A Novel Lossless Robust Reversible Watermarking Method for Copyright Protection of Images

Sidham Abhilash¹, S M Shamseerdula²
¹ M.Tech (CSP), GPREC, Kurnool.
² Asst Professor, GPREC, Kurnool.

Abstract
Robust reversible watermarking (RRW) methods are popular in multimedia for protecting copyright, while preserving high 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 are used to embed watermarks, e.g., secret information, 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, e.g., medical and military images. Although researchers proposed some RW methods for various media, e.g., images, audios, videos, and 3-D meshes; 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. Recently, a dozen of RRW methods for digital images have been proposed, which can be classified into two groups: histogram rotation (HR)-based methods and histogram distribution constrained (HDC) methods.

The HR-based methods, accomplish robust loss-less embedding by slightly rotating the cancooids 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, 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. 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.

3) **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 SQH Shifting and Clustering, but also duly introduce perceptual characteristics of HVS to the RRW framework; forming a novel yet pragmatic RRW framework.

**II. RELATED WORKS**

In this section, we briefly introduce the GSQH driven method 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, bag-of-words, and digital watermarking. For RW methods, SQH has recently received considerable attention due to stability and simplicity, e.g., arithmetic average of difference (AAD) histogram, difference histogram, and prediction error histogram. In particular, we proposed a GSQH driven method, 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 \( n^t \) 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., \( X = \{ (x_1, n_1), \ldots, (x_i, n_i), \ldots, (x_m^*, n_m^*) \} \), where \( x_i \) represents the different values of the AAD, and \( n_i \) is the corresponding frequency of \( x_i \) in SQH. Let \( x_s \) and \( x_r \) be the two peaks points of SQH, wherein

\[
\begin{align*}
 r &= \arg \max_{i} \{ i \} \quad \text{and} \\
 i &= \arg \max_{i} \{ m \} \quad \text{and} \\
 i &= \arg \min_{i} \{ \leq \} \\
 i &= \arg \min_{i} \{ \leq \} \\
 &\text{Suppose } x_i \leq x_r \text{ , then the embedding is done according to} \\
 &s_l = x_i - z, \quad \text{if } s_l < x_i - z \\
 &w = s_l - b_2 (z + 1), \quad \text{if } x_i - z \leq s_l \leq x_i \\
 &s_k = 0, \quad \text{if } x_i < s_k < x_s \\
 &s_l + b_2 (z + 1), \quad \text{if } x_r \leq s_k \leq x_s + z \\
 &s_k + z + 1, \quad \text{if } s_k > x_s + z \\
 \end{align*}
\]

in which \( s_l \) and \( s_k \) are the AADs of the 4th block in the host and watermarked images, respectively, \( b_2 \in \{ 0, 1 \} \) is the 4th watermark bit, and \( z \geq 0 \) is a scale factor. Correspondingly, the watermarking extraction 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 briefly 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 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.
is defined as

\[ r = \gamma, \quad 1 \leq k \leq w, \quad 1 \leq s_k \leq \gamma + 1 \]

Where \( b_i' \) is the 4th 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 block. Here, \( _- (\rho, o), \_ (\rho, i, j) \) 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 \( c_{\rho^o} (i, j) \) at \( (i, j) \) in the sub-band \( c_{\rho}^o \) with resolution level \( \rho \in \{0, 1, 2, 3\} \) and orientation \( o \in \{L, H, LL, HL, HH\} \), the JND threshold is denoted by \( JND_{\rho^o} (i, j) = _- (\rho, o) \_ (\rho, i, j) \_ (\rho, i, j)^{\rho^o}. \) (5)

\[ _- (\rho, o) = \sqrt{2}, \quad \text{if } o = HH \]
\[ = 1, \quad \text{otherwise} \]

and

\[ _- (\rho, i, j) = 1 - (\rho, i, j), \quad \text{if } (\rho, i, j) < 0.5 \]
\[ = 0.1, \quad \text{if } (\rho, i, j) < 0.5 \]

\[ = 1 + (\rho, i, j), \quad \text{otherwise} \]

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., \( c_{3}^{HH} \), 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) and the blocks of interest in the sub-band \( c_{3}^{HH} \) 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.
A. PIPA

In RRW, how to handle both overflow and underflow of pixels is important for reversibility. Xuan et al. 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. Based on this, we investigate the effects of changing wavelet coefficients on pixels from two aspects.

For convenience, we denote the wavelet coefficient $c_{0}^{\text{ox}}(i, j)$ in the sub-band $c_{0}^{\text{ox}}$ as $c_{i, j}^{\text{ox}}$, wherein $1 \leq i \leq M$, $1 \leq j \leq N$, and $M \times N$ is the size of $c_{0}^{\text{ox}}$.

1) Single Sub-Band and Single Wavelet Coefficient:

Given the watermark strength $\lambda$, we consider the changes of pixels when an arbitrary wavelet coefficient in $c_{0}^{\text{ox}}$ is changed. In particular, if $c_{i, j}^{\text{ox}} \leftarrow c_{i, j}^{\text{ox}} + \lambda$, 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 $c_{0}^L$, $c_{0}^H$, and $c_{0}^{HH}$ sub-bands are both increased and decreased. Based on this,
two neighboring blocks in $c_0^{L\downarrow}$ and $c_0^{L\uparrow}$ 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 $\lambda$. 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 visibility 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 $S = [S_1, \ldots, S_k, \ldots, S_n]$ be the MWC’s in the sub band, then the MWC of the $k$th blocks, $S_k$ is defined as:

$$S_k = \{x \in \{x_k, x_r\} \mid \text{represents the aforementioned two peak points and } \delta \text{ is a predefined constant for threshold control. When } \delta \geq \max \{d(x_k, \min(S)), d(x_r, \max(S))\}, \text{ 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 $\rho$, denoted as

$$D. Extraction:**

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 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 the partition as a clustering problem with a certain number of clusters and adopt k-means clustering algorithm 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 of the obtained MWCs, $F = \{f_1, \ldots, f_m\}$ be the set of clusters, and $g_1$, $g_2$, $\ldots$, $g_s$ be the set of clusters, be the set of clusters of clusters. The above classification process is summarized in Table V. Particularly, the initial cluster centers are given by considering the features of the embedding process, e.g., $F = \{r \min(S'), 0, R_{\text{med}}(S')\}$ for $\mu = 3$, to improve the efficiency of classification. Based on the results of classification, the embedded watermarks can be extracted by

**IV. Result**

The watermarking performed with the RRW method is more robust than the bit-plane slicing, as it withstands nearly all attacks. Cropping was the most detrimental to the extracted watermark, while image enhancement (Unsharp filter) provided the best extracted watermark. Bit-plane slicing was extremely vulnerable to any sort of smoothing or averaging filters, as the watermark failed to be appropriately discernable after these attacks in all cases. Bit-plane slicing responded the best to scalar changes and performed acceptably when attacked with cropping.
While not as robust as the RRW watermarking method, bit-plane slicing performs acceptably as a fragile watermark as long as smoothing or averaging is not involved. It does allow one to discern how much an image has been tampered with to some degree, but only in approximately half of the attacks. Table 1. Results of various morphological attacks on the images in Figure. Images that was watermarked and attacked. Experimental results of the RRW method can be seen in the Figure.

V. 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 to further improve
robustness. In addition, it is valuable to integrate the merits of sparse representation and probabilistic graphical model [50] into the designing of image watermarking.

REFERENCE
