Stereo image watermarking scheme for authentication with self-recovery capability using inter-view reference sharing

Ting Luo · Gangyi Jiang · Xiaodong Wang · Mei Yu · Feng Shao · Zongju Peng

Abstract Advances in three dimensional video is a strong stimulus for research in authentication of stereo image to avoid illegal modification. In this paper, a stereo image watermarking scheme is proposed for authentication with self-recovery capability using inter-view reference sharing. A mechanism of inter-view reference sharing in stereo image pairs is designed to reduce bits for recovery reference generation compared with independent references. Discrete wavelet transform coefficients are employed to generate the references, and two reference copies of each block embedded in two different mapping blocks are prepared for recover tamper. Moreover, detail information from high frequency coefficients is also embedded so as to improve the quality of tamper recovery. For the purpose of resisting collage attack and increasing the probability of tamper detection, disparities between pairs of matched blocks are checked to conduct tamper detection. Experimental results show that the proposed scheme can detect tampered blocks with the probabilities of more than 99 % after collage attack. When stereo images are cropped from 10 to 70 % with randomly tampering, they are recovered without losing main visual information and qualities

T. Luo · G. Jiang (✉) · X. Wang · M. Yu · F. Shao · Z. Peng
Faculty of Information Science and Engineering,
Ningbo University, Ningbo, China
e-mail: jianggangyi@126.com

T. Luo
e-mail: luoting@nbu.edu.cn

X. Wang
e-mail: wangxiaodong@nbu.edu.cn

M. Yu
e-mail: yumei2@126.com

F. Shao
e-mail: shaofeng@nbu.edu.cn

Z. Peng
e-mail: pengzongju@nbu.edu.cn

M. Yu
National Key Lab of Software New Technology,
Nanjing University, Nanjing 210093, China

of recovery are better than those of existing monocular image watermarking schemes extended to stereo images.

Keywords Three dimensional video · Stereo image watermarking · Block matching · Inter-view reference sharing

1 Introduction

Recently, three dimensional (3D) video and imaging technologies of creating the illusion of depth present new advances in video areas by offering viewers a stunning 3D experience [18, 19]. Stereo images captured for the same scene from two different cameras are a typical and efficient representation of 3D scenes. Due to wide availability of network access, stereo images are transmitted and exchanged between different terminals. However, the advent of image editing poses a serious threat to the integrity of stereo images, and people can easily manipulate or forge them without the permission of original authors [7]. Consequently, researches on the authentication of stereo images have become increasingly active, and watermarking technique is one of effective ways to keep the integrity of stereo images [20].

Authentication watermarking technologies are classified as semi-fragile and fragile ones [4]. Semi-fragile watermarking is relatively robust to common image processing, such as JPEG compression and filtering, but fragile to malicious attacks such as cropping and content replacement. Qin et al. presented two authentication measures with evaluating similarity between extracted and embedded watermarks, and measuring clustering level of tampered pixels, respectively [16]. Liu et al. used quantizing Zernike moment magnitudes of images to distinguish malicious attack from content-based image processing [14]. However, one of most challenging issues for semi-fragile watermarking is that no clear boundary between malicious and non-malicious attacks is authorized, such as what ratios of JPEG compression is malicious, and sometimes it is quite subjective. By contrast, fragile watermarking is sensitive to any pixel modification for avoiding missing any tamper [2]. He. et al. proposed a neighborhood-characteristic statistical fragile watermarking to locate tampered pixels with less than 14 % tampered areas [5]. But it cannot recover tamper, and authentication watermarking schemes are expected to have self-recovery capabilities.

Lin et al. presented a hierarchical watermarking scheme for tamper detection and self-recovery [12], in which images are divided into non-overlapping blocks, two least significant bits (LSB) of pixels are allocated for watermark embedding, and a recovery reference copy of each block is embedded into a unique mapping block. It cannot recover tampered blocks whose reference embedded into the other block is destroyed, which is called tampering coincidence problem. Thus, if tampered regions are extensive, many tampered blocks cannot be recovered from embedded recovery references. The problem also exists in [21]. To solve the problem, two reference copies of each block were embedded into two different mapping blocks in [10, 11]. Lee et al. improved Lin's scheme for qualities of tamper recovery in case large percents of images were modified [10]. However, because three LSBs of each pixel of images are employed to embed watermark, the quality of watermarked images is degraded. Li et al. also solved the coincidence problem with two LSBs [11], without degrading watermarked images much as Lee's scheme. However, the scheme uses 8×8 blocks rather than 4×4 of Lee's scheme [10], and if a pixel in a block is modified, the whole block is considered as tampered. As a result, the scheme increases fake pixels which are undamaged but detected as tampered. Accordingly, it decreases qualities of recovered images. In addition, limited coefficients of discrete cosine transform (DCT) are

employed to generate recovery references so that much detail information is lost. This affects qualities of recovered images as well. Average intensities are another way to be used to generate references in [12] and [10], and quality of reconstructed blocks is worse in complex texture regions than in flat regions. For authentication of integrity, tamper detection of each block is independent in [12] and [10], and thus, they cannot resist collage attack which is defined in [3]. In [11], two tamper detection methods with dependent blocks can identify collage attack, and fusion of two detection results is used to improve accuracy. However, the fusion will turn tampered blocks identified to undetected so that it decreases the probability of tamper detection.

In previous studies of authentication watermarking schemes with self-recovery, the focus was mainly placed on monocular images captured from only a single camera without depth. Stereo images are quite different from monocular images, and consist of left and right views [17]. More importantly, the two views are not independent of each other, but have high content correlations, which is inherent relationship of stereo images [22]. Therefore, recovery references of similar left and right view blocks can be generated together, so that bits of references can be reduced compared with generation for independent blocks. Moreover, tamper of stereo images is often symmetric for avoiding being observed visually, and relationship between tamper in left and right views can help tamper detection for stereo images. However, few efforts have been devoted to stereo image authentication watermarking. Most existing stereo image watermarking schemes were designed for copyright protection. Hwang et al. proposed stereo watermarking schemes based on DCT and discrete wavelet transform (DWT), respectively [8, 9], and they did not make use of relationships of stereo images. If relationships between two views are employed, performances of stereo image watermarking schemes can be improved [1]. Niu et al. achieved a trade-off between imperceptivity and robustness according to visual sensitivity of stereo vision [15]. Texture images plus depth images are another representation of 3D images, based on which Lin et al. presented a robust 3D image watermarking scheme to embed a watermark into texture images. Besides texture images, the watermark can also be extracted from left and right views which are rendered from texture images plus depth images [13]. However, 3D warping causes appearance of holes on rendered left and right views due to occlusion parts, which damages 3D vision. Thus stereo images consisting of two views are better representation of 3D images.

The problems in terms of degradation of watermarked images, tamper detection and recovery are still not resolved, if existing monocular image authentication watermarking schemes are directly applied to stereo images. Relationship between left and right views of stereo image is mined to overcome those problems in this paper. Firstly, block matching is employed for authentication, and it can resist collage attack and increase the probability of tamper detection. Secondly, reference sharing between left and right view blocks reduces bits of references. Moreover, coefficients of DWT are utilized to generate references without missing detail information for tamper self-recovery based on block matching and mapping. This paper proposes a stereo image watermarking for authentication with self-recovery capability using inter-view reference sharing. Inter-view relationship as one of important inherent characteristics between two views is utilized to achieve that most blocks in one view are matched with a unique similar block in the other view. Then, a mechanism of inter-view reference sharing is analyzed that a reference of one block is shared with its matched block, based on which the number of bits for keeping references is decreased and quality of watermarked stereo images is ensured. Detail DWT coefficients are not only used to detect tamper, but also reconstruct tampered blocks, so as to improve quality of recovered stereo image. Moreover, since tamper will destroy some of original relationship of block matching,

the relationship is checked again in tamper detection to improve accuracy of tamper detection. Experimental results demonstrate that the proposed scheme can detect successfully general pasting, cropping, and collage attack. When stereo images are cropped from 10 to 70 % sizes with randomly tampering, the proposed scheme outperforms existing monocular schemes [10, 12, 13], if they are extended to stereo images.

The remainder of this paper is organized as follows. Section 2 describes the mechanism of inter-view reference sharing in stereo image pairs. The proposed scheme is given in Section 3. Section 4 presents experimental results and discussions. Finally the conclusion is given in section 5.

2 Mechanism of inter-view reference sharing

For the purpose of recovering tamper, recovery reference bits are generated with approximation information of image blocks. Since there are high content correlations between left and right views of a stereo image, the mechanism of inter-view reference sharing is presented to save reference bits, in which reference of a block in one view is shared with its matched block in other view.

2.1 Reference generation

Let each view of a stereo image pair be of size $N_1 \times N_2$, where N_1 and N_2 are assumed to be a multiple of four. Left and right views of a stereo image are divided into non-overlapping blocks of 4×4 pixels denoted as $B_{i,j}^L$ and $B_{i,j}^R$, respectively, where $i=1,2,…, N_1/4$ and $j= 1,2,…, N_2/4$. As references of blocks should approximate themselves, approximation coefficients of DWT are employed. Each 4×4 block is decomposed by using one level of DWT, before which the two LSBs of each pixels of the block are firstly set to 0 since they will be used to embed watermark. The four sub-bands of DWT named LL, LH, HL, and HH are denoted as $D_{i,j}^\theta = \{D_{i,j}^\theta(k)|k = 1, 2, 3, 4\}$. Here, $\theta=1,2,3$, and 4 represent the four sub-bands LL, LH, HL, and HH, k indicates the kth DWT coefficient in each of the four sub-bands. $D_{i,j}^{\theta,L}$ and $D_{i,j}^{\theta,R}$ are with respect to left and right views, respectively. Since HH is not used for tampered block reconstruction, it is not taken into account in the proposed scheme. LH and HL contain horizontal and vertical edge information, respectively, and coefficients with small positive or negative numbers in LH and HL represent positively or negatively correlation along corresponding edges. Thus, if positive or negative signs of coefficients are recorded for detail information, it will be useful for tamper detection and recovery. Based on this point of view, the detail coefficient setting rule for reference and authentication generation denoted by Eq. (1) is utilized.

$$D_{i,j}^{m,\theta}(k) = \begin{cases} 0, & if\ D_{i,j}^\theta(k) \geq 0 \\ 1, & otherwise \end{cases}, \quad for\ \theta = 2,\ 3 \qquad (1)$$

Coefficients of LL are quantified as

$$X_{i,j}(k) = Round\left(D_{i,j}^1(k)/Q\right) \qquad (2)$$

where $Round(\cdot)$ returns the nearest integer to the parameter. $X_{i,j}^L$ and $X_{i,j}^R$ are with respect to left and right views, respectively. Let $S_{i,j}(k)$ with 5 bits be the binary representation of $X_{i,j}(k)$,

to quantize $X_{i,j}(k)$ into 5 bits, Q is set to 17. 20 bits denoted as $S_{i,j}$, $S_{i,j} = S_{i,j}(1)\|S_{i,j}(2)\|S_{i,j}(3)\|S_{i,j}(4)$, are employed to generate a reference, wherein "$\|$" denotes concatenation of bits. $S_{i,j}^L$ and $S_{i,j}^R$ are references of blocks in left and right views without detail information, respectively.

2.2 Inter-view reference sharing based on block matching

Block matching is defined as a block in one view of a stereo image pair is matched with a unique block in the other view with high correlations. Suppose stereo images are captured by parallel cameras, that is, only horizontal disparity is taken into account, and left and right views are chosen as the reference image and searching image, respectively. Let $d(i,j)$ denote a disparity for pairs of matched blocks, and is computed as

$$d(i, j) = \underset{d_t \in Z}{\arg\max} \left(count\left(D_{i,j}^{m,2,L}, \ D_{i-d_t,j}^{m,2,R} \right) \right) \qquad (3)$$

where $count(\cdot)$ returns the number of pairs satisfying that the corresponding coefficients of the two sub-bands are equal, $D_{i,j}^{m,2,L}$ and $D_{i,j}^{m,2,R}$ are $D_{i,j}^{m,2}$ with respect to left and right views respectively, and Z is defined as

$$Z = \left\{ d_t | \forall k \left(abs(X_{i,j}^L(k) - X_{i-d_t,j}^R(k)) \ \leq \ 1 \right), -\beta \leq d_t \leq \beta \right\} \qquad (4)$$

where $abs(\cdot)$ returns the absolute value, and β is the parameter of the searching range. The aim of Eq. (3) is to add horizontal image details to reconstruct tampered blocks. Thus, if one of a matched block pair is tampered, LH of DWT decomposed non-tampered block in the pair is used as LH of the tampered block.

Most blocks in left and right views are established as one-to-one matching, and their disparities are recorded. However, a few blocks in one view cannot find its matched blocks in the other view due to existing occlusion parts, and their disparities are set to $\beta+1$. Thus, blocks are divided into matched and unmatched types, and a disparity map is generated and denoted as D_p, which is a secret key.

Let $\beta=15$, then 5 bits can record one disparity value. Thus, compared with stereo images, the disparity map D_p takes up $5/(2 \times 8 \times 4 \times 4) = 1.95\ \%$. In addition, D_p can be compressed losslessly, thus bits of D_p can be decreased further. Ratios of D_p to different stereo images are listed in Table 1, in which D_p is compressed with Huffman coding. It is seen that the ratios are no more than 1.78 %, which can be acceptable for practical applications.

Let $B_{i,j}^L$ and $B_{i-d(i,j),j}^R$ denote a pair of matched blocks, $X_{i,j}^L(k)$ and $X_{i-d(i,j),j}^R(k)$ are their quantization values of LL, respectively. According to Eq. (4), for a pair of matched blocks, two bits are enough to record the difference between $X_{i,j}^L(k)$ and $X_{i-d(i,j),j}^R(k)$, denoted as $S_{i,j}^b(k)$, where the first bit is the sign, which is 0 if it is positive or 1 if it is negative. The other bit is the absolute difference, 0 or 1. The inter-view reference sharing for

Table 1 Ratios of D_p to stereo images [.%]

Alt Moabit	Laundry	Bowling	Book Arrival	Akko	Art
1.71	1.61	1.63	1.78	1.52	1.44

a pair of matched blocks is represented as.

$$C_{i,j} = S_{i,j} || S_{i,j}^b \tag{5}$$

where $S_{i,j}^b$ is defined as

$$S_{i,j}^b = S_{i,j}^b(1) || S_{i,j}^b(2) || S_{i,j}^b(3) || S_{i,j}^b(4) \tag{6}$$

It is noted that if the reference of a pair of matched blocks is not shared with each other, 2×20 bits are needed, but under the inter-view reference sharing, 28 bits are enough, therefore 12 bits are saved for a pair of matched blocks. A block in left view as shown in Fig. 1(a) is decomposed by using one level of DWT, and Fig. 1(d) shows the corresponding LL sub-band. The coefficients of LL are quantized as 20, 20, 20 and 20, respectively, $S_{i,j} = \{10100\ 10100\ 10100\ 10100\}$. In Fig. 1(b, e), the quantized values of LL in right view are 20, 19, 21 and 20, respectively. The left and right image blocks in Fig. 1 are matched according to Eq. (4), the corresponding $S_{i,j}^b = \{00\ 01\ 11\ 00\}$, and $C_{i,j} = \{10100\ 10100\ 10100\ 10100\ 00\ 01\ 11\ 00\}$.

3 The proposed scheme

3.1 Watermark embedding

Figure 2 shows the main flowchart of the proposed watermark embedding scheme. Stereo image blocks are divided into matched and unmatched types based on block matching. The mechanism of inter-view reference sharing is employed for reference generation of each pair of matched blocks. Watermark bits including authentication bits and reference bits are embedded into predefined embedding position, and watermarked stereo images are achieved.

Each view of a stereo image to be watermarked is divided into non-overlapping blocks of 4×4 pixels, and two LSBs of each pixel within all blocks are set to zero first. After that, each block is transformed by using one level of DWT, and LL is quantized according to Eq. (2), while $D_{i,j}^{m,2}$ and $D_{i,j}^{m,3}$ are obtained according to coefficients of LH and HL based on the
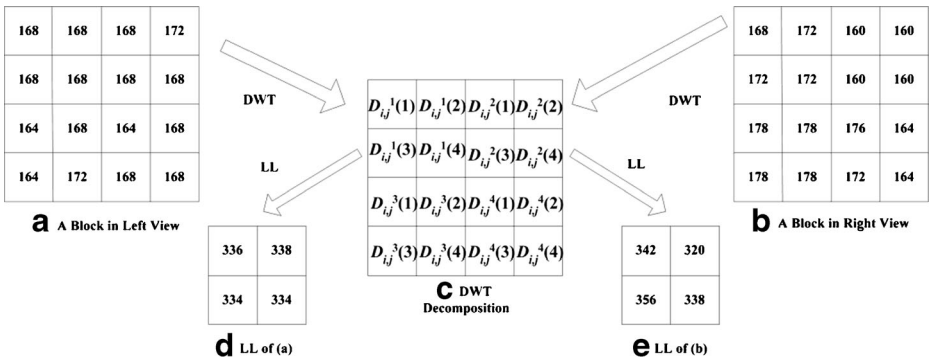

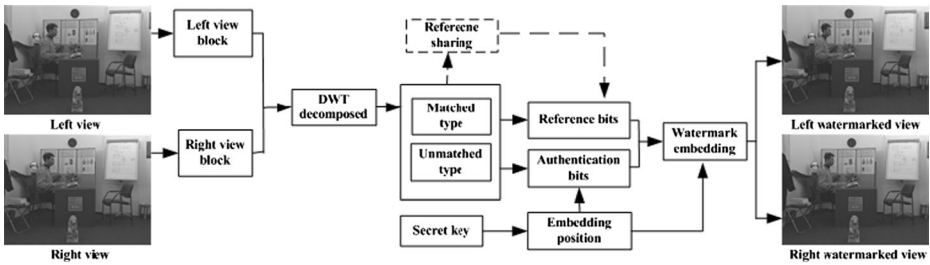Fig. 1 Sample of one level of DWT decomposition

Fig. 2 Flowchart of watermark embedding

detail coefficient setting rule denoted by Eq. (1). Then, all blocks in left and right views can be classified into matched or unmatched type using the block matching method mentioned in subsection 2.2.

Before watermark embedding, block mappings are established to determine embedding positions, and authentication and recovery reference bits of blocks are both embedded into their mapping blocks for tamper detection and recovery, respectively. Block mapping is established among blocks with the same type. Let N_m and N_u be the number of blocks for matched and unmatched types in each view, respectively, where $N_m + N_u = (N_1/4) \times (N_2/4)$. Matched blocks and unmatched blocks are labeled with a number from 1 to N_m and 1 to N_u, respectively, from left to right and top to bottom in both of views. Each matched block only has one unique mapping block in the different view. A secret key denoted as k_1 is used to generate two pseudo-random sequences with no-repeating integer values between 1 and N_m for matched blocks [6], denoted as r^L and r^R, respectively, where $r^L = \{r_t^L | t = 1, 2, \ldots, N_m\}$ and $r^R = \{r_t^R | t = 1, 2, \ldots, N_m\}$. r^L represents positions of mapping blocks in right view for all matched blocks in left view, that is, the matched block r_t^L in right view is the mapping block for the matched block t in left view. For example, if $r_2^L=10$, the 10th matched block in right view is the mapping block for the 2th matched block in left view. Similarly, r^R denotes mapping blocks in left view for all matched blocks in right view. Different from matched block, each unmatched block has two mapping blocks locating in two respective views. Four pseudo-random sequences with no-repeating integer values between 1 and N_u are generated with a secret key denoted as k_2 for the unmatched blocks. Two sequences indicate positions of mapping blocks in left and right views for unmatched blocks in left view, respectively, and the other two sequences represent positions of mapping blocks in left and right views for unmatched blocks in right view. Thus, one unmatched block is the same mapping block for two unmatched blocks in left and right views, respectively.

Since authentication and reference bits are generated and embedded in different ways for matched and unmatched blocks, they are described respectively as follows.

For a matched block, two LSBs of each pixel, that is, 32 bits of each block are allocated for watermark embedding, where reference and authentication bits are fixed as 28 and 4 bits, respectively. The watermark generation and embedding process for matched blocks consists of the following four steps.

Step a-1.   Suppose $B_{i,j}^L$ and $B_{i-d(i,j),j}^R$ are a pair of matched blocks, and their mapping blocks are obtained and denoted as $B_{i',j'}^R$ and $B_{i'',j''}^L$ using the secret key k_1, respectively, where $i',i''=1,2,\ldots,N_1/4$ and $j',j''=1,2,\ldots,N_2/4$. Reference bits $C_{i,j}$ for $B_{i,j}^L$ and $B_{i-d(i,j),j}^R$ is computed using Eq. (5).

Step a-2. In order to resist collage attack, authentication bit generation is not block independent. Authentication bits for $B_{i,j}^L$ embedded into $B_{i',j'}^R$ are denoted as $A_{i',j'}^R$ and computed as

$$A_{i',j'}^R = D_{i,j}^{m,3,L} \oplus D_{i',j'}^{m,3,R} \tag{7}$$

where $\oplus$ is the binary XOR operator, $D_{i,j}^{m,3,L}$ is $D_{i,j}^{m,3}$ of block $B_{i,j}^L$ in left view, and $D_{i',j'}^{m,3,R}$ is $D_{i,j}^{m,3}$ of block $B_{i',j'}^R$ in right view. Similarly, let $D_{i-d(i,j),j}^{m,3,R}$ be $D_{i,j}^{m,3}$ of block $B_{i-d(i,j),j}^R$ in right view, and $D_{i'',j''}^{m,3,L}$ be $D_{i,j}^{m,3}$ of block $B_{i'',j''}^L$ in left view, authentication bits of $B_{i-d(i,j),j}^R$ embedded in $B_{i'',j''}^L$ are denoted as $A_{i'',j''}^L$ and computed as

$$A_{i'',j''}^L = D_{i-d(i,j),j}^{m,3,R} \oplus D_{i'',j''}^{m,3,L} \tag{8}$$

Step a-3. Embedding bits in $B_{i',j'}^R$ and $B_{i'',j''}^L$ are denoted as $W_{i',j'}^R$ and $W_{i'',j''}^L$, respectively. Here $W_{i',j'}^R = A_{i',j'}^R || C_{i,j}$, and two LSBs of each pixel in $B_{i',j'}^R$ are replaced by

$$I_{i',j',m,n}^R = I_{i',j',m,n}^R + W_{i',j'}^R(z) \times 2^0 + W_{i',j'}^R(z+4) \times 2^1 \tag{9}$$

where I^R denotes pixel value of a block in right view whose two LSBs are already set to 0, and (m,n) denotes the index of pixels in blocks, here, $1 \le m$, $n \le 4$. z is the index of bits and $z = n + (m-1) \times 8$. $W_{i'',j''}^L$, denoted by $W_{i'',j''}^L = A_{i'',j''}^L || C_{i,j}$, is embedded into $B_{i'',j''}^L$ similar as Eq. (9), except that $W_{i',j'}^R$ and $I_{i',j',m,n}^R$ are replaced with $W_{i'',j''}^L$ and $I_{i'',j'',m,n}^L$. Benefiting from the mechanism of inter-view reference sharing, although each matched block only has one unique mapping block, reference copies are embedded twice in two different blocks.

Step a-4. Repeat the above three steps until watermark embedding of all matched blocks is finished.

For example, $B_{i,j}^L$ and $B_{i-d(i,j),j}^R$ are a pair of matched blocks, and $B_{i',j'}^R$ is the mapping block of $B_{i,j}^L$. Suppose $C_{i,j}$ is {10100 10100 10100 10100 00 01 11 00} based on Eq. (5), and according to the coefficient setting rule, $D_{i,j}^{m,2,L} = \{1011\}$, $D_{i,j}^{m,3,L} = \{1001\}$, $D_{i-d(i,j),j}^{m,2,R} = \{1011\}$, and $D_{i',j'}^{m,3,R} = \{1011\}$. Thus, $A_{i',j'}^R$ is {0010} by using Eq. (7). Finally, $W_{i',j'}^R = \{001010100\ 10100\ 10100\ 10100\ 00\ 01\ 11\ 00\}$ is embedded into $B_{i',j'}^R$.

For an unmatched block, three LSBs of each pixel, that is, 48 bits of each block totally are allocated for watermark embedding. Since an unmatched block is the mapping block of other two unmatched blocks in two respective views, 24 bits of each unmatched block are used to keep reference of one block in the same view, and 20 bits to keep reference of another block in the other view. The left 4 bits are utilized for embedding authentication bits. Thus, for each of unmatched blocks in left and right views, firstly set three LSBs of its each pixel to zero, and then transform the block with one level of DWT so as to obtain its $D_{i,j}^{u,2}$ and $D_{i,j}^{u,3}$ with respect to LH and HL, denoted by

$$D_{i,j}^{u,\theta}(k) = \begin{cases} 0, & \text{if } D_{i,j}^\theta(k) \ge 0 \\ 1, & \text{otherwise} \end{cases}, \quad \text{for } \theta = 2,3 \tag{10}$$

Suppose the unmatched blocks in left and right views denoted as $B_{i,j}^L$ and $B_{i',j'}^R$, respectively, have the same mapping block. The embedding process for unmatched block is portrayed as follows.

Step b-1. Let block $B_{i',j'}^L$ in left view be the same mapping block of $B_{i,j}^L$ and $B_{i'',j''}^R$, $S_{i,j}^L || D_{i,j}^{u,2,L}$ with 24 bits and $S_{i'',j''}^R$ with 20 bits which will be embedded in $B_{i',j'}^L$, are references of $B_{i,j}^L$ and $B_{i'',j''}^R$, respectively. Similarly, let block $B_{i''',j'''}^R$ in right view be the same mapping block of $B_{i,j}^L$ and $B_{i'',j''}^R$, then $S_{i,J}^l$ and $S_{i'',j''}^R || D_{i'',j''}^{u,2,R}$ which will be embedded in $B_{i''',j'''}^R$, are another copies of references of $B_{i,j}^L$ and $B_{i'',j''}^R$, respectively.

Step b-2. For the mapping block $B_{i',j'}^L$, the embedded authentication bits denoted as $A_{i',j'}^L$ are computed as

$$A_{i',j'}^L = D_{i,j}^{u,3,L} \oplus D_{i',j'}^{u,3,L} \tag{11}$$

For the mapping block $B_{i''',j'''}^R$, the embedded authentication bits denoted as $A_{i''',j'''}^R$ are calculated as

$$A_{i''',j'''}^R = D_{i'',j''}^{u,3,R} \oplus D_{i''',j'''}^{u,3,R} \tag{12}$$

Step b-3. For $B_{i',j'}^L$ in left view, the 48 bits embedded watermark $W_{i',j'}^L$ is defined as $W_{i',j'}^L = S_{i,j}^L || D_{i,j}^{u,2,L} || S_{i'',j''}^R || A_{i',j'}^L$, then three LSBs of each pixel in $B_{i',j'}^L$ are replaced by.

$$I_{i',j',m,n}^L = I_{i',j',m,n}^L + W_{i',j'}^L(z) \times 2^0 + W_{i',j'}^L(z+4) \times 2^1 + W_{i',j'}^L(z+8) \times 2^2 \tag{13}$$

where I^L denotes the pixel value in left view, and $z = n + (m-1) \times 12$. Similarly, if the mapping block is $B_{i''',j'''}^R$, let embedding bits be $W_{i''',j'''}^R$ and three LSBs of $I_{i''',j''',m,n}^R$ in $B_{i''',j'''}^R$ are replaced similar as Eq. (13), where $W_{i',j'}^L$ is instead of $W_{i''',j'''}^R$ and $W_{i''',j'''}^R = S_{i'',j''}^R || D_{i'',j''}^{u,2,R} || S_{i,j}^L || A_{i''',j'''}^R$.

Step b-4. Repeat the above three steps until watermark embedding for all unmatched blocks is finished.

For example, suppose $B_{i',j'}^L$ is the same mapping block of $B_{i,j}^L$ and $B_{i'',j''}^R$. Three LSBs of pixels within $B_{i,j}^L$ and $B_{i'',j''}^R$ are set to 0, and then one level of DWT is applied to them. As an example, if $S_{i,j}^L = \{11001\ 11001\ 11001\ 11011\}$, $S_{i'',j''}^R = \{10001\ 10000\ 10010\ 10011\}$, $D_{i,j}^{u,2,L} = \{1100\}$, $D_{i,j}^{u,3,L} = \{1100\}$ and $D_{i',j'}^{u,3,L} = \{1110\}$. Since according Eq. (11) $A_{i',j'}^L = \{0010\}$. $W_{i',j'}^L = \{11001\ 11001\ 11001\ 11011\ 110010001\ 10000\ 10010\ 100110010\}$ will be embedded into three LSBs of $B_{i',j'}^L$.

After finishing watermark embedding, watermarked stereo images are achieved.

3.2 Tamper detection

To implement tamper detection, first divide the received watermarked stereo image into non-overlapping blocks with the size of 4×4. According to the secret key D_p, all blocks are

classified into the matched or unmatched types. Then, watermark consisting authentication and recovery reference bits are extracted from each block, which are the inverse of Eq. (9) and Eq.(13) for the matched and unmatched blocks, respectively. Authentication bits are extracted from blocks to check whether the corresponding blocks are tampered or not, and recovery reference bits are used for tamper recovery. Let two tamper detection masks for left view blocks be $F_{1,i,j}^{L}$ and $F_{2,i,j}^{L}$, respectively, $F_{1,i,j}^{R}$ and $F_{2,i,j}^{R}$ be tamper detection masks for right view. The process of tamper detection is described as follows.

Step c-1. Firstly, two or three LSBs of each pixel within matched or unmatched blocks are set to zero, respectively. Then each block is decomposed by using one level of DWT, and $\widehat{D}_{i,j}^{m,\theta}(k)$ and $\widehat{D}_{i,j}^{u,\theta}(k)$ with respect to LH and HL sub-bands of matched and unmatched blocks are obtained to 0 or 1 similar as described by Eq. (1) and Eq. (10).

Step c-2. For a pair of matched blocks $\widehat{B}_{i,j}^{L}$ and $\widehat{B}_{i-d(i,j),j}^{R}$, and their mapping blocks $\widehat{B}_{i',j'}^{R}$ and $\widehat{B}_{i'',j''}^{L}$ are gained using secret key k_1, respectively. $\widehat{A}_{i',j'}^{R}$ and $\widehat{A}_{i'',j''}^{L}$ are extracted from $\widehat{B}_{i',j'}^{R}$ and $\widehat{B}_{i'',j''}^{L}$, respectively. XOR operator is performed on HL of decomposed $\widehat{B}_{i,j}^{L}$ and $\widehat{B}_{i',j'}^{R}$ to compute $\widetilde{A}_{i',j'}^{R}$. If $\widetilde{A}_{i',j'}^{R}$ and $\widehat{A}_{i',j'}^{R}$ are not equal, $\widehat{B}_{i,j}^{L}$ and $\widehat{B}_{i',j'}^{R}$ are marked as invalid, and $F_{1,i,j}^{L}$ and $F_{1,i',j'}^{R}$ are set to 1, otherwise they are set to 0.

In the same way, $\widetilde{A}_{i'',j''}^{L}$ is calculated by applying the XOR operator of adjusted HL of decomposed $\widehat{B}_{i-d(i,j),j}^{R}$ and $\widehat{B}_{i'',j''}^{L}$. $\widetilde{A}_{i'',j''}^{L}$ and $\widehat{A}_{i'',j''}^{L}$ are compared, and if they are not equal, $F_{2,i'',j''}^{L}$ and $F_{2,i-d(i,j),j}^{R}$ will be set to 1, otherwise they are set to 0. Step c-2 is repeated until all matched blocks are checked whether they are valid or not.

For example, following the matched block example in subsection 3.1, $\widehat{B}_{i,j}^{L}$ is supposed to be tampered, and suppose $\widehat{D}_{i,j}^{m,3,L}=\{1111\}$ is computed. $\widehat{A}_{i',j'}^{R}=\{0010\}$ is extracted from $\widehat{B}_{i',j'}^{R}$, and $\widehat{D}_{i',j'}^{m,3,R}=\{1011\}$ is computed from $\widehat{B}_{i',j'}^{R}$. Since $\widetilde{A}_{i',j'}^{R}=\widehat{D}_{i,j}^{m,3,L}\oplus\widehat{D}_{i',j'}^{m,3,R}$, $\widetilde{A}_{i',j'}^{R}$ is determined to be $\{0100\}$. $\widetilde{A}_{i',j'}^{R}$ is not equal with $\widehat{A}_{i',j'}^{R}$, and thus, $\widehat{B}_{i,j}^{L}$ and $\widehat{B}_{i',j'}^{R}$ are marked as invalid.
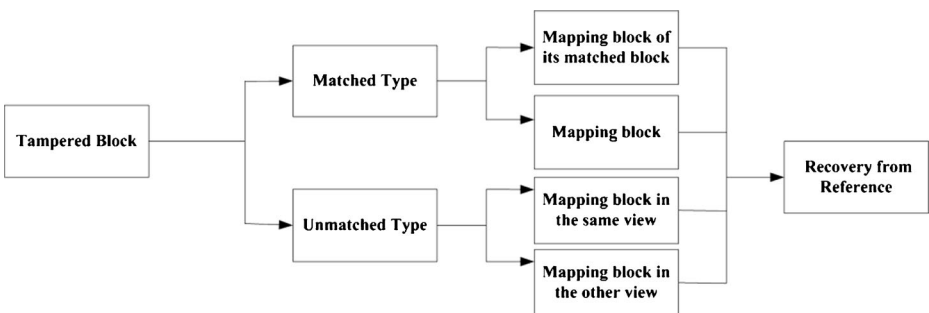

Fig. 3 Flowchart of tampered block recovery

Step c-3. For an unmatched block $\widehat{B}_{i,j}^{L}$ in left view, suppose $\widehat{B}_{i',j'}^{L}$ is its mapping block in the same view using secret key k_2, and $\widehat{A}_{i',j'}^{L}$ are extracted from $\widehat{B}_{i',j'}^{L}$. $\widetilde{A}_{i',j'}^{L}$ is computed from HL of $\widehat{B}_{i,j}^{L}$ and $\widehat{B}_{i',j'}^{L}$, where $\widehat{A}_{i',j'}^{L} = \widehat{D}_{i,j}^{u,3,L} \oplus \widehat{D}_{i',j'}^{u,3,L}$. If $\widehat{A}_{i',j'}^{L}$ is not equal with $\widetilde{A}_{i',j'}^{L}$, $F_{1,i,j}^{L}$ and $F_{i',j'}^{L}$ are set to 1, otherwise they are set to 0.

For an unmatched block $\widehat{B}_{i'',j''}^{R}$ and its mapping block $\widehat{B}_{i''',j'''}^{R}$ in right view, $\widehat{A}_{i''',j'''}^{R}$ is compared with $\widetilde{A}_{i''',j'''}^{R}$ computed using $\widehat{A}_{i''',j'''}^{R} = \widehat{D}_{i'',j''}^{u,3,R} \oplus \widehat{D}_{i''',j'''}^{u,3,R}$, and if they are not equal, $F_{1,i'',j''}^{R}$ and $F_{2,i''',j'''}^{R}$ are set to 1, otherwise they are set to 0. The step c-3 is repeated to check all unmatched blocks.

For example, following the above example of unmatched block in subsection 3.1, suppose $\widehat{B}_{i,j}^{L}$ is tampered, and $\widehat{D}_{i,j}^{u,3,L} = \{0011\}$ is computed. $\widehat{A}_{i',j'}^{L} = \{0010\}$ is extracted from $\widehat{B}_{i',j'}^{L}$ and $\widehat{D}_{i',j'}^{u,3,L} = \{1110\}$ is computed from $\widehat{B}_{i',j'}^{R}$. Thus, $\widetilde{A}_{i',j'}^{L} = \{1101\}$. Since $\widetilde{A}_{i',j'}^{L}$ is not equal with $\widehat{A}_{i',j'}^{L}$, both $\widehat{B}_{i,j}^{L}$ and $\widehat{B}_{i',j'}^{L}$ are marked as invalid.

Step c-4. After the initial F_{1}^{L}, F_{2}^{L}, F_{1}^{R} and F_{2}^{R} are obtained, morphological erosion and dilation are operated on the four masks. Then fusion is implemented on each of the two views, denoted by $F^{L} = F_{1}^{L} \& F_{2}^{L}$ and $F^{R} = F_{1}^{R} \& F_{2}^{R}$, where & is the binary AND operator. For example, in above examples of tamper detection $\widehat{B}_{i,j}^{L}$ is still marked as invalid while other blocks turns to valid after the fusion operation.

Step c-5. For each pair of matched blocks, whether they are still matched in the tampered stereo image is checked according to block matching method described in subsection 2.2, if they are not matched, both of them are marked as invalid, and the corresponding F^{L} and F^{R} are also updated to 1.

Step c-6. Erosion and dilation are operated on the updated F^{L} and F^{R} again. Finally, blocks are authenticated according to F^{L} or F^{R}.

3.3 Tamper self-recovery

If blocks are determined to be invalid according to F^{L} and F^{R}, the tamper self-recovery process as illustrated in Fig. 3 is implemented, where reference bits are extracted from either of two corresponding mapping blocks to reconstruct tampered blocks.

For a pair of matched blocks $\widehat{B}_{i,j}^{L}$ and $\widehat{B}_{i-d(i,j),j}^{R}$, if $\widehat{B}_{i,j}^{L}$ is invalid, the recovery steps are as follows.

Step d-1. The mapping block $\widehat{B}_{i',j'}^{R}$ of $\widehat{B}_{i,j}^{L}$ is obtained using secret key k_1. If $\widehat{B}_{i',j'}^{R}$ is not tampered, $\widehat{C}_{i,j}$ and $\widehat{A}_{i',j'}^{R}$ are extracted from $\widehat{B}_{i',j'}^{R}$. Moreover, set $\widehat{D}_{i,j}^{m,3,L}$ to be $\widehat{D}_{i',j'}^{m,3,R} \oplus \widehat{A}_{i',j'}^{R}$. If $\widehat{B}_{i-d(i,j),j}^{R}$ is valid, $\widehat{D}_{i,j}^{m,2,L}$ is set to $\widehat{D}_{i-d(i,j),j}^{m,2,R}$ extracted from $\widehat{B}_{i-d(i,j),j}^{R}$, otherwise, $\widehat{D}_{i,j}^{m,2,L}$ is not obtained, and then go to step d-2. If $\widehat{B}_{i',j'}^{R}$ is tampered, the mapping block $\widehat{B}_{i'',j''}^{L}$ of $\widehat{B}_{i-d(ij),j}^{R}$ is checked, and if it is valid, only $\widehat{C}_{i,j}$ is obtained from $\widehat{B}_{i'',j''}^{L}$, and then go to step d-2. If $\widehat{B}_{i'',j''}^{L}$ is tampered as well, $\widehat{B}_{i,j}^{L}$ cannot be recovered from recovery references, and the spatial interpolation method introduced in the follows will be employed.

Step d-2.   The former 20 bits of $\widehat{C}_{i,j}$ are used for recovery, where every 5 bits generate an integer denoted as $\widehat{X}^L_{i,j}(k)$. Then LL coefficients are computed as

$$\widehat{D}^{1,L}_{i,j}(k) = \widehat{X}^L_{i,j}(k) \times Q \tag{14}$$

Step d-3.   If $\widehat{D}^{m,2,L}_{i,j}$ (or $\widehat{D}^{m,3,L}_{i,j}$) is 0, the corresponding $\widehat{D}^{2,L}_{i,j}$ (or $\widehat{D}^{3,L}_{i,j}$) with respect to coefficient of LH (or HL) is set to α, otherwise, set to $-\alpha$. α is the parameter and discussed in the follows. If $\widehat{D}^{m,2,L}_{i,j}$ or $\widehat{D}^{m,3,L}_{i,j}$ can not be obtained, the coefficients of LH or HL will be set to 0. Coefficients of HH are always set to 0.

Step d-4.   Finally $\widehat{B}^L_{i,j}$ is recovered by using the reverse of DWT.

If $\widehat{B}^R_{i-d(i,j),j}$ is tampered, $\widehat{D}^{m,2,R}_{i-d(i,j),j}$ and $\widehat{D}^{m,3,R}_{i-d(i,j),j}$ are obtained in the similar way as the step d-1 and step d-3. The only different is in step d-2, the latter 8 bits of $\widehat{C}_{i,j}$ are used to compute $\widehat{X}^R_{i-d(i,j),j}(k)$ from $\widehat{X}^L_{i,j}(k)$, and the method is the reverse of inter-view reference sharing generation. LL is computed according to Eq. (14) as well, and then the reverse of DWT is applied to recover $\widehat{B}^R_{i-d(i,j),j}$. All matched blocks are recovered in the above way.

Following the above matched block example, $\widehat{B}^L_{i,j}$ needs to be recovered. $\widehat{D}^{1,L}_{i,j} = \{20\ 20\ 20\ 20\} \times Q$ is computed based on Eq. (14), $\widehat{D}^{m,2,L}_{i,j} = \{1011\}$, and $\widehat{D}^{m,3,L}_{i,j} = \{1001\}$. Moreover, according to $\widehat{D}^{m,2,L}_{i,j}$ and $\widehat{D}^{m,3,L}_{i,j}$, $\widehat{D}^{2,L}_{i,j}$ and $\widehat{D}^{3,L}_{i,j}$ are obtained to be $\{-\alpha\ \alpha\ -\alpha\ -\alpha\}$ and $\{-\alpha\ \alpha\ \alpha\ -\alpha\}$, respectively. Coefficients of HH are set to 0, and $\widehat{B}^L_{i,j}$ is recovered by using the reverse of DWT.

For an invalid unmatched block, the recovery process is different from that of a matched block.

Step e-1   The first mapping block in the same view of the invalid unmatched block is gained. If it is tampered, go to step e-2. Otherwise, the former 24 bits of recovery embedded in the mapping block are extracted, which is the reverse of step b-3, wherein 20 bits and 4 bits are information of LL and LH sub-bands of the invalid unmatched block, respectively. Moreover, information of HL sub-band of the invalid block is computed by applying XOR operation on authentication bits from the mapping block and $\widehat{D}^{u,3}$ obtained with the detail coefficient setting rule from HL sub-band of the decomposed mapping block. Then coefficients of HL and LH of the invalid unmatched block will be set to α or $-\alpha$, similar as what is done in step d-3. Coefficients of HH of the invalid unmatched block are set to 0. Then go to step e-3 directly.

Step e-2   The second mapping block in the other view of the invalid unmatched block is gained. If it is tampered, the spatial interpolation recovery method is employed. Otherwise, the latter 20 bits of recovery bits for block recovery is obtained, which is the reverse of step b-3 as well. Coefficients of LL sub-band of the invalid unmatched block is achieved based on Eq. (14), and coefficients of other three sub-bands are set to 0.

Step e-3   The invalid unmatched block is recovered by using the reverse of DWT. Repeat step e-1 to step e-3 until all invalid unmatched blocks are operated.
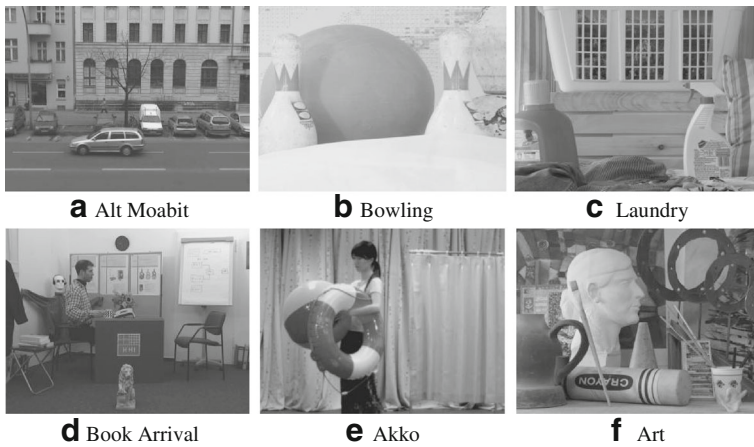
Fig. 4 Left views of test stereo images. **a** Alt Moabit, **b** Bowling, **c** Laundry, **d** Book Arrival, **e** Akko, **f** Art

For example, following the above unmatched block example, $\widehat{B}_{i,j}^L$ is to be recovered. $\widehat{S}_{i,j}^L$ and $\widehat{D}_{i,j}^{u,2,L} = \{1100\}$ are extracted from $\widehat{B}_{i',j'}^L$, and $\widehat{D}_{i,j}^{1,L} = \{25\ 25\ 25\ 27\} \times Q$ is computed based on Eq. (14). Since $\widehat{D}_{i',j'}^{u,3,L} = \{1110\}$ and $\widehat{A}_{i',j'}^{L} = \{0010\}$, $\widehat{D}_{i,j}^{u,3,L} = \{1100\}$. Moreover, according to $\widehat{D}_{i,j}^{u,2,L}$ and $\widehat{D}_{i,j}^{u,3,L}$, $\widehat{D}_{i,j}^{2,L}$ and $\widehat{D}_{i,j}^{3,L}$ are set to $\{-\alpha\ -\alpha\ \alpha\ \alpha\}$ and $\{-\alpha\ -\alpha\ \alpha\ \alpha\}$, respectively. Coefficients of HH are set to 0, and then $\widehat{B}_{i,j}^L$ is recovered by using the reverse of DWT.

After recovery from references, there may be a few invalid blocks without being recovered. Pixels in such invalid blocks will be recovered by spatial interpolation with their nearest pixels, denoted by

$$I_{x,y} = \lambda \times \left(u_1 \times I_{x,y-1} + (1 - u_1) \times I_{x,y+1} \right) + (1 - \lambda)$$
$$\times \left(u_2 \times I_{x-1,y} + (1 - u_2) \times I_{x+1,y} \right) \qquad (15)$$

where $1 \leq x \leq N_1$ and $1 \leq y \leq N_2$. u_1 and $1-u_1$ are used to weigh the contributions from the pixels on either side of the tampered pixel, and u_2 and $1-u_2$ are used to weigh the contributions from those above and below the tampered pixel. u_1 and u_2 are functions of the distances between the lost pixel and its closest valid neighbor pixels lying on the same row and column, respectively. λ determines weights of neighbor pixels on the same row and column. Edge detection is employed in two views, and if nearest neighbor pixels locate on horizontal or vertical edges, $\lambda=1$ or $\lambda=0$, otherwise, $\lambda=0.5$ for flat regions.

Table 2 Ratios of matched type of blocks in tested stereo image pairs [.%]

Alt Moabit	Bowling	Laundry	Book Arrival	Akko	Art
81.87	87.29	82.31	85.29	82.79	80.16

a Alt Moabit   **b** Bowling   **c** Laundry

d Book Arrival   **e** Akko   **f** Art
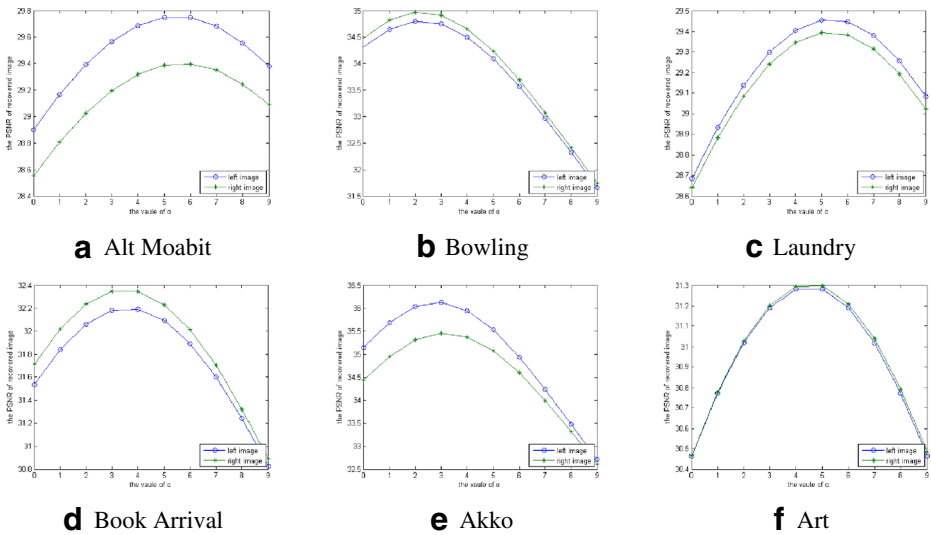
Fig. 5 Relationships between α and PSNR of recovered images. **a** Alt Moabit, **b** Bowling, **c** Laundry, **d** Book Arrival, **e** Akko, **f** Art

The parameter α is relative to the quality of recovery. In order to determine how α is related to quality of recovery, function $f(\alpha)$ is established for peak signal-to-noise ratio (PSNR) of reconstructed blocks, $f(\alpha) = 10 \times \log_{10}(255 \times 255/g(\alpha))$, and $g(\alpha)$ is defined as

$$g(\alpha) = \frac{1}{16} \sum_{m=1}^{4} \sum_{n=1}^{4} \left(I_{i,j,m,n} - \widehat{I}_{i,j,m,n} \right)^2 \qquad (16)$$

where I and $\widehat{I}$ are original pixels and the reconstructed pixels, respectively. $\widehat{I}$ is obtained from the reverse of DWT, wherein coefficients of HH are set to 0.

$$\widehat{I}_{i,j} = u_{21}(u_{12}(X_{i,j}(k) \times Q)^*h_0 + \alpha \times u_{12}(D^2{}_{i,j})^*h_1)^*h_0 + \\ u_{21}(\alpha \times u_{12}(D^3{}_{i,j})^*h_0)^*h_1 \qquad (17)$$

where * is the convolution, h_0 and h_1 are the low and high pass filter, respectively. $u_{12}(\cdot)$ and $u_{21}(\cdot)$ denote row-based upsampling and columns-based upsampling, respectively. Suppose $P_{1,i,j} = u_{21}(u_{12}(X_{i,j}(k) \times Q)^*h_0)^*h_1$, and $P_{2,i,j} = u_{21}(u_{12}(D^2_{i,j})^*h_1)^*h_0 + u_{21}(u_{12}(D^3_{i,j})^*h_0)^*h_1$, $g(\alpha)$ is simplified as

$$g(\alpha) = \frac{1}{16} \sum_{m=1}^{4} \sum_{n=1}^{4} (I_{i,j,m,n} - P_{1,i,j,m,n} - \alpha \times P_{2,i,j,m,n})^2 \qquad (18)$$

where $g(\alpha)$ is the quadratic equation, thus $f(\alpha)$ gets maximum value relative to a unique α. The parameter α will be discussed in the experiment section as well.

4 Experimental results and discussion

In order to substantiate the effectiveness of the proposed scheme, a series of experiments are tested on six stereo images with the size of 480×640, which are the first frames of 'Alt
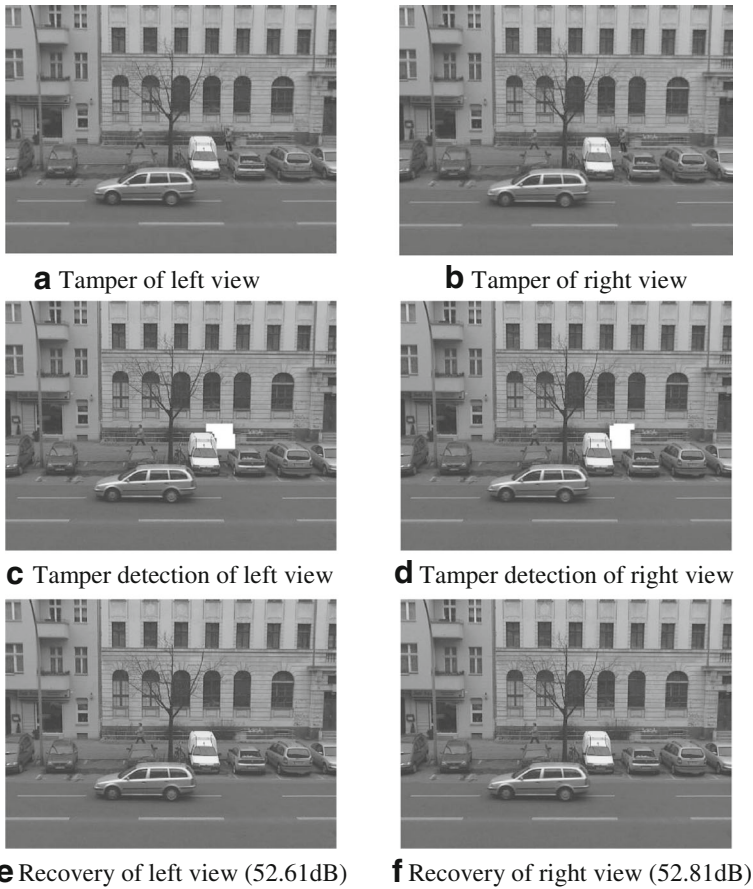
a Tamper of left view   **b** Tamper of right view

c Tamper detection of left view   **d** Tamper detection of right view

e Recovery of left view (52.61dB) **f** Recovery of right view (52.81dB)

Fig. 6 Tamper, detection and recovery of 'Alt Moabit'. **a** Tamper of *left view*, **b** Tamper of *right view*, **c** Tamper detection of *left view*, **d** Tamper detection of *right view*, **e** Recovery of *left view* (52.61 dB), **f** Recovery of *right view* (52.81 dB)

Moabit', 'Bowling', 'Laundry', 'Book Arrival', 'Akko' and 'Art', and their left views are shown in Fig. 4. Moreover, Lin's scheme in [12], Lee's scheme in [10] and Li's scheme in [11] are also applied to these stereo images wherein the two views are considered to be independent with each other. The three comparison schemes are named as Lin's scheme, Lee's scheme, and Li's scheme hereinafter.

Table 3 Comparisons in tamper detection and recovery of 'Alt Moabit'

	Proposed scheme		Lin's scheme		Lee's scheme		Li's scheme	
	Left	Right	Left	Right	Left	Right	Left	Right
$PTD(.\%)$	100	100	100	100	98.23	98.75	91.12	90.15
$TR(.dB)$	52.61	52.81	52.65	52.54	51.93	52.01	47.12	47.23

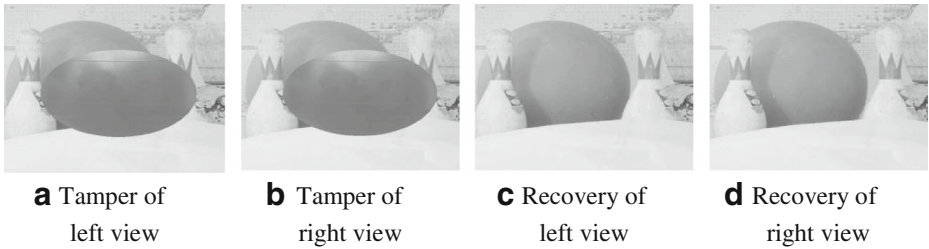| **a** Tamper of | **b** Tamper of | **c** Recovery of | **d** Recovery of |
| left view | right view | left view | right view |

Fig. 7 Tamper and recovery of 'Bowling'. **a** Tamper of *left view*, **b** Tamper of *right view*, **c** Recovery of *left view*, **d** Recovery of *right view*

Table 2 shows that more than 80 % of blocks are matched type for these different test stereo images, thus most of blocks only use two LSBs of each pixel to embed watermark, which means that the transparency of watermark can be satisfied. PSNRs of the watermarked images obtained with the proposed scheme are around 42 dB and higher than that of Lee's scheme with 40 dB. Although PSNRs of Lin's and Li's schemes are higher with around 44 dB and 43 dB, respectively, PSNR of the proposed scheme is satisfied for watermark transparency.

4.1 Parameter α

As information from LH and HL sub-bands provides details to the recovered stereo images, the parameter α is important to quality of recovery. Stereo images are reconstructed by using Eq. (17), and PSNRs between the original and reconstructed images are computed. In order to optimize PSNR, relationships between α and PSNR are built for the six test stereo images shown in Fig. 5, where PSNR has the maximum relative to a unique proper α according with $f(\alpha)$. α is determined as the value that achieves the maximum average PSNR of left and right views of the watermarked stereo image, and it is also a secret key for the watermarked stereo image in the proposed scheme.

4.2 General pasting

Stereo images are modified symmetrically, otherwise, tamper will be discovered visually and easily. Thus, in the following experiments, both of views are tampered simultaneously. Firstly, an extra person is pasted on 'Alt Moabit' with tamper ratios of around 0.4 % as illustrated in Fig. 6(a, b), and users hardly differentiate the fake one from the real one. The probabilities of tamper detection (*PTD*) of the proposed scheme for both views reach 100 %

Table 4 Comparisons in tamper detection and recovery of 'Bowling'

	Proposed scheme		Lin's scheme		Lee's scheme		Li's scheme	
	Left	Right	Left	Right	Left	Right	Left	Right
PTD(.%)	99.43	99.17	99.89	99.93	98.02	98.12	91.73	93.27
TR(.dB)	39.12	38.89	36.71	36.83	36.88	36.32	30.32	30.02

a Tamper of left view   **b** Tamper of right view   **c** Recovery of left view   **d** Recovery of right view
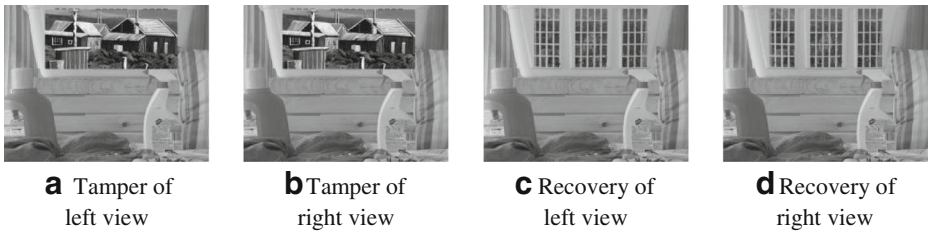
Fig. 8 Tamper and recovery of 'Bowling'. **a** Tamper of *left view*, **b** Tamper of *right view*, **c** Recovery of *left view*, **d** Recovery of *right view*

as shown in Fig. 6(c, d), where pixels of the identified tampered blocks are set to 255. *PTD* of the proposed scheme is same as that of Lin's scheme, and higher than those of the other two schemes as listed in Table 3. On the other hand, the tamper is recovered without perception as shown in Fig. 6(e, f), PSNRs of recovered left and right views relative to the watermarked images are 52.61 dB and 52.81 dB, respectively, higher than those of Lee's scheme and Li's scheme, and close to that of Lin's scheme. Therefore, the proposed scheme is able to detect tiny tamper correctly and recover it close to the original.

In the second experiment, tamper is expanded and a 'Ball' is pasted on smooth regions in both of views with tamper ratios of around 33 %, as shown in Fig. 7(a, b). Except Li's scheme, other three schemes detect tamper with *PTD* of around 99 %, as shown in Table 4. This suggests that fusion operation decreases *PTD* in Li's scheme with *PTD* of less than 94 %, and in the proposed scheme block matching is checked for increasing *PTD*. Figure 7(c, d) are final recoveries of the proposed scheme, and they are close to original stereo images visually. PSNRs of the stereo images recovered by the proposed scheme are much higher than that with respect to other three schemes of more than 2 dB.

In the third experiment, the texture background of 'Laundry' is modified as shown in Fig. 8(a, b). Similar as the results of the second experiment, more than 99 % of tamper can be identified by Lin's scheme, Lee's scheme and the proposed scheme, as shown in Table 5, *PTD* of Li's scheme is the lowest. The recovered quality of the proposed scheme is the best with PSNR of 31.99 dB and 31.39 dB for left and right views, respectively, and it out-performs other three schemes of at least more than 1 dB. Recoveries of the texture background are clear and almost unperceivable by human eyes as shown in Fig. 8(c, d).

In the above three experiments, the proposed scheme detects tamper from tiny to extensive and from smooth region to texture background correctly, and provides better recovery capabilities compared with the other three schemes.

Table 5 Comparisons in tamper detection and recovery of 'Laundry'

	Proposed scheme		Lin's scheme		Lee's scheme		Li's scheme	
	Left	Right	Left	Right	Left	Right	Left	Right
PTD(.%)	99.58	99.45	99.90	99.85	99.57	99.42	90.12	90.23
TR(.dB)	31.99	31.39	28.46	28.37	30.89	30.72	28.12	28.34

a Tamper of left view   **b** Tamper of right view

c Detection of   **d** Detection of   **e** Recovery of   **f** Recovery of
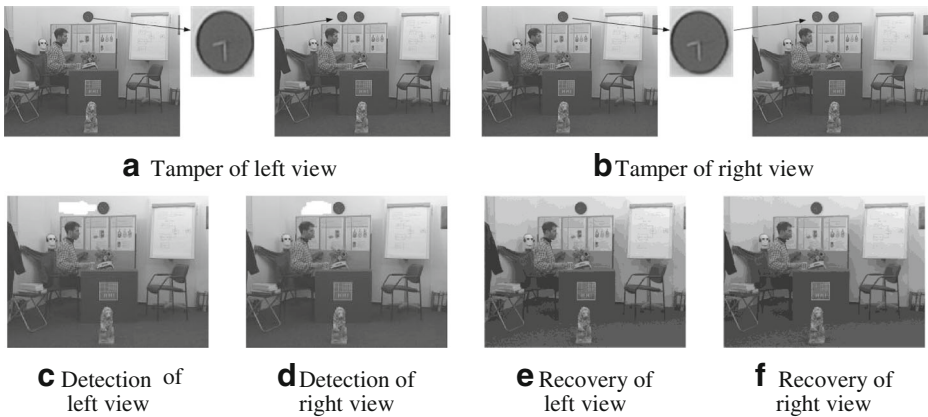left view   right view   left view   right view

Fig. 9 Tamper, detection, recovery of 'Book Arrival'. **a** Tamper of *left view*, **b** Tamper of *right view*, **c** Detection of *left view*, **d** Detection of *right view*, **e** Recovery of *left view*, **f** Recovery of *right view*

4.3 Collage attack

Collage attack is another common way to modify stereo images. Figures 9(a, b) show that the 'clock' copied from the watermarked 'Book arrival' is pasted on the same views. Lin's scheme and Lee's scheme fail to detect them with *PTDs* of 0 as shown in Table 6 due to block independence of authentication bits. The proposed scheme detects the tamper of both views as illustrated in Fig. 9(c, d) with *PTDs* of 99.89 % and 99.79 %, respectively, which are much higher than those of Li's scheme with 87.65 % and 89.12 %. However, some valid blocks are identified as tampered blocks by mistake as shown in Fig. 9(c, d), and the proposed scheme with high recovery capability hides this kind of weakness and recovers the tamper imperceptibly as shown in Fig. 9(e, f). PSNRs of tamper recoveries (*TR*) for both views arrive to 57.21 dB and 57.81 dB, respectively, which are nearly 9 dB higher than those of Li's scheme. Lin's scheme and Lee's scheme can not recover this kind of tamper due to the failure of tamper detection.

'Akko' and 'Art' are watermarked using the same watermarking schemes, and "Akko" is modified by pasting a sculpture head of 'Art' on it as shown in Fig. 10(a, b). Figure 10(c, d, e, f) show the tamper detection results of the proposed scheme, Lin's scheme, Lee's scheme, and Li's scheme in left view, respectively. Lin's scheme and Lee's scheme can not detect the tamper either, only edges of tamper are identified, and the corresponding *PTRs* are less than 8 % as shown in Table 7. *PTRs* of the proposed scheme achieve 99.41 % and 99.22 % for left and right views, respectively, which are much higher than those of Li's scheme. Although,

Table 6 Comparisons in tamper detection and recovery of 'Book Arrival'

	Proposed scheme		Lin's scheme		Lee's scheme		Li's scheme	
	Left	Right	Left	Right	Left	Right	Left	Right
PTD(.%)	99.89	99.79	0	0	0	0	87.65	89.12
TR(.dB)	57.21	57.81	30.62	30.64	30.87	30.86	48.23	48.56

a Tamper of left view   **b** Tamper of right view   **c** Left view detection of the proposed scheme   **d** Left view detection of Lin's scheme

e Left view detection of Lee's scheme   **f** Left view detection of Li's scheme   **g** Left view recovery of the proposed scheme   **h** Left view recovery of Li's scheme
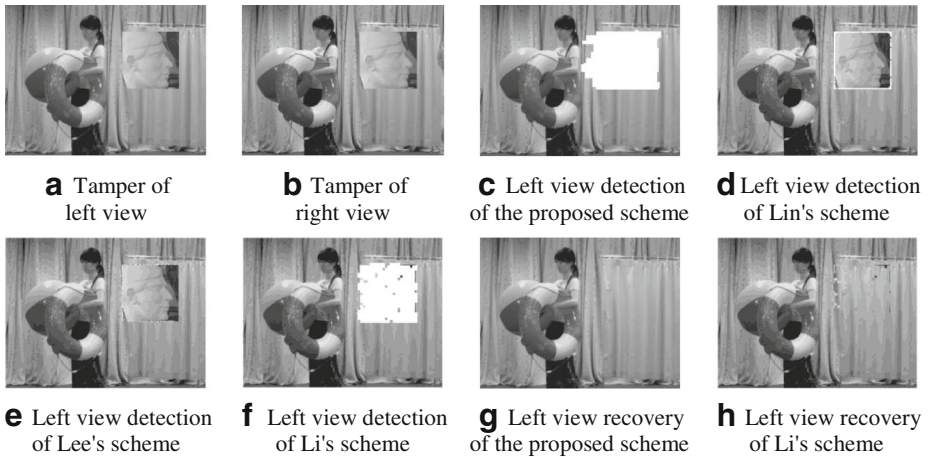
Fig. 10 Tamper, detection and recovery of 'Akko'. **a** Tamper of *left view*, **b** Tamper of *right view*, **c** *Left view* detection of the proposed scheme, **d** *Left view* detection of Lin's scheme, **e** *Left view* detection of Lee's scheme, **f** *Left view* detection of Li's scheme, **g** *Left view* recovery of the proposed scheme, **h** *Left view* recovery of Li's scheme

some valid blocks are still falsely identified as tampered as shown in Fig. 10(c), PSNR of the proposed scheme is more than 10 dB higher than that of Li's scheme. Fig. 10(g, h) show the left view recoveries of Li's scheme and the proposed scheme, respectively, and the visual quality of the proposed scheme is still better.

In the proposed scheme, if a pair of matched blocks are not matched after modification of the stereo image, both of which are identified as invalid. It leads to a few non-tampered blocks to be identify as invalid blocks by mistake, but the ratio of mistake is acceptable and high capability of block reconstruction ensures the quality of stereo image recovery. Based on experiments, the proposed scheme resists collage attack with better recovery capability than other three schemes.

4.4 Cropping with different sizes

In order to demonstrate superiority in recovery of the proposed scheme again, Lin's scheme, Lee's scheme and Li's scheme are further extended to use the disparity map D_p for recoveries. To achieve better recovery results, Eq.(4) is modified by using '=0' instead of '≤1' to compute D_p for these three schemes, which is employed for recovery after using

Table 7 Comparisons in tamper detection and recovery of 'Akko'

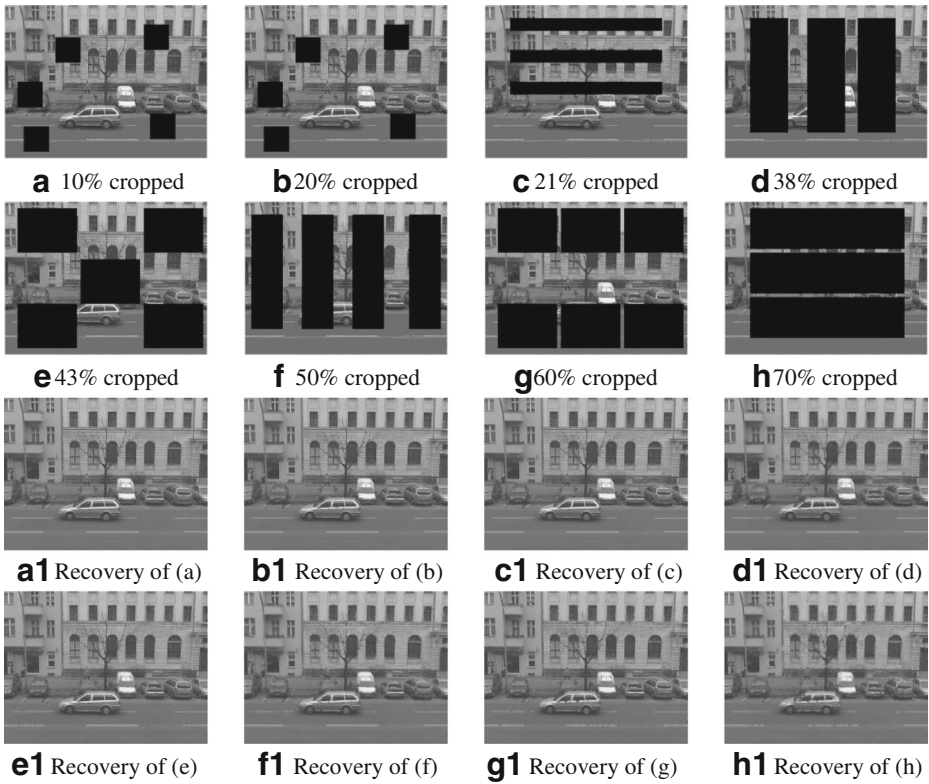|	Proposed scheme		Lin's scheme		Lee's scheme		Li's scheme	
	Left	Right	Left	Right	Left	Right	Left	Right
$PTR(.\%)$	99.41	99.22	7.81	7.93	5.44	5.14	95.24	94.91
$TR(.dB)$	44.27	44.26	27.23	27.56	26.11	26.23	33.37	33.12

Fig. 11 Recoveries of randomly tampering for 'Alt Moabit'. **a** 10 % cropped, **b** 20 % cropped, **c** 21 % cropped, **d** 38 % cropped, **e** 43 % cropped, **f** 50 % cropped, **g** 60 % cropped, **h** 70 % cropped. *a1* Recovery of (**a**), (*b1*) Recovery of (**b**), (*c1*) Recovery of (**c**), (*d1*) Recovery of (**d**), (*e1*) Recovery of (**e**), (*f1*) Recovery of (**f**), (*g1*) Recovery of (**g**), (*h1*) Recovery of (**h**)

references. The improved three schemes are denoted as Lin's scheme + D_p, Lee's scheme + D_p, Li's scheme + D_p, respectively.

Cropping is a general tamper as well and the cropped pixels are set to 0 in the experiments. 'Alt Moabit' is cropped from 10 % to 70 % with randomly tampering, and Fig. 11(a-h) show the
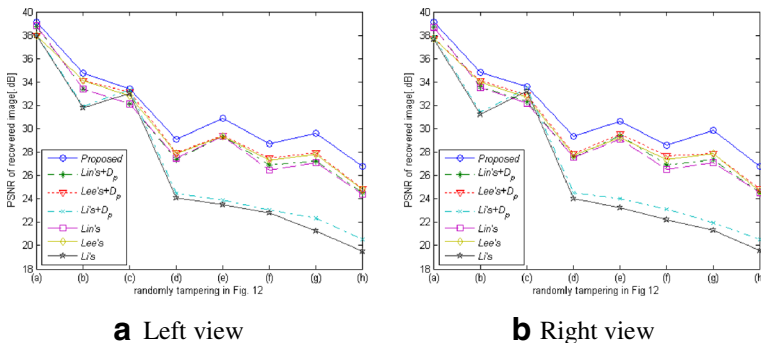

Fig. 12 PSNR comparisons of different schemes for 'Alt Moabit' cropping. **a** Left view, **b** Right view
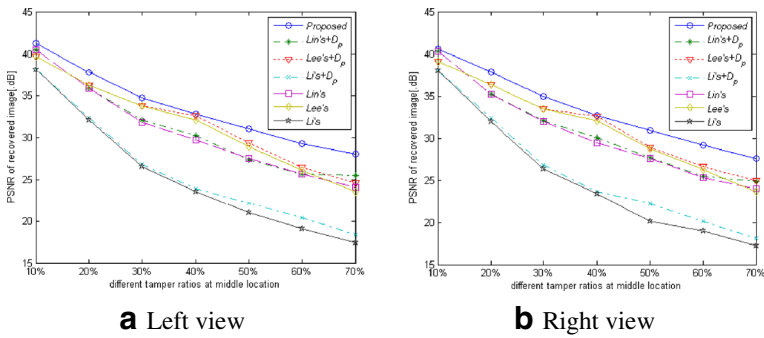
a Left view   **b** Right view

Fig. 13 PSNR comparisons of different schemes for 'Art' cropping. **a** Left view, **b** Right view

a Watermarked

b Lin's scheme (25.63dB)

c Lee's scheme (26.11dB)

d Li's scheme (19.11dB)

e Proposed scheme (29.06dB)

f Lin's scheme+D_p (25.69dB)

g Lee's scheme+D_p (26.42dB)

h Li's scheme+D_p (20.43dB)

a1 Local of (a)

b1 Local of (b)

c1 Local of (c)

d1 Local of (d)

e1 Local of (e)

f1 Local of (f)

g1 Local of (g)
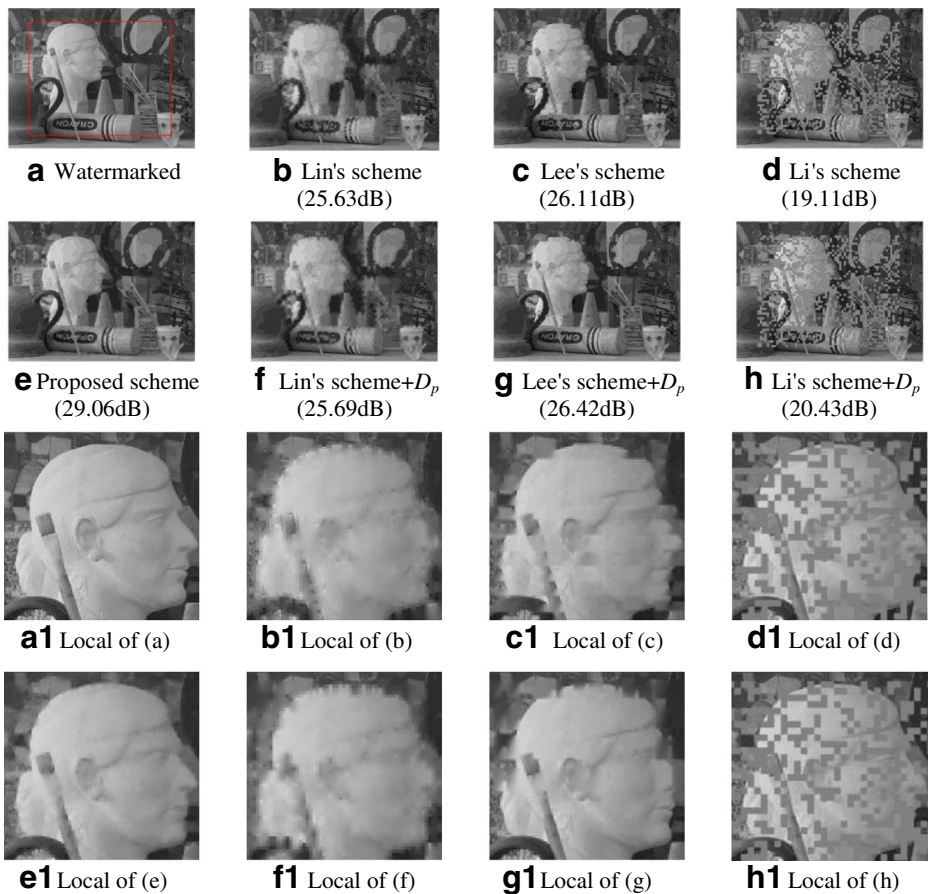
h1 Local of (h)

Fig. 14 Visual comparisons of recovery for 60 % of tamper in left view of 'Art'. **a** Watermarked, **b** Lin's scheme (25.63 dB), **c** Lee's scheme (26.11 dB), **d** Li's scheme (19.11 dB), **e** Proposed scheme (29.06 dB), **f** Lin's scheme + D_p (25.69 dB), **g** Lee's scheme + D_p (26.42 dB), **h** Li's scheme + D_p (20.43 dB). *a1* Local of (**a**), (*b1*) Local of (**b**), (*c1*) Local of (**c**), (*d1*) Local of (**d**), (*e1*) Local of (**e**), (*f1*) Local of (**f**), (*g1*) Local of (**g**), (*h1*) Local of (**h**)

tampered left view of 'Alt Moabit', where texture or background information is removed. PTRs of the seven schemes are all near to 99 %. The tampered images are all recovered without loss of main visual information, which are still recognized visually as illustrated in Fig. 11(*a1-h1*). PSNRs between recovered images and watermarked image for the seven schemes are illustrated in Fig. 12. Lin's scheme + D_p, Lee's scheme + D_p and Li's scheme + D_p perform better than their original schemes without D_p, which proves that relationship between left view and right view of a stereo image is able to improve recovery performance. However, PSNRs of the proposed scheme are the highest among the seven schemes. As the tamper size is increasing to 60 % or 70 %, the proposed scheme achieves better recover qualities with more than about 1.5 dB higher than those of other six schemes.

   In order to verify the performance of the proposed scheme on complex texture images, 'Art' with more complex texture is cropped from 10 % to 70 % at its middle locations. The tampered regions are detected with PRT of around 99 %, and PSNRs of the proposed scheme are more than 28 dB. Although other three schemes with D_p are superior to their original schemes as shown in Fig. 13, the corresponding PSNRs are still lower than that of the proposed scheme, and the greater the cropped ratio is, the bigger the gap is. Compared with experiments of 'Alt Moabit', the proposed scheme shows its superiority in recovery of images with more complex texture, because reference generation by using DWT with some detail information is better than average pixel used in Lin's scheme and Lee's scheme. Besides PSNR, the subjective visual perception of the recovered image obtained with the proposed scheme is also better than those of other six schemes. Fig. 14 shows recovered left views with respect to the seven schemes wherein 60 % of the original stereo image marked with a red rectangle in Fig. 14(a) is cropped. The recoveries of Lin's scheme and Lin's scheme + D_p are extremely blurred and is difficult to be recognized as shown in Fig. 14(*b1, f1*). The recoveries of Lee's scheme and Lee's scheme + D_p are better, but are still vague on textured information, such as eyes of the sculpture and pen as shown in Fig. 14(*c1, g1*). Lee's scheme + D_p performs better than Lee's scheme, for example some parts of pen are clearer, which is benefited from the use of disparity for recovery. Li's scheme is the worst, the image is totally destroyed as illustrated in Fig. 14(*d1*). Li's scheme + D_p just improves its original version a little. Fig. 14(*e1*) shows the recovery of the proposed scheme, wherein the eyes, nose, etc., are clearer. From the two cropping experiments with different tampering sizes, it is proved that the proposed scheme performs better for extensive tamper than tiny modification, and outperforms the other six schemes, even though the three schemes are improved with the use of D_p.

5 Conclusions

In this paper, an effective stereo image watermarking scheme with high self-recovery capability is presented. The mechanism of inter-view reference sharing decreases bits allocated for watermark embedding, which ensures quality of watermarked stereo images. Two reference copies of each block are embedded into the stereo image itself, so as to improve quality of tamper recovery. Moreover, the proposed scheme performs well on recovery of complex texture images due to references generated with DWT coefficients instead of traditional average pixel of a block. Besides detection of general tamper, it resists collage attack because authentication bits are not block independent and pairs of matched blocks are checked again whether they are still matched after being tampered. The experimental results demonstrate that the proposed scheme is able to recover stereo images with the tamper ratio from 10 % to 70 % imperceptibly, and PSNRs of recoveries are higher than those of extended Lin's, Lee's and Li's schemes especially at large

tamper ratio such as 60 % or more. The usage of high content correlation between the two views of stereo images improves performance of the proposed scheme well, and more characteristics of stereo images will be mined further in the future work, such as stereo vision.

References

1. Campisi P (2008) Object-oriented stereo-image digital watermarking. In: Journal of electronic imaging, vol. 17, SPIE, 043024
2. Celik M, Sharma G, Saber E, Tekalp A (2002) Hierarchical watermarking for secure image authentication with localization. In: IEEE transaction on image processing, vol. 11, IEEE, pp. 585–595
3. Fridrich J, Goljan M, Memon N (2002) Cryptanalysis of the Yeung-Mintzer fragile watermarking technique. In: Journal of electronic imaging, vol. 11, SPIE, pp. 262–274
4. Haouzia A, Noumeir R (2008) Methods for image authentication: a survey. In: Multimedia tools and applications, vol. 39, Springer, pp. 1–46
5. He H, Zhang J, Tai H (2011) A neighborhood-characteristic-based detection model for statistical fragile watermarking with localization. In: Multimedia tools and applications, vol. 52, Springer, pp 307–324
6. Huo Y, He H, Chen F (2012) Alterable-capacity fragile watermarking scheme with restoration capability. In: Optics communications, vol. 285, Elsevier, pp. 1759–1766
7. Hur N, Lee H, Lee G, Lee S, Gotchev A, Park S (2011) 3DTV broadcasting and distribution systems. In: IEEE transaction on broadcasting, vol. 57, IEEE, pp. 395–407
8. Hwang D, Bae K, Lee M, Kim S (2003) Real-time stereo image watermarking using discrete cosine transform and adaptive disparity maps. In: Proceedings of SPIE 5241, SPIE, pp. 233–242
9. Hwang D, Base K, Kim E (2004) Stereo image watermarking scheme based on discrete wavelet transform and adaptive disparity estimation. In: Proceedings of SPIE 5208, SPIE, pp. 196–205
10. Lee T, Lin S (2008) Dual watermark for image tamper detection and recovery. In: Pattern recognition, vol. 41, Elsevier, pp. 3497–3506
11. Li C, Wang Y, Ma B, Zhang Z (2011) A novel self-recovery fragile watermarking scheme based on dual–redundant-ring structure. In: Computer & electrical engineering, vol. 37, Elsevier, pp. 927–940
12. Lin P, Hsieh C, Huang P (2005) A hierarchical digital watermarking method for image tamper and recovery. In: Pattern recognition, vol. 38, Elsevier, pp. 2519–2529
13. Lin Y, Wu J (2011) A novel blind watermarking scheme for depth-image-based rendering 3D images. In: IEEE Transaction on Broadcasting, vol. 57, IEEE, pp. 602–611
14. Liu H, Yao X, Huang J (2010) Semi-fragile Zernike moment-base image watermarking for authentication. In: EURASIP Journal on advances in signal processing, vol. 7, Springer, pp. 1–17
15. Niu Y, Souidene W, Beghdadi A (2011) A visual sensitivity model based stereo image watermarking scheme. In: Third European workshop on visual information processing, IEEE, pp. 211–215
16. Qi X, Xin X (2011) A quantization-based semi-fragile watermarking scheme for image content authentication. In: Journal of visual communication and image representation, vol. 22, Elsevier, pp. 187–200
17. Sabater N, Almansa A, Morel J (2012) Meaningful matches in stereovision. In: IEEE transaction on pattern analysis and machine intelligence, vol. 34, IEEE, pp. 930–942
18. Shao F, Jiang G, Yu M, Chen K, Ho Y-S (2012) Asymmetric coding of multi-view video plus depth based 3-D video for view rendering. In: IEEE Transactions on Multimedia, vol. 14, IEEE, pp. 157–167
19. Tanimoto M, Tehrani M, Fujii T, Yendo T (2012) FTV for 3-D spatial communication. In: Proceedings of the IEEE, vol. 100, IEEE, pp. 905–917
20. Vleeschouwer C, Delaigle J, Macq B (2002) Invisibility and application functionalities in perceptual watermarking an overview. In: Proceedings of IEEE, vol. 90, IEEE, pp. 64–77
21. Zhang X, Wang S (2009) Fragile watermarking scheme using a hierarchical mechanism. In: Signal processing, vol. 89, Elsevier, pp. 675–679
22. Zhou Y, Jiang G, Yu M, Peng Z, Shao F (2012) Region-based error concealment of right-view frames for stereoscopic video transmission. In: Computers and electrical engineering, vol. 38, Elsevier, pp. 217–230

Ting Luo received his BSc degree in computer science from Ningbo University, Ningbo, China, in 2003, and his MSc degree in business information technology from Middlesex University, London, UK in 2004. He is currently working towards his PhD degree in Faculty of Information Science and Engineering Ningbo University and as a teacher in the College of Science and Technology Ningbo University as well. His research interests include multimedia security, image processing, data hiding and pattern recognition.

Gangyi Jiang received his M.S. degree from Hangzhou University, Hangzhou, China, in 1992, and received the Ph.D. degree from Ajou University, Korea, in 2000. From 2000, he joined the Faculty of Information Science and Engineering, Ningbo University, China, as a Professor. His research interests include digital video compression and communications, 3D image/video processing and visual perception.

Xiaodong Wang received his BSc degree in applied physical from Ningbo University, Ningbo, China, in 1992, and his MSc degree in software engineering from Zhejiang University, Hangzhou, China, in 2007. From 1992, he joined in Ningbo University, China. Since 2008, he works as an Associate Professor in Faculty of Information Science and Engineering. His research interests include network communications, 3D video processing and transmission system.

Mei Yu received the M.S. degree from Hangzhou Institute of Electronics Engineering, Hangzhou, China, in1993, and the Ph.D. form Ajou University, Korea, in 2000. Then, she joined in Ningbo University, Ningbo, China. Since 2005, she works as a Professor in Faculty of Information Science and Engineering, Ningbo University, China. Her research interests include image/video coding, stereoscopic image/video quality assessment and visual perception.

Feng Shao received his B.S. and Ph.D degrees from Zhejiang University, Hangzhou, China, in 2002 and 2007, respectively, all in Electronic Science and Technology. He is currently an associate professor in Faculty of Information Science and Engineering, Ningbo University, China. His research interests include video coding, image processing and perception, etc.

Zongju Peng received his B.S. degree in computer science from Sichuan University, Chengdu, China, in 1998, and Ph.D. degree from Institute of Computing Technology, Chinese Academy of Science in 2010. He is currently an associate professor in Faculty of Information Science and Engineering, Ningbo University, China. His research concentrates on image/video compression.