



A NOVEL ROBUST REVERSIBLE WATERMARKING FOR COPYRIGHT PROTECTION OF IMAGES

Dr SK Moulali

Associate Professor, Department of CSE

Sri Sai Institute of Technology and Science, Rayachoti

Email: moulaliphd@gmail.com

Abstract— Robust reversible watermarking (RRW) methods are popular in multimedia for protecting copyright, while preserving intactness of host images and providing robustness against unintentional attacks. However, conventional RRW methods are not readily applicable in practice. That is mainly because: 1) they fail to offer satisfactory reversibility on large-scale image datasets; 2) they have limited robustness in extracting watermarks from the watermarked images destroyed by different unintentional attacks; and 3) some of them suffer from extremely poor invisibility for watermarked images. Therefore, it is necessary to have a framework to address these three problems, and further improve its performance. This paper presents a novel pragmatic framework, wavelet-domain statistical quantity histogram shifting and clustering (WSQH-SC). Compared with conventional methods, WSQH-SC ingeniously constructs new watermark embedding and extraction procedures by histogram shifting and clustering, which are important for improving robustness and reducing run-time complexity. Additionally, WSQH-SC includes the property-inspired pixel adjustment to effectively handle overflow and underflow of pixels. This results in satisfactory reversibility and invisibility. Furthermore, to increase its practical applicability, WSQH-SC designs an enhanced pixel-wise masking to balance robustness and invisibility. We perform extensive experiments over natural, medical, and synthetic aperture radar images to show the effectiveness of WSQH-SC by comparing with the histogram rotation-based and histogram distribution constrained methods.

Index Terms— Integer wavelet transform, k -means clustering, masking, robust reversible watermarking (RRW).

I. INTRODUCTION

REVERSIBLE WATERMARKING (RW) methods [1] are used to embed watermarks [2], e.g., secret information [3], into digital media while preserving high intactness and good fidelity of host media. It plays an important role in protecting copyright and content of digital media for sensitive applications,

S.M. Shamsheerdaula

Asst Prof. ECE Dept

G. Pulla Reddy Engg College, Kurnool

shamsheerdaula@gmail.com

e.g., medical and military images. Although researchers proposed some RW methods for various media, e.g., images [4], [5], audios [6], videos [7],

And 3-D meshes [8], they assume the transmission channel is lossless. The robust RW (RRW) is thus a challenging task. For RRW, the essential objective is to accomplish watermark embedding and extraction in both lossless and lossy environment. As a result,

RRW is required to not only recover host images and

Watermarks without distortion for the lossless channel, but also resist unintentional attacks and extract as many watermarks as possible for the noised channel [9]. Recently, a dozen of RRW methods for digital images have been proposed [10]–[13], which can be classified into two groups [9]: histogram rotation (HR)-based methods and histogram distribution constrained (HDC) methods.

The HR-based methods [10], [11] accomplish robust loss-less embedding by slightly rotating the cancrroids vectors of two random zones in the no overlapping blocks. Due to the close correlation of neighboring pixels, these methods were reported to be robust against JPEG compression. However, they are sensitive to “salt-and-pepper” noise, which leads to poor visual quality of watermarked images, and impedes lossless recovery of host images and watermarks. To solve this problem, the HDC methods have been developed in spatial-and wavelet-domains [12], [13], which divide image blocks into different types and embed the modulated watermarks for each type based on histogram distribution. Unfortunately, these methods suffer from unstable reversibility and robustness according to [9]. In summary, the above analysis shows that both kinds of RRW methods are not readily applicable in practice. Therefore, a novel pragmatic RRW framework with the following three objectives is of great demand:

- 1) Reversibility, i.e., how to handle both overflow and underflow of pixels;
- 2) Robustness, i.e., how to resist unintentional attacks; and
- 3) Invisibility, i.e., how to make a trade-off between robustness and invisibility.

In this paper, motivated by the excellent spatial-frequency localization properties of wavelet transform, we develop a novel RRW framework in the wavelet domain. This framework uses the statistical quantity histogram (SQH) as the embedding carrier inspired by our previous work, the generalized SQH (GSQH) driven method [14], and constructs new watermark embedding and extraction processes by histogram shifting and clustering. In this framework, we carefully design the three key components, which are the property inspired pixel adjustment (PIPA), the SQH shifting and clustering, and the enhanced pixel-wise masking (EPWM), to effectively solve the aforementioned three problems. In particular:

- 1) *PIPA*: To successfully avoid both overflow and under-flow of pixels, we develop PIPA to investigate the intrinsic relationship between wavelet coefficient and pixel changes in order to determine how to duly change wavelet coefficients during the embedding process. By taking the scale and region of wavelet coefficient changes into account, PIPA preprocesses the host images accordingly by adjusting the pixels possible to overflow and underflow into a reliable range before embedding. Finally, the preprocessed host images are used to embed watermarks.
- 2) *SQH Shifting and Clustering*: To better resist unintentional attacks, we build SQH with threshold constraint by deeply studying characteristics of the wavelet coefficients, design the watermark embedding process by bi-directionally shifting SQH, and adopt the k -means clustering algorithm to recover watermarks by creatively modeling the extraction process as a classification problem. Besides from superior robustness, this way simplifies watermark embedding and extraction, and reduces the run-time complexity of the proposed framework.

EPWM: To effectively balance robustness and invisibility, we consider the local sensitivity of human visual system (HVS) in wavelet domain, and design an EPWM to precisely evaluate the just noticeable distortion (JND) thresholds of wavelet coefficients, which thereafter are used to adaptively optimize watermark strength. Because the SQH shifting and clustering are employed for watermark embedding and extraction, respectively, we term the proposed framework the wavelet-domain SQH shifting and clustering or WSQH-SC for short. WSQH-SC has the following advantages: 1) it not only offers the robust and lossless watermark embedding and extraction processes by integrating SQH shifting and clustering, but also duly introduces perceptual characteristics of HVS to the RRW field, forming a novel yet pragmatic RRW framework; 2) it out performs the two representative kinds of RRW methods in terms of reversibility, robustness, invisibility, capacity, and run-time complexity; 3) it is widely applicable to different kinds of images; and 4) it is readily applicable in practice with the help of strong robustness and optimal performance trade-off..The rest of this paper is organized as follows. Section II briefs the related works for readers to better understand the proposed framework. In Section III, we detail the proposed framework with four modules including PIPA, SQH constructions, EPWM-based embedding, and extraction based on

Notation	Description	Notation	Description
I	Host image	$h \times w$	Block size
$2M \times 2N$	Image size	δ	Threshold
$p_k(i, j)$	Wavelet coefficient at (i, j) in the k th block	x_l, x_r	Peak points
c^ω_ρ	Sub-band with orientation ω and resolution level ρ	η	Adjustment Scale
$c^\omega(i, j)_\rho$	Wavelet coefficient at (i, j) in sub-band c^ω_ρ	b_k	k th Watermark Bit
$JND_\rho(i, j)$	JND threshold of $c^\omega(i, j)_\rho$	λ	Watermark strength
S_k	Mean of wavelet coefficients in the k th block	I^w	Watermarked image

Table: IMPORTANT NOTATIONS

k -means clustering. The experimental results in Section IV and Section V thoroughly demonstrate the effectiveness and superiority of the proposed framework, and Section VI concludes this paper.

For convenience, Table I lists important notations used in the rest of this paper.

II. RELATED WORKS

In this section, we briefly introduce the GSQH driven method [14] and discuss its useful inspirations to our novel framework. Thereafter, a popular pixel-wise masking (PWM) model is presented to lay the groundwork for the proposed EPWM.

A. GSQH Driven Method:

The histogram plays an important role in many practical models and applications, e.g., histogram of oriented gradient features [15], bag-of-words [16], [17], and digital watermarking [18]–[20]. For RW methods, SQH has recently received considerable attention due to stability and simplicity, e.g., arithmetic average of difference (AAD) histogram [21], difference histogram [22], and prediction error histogram [23]. In particular, we proposed a GSQH driven method [14], which embeds and extracts watermarks by SQH shifting. The following is a brief review of this method.

Given a t -bit host image I with n^* no overlapping blocks, its SQH can be generated by calculating the AAD of each block. For convenience, we denote the SQH by a set of data pairs, i.e., $X = \{(x_1, n_1), \dots, (x_i, n_i), \dots, (x_{m^*}, n_{m^*})\}$, where x_i represents the different values of the AAD, and n_i is the corresponding frequency of x_i in SQH. Let x_r and x_l be the two peak points of SQH, wherein

$$r = \arg \max_i n_i, 1 \leq i \leq m^* \quad (1)$$

and

$$l = \arg \max_i n_i, 1 \leq i \leq m', i \leq r. \quad (2)$$

Suppose $x_l \leq x_r$, then the embedding is done according to

$$s_k^w = \begin{cases} s_k - z - 1, & \text{if } s_k < x_l - z \\ s_k - b_k(z + 1), & \text{if } x_l - z \leq s_k \leq x_l \\ 0, & \text{if } x_l < s_k < x_r \\ s_k + b_k(z + 1), & \text{if } x_r \leq s_k \leq x_r + z \\ s_k + z + 1, & \text{if } s_k > x_r + z \end{cases} \quad (3)$$

in which s_k and s_k^w are the AADs of the k th block in the host and watermarked images, respectively, $b_k \in \{0, 1\}$ is the k th watermark bit, and $z \geq 0$ is a scale factor. Correspondingly, the watermarking extraction is defined as

$$b_k^r = \begin{cases} 0, & \text{if } x_l - z \leq s_k^w \leq x_l \text{ or } x_r \leq s_k^w \leq x_r + z \\ 1, & \text{if } x_l - 2z - 1 \leq s_k^w \leq x_l - z - 1 \text{ or } x_r + z + 1 \leq s_k^w \leq x_r + 2z + 1 \end{cases} \quad (4)$$

Where b_k^r is the k th extracted watermark bit. Finally, the host image can be recovered without distortion using the inverse operation of (3) when the watermarked image is not degraded by unintentional attacks. Extensive experimental results suggest that the GSQH driven method has its pros and cons. On one hand, it combines GSQH and histogram shifting together to obtain good performance. On the other hand, however, it has three shortcomings: 1) it uses the AADs of all of the blocks, both reliable and unreliable, to generate the SQH of the host image, which increases complexity of watermark embedding; 2) it fails to consider the optimization of watermark strength; and 3) it suffers from unstable robustness against JPEG compression. By taking these pros and cons into account, we therefore integrate PIPA, SQH shifting, clustering, and EPWM into a novel RRW framework, which effectively overcomes the above shortcomings and makes our work intrinsically different from existing RRW methods

$$x_r + z + 1 \leq s_k \leq x_r + 2z + 1$$

B.PWM

The past years have witnessed the significance of HVS in various applications and many visual masking algorithms revealing the perceptual characteristics of HVS have been applied to digital watermarking. In particular a PWM algorithm proposed by Bami et al. has received much publicity, which computes the JND threshold of each

Here, $_{\rho}$, $_{\omega}$, $_{\rho, i, j}$ and $_{\rho, i, j}$ evaluate resolution, brightness, and texture sensitivities, respectively, defined by

wavelet coefficient based on resolution sensitivity, brightness sensitivity, and texture sensitivity. Given the wavelet coefficient $c_{\rho}^{\omega}(i, j)$ at (i, j) in the sub-band c_{ρ}^{ω} with resolution level $\rho \in \{0, 1, 2, 3\}$ and orientation $\omega \in \{LL, LH, HL, HH\}$, the JND threshold is denoted by

$$JND_{\rho}^{\omega}(i, j) = _{\rho, \omega} _{\rho, i, j} _{\rho, i, j}^{0.02}. \quad (5)$$

$$_{\rho, \omega} = \begin{cases} \sqrt{_{\rho, \omega}} & \text{if } \rho = 0 \\ 2, & \text{if } \omega = HH \\ 1, & \text{otherwise} \end{cases} \cdot \begin{cases} 1.00, & \text{if } \rho = 0 \\ 0.32, & \text{if } \rho = 1 \\ 0.16, & \text{if } \rho = 2 \\ 0.10, & \text{if } \rho = 3 \end{cases} \quad (6)$$

and

$$-(\rho, i, j) = \begin{cases} 1 - (\rho, i, j), & \text{if } (\rho, i, j) < 0.5 \\ + (\rho, i, j), & \text{otherwise} \end{cases} \quad (7)$$

$$\begin{aligned} c_{k+p}^{\omega} &= x + \frac{i}{2^k}, y + \frac{j}{2^k} \\ \times \text{Var } c_3^{LL} &= 1 + x + \frac{i}{2^{3-\rho}}, 1 + y + \frac{j}{2^{3-\rho}} \quad \begin{matrix} x = 0, 1 \\ y = 0, 1. \end{matrix} \end{aligned} \quad (9)$$

In summary, PWM estimates how HVS perceives disturbances in images by considering the resolution, brightness, and texture sensitivities. However, it is not precise enough because the low-pass sub-band at the forth resolution level, i.e., c_3^{LL} , has less image content, which ends up with the approximate estimation of texture and brightness. To solve this problem, we design the EPWM to better depict local sensitivity of HVS, which not only improves texture and brightness sensitivities but also optimizes the sensitivity weight. In the next section, EPWM is used to adaptively adjust watermark strength, which is helpful for increasing practical applicability of the proposed framework.

III. PROPOSED FRAMEWORK

In this section, we introduce a new RRW framework, i.e., WSQH-SC, which accomplishes the robust lossless embedding task by incorporating the merits of SQH shifting, k -means clustering and EPWM. WSQH-SC comprises two processes watermark embedding and extraction. In view of their similarity, Fig. 1 only shows the diagram of the embedding process in which the three modules are termed: 1) PIPA; 2) SQH construction; and 3) EPWM-based embedding, and they are detailed in the following three subsections. To be specific WSQH-SC first investigates the wavelet sub-band properties in depth and exploits PIPA to preprocess the host image which is of great importance to avoid both overflow and underflow of pixels during the embedding process. Afterward the host image is decomposed by the 5/3 integer wavelet transform (IWT) [30] and the blocks of interest in the sub-band c_0^H are selected to generate the SQH with the help of the threshold constraint. Finally, watermarks can be embedded into the selected blocks by histogram shifting, wherein EPWM is designed to adaptively control watermark strength. After the IWT reconstruction, the watermarked image is obtained

$$\begin{aligned} (\rho, i, j) &= \frac{1}{c_3^{LL}} \left(1 + \frac{i}{2^{3-\rho}}, 1 + \frac{j}{2^{3-\rho}} \right) \\ (\rho, i, j) &=_{k=0} \frac{3-\rho}{2} \omega \in \{LH, HL, HH\} \quad x=0, y=0 \end{aligned}$$

A. PIPA

In RRW, how to handle both overflow and underflow of pixels is important for reversibility. Xuan *et al.* [31] proposed a pixel adjustment strategy to tackle this problem. Unfortunately, it cannot be directly applied to wavelet domain because the adjustment scale related to wavelet transform is unknown. As a consequence, we develop PIPA to handle this problem. Firstly, PIPA deeply exploits the intrinsic relationship between wavelet coefficient and pixel changes. Secondly, by taking the scale and region of wavelet coefficient changes into consideration, PIPA determines the adjustment scale and employs the pixel adjustment strategy to preprocess the host images. To better present the technical details of PIPA, Table III gives the 5/3 filter coefficients [32]. Based on this, we investigate the effects of changing wavelet coefficients on pixels from two aspects.

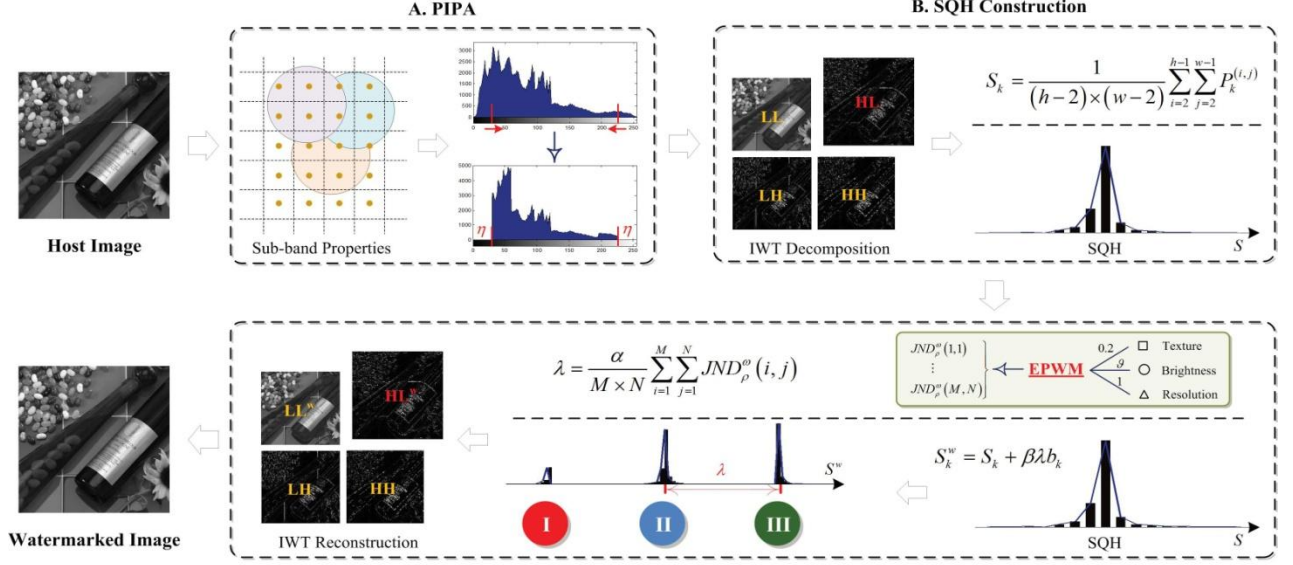


Fig. 1. Embedding process of the proposed wavelet-domain SQH shifting and clustering framework. (a) PIPA. (b) SQH construction. (c) EPWM based embedding.

For convenience, we denote the wavelet coefficient $c_0^\omega(i, j)$ in the sub-band c_0^ω as $c_{i,j}^\omega$, wherein $1 \leq i \leq M$, $1 \leq j \leq N$, and $M \times N$ is the size of c_0^ω .

1) Single Sub-Band and Single Wavelet Coefficient:

Given the watermark strength λ , we consider the changes of pixels when an arbitrary wavelet coefficient in c_0^ω is changed. In particular, if $c_{i,j}^\omega \leftarrow c_{i,j}^\omega + \lambda$, the corresponding changes of pixels in terms of scale and region are shown in Table III, wherein the affected region is represented by the location of the center, $v_L = [1, 2, 1]$ and $v_H = [1, 2, -6, 2, 1]$.

From Table III, we can derive three properties: 1) intra-band correlation, i.e., the pixel regions affected by the neigh-boring wavelet coefficients in a sub-band are overlapped;

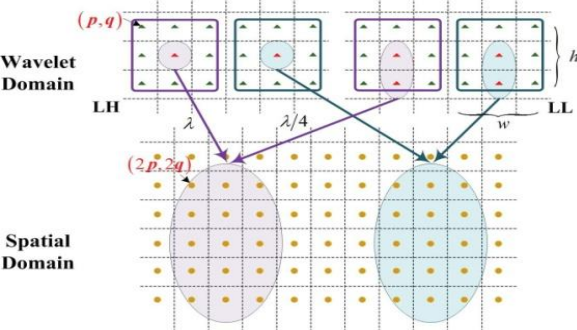
2) **inter-band correlation**, i.e., the regions affected by the wavelet coefficients in different sub-bands are also overlapped; and 3) bi-directional change, i.e., the grayscale values of pixels affected by the wavelet coefficients in the c_0^{HL} , c_0^{LH} , and c_0^{HH} sub-bands are both increased and decreased. Based on this, Fig.2. Example of the effects of changing wavelet coefficients on pixels based on multiple sub-bands and multiple wavelet coefficients.

we conclude that it is impractical to use all of the wavelet coefficients in a sub-band for watermark embedding. That is because it is virtually impossible to determine the adjustment scale and use the pixel adjustment strategy to solve both overflow and underflow problems. Inspired by [13], we aim to search for a new solution by further investigating the effects of changing wavelet coefficients on pixels based on multiple sub-bands and multiple wavelet coefficients.

3) Multiple Sub-Bands and Multiple Wavelet Coefficients:

Considering an arbitrary block with the top left corner at (p, q) in c_0^ω , $1 \leq p < M$, $1 \leq q < N$, we investigate the changes of wavelet coefficients and pixels in two special cases, as shown in Table IV. Here, $v_F = [0, 0, 1, 0, 1, 0, \dots, 1, 0, 0]_{2 \times h-1}$, $v_G = [1, 2, \dots, 2, 1]_{2 \times w-3}$ and the affected region of pixels is

Denoted by the location of its top left corner. To further illus-trade such effects, Fig. 2 shows an example in which the block size is 3×3 , and the wavelet coefficients of two neighboring blocks in c_0^{LL} and c_0^{LH} are changed simultaneously.



With Table IV and Fig. 2, we can deduce that: 1) the affected pixel regions are no overlapped when the wavelet coefficients of neighboring blocks are changed at the same time and 2) the pixel changes are mono directional and

the maximum change scale equals λ . In this case, we can easily Determine the adjustment scale and use the pixel adjustment strategy to preprocess host images. In particular, given a t -bit host image I with the size of $2M \times 2N$, the pixel adjustment is performed by

B.SQH Construction

In this subsection, we consider the SQH construction task with a threshold constraint. Inspired by characteristics of the wavelet coefficients, we focus on the mean of wavelet coefficients (MWC) histogram by taking the following two properties into account : 1) it is designed in high pass sub band of wavelets decomposition to which HVS is less sensitive, leading to high invisibility of watermarked images and 2) it has almost a zero mean and Laplacian like distribution based on the experimental study of wavelet high pass is stable for different images. In particular, an MWC histogram is generated based on the following procedure.

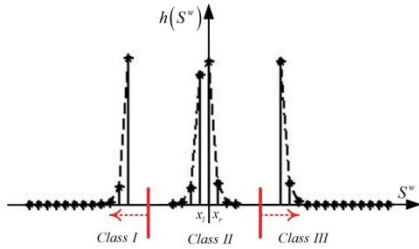
Consider a given host image I , we first decompose I using 5/3 IWT to obtain the sub band then divide into n non overlapping blocks. Let $S = [S_1, \dots, S_k, \dots, S_n]$ be the MWC's in the sub band, then the MWC of the k th blocks, S_k is defined as

To consider the MWC histogram our concern is the possibility of utilizing the blocks of interest in a sub band which will be helpful of simplifying the embedding process . In view of the histogram distribution of MWC, only the peak and its neighbor in the histogram are mostly useful for the embedding task . Therefore, a threshold constraint is applied to the block to retain those of interest ,each of which satisfies the following condition where $d(\cdot)$ computes the Euclidean distance of two elements $x \in \{x_l, x_r\}$ represents the aforementioned two peak points and δ is a predefined constant for threshold control. When $\delta \geq \max \{d(x_l, \min(S)), d(x_r, \max(S))\}$, all of the blocks will be retained for embedding, which is a special case of this constraint. Moreover, with the help of the threshold constraint the capacity can be controlled flexibly.

C. EPWM-Based Embedding:

It has been well acknowledged that a balance between invisibility and robustness is important for robust watermarking methods. Although many efforts have been made to design lossless embedding models, little progress has been made in this trade-off. Therefore, we develop EPWM to tackle this problem by utilizing the JND thresholds of wavelet coefficients to optimize watermark strength. In view of the disadvantages of PWM discussed in Section II, EPWM focuses on improving the local sensitivity of images to noise by mainly estimating brightness and texture sensitivities in a more precise way

Motivated by the benefits of luminance masking [34], we first redefine the brightness sensitivity in by calculating the luminance masking of the low-pass sub-band at resolution level ρ , denoted as



D. Extraction Based on k-Means Clustering:

If watermarked images are transmitted through an ideal channel, we can directly adopt the inverse operation of (19) to recover host images and watermarks. However, in the real environment, degradation may be imposed on watermarked images due to unintentional attacks, e.g., lossy compression and random noise. Therefore, it is essential to find an effective watermark extraction algorithm so that it can resist unintentional attacks in the lossy environment. Based on the aforementioned embedding model in (19), the MWC histogram of watermarked images are divided into three parts shown in Fig. 3, in which the center part corresponds to watermark bit “0” and others to bit “1.” To extract the embedded watermarks, the key issue is to partition these parts dynamically. In the lossy environment, this is very difficult because the histogram distribution of MWC is destroyed by unintentional attacks, as reported in [14]. In this paper, by investigating the effects of unintentional attacks on histogram we treat



Particularly; the initial cluster centers are given by considering the features of the embedding process. e.g. $F = \{\tau \min(S^w), 0, R_{\max}(S^w)\}$ for $\mu = 3$, to improve the efficiency of classification. Based on the results of classification, the embedded watermarks can be extracted by

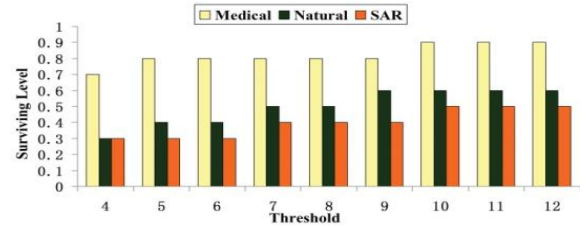
the partition as a clustering problem with a certain number of clusters and adopt k-means clustering algorithm [37], [38] to tackle this problem for simplicity. Similar to the embedding process, we first decompose the watermarked image with 5/3 IWT and construct the MWC histogram by calculating the MWCs of blocks of interest in the sub-band c^{HL} . Let $S^w = S_1^w \dots S_m^w$ the obtained MWCs, $F = f_1, \dots, f_\mu$ be the cluster centers, and $g = g_1, \dots, g_\mu$ be the set of clusters, be the set of clusters of clusters. The above classification process is summarized in Table V.

EXPERIMENTS FOR EPWM

To verify the effectiveness of the proposed EPWM, we compare it with the PWM [29] and the improved PWM (IPWM) compare it with the PWM [29] and the improved PWM(IPWM), e.g., *Lena*, *Barbara*, *Boat*, and 29 undistorted images from the laboratory for image and video engineering (LIVE) II database [39]. In our experiments, the wavelet transform with the “Daubechies” filter with three taps is applied to decompose a given host image I into high- and low-pass sub-bands. Based on this, the JND thresholds of different sub-bands can be computed with (18). For the sake of simplicity, Fig. 4 shows the examples of different masking models in which ϑ is 0.5 in EPWM and the JND thresholds of wavelet coefficients in the sub-band c_0^{LL} are used for comparison. From Fig. 4, we can see that the proposed EPWM estimates the brightness and texture sensitivities with greater precision than PWM and IPWM. To further demonstrate the superiority of EPWM for different applications, we conducted two groups of experiments based on the suggestions in [40], in which the applications in noise shaping and image quality assessment were considered, respectively. To be specific, the noise shaping experiment compares the amount of noise hidden into an image to evaluate the accuracy of masking models, in which the noise is determined by the JND thresholds of different masking models. The experiment of image quality assessment performs the evaluation by applying the masking models under comparison to an objective image quality assessment method and comparing the consistency of the obtained objective scores with the given subjective viewing ones. In the following subsections, we will make an empirical study of the experimental results.

A. Results of Noise Shaping

As discussed in [40], a better masking model should be able to hide more noise into an image at a certain level of perceptual quality. Therefore, we carried out this experiment by hiding the noise determined by JND thresholds of different masking models into a given host image I . In particular, we first decomposed I with wavelet transform and embedded a random sequence $u \in \{1, -1\}$ into the sub-band c_0^{HL} by where $JND_0^{HL}(i, j)$ was calculated by PWM, IPWM, and EPWM, respectively. Afterward, the wavelet reconstruction was used to obtain the noised image. Based on this, we compared the host image I and its noised variation $\hat{I}$ from two aspects, human observers’ observation and peak signal-to-noise ratio (PSNR). When a human observer cannot discern significant quality difference among the noised images obtained by the aforementioned three masking models, we consider these noised images to have the same level of perceptual quality. In this case, we compare the PSNRs of the noised image versus the host one to evaluate the amount of noise. The lower the PSNR is, the greater the amount of noise is, and thus the more precise the masking model is. Fig. 5 illustrates the examples of the images noised by three masking models, wherein the tuning parameter ϑ equals 0.8 in EPWM. As shown, these noised images have the same level of perceptual quality since no



significant quality difference can be observed. The corresponding results of PSNR on the synthetic dataset are shown in Tables VIII and IX. On average, it can be that EPWM outperforms PWM by 0.46 dB and IPWM by 1.07



dB based on standard images, and it is superior to PWM by 0.85 dB and to IPWM by 1.1 dB based on LIVE II database, respectively. Therefore, we can conclude that EPWM is more precise than PWM and IPWM. Another merit of EPWM is that it can regulate the noise energy with the help of the tuning parameter ϑ , as shown in Fig. 6. With the increase of ϑ , the PSNR is decreased a little, so more noise can be hidden into the given image. Based on the empirical study, it is suggested that the tuning parameter be from 0.4 to 0.8 in view of the good trade-off between the noise energy and JND thresholds.

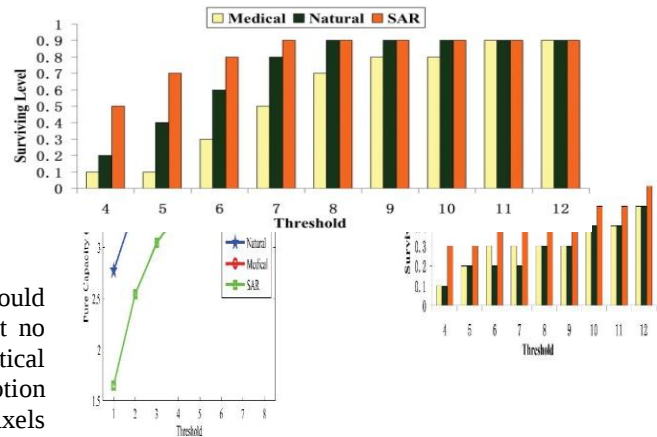
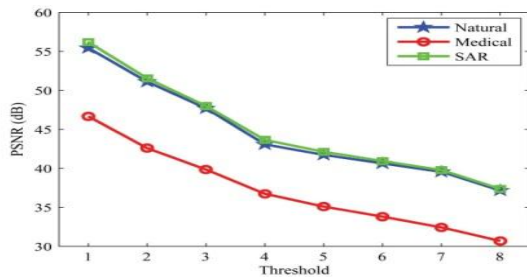
A. Parameter Sensitivity Analysis:

In this experiment, we study how free parameters affect the performance of WSQH-SC and how to set suitable parameters to achieve a reasonable result for different purposes. In the proposed framework, we have three free parameters, i.e., block size $h \times w$, threshold δ and watermark strength λ . Because $\geq \delta$ and reversibility is independent of the aforementioned three parameters, we take a simple case, i.e., $\lambda = \delta$, into account, and thus pay more attention to the effects of threshold and block size on capacity, invisibility, and robustness.

1) Threshold:

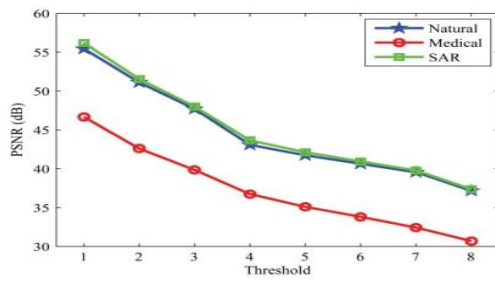
In this part, we first investigate the effects of threshold on capacity, invisibility and robustness, respectively. Then, the optimization of the threshold is discussed based on the proposed EPWM to make a trade-off between invisibility and robustness. Given the host image I with the size of $2M \times 2N$, the pure capacity $\mathbf{C}$ can be computed by set. In view of the Laplacian-like distribution of MWC, we can see that the number of S_k satisfying (28) is increased with the increase of δ for a given block size $h \times w$, which means

the capacity is improved. Fig. 7 gives the average experimental results of three kinds of images for the block size of 8×8 and threshold δ from 1 to 8 with a step of 1. It can also be seen that medical images achieve the highest capacity because neighboring pixels in medical images are more closely correlated than those in other images and thus more S_k s are included for the same threshold. Given the block size $h \times w$ and wavelet sub-band c_0^H , we study the influence of δ on invisibility. With Table IV, we can equivalently rewrite MSE by in which C is the capacity for a given δ . With (25) and (29), it can be seen that PSNR drops with $C\delta^2$ becoming ever larger. For empirical justification, the experiments are conducted in the same experimental



environment and the results are shown in Fig. 8. It should be noted that (29) is derived on the assumption that no pixels are changed by PIPA in the first step. In practical scenarios, natural and SAR images satisfy this assumption averagely, however, for medical images, many pixels need to be changed. As a result, the PSNR of medical images is obviously lower than that of others. To examine the effects of δ on robustness, Figs. 9–11 report the average robustness against the three kinds of unintentional attacks, respectively, where the block size is 8×8 , and δ changes from 4 to 12 with a step of 1. It can be seen that the robustness is gradually increased with the increase of δ . In particular, the robustness against JPEG compression for SAR images is increased more greatly than that for other images, up to 0.5 when $\delta = 9$. As for JPEG2000 compression, the robustness for medical images is the highest, up to 0.8 when $\delta = 5$, while the robustness for natural and SAR images are 0.4 and 0.3 for the same δ , higher than that against JPEG compression. By contrast, a stronger robustness against AGN is obtained for natural, medical and SAR images. For example, it is as high as 0.9 for natural and SAR images when $\delta = 8$.

According to the above results, we can see that the robustness increases with the threshold becoming ever larger while the invisibility drops at the same time. Therefore, it is of great importance to make a trade-off between them in practical applications. In this paper, we develop EPWM to serve this purpose by optimizing watermark

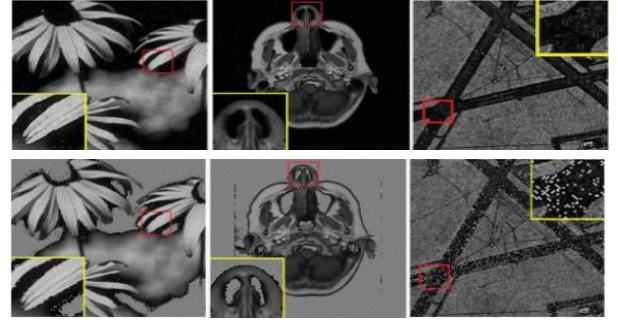


statistical average results for the block size of 8×8 are shown in Table XII, which show that EPWM applied to the proposed WSQH-SC can perform well for natural, medical, and SAR images with good robustness and acceptable invisibility. 2) **Block Size:** Besides the threshold, the block size is another important parameter affecting capacity, invisibility, and robustness. In our experiments, we will make an analysis of its effects, wherein the block size is specified from 4×4 to 16×16 with a step of 4, and the threshold is 8 for both capacity and invisibility tests and is 16 for the robustness test. According to (28), the pure capacity C is decreased with the increase of block size for a fixed threshold. This is because the number of blocks satisfying the threshold constraint is reduced. Fig. 12 presents the experimental results for natural, medical, and SAR images, which demonstrates that the smaller the block is, the higher the capacity is, and vice versa. Given a fixed threshold δ , suppose all blocks are used for watermark embedding, i.e., $C = M/h_- \cdot N/w_-$. We study the effect of block size $h \times w$ on invisibility.

PSNR becomes smaller with the increase of h and w . Fig. 13 provides the experimental justification, where medical images have the lowest PSNR due to the influence of PIPA.

Table XIII illustrates the effect of the block size on robustness, from which we can see that the robustness is improved to some extent with the increase of the block size. For each kind of images, the robustness against JPEG2000 and AGN remains unchanged for block sizes of 8×8 to 16×16 , while the robustness against JPEG compression is increased by up to 0.7 with the block size

increasing from 4×4 to 12×12 .



B. Performance Evaluation:

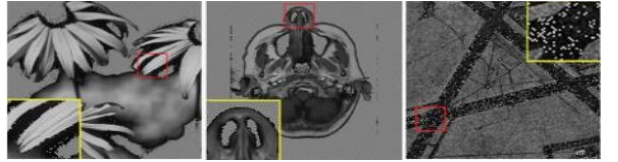
In this section, we evaluate the proposed framework, i.e., WSQH-SC in comparison with the three classical RRW methods including HR [11], HDC (1) [12], and HDC (2) [13] in terms of robustness, reversibility, invisibility, capacity, and run-time complexity. For the sake of fair comparison, we also apply these compared methods to the aforementioned 300 test images in the same experimental environment. In particular, BCH (15, 11, 1) is used in HDC (1) and HDC (2).

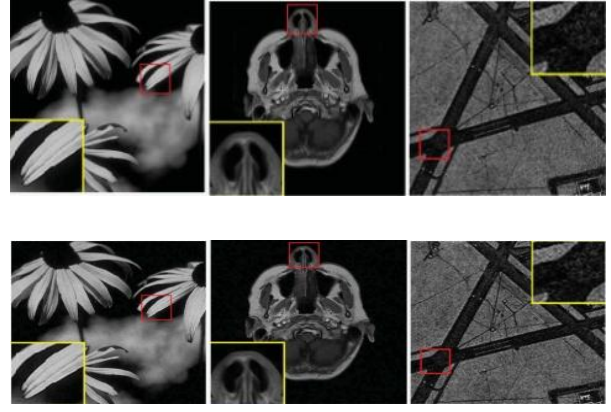
1. **Robustness:** We compare the robustness of the afore-mentioned methods when the block size is 8×8 and the watermark strength is 16: 1) the embedding level is 8 in HR;

$K = 8$ in HDC (1); and 3) $\lambda = 16$ in WSQH-SC. The strength in HDC (2) is $T + 1$. Refer to [12] and [13] for the definitions of K and T . In addition, by taking the effect of major voting on robustness into account, we repeatedly embedded the watermarks with 100 bits into host images for spatial-domain methods, i.e., HR and HDC (1) and embedded the watermarks with 10 bits for wavelet-domain ones, i.e., HDC (2) and WSQH-SC.

Table XIV reports the experimental results, from which we conclude that the proposed WSQH-SC is superior to the other methods. In particular, we observe that: 1) WSQH-SC has higher robustness against JPEG2000 compression than HR by 0.5 for medical images; 2) WSQH-SC outperforms HDC (1) greatly with the robustness against AGN; and 3) the performance of HDC (2) is poor and the robustness against JPEG compression is as low as 0 for medical images, much worse than that of WSQH-SC.

(2) **Reversibility:** In this part, we use the aforementioned IER to evaluate reversibility, which shows whether host images and watermarks can be recovered without distortion in the lossless environment. The experimental results are presented in Table XV when the block size is fixed at 8×8 and the threshold δ equals 4 and 8. As shown, HR fails to recover a few of the host images due to both overflow and underflow of pixels. Almost all of the host images cannot be recovered using HDC (1), as analyzed in [45]. In this case, the extracted watermarks by HDC (1) may contain errors. HDC (2) also performs poorly, especially for medical images, which is related to insufficient consideration of the effects of changing wavelet coefficients on pixels. Therefore, a conclusion can be drawn that the

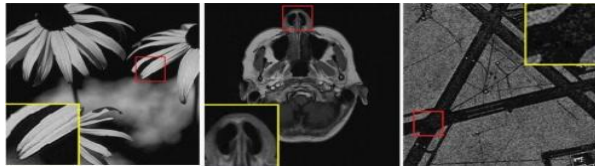




proposed WSQH-SC achieves better reversibility in comparison with others. It should be noted that the results of HDC (2) are the same because the watermark strength is determined by host images and irrelevant to the given thresholds.

3) **Invisibility:** Invisibility is used to evaluate the distortion of watermarked images versus host ones. Fig. 14 illustrates some examples of watermarked images obtained by the compared methods, where the regions of interest are zoomed to show the difference clearly. In this experiment, the block size is 4×4 , and the watermark strength is 8 for HR, HDC (1) and WSQH-SC. It shows that “salt-and-pepper” noise greatly degrades the visual quality of the watermarked images in HR, and results in the worst PSNRs for the three test images. From the perspective of human observers, WSQH-SC achieves a similar visual quality to HDC (1) and a slightly better visual quality than HDC (2) due to the blocking effect in HDC (2). The objective comparison suggests that the PSNR of WSQH-SC for the SAR image is higher than that of the others, and that for the medical image is lower because of PIPA. In summary, the invisibility of WSQH-SC is superior to HR and comparable to HDC (1) and HDC (2) in most cases.

Capacity: To compare the capacity of WSQH-SC with the above three methods, we calculated their pure capacities based on the 300 test images. The average results are illustrated in Table XVI, wherein the block size is 8×8 and the threshold is 8 in WSQH-SC. For convenience, we classify the methods into two categories, i.e., those in spatial- and wavelet-domains. It can be seen that the spatial-domain methods achieve higher capacity than those developed in the wavelet domain. This is because these methods employ the whole host image to embed watermarks while the latter uses only one sub-band. Based on the comparison between the wavelet-domain methods, the capacity of the WSQH-SC is higher than that of HDC (2) because of the usage of error correction coding in HDC (2).



Run-Time Complexity:

We study the run-time complexity of WSQH-SC and compare it with that of the aforementioned methods. To further verify the effects of the k -means clustering algorithm on the run-time, we construct a new method, termed WSQH-S, which has the same embedding process as WSQH-SC and employs the inverse operation of (19) for watermark extraction. To fairly compare different methods, we carry out this experiment on the same computer with

a 2.93-GHz Intel(R) Core(TM) i3 CPU and 2 GB memory. We use tic and toc commands in MATLAB to record the time cost in seconds. Table XVII shows the average results of the 300 test images, wherein the parameters are the same as those for the robustness test. In addition, the JPEG compression with the quality factor of 100 is used here. From the results, we observe that the proposed WSQH-SC outperforms the classical RRW methods. And its run-time of watermark extraction is slightly longer than that of WSQH-S due to the usage of the k -means clustering algorithm. In particular, HDC (1) and HDC (2) are much slower than WSQH-SC due to the Arnold permutation. HR computes the centroids of two zones in a block, and results in a slightly longer run-time. In summary, the experimental results show that: 1) the WSQH-SC is superior to the classical RRW methods in terms of run-time complexity and 2) it is worthwhile to improve the robustness with a low run-time cost.

VI. CONCLUSION

In this paper, we have developed a novel yet pragmatic framework for RRW. It includes carefully designed PIPA, SQH shifting and clustering, and EPWM, each of which handles a specific problem in RRW. PIPA preprocesses host images by adjusting the pixels into a reliable range for satisfactory reversibility. SQH shifting and clustering constructs new watermark embedding and extraction processes for good robustness and low run-time complexity. EPWM precisely estimates the local sensitivity of HVS and adaptively optimizes the watermark strength for a trade-off between robustness and invisibility. In contrast to representative methods, thorough experimental results on natural, medical and SAR images demonstrate that the proposed framework: 1) obtains comprehensive performance in terms of reversibility, robustness, invisibility, capacity and run-time complexity; 2) is widely applicable to different kinds of images; and 3) is readily applicable in practice. In future, we will combine the proposed framework with the local feature [46]–[48] to further improve robustness. In addition, it is valuable to integrate the merits of sparse representation [49] and probabilistic graphical model [50] into the designing of image watermarking.

REFERENCE

1. J. Fridrich, M. Goljan, and R. Du, “Lossless data embedding-new paradigm in digital watermarking,” *EURASIP*
2. *J. Appl. Signal Process.*, vol. 2002, no. 1, pp. 185–196, Jan. 2002.
3. Z. Zhao, N. Yu, and X. Li, “A novel video watermarking scheme in compression domain based on fast motion estimation,” in *Proc. Int. Conf. Commun. Technol.*, vol. 2, 2003, pp. 1878–1882.
4. X. Li, “Watermarking in secure image retrieval,” *Pattern Recog. Lett.*, vol. 24, no. 14, pp. 2431–2434, Oct. 2003.
5. M. Celik, G. Sharma, A. Tekalp, and E. Saber, “Lossless generalized-LSB data embedding,” *IEEE Trans. Image Process.*, vol. 14, no. 2, pp. 253–266, Feb. 2005.
6. J. Tian, “Reversible watermarking using a difference expansion,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 8, pp. 890–896, Aug. 2003.
7. M. Van der Veen, F. Bruekers, A. Van Leest, and S. Cavin, “High capacity reversible watermarking for audio,” *Proc. SPIE*, vol. 5020, no. 1, pp. 1–11, Jan. 2003.
8. R. Du and J. Fridrich, “Lossless authentication of MPEG-2 video,” in *Proc. IEEE Int. Conf. Image Process.*, vol. 2, Dec. 2002, pp. 893–896.
9. J. Dittmann and O. Benedens, “Invertible authentication for 3-D-meshes,” *Proc. SPIE*, vol. 5020, no. 653, pp. 653–664, Jan. 2003.
11. L. An, X. Gao, C. Deng, and F. Ji, “Robust lossless data hiding: Analysis and evaluation,” in *Proc. Int. Conf. High Perform. Comput. Simul.*, 2010, pp. 512–516.
12. C. De Vleeschouwer, J. Delaigle, and B. Macq, “Circular interpretation of histogram for reversible watermarking,” in *Proc. IEEE Workshop Multimedia Signal Process.*, 2001, pp. 345–350.
13. C. De Vleeschouwer, J. Delaigle, and B. Macq, “Circular interpretation of bijective transformations in lossless watermarking for media asset management,” *IEEE Trans. Multimedia*, vol. 5, no. 1, pp. 97–105, Mar. 2003.
14. Z. Ni, Y. Shi, N. Ansari, W. Su, Q. Sun, and X. Lin, “Robust lossless image data hiding designed for semi-fragile image authentication,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 4, pp. 497–509, Apr. 2008.
15. D. Zou, Y. Shi, Z. Ni, and W. Su, “A semi-fragile lossless digital watermarking scheme based on integer wavelet transform,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 10, pp. 1294–1300, Oct. 2006.
16. X. Gao, L. An, Y. Yuan, D. Tao, and X. Li, “Lossless data embedding using generalized statistical quantity histogram,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 8, pp. 1061–1070, Aug. 2011.
17. N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in *Proc. IEEE Conf. Comput. Vision Pattern Recog.*, vol. 1, 2005, pp. 886–893.
18. L. Fei-Fei and P. Perona, “A Bayesian hierarchical model for learning natural scene categories,” in *Proc. IEEE Conf. Comput. Vision Pattern Recog.*, vol. 2, Jun. 2005, pp. 524–531.
19. S. Zhang, Q. Tian, G. Hua, Q. Huang, and S. Li, “Descriptive visual words and visual phrases for image applications,” in *Proc. ACM Int. Conf. Multimedia*, 2009, pp. 75–84.
20. C. Deng, X. Gao, X. Li, and D. Tao, “Local histogram based geometric invariant image watermarking,” *Signal Process.*, vol. 90, no. 12, pp. 3256–3264, Dec. 2010.