IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 21, 2020, accepted September 3, 2020, date of publication September 8, 2020,
date of current version September 21, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3022652

JND-Guided Perceptually Color Image Watermarking in Spatial Domain

WENBO WAN[1], KAI ZHOU[1], KAI ZHANG[1], YANTONG ZHAN[1], AND JING LI[1,2], (Member, IEEE)
[1]School of Information Science and Engineering, Shandong Normal University, Jinan 250014, China
[2]School of Mechanical and Electrical Engineering, Shandong Management University, Jinan 250014, China

Corresponding author: Jing Li (lijingjdsun@hotmail.com)

This work was supported in part by the Natural Science Foundation of China under Grant 61601268, Grant 61803237, Grant 61901246, and Grant U1736122; in part by the China Postdoctoral Science Foundation under Grant 2019TQ0190 and Grant 2019M662432; in part by the Natural Science Foundation for Distinguished Young Scholars of Shandong Province under Grant JQ201718; and in part by the Shandong Provincial Key Research and Development Plan under Grant 2017CXGC1504.

ABSTRACT Watermarking is an effective solution for copyright protection and forensics tracking via hiding information into the image signal. Recently, spatial-embedding watermarking methods from different transform domain have been proposed rapidly and effectively protect the copyright of the color image. Compared to the individual frequency domain method, it has both advantages of spatial domain and frequency domain. Here, we proposed a novel spatial-embedding watermarking method based on an attended just noticeable difference (JND) model with color complexity, to achieve a good tradeoff between robustness and perceptual quality. In particular, at the spatial embedding level, the direct current (DC) coefficients are selected for embedding and the perceptual JND model is used to guide the amount of pixel modification to improve visual quality. Different from the previous JND model, we proposed an attended JND model that considers color complexity, which is more consistent of the human visual perception model. Compared with the other JND models, the proposed JND model is more suitable for the watermarking framework. Experimental results show that the proposed watermarking scheme has better performance than other existing watermarking schemes. This greatly benefits the practical implementations of the spatial-embedding watermarking methods.

INDEX TERMS Image watermarking, dc coefficient, color complexity, just noticeable difference, visual saliency.

I. INTRODUCTION

With the rapid development of the internet and multimedia technologies, the distribution of digital media (e.g. audio, image, video) becomes much easier. Thus, the protection of multimedia copyright is an urgent problem. Digital watermarking, hiding proprietary information into the original product, is an effective solution for multimedia copyright protection [1]–[5].

Digital watermarking, especially robust image watermarking, has attracted a lot of attention during the past decades. In general, robust watermarking is classified into two categories: spatial domain and frequency domain watermarking schemes [6]–[9]. The spatial domain watermarking schemes carry out watermark embedding by modifying the values of image pixels directly. On the other hand, if the watermark is embedded in the frequency domain, the pixel values are first transformed to the frequency coefficients and then the selected coefficients will be modified to embed the watermark. Commonly used transforms include discrete cosine transform (DCT) [10], discrete fourier transform (DFT) [11], discrete wavelet transform (DWT) [12] and singular value decomposition (SVD) [13], etc.

Currently, digital images are commonly stored and transmitted in JPEG format. Hence, most image watermarking methods are designed in DCT domain as DCT has a strong "energy compaction" property, capable of achieving high quality at high data compression ratios. In recent years, a large amount of DCT-based watermarking methods have been proposed. Roy and Pal [9] proposed a blind watermarking scheme that selects the middle band coefficients to embed watermark. But the visual quality of the watermarked image has serious distortion. Das *et al.* [10] presented a watermarking scheme based on the inter-block coefficient correlation

The associate editor coordinating the review of this manuscript and approving it for publication was Feng Shao.

in DCT. This scheme embedded watermark in DCT coefficients according to the different adjacent blocks. However, the robustness of watermarking scheme should be improved to satisfy a good resistance for noise attack and gaussian filter. Studies from Hang *et al.* [14] showed that, it is feasible to embed the watermark into direct current (DC) component of the DCT domain. Recently, a series of new methods [15]–[17] were proposed, which combine the advantages of spatial domain convenience and frequency domain robustness. These methods based on the combination of spatial embedding and DC quantization, where the pixel value in spatial domain was updated directly according to the DC modification from watermark embedding. So, the watermarking in frequency domain can distribute the energy of embedded data over the spatial domain. Though these existing methods have greatly improved the robustness performance of the watermarking method, it is worthy of noting that the modified pixel values with an uniform averaged sum for a fixed block. There is still a remarkable gap between visual distortion and the embedding strength. So, in this work, we focus on spatial domain embedding with the guideline from the perceptual quality measurement [18], [19].

Reliable robust watermarking methods aim to obtain as less distortion as possible, and acheieve a higher robustness performance. Consequently, human visual system (HVS), which can efficiently represent the visual distortions of an image, is usually designed for watermark embedding [20]. The visual threshold indicates that the maximum distortion in pixels or subbands that the HVS cannot perceive, which is called as just noticeable difference (JND). The JND model can be employed to achieve a better tradeoff between imperceptibility and robustness [21]–[26]. According to this assumption, Hu and Chang [24] proposed a DCT domain watermarking method which selected five low frequency DCT coefficients to embed watermark, and a weighting strategy based on JND model was developed to handle coefficient modifications. Wang *et al.* [26] proposed an enhanced JND model by formulating the orientation diversity as a factor to determine the contrast masking effect, and the JND model was used to control the watermark strength.

Most of existing perceptual watermarking algorithms regularly measure the degradations from the monochrome image [27]. However, the HVS presents obvious color differences measurement, with which the HVS focuses only on these color complex regions for detailed perception and withdraws the other regions. Besides, according to the visual cognition theory, attended regions play more important roles than the other unattended regions for visual perception, and degradations on attended regions have a larger influence on the watermarked images. Recently, the color measurement and the attention mechanism have already been taken into account for watermarking algorithms, in which the color and salient map is created to highlight these attended regions during watermark embedding.

In this work, we proposed a new color image watermarking in spatial domain based on an attended JND model with color complexity. The main contributions in this article are summarized as follows:

1) Herein, we present an innovative spatial watermarking method suitable for color images. The pixel values are directly modified with the visual guidance of the proposed JND estimation in the embedding procedure. The changed pixels, which are consistent with the HVS properties, are obtained.
2) A new attended JND model is designed with color complexity factor which considering structural information and applied to watermarking method. The proposed JND model have better robustness than existing JND models in the watermarking scheme.The robustness of the proposed method has increased while maintaining high transparency of the watermark in the image.
3) The proposed watermarking method has a tradeoff between invisibility and robustness. At the same time, the proposed watermarking algorithm is relatively robust against some common image processing (such as noise addition, JPEG compression, filtering, and volumetric attack) and geometrical distortions (such as cropping, scaling, and rotation). Experimental results show that the proposed scheme performs well under common signal processing operations.

The remainder of the paper is organized as follows: In Sectionn II, we will introduce the related work on DCT transform and traditional pixel JND model. Section III shows the attended pixel domain JND model with color complexity. A spatial perceptual-embedding watermarking scheme is described in Section IV. We present the experimental results and discussions in Section V, and the conclusion are given in Section VI.

II. RELATED WORK

A. DCT

DCT, a kind of compression kernel in JPEG standard, has been widely used in watermarking methods. In general, an image can be transformed from the spatial domain to frequency domain by 2D block DCT, which can also be inverted to the original image.

JPEG compression is obtained by DCT transform of image block, so we take an 8×8 image block as an example to illustrate. For k-th image block of image, $B^k(i,j)$ $(i = 0, 1, 2, \ldots, 7, j = 0, 1, 2, \ldots, 7)$, 2-D DCT is given as follows,

$$D^k(u,v) = \partial_u \partial_v \sum_{i=0}^{7} \sum_{j=0}^{7} B^k(i,j) \times \cos \frac{\pi (2i+1) u}{2 \times 8} \times \cos \frac{\pi (2j+1) v}{2 \times 8}, \quad (1)$$

where u and v are the horizontal and the vertical frequency $(u = 0, 1, 2, \ldots, 7, v = 0, 1, 2, \ldots, 7)$, and $D^k(u,v)$ are DCT coefficients of image block $B^k(i,j)$,

$$\partial_u = \begin{cases} \sqrt{1/8} & \text{if } u = 0 \\ \sqrt{2/8} & \text{if } 1 \le u < 7, \end{cases} \quad (2)$$

$$\partial_v = \begin{cases} \sqrt{1/8} & \text{if } v = 0 \\ \sqrt{2/8} & \text{if } 1 \le v < 7. \end{cases} \quad (3)$$

Based on Eq. (1), DC coefficient X^k can be given as follows,

$$X^k = D(0,0) = \frac{1}{\sqrt{64}} \times \sum_{i=0}^{7}\sum_{j=0}^{7} B^k(i,j). \quad (4)$$

It is shown that the DC coefficient can be obtained directly in spatial domain instead of DCT transform.

And inverse 2-D DCT of the k-th image block $B^k(i,j)$ is described as following,

$$B^k(i,j) = \sum_{i=0}^{7}\sum_{j=0}^{7} \partial_u \partial_v D^k(u,v) \times \cos\frac{\pi(2i+1)u}{2\times 8} \times \cos\frac{\pi(2j+1)v}{2\times 8}. \quad (5)$$

Based on Eq. (4) and Eq. (5), the inverse DCT of the k-th image block in Eq. (4) will be rewritten by

$$B^k(i,j) = \frac{X^k}{N} + A^k(i,j), \quad (6)$$

where N is the dimension of the k-th block (is 8 in this case), $A^k(i,j)$ represents the reconstructed image block from all the set of AC coefficients.

B. SPATIAL-UNIFORM EMBEDDING-BASED WATERMARKING

Recently, Su and Chen [15] proposed a robust color image watermarking methods based on DC coefficient, wherein the algorithm is executed in spatial domain, and it not only has simple and quick performance of the spatial domain but also has high robustness of DCT domain. In Su's method [15], when embedding watermark information into DC coefficient of k-th the image block, the modification of the image block's DC coefficient can be defined as E^k, and the modifited DC coefficient with modification E^k can be obtained by Eq. (7),

$$X_w^k = X^k + E^k, \quad (7)$$

where X^k is DC coefficient of the k-th block, X_w^k is the watermarked DC coefficient, and E^k is the modification. The embedding processing can be found in Ref. [15]. According to Eq. (6), the watermark bits are embedded into X^k, and the uniform changing with E^k/N for the value of each pixel is implemented. Obviously, it can be described as,

$$\begin{aligned} B^{k*}(i,j) &= \frac{X_w^k}{N} + A^k(i,j) \\ &= \frac{X^k + E^k}{N} + A^k(i,j) \\ &= B^k(i,j) + \frac{E^k}{N}, \end{aligned} \quad (8)$$

where E^k/N is the uniform changing for each pixel in the spatial domain, $B^{k*}(i,j)$ is pixel of the watermarked image block's in spatial domain. It can be shown that the pixel values are modified with a uniform change in Su' method [15]. However, on the other hand, it is obviously that HVS cannot perceive all the changes in the pixels due to its underlying physiological and psychological mechanism, so some perceptual redundancies exist equally. Firstly, it ensures that only the important visually information is protected. Secondly, better perceptual performance can be achieved by discarding perceptually unnecessary information. So how to build perceptual changes in watermarking becomes a challenging research task.

C. JND MODELING

The JND refers to the maximum distortion which cannot be perceived by the human eyes. Inspired by the research of cognitive scientists, numerous JND estimation models have been proposed. The JND models are divided into two categories by calculating the JND threshold domain. One category is the pixel-wise domain which can directly calculate the JND threshold for image pixel [18], [19]. Other category is the subband-domain, for example, DCT domain [28]. The subband-domain JND model first transfer the image into subband domain and the JND threshold can be calculated on each subband. Pixel-wise domain JND model does not need the transformation to subbands, it is more convenient than subband-domain JND to provide a JND threshold. JND gives us a promising way to guide the pixels changes.

In other word, JND refers to the minimum visibility threshold when visual content can be distinguished. JND can be determined by subjective observation and testing, but it will cost a lot of human and material resources. Therefore, through psychological research, researchers can model visual redundancy through various external environments and internal generation mechanism of brain information, and get objective JND model. Chou and Li [29] proposed pixel-wise JND profile, in which luminance masking and contrast masking functions are proposed in the spatial domain. Wu *et al.* [18] proposed a JND model by considering the regular pattern complexity and contrast masking as a factor for spatial masking (SM) effect. Put the k-th image block as an example, thus, the traditional pixel JND model in k-th image block can be expressed by the nonlinear additivity model for masking (NAMM) [18] model,

$$JND^k(i,j) = L_A^k(i,j) + S_M^k(i,j) - 0.3 \times \left\{ L_A^k(i,j), S_M^k(i,j) \right\}, \quad (9)$$

where $L_A^k(i,j)$ is luminance adaptation for k-th image block, $S_M^k(i,j)$ is the spatial masking effect for k-th image block.

1) LUMINANCE ADAPTATION

As we all know, the HVS has different sensitivity with different luminance, and luminance adaptation needs to be considered. Therefore, the visual threshold based the luminance

adaptation model [29] can be expressed by,

$$L_A^k(i,j) = \begin{cases} 17 \times \left(1 - \sqrt{\frac{I_b^k(i,j)}{127}}\right), & \text{if } I_b^k(i,j) < 127 \\ \frac{3 \times \left(I_b^k(i,j) - 127\right)}{128} + 3, & \text{if } I_b^k(i,j) \geq 127, \end{cases} \tag{10}$$

where $I_b^k(i,j)$ is the background luminance, which is calculated as the mean luminance value of a surrounding region that (i,j) located.

2) SPATIAL MASKING

In [18], Wu *et al.* proposed a pixel-based JND model based on pattern complexity. It is well known that the JND model can be impacted by spatial marking effect. Through different masking effects, a total spatial masking effect of the k-th image block can be given as follow,

$$S_M^k(i,j) = \max\left(P_{CM}^k(i,j), C_M^k(i,j)\right), \tag{11}$$

where the $S_M^k(i,j)$ is the total spatial masking effect, $P_{CM}^k(i,j)$ is the pattern masking effect of the k-th image block, $C_M^k(i,j)$ is the contrast masking effect of the k-th image block. The visual pattern masking effect can be obtained by pattern complexity function and luminance contrast function,

$$P_{CM} = f(P_C) \cdot f(L_C), \tag{12}$$

where P_{CM} is the pattern masking effect, $f(P_C)$ is the pattern complexity function, $f(L_C)$ is the luminance contrast function. Thus, $f(P_C)$can be modeled as,

$$f(P_C) = a1 \cdot \frac{P_C^{a2}}{P_C^2 + a3}, \tag{13}$$

where three constants are set as $a1 = 0.8$, $a2 = 2.7$, $a3 = 0.01$ in [30], and the pattern complexity P_C is calculated by its histogram in [18]. The function $f(L_C)$ follows logarithmic form,

$$f(L_C) = \log_2(1 + L_C), \tag{14}$$

where L_C is luminance contrast which can be calculated by the gradient magnitude.

$$L_C^k(i,j) = \sqrt{G_v^2 + G_h^2}, \tag{15}$$

The G_v and G_h of the image block $B^k(i,j)$ can be calculated by the Prewitt edge kernels filters,

$$G_v^k(i,j) = B^k(i,j) * K_v, G_h^k(i,j) = B^k(i,j) * K_h, \tag{16}$$

where $*$ denotes the convolution operation, K_v and K_h are the Prewitt edge kernels. The contrast masking function C_M is fitted with the data of visibility thresholds from masking experiment on luminance contrast [18], [30], and it is defined as follows,

$$C_M = 0.115 \times \frac{a4 \cdot L_C^{2.4}}{L_C^2 + a5}, \tag{17}$$

where the two parameters $a4$, $a5$ are set to 16, 26 respectively, as discussed in [18].

The JND has been considered a suitable solution for controlling the watermark strength and generating robust watermarking schemes with distortions that are below the sensitivity threshold. However, JND assumes the same attention level for all image regions, which does not reflect the behavior of an observer. Visual saliency (VS) is an important feature of human visual system and the study of visual saliency was reported in early [31]. Human beings can show visual gaze, which is to keep the visual gaze on a single place. Inspired by this visual perception phenomenon, some visual saliency models were proposed. The visual saliency detection model simulates the perception scene of human visual system, and generally focuses on some areas of human interest, which affects JND threshold to a certain extent.

As we all know, color plays an extremely important role in our understanding of the world. Reference [32] points out that same color in two blocks can be recognized as the same color only when the color and direction structure information of its adjacent areas are identical. Therefore, there is a certain masking effect between adjacent regions. For color images, some existing JND models do not consider the influence of chrominance information, but simply use the luminance information to evaluate the JND threshold, which is a disadvantage. Therefore, for color image, JND threshold should consider the influence of color information.

III. PROPOSED JND MODEL

There are many JND models considering various factors that have been proposed in recent years. However, many existing JND models only consider the factors of the image itself, such as background brightness, texture region, structure information, and so on. Although some JND models consider complex features, they often only consider the influence of single complex features on JND. Therefore, we propose an attended JND model considering color complexity.

As we all know, the human sensitivity will be influenced by color change. The consistent color difference of color perception covers a variety of visual information, which plays an important role in the analysis and understanding of image content. The research of cognitive science shows that HVS can extract repetition patterns of visual content representation well. We regard pattern complexity as an important factor that affects JND threshold: in a regular pattern, interaction is relatively simple, masking effect is limited; in an irregular pattern, interaction is complex, masking effect is strong. For a given image, the VS map calculated by the appropriate VS model can reflect the significance of the local region to HVS, and the JND threshold corresponding to the region with large significance characteristics will be reduced accordingly.

A. COLOR COMPLEXITY

Color, an important part of HVS, is the basis of the perception of the world. The differences between colors will lead to different visual perceptions and different visual masking. In recent applications, the watermark information is embedded into the color image. Su *et al.* [16] used approximate Schur decomposition to each channel of RGB three channels, and each channel selected coefficient to embed watermark. Jia employed the left singular matrix of SVD to embed watermark into each color channel [33]. Above watermarking algorithms only embedding watermark into the individual channel of RGB, without considering the correlation between RGB channels, the robustness is limited compared with watermarking algorithms for grayscale images. Due to there are strong correlations between RGB channels, some watermarking methods were proposed to employ this correlation. Zhang *et al.* [34] used Tucker decomposition to color image, and employed the correlations between chromaticity to embed watermark information, which improved robustness.

And a series of watermarking schemes select Y channel in the YCbCr color space (Y channel denotes the luminance component, Cb and Cr channels are denoted to represent the color information) as the embedding channel for color watermarking method. Su *et al.* [6] directly calculated the DC coefficients of Y channel without DCT transform, and embedded watermark information by modified the each pixel of the block. Koju and Joshi [35] embedded watermark information to the Y channel, which has good robustness and imperceptibility. However, the visual perception of color images depends not only on luminance but also on chromaticity. Different from the previous watermarking algorithm that only selects the Y channel of as the embedding channel, we also used a color measurement method to measure color complexity. The transformation between RGB and YCbCr is defined as follows:

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0.2990 & 0.5870 & 0.1140 \\ -0.169 & -0.331 & 0.5000 \\ 0.5000 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix}, \quad (18)$$

In the psychophysical JND modeling, considering the component of color in YCbCr, using Cb channel Cb to get a color masking factor to better adjust JND model. It can be seen that Cb can reflect the change of color image [26]. Generally, since adjacent pixels exhibit regular structure, the structure is constructed by using the empirical distributions of pairwise products of neighboring values. Thus current pixel with four structure features among horizontal, vertical, main diagonal and sub diagonal directions can be obtained. It can be clearly seen in Figure 1.

$$S_H(i,j) = \prod_{\lambda=-1}^{1} Cb(i, j+\lambda), \quad (19)$$

$$S_V(i,j) = \prod_{\lambda=-1}^{1} Cb(i+\lambda, j), \quad (20)$$

Cb (i-1,j-1)	Cb (i-1,j)	Cb (i-1,j+1)
Cb (i,j-1)	Cb (i,j)	Cb (i,j+1)
Cb (i+1,j-1)	Cb (i+1,j)	Cb (i+1,j+1)

FIGURE 1. The structure information in Cb channel are computed along four orientations-horizontal, vertical, main diagonal, and sub diagonal in a 3 × 3 window.

$$S_{D^+}(i,j) = \prod_{\lambda=-1}^{1} Cb(i+\lambda, j+\lambda), \quad (21)$$

$$S_{D^-}(i,j) = \prod_{\lambda=-1}^{1} Cb(i+\lambda, j-\lambda), \quad (22)$$

where $S_H(i,j)$, $S_V(i,j)$, $S_{D^+}(i,j)$, $S_{D^-}(i,j)$ respectively represents horizontal, vertical, main diagonal and sub diagonal direction. λ represents the value selected for calculation in a 3 × 3 window. The each pixel color complexity C_{bs} from extracted each pixel with four structure images $[S_H, S_V, S_{D^+}, S_{D^-}]$, and it can be given by

$$C_{bs} = \max(S_H, S_V, S_{D^+}, S_{D^-}), \quad (23)$$

where C_{bs} is a max direction feature of color complexity from four structure images. max $(\cdot)$ means the maximal direction value of four structural feature images.

B. SPATIAL MASKING EFFECT WITH COLOR COMPLEXITY

In this part, we proposed a new spatial masking function which consider the masking effect of different visual contents.

In the color image, the color of the adjacent position with the same direction structure information is always consistent. To judge whether two image blocks have the same color, we should not only look at the color difference between the two blocks, but also consider the structure information of the two blocks. In other words, image blocks with the same color will show different color differences in the visual system under different structural information. And the structural information usually be used to calculate the pattern complexity. The experimental results from [36] showed that most observers tend to focus on some regions with the same pattern complexity in image, and the judgment of regional pattern complexity is generally determined by the color difference of these regions. In a word, color difference has a great influence on pattern complexity in visual content. Therefore, the chromaticity difference of color image will have a certain impact on the visual pattern complexity to a certain extent. The improved visual pattern masking effect can be obtained by pattern complexity function, luminance contrast function

and color complexity function,

$$P^*_{CM} = f\left(P_C\right) \cdot f\left(L_C\right) \cdot f\left(C_C\right), \tag{24}$$

where $f\left(C_C\right)$ is the color complexity function which can be established with the linear transducer as,

$$f\left(C_C\right) = 1 + \left(\coprod C_{bs} - 0.05\right) \cdot 0.12, \tag{25}$$

where $\coprod$ is the normalization processing. And an improved spatial masking which consider color difference will be calculated,

$$S^{k*}_M\left(i,j\right) = \max\left(P^{k*}_{CM}\left(i,j\right), C^k_M\left(i,j\right)\right), \tag{26}$$

C. VISUAL SALIENCY MODULATION

If an area has certain characteristics that make it stand out from the surrounding area and attract attention, it is called a VS area [37], [38]. The existing VS model can generate a saliency map, which can predict the saliency region in a given image. The JND threshold value in the region with a more significant image will be reduced accordingly, so visual significance can adjust JND threshold value well to avoid the problem of the too high threshold value. As a simple but effective saliency estimator, the classical saliency model for visual saliency-based index (VSI) [39] is employed. Given an color image block, the saliency model can be obtained by frequency prior, location prior and color prior in Zhang *et al.* [39].

$$S^k_{i,j} = S^k_F\left(i,j\right) \cdot S^k_D\left(i,j\right) \cdot S^k_C\left(i,j\right), \tag{27}$$

where (i,j) is the location of the k-th image block, $S^k_F\left(i,j\right)$ is the frequency prior, $S^k_D\left(i,j\right)$ is the location prior and $S^k_C\left(i,j\right)$ is the color prior. Complete calculation about saliency detection can be found in [39].

Considering the above complexity factors, we propose a JND model in pixel domain based on the multiple complexity features which impact the visual perception. The JND threshold for each pixel of the k-th image block is calculated using the NAMM [18] as

$$\begin{aligned} JND^{k*}\left(i,j,\beta\right) = &\left\{L^k_A\left(i,j\right) + S^{k*}_M\left(i,j\right) - 0.3\right. \\ &\left.\times \min\left[L^k_A\left(i,j\right), S^{k*}_M\left(i,j\right)\right]\right\} \cdot \eta\left(S^k_{i,j}, \beta\right), \end{aligned} \tag{28}$$

where (i,j) is the location of k-th image block, $L^k_A\left(i,j\right)$ is the luminance adaptation, $S^{k*}_M\left(i,j\right)$ is the spatial masking effect with color effect, and the $\eta(S^k_{i,j}, \beta)$ describes the salience modulation as follows,

$$\eta(S^k_{i,j}, \beta) = \beta 1\left(\frac{1}{2} - \frac{1}{1+\exp(\beta 2(S^k_{i,j} - \beta 3))}\right) + \beta 4. \tag{29}$$

where $S^k_{i,j}$ is a sigmoid function which can be defined with four coefficients in a vector $\beta = [\beta 1, \beta 2, \beta 3, \beta 4]$. In this section, due to multiple complexity features will impact the visual perception, we proposed an improved pixel JND model considering some visual features. Then, to avoid the spatial distortion with uniform changes caused by embedding watermark information into DC coefficients, the proposed JND model in Eq. (28), which reveals the visibility thresholds of the HVS for spatial pixels, can be directly used to guide the watermark embedding for each pixel.

IV. PROPOSED SCHEME

A. JND-GUIDED EMBEDDING IN SPATIAL DOMAIN

In this section, the DC coefficient was chosen as the watermark embedding position just like [14]–[16]. When the watermark information is embedded into the DC coefficients, the corresponding pixel values in the spatial domain will change as Eq. (8). Reference [15] make the pixel value modification amount of each block the same, which will cause some visual distortion in spatial domain. Each pixel in the spatial domain has the same amount of modification, and the texture, smoothness, edge features of the image itself are not well-considered, nor are they related to HVS visual perception features, which inevitably leads to distortion in the spatial domain. Therefore, it is necessary to use the JND model which accords with the characteristics of human visual perception to guide the amount of modification of each pixel and make it meet the perception characteristics of HVS. The flow of watermarking scheme can be given in Figure 2.

In our work, the JND model of the pixel domain is calculated directly from each pixel of the visual content, and JND model guides the amount of modification of each pixel in each image block,

$$B^{k*}(i,j) = B^k(i,j) + \frac{JND^{k*}(i,j)}{\sum\sum\left[JND^{k*}(i,j)\right]} \cdot (N \cdot E^k), \tag{30}$$

where $B^{k*}(i,j)$ is pixel of the watermarked image block in spatial after embedding watermark. (i,j) are the spatial indexes ($i = 0,1,2,\ldots,7, j = 0,1,2,\ldots,7$) in a block. E^k, which can be caululated by Eq. (32), is the modification of DC coefficient. $N \cdot E^k$ is the total modification of the k-th block in the spatial domain when the watermark information is embedded in the DC coefficient, $\sum\sum\left[JND^{k*}(i,j)\right]$ is the total JND threshold of the k-th block. Therefore, the JND model in the pixel domain is used to guide the modification amount of each pixel so that the watermarked image can better meet the characteristics of human eyes while getting a better robustness performance.

B. WATERMARK EMBEDDING

For a RGB color image, we select the Y channel to embed watermark information. Different from [15], we use dither modulation (DM) as embedding algorithm. Figure 2 gives the flow of the watermarking scheme. Firstly, the RGB color image I is converted into YCbCr color space by Eq. (18). The Y channel is divided into non-overlapped blocks with size 8×8. For k-th block, the DC coefficient can be obtained by Eq. (4), and the k-th image block's DC coefficient is then used as the host to embed the watermark bit m by dither

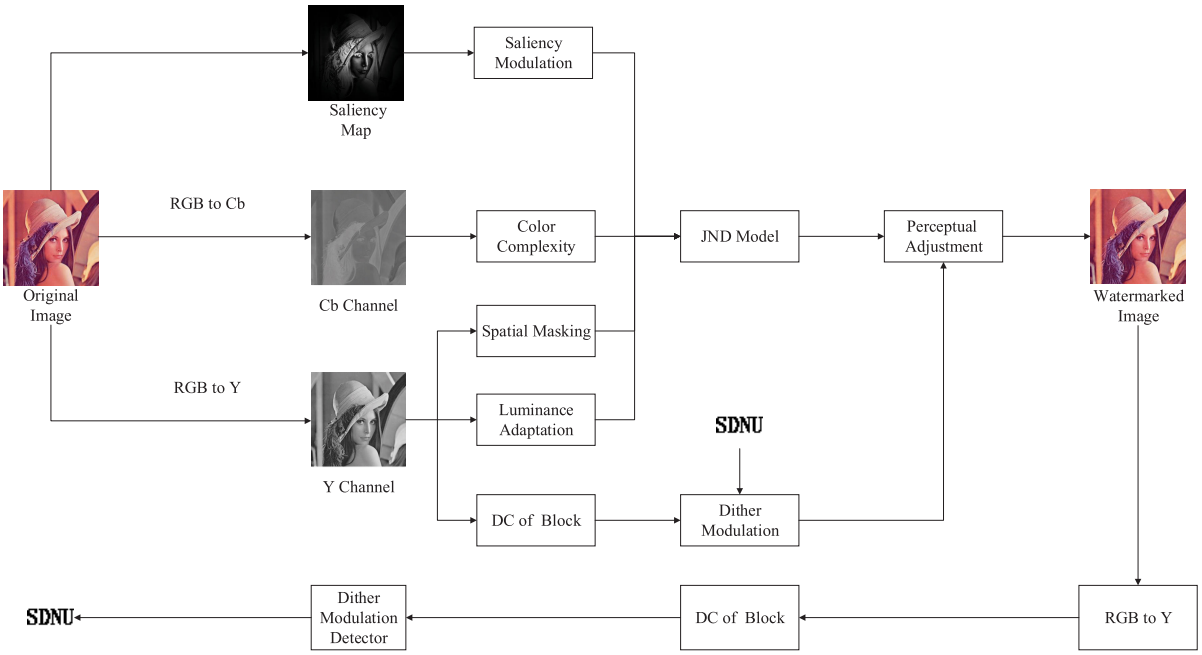


FIGURE 2. Flow-chart of the proposed watermarking algorithm.

modulation (DM) as follows:

$$X_w^k = Q\left(X^k, \Delta, m, d_m\right) = \Delta \cdot round\left(\frac{X^k + d_m}{\Delta}\right) - d_m, \quad m \in \{0, 1\}, \tag{31}$$

where X^k is the DC coefficient of the k-th image block, Δ is the quantization step, d_m is the dither signal corresponding to the watermark bit m. The watermarked signal X_w^k is obtained by the quantized data. Therefore, the modification E_k in Eq. (30) is obtained,

$$E^k = X_w^k - X^k. \tag{32}$$

The sum modification $N \cdot E^k$ for the k-th block can distribute the energy of embedded data over the spatial block. Consequently, $N \cdot E^k$ can be distributed over all pixels in the spatial block with the guidance of the perceptual JND in Eq. (30).

Repeating the same operation for the non-overlapped blocks, the watermark information is embedded in Y channel. Then YCbCr watermarked image transforms to RGB color watermarked image I_w. The main steps of watermark embedding scheme can be described as Algorithm 1 showed.

C. WATERMARK EXTRACTION

Firstly, the RGB color watermarked image I'_w, which is obtained after the image transmission with I_w, is converted into YCbCr image, and the watermarked channel Y is divided into non-overlapped 8×8 pixels blocks. In the extraction process, for k-th block, the DC coefficient $X_w^{k'}$ can be obtained by Eq. (4). Then, the watermark can be detected from $X_w^{k'}$ according to the minimum distance as follows:

$$m' = \underset{m \in \{0,1\}}{argmin} \left| X_w^{k'} - Q\left(X_w^{k'}, \Delta, m, d_m\right) \right|, \tag{33}$$

where m' is the extracted watermark information of the block, Δ is the quantization step, d_m is the dither signal corresponding to the watermark bit m. Repeating the same operation for the non-overlapped blocks, the watermark is reconstructed. The main steps of watermark extracting scheme can be described as Algorithm 2 showed.

V. EXPERIMENTS AND RESULTS ANALYSIS

In this section, we show and discuss the experiment results. To prove the goodness of our proposed watermarking scheme, we perform 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.

The images from databases CVG-UGR [40], are used as the host images. Six 24-bit 512×512 color images, as shown in Figure 3, are tested and displayed in this experiment. In this article, original watermark is the binary image with size 64×64, as shown in Figure 4.

For evaluating the imperceptibility of the proposed method, the peak signal-to-noise ratio (PSNR) and the visual saliency-based index (VSI) are utilized as the performance metrics. PSNR in Eqs. (34) is adapted to measure the similarity between the host image and watermarked image,

$$PSNR = 10 \log_{10}\left(\frac{255^2}{MSE}\right), \tag{34}$$

where MSE is the mean square error for color original image and watermarked image.

Algorithm 1 Watermark Embedding

Input: Host image (I); watermark (m)
Output: Watermarked image (I_w)
Begin
Step 1: The RGB color image is converted into YCbCr color space. The Y channel is regard as the watermark embedding channel, and Cb channel can be select as a reference image for watermarking JND processing;
Step 2: For Cb channel, structural information with different directions can be calculated by Eqs. (19) (20) (21)and (22);
Step 3: Divide the final structural feature map into 8×8 non-overlapped block. Calculate the color complexity of each 8×8 block for JND profile by Eq. (23);
Step 4: Divide the Y channel image into 8×8 non-overlapped blocks;
Step 5: The quantization step Δ can be set to control image quality;
Step 6: One bit of binary watermark m is embedded into an image block, according to the rules as follows:
for *each block* **do**
1. Estimate the perceptual JND factors: the luminance adaptation, spatial masking (color complexity masking, pattern complexity masking and contrast masking), and saliency model by Eqs. (10) (26), and (27), respectively;
2. Obtain final JND value of each block by Eq. (28);
3. Obtain the DC coefficient by Eq. (4). And one bit of the watermark message m is embedded into the DC coefficient by Eqs. (31) and (32);
4. Generate the modified block B^*;

end
Step 7: Generate the watermarked image Y by collecting all the modified blocks B^*;
Step 8: Generate the watermarked color image by concatenating the modified Y with Cb and Cr image channel and then convert the color space from YCbCr to RGB;
Return Watermarked image I_w;

Algorithm 2 Watermark Extraction

Input: Watermarked image (I'_w)
Output: Watermark message (m')
Begin
Step 1: Transform the color image from RGB color space to YCbCr color space. Select the channel Y as the main host image;
Step 2: Divide the host image into 8×8 non-overlapped blocks;
for *each block* **do**
1. Obtain the DC coefficient by Eq. (4);
2. And one bit of the watermark message is extracted by Eq. (33);

$$m' = \underset{m \in \{0,1\}}{argmin} \left| X_w^{k'} - Q\left(X_w^{k'}, \Delta, m, d_m\right) \right|$$

end
Return Watermark message m';

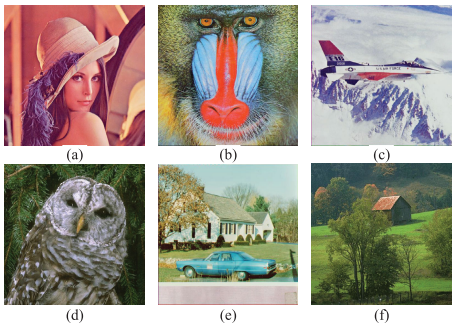


FIGURE 3. The color host images: (a) Lena, (b) Baboon, (c) Avion, (d) Bardowl, (e) House, (f) Barnfall.

FIGURE 4. The binary watermark.

As an image quality assessment metric, VSI has good visual performance to measure the image quality between original image and distortion image. Thus, we used VSI for watermarked image and original image.

$$VSI = \frac{\sum T\left(\max\left(S_1, S_2\right)\right)}{\sum\left(\max\left(S_1, S_2\right)\right)}, \tag{35}$$

where S_1, S_2 represent the VS map extracted from original images I and watermarked image I'_W; T is the local similarity of I and I'_W, as follows,

$$T = T_{VS}\left(x, y\right) \cdot \left[T_G\left(x, y\right)\right]^{0.4} \cdot \left[T_C\left(x, y\right)\right]^{0.02}, \tag{36}$$

where $T_{VS}\left(x, y\right)$ is the similarity map of saliency images of I and I'_W, $T_G\left(x, y\right)$ is the similarity map of gradient images of I and I'_W, $T_C\left(x, y\right)$ is the similarity map of chrominance components images of I and I'_W, as described in [39].

As an evaluation criterion, the bit error rate (BER) estimates the error rate of the watermark. Therefore, BER can be utilized for evaluating the robustness performance. The BER closed to 0 proves that the watermarking algorithm has better robustness performance, and the equation is as follows:

$$BER = \frac{\sum\sum m \oplus m'}{Area}, \tag{37}$$

where m is the original watermark and m' is the extracted watermark, $Area$ is the size of the watermark image (64×64).

A. EVALUATION OF PROPOSED JND MODEL

We verify the effectiveness by injecting noise into the pixels based on the corresponding JND value

$$\hat{I}\left(x, y, \beta\right) = I\left(x, y\right) + \delta \cdot rand\left(x\right) \cdot JND^*\left(x, y, \beta\right), \tag{38}$$

where $\hat{I}(x, y, \beta)$ is the JND noise contaminated image. $I(x, y)$ is the original image, δ regulates the energy of JND noise, which makes the same energy for different JND models, and rand(x) takes random value of $+1$ or -1.

1) PARAMETERS SETTING

Based on Eqs. (34), (35) and (38), the four coefficients in Eq. (39) can be obtained by a cost function in [41], and δ is set to 1.

$$Q\left(\hat{I}(x, y, \beta) | I(x, y,), S_{x,y}\right) = \frac{PSNR\left(\hat{I}(x, y, \beta) | I(x, y)\right)}{\gamma VSI\left(\hat{I}(x, y, \beta) | I(x, y), S_{x,y}\right)}, \quad (39)$$

$$\beta^* = \underset{\beta}{argmin} Q\left(\hat{I}(x, y, \beta) | I(x, y), S_{x,y}\right), \quad (40)$$

where $I(x, y)$ is the original image, and $\hat{I}(x, y, \beta)$ is the noise image which be obtained by Eq. (38). $PSNR\left(\hat{I}(x, y, \beta) | I(x, y)\right)$ is the PSRN value of $\hat{I}(x, y, \beta)$ and $I(x, y)$. $VSI\left(\hat{I}(x, y, \beta) | I(x, y)\right)$ is the VSI value of $\hat{I}(x, y, \beta)$ and $I(x, y)$. γ is a regularization term, which is set to 30 here [41]. From the above formula, we can see that when the threshold value of JND increases, the value of PSNR will decrease, the value of VSI will increase, and the image quality of distorted image will improve. Therefore, to get a better visual quality, the distorted image will get a minimum cost function and a maximum JND threshold by adjusting β. Through the experiment, we get the optimal value β, $\beta = [0.561, 4.61, 0.62, 1.06]$. In 184 runs out of 200, the cost function values we get are equal to the minimum cost function. That is to say, the optimization procedure yielded the same optimum point in 92% of the conducted runs.

2) VISUAL DISTORTION EVALUTION

Since the JND threshold reveals the visibility of the HVS, it can be used for guidance the noise for image. It is well known that the better JND model will gives better perceptual quality under same noise. In this section, we compare the proposed JND model with three pixel domain JND models, which are Wu *et al.*'s model (Wu 2017) [18], Wu *et al.*'s model(Wu 2013) [42], and Jakhetiya *et al.*'s (Jakhetiya 2018) [19].

With the help of Eq. (37), the same level of noise (with PSNR=35 dB) is injected into the testing images, and VSI is adopted to reflect the visual quality of the distorted image obtained by injecting noise. And the comparison results can be given by Figure 5, though the proposed JND model cannot has better visual quality in 'Lena' than Refs. [18] and [19], our model still has better perceptual quality than Ref. [42]. It can be seen that the proposed JND model has good visual quality.

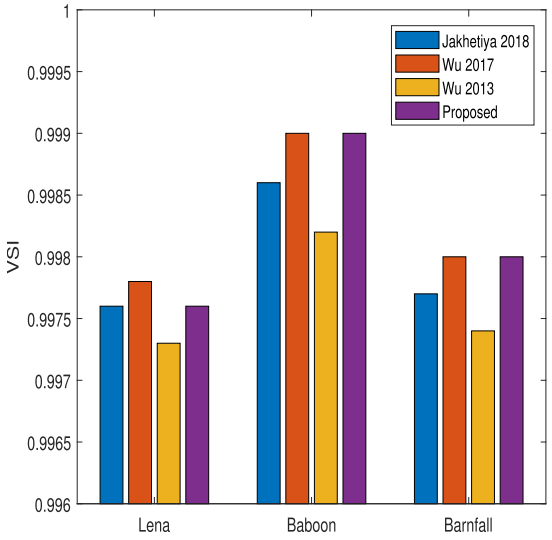


FIGURE 5. VSI comparison of different JND models within PSNR = 35dB.

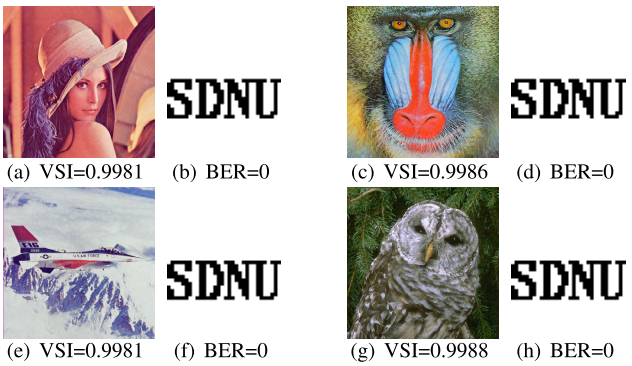
3) EVALUATION OF DIFFERENT JND MODELS FOR WATERMARKING METHODS

In this article, we use the JND model in the watermarking method. The JND model has adopted to guide the modification of the pixel when embedding the watermark, which can fit the HVS. In this section, the different JND models [18] (Wu 2017) and [19] (Jakhetiya 2018), which are used in watermarking scheme, can be compared to the robustness performance. The PSNR regulates to 42dB between the original image and watermarked image, which cannot be observed for HVS. Testing images are standard color images with a dimension of 512×512 as shown in Figure 3. And the watermark information is embedded into Y channel. A binary watermark with size 64×64 is used to embed into the cover images as shown in Figure 4.

To compare the performance of the proposed JND and the other pixel domain JND models within the watermarking framework, different kinds of attacks such as Gaussian Noise (GN) with mean zero and different variance, JPEG compression, where the JPEG quality factor varies from 30 to 40, and Salt & Pepper noise (SPN) with different factors were used to evaluate the robustness of the proposed JND model. Table 1 shows the average BER values of watermarked images attacked by Gaussian Noise, JPEG compression attacks and Salt & Pepper noise. When the watermarked images are attacked by Gaussian Noise with variance higher than 0.1%, the BER of our JND model is significantly lower than the other two JND models. This indicates that the proposed JND model performed much better than others. As shown in Table 1, performance emerges in the three JND models within the watermarking algorithms show about JPEG compression attacks. When the compression quality is 30, the average BER value of the proposed model is 1.7% lower than the model of Ref. [18], and the average BER is 2.7% lower than the Ref. [19]. When the watermarked

TABLE 1. BER comparison with PSNR = 42 dB under different image attacks.

Attack	Lena			Baboon			Avion		
	[18]	[19]	Proposed	[18]	[19]	Proposed	[18]	[19]	Proposed
GN(0.0010)	**0.0115**	0.0166	0.0134	0.0225	0.0344	**0.0188**	0.0184	0.0137	**0.0181**
GN(0.0015)	0.0344	0.0413	**0.0340**	0.0576	0.0828	**0.0437**	0.0439	0.0491	**0.0349**
JPEG(30)	0.0530	0.0532	**0.0391**	0.0769	0.1199	**0.0605**	0.0686	0.0547	**0.0459**
JPEG(40)	**0.0027**	0.0042	0.0075	0.0073	0.0068	**0.0034**	**0.0049**	0.0056	0.0061
SPN(0.0010)	0.0144	0.0125	**0.0090**	0.0146	0.0198	**0.0090**	0.0217	0.0205	**0.0117**
SPN(0.0015)	0.0200	0.0178	**0.0142**	0.0205	0.0264	**0.0139**	0.0293	0.0247	**0.0166**



(a) VSI=0.9981 (b) BER=0 (c) VSI=0.9986 (d) BER=0

(e) VSI=0.9981 (f) BER=0 (g) VSI=0.9988 (h) BER=0

FIGURE 6. The watermarked image's VSI, BER values without attacks. (a), (c), (e) and (g) is the watermarked image with VSI, and (b), (d), (f) and (h) is extrated watermark without attack.

images are contaminated by Salt & Pepper noise attacks with different factors, the average BER values of three JND models are presented in Table 1. It can be clearly seen that the proposed model always has the lowest BER for different Salt & Pepper noise intensities. Above all, the watermarking frame based on the proposed JND model has excellent robustness performance.

B. IMPERCEPTIBILITY TEST

In this part, to test the invisibility of the method, we use subjective tests. PSNR and VSI, which are consistent with HVS, are the object tests for watermarked images. Because the visual quality of watermarked image can be changed by the steps, therefore, the PSNR value between the watermarked image and the original image is fixed at 42dB. And the VSI and BER values can be obtained from our proposed method in Figure 6.

As we all know, the closer the VSI value to 1, the better the image quality. When VSI = 0.9920, the difference between the watermarked image and the original image cannot be seen by the human eyes. Figure 6, the VSI values from our method are closer to 1, which has better visual quality. Moreover, it can be seen from the BER, the watermark can be entirely extracted under without any attacks by our method.

C. ROBUSTNESS WITHIN INDIVIDUAL QUANTIZATION-BASED WATERMARKING TECHNIQUES

It is clear that the proposed image watermarking method has good visual perceptual quality with HVS in Figure 6. Moreover, robustness is an important metric for the watermarking scheme. In this section, in order to show the robustness of the proposed method, some basic attacks are selected to test the robustness, and we compared it with other quantization-based methods [21], [22], [43], [44] and [45]. To ensure the fairness of the experiment, the watermark is embedded into the Y channel of the color image, and we adjust the quantization step size so that the PSNR = 42dB of the watermarked image. In order to better demonstrate the robustness of the proposed method, we also show the average BER values of the testing watermarked images under some common attacks.

Noise attack is the most common image processing technology to verify robustness. In this article, Gaussian noise attack and Salt & Pepper noise attack have been taken as noise attacks to test the robustness of the proposed method. A Gaussian white noise with a mean of 0 and different standard deviations are used to attack the watermarked image. In Table 2, the average BER values can be given for the different Gaussian noise. It can be clearly seen that our method has lower BER values which attacked by Gaussian noise. Table 2 shows the average comparison results after adding the Salt & Pepper noise with different noise quantity. Although the average BER of our method is approximately the same as the Ref. [22] when the noise factor is 0.05%, but the experimental results of other noise factors, the average BER of our method is lower than Ref. [21], Ref. [22], Ref. [43], Ref. [44] and Ref. [45]. As can be seen from Table 2, the proposed scheme had good robustness against the noise attacks and exhibited the best performance.

The watermark robustness against the attack of lossy compression is an important performance to be evaluated. JPEG compression is one of the standard compression techniques that is used to compress the image. In this experiment, the watermarked image is attacked by JPEG compression. Table 3 lists the part experimental results of sample images with different compression factors, respectively. When the compression quality is 60, the average BER value of the

TABLE 2. The average BER comparison with PSNR = 42dB under noise attacks.

Attack	Parameters	Ref. [21]	Ref. [22]	Ref. [43]	Ref. [44]	Ref. [45]	Proposed
Gaussian noise	(0.0005)	0.0105	0.0046	0.1045	0.0181	0.0088	**0.0011**
	(0.0010)	0.0349	0.0242	0.1561	0.0592	0.1444	**0.0138**
	(0.0015)	0.0693	0.0481	0.1856	0.1038	0.1890	**0.0325**
Salt & Pepper noise	(0.0005)	0.0068	0.0054	0.0567	0.0074	0.0201	**0.0051**
	(0.0010)	0.0162	0.0132	0.0673	0.0179	0.0382	**0.0093**
	(0.0015)	0.0235	0.0176	0.0762	0.0262	0.0409	**0.0151**

TABLE 3. The average BER comparison with PSNR = 42dB under JPEG compression attacks.

Attack	Parameters	Ref. [21]	Ref. [22]	Ref. [43]	Ref. [44]	Ref. [45]	Proposed
JPEG	(40)	0.0318	0.0170	0.1699	0.0544	0.1677	**0.0055**
	(50)	0.0089	0.0066	0.1149	0.0203	0.1182	**0.0012**
	(60)	0.0031	0.0024	0.0869	0.0088	0.0833	**0.0005**

TABLE 4. The average BER comparison with PSNR = 42dB under filtering attacks.

Attack	Parameters	Ref. [21]	Ref. [22]	Ref. [43]	Ref. [44]	Ref. [45]	Proposed
Gaussian filter	(3×3)	0.0215	0.0267	0.0577	0.0149	0.0323	**0.0135**
Median filter	(3×3)	0.1628	0.1750	0.2462	0.1736	0.1951	**0.0731**

proposed model is nearly to zero. Thus, our method has better performance than others when against the JPEG compression attacks.

Filtering attack is one of the classical attacks for the watermarked image. Since the watermark can be removed from the watermarked image by the filter, the Gaussian filtering (GF) and the Median filtering (MF) are used to process the watermarked image. The robust to Gaussian filtering attack and Median filtering attack with window size 3×3. Table 4 shows the average experiment results by the GF attack and MF. On total, the proposed method has the best robustness performance under filtering attacks.

The watermarked image is first rotated clockwise by a certain number of degrees, and then counterclockwise by the same number of degrees. In this experiment, the watermarked image is first rotated 15°, 30°, 60° clockwise, and then

TABLE 5. The average BER comparison with PSNR = 42dB under Rotation.

Attack	Parameters	Ref. [21]	Ref. [22]	Ref. [43]	Ref. [44]	Ref. [45]	Proposed
Rotation	(15°)	0.0078	0.0084	0.0461	0.0075	0.1091	**0.0015**
	(30°)	0.0052	0.0047	0.0412	0.0044	0.0204	**0.0007**
	(60°)	0.0056	0.0054	0.0426	0.0048	0.0202	**0.0004**

TABLE 6. The average BER comparison with PSNR = 42dB under combined attacks.

Attack	Parameters	Ref. [21]	Ref. [22]	Ref. [43]	Ref. [44]	Ref. [45]	Proposed
JPEG+ Gaussian noise	(50+0.0010)	0.0770	0.0516	0.2275	0.1123	0.1990	**0.0452**
	(50+0.0015)	0.1152	0.0784	0.2717	0.1539	0.2259	**0.0724**
JPEG+ Salt & Pepper noise	(50+0.0010)	0.0272	0.0216	0.1333	0.0405	0.1354	**0.0187**
	(50+0.0015)	0.0349	0.0271	0.1418	0.0493	0.1410	**0.0255**

restored to its original shape counterclockwise. Table 5 shows the average BER values at different rotation angles, and the average BER values do not exceed 0.2% from our method, which can prove our method has better performance. Generally, the watermarked image will be contaminated by multiple attacks. We further compared some combined attacks the same as Table 6. The testing images attacked by JPEG quality = 50, and then attacked by the noise attacked. It can be noted that our method shows better robustness performance.

Table 7 demonstrates the comparsion average BER values between our scheme and other methods for different image attacks with filxed image quality, VSI = 0.9970. It is obviously seen that our proposed method has better robustness to against the process of adding Gaussian noise and noise. For JPEG compression, we only list the experimental results with QF of 40 and 50. It can be seen from Table 7 that our method has a lower BER than other watermarking algorithms. The value of BER obtained by proposed method do not exceed 0.1% for GF with window size 3×3. When the watermarked image is attacked by MF with window 3×3, the average BER of our method is about 0.5% lower than that of Ref. [44]. And for rotation, the value of average BER obtained by proposed method do not exceed 0.06% when the rotation angle is 15°, 30°. Besides, the robustness performance of our method is obviously better than others under combined image attacks.

D. ROBUSTNESS WITH SPATIAL-UNIFORM EMBEDDING-BASED WATERMARKING METHOD

In this section, the proposed method compares with the spatial-uniform watermarking method proposed in Refs. [15] and [16]. In Ref. [15], a simple preprocessing of a 32×32 binary watermark is proposed to obtain the actual 64×64 bits, which are embedded into 64×64 blocks. Another selective mechanism is used in Ref. [16], where 32×32 blocks are adaptive selected from the original 128×128 blocks. In this section, we compare the proposed scheme with Refs. [15] and [16]. To ensure the fairness of the experiment, the proposed method compares with other existing image watermarking Ref. [15], with the same image quality. The binary watermark with size 32×32, and the Y channel of the color image are utilized to embed watermark. And same as Ref. [15], an optimum sub-watermark can be obtained by combining 4 sub-watermarks here. The BER was computed to make the objective performance comparison.

To verify the robustness of our method, watermarked images are attacked by common image processing operations (such as adding noise, JPEG compression, filtering attacks, and combined attack) and geometrical distortions (such as scaling, rotation, and cropping). We test watermarked images by different attacks with different factors. Table 8 demonstrates the average BER values for different image attacks

TABLE 7. The average BER comparison with VSI = 0.9970 under different attacks.

Attack	Parameters	Ref. [21]	Ref. [22]	Ref. [43]	Ref. [44]	Ref. [45]	Proposed
Gaussian noise	(0.0010)	0.0119	0.0070	0.0998	0.0449	0.0955	**0.0035**
	(0.0015)	0.0258	0.0158	0.1186	0.0623	0.1190	**0.0106**
Salt & Pepper noise	(0.0010)	0.0090	0.0058	0.0684	0.0302	0.0457	**0.0024**
	(0.0015)	0.0129	0.0080	0.0718	0.0353	0.0515	**0.0042**
JPEG	(40)	0.0026	0.0015	0.0953	0.0430	0.1122	**0.0011**
	(50)	0.0007	0.0010	0.0819	0.0337	0.0808	**0.0006**
Gaussian filter	(3×3)	0.0112	0.0121	0.0289	0.0058	0.0290	**0.0059**
Median filter	(3×3)	0.1442	0.1566	0.1483	0.1052	0.1174	**0.0597**
Rotation	(15°)	0.0034	0.0031	0.0271	0.0024	0.1878	**0.0005**
	(30°)	0.0015	0.0021	0.0246	0.0011	0.0200	**0.0004**
JPEG+ Gaussian noise	(50+0.0005)	0.0127	0.0064	0.0777	0.0167	0.0810	**0.0042**
	(50+0.0010)	0.0306	0.0156	0.1000	0.0347	0.1007	**0.0122**
	(50+0.0015)	0.0487	0.0274	0.1209	0.0521	0.1179	**0.0227**
JPEG+ Salt&Pepper noise	(50+0.0005)	0.0079	0.0048	0.0612	0.0084	0.0594	**0.0031**
	(50+0.0010)	0.0130	0.0082	0.0665	0.0140	0.0652	**0.0056**
	(50+0.0015)	0.0173	0.0107	0.0695	0.0186	0.0694	**0.0087**

with filxed image quality, PSNR = 42dB. The BER of Gaussian noise with variance 0.0025 and the BER of Salt & Pepper noise (SPN) with the noise quantity 0.0025 are given in Table 8. It can be clearly seen that our method produces a lower BER than other methods for noise attack. For JPEG compression, we show the results of JPEG compression with JPEG factor=30 in Table 8. Although the average BER obtained by using our method is 0.19% higher than that of Ref. [16], it is 0.2% lower than that of Ref. [15]. In the filtering attack, our method and Ref. [15] has good robustness performance. The scaling operations, as an image geometric attack, is often used in image processing. At first, the

FIGURE 7. Results under different types of attacks and recovered watermark from Lena image (a) Gaussian Noise (var = 0.002) (b) Gaussian Noise (var = 0.003) (c) Salt&Pepper noise (density = 0.002) (d) Salt&Pepper noise (density = 0.003)(e) JPEG with QF = 30 (f) JPEG with QF = 50 and Gaussian noise (0.002) (g) Gaussian filter (3, 3) (h) Median filter (3, 3) (i) Centred Cropping 1/16 (128 × 128 by white) (j) Left upper cropping 1/16 (128 × 128 by white) (k) Centred Cropping 1/16 (128 × 128 by black) (l) Left upper cropping 1/16 (128 × 128 by black) (m) Scaling 0.25 (n) Scaling 0.33 (o) Image rotation (angle = 25°) (p) Image rotation (angle = 60°).

TABLE 8. Average BER values with PSNR = 42dB between Ref. [15], Ref. [16], and our proposed scheme under different attacks.

	GN 0.0025	SPN 0.0025	JPEG 30	Scaling 0.25	MF (3,3)	GF (3,3)	Cen-Crop 1/16	Rot (25°)	J+GN (50+0.002)
Ref. [15]	0.0439	0.0439	0.0059	0.0264	0.0078	**0.0000**	**0.0000**	0.0684	0.0479
Ref. [16]	0.0166	0.1387	**0.0020**	0.2940	0.0264	0.0737	0.1117	0.1680	0.0225
Proposed	**0.0107**	**0.0303**	0.0039	**0.0146**	**0.0020**	**0.0000**	**0.0000**	**0.0537**	**0.0107**

watermarked image is scaled down from 25% to 200% with an increment size 25%, respectively. We give the BER when the watermarked image is reduced by 25% and then restored to the original size. The BER of our method is about 1% lower than that of Ref. [15]. For cropping operations, we only give BER after the center-crop (Cen-Crop) by 128 × 128 white. It can be clearly seen that our method and Ref. [15] can resist this image processing operation. We give the experimental result of rotating (Rot) the watermarked image by 25°. The watermarked image is first rotated clockwise by 25°, and then counterclockwise by the same degree. Our method has the lower average BER than others. Moreover, the watermarked images also will be attacked by the combined attacks. The BER is not exceed 1.1% under JPEG=50 and GN with 0.002 (J+GN) in our method. Thus, the robustness performance of our method is obviously better than Ref. [15] and Ref. [16].

In order to better demonstrate the robustness of our method, 'Lena' watermarked image with image quality of PSNR = 42dB, and then attacked by various image processing opetations and the extracted watermark information are shown in Figure 7.

At the same time, we also tested the robustness of the watermarked image with the image quality maintained at VSI=0.9970. The average BER of watermarking algorithms are compared through various common image processing of watermarked image. For different attacks with various factors, we only give some experimental results in Table 9. The BER of Gaussian noise with variance 0.003 and the BER of Salt & Pepper noise with the noise quantity 0.01 are given in Table 9. Our proposed method can resist the noise attacks than other methods from BER values. For JPEG compression, we only give a BER with quality factor of 20. The BER

TABLE 9. Average BER values with VSI = 0.9970 between Ref. [15], Ref. [16], and our proposed scheme under different attacks.

	GN 0.003	SPN 0.01	JPEG 20	Scaling 0.25	MF (3,3)	GF (3,3)	Cen-Crop 1/16	Rot (25°)	J+GN (50+0.002)
Ref. [15]	0.0057	0.0078	0.0228	**0.0164**	0.0117	**0.0000**	**0.0000**	0.0546	0.0016
Ref. [16]	0.0213	0.0615	0.0376	0.3032	0.0667	0.0068	0.0390	0.1567	0.0094
Proposed	**0.0034**	**0.0068**	**0.0091**	0.0166	**0.0105**	**0.0000**	**0.0000**	**0.0522**	**0.0006**

TABLE 10. The average BER comparison with PSNR = 42dB under different attacks.

Attack	Ref. [21]	Ref. [22]	Ref. [43]	Ref. [44]	Ref. [45]	Proposed
Gaussian filter 3×3	0.0215	0.0267	0.0577	0.0149	0.0323	**0.0135**
Median filter 3×3	0.1628	0.1750	0.2462	0.1736	0.1951	**0.0731**
Wiener filter 3×3	0.1327	0.1440	0.1980	0.1099	0.1412	**0.0200**
Sharpen 0.5	0.4848	0.4293	0.5822	0.4323	**0.2304**	0.2903
Rotation 30°	0.0060	0.0060	0.0412	0.0045	0.0205	**0.0010**
Scaling 2	0.0028	0.0048	0.0316	0.0012	0.0145	**0.0000**
Cen-Crop 1/16	0.0386	0.0433	0.0773	0.0378	0.0457	**0.0350**
Lef-Up-Crop 1/16	0.0347	0.0419	0.0718	0.0327	0.0407	**0.0316**
JPEG compression 40	0.0318	0.0170	0.1699	0.0544	0.1677	**0.0055**
Template remove	0.1580	0.1636	0.221	0.1355	0.1817	**0.0677**

obtained by our method is about 0.1% lower than that of Ref. [15] and 0.2% lower than that of Ref. [16]. Our method and Ref. [15] both show good robustness under filtering attacks and cropping attacks. For Scaling attack, we list the BER of watermark information after the watermarked image is scaled to 25%, and then the BER of the original size is restored. The method we proposed and Ref. [15] get a similar bit error rate. Table 9 shows the experimental result of watermarked images at 25°. The BER of our method is 10% lower than that of Ref. [16]. The BER of our method not exceed 0.01%, which can prove the proposed method can effectively against JPEG + Gaussian noise attacks.

E. ROBUSTNESS WITH QUANTIZATION-BASED WATERMARKING METHODS ON BENCHMARK

Table 10 lists the average BER values between our scheme and some quantization-based watermarking schemes [21], [22], [43], [44] and [45] on the software benchmark. There are different image attacks on the software benchmark, and we give experimental results on table 10. For filter attacks, we used the Gaussian filter, Median filter, and Wiener filter with a 3×3 window. In table 10, the proposed method has obvious performance than others for filter attacks. In the procedure of the sharpening, the watermarked image is attacked by different radii, and we only give the experiment result under radii with 0.5. The average BER value is higher than Ref. [44], but lower other watermarking methods. The watermarked image also has been attacked by Rotation and Scaling. The lower average results can be obtained by our proposed method under these two image attacks. For cropping attack, we respectively list average BER values after center cropping (Cen-Crop) with 128×128 white and left upper cropping (Lef-Up-Crop) with 128×128 white. The proposed method has better robustness than others for cropping attacks. For JPEG compression, we show the result for JPEG factor=40. Compared with the other methods, the proposed method can resist the JPEG compression attack to a

certain extent. For the template removal attack, our method is robust than others. Compared with the other watermarking methods, we not only use the Y channel as the embedding position but also consider the relationship between the chromaticity of the color image. Meanwhile, using the chromaticity information to construct a visual perception model that guides the embedding of watermark information. It can be seen that the proposed method has better robustness than others.

VI. CONCLUSION

This article proposes a new spatial-perceptual embedding with the guidance of the JND model for color image watermarking. The DC coefficient of the block is used as the cover coefficient for watermark embedding, and the DC coefficients are quantized by DM. Different from the previous JND model, we proposed an attended JND model with color complexity, which is more consider the human visual perception model. The pixel modification amount is modified by the pixel domain JND thresholds so that the watermarked image is more consistent with the perceptual characteristics of HVS. Experimental results have demonstrated that the proposed JND model has better robustness than other existing JND models in watermarking scheme. And the proposed watermarking method is robust against common image processing attacks, such as noise addition, compression, and volumetric attack. In the future, we will not only consider the pixel update value in the watermarking framework from the HVS system but also consider implementing pixel updates in the watermarking framework together with other image enhancement tasks.

REFERENCES

[1] I. Cox, M. Miller, J. Bloom, J. Fridrich, and T. Kalker, *Digital Watermarking and Steganography*. San Mateo, CA, USA: Morgan Kaufmann, 2007.

[2] B. Chen, J. Wang, Y. Chen, Z. Jin, H. J. Shim, and Y. Shi, "High-capacity robust image steganography via adversarial network," *KSII Trans. Internet Inf. Syst.*, vol. 14, no. 1, pp. 366–381, 2020.

[3] J. Li, H. Zhang, W. Wan, and J. Sun, "Two-class 3D-CNN classifiers combination for video copy detection," *Multimedia Tools Appl.*, vol. 79, no. 7, pp. 4749–4761, 2020.

[4] 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," *Appl. Soft Comput.*, vol. 68, pp. 360–376, Jul. 2018.

[5] J. Zong, L. Meng, H. Zhang, and W. Wan, "JND-based multiple description image coding," *Appl. Soft Comput.*, vol. 68, pp. 360–376, 2018.

[6] Q. Su, Y. Niu, Q. Wang, and G. Sheng, "A blind color image watermarking based on DC component in the spatial domain," *Optik*, vol. 124, no. 23, pp. 6255–6260, Dec. 2013.

[7] B. Chen, C. Zhou, B. Jeon, Y. Zheng, and J. Wang, "Quaternion discrete fractional random transform for color image adaptive watermarking," *Multimedia Tools Appl.*, vol. 77, no. 16, pp. 20809–20837, 2018.

[8] 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," *Inf. Sci.*, vols. 436–437, pp. 247–267, Apr. 2018.

[9] S. Roy and A. K. Pal, "A blind DCT based color watermarking algorithm for embedding multiple watermarks," *AEU-Int. J. Electron. Commun.*, vol. 72, pp. 149–161, Feb. 2017.

[10] C. Das, S. Panigrahi, V. K. Sharma, and K. K. Mahapatra, "A novel blind robust image watermarking in DCT domain using inter-block coefficient correlation," *AEU-Int. J. Electron. Commun.*, vol. 68, no. 3, pp. 244–253, Mar. 2014.

[11] M. Cedillo-Hernandez, F. Garcia-Ugalde, M. Nakano-Miyatake, and H. Perez-Meana, "Robust watermarking method in DFT domain for effective management of medical imaging," *Signal, Image Video Process.*, vol. 9, no. 5, pp. 1163–1178, Jul. 2015.

[12] C. Pradhan, S. Rath, and A. K. Bisoi, "Non blind digital watermarking technique using DWT and cross chaos," *Procedia Technol.*, vol. 6, pp. 897–904, Jan. 2012.

[13] M. Ali, C. W. Ahn, and M. Pant, "A robust image watermarking technique using SVD and differential evolution in DCT domain," *Optik*, vol. 125, no. 1, pp. 428–434, Jan. 2014.

[14] J. Huang, Y. Q. Shi, and Y. Shi, "Embedding image watermarks in DC components," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 6, pp. 974–979, Sep. 2000.

[15] Q. Su and B. Chen, "Robust color image watermarking technique in the spatial domain," *Soft Comput.*, vol. 22, no. 1, pp. 91–106, Jan. 2018.

[16] Q. Su, Z. Yuan, and D. Liu, "An approximate Schur decomposition-based spatial domain color image watermarking method," *IEEE Access*, vol. 7, pp. 4370–4385, 2018.

[17] Q. Su, D. Liu, Z. Yuan, G. Wang, X. Zhang, B. Chen, and T. Yao, "New rapid and robust color image watermarking technique in spatial domain," *IEEE Access*, vol. 7, pp. 30398–30409, 2019.

[18] J. Wu, L. Li, W. Dong, G. Shi, W. Lin, and C.-C.-J. Kuo, "Enhanced just noticeable difference model for images with pattern complexity," *IEEE Trans. Image Process.*, vol. 26, no. 6, pp. 2682–2693, Jun. 2017.

[19] V. Jakhetiya, W. Lin, S. Jaiswal, K. Gu, and S. C. Guntuku, "Just noticeable difference for natural images using RMS contrast and feed-back mechanism," *Neurocomputing*, vol. 275, pp. 366–376, Jan. 2018.

[20] W. Wan, J. Wang, J. Li, L. Meng, J. Sun, H. Zhang, and J. Liu, "Pattern complexity-based JND estimation for quantization watermarking," *Pattern Recognit. Lett.*, vol. 130, pp. 157–164, Feb. 2020.

[21] C. Wang, M. Xu, W. Wan, J. Wang, J. Li, and J. Sun, "Robust image watermarking via perceptual structural regularity-based JND model," *KSII Trans. Internet Inf. Syst.*, vol. 13, no. 2, pp. 1080–1099, 2019.

[22] W. Wan, J. Wang, M. Xu, J. Li, J. Sun, and H. Zhang, "Robust image watermarking based on two-layer visual saliency-induced JND profile," *IEEE Access*, vol. 7, pp. 39826–39841, 2019.

[23] X. Lu, H. Zhang, J. Sun, Z. Wang, P. Guo, and W. Wan, "Discriminative correlation hashing for supervised cross-modal retrieval," *Signal Process., Image Commun.*, vol. 65, pp. 221–230, Jul. 2018.

[24] H.-T. Hu and J.-R. Chang, "Dual image watermarking by exploiting the properties of selected DCT coefficients with JND modeling," *Multimedia Tools Appl.*, vol. 77, no. 20, pp. 26965–26990, Oct. 2018.

[25] 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 Appl.*, vol. 79, no. 7, pp. 4907–4930, 2018.

[26] J. Wang, W. B. Wan, X. X. Li, J. D. Sun, and H. X. Zhang, "Color image watermarking based on orientation diversity and color complexity," *Expert Syst. Appl.*, vol. 140, Feb. 2020, Art. no. 112868.

[27] J. Wang and W. Wan, "A novel attention-guided JND model for improving robust image watermarking," *Multimedia Tools Appl.*, vol. 79, nos. 33–34, pp. 24057–24073, Sep. 2020, doi: 10.1007/s11042-020-09102-2.

[28] S. H. Bae and M. Kim, "A new DCT-based JND model of monochrome images for contrast masking effects with texture complexity and frequency," in *Proc. IEEE Int. Conf. Image Process.*, Jun. 2013, vol. 1, no. 1, pp. 431–434.

[29] C. H. Chou and Y. C. Li, "A perceptually tuned subband image coder based on the measure of just-noticeable-distortion profile," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 5, no. 6, pp. 467–476, Dec. 1995.

[30] J. Wu, W. Lin, G. Shi, X. Wang, and F. Li, "Pattern masking estimation in image with structural uncertainty," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 4892–4904, Dec. 2013.

[31] H. Tian, Y. Fang, Y. Zhao, W. Lin, R. Ni, and Z. Zhu, "Salient region detection by fusing bottom-up and top-down features extracted from a single image," *IEEE Trans. Image Process.*, vol. 23, no. 10, pp. 4389–4398, Oct. 2014.

[32] B. Ortiz-Jaramillo, A. Kumcu, L. Platisa, and W. Philips, "Evaluation of color differences in natural scene color images," *Signal Process., Image Commun.*, vol. 71, pp. 128–137, Feb. 2019.

[33] S.-L. Jia, "A novel blind color images watermarking based on SVD," *Optik*, vol. 125, no. 12, pp. 2868–2874, Jun. 2014.

[34] F. Zhang, T. Luo, G. Jiang, M. Yu, H. Xu, and W. Zhou, "A novel robust color image watermarking method using RGB correlations," *Multimedia Tools Appl.*, vol. 78, no. 14, pp. 20133–20155, Jul. 2019.

[35] R. Koju and S. R. Joshi, "Comparative analysis of color image watermarking technique in RGB, YUV, and YCbCr color channels," *Nepal J. Sci. Technol.*, vol. 15, no. 2, pp. 133–140, Feb. 2015.
[36] H. Liu, M. Huang, G. Cui, M. R. Luo, and M. Melgosa, "Color-difference evaluation for digital images using a categorical judgment method," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 30, no. 4, pp. 616–626, 2013.
[37] A. Borji and L. Itti, "State-of-the-Art in visual attention modeling," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 185–207, Jan. 2013.
[38] 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, Nov. 2016.
[39] L. Zhang, Y. Shen, and H. Li, "VSI: A visual saliency-induced index for perceptual image quality assessment," *IEEE Trans. Image Process.*, vol. 23, no. 10, pp. 4270–4281, Oct. 2014.
[40] University of Granada. *Computer Vision Group. CVG-UGR Image Database*. Accessed: Mar. 13, 2017. [Online]. Available: http://decsai.ugr.es/cvg/dbimagenes/
[41] H. Hadizadeh, "A saliency-modulated just-noticeable-distortion model with non-linear saliency modulation functions," *Pattern Recognit. Lett.*, vol. 84, pp. 49–55, Dec. 2016.
[42] J. Wu, G. Shi, W. Lin, A. Liu, and F. Qi, "Just noticeable difference estimation for images with free-energy principle," *IEEE Trans. Multimedia*, vol. 15, no. 7, pp. 1705–1710, Nov. 2013.
[43] Q. Li and I. J. Cox, "Improved spread transform dither modulation using a perceptual model: Robustness to amplitude scaling and JPEG compression," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Processing-ICASSP*, vol. 2, Apr. 2007, p. II-185.
[44] X. Li, J. Liu, J. Sun, X. Yang, and W. Liu, "Step-projection-based spread transform dither modulation," *IET Inf. Secur.*, vol. 5, no. 3, pp. 170–180, 2011.
[45] L. Ma, D. Yu, G. Wei, J. Tian, and H. Lu, "Adaptive spread-transform modulation using a new perceptual model for color image," *IEICE Trans. Inf. Syst.*, vol. 93, no. 4, pp. 843–857, 2010.

WENBO WAN received the Ph.D. degree from Shandong University, Jinan, China, in 2015, supervised by Prof. Ju Liu. From June 2019 to October 2019, he was a Visiting Researcher with the Department of Computer Science, City University of Hong Kong, Hong Kong. He is currently an Associate Professor with the School of Information Science and Engineering, Shandong Normal University. His research interests include multimedia security, multimedia quality assessment, and image/video watermarking.

KAI ZHOU received the bachelor's degree from the Science and Information College, Qingdao Agricultural University, Qingdao, Shandong, in 2017. He is currently pursuing the master's degree in communication and information systems with Shandong Normal University. His research interests include image processing and image/video watermarking.

KAI ZHANG received the Ph.D. degree from Xidian University, Xi'an, China, in 2018. He is currently a Lecturer with the School of Information Science and Engineering, Shandong Normal University. His research interests include artificial intelligence, machine learning, multisource remote sensing image fusion, and intelligent interpretation.

YANTONG ZHAN received the Ph.D. degree from the China University of Mining and Technology, Beijing, China, in 2019. She is currently a Lecturer with the School of Information Science and Engineering, Shandong Normal University. Her research interests include artificial intelligence and pattern recognition, and industrial field intelligent detection system research and development.

JING LI (Member, IEEE) received the Ph.D. degree from Shandong Normal University, Jinan, China, in 2020. She is currently an Associate Professor with the School of Mechanical and Electrical Engineering, Shandong Management University. Her research interests include multimedia security and multimedia processing and retrieval.