A Robust Reversible Watermarking Scheme Using Attack-Simulation-Based Adaptive Normalization and Embedding

Yichao Tang, Chuntao Wang, *Member, IEEE*, Shijun Xiang, *Member, IEEE*, and Yiu-Ming Cheung, *Fellow, IEEE*

***Abstract*— For copyright protection and perfect recovery of the original image in case of no attacks, it is necessary to develop robust reversible watermarking (RRW) methods that counteract both common signal processing (CSP) and geometric deformation (GD) attacks (RRW-CG). However, to the best of our knowledge, none of the existing RRW methods exploit target attacks as prior knowledge to improve their robustness and embedding capacity. To this end, we propose a two-stage RRW-CG scheme with attack-simulation-based adaptive normalization and embedding. Specifically, the polar harmonic transform (PHT) moments are taken as watermark carriers, and their stability with respect to target attacks is evaluated by performing attack simulation tests on large-scale images. This enables the adaptive normalization of PHT moments to improve the watermark robustness. The PHT moments with high stability are then chosen as watermark carriers, and the conventional spread transform dither modulation (STDM) with one quantization level is optimized to form the enhanced version with multiple quantization levels, in which the embedding strength is determined adaptively via attack simulation tests on the candidate watermarked image. This in turn improves the watermark robustness and increases the embedding capacity. After the robust watermark has been embedded, errors caused by robust watermarking are used as the auxiliary information and then inserted into the robustly watermarked image via the recursive code-based reversible watermarking technique, ensuring the reversibility in case of no attacks. Extensive experimental simulation results show that the proposed scheme outperforms the state-of-the-art RRW methods in terms of robustness against CSP such as AWGN, JPEG, JPEG2000, mean filtering, and median filtering as well as GD including rotation and scaling under the same invisibility, reversibility, and embedding capacity. This indicates that, by exploiting target attacks as prior knowledge and designing the attack-simulation-based adaptive normalization and embedding, the proposed novel RRW is feasible and effective.**

***Index Terms*— Robust reversible watermarking, two-stage embedding, geometric deformation, polar harmonic transform, attack simulation.**

Manuscript received 17 August 2023; revised 24 January 2024; accepted 26 February 2024. Date of publication 4 March 2024; date of current version 2 May 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 62172165, Grant 62272197, and Grant 62062044; in part by the Natural Science Foundation of Guangdong Province under Grant 2022A1515010325; in part by Guangzhou Basic and Applied Basic Research Project under Grant 202201010742; in part by NSFC/Research Grants Council (RGC) Joint Research Scheme under Grant N_HKBU214/21; in part by the General Research Fund of RGC under Grant 12201321, Grant 12202622, and Grant 12201323; and in part by RGC Senior Research Fellow Scheme under Grant SRFS2324-2S02. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Roberto Caldelli. *(Corresponding author: Chuntao Wang.)*

Yichao Tang and Chuntao Wang are with the College of Mathematics and Informatics, South China Agricultural University, Guangzhou 510642, China, also with the Key Laboratory of Smart Agricultural Technology in Tropical South China, Ministry of Agriculture and Rural Affairs, Guangzhou 510642, China, and also with Guangzhou Key Laboratory of Intelligent Agriculture, Guangzhou 510642, China (e-mail: yichao_tang@foxmail.com; wangct@scau.edu.cn).

Shijun Xiang is with the School of Information Science and Technology, Jinan University, Guangzhou 510632, China (e-mail: shijun_xiang@qq.com).

Yiu-Ming Cheung is with the Department of Computer Science, Hong Kong Baptist University, Hong Kong (e-mail: ymc@comp.hkbu.edu.hk).

Digital Object Identifier 10.1109/TIFS.2024.3372811

I. Introduction

With the widespread applications and dissemination of digital images, it is essential to protect their copyright and integrity. Digital watermarking is a feasible solution to this problem, which embeds information into an image to achieve objectives such as copyright protection [1], [2], covert communication [3], [4], integrity authentication [5], [6], [7], [8], and so on. Depending on the robustness against attacks and the reversibility of the embedding process, digital watermarking techniques are divided into three categories: robust watermarking, reversible watermarking, and robust reversible watermarking (RRW).

Robust watermarking [9], [10], [11], [12] resists various attacks but causes permanent distortions of the original image; it is applied for scenarios in which the embedded watermark needs to be preserved, e.g., commercial photography and illustrations. Reversible watermarking [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26] can recover the original image from the watermarked one only if no attack has imposed on the watermarked image; it is appropriate for scenarios in which the original image needs to be recovered, including medical images and remote sensing images.

RRW integrates both the robust and reversible watermarking techniques. Specifically, RRW recovers the original image and extracts the watermark perfectly if the watermarked image has not been attacked; otherwise, only the watermark can be extracted. This makes it more adaptable for scenarios that require both copyright protection and perfect recovery of the original image, especially in the situations of digital artworks and high-fidelity images. According to their resistance to different types of attacks, RRW can be classified into two categories: RRW that resists common signal processing (CSP) attacks (RRW-CSP) [27], [28], [29], and RRW that counteracts both CSP and geometric deformation (GD) attacks (RRW-CG) [30], [31], [32], [33], [34], [35], [36]. RRW-CSP is able to resist CSP such as noises, compression, and filtering, but it

fails to withstand GD such as rotation or scaling. By contrast, RRW-CG can counteract both CSP and GD, but it leads to a lower embedding capacity or more distortions.

In the early study on RRW-CSP, De Vleeschouwer et al. [37] proposed a histogram rotation technique, which randomly divides pixels in each embedding block into two groups, maps them to the corresponding circular histograms, and finally performs watermark embedding by rotating these circular histograms. Although this method is resistant to JPEG, it leads to overflow/underflow problem. To address this issue, Ni et al. [38] presented another RRW-CSP that classifies the embedding blocks according to the grayscale value distribution in each block and assigns corresponding embedding strategies, which achieves robustness against JPEG and JPEG2000. In recent years, many RRW-CSP methods have been developed in the spatial [39], transformed [40], and miscellaneous domains [41], [42]. Recently, Kumar and Jung [28] constructed a two-layer RRW-CSP that decomposes an image into high- and low-significant planes to embed the watermark for copyright protection and the auxiliary information for the original image recovery, respectively. As JPEG typically affects the least-significant plane, this method well resists JPEG. However, this method exhibits weak robustness against noise attacks. Besides, Liang et al. [29] implemented an RRW-CSP for encrypted images by leveraging the homomorphic multiplication property of the Paillier cryptosystem to shift the statistical histogram in the encrypted domain for watermark embedding. This method is robust to additive white Gaussian noise (AWGN), JPEG, and JPEG2000, but it yields the watermarked images with undesirable visual quality.

Compared with RRW-CSP that only resists CSP, RRW-CG can counteract both CSP and GD, which makes it more adaptable. The first RRW-CG was proposed by Chrysochos et al. [30], which pairs histogram bins of a grayscale image and then embeds the watermark by swapping the pixel values of the paired bins. Although this method resists against GD such as rotation, flipping, and cropping, it fails to resist filtering attacks and provides a low embedding capacity. Later, Chang et al. [31] developed an RRW-CG that embeds the watermark by modifying differences of the discrete cosine transform (DCT) coefficients of two selected subsampling images. This method is resistant to noises, compression, rotation, and scaling, but its overall robustness is unsatisfactory. Recently, Hu and Xiang [32], [33] proposed two RRW-CG methods that use the quantization index modulation (QIM) to embed watermarks into magnitudes of Zernike moments and polar harmonic transform (PHT) moments, respectively. As magnitudes of both Zernike moments and PHT moments are highly resistant to CSP and GD, these two methods provide promising performance against both CSP and GD attacks.

The existing literature has demonstrated extensive researches on RRW-CSP, while RRW-CG has received less attention. Nevertheless, RRW-CG deserves more exploration because it can resist various types of CSP and GD attacks. Although several RRW-CG methods in the literature have achieved desirable robustness against both CSP and GD attacks, their robustness at the same capacity or vice versa can be further improved, which are essential for copyright protection and perfect recovery of digital artworks and high-fidelity images. For example, these methods do not exploit in the embedding side the common attack types as prior knowledge to further enhance the RRW-CG's performance.

To overcome these shortcomings, we therefore exploit target attacks such as CSP as the prior knowledge and propose a two-stage RRW-CG scheme using the attack-simulation-based adaptive normalization and embedding, thereby achieving the objective of high-robustness and high-capacity on the condition of invisibility and reversibility. Specifically, as PHT moments have lower computational complexity and higher numerical stability than the other geometric moments such as Zernike moments and the orthogonal Fourier-Mellin moments, this scheme adopts PHT moments as the watermark carrier to resist GD such as rotation and scaling. In addition, inspired by the spread spectrum watermarking that improves the watermark robustness by spreading each watermark bit into multiple carriers, the proposed scheme optimizes the conventional spread transform dither modulation (STDM) to perform the robust watermarking. In the STDM-based embedding process, the common target attack types and intensities are taken as prior knowledge, and the adaptive normalization and embedding are then developed to implement the robust watermarking. In more detail, the adaptive normalization determines the moment-by-moment weights by the stability of PHT moments over different simulated target attacks and implements the moment normalization via the obtained adaptive weights. The adaptive watermark embedding performs attack simulation tests on the candidate robustly watermarked image and decides the optimal embedding strength for each watermark bit based on its resistance to the simulated attacks, aiming to further improve the robustness and capacity. After the robust watermark has been inserted via the STDM-based embedding process at the first stage, a recursive code-based reversible watermarking [43] is adopted at the second stage to insert the generated auxiliary information for recovering the original image, where the recursive code is adopted for its rate-distortion bound approaching characteristics. By integrating these strategies, the proposed scheme may achieve higher robustness and larger capacity under the invisibility and reversibility conditions. Extensive experimental simulation results show that the proposed scheme resist against CSP such as AWGN, JPEG, JPEG2000, mean filtering, and median filtering, as well as GD including rotation and scaling. Compared with the existing state-of-the-art RRW methods, the proposed scheme obtains higher robustness against CSP and GD attacks under the same invisibility, reversibility, and embedding capacity. This indicates the effectiveness of the proposed scheme.

The main highlights and contributions of the proposed scheme are as follows.

1) Develop an adaptive normalization strategy. Compared with the conventional fixed normalization weight [32], [33], [35], this strategy imposes the priorly known attacks on large-scale images to assess the stability of PHT moments with different orders and repetitions and then adaptively determines the normalization weight for a PHT moment based on the assessed stability. As the prior

knowledge has been well exploited, higher robustness is expected to achieve.

2) Design an adaptive watermark embedding technique. Rather than using the conventional single-level quantizer [32], [33], [34], [35], [36] to yield the fixed watermark embedding strength, this technique implements target attacks on the candidate robustly-watermarked image and then adaptively determines the embedding strength according to the watermark detection performance of the attacked image via the elaborately developed multi-level quantizer. Consequently, this technique enhances the robustness under the same embedding distortion and reduces the amount of auxiliary information for recovering the original image.
3) Propose an RRW-CG scheme that achieves higher robustness and larger embedding capacity under the constraints of invisibility and reversibility. And it also outperforms the state-of-the-art methods.

The remainder of this paper is organized as follows. Section II reviews the existing two-stage RRW methods and introduces PHT moments and STDM techniques. Section III describes the proposed two-stage RRW-CG scheme. The experimental simulation results and analysis are presented in Section IV. Section V finally draws the conclusion.

II. Related Works

This paper presents a two-stage RRW method that employs PHT moments as the watermark carrier and the optimized STDM technology for watermark embedding. Thus, this section briefly reviews the existing two-stage RRW methods and introduces the PHT moments and STDM techniques.

A. Two-Stage RRW Methods

Coltuc and Chassery [27] proposed the first two-stage RRW method, which embeds a watermark into DCT coefficients of a grey image to generate a robustly watermarked image, and then reversibly inserts the differences between the original image and the robustly watermarked one into the robustly watermarked image. However, this method degrades the watermark robustness by directly embedding the auxiliary information into the robustly watermarked image. Wang et al. [44] addressed this issue by designing an independent-domain-based two-stage RRW that divides the original image into two independent regions of low- and high-frequencies for embedding the watermark and the auxiliary information separately. This method improves the watermark robustness, but it overlooks the effect of the carrier stability on the performance. Later, Xiong et al. [42] developed a two-stage RRW based on secure multi-party computation for multi-party copyright protection. This method embeds the watermark and the auxiliary information into the encrypted domain of an image using the patchwork and prediction-error expanding (PEE) techniques, respectively, which preserves image privacy and ensures watermark robustness. However, this method involves complex encryption and decryption operations and leads to the undesirable visual quality of watermarked images. All these two-stage RRW methods are vulnerable to GD such as rotation and scaling and thus they are classified as RRW-CSP.

Attempts to solving these problems give rise to RRW-CG. For example, by using Zernike moments and PHT moments with geometric invariance as watermark carriers, Hu and Xiang [32], [33] proposed two QIM-based RRW-CG methods. These methods are robust to rotation and scaling, but their embedding capacity and noise resistance need to be further improved. To tackle these challenges, we improved Hu's methods in our previous work [34] by using an optimized embedding strategy and a rounding error compensation, leading to higher watermark robustness and embedding capacity. However, it does not leverage common attack types as prior knowledge to further enhance the performance. Moreover, Fu et al. [35] introduced an RRW-CG method that applies median filtering to remove noises in the watermarked image. This method improves the robustness against noise attacks, but it does not achieve desirable results against other types of attacks. In brief, these mentioned methods are all robust to both CSP and GD.

B. Polar Harmonic Transform

PHT is a general name for three types of the orthogonal moments based on harmonic functions proposed by Yap et al. [45]. They are Polar Complex Exponential Transform (PCET), Polar Cosine Transform (PCT), and Polar Sine Transform (PST). These moments are defined without any complicated factorial/gamma terms and long summations, which implies that PHT has good computational complexity and numerical stability.

Suppose $f(x, y)(x, y \in [1, K])$ is an image function of size $K \times K$, which can be defined on a unit circle as $f(x_s, y_t)\left(x_s^2 + y_t^2 \leq 1\right)$, and $f(x_s, y_t)$ can be expressed as a linear combination of PHT moments M_{nl} and PHT basis functions $V_{nl}(x_s, y_t)$, i.e.,

$$f(x_s, y_t) = \sum_{n=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} M_{nl} V_{nl}(x_s, y_t) \tag{1}$$

where M_{nl} is a PHT moment of order n with repetition l. M_{nl} is computed from the inner product of $f(x_s, y_t)$ and $V_{nl}(x_s, y_t)$, i.e.,

$$M_{nl} = \sum_{s=0}^{K-1} \sum_{t=0}^{K-1} f(x_s, y_t)[V_{nl}(x_s, y_t)]^* \Delta x_s \Delta y_t \tag{2}$$

where $(\cdot)^*$ denotes the complex conjugate, $(x_s, y_t) = \left(\frac{2s+1-K}{K}, \frac{2t+1-K}{K}\right)$, and $\Delta x_s = \Delta y_t = \frac{2}{K}$.

Since M_{nl} is defined on the unit circle, $V_{nl}(x_s, y_t)$ can be decomposed into the product of the radial basis function $R_n(r)$ and the angular basis functions $A_l(\theta) = e^{il\theta}\left(i = \sqrt{-1}\right)$, i.e.,

$$V_{nl}(x_s, y_t) = R_n(r) A_l(\theta) \tag{3}$$

where $r = \sqrt{x_s^2 + y_t^2}$ and $\theta = \tan^{-1}(y_t / x_s)$.

The existing unit circle-based orthogonal moments have the same definition except for the radial basis functions [46]. The

radial basis functions of PCET, PCT, and PST are defined as

$$R_n^E(r) = \frac{1}{\sqrt{\pi}} e^{i2\pi nr^2}, \tag{4}$$

$$R_n^C(r) = \begin{cases} \frac{1}{\sqrt{\pi}} & n = 0 \\ \sqrt{\frac{2}{\pi}} cos(\pi nr^2) & n > 0 \end{cases}, \tag{5}$$

$$R_n^S(r) = \sqrt{\frac{2}{\pi}} sin(\pi nr^2). \tag{6}$$

C. Spread Transform Dither Modulation

Chen and Wornell [47] proposed QIM and STDM, which are robust watermarking techniques that quantize the dithered signals to implement the watermark embedding. STDM is an extension of QIM that improves the robustness against random noise attacks by spreading one watermark bit into multiple carriers. Instead of directly quantizing the original image coefficients, STDM first projects the pre-embedded vector x onto a random direction vector u, and then applies dither modulation on the projected scalar to embed the watermark w.

In STDM, the watermark is inserted as

$$y = x + \left(Q^w\left(x^T u, \Delta\right) - x^T u\right) \cdot u, w \in \{0, 1\} \tag{7}$$

where $(\cdot)^T$ is a transpose operation, Δ is a quantization step, $Q^w(\cdot)$ is a single-level quantizer expressed as

$$Q^w\left(x^T u, \Delta\right) = \left[\frac{x^T u + \beta(w)}{\Delta}\right]\Delta - \beta(w) \tag{8}$$

where $[\cdot]$ is a rounding function and $\beta(w)$ is a dither value corresponding to the watermark w with $\beta(1) = \beta(0) + \Delta/2$.

The embedded watermark can be extracted from the watermarked image using the minimum distance decoding, which is conducted as

$$\tilde{w} = arg \min_{w \in \{0,1\}} \left|\tilde{y}^T u - Q^w\left(\tilde{y}^T u, \Delta\right)\right| \tag{9}$$

where $\tilde{y}$ is the vector obtained from the received image, $\tilde{w}$ is the extracted watermark.

III. Proposed Two-Stage RRW-CG Scheme

This section presents the proposed two-stage RRW-CG scheme that develops the adaptive normalization and embedding techniques and optimizes the conventional STDM. The proposed scheme comprises three parts: two-stage watermark embedding, integrity authentication, and watermark extraction and image reconstruction. In the two-stage watermark embedding part, the robust watermark and the auxiliary information are embedded into the original image by the robust and reversible embedding stages, respectively. The integrity authentication part checks whether the received image has been attacked or not. The third part performs watermark extraction and image reconstruction in case of no attacks, or only extracts the watermark if the received image has been attacked. Details of these three parts are described in Sections III-A, III-B, and III-C, respectively.

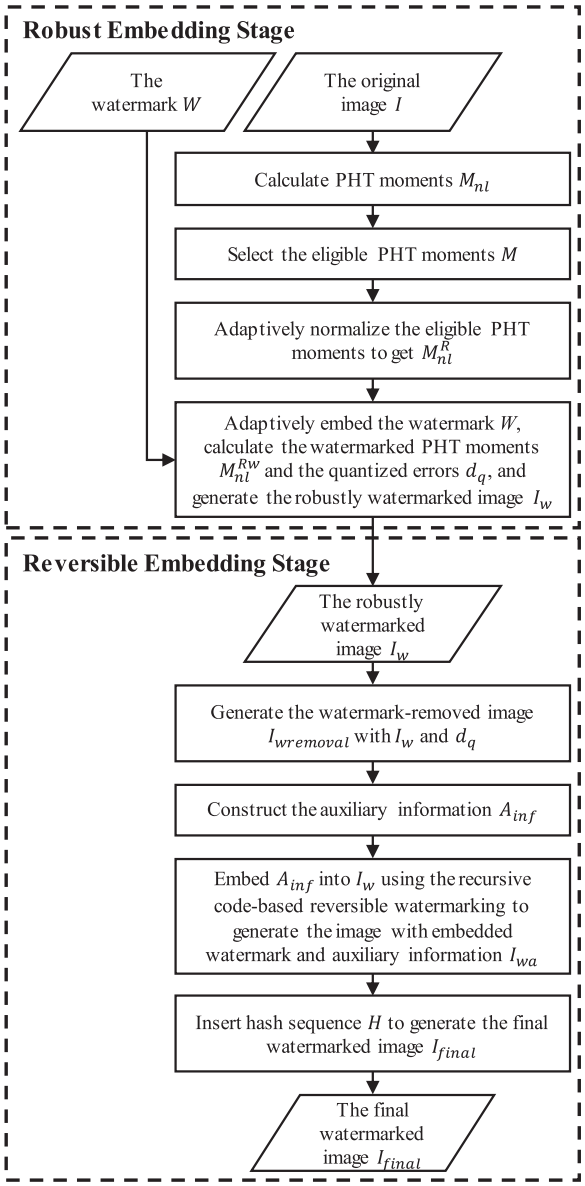


Fig. 1. Flowchart of the two-stage watermark embedding of the proposed RRW-CG scheme. M_{nl} denotes the PHT moment with order n and repetition l, M_{nl}^R represents the normalized PHT moment, and M_{nl}^{Rw} is the watermarked normalized PHT moment.

A. Two-Stage Watermark Embedding

The two-stage watermark embedding consists of the robust embedding stage and the reversible embedding stage, which embeds the robust watermark for protecting image copyright and the auxiliary information for integrity authentication, respectively. The auxiliary information consists of the quantized errors due to watermark embedding, the rounding errors from image reconstruction, and the least significant bits (LSBs) for the integrity authentication. Fig. 1 illustrates the flowchart of the proposed two-stage watermark embedding process, and details are described below.

1) Robust Embedding Stage: The robust embedding stage includes four portions: PHT moment calculation, PHT moment selection, the adaptive moment normalization, and the adaptive watermark embedding. Among them, both the adaptive moment normalization and watermark embedding use com-

mon target attacks as prior knowledge, but they apply attack simulation tests in different ways and for different purposes. The former performs attack simulation tests on large-scale images to analyze the stability of PHT moments with different orders and repetitions under target attacks and then determines adaptively the normalization weights according to the moment stability. The latter imposes attack simulation tests on each candidate robustly watermarked image and then decides adaptively the optimal embedding strength for each watermark bit according to the watermark detection performance of the attacked watermarked image. It can be expected that the more the common target attacks are applied in the attack simulations, the higher the resistance to practical attacks could be, but the larger the computational complexity would be in the robust watermark embedding. To obtain a good trade-off, we choose to select several typical attacks including AWGN, JPEG2000, and mean filtering as the target attacks in the attack simulation applied in both the adaptive normalization and watermark embedding parts. These target attacks can typically represent various CSP attacks in practical applications and thus achieve desirable performance, as will be demonstrated by the experimental simulation results in Section IV. The details of the four portions of the embedding stage are described as follows.

a) PHT Moment Calculation: Given an image I of size $K \times K$, calculate the PHT basis functions $V_{nl}(x_s, y_t)$ and the corresponding PHT moments M_{nl} by Eqs. (2), and (3), respectively.

Due to different limitations of the three PHT moments on order n and repetition l [45], this work calculates the PCET moments that satisfy $|n| \leq N$ and $|l| \leq L$, the PCT moments that meet $0 \leq n \leq N$ and $|l| \leq L$, and the PST moments with $1 \leq n \leq N$ and $|l| \leq L$, where N and L are the maximum order and repetition, respectively.

b) PHT Moment Selection: The mapping of a digital image to a unit circle introduces numerical and geometric errors in some PHT moments' calculations. To avoid the degradation of watermark robustness caused by these errors, Li et al. [48] discovered that only PHT moments that satisfy $l \neq 4m$, $m \in \mathbb{Z}$ are accurate and suitable for watermark embedding. Thus, the PCET moments, $S^E = \left\{M^E_{nl}, 0 \leq n \leq N, l \neq 4m\right\}$, the PCT moments, $S^C = \left\{M^C_{nl}, 0 \leq n \leq N, l \geq 0, l \neq 4m\right\}$, and the PST moments, $S^S = \left\{M^S_{nl}, 1 \leq n \leq N, l \geq 0, l \neq 4m\right\}$, are employed for watermarking in this paper, where PCET, PCT, and PST are three implemented versions of PHT moments, as described in Section II-B.

Assume that L watermark bits are embedded in an image, each of which is inserted into P PHT moments. Then, $L_p = L \times P$ PHT moments are selected from S^E, S^C, or S^S via a secret key, KEY, forming $M = \left\{M_{n1,l1}, M_{n2,l2}, \ldots, M_{nL_p,lL_p}\right\}$.

c) Adaptive Moment Normalization: Normalization is a technique that adjusts magnitudes of the to-be-handled geometric moments to achieve scale invariance [49]. As different geometric moments have varying stability under various attacks, they should be normalized adaptively according to their stability characteristics to improve the robustness. To this end, we adaptively determine normalization weights according to the difference stability of PHT moments and then normalize the concerned PHT moment with the determined weight. The adaptive normalization is performed as

$$M^R_{ni,li} = \frac{M_{ni,li}}{M_{00}} \times \frac{10^3}{T_{ni,li}} \tag{10}$$

where $M^R_{ni,li}$ is the normalized PHT moment, M_{00} is the PHT moment with zero-order and zero-repetition, $T_{ni,li}$ is an adaptive normalization weight, and 10^3 is a constant for computational convenience that allows the normalized PHT moments to use integer quantization steps Δ.

To achieve high robustness against attacks, weight $T_{ni,li}$ should be set according to their resistance to various attacks. As PHT moments themselves are resilient to rotation and scaling, the concerned attacks are mainly the common CSP attacks that can be known as prior knowledge to the embedder. As a result, $T_{ni,li}$ can be adaptively determined according to PHT moments' resilience to the common CSP attacks represented by AWGN, JPEG2000, and mean filtering, which in turn achieves robustness to noises, compression, and filtering. That is, by imposing the common CSP attacks on large-scale images, characteristics of PHT moments against the common CSP attacks can be obtained and thus $T_{ni,li}$ can then be adaptively decided accordingly.

As there are a lot of CSP attacks, it is hard to directly set $T_{ni,li}$ adaptively according to the simulated CSP attacks. Alternatively, we turn to characterize variations of PHT moments over the simulated CSP attacks as the momment stability and then represent the stability with a quantitive model. In more detail, motivated by the robustness and flexibility of the bi-square polynomial fitting method [50], we thus fit the average variations of PHT moments under target attack simulation tests on large-scale images a polynomial function. The fitting accuracy depends on the polynomial order, i.e., the large the order, the higher the accuracy is, and the more the polynomial coefficients could be. To balance the accuracy and the polynomial coefficient number, we adopt the following polynomial function, i.e.,

$$\begin{aligned} T_{ni,li} = {} & p_1 + p_2 n_i + p_3 l_i + p_4 {n_i}^2 + p_5 n_i l_i + p_6 {l_i}^2 \\ & + p_7 {n_i}^3 + p_8 {n_i}^2 l_i + p_9 n_i {l_i}^2 + p_{10} {l_i}^3 \end{aligned} \tag{11}$$

where $T_{ni,li}$ is the adaptively fitted weight, and n_i and l_i denote the order and repetition of the PHT moment $M_{ni,li}$ that will be subjected to the adaptive normalization, respectively. p_1 to p_{10} are the polynomial coefficients derived from the fitting operation, whose values will be determined via the numerical simulation in Section IV-B.

d) Adaptive Watermark Embedding: Even with the adaptive normalization, the stability of the normalized PHT moments with the same order and repetition of different images still vary significantly under the same attack. This implies that using a fixed embedding strength for all watermark bits may degrade the robustness. A more effective approach is to determine adaptively the appropriate embedding strength for each watermark bit for different images. Specifically, the robust watermark is embedded using a small embedding strength, yielding a candidate watermarked image. Various CSP attacks represented by AWGN, JPEG2000, and

mean filtering are then performed on the candidate watermarked image. If the robust watermark can be extracted correctly, the used embedding strength can be considered as a practically optimal one for the to-be-embedded watermark bit; otherwise, the embedding strength is increased step by step. Through this adaptive embedding strategy, each robust watermark bit can be inserted with the practically minimum embedding strength on the condition of desirable robustness, which essentially leads to higher robustness at the same distortion and vice versa.

This adaptive embedding may require multiple embedding strength. However, the conventional STDM is a single-level quantizer and thus cannot be directly applied to this embedding scenario. To tackle this problem, we optimize the conventional STDM by modifying the single-level quantizer to become a multiple-level one, where different quantization steps correspond to different embedding strengths. The larger the quantization level, the higher the embedding strength, and vice versa. As a result, the multi-level quantizer can adaptively assign the optimal embedding strength for each watermark bit, enhancing the robustness and invisibility. The specific embedding process is as follows.

Step 1. This step embeds watermarks using a multi-level quantizer and generates a candidate watermarked image. To assign different embedding strengths for different watermark bit, we design a multi-level quantizer with Z embedding strength levels, where a larger levels correspond to higher embedding strength. To determine the optimal embedding strength level for each watermark bit, we use a step-by-step incremental strategy. That is, the robust watermark is first embedded with the smallest embedding strength level, the robust watermark is then extracted from the degraded watermarked image that is imposed with the simulated attacks, the embedding strength level is increased in case of incorrect watermark extraction and remains otherwise. This process is repeated until the robust watermark embedding is completed. To further enhance the robustness and invisibility, a watermark bit is inserted via STDM into a P PHT moment, says $M_i^R = \left\{\left|M^R_{n((i-1)P+1),l((i-1)P+1)}\right|, \left|M^R_{n((i-1)P+2),l((i-1)P+2)}\right|, \ldots, \left|M^R_{n(iP),l(iP)}\right|\right\}$, as illustrated below.

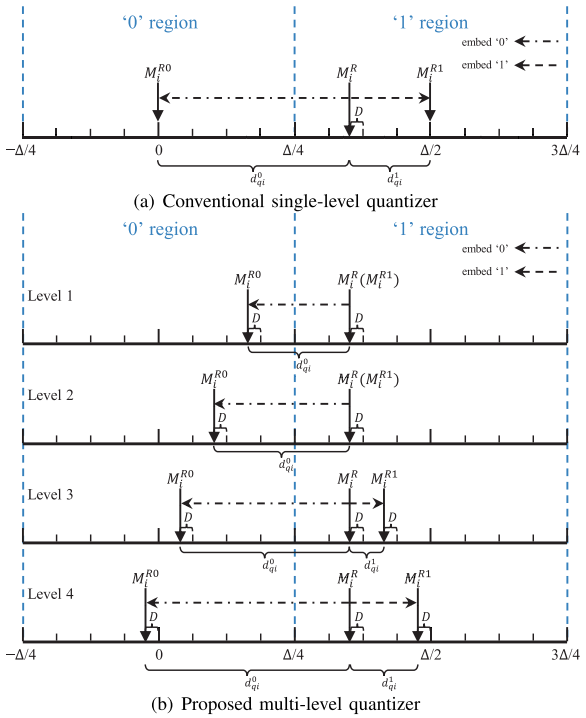
$$M_i^{Rw} = M_i^R + \left(Q_z^w\left(M_i^{RT}u, \Delta\right) - M_i^{RT}u\right)\cdot u, \quad w \in \{0,1\} \tag{12}$$

where $(\cdot)^T$ is a transpose operation, $Q_z^w(\cdot)$ is a multi-level quantizer for the z-th embedding strength level, which is expressed as (13), shown at the bottom of the next page.

After the robust watermark is embedded, distortions caused by the i-th watermark bit is denoted the quantized error d_{qi}, which is used to recover the original image. It is calculated as

$$d_{qi} = Q_z^w\left(M_i^{RT}u, \Delta\right) - M_i^{RT}u \tag{14}$$

Fig. 2 shows that the proposed multi-level STDM only quantizes the integer part, which reduces the amount of quantization errors. In contrast, the conventional single-level STDM quantizes both the integer and decimal parts, which requires more bits to represent quantized errors.



(a) Conventional single-level quantizer

(b) Proposed multi-level quantizer

Fig. 2. Illustration of the optimized STDM-based watermarking method. (a) The conventional single-level quantizer with $\beta_i(0) = 0$ and $\beta_i(1) = \Delta/2$, where $\Delta = 16$; (b) The proposed multi-level quantizer that forces the quantized error d_{qi}^j to be an integer. The D in the figure denotes the decimal part of the normalized PHT moment, which corresponds to the $\left|[x^T u] - x^T u\right|$ terms in Eq. (13).

To reconstruct the robustly watermarked image, an inverse normalization for the normalized PHT moment $M_{ni,li}^{Rw}$ is required. It is computed as

$$M_{ni,li}^{w} = M_{ni,li}^{Rw} \times M_{00} \times \frac{T_{ni,li}}{10^3} \tag{15}$$

where $M_{ni,li}^{w}$ is the PHT moment of a candidate watermarked image.

According to the study by Li et al. [48], the direct image reconstruction using PHT moments causes significant image quality degradation. To tackle this issue, we turn to generate the compensation image I_{com} [48] and then add it to the original image, yielding the corresponding candidate watermarked image I_{cw}. In more detail, it is processed as

$$\begin{aligned} I_{cw} &= I + I_{com} \\ &= I + \left[\sum_{i=1}^{L_p}\left(\left(M_{ni,li}^{w} - M_{ni,li}\right)V_{ni,li} + \left(M_{ni,li}^{w*} - M_{ni,li}^{*}\right)V_{ni,li}^{*}\right)\right] \end{aligned} \tag{16}$$

where $M_{ni,li}^{w*}$, $M_{ni,li}^{*}$, and $V_{ni,li}^{*}$ are the conjugates of $M_{ni,li}^{w}$, $M_{ni,li}$, and $V_{ni,li}$, respectively. By modifying both PHT moments and their conjugate moments, the pixel values of the candidate watermarked image are ensured to be real numbers. According to the PHT definition, the conjugate moment of $M_{ni,li}$ is expressed as

$$\begin{cases} M_{ni,li}^{*} = M_{-ni,-li} & for\ PCET \\ M_{ni,li}^{*} = M_{ni,-li} & for\ PCT\ and\ PST \end{cases} \tag{17}$$

Step 2. This step assigns the optimal embedding strength for each watermark bit by simulating various CSP attacks on the candidate watermarked image I_{cw}. To enhance the robustness of each embedded watermark bit, we mainly perform three attack simulations: AWGN with variance $\sigma^2 = 0.03$, JPEG2000 with compression ratio $C_r = 100$, and mean filtering with window size, W_s, 5×5. These attacks and parameters are chosen based on the attack types and their maximum attack intensities that the robust watermark is expected to withstand in practical applications. For example, we set AWGN variance $\sigma^2 = 0.03$ to ensure high watermark robustness under this attack intensity. Then, we extract the watermark from the attacked image $\hat{I}_{cw}$ and increase its embedding strength if any watermark bit is incorrect, or keep it unchanged otherwise. It is noted that although the proposed method assigns the optimal embedding strength that is able to resist the target simulated attacks, it does not ensure that the watermarked image definitely counteract practical attacks. This is because the simulated attacks may deviate the practical ones.

To extract the watermark from each image in $\hat{I}_{cw}$, its PHT moments $\hat{M}_{nl}^{w}$ are first calculated and selected via Eq. (2) and the secret key KEY, respectively, the normalized PHT moments $\hat{M}_{nl}^{Rw}$ are subsequently computed with Eq. (10), and the watermark is finally detected as

$$\hat{w}_i = arg \min_{\hat{w}_i \in \{0,1\}} \left| \hat{M}_i^{RwT} u - Q^w \left(\hat{M}_i^{RwT} u, \Delta \right) \right| \tag{18}$$

If $\hat{w}_i$ is not equal to the original one and the corresponding embedding strength level z is less than or equal to Z, go to *Step* 3; otherwise, go to *Step* 4.

Step 3. Record the positions from which the extracted watermark is incorrect and increase the embedding strength by one level for these positions. Then, re-embed the watermark and return to *Step* 2.

Step 4. The adaptive watermark embedding is complete, and the final candidate watermarked image I_{cw} is taken as the robustly watermarked image I_w.

2) Reversible Embedding Stage: The reversible embedding stage consists of four parts: watermark-removed image generation, the auxiliary information construction, the auxiliary information embedding, and the integrity authentication sequence insertion. Below are details.

a) Watermark-removed Image Generation: To reduce differences between the robustly watermarked image I_w and the original one I, we generate a watermark-removed image $I_{wremoval}$ that approximates I using I_w and the quantized error d_q. As both the embedder and extractor can obtain $I_{wremoval}$, this enables the recovery of I.

The generation of $I_{wremoval}$ first requires calculating the normalized PHT moments M_{nl}^{Rrw} of I_w using Eqs. (2), and (10). Then, the normalized PHT moments M_{nl}^{Rr} of $I_{wremoval}$ is obtained as

$$M_i^{RrT} u = M_i^{RrwT} u - d_{qi} \tag{19}$$

where M_i^{Rr} and M_i^{Rrw} are the normalized PHT moments of $I_{wremoval}$ and I_w, respectively.

Subsequently, the PHT moments M_{nl}^{r} of $I_{wremoval}$ is computed via Eq. (15). Similar to the generation of the candidate watermarked image, the generation of the watermark-removed image $I_{wremoval}$ also requires obtaining a compensation image I'_{com} first. Therefore, $I_{wremoval}$ is generated as

$$\begin{aligned} I_{wremoval} &= I_w + I'_{com} \\ &= I_w + \left[\sum_{i=1}^{L_p} \left(\left(M_{ni,li}^{r} - M_{ni,li}^{rw} \right) V_{ni,li}^{r} \right. \right. \\ &\quad \left. \left. + \left(M_{ni,li}^{r*} - M_{ni,li}^{rw*} \right) V_{ni,li}^{r*} \right) \right] \end{aligned} \tag{20}$$

b) Auxiliary Information Construction: To achieve the reversibility, it is necessary to construct the auxiliary information for recovering the original image, which consists of quantized errors, rounding errors, and LSBs for the integrity authentication.

The quantized error d_q is the distortion generated by robust watermark embedding, which is calculated by Eq. (14). Using d_q as part of the auxiliary information lets the receiver yield the watermark-removed image $I_{wremoval}$.

The rounding errors d_r are the differences between the original image I and the watermark-removed image $I_{wremoval}$, which is caused by the distortions due to the rounding functions in Eqs. (16), and (20). Using d_r as part of the auxiliary information enables the receiver recover the original image I from $I_{wremoval}$. The d_r is computed as

$$d_r = I - I_{wremoval} \tag{21}$$

To verify the integrity of the received image, a hash operation is applied, yielding a hash sequence H of length L_h. As H needs to be stored in LSBs of the predefined pixels of the image with the embedded watermark and auxiliary information, I_{wa}, the original LSBs b_l need to be preserved. Thus, b_l is taken as part of the auxiliary information.

In summary, the auxiliary information is constructed as $A_{inf} = \{d_q, d_r, b_l\}$.

c) Auxiliary Information Embedding: Considering that the recursive code technique can approach the rate-distortion bound of reversible watermarking and achieve the optimal embedding, we reversibly embed the auxiliary information A_{inf} into the robustly watermarked image I_w using a recursive code-based reversible watermarking [43]. To preserve the integrity verfication and watermark robustness, A_{inf} is embedded outside the inscribed circle of I_w, which excludes

$$Q_z^w \left(x^T u, \Delta \right) = \begin{cases} Q^w \left(x^T u, \Delta \right) - \dfrac{Z - z}{Z} \dfrac{\Delta}{4} - \left| [x^T u] - x^T u \right|, & \left| x^T u - Q^w \left(x^T u, \Delta \right) \right| > \dfrac{\Delta}{4} \text{ and } x^T u \le Q^w \left(x^T u, \Delta \right) \\ Q^w \left(x^T u, \Delta \right) + \dfrac{Z - z}{Z} \dfrac{\Delta}{4} + \left| [x^T u] - x^T u \right|, & \left| x^T u - Q^w \left(x^T u, \Delta \right) \right| > \dfrac{\Delta}{4} \text{ and } x^T u > Q^w \left(x^T u, \Delta \right) \\ x^T u, & otherwise \end{cases} \tag{13}$$

the first L_h pixels. This leads to the image with the embedded robust watermark and the auxiliary information, says I_{wa}.

If the area outside the inscribed circle of I_w cannot accommodate all A_{inf}, some A_{inf} must be embedded inside the circle. As the embedding strength of A_{inf} is much lower than that of the robust watermark w, it can be treated as high-frequency noises that have little influence on the watermark robustness.

d) Integrity Authentication Sequence Insertion: To authenticate the image integrity, we extract LSBs of the first L_h pixels of Image I_{wa} and set them to zeros, resulting in an image for the integrity detection, says I_{id}. Then, a hash operation (e.g., SHA-256) is applied on I_{id} to obtain an L_h-bit hash sequence H. Finally, the LSBs of the first L_h pixels of I_{id} are replaced with H to generate the final watermarked image I_{final}.

B. Integrity Authentication

Integrity authentication can be used to determine whether the received image I_{recv} has been attacked or not. Specifically, the first L_h pixels' LSBs of I_{recv} are extracted to obtain a sequence H_{extr}, and these pixels' LSBs are set as zeros to get an image $\tilde{I}_{id}$. The same hash operation is performed on $\tilde{I}_{id}$ to generate another hash sequence H_{cmpt}. If H_{extr} is equal to H_{cmpt}, it then indicates that I_{recv} has not been attacked; otherwise, I_{recv} has been attacked. Since the image and hash sequence do not have a one-to-one correspondence, two exceptions may exist: 1) both H_{extr} and H_{cmpt} stay the same; or 2) H_{extr} and H_{cmpt} change in the same way. However, these two exceptions are exceedingly rare.

C. Watermark Extraction and Image Reconstruction

After the integrity authentication, the receiver performs watermark extraction and/or image reconstruction. If I_{recv} has been attacked, only the watermark w_{extr} is extracted, whereas the original image I cannot be recovered; otherwise, both w_{extr} and I are extracted and recovered from I_{recv}. Fig. 3 gives the procedure for watermark extraction and image reconstruction.

When I_{recv} has been attacked, the embedded auxiliary information cannot be retrieved, and so thus image I. Thus, only watermark can be extracted, which is performed as follows. First, the PHT moments $\tilde{M}_{nl}^{w}$ of I_{recv} are calculated and selected using Eq. (2) and the secret key KEY, respectively. Then, the normalized PHT moments $\tilde{M}_{nl}^{Rw}$ are obtained with Eq. (10). Finally, the watermark $\tilde{W}$ is extracted as

$$\tilde{w}_i = arg \min_{\tilde{w}_i \in \{0,1\}} \left| \tilde{M}_i^{RwT} u - Q^w \left(\tilde{M}_i^{RwT} u, \Delta \right) \right| \tag{22}$$

where $\tilde{w}_i$ represents the i-th extracted watermark bit.

If I_{recv} has not been attacked, the wateramrk extraction process is the same as that in case of under attacks. When the auxiliary information $\tilde{A}_{inf}$ is extracted from I_{recv}, the extracted $\tilde{A}_{inf}$ can be used to recover the original image I. Specifically, after extracting $\tilde{A}_{inf}$, the LSBs of the first L_h pixels are replaced with b_l to obtain the image $\tilde{I}_w$, which is the same as the robustly watermarked image I_w. Then, I_w and the extracted quantized errors d_q are used to generate the watermark-removed image $I_{wremoval}$ by Eqs. (2), (10), (19), and (20). Finally, I can be recovered perfectly as

$$I = I_{wremoval} + d_r \tag{23}$$

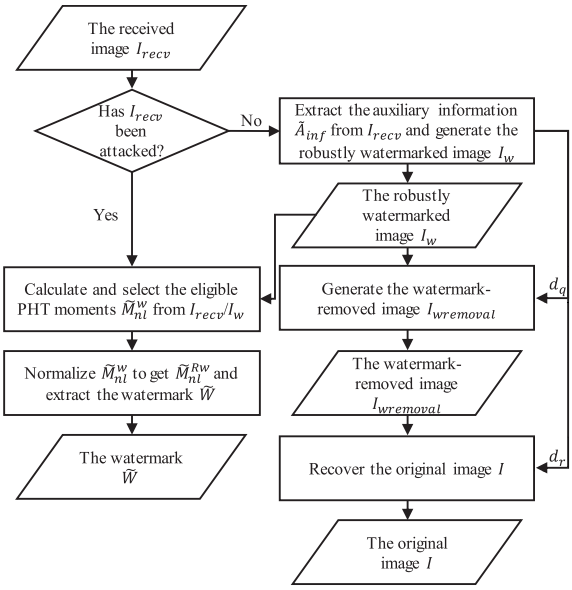


Fig. 3. Flowchart of watermark extraction and image reconstruction of the proposed RRW-CG scheme. $\tilde{M}_{nl}^{w}$ denotes the PHT moment from the received image I_{recv} or the robustly watermarked image I_w.

IV. Experimental Results and Analysis

In this section, we investigate the proposed two-stage RRW-CG scheme using the optimized STDM, the adaptive normalization, and the adaptive watermark embedding. As PHT is the general name of PCET, PCT, and PST, the RRW methods based on these three PHT moments are tested, which are denoted RRW-PCET, RRW-PCT, and RRW-PST for notational convenience, respectively. The preferable parameter settings of the proposed scheme are first determined through experimental simulations, and the effectiveness of each developed technique in the scheme is then evaluated. Subsequently, the visual quality of the watermarked images is illustrated. The robustness of the proposed scheme is finally compared with the state-of-the-art two-stage RRW methods. These are presented in Sections from IV-A to IV-E, respectively.

A. Parameter Settings

In subsequent experiments, we randomly test 1000 images from the BOWS2 database [51] to seek the optimal parameters and verify the effectiveness of each technique in the proposed scheme. We also randomly select another 1000 images from the same database to conduct comparative experiments to evaluate the robustness of the proposed scheme.

The proposed scheme involves five parameters: 1) The maximum order N of PHT moments, which determines the to-be-embedded watermark length; 2) The quantization step Δ, which balances the robustness and invisibility of the watermark; 3) The spreading factor P, which determines the number of carriers for embedding one watermark bit; 4) The maximum level of embedding strength Z, which is used to

TABLE I
PARAMETER SETTINGS FOR 256-BIT WATERMARK

Methods	Parameter settings
WANG [44]	blocks size 32×16; $\zeta = 2.9$
HU-Zernike [32]	$N = 36$; $\Delta = 14$; $T = 1000$
HU-PHT [33]	$N = 19$; $\Delta = 40$; $T = 100$
TANG-PZM [34]	$N = 26$; $\Delta = 32$; $T_{start} = 2400$; $\gamma = 10$
FU-FrZM [35]	$N = 26$; $\Delta = 122500$; $\alpha = 1.14$
RRW-PCET	$N = 27$; $\Delta = 50$; $P = 2$; $Z = 4$
RRW-PCT	$N = 38$; $\Delta = 70$; $P = 2$; $Z = 4$
RRW-PST	$N = 38$; $\Delta = 60$; $P = 2$; $Z = 4$

TABLE II
PARAMETER SETTINGS FOR 512-BIT WATERMARK

Methods	Parameter settings
WANG [44]	blocks size 16×16; $\zeta = 2.6$
HU-PHT [33]	$N = 26$; $\Delta = 28$; $T = 100$
TANG-PZM [34]	$N = 36$; $\Delta = 32$; $T_{start} = 2800$; $\gamma = 10$
FU-FrZM [35]	$N = 36$; $\Delta = 100500$; $\alpha = 1.14$
RRW-PCET	$N = 38$; $\Delta = 34$; $P = 2$; $Z = 4$
RRW-PCT	$N = 53$; $\Delta = 46$; $P = 2$; $Z = 4$
RRW-PST	$N = 53$; $\Delta = 38$; $P = 2$; $Z = 4$

adjust the embedding strength step between adjacent levels; and 5) The adaptive normalization weight $T_{ni,li}$, which is related to the stability of PHT moments with different orders and repetitions. The optimal parameter settings are determined by experimental simulations, aiming to achieve a preferable trade-off between robustness and invisibility. The first two parameters are decided by the to-be-embedded watermark length, while the last three parameters are obtained via the experimental simulations.

To explore the performance under different embedding capacities, we consider the 256- and 512-bit watermark cases. To further evaluate the proposed scheme, it is compared with five state-of-the-art two-stage RRW methods, namely WANG [44], HU-Zernike [32], HU-PHT [33], TANG-PZM [34], and FU-FrZM [35]. The optimal parameter settings for these compared methods are employed from the original paper or derived from the rules given in the corresponding methods, as shown in Tables I and II. For a fair comparison, the parameter settings that make PSNRs of the watermarked images around 39 dB are chosen. It is noted that PSNRs by the proposed scheme are generally larger than those by the compared methods and thus achieving higher robustness than the compared methods would indicate the effectiveness of the proposed scheme. By the way, as HU-Zernike employs Zernike moments with weak numerical stability, it is unable to embed a 512-bit watermark and thus omitted from Table II.

B. Calculation of the Adaptive Normalization Weight

As described in Section II-A, the adaptive normalization weight $T_{ni,li}$ of the proposed scheme is determined via Eq. (11). To determine desirable polynomial coefficients in Eq. (11), we perform attack simulation tests to fit these coefficients.

In the simulation, we randomly select 1000 images from the BOWS2 database [51] and impose the same attack types and intensities as those used for the adaptive watermark embedding, i.e., AWGN with variance $\sigma^2 = 0.03$, JPEG2000 with compression ratio $C_r = 100$, and mean filtering with window size, W_s, 5×5. It is noted that these attacks are used to determine the polynomial coefficients in Eq. (11), whereas the attack simulation tests in the adaptive watermark embedding process to decide the optimal embedding strength for each watermark bit, which are performed individually.

By using the bi-square polynomial fitting method [50], the polynomial coefficients fitted from the simulated attacks are calculated as

PCET moments: $p_1 = 0.04$, $p_2 = -4.27e-04$, $p_3 = 5.95e-06$, $p_4 = 2.65e-05$, $p_5 = -2.21e-07$, $p_6 = -1.97e-06$, $p_7 = -5.09e-07$, $p_8 = 4.83e-09$, $p_9 = 2.96e-07$, and $p_{10} = -3.48e-09$.

PCT moments: $p_1 = 0.06$, $p_2 = -1.71e-03$, $p_3 = -9.38e-04$, $p_4 = 5.60e-05$, $p_5 = 5.71e-05$, $p_6 = 1.88e-05$, $p_7 = -5.70e-07$, $p_8 = -7.66e-07$, $p_9 = -4.98e-07$, and $p_{10} = -1.30e-07$.

PST moments: $p_1 = 0.06$, $p_2 = -1.65e-03$, $p_3 = -1.04e-03$, $p_4 = 5.11e-05$, $p_5 = 5.53e-05$, $p_6 = 2.16e-05$, $p_7 = -5.07e-07$, $p_8 = -7.01e-07$, $p_9 = -4.55e-07$, and $p_{10} = -1.56e-07$.

To evaluate the fitting accuracy, we calculate the R-squared values between T_{nl} and each variation of PCET, PCT and PST. The values are 0.955, 0.961, and 0.970, respectively, suggesting that T_{nl} from the fitting operation accurately characterizes variations of the three types of PHT moments over the simulated attacks and thus the stability of the three types of PHT moments.

To sum up, the adaptive normalization weight $T_{ni,li}$ of each PHT moment can be calculated based on its order and repetition. The larger the $T_{ni,li}$ is, the weaker the stability of the PHT moment is, and the larger the watermark embedding strength should be. According to Eqs. (10), (12) and (15), the watermark embedding strength measured by the embedding distortion in case of the fixed quantization step Δ is inversely proportional to $T_{ni,li}$. Thus, the normalization equation in Eq. (10) is constructed to implement the adaptive normalization in our work.

C. Effectiveness of the Developed Techniques

The proposed scheme applies STDM and develops the adaptive normalization and embedding. To assess their effectiveness, we evaluate the robustness of the watermarked images with a 256-bit watermark against various attacks in this section.

1) Effectiveness of STDM: The proposed scheme employs the optimized STDM for watermark embedding, which differs from QIM in the spreading factor P and distinguishes from the conventional STDM in the number of quantization levels. That is, QIM can be regarded as a special case of STDM with $P = 1$.

To assess the robustness of the proposed scheme under different spreading factors, we conduct the experimental simulations by taking RRW-PCET as an example, which is similar for RRW-PST and RRW-PCT. To represent different spreading degrees, three spreading factors, $P = 1$, 2, and 4, are used. In the simulation, five types of attacks including AWGN, JPEG, JPEG2000, mean filtering, and median filtering

TABLE III

ROBUSTNESS PERFORMANCE OF THE PROPOSED RRW-PCET METHOD IN TERMS OF BER AGAINST AWGN, JPEG, JPEG2000, MEAN FILTERING, AND MEDIAN FILTERING WITH DIFFERENT SPREADING FACTOR P. THE σ^2, Q_f, C_r, AND W_s DENOTE AWGN VARIANCE, JPEG QUALITY FACTOR, JPEG2000 COMPRESSION RATIO, AND FILTERING WINDOW SIZE, RESPECTIVELY

Attack		AWGN (σ^2)			JPEG (Q_f)			JPEG2000 (C_r)			mean filtering (W_s)		median filtering (W_s)		PSNR
Parameter		0.013	0.021	0.029	10	30	50	60	80	100	3×3	5×5	3×3	5×5	
$P=1$	Lena	0.39	1.56	5.47	3.91	0.00	0.00	2.34	3.91	7.81	0.00	0.00	0.00	1.17	39.64
	Average	1.08	4.15	9.42	3.97	0.13	0.03	3.75	6.17	8.48	0.00	1.25	0.21	2.17	39.23
$P=2$	Lena	0.39	1.95	4.69	2.73	0.00	0.00	1.17	3.13	6.25	0.00	0.00	0.00	0.00	39.43
	Average	**0.95**	**3.94**	**8.98**	**3.82**	**0.10**	**0.02**	**3.68**	**5.81**	**8.51**	**0.00**	**0.81**	**0.15**	**1.92**	**39.09**
$P=4$	Lena	0.78	2.73	6.25	4.08	0.08	0.00	4.32	6.59	9.31	0.00	0.00	0.00	0.39	39.18
	Average	1.02	3.98	9.15	4.02	0.06	0.01	4.53	6.79	10.80	0.00	0.92	0.18	2.02	38.96

TABLE IV

ROBUSTNESS PERFORMANCE OF THE PROPOSED RRW-PCET METHOD IN TERMS OF BER AGAINST AWGN, JPEG, JPEG2000, MEAN FILTERING, AND MEDIAN FILTERING WITH DIFFERENT NORMALIZATION STRATEGIES. THE σ^2, Q_f, C_r, AND W_s DENOTE AWGN VARIANCE, JPEG QUALITY FACTOR, JPEG2000 COMPRESSION RATIO, AND FILTERING WINDOW SIZE, RESPECTIVELY

Attack		AWGN (σ^2)			JPEG (Q_f)			JPEG2000 (C_r)			mean filtering (W_s)		median filtering (W_s)		PSNR
Parameter		0.013	0.021	0.029	10	30	50	60	80	100	3×3	5×5	3×3	5×5	
T^f	Lena	0.39	2.34	5.08	4.30	0.00	0.00	3.13	7.42	12.50	0.00	0.39	0.39	3.13	39.18
	Average	1.44	4.58	10.04	4.64	0.15	0.04	4.53	7.13	10.48	0.01	0.83	0.40	2.25	38.82
T^a_{nl}	Lena	0.39	1.95	4.69	2.73	0.00	0.00	1.17	3.13	6.25	0.00	0.00	0.00	0.00	39.43
	Average	**0.95**	**3.94**	**8.98**	**3.82**	**0.10**	**0.02**	**3.68**	**5.81**	**8.51**	**0.00**	**0.81**	**0.15**	**1.92**	**39.09**

are applied to the watermarked images. Parameters of these attacks are: the AWGN variance σ^2 from 0.013 to 0.029 with step 0.008; the JPEG quality factor Q_f from 10 to 50 with step 20; the JPEG2000 compression ratio C_r from 60 to 100 with step 20; and the mean filtering and median filtering window sizes W_s are 3×3 and 5×5, respectively.

Table III shows bit error rates (BERs) for Image Lena and the average BER of 1000 images under these attacks. The results indicate that STDM with $P = 2$ achieves the optimal trade-off between robustness and invisibility, demonstrating the effectiveness of the spreading strategy for improving the watermark robustness. By spreading one watermark bit into two PHT moments, the invisibility is slightly reduced, while the robustness is significantly improved. However, $P = 4$ performs weakly because the high order PCET moments capture the high-frequency information of the image, which are more vulnerable to compression attacks.

2) Effectiveness of Adaptive Normalization: To demonstrate the effectiveness of the adaptive normalization, we adopt two normalization strategies: using the fixed normalization weights T^f, and taking the adaptive normalization weights T^a_{nl}, and then compare their robustness against AWGN, JPEG, JPEG2000, mean filtering, and median filtering on the condition of similar PSNRs. In the simulation, T^a_{nl} are determined by Eq. (11), while T^f are adjusted image-by-image to result in the close PSNRs. The AWGN variance σ^2 ranges from 0.013 to 0.029 with an interval of 0.008; the JPEG quality factor Q_f ranges from 10 to 50 with step 20; the JPEG2000 compression ratio C_r ranges from 60 to 100 with step 20; and window sizes W_s of the mean filtering and median filtering are 3×3 and 5×5, respectively.

Table IV shows BERs for Image Lena and the average BER of 1000 test images. It is found that the adaptive normalization using T^a_{nl} leads to lower BERs than the fixed normalization with T^f under the same PSNR, demonstrating the effectiveness of the adaptive normalization.

3) Effectiveness of Adaptive Watermark Embedding: To evaluate the adaptive watermark embedding, we examine the robustness under two scenarios: with or without the adaptive watermark embedding, which are denoted S_1 and S_2 for notational convenience. For each test image, a 256-bit watermark is embedded and the quantization step Δ is adjusted to make PSNRs of the watermarked images in both scenarios similar. In addition, we also investigate how the maximum level of embedding strength Z affects the robustness.

As shown in Table V, scenario S_1 outperforms S_2 in terms of robustness under various attacks. This is because S_1 adapts the embedding strength of different watermark bits to the stability of different PHT moments over various attacks and thus achieves higher robustness than S_2 at the same distortion.

Table V also indicates that increasing the quantization level Z enhances robustness of the watermarked image. Nevertheless, the robustness improvement is negligible when Z is greater than 4. Thus, we choose $Z = 4$ as the practically optimal setting in our experimental simulations to balance robustness and computation.

In addition, the optimized STDM with $Z = 4$ quantization levels can not only achieve the adaptive watermark embedding but also reduce the amount of quantized errors due to robust watermark embedding. By comparing the conventional single-level STDM in Eq. (8) with the proposed multi-level one in Eq. (13), it is found that the multi-level quantizer only quantizes the integer part of magnitude of the normalized PHT moments while preserving the decimal part, which significantly reduces the amount of quantized errors. Table VI shows the bit number representing the quantized error under the aforementioned two scenarios, in which each bit number is averaged over 1000 test images.

It is shown from Table VI that compared with the single-level quantizer, the multi-level quantizer reduces the average bit number for the quantized error d_q by 89%, from 23027.3 to 2531.2 bits. Therefore, the bit number representing the quantized errors due to the multi-level quantizer is significantly

TABLE V

ROBUSTNESS PERFORMANCE OF THE PROPOSED RRW-PCET METHOD IN TERMS OF BER AGAINST AWGN, JPEG, JPEG2000, MEAN FILTERING, AND MEDIAN FILTERING WITH SCENARIOS S_1 AND S_2. THE σ^2, Q_f, C_r, AND W_s DENOTE AWGN VARIANCE, JPEG QUALITY FACTOR, JPEG2000 COMPRESSION RATIO, AND FILTERING WINDOW SIZE, RESPECTIVELY

Attack			AWGN (σ^2)			JPEG (Q_f)			JPEG2000 (C_r)			mean filtering (W_s)		median filtering (W_s)		PSNR
Parameter			0.013	0.021	0.029	10	30	50	60	80	100	3×3	5×5	3×3	5×5	
S_1	$Z=2$	Lena	2.34	2.73	6.25	1.95	0.00	0.00	2.73	3.91	7.42	0.00	0.00	0.00	0.39	38.94
		Average	1.93	5.03	11.87	4.05	0.18	0.04	5.09	7.61	11.23	0.06	1.80	0.27	2.13	38.54
	$Z=4$	Lena	0.39	1.95	4.69	2.73	0.00	0.00	1.17	3.13	6.25	0.00	0.00	0.00	0.00	39.43
		Average	0.95	3.94	8.98	3.82	0.10	0.02	3.68	5.81	8.51	0.00	0.81	0.15	1.92	39.09
	$Z=6$	Lena	0.78	1.95	3.52	2.73	0.00	0.00	1.56	2.73	5.47	0.00	0.39	0.00	0.78	39.64
		Average	**1.07**	**3.80**	**7.69**	**3.81**	**0.13**	**0.01**	**3.60**	**5.90**	**8.42**	**0.00**	**0.49**	**0.10**	**1.85**	**39.20**
S_2		Lena	1.17	4.69	8.98	2.34	0.00	0.00	2.73	5.47	7.42	0.00	1.56	0.00	2.73	39.41
		Average	3.02	6.68	12.53	3.68	0.10	0.00	3.73	5.85	8.79	0.05	1.13	0.22	2.98	38.99

TABLE VI

THE BIT NUMBER REPRESENTING THE QUANTIZED ERROR d_q FOR DIFFERENT QUANTIZERS

Quantizer	Conventional single-level quantizer	**Proposed multi-level quantizer**
Bit number for Lena	22013	**2558**
Average bit number	23027.3	**2531.2**

TABLE VII

TARGET ATTACK TYPES FOR DIFFERENT SCENARIOS, WHERE NONE IMPLIES THAT THE ADAPTIVE NORMALIZATION AND WATERMARK EMBEDDING STRATEGIES ARE NOT USED

Scenarios	Target attack types
S_3	None
S_4	AWGN
S_5	AWGN, JPEG2000
S_6	AWGN, JPEG2000, mean filtering
S_7	AWGN, JPEG2000, mean filtering, JPEG
S_8	AWGN, JPEG2000, mean filtering, JPEG, median filtering

lower than that of the single-level quantizer. This demonstrates that the optimized STDM with a multi-level quantizer effectively reduces quantized errors compared with the conventional single-level quantizer.

4) Effectiveness of Attack Simulation on the Proposed Scheme: Both the adaptive normalization and watermark embedding strategies of the proposed scheme employ the attack simulation test. To evaluate the effectiveness of the attack simulation, we consider six scenarios S_3, S_4, ..., S_8, as shown in Table VII. In the simulation, the quantization step Δ is adjusted to make PSNRs of the watermarked images for all scenarios similar and the other parameters except the target attacks are set the same. The robustness performance of different scenarios against AWGN, JPEG, JPEG2000, mean filtering, and median filtering is given in Table VIII.

It is observed from Table VIII that compared with scenarios S_4 to S_8 that use the attack simulation, scenario S_3 without leveraging the attack simulation exhibits worse performance in terms of BER against AWGN. This is because S_3 neither employs the adaptive normalization weights obtained from the stability of PHT moments nor adaptively determines the embedding strength according to the watermark detection performance of the attacked image.

By comparing scenarios S_4, S_5, and S_6, it can be found that the overall robustness of the proposed scheme is improved by increasing the target attack types, as this enables the scheme to adapt to more attack situations. However, increasing the target attack types also degrades the scheme's resistance to AWGN, which is because the embedding strength in scenarios S_5 and S_6 is optimized for all types of target attacks rather than AWGN and thus the optimal embedding strength is not necessarily the best for AWGN. In addition, by comparing scenarios S_6, S_7, and S_8, it is observed that increasing target attack types can further enhance the robustness, but it also increases the computational complexity.

The comparison between scenarios S_4 and S_5 against JPEG reveals that even though S_5 does not take JPEG as the target attack, it really enhances the robustness against JPEG by adopting JPEG2000 as the target attack. Similarly, S_6 improves the robustness to median filtering by using mean filtering as the target attack. These results suggest that JPEG2000 and mean filtering can represent JPEG and median filtering to some extent, respectively, when selecting the target attack types. Therefore, we choose S_6 containing AWGN, JPEG2000, and mean filtering as the practically optimal target attack types to improve the robustness to noises, compression, and filtering.

In summary, the proposed scheme significantly improves the robustness using target attacks for implementing the attack simulation and trades off the computational and robustness by adopting the representative attacks.

D. Visual Quality of Watermarked Images

This section evaluates the visual quality and PSNRs of the watermarked images. Parameters in Section IV-A are used to embed a 256-bit watermark into each test image and PSNRs of both the robustly watermarked image I_w and the final watermarked image I_{final} are then calculated accordingly.

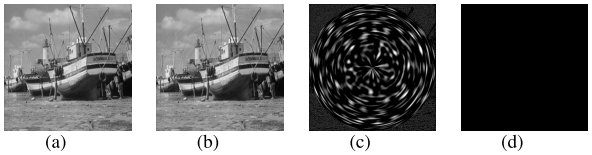
For illustration, Image Boat is taken as an example. Fig. 4 displays the original image I, the final watermarked image I_{final}, the difference between I and I_{final}, and the difference between I and the image recovered from I_{final}. The results show that I_{final} has a good visual quality and is hard to distinguish from I by the naked eyes. Similar results are also obtained for other images. Moreover, Fig. 4 additionally demonstrates that, the proposed scheme is able to recover the original image perfectly in case of no attacks.

Table IX summarizes PSNRs of I_w and I_{final} for Image Lena as well as the average PSNRs of them for 1000 test images. It is observed that the proposed RRW-PCET method achieves average PSNRs of 40.25 and 39.09 dB for I_w

TABLE VIII

ROBUSTNESS PERFORMANCE OF THE PROPOSED RRW-PCET METHOD WITH SCENARIOS S_3 TO S_8. THE σ^2, Q_f, C_r, AND W_s DENOTE AWGN VARIANCE, JPEG QUALITY FACTOR, JPEG2000 COMPRESSION RATIO, AND FILTERING WINDOW SIZE, RESPECTIVELY

Attack		AWGN (σ^2)			JPEG (Q_f)			JPEG2000 (C_r)			mean filtering (W_s)		median filtering (W_s)		PSNR
Parameter		0.013	0.021	0.029	10	30	50	60	80	100	3×3	5×5	3×3	5×5	
S_3	Lena	1.17	8.59	17.19	0.78	0.00	0.00	3.91	4.69	10.94	0.00	0.39	0.00	0.00	39.36
	Average	2.11	7.68	15.38	0.86	0.00	0.00	4.91	8.30	12.15	0.03	1.87	0.05	1.38	38.82
S_4	Lena	0.39	0.78	3.52	10.16	0.39	0.00	8.59	15.23	21.48	0.39	10.94	0.78	6.64	39.36
	Average	0.53	2.16	4.74	6.62	0.67	0.34	11.83	15.00	17.96	1.11	9.02	1.31	7.89	38.93
S_5	Lena	0.78	2.34	5.08	3.91	0.39	0.00	1.17	3.91	7.03	0.39	3.91	0.78	2.73	39.27
	Average	0.91	3.72	8.53	4.08	0.31	0.20	3.88	5.95	8.21	0.34	4.03	0.75	3.89	38.88
S_6	Lena	0.39	1.95	4.69	2.73	0.00	0.00	1.17	3.13	6.25	0.00	0.00	0.00	0.00	39.43
	Average	**0.95**	**3.94**	**8.98**	**3.82**	**0.10**	**0.02**	**3.68**	**5.81**	**8.51**	**0.00**	**0.81**	**0.15**	**1.92**	**39.09**
S_7	Lena	0.78	1.56	3.13	1.95	0.00	0.00	2.73	3.52	7.42	0.00	0.00	0.39	1.17	39.30
	Average	0.93	3.97	9.09	1.20	0.08	0.05	3.45	5.67	7.73	0.00	0.90	0.20	2.23	39.05
S_8	Lena	0.78	1.56	3.52	0.78	0.00	0.00	3.13	3.91	7.81	0.00	0.00	0.00	0.00	39.38
	Average	0.90	3.98	9.25	1.51	0.05	0.02	2.92	5.13	7.52	0.00	0.90	0.10	0.88	38.96



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

Fig. 4. Illustration of the visual quality of the watermarked image using the proposed RRW-PCET method. (a) the original image I; (b) the final watermarked image I_{final} (PSNR = 39.06 dB); (c) the difference between (a) and (b), in which the magnitude is scaled 30 times for visual convenience; and (d) the difference between (a) and the image recovered from (b).

TABLE IX

PSNRS IN TERMS OF DB OF THE ROBUSTLY WATERMARKED IMAGE I_w AND THE FINAL WATERMARKED IMAGE I_{final} BY THE PROPOSED RRW-PCET METHOD

Watermarked Image	I_w	I_{final}
PSNR for Lena	40.07	39.43
Average PSNR	40.25	39.09

and I_{final}, respectively. Embedding the auxiliary information decreases PSNRs of I_w, which reduces by only about 1.2 dB for 1000 test images and thus acceptable in practice.

E. Robustness Against CSP and GD Attacks

To evaluate the effectiveness of the three proposed methods i.e., RRW-PCET, RRW-PCT, and RRW-PST, we compare them with five state-of-the-art two-stage RRW methods including WANG [44], HU-Zernike [32], HU-PHT [33], TANG-PZM [34], and FU-FrZM [35] in this section.

In the comparative experiments, we select 1000 images from the BOWS2 database [51] that are different from those used in the above experiments, insert the same 256- and 512-bit watermark using the proposed two-stage RRW methods, and adjust the embedding strength of each RRW method according to the parameter settings in Tables I and II. Tables X and XI show the PSNRs of the watermarked images. The three proposed methods, i.e., RRW-PCET, RRW-PCT, and RRW-PST, produce higher PSNRs than the other compared methods. Therefore, the proposed scheme can demonstrate its superiority if it obtains a lower BER. However, HU-Zernike cannot embed a 512-bit watermark because of the weak numerical stability of higher order Zernike moments.

The robustness of the compared seven methods is evaluated by testing their resistance to both CSP and GD. WANG uses

TABLE X

PSNRS IN TERMS OF DB OF THE WATERMARKED IMAGES AFTER EMBEDDING A 256-BIT WATERMARK BY ALL COMPARED METHODS

Methods	WANG [44]	HU-Zernike [32]	HU-PHT [33]
PSNR for Lena	38.30	38.45	38.43
Average PSNR	38.58	37.87	38.66
Methods	TANG-PZM [34]	FU-FrZM [35]	RRW-PCET
PSNR for Lena	38.84	38.69	39.43
Average PSNR	38.59	38.88	39.09
Methods	RRW-PCT	RRW-PST	
PSNR for Lena	40.01	39.04	
Average PSNR	39.26	39.05	

TABLE XI

PSNRS IN TERMS OF DB OF THE WATERMARKED IMAGES AFTER EMBEDDING A 512-BIT WATERMARK BY ALL COMPARED METHODS

Methods	WANG [44]	HU-PHT [33]	TANG-PZM [34]
PSNR for Lena	38.47	38.76	38.91
Average PSNR	38.76	38.68	38.86
Methods	FU-FrZM [35]	RRW-PCET	RRW-PCT
PSNR for Lena	38.60	39.10	39.31
Average PSNR	38.77	39.22	39.08
Methods	RRW-PST		
PSNR for Lena	38.95		
Average PSNR	39.16		

the statistical histogram of image blocks as the carrier for watermark embedding, which makes it unable to withstand GD. Therefore, we only test its robustness against CSP. The experimental results of each method against CSP and GD are presented below.

1) Robustness Against CSP: We first investigate the performance of the seven two-stage RRW methods against three CSPs. In the simulation, AWGN, JPEG, JPEG2000, mean filtering, and median filtering attacks are imposed on each watermarked image. The AWGN variance σ^2 ranges from 0.013 to 0.029 with step 0.008; the JPEG quality factor Q_f varies from 10 to 50 with step 20; the JPEG2000 compression ratio C_r is from 60 to 100 with step 20; and window sizes W_s of mean filtering and median filtering are 3×3 and 5×5, respectively. Tables XII and XIII summarize the robustness of the eight two-stage RRW methods against AWGN, JPEG,

TABLE XII

ROBUSTNESS OF ALL COMPARED METHODS IN TERMS OF BER AGAINST AWGN, JPEG, JPEG2000, MEAN FILTERING, AND MEDIAN FILTERING FOR 256-BIT CASE. THE σ^2, Q_f, C_r, AND W_s DENOTE AWGN VARIANCE, JPEG QUALITY FACTOR, JPEG2000 COMPRESSION RATIO, AND FILTERING WINDOW SIZE, RESPECTIVELY

Attack		AWGN (σ^2)			JPEG (Q_f)			JPEG2000 (C_r)			mean filtering (W_s)		median filtering (W_s)		PSNR
Parameter		0.013	0.021	0.029	10	30	50	60	80	100	3×3	5×5	3×3	5×5	
WANG [44]	Lena	0.39	2.73	6.64	45.31	27.34	9.77	46.48	47.66	48.44	33.59	45.70	26.17	41.80	38.30
	Average	0.94	2.36	5.40	31.52	14.17	4.31	41.58	42.95	44.17	25.53	42.95	16.92	39.64	38.58
HU-Zernike [32]	Lena	8.20	14.45	23.83	3.52	0.00	0.00	7.03	8.98	13.28	0.00	0.39	0.00	0.78	38.45
	Average	5.88	12.75	17.93	3.61	0.02	0.00	6.41	9.13	12.72	0.04	1.68	0.18	3.04	37.87
HU-PHT [33]	Lena	4.30	10.55	19.14	1.56	0.00	0.00	3.13	3.52	7.42	0.00	0.00	0.00	0.78	38.43
	Average	4.12	11.07	18.48	1.31	0.00	0.00	3.99	6.75	10.24	0.00	0.61	0.30	2.05	38.66
TANG-PZM [34]	Lena	2.73	10.61	17.58	3.91	0.00	0.00	3.52	5.08	7.42	0.00	0.39	0.00	0.00	38.84
	Average	5.33	12.75	18.05	3.25	0.03	0.00	4.72	6.97	9.71	0.01	0.45	0.16	2.16	38.59
FU-FrZM [35]	Lena	2.34	6.64	12.89	0.39	0.00	0.00	0.78	3.13	6.64	0.00	0.00	0.00	0.00	38.69
	Average	4.30	10.71	17.18	1.87	0.01	0.01	3.84	6.13	8.68	0.01	0.57	0.12	1.66	38.88
RRW-PCET	Lena	0.39	1.95	4.69	2.73	0.00	0.00	1.17	3.13	6.25	0.00	0.00	0.00	0.00	39.43
	Average	**0.95**	**3.94**	**8.98**	**3.82**	**0.10**	**0.02**	**3.68**	**5.81**	**8.51**	**0.00**	**0.81**	**0.15**	**1.92**	**39.09**
RRW-PCT	Lena	0.39	1.56	4.69	1.56	0.00	0.00	3.13	3.52	6.25	0.00	0.00	0.00	0.00	40.01
	Average	**0.76**	**3.50**	**8.40**	**3.37**	**0.08**	**0.01**	**3.42**	**5.04**	**7.70**	**0.00**	**0.25**	**0.17**	**1.96**	**39.26**
RRW-PST	Lena	0.39	2.34	5.86	0.39	0.00	0.00	0.78	1.17	5.47	0.00	0.00	0.00	0.00	39.04
	Average	**0.63**	**3.08**	**7.76**	**3.01**	**0.10**	**0.02**	**2.92**	**4.53**	**6.68**	**0.00**	**0.19**	**0.15**	**1.35**	**39.05**

TABLE XIII

ROBUSTNESS OF ALL COMPARED METHODS IN TERMS OF BER AGAINST AWGN, JPEG, JPEG2000, MEAN FILTERING, AND MEDIAN FILTERING FOR 512-BIT CASE. THE σ^2, Q_f, C_r, AND W_s DENOTE AWGN VARIANCE, JPEG QUALITY FACTOR, JPEG2000 COMPRESSION RATIO, AND FILTERING WINDOW SIZE, RESPECTIVELY

Attack		AWGN (σ^2)			JPEG (Q_f)			JPEG2000 (C_r)			mean filtering (W_s)		median filtering (W_s)		PSNR
Parameter		0.013	0.021	0.029	10	30	50	60	80	100	3×3	5×5	3×3	5×5	
WANG [44]	Lena	4.69	10.35	10.35	40.63	21.29	5.27	44.14	46.09	45.90	29.10	45.31	21.09	42.97	38.47
	Average	4.53	11.50	10.41	32.02	11.91	3.51	42.67	44.21	45.39	25.37	44.96	17.31	42.31	38.76
HU-PHT [33]	Lena	11.72	18.55	32.81	10.35	0.00	0.00	9.96	10.94	22.46	0.00	3.52	0.20	2.73	38.76
	Average	12.81	20.57	30.65	6.26	0.01	0.00	12.62	17.36	22.59	0.11	5.29	0.44	5.57	38.68
TANG-PZM [34]	Lena	10.35	19.73	26.37	5.27	0.00	0.00	7.03	8.01	17.77	0.00	2.15	0.00	1.37	38.91
	Average	15.17	24.34	30.99	8.44	0.23	0.03	11.32	15.48	20.42	0.11	3.67	0.45	4.95	38.86
FU-FrZM [35]	Lena	12.50	18.95	25.39	6.84	0.00	0.00	6.45	9.18	17.19	0.00	1.95	0.00	0.39	38.60
	Average	13.65	23.34	30.09	6.89	0.04	0.00	10.52	14.54	18.91	0.17	3.63	0.38	4.22	38.77
RRW-PCET	Lena	4.49	10.74	16.41	4.30	0.00	0.00	10.94	12.50	18.16	0.00	2.54	0.00	0.78	39.10
	Average	**5.08**	**10.46**	**15.37**	**6.25**	**0.21**	**0.04**	**10.28**	**14.45**	**18.88**	**0.03**	**6.10**	**0.37**	**4.81**	**39.22**
RRW-PCT	Lena	5.27	11.33	16.02	6.05	0.00	0.00	7.23	11.52	18.16	0.00	1.76	0.00	0.20	39.31
	Average	**5.09**	**10.54**	**15.38**	**5.25**	**0.20**	**0.04**	**10.06**	**14.26**	**18.22**	**0.05**	**4.28**	**0.35**	**4.21**	**39.08**
RRW-PST	Lena	4.49	11.33	16.80	3.52	0.20	0.00	9.77	14.06	19.53	0.00	1.17	0.00	0.00	38.95
	Average	**4.23**	**9.50**	**14.01**	**5.34**	**0.18**	**0.03**	**10.15**	**14.10**	**18.43**	**0.04**	**3.16**	**0.39**	**3.66**	**39.16**

JPEG2000, mean filtering, and median filtering attacks in case of 256- and 512-bit watermarks.

Table XII shows BERs of the compared methods under various AWGN variances, JPEG quality factors, JPEG2000 compression ratios, and filtering window sizes for 256-bit watermarks. For AWGN, the three proposed methods achieve lower BERs than the other four geometric-moment-based methods including HU-Zernike, HU-PHT, TANG-PZM, and FU-FrZM. For instance, when the AWGN variance σ^2 is 0.029, the average BERs of RRW-PCET, RRW-PCT, and RRW-PST are 8.98%, 8.40%, and 7.76%, respectively, while those of HU-Zernike, HU-PHT, TANG-PZM, and FU-FrZM are 17.93%, 18.48%, 18.05%, and 17.18%, respectively. This demonstrates that our methods are more robust to AWGN than the other geometric-moment-based methods. However, WANG outperforms our methods because it uses the low-frequency components of Haar wavelet and thus is more resistant to AWGN. For JPEG, the average BERs of all methods except WANG are below 10% even when the JPEG quality factor Q_f is 10. Among them, the average BERs of the three proposed methods are slightly larger than those of HU-PHT and FU-FrZM, which is because these three methods do not include JPEG attack in the attack simulation tests and thus cannot assign the optimal embedding strength for each watermark bit under JPEG attack. However, the use of JPEG2000 in the attack simulation tests improves the compression resistance, which makes the three proposed methods exhibit higher robustness than HU-Zernike and TANG-PZM. For JPEG2000, the three proposed methods perform better than the other compared methods. Even at the severest JPEG2000 compression ratio C_r of 100, the average BERs of these three proposed methods are lower than 10%, which implies the effectiveness of the adaptive normalization and embedding. For mean filtering, all geometric-moment-based methods achieve excellent robustness. Among our methods, RRW-PCT and RRW-PST outperform RRW-PCET, which is because PCT and PST moments are more stable than PCET moments in resisting filtering attacks [36]. For median filtering, the average BERs of our methods are similar to those of other geometric-moment-based methods, in which the performance of RRW-PST is better. This demonstrates that the mean filtering used in the attack simulation tests also improves the robustness of our methods against other filtering attacks.

Table XIII presents BERs of the compared methods under various AWGN variances, JPEG quality factors, JPEG2000 compression ratios, and filtering window sizes for 512-bit watermarks. It is noted that BERs of all methods increase for the 512-bit watermark case compared with the 256-bit

TABLE XIV

ROBUSTNESS OF ALL COMPARED METHODS IN TERMS OF BER AGAINST ROTATION AND SCALING FOR 256-BIT CASE. THE R_a AND S_f DENOTE ROTATION ANGLE AND SCALING FACTOR, RESPECTIVELY

Attack		Rotation (R_a)					Scaling (S_f)						PSNR
Parameter		10	30	50	70	90	0.5	0.8	1.1	1.4	1.7	2.0	
HU-Zernike [32]	Lena	0.00	0.00	0.00	0.00	0.00	0.39	0.00	0.00	0.00	0.00	0.00	38.45
	Average	0.45	0.26	0.15	0.29	0.00	2.30	0.00	0.00	0.02	0.05	0.12	37.87
HU-PHT [33]	Lena	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	38.43
	Average	0.00	0.00	0.00	0.00	0.00	0.08	0.00	0.00	0.00	0.00	0.00	38.66
TANG-PZM [34]	Lena	0.00	0.00	0.00	0.00	0.00	0.39	0.00	0.00	0.00	0.00	0.00	38.84
	Average	0.17	0.11	0.07	0.08	0.00	1.10	0.00	0.00	0.00	0.01	0.03	38.59
FU-FrZM [35]	Lena	0.00	0.00	0.00	0.00	0.00	1.17	0.00	0.00	0.00	0.00	0.00	38.69
	Average	0.11	0.02	0.01	0.01	0.00	0.74	0.13	0.01	0.01	0.02	0.01	38.88
RRW-PCET	Lena	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	39.43
	Average	**0.02**	**0.02**	**0.00**	**0.02**	**0.00**	**0.10**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**39.09**
RRW-PCT	Lena	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	40.01
	Average	**0.04**	**0.03**	**0.00**	**0.03**	**0.00**	**0.24**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**39.26**
RRW-PST	Lena	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	39.04
	Average	**0.04**	**0.01**	**0.00**	**0.01**	**0.00**	**0.16**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**39.05**

TABLE XV

ROBUSTNESS OF ALL COMPARED METHODS IN TERMS OF BER AGAINST ROTATION AND SCALING FOR 512-BIT CASE. THE R_a AND S_f DENOTE ROTATION ANGLE AND SCALING FACTOR, RESPECTIVELY

Attack		Rotation (R_a)					Scaling (S_f)						PSNR
Parameter		10	30	50	70	90	0.5	0.8	1.1	1.4	1.7	2.0	
HU-PHT [33]	Lena	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	38.76
	Average	0.00	0.00	0.00	0.00	0.00	0.23	0.00	0.00	0.00	0.00	0.00	38.68
TANG-PZM [34]	Lena	0.20	0.98	0.00	0.00	0.00	0.59	0.00	0.00	0.20	0.00	0.20	38.91
	Average	0.99	0.81	0.54	0.80	0.00	3.75	0.04	0.05	0.15	0.15	0.34	38.86
FU-FrZM [35]	Lena	0.00	0.00	0.00	0.00	0.00	1.95	0.20	0.00	0.00	0.00	0.00	38.60
	Average	0.20	0.07	0.05	0.05	0.00	3.23	0.78	0.07	0.00	0.07	0.09	38.77
RRW-PCET	Lena	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	39.10
	Average	**0.05**	**0.07**	**0.01**	**0.04**	**0.00**	**0.19**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**39.22**
RRW-PCT	Lena	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	39.31
	Average	**0.04**	**0.08**	**0.00**	**0.06**	**0.00**	**0.42**	**0.00**	**0.00**	**0.00**	**0.01**	**0.02**	**39.08**
RRW-PST	Lena	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	38.95
	Average	**0.04**	**0.07**	**0.01**	**0.04**	**0.00**	**0.31**	**0.00**	**0.00**	**0.00**	**0.01**	**0.02**	**39.16**

watermark case. This is because the embedding strength is reduced to remain the same PSNRs and thus the robustness is degraded. For AWGN, the six geometric-moment-based methods have a notable increase in BER than WANG as they modify more geometric moments sensitive to AWGN. Compared with the other geometric-moment-based methods, our methods have lower BERs under different embedding capacity, indicating the effectiveness of the adaptive watermark embedding. For JPEG, the three proposed methods exhibit a slight improvement over TANG-PZM and FU-FrZM. This suggests that the adaptive normalization can effectively improve the robustness. For JPEG2000, our methods have strong robustness. Among them, BERs of RRW-PCET are slightly higher than those of RRW-PCT and RRW-PST because of the weaker numerical stability of PCET moments in comparison with PCT and PST moments [45]. For mean filtering, the average BERs of all methods except WANG are below 10%. This is because these methods employ global geometric moments capturing the image's global features as watermark embedding carriers and thus achieves resistance to filtering attacks effectively. For median filtering, the proposed three methods have similar robustness to the other three geometric-moment-based approaches, which further illustrates the effectiveness of the attack simulation test.

2) Robustness Against GD: As WANG embeds the watermark into the statistical histogram of image blocks, it can not counteract GD attacks and thus not assessed in this subsection. To evaluate the effectiveness of the other seven two-stage RRW methods against GD, we impose rotation and scaling on each watermarked image. The rotation angle R_a ranges from 10 to 90 degrees with step 20, and the scaling factor S_f varies from 0.5 to 2 with step 0.3. The robustness performance of these methods against rotation and scaling is summarized in Tables XIV and XV, where 256- and 512-bit watermarks are embedded.

Table XIV shows BERs of the compared methods under different rotation angles and scaling factors for 256-bit watermarks. All seven methods exhibit high robustness against rotation attacks. BERs of HU-Zernike and TANG-PZM are slightly higher than those of HU-PHT, FU-FrZM, and the three proposed methods, due to the weaker numerical stability of integer-order Zernike moments and PZMs. For scaling attacks, HU-PHT and the three proposed methods achieve excellent performance because of the scaling invariance property of PHT moments [45]. Their BERs remain close to zero across scaling factors from 0.5 to 2.

Table XV presents BERs of the compared methods under different rotation angles and scaling factors for 512-bit watermarks. For rotation attacks, the robustness of TANG-PZM significantly degrades for the 512-bit watermark case compared with the 256-bit watermark case because TANG-PZM adopts higher order PZMs that are more sensitive to rotation attacks than higher order PHT moments and fractional-order Zernike moments. For scaling attacks, BERs of TANG-PZM and FU-

FrZM are higher when the scaling factor decreases, due to the weak numerical stability of PZMs and fractional-order Zernike moments.

V. Conclusion

This paper presents a novel two-stage RRW-CG scheme that uses the attack-simulation-based adaptive normalization and embedding. The proposed scheme consists of three parts: 1) two-stage watermark embedding, which embeds a robust watermark for copyright protection and the auxiliary information for integrity authentication; 2) integrity authentication, which checks whether the received image has been attacked or not; and 3) watermark extraction and image reconstruction, which recovers both the robust watermark and the original image in case of no attack, or only extracts the robust watermark otherwise. The first part includes the robust and reversible embedding stages, in which the robust embedding stage develops the attack-simulation-based adaptive normalization and embedding while the reversible embedding stage hides the auxiliary information consisting of quantization and rounded errors and LSBs for integrity authentication with the rate-distortion approaching recursive code. Extensive experimental simulation results show that the proposed scheme has strong robustness against CSP such as AWGN, JPEG, JPEG2000, mean filtering, and median filtering, as well as GD including rotation and scaling in cases of 256- and 512-bit watermarks. Compared with the state-of-the-art methods, the proposed scheme obtains higher robustness under the same excellent invisibility, reversibility, and embedding capacity.

Although the proposed scheme has achieved promising performance against CSP and GD attacks, it has shortcomings such as the increase of computational complexity and the limited adaptability of the attack simulation test to various known or unknown practical attacks. The former is due to the fact that the step-by-step increasing of the embedding strength requires multiple rounds of watermark embedding and detection. The latter is because the attack simulation test cannot optimize the embedding strength for all types of practical attacks although it well resists most common attacks and obtains the satisfactory performance. These would be further investigated in our future work.

In addition, as the proposed scheme uses the global geometric moments as the watermark carrier, it has a limitation in resisting local geometric deformation attacks. To overcome this limitation, in the future research local geometric moments could be taken as the carrier and the watermark could be repeatedly embedded in the local geometrical moments from different regions of the image to improve robustness against local geometric deformations.

References

[1] A. Anand and A. K. Singh, "Health record security through multiple watermarking on fused medical images," *IEEE Trans. Computat. Social Syst.*, vol. 9, no. 6, pp. 1594–1603, Dec. 2022.

[2] G. Zhang, L. Zheng, Z. Su, Y. Zeng, and G. Wang, "M-sequences and sliding window based audio watermarking robust against large-scale cropping attacks," *IEEE Trans. Inf. Forensics Security*, vol. 18, pp. 1182–1195, 2023.

[3] Y. Zhang, X. Luo, J. Wang, Y. Guo, and F. Liu, "Image robust adaptive steganography adapted to lossy channels in open social networks," *Inf. Sci.*, vol. 564, pp. 306–326, Jul. 2021.

[4] P. Fan, H. Zhang, and X. Zhao, "Adaptive QIM with minimum embedding cost for robust video steganography on social networks," *IEEE Trans. Inf. Forensics Security*, vol. 17, pp. 3801–3815, 2022.

[5] P. Chowdhuri and B. Jana, "Hiding data in dual color images reversibly via weighted matrix," *J. Inf. Secur. Appl.*, vol. 50, Feb. 2020, Art. no. 102420.

[6] T. Zhang, X. Li, W. Qi, and Z. Guo, "Location-based PVO and adaptive pairwise modification for efficient reversible data hiding," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 2306–2319, 2020.

[7] W. He, G. Xiong, and Y. Wang, "Reversible data hiding based on adaptive multiple histograms modification," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 3000–3012, 2021.

[8] F. Yan, H. Huang, and X. Yu, "A multiwatermarking scheme for verifying medical image integrity and authenticity in the Internet of Medical Things," *IEEE Trans. Ind. Informat.*, vol. 18, no. 12, pp. 8885–8894, Dec. 2022.

[9] P. Pal, B. Jana, and J. Bhaumik, "Watermarking scheme using local binary pattern for image authentication and tamper detection through dual image," *Secur. Privacy*, vol. 2, no. 2, p. e59, Mar. 2019.

[10] S. Haddad, G. Coatrieux, A. Moreau-Gaudry, and M. Cozic, "Joint watermarking-encryption-JPEG-LS for medical image reliability control in encrypted and compressed domains," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 2556–2569, 2020.

[11] P. Chowdhuri, B. Jana, and D. Giri, "Secured steganographic scheme for highly compressed color image using weighted matrix through DCT," *Int. J. Comput. Appl.*, vol. 43, no. 1, pp. 38–49, Jan. 2021.

[12] O. Evsutin and K. Dzhanashia, "Watermarking schemes for digital images: Robustness overview," *Signal Process., Image Commun.*, vol. 100, Jan. 2022, Art. no. 116523.

[13] J. Biswapati, "Dual image based reversible data hiding scheme using weighted matrix," *Int. J. Electron. Inf. Eng.*, vol. 5, no. 1, pp. 6–19, Sep. 2016.

[14] J. Biswapati, G. Debasis, and M. S. Kumar, "Weighted matrix based reversible data hiding scheme using image interpolation," in *Computational Intelligence in Data Mining—Volume 2*, vol. 411. New Delhi, India: Springer, 2016, doi: 10.1007/978-81-322-2731-1_22.

[15] G. Debasis, J. Biswapati, and M. S. Kumar, "Dual image based reversible data hiding scheme using three pixel value difference expansion," in *Information Systems Design and Intelligent Applications*, vol. 434. New Delhi, India: Springer, 2016, doi: 10.1007/978-81-322-2752-6_40.

[16] B. Jana, "High payload reversible data hiding scheme using weighted matrix," *Optik*, vol. 127, no. 6, pp. 3347–3358, Mar. 2016.

[17] B. Jana, D. Giri, and S. K. Mondal, "Partial reversible data hiding scheme using (7, 4) Hamming code," *Multimedia Tools Appl.*, vol. 76, no. 20, pp. 21691–21706, Oct. 2017.

[18] B. Jana, D. Giri, and S. K. Mondal, "Dual image based reversible data hiding scheme using (7,4) Hamming code," *Multimedia Tools Appl.*, vol. 77, no. 1, pp. 763–785, Jan. 2018.

[19] P. Pal, P. Chowdhuri, and B. Jana, "Weighted matrix based reversible watermarking scheme using color image," *Multimedia Tools Appl.*, vol. 77, no. 18, pp. 23073–23098, Sep. 2018.

[20] B. Jana, "Reversible data hiding scheme using sub-sampled image exploiting Lagrange's interpolating polynomial," *Multimedia Tools Appl.*, vol. 77, no. 7, pp. 8805–8821, Apr. 2018.

[21] S. Meikap and B. Jana, "Directional PVO for reversible data hiding scheme with image interpolation," *Multimedia Tools Appl.*, vol. 77, no. 23, pp. 31281–31311, Dec. 2018.

[22] S. Mukherjee and B. Jana, "A novel method for high capacity reversible data hiding scheme using difference expansion," *Int. J. Natural Comput. Res.*, vol. 8, no. 4, pp. 13–27, Oct. 2019.

[23] J. He, J. Chen, and S. Tang, "Reversible data hiding in JPEG images based on negative influence models," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 2121–2133, 2020.

[24] W. He and Z. Cai, "An insight into pixel value ordering prediction-based prediction-error expansion," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 3859–3871, 2020.

[25] P. Pal, B. Jana, and J. Bhaumik, "A secure reversible color image watermarking scheme based on LBP, Lagrange interpolation polynomial and weighted matrix," *Multimedia Tools Appl.*, vol. 80, no. 14, pp. 21651–21678, Jun. 2021.

[26] P. K. Singh, B. Jana, and K. Datta, "Superpixel based robust reversible data hiding scheme exploiting Arnold transform with DCT and CA," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 34, no. 7, pp. 4402–4420, Jul. 2022.

[27] D. Coltuc and J.-M. Chassery, "Distortion-free robust watermarking: A case study," *Proc. SPIE*, vol. 6505, pp. 585–592, Feb. 2007.

[28] R. Kumar and K.-H. Jung, "Robust reversible data hiding scheme based on two-layer embedding strategy," *Inf. Sci.*, vol. 512, pp. 96–107, Feb. 2020.

[29] X. Liang, S. Xiang, L. Yang, and J. Li, "Robust and reversible image watermarking in homomorphic encrypted domain," *Signal Process., Image Commun.*, vol. 99, Nov. 2021, Art. no. 116462.

[30] E. Chrysochos, V. Fotopoulos, A. N. Skodras, and M. Xenos, "Reversible image watermarking based on histogram modification," in *Proc. 11th Panhellenic Conf. Inf. (PCI)*, 2007, pp. 93–104.

[31] C.-C. Chang, P.-Y. Lin, and J.-S. Yeh, "Preserving robustness and removability for digital watermarks using subsampling and difference correlation," *Inf. Sci.*, vol. 179, no. 13, pp. 2283–2293, Jun. 2009.

[32] R. Hu and S. Xiang, "Cover-lossless robust image watermarking against geometric deformations," *IEEE Trans. Image Process.*, vol. 30, pp. 318–331, 2021.

[33] R. Hu and S. Xiang, "Lossless robust image watermarking by using polar harmonic transform," *Signal Process.*, vol. 179, Feb. 2021, Art. no. 107833.

[34] Y. Tang, S. Wang, C. Wang, S. Xiang, and Y.-M. Cheung, "A highly robust reversible watermarking scheme using embedding optimization and rounded error compensation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 33, no. 4, pp. 1593–1609, Apr. 2023.

[35] D. Fu, X. Zhou, L. Xu, K. Hou, and X. Chen, "Robust reversible watermarking by fractional order Zernike moments and pseudo-zernike moments," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 33, no. 12, pp. 7310–7326, Dec. 2023.

[36] Y. Tang, K. Li, C. Wang, S. Bian, and Q. Huang, "A two-stage robust reversible watermarking using polar harmonic transform for high robustness and capacity," *Inf. Sci.*, vol. 654, Jan. 2024, Art. no. 119786.

[37] C. De Vleeschouwer, J.-F. 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.

[38] Z. Ni, Y. Q. 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.

[39] W. Wang, J. Ye, T. Wang, and W. Wang, "Reversible data hiding scheme based on significant-bit-difference expansion," *IET Image Process.*, vol. 11, no. 11, pp. 1002–1014, Nov. 2017.

[40] S. Xiang and Y. Wang, "Distortion-free robust reversible watermarking by modifying and recording IWT means of image blocks," *Digital-Forensics and Watermarking*, vol. 9569. Cham, Switzerland: Springer, 2016, doi: 10.1007/978-3-319-31960-5_28.

[41] F. Peng, W.-Y. Jiang, Y. Qi, Z.-X. Lin, and M. Long, "Separable robust reversible watermarking in encrypted 2D vector graphics," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 8, pp. 2391–2405, Aug. 2020.

[42] L. Xiong, X. Han, C. Yang, and Y. Shi, "Robust reversible watermarking in encrypted image with secure multi-party based on lightweight cryptography," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 1, pp. 75–91, Jan. 2022.

[43] D. Hou, W. Zhang, Y. Yang, and N. Yu, "Reversible data hiding under inconsistent distortion metrics," *IEEE Trans. Image Process.*, vol. 27, no. 10, pp. 5087–5099, Oct. 2018.

[44] X. Wang, X. Li, and Q. Pei, "Independent embedding domain based two-stage robust reversible watermarking," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 8, pp. 2406–2417, Aug. 2020.

[45] P.-T. Yap, X. Jiang, and A. C. Kot, "Two-dimensional polar harmonic transforms for invariant image representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 7, pp. 1259–1270, Jul. 2010.

[46] S. Qi, Y. Zhang, C. Wang, J. Zhou, and X. Cao, "A survey of orthogonal moments for image representation: Theory, implementation, and evaluation," *ACM Comput. Surv.*, vol. 55, no. 1, pp. 1–35, Jan. 2023.

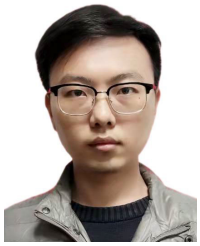
[47] B. Chen and G. W. Wornell, "Quantization index modulation: A class of provably good methods for digital watermarking and information embedding," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1423–1443, May 2001.

[48] L. Li, S. Li, A. Abraham, and J.-S. Pan, "Geometrically invariant image watermarking using polar harmonic transforms," *Inf. Sci.*, vol. 199, pp. 1–19, Sep. 2012.

[49] M. Alghoniemy and A. H. Tewfik, "Geometric invariance in image watermarking," *IEEE Trans. Image Process.*, vol. 13, no. 2, pp. 145–153, Feb. 2004.

[50] P. W. Holland and R. E. Welsch, "Robust regression using iteratively reweighted least-squares," *Commun. Statist., Theory Methods*, vol. 6, no. 9, pp. 813–827, Jan. 1977.

[51] P. Bas and T. Filler. (Jul. 2007). *Break our Watermarking System*. [Online]. Available: http://bows2.ec-lille.fr/



Yichao Tang received the B.S. and M.S. degrees from East China Jiaotong University in 2016 and 2020, respectively. He is currently pursuing the Ph.D. degree with South China Agricultural University.

His research interests include information security, robust watermarking, reversible data hiding, and AIGC security.

Chuntao Wang (Member, IEEE) received the B.S. and Ph.D. degrees from Sun Yat-sen University in 2002 and 2007, respectively.

From October 2007 to September 2008, he was a Post-Doctoral Fellow with Korea University, South Korea. From November 2008 to November 2010, he was a Post-Doctoral Researcher with Sun Yat-sen University. Currently, he is a Full Professor with the School of Mathematics and Informatics, South China Agricultural University. His research interests include information hiding, multimedia signal processing, and agricultural artificial intelligence.

Shijun Xiang (Member, IEEE) received the B.S. degree from Chang'an University in 1997, the M.S. degree from Guizhou University in 2000, and the Ph.D. degree from Sun Yat-sen University, China, in 2006.

From 2006 to 2007, he was a Post-Doctoral Researcher with Korea University, Seoul, South Korea. He is currently a Full Professor with the College of Information Science and Technology, Jinan University, Guangzhou, China. He has authored or coauthored more than 100 peer-reviewed articles, including IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, IEEE TRANSACTIONS ON IMAGE PROCESSING, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY , and IEEE TRANSACTIONS ON MULTIMEDIA. His current research interests include robust watermarking, reversible data hiding, and AIGC security.

Yiu-Ming Cheung (Fellow, IEEE) received the Ph.D. degree from the Department of Computer Science and Engineering, The Chinese University of Hong Kong, Hong Kong. He is currently a Chair Professor with the Department of Computer Science, Hong Kong Baptist University, Hong Kong. His current research interests include machine learning and visual computing, and their applications in data science, pattern recognition, multi-objective optimization, and information security. He is a fellow of AAAS, IET, BCS, and AAIA. He is the Editor-in-Chief of IEEE TRANSACTIONS ON EMERGING TOPICS IN COMPUTATIONAL INTELLIGENCE. Also, he serves as an Associate Editor for IEEE TRANSACTIONS ON CYBERNETICS, *Pattern Recognition*, *Knowledge and Information Systems*, *Pattern Recognition Letters*, and *Neurocomputing*, just to name a few. The details can be found at: https://www.comp.hkbu.edu.hk/ ymc.