," *IET Information Security*, vol. 13, no. 2, pp. 87–95, 2019.
- [11] H. Wang, D. He, J. Shen, Z. Zheng, X. Yang, and M. H. Au, "Fuzzy matching and direct revocation: a new CP-ABE scheme from multilinear maps," *Soft Computing*, vol. 22, no. 7, pp. 2267–2274, 2018.
- [12] L. Zou, J. Sun, M. Gao, W. Wan, and B. B. Gupta, "A novel coverless information hiding method based on the average pixel value of the sub-images," *Multimedia Tools and Applications*, vol. 78, no. 7, pp. 7965–7980, 2019.
- [13] Z. Zhou, Q. M. J. Wu, Y. Yang, and X. Sun, "Region-level visual consistency verification for large-scale partial-duplicate image search," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 16, no. 2, pp. 1–25, 2020.
- [14] H. Liu, B. Xu, D. Lu, and G. Zhang, "A path planning approach for crowd evacuation in buildings based on improved artificial bee colony algorithm," *Applied Soft Computing*, vol. 68, pp. 360–376, 2018.
- [15] H. Liu, B. Liu, H. Zhang, L. Li, X. Qin, and G. Zhang, "Crowd evacuation simulation approach based on navigation knowledge and two-layer control mechanism," *Information Sciences*, vol. 436–437, pp. 247–267, 2018.
- [16] E. Yu, J. Sun, J. Li, X. Chang, X. H. Han, and A. G. Hauptmann, "Adaptive semi-supervised feature selection for cross-modal retrieval," *IEEE Transactions on Multimedia*, vol. 21, no. 5, pp. 1276–1288, 2019.
- [17] J. Liu and K. She, "A hybrid approach of DWT and DCT for rational dither modulation watermarking," *Circuits, Systems, and Signal Processing*, vol. 31, no. 2, pp. 797–811, 2012.
- [18] G. S. Kalra, R. Talwar, and H. Sadawarti, "Adaptive digital image watermarking for color images in frequency domain," *Multimedia Tools and Applications*, vol. 74, no. 17, pp. 6849–6869, 2015.
- [19] S. Masmoudi, M. Charfeddine, and C. Ben Amar, "A semi-fragile digital audio watermarking scheme for MP3-encoded signals using Huffman data," *Circuits, Systems, and Signal Processing*, vol. 39, no. 6, pp. 3019–3034, 2020.
- [20] L. Ma, D. Yu, G. Wei, J. Tian, and H. Lu, "Adaptive spread-transform dither modulation using a new perceptual model for color image watermarking," *IEICE Transactions on Information and Systems*, vol. E93-D, no. 4, pp. 843–857, 2010.
- [21] W. Wan, J. Liu, J. Sun, and D. Gao, "Improved logarithmic spread transform dither modulation using a robust perceptual model," *Multimedia Tools and Applications*, vol. 75, no. 21, pp. 13481–13502, 2016.
- [22] C. Wang, M. Xu, W. Wan et al., "Robust image watermarking via perceptual structural regularity-based JND model," *KSII Transactions on Internet and Information Systems*, vol. 13, no. 2, pp. 1080–1099, 2019.
- [23] H. Wang, S. Wang, T. Li, H. Yin, and L. Yu, "Surprise based JND estimation for images," *IEEE Signal Processing Letters*, vol. 14, no. 8, pp. 181–185, 2015.
- [24] X. Shang, Y. Wang, L. Luo, Y. Zuo, and Z. Zhang, "Fast mode decision for multiview video coding based on just noticeable distortion profile," *Circuits, Systems, and Signal Processing*, vol. 34, no. 1, pp. 301–320, 2015.
- [25] A. B. Watson, "DCT quantization matrices optimized for individual images," *Proceedings of SPIE*, vol. 1913, pp. 202–216, 1993.
- [26] Y. C. Chung, J. M. Wang, R. R. Bailey, S. W. Chen, and S. L. Chang, "A non-parametric blur measure based on edge analysis for image processing applications," *IEEE Conference on Cybernetics and Intelligent Systems*, pp. 356–360, 2004.
- [27] R. Ferzli and L. J. Karam, "A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB)," *IEEE Transactions on Image Processing*, vol. 18, no. 4, pp. 717–728, 2009.
- [28] Z. Zhang, Y. Liu, Z. Xiong, J. Li, and M. Zhang, "Focus and blurriness measure using reorganized DCT coefficients for an autofocus application," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 1, pp. 15–30, 2018.
- [29] S. Xing, F. Liu, Q. Wang, X. Zhao, and T. Li, "A hierarchical attention model for rating prediction by leveraging user and product reviews," *Neurocomputing*, vol. 332, pp. 417–427, 2019.
- [30] S. H. Bae and M. Kim, "A novel DCT-based JND model for luminance adaptation effect in DCT frequency," *IEEE Signal Processing Letters*, vol. 20, no. 9, pp. 893–896, 2013.
- [31] W. Wan, J. Wu, X. Xie, and G. Shi, "A novel just noticeable difference model via orientation regularity in DCT domain," *IEEE Access*, vol. 5, pp. 22953–22964, 2017.
- [32] Zhenyu Wei and K. N. Ngan, "Spatio-temporal just noticeable distortion profile for grey scale image/video in DCT domain," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 19, no. 3, pp. 337–346, 2009.
- [33] W. Wan, J. Wang, J. Li, J. Sun, H. Zhang, and J. Liu, "Hybrid JND model-guided watermarking method for screen content

- images,” *Multimedia Tools and Applications*, vol. 79, no. 7-8, pp. 4907–4930, 2020.
- [34] J. Zong, L. Meng, H. Zhang, and W. Wan, “JND-based multiple description image coding,” *KSII Transactions on Internet and Information Systems*, vol. 11, no. 8, pp. 3935–3949, 2017.
- [35] B. Chen and G. W. Wornell, “Quantization index modulation: a class of provably good methods for digital watermarking and information embedding,” *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1423–1443, 2001.
- [36] J. Wang and W. Wan, “A novel attention-guided JND model for improving robust image watermarking,” *Multimedia Tools and Applications*, vol. 79, no. 33-34, pp. 24057–24073, 2020.
- [37] W. Wan, J. Wang, J. Li et al., “Pattern complexity-based JND estimation for quantization watermarking,” *Pattern Recognition Letters*, vol. 130, pp. 157–164, 2020.
- [38] C. T. Vu, T. D. Phan, and D. M. Chandler, “A spectral and spatial measure of local perceived sharpness in natural images,” *IEEE Transactions on Image Processing*, vol. 21, no. 3, pp. 934–945, 2012.
- [39] E. C. Larson and D. M. Chandler, “Most apparent distortion: full-reference image quality assessment and the role of strategy,” *Journal of Electronic Imaging*, vol. 19, no. 1, p. 011006, 2010.
- [40] “LIVE Image Quality Assessment Database,” <http://live.ece.utexas.edu/research/quality/subjective.htm>.