RESEARCH **Open Access**

CrossMark

A hybrid NSCT domain image watermarking scheme

Kai-Lung Hua[1], Bi-Ru Dai[1], Kathiravan Srinivasan[2*] (iD), You-Hsiang Hsu[1] and Vishal Sharma[3]

Abstract

At present, dealing with the piracy and tampering of images has become a notable challenge, due to the presence of smart mobile gadgets. In this paper, we propose a novel watermarking algorithm based on non-subsampled contourlet transform (NSCT) for improving the security aspects of such images. Moreover, the fusion of feature searching approach with watermarking methods has gained prominence in the current years. The scale-invariant feature transform (SIFT) is a technique in computer vision for detecting and illustrating the local features in images. Nevertheless, the SIFT algorithm can extract feature points with high invariance that are resilient to several issues like rotation, compression, and scaling. Furthermore, the extracted feature points are embedded with watermark using the NSCT approach. Subsequently, the tree split, voting, rotation searching, and morphology techniques are employed for improving the robustness against the noise. The proposed watermarking algorithm offers superior capability, better capture quality, and tampering resistance, when compared with existing watermarking approaches.

Keywords: Watermarking algorithm, SIFT, NSCT, Morphology, Mobile images

1 Introduction

In the current scenario, digital watermarking approaches have gained a lot of significance, due to the rapid progress in technology. Additionally, exchange and sharing of images have become easier and faster due to the advent of smart phones and gadgets. Furthermore, this technology has been considered as a vital approach for safeguarding the copyright and intellectual property of images from severe privacy and security issues. Nevertheless, when the image resolution gets higher and the volume turns bulkier, handling such images is always a challenging task. Subsequently, it becomes essential to compress the images; while compressing, a lot secretive information can probably get distorted. Hence, capturing the identifiable information from the image turns out to be a difficult task. Consequently, there is a need for the present watermarking approaches to be strong enough to deal with the compression issues. However, the digital watermarking approaches should also be able to achieve transparency, robustness, and better capacity to represent their unique identities. Data fusion can be also considered as an important aspect in image watermarking in video recognition systems [1, 2]. Deep learning can also enhance the image-based watermarking and can be used to formulate ranking algorithms for image recognition [3, 4]. Several watermarking techniques based on discrete cosine transform (DCT) [5–10] and discrete wavelet transform (DWT) [11–17] have been already established. Patra [7] established a watermarking scheme based on the Chinese remainder theorem (CRT), which was deployed in the DCT domain, and this approach was more effective against the JPEG compression attacks. Ababneh [12–14] established a compensated signature embedding (CSE) framework that could resist against attacks by the JPEG 2000 compression. Nevertheless, the abovementioned techniques were unable to solve the problems of rotation and scaling. Huynh-The and team [18] established a digital image watermarking scheme based on a coefficient quantization method that intuitively encodes the owner's data for each color channel to enhance imperceptibility and robustness of the concealed data. Wang and Colleagues [19] established a strong color image watermarking model based on local quaternion exponent moments that resisted the desynchronization attacks. Abdelhakim

*Correspondence: kathiravan@niu.edu.tw
[2]Department of Computer Science and Information Engineering, National Ilan University, Yilan County, Taiwan (ROC)
Full list of author information is available at the end of the article

Hua *et al. EURASIP Journal on Image and Video Processing* (2017) 2017:10
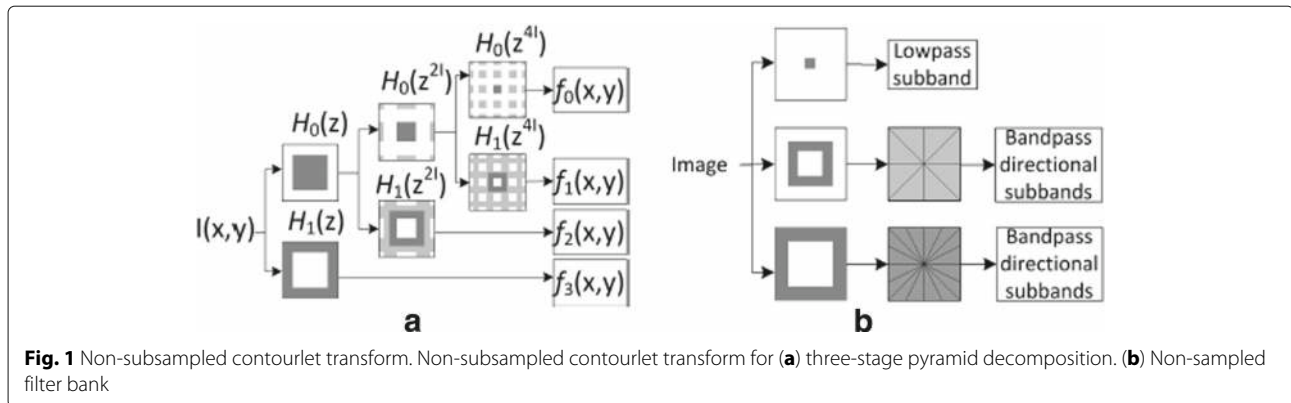
Page 2 of 17

and team [20] established a scheme, in which the embedding strength parameters for per-block image watermarking in the DCT domain are optimized. In this approach, the Bees algorithm was chosen as the optimization model and the fitness function was employed to exactly fit the optimization issues. Choi and Pun [21] presented a strong reversible watermarking model, in which the bit plane manipulation was deployed to conceal watermark bits in bit planes that are resistant to attacks. Moosazadeh and Andalib [22] established a digital image watermarking approach in YCbCr color space and DCT domain. Moreover, their scheme employed the coefficient exchange for embedding the watermark bits and also the genetic algorithm was used for choosing the target Y component coefficients of the host image. Kadu and team [23] presented a very proficient approach for copyright protection that was based on a modest and a competent embedding method for DWT-based video watermarking. They utilized this model in indoor video watermarking applications. Presently, the notion of feature searching has been widely employed in the digital watermarking approaches for enhancing its robust capacity. Feature searching is an important aspect of modern day sign-board reading systems. Detection of low-resolution images from weakly labeled street images can be done using efficient learning and recognition system as the one developed by Tsai et al. [24]. The scale invariant feature transform (SIFT) is one of the widely used feature searching techniques [25]. Furthermore, SIFT can determine some feature points even under distinct vicious distortions and the watermark is embedded in this feature region. Therefore, when the image gets distorted, it is definitely feasible to determine the feature regions with embedded information. Apart from selecting the feature region, identifying a suitable ambience for embodiment is also an essential task. Recently, the non-subsampled contourlet transform (NSCT) is an emerging approach that can be utilized for watermarking [26]. When compared with DCT, DWT, and alternate transforms, the NCST has superior capacity and offers a large amount of coefficients for the watermarking process. Li [27] established a scheme, in which they amalgamated the SIFT and NSCT approaches and they employed the notion of quantization to embed the watermark, thereby it has a greater capacity. Nevertheless, while deploying the NSCT approach to embed the information, the region surrounding the texture portion gets distorted by high-frequency information. Consequently, the resultant watermark turns out to be ruined due to the high-frequency noises, thus conceding the capture quality. In this work, a novel work amalgamating the SIFT and NSCT approaches with the tree split, voting, rotation, searching, and morphology, thereby providing a splendid and an efficacious model with high capture quality. In Section 2, the SIFT and NSCT approaches are illustrated.

Section 3 presents the in-depth details of the proposed method. The experimental results and conclusion are discussed in Sections 4 and 5, respectively. Overall, the main contributions of this paper can be summarized as follows:

- Design, implementation, and evaluation of a novel robust image watermarking scheme.
- The proposed watermarking scheme first perform a quadtree decomposition on lowpass subband of NSCT domain to avoid the relatively high-frequency texture. Second, the proposed method employ max-pooling technique to retrieve the fused watermark from each subregions to enhance the capture quality and tampering resistance. Third, a circulation procedure is proposed to offer rotation-tamper-proof ability. Finally, a morphology step is included to refine the extracted watermark.
- The proposed watermarking algorithm offers superior capability, better capture quality, and tampering resistance, when compared with existing watermarking approaches.

2 Materials and methods

Lowe established the SIFT method, in which the notion is to capture the feature points, not ruined by image processing, despite the fact that the image is under a distinct scale (either zooming or shrinking) [25]. As soon as the images are processed by means of a Gaussian function, the blurred version of the image shall be the best fit to characterize the scaling space. Primarily, to capture the feature points fruitfully, the difference of the Gaussian and the pyramid depiction has to be exploited for simulating the scaling space. Furthermore, in the region, extremes are utilized as feature candidate points, thereby computing the stability of these neighboring pixels. Subsequently, the pixels with low stability are discarded and finally the orientation of these feature points is determined. Moreover, every feature point offers the information about its coordinate, scale, and orientation, after the SIFT computation [25]. The non-subsampled contourlet transform includes two major steps: (1) non-subsampled pyramid (NSP) and (2) non-subsampled directional filter bank (NSDFB). This approach is analogous to Laplacian pyramid, with a subband decomposition of L stages as shown in Fig. 1a, and there is no requirement of downsampling. The NSP process results in the decomposition of a lowpass subband and L highpass sub-bands; further, NSDFB is applied to the highpass sub-bands, as shown in Fig. 1b. Moreover, only the lowpass sub-bands are computed in this work [26]. When compared with other transformed low frequencies, the NSCT lowpass sub-bands have several coefficients that can be employed for watermarking. Nevertheless, the lowpass sub-bands are usually blurred images which are processed by a filter. Furthermore, the

Hua *et al. EURASIP Journal on Image and Video Processing*   (2017) 2017:10

Page 3 of 17

Fig. 1 Non-subsampled contourlet transform. Non-subsampled contourlet transform for (**a**) three-stage pyramid decomposition. (**b**) Non-sampled filter bank

lowpass sub-bands contain smooth low-frequency information, in addition to a small number of high-frequency information. Hence, when the watermarking is performed in that region, the information gets effortlessly dispersed by high-frequency noises, during the process of computing and capturing it.
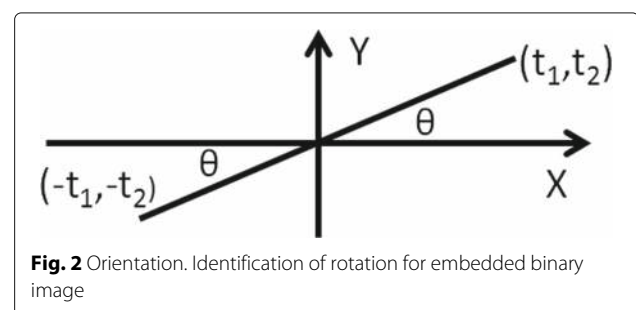
3   Proposed model

After obtaining the feature details from the SIFT process, the coordinate of the feature points is chosen to be the center, thereby creating a rectangular region for the watermark embedding. The measurement of every part of the region is assessed by scales. However, if the assessed region is very petite, so the capacity for the watermark embedding is not sufficient. On the other hand, if the assessed region is very large, the image will undergo severe tampering. Furthermore, the length and width of the feature region are fixed as 4s+1 (making use of the feature point as the center and 2s as the radius). Nevertheless, the SIFT approach computes several feature points, and not all feature points are appropriate for information embedding. Hence, the feature points need to be filtered. The following particulars are assumed to be the filtering constraints: (1) the feature points may overlap with the neighboring regions, if the value of S is very large; (2) the image undergoes severe tampering, if the value of S is very petite; (3) since the value of S varies with the resolution, a specific range of S is chosen to be the feature region, which results in missing feature points, during the image scaling process. Considering the entire scenario, primarily, the threshold D is defined (it varies with the image resolution)and every feature point is sorted based on its value. Besides, the values of *S* greater than the value of *D* are discarded and obtain *N* (the user can choose the value of *N*) feature regions with decreasing orders of *S*. The image resolution is fixed as $m \times n$ and *m* represents the longest portion and n indicates the shortest portion of the image. Based on our examination, there will be an overlap, when the side length of the radius is greater than $\frac{2m}{15}$. Therefore, the value of *D* is fixed to be $\frac{m}{15}$; and if

still there is an overlap, further, the feature region with a larger scale is chosen. After fixing the feature regions, then it is processed distinctly using NSCT approach. Moreover, the lowpass sub-band is embedded and its resolution is similar to the matching regions.

3.1   Embed phase

3.1.1   Orientation identify

The SIFT outcomes provide the orientation information of every feature point; however, based on Lowe Lowe [25], 15% of the feature points might have more than one orientation. Moreover, in this work, merely a single orientation information is required for positioning; therefore, the other orientations produced by SIFT approach are discarded. In order to obtain the orientation information of the feature regions, the formula (1) is employed for computing the gradient of feature regions in the *x* and *y* axis t_1, t_2, f_0 as the post-NSCT low pass sub-bands;$f_0(x, y)$ as the intensity value of the matching position; $f_0(\bar{x}, \bar{y})$ as the center of that region. Subsequently, the value of θ is obtained by deploying the formula (2). In the result of t_1, t_2, the angle is identical to $-t_1, -t_2$; however, the orientation is 180° opposite as in Fig. 2. While making a single orientation, (θ is on the first/fourth quadrant), when the result is t1<0, $\theta + \pi$ is the only orientation of ϕ that rotates the embedded binary image clockwise towards the only orientation, and this phenomenon is referred to as discrete rotation.

Fig. 2 Orientation. Identification of rotation for embedded binary image

Hua *et al. EURASIP Journal on Image and Video Processing* (2017) 2017:10

Page 4 of 17

$$t_1 = \sum_{(x,y)\in A} f_0(x,y)/(x-\bar{x})$$
$$t_2 = \sum_{(x,y)\in A} f_0(x,y)/(y-\bar{y}) \qquad (1)$$
$$A = \{(x,y)| \sqrt{(x-\bar{x})^2 + (y-\bar{y})^2} <= 2s\}$$

$$\theta = \arctan(t2/t1)$$
$$\phi = \begin{cases} \theta + \pi &, \text{if } t_1 < 0 \\ \theta &, \text{otherwise} \end{cases} \qquad (2)$$

3.1.2 Embedding the watermark

As stated earlier, in the NSCT approach, the captured image has blurred texture caused by the relatively high-frequency components. Therefore, the tree split approach is employed to safeguard the watermark from being embedded into the high-frequency region. However, voting is the technique to capture the feature points, for achieving better accuracy. Primarily, the edge detection is performed on the lowpass sub-band of the feature region for determining the location of the high-frequency details, by employing the canny edge detector [28]. If the region has no high-frequency details, then the region is not segregated. On the other hand, if the region contains the high-frequency information, then the region is segregated into 2×2 blocks and all blocks are completely analyzed. The analysis is performed by using the recursive method; it gets accomplished, when the segregated block size is lesser than half size of the minutest feature region and this process is referred to as "tree split." Once the segregation process gets over, the watermarked image is modified into the similar size as each block by means of bicubic interpolation. All bits are embedded into analogous coefficients of f_0. The block size is adjusted, to prevent the watermark from getting extensively distorted. Hence, the block size should be

identical to the half size of the smallest feature region. The quantization process is similar to the technique used by Li [27], due to fact that the coefficient of lowpass sub-band ranges between 0~255, and Δ is the quantization step. The coefficient can be computed using the formula $255/\Delta$ (255 is divided by Δ) to obtain the result. When the resultant coefficient is a odd number, it is substituted in sub-formula 1 and when the resultant coefficient is an even number, it is substituted in sub-formula 0. In formula (3), $f_0(x,y)$ represents the number of coefficients.

$$Q(x,y) = \begin{cases} 0, \text{ if } k\Delta \le f_0(x,y) < (k+1)\Delta \text{ for } k = 0, 2, \dots \\ 1, \text{ if } k\Delta \le f_0(x,y) < (k+1)\Delta \text{ for } k = 1, 3, \dots \end{cases}$$
$$(3)$$

The original coefficient is assigned to the corresponding number by using the embedded information. The coefficients are represented in the numbers as depicted in formula (3). Consequently, the coefficient has to be set in the center of the corresponding quantization. Hence, the quantization noise should be computed before finding the deviation in formula (4).

$$r(x,y) = f_0(x,y) - \lfloor f_0(x,y)/\Delta \rfloor * \Delta \qquad (4)$$

Furthermore, the outcome of formulas (3) and (4) is employed to compute formula (5). w_i indicates the bit of the watermark process. The ultimate result of the coefficient is given by formula (6). Figure 3 illustrates the embedded diagram.
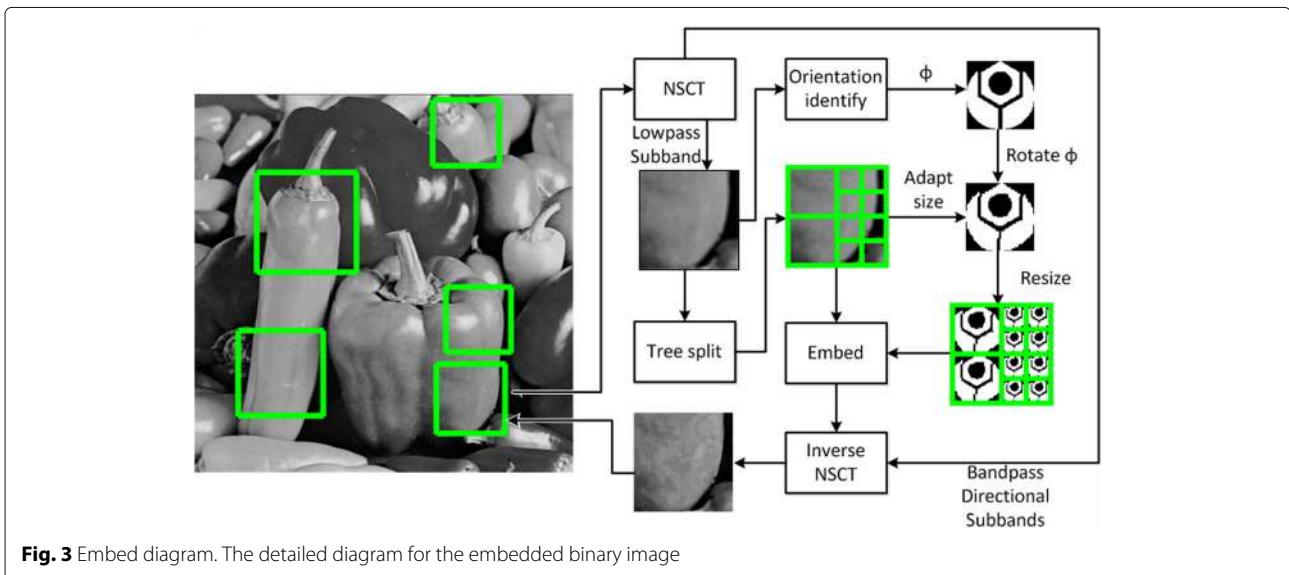
Fig. 3 Embed diagram. The detailed diagram for the embedded binary image

$$u(x,y) = \begin{cases} -r(x,y) + 0.5\Delta, & \text{if } Q(x,y) = w_i \\ -r(x,y) + 1.5\Delta, & \text{if } Q(x,y) \neq w_i, r(x,y) > 0.5\Delta \\ -r(x,y) - 0.5\Delta, & \text{if } Q(x,y) \neq w_i, r(x,y) \leq 0.5\Delta \end{cases}$$

$$\tag{5}$$

$$\grave{f}_0(x,y) = f_0(x,y) + u(x,y) \tag{6}$$
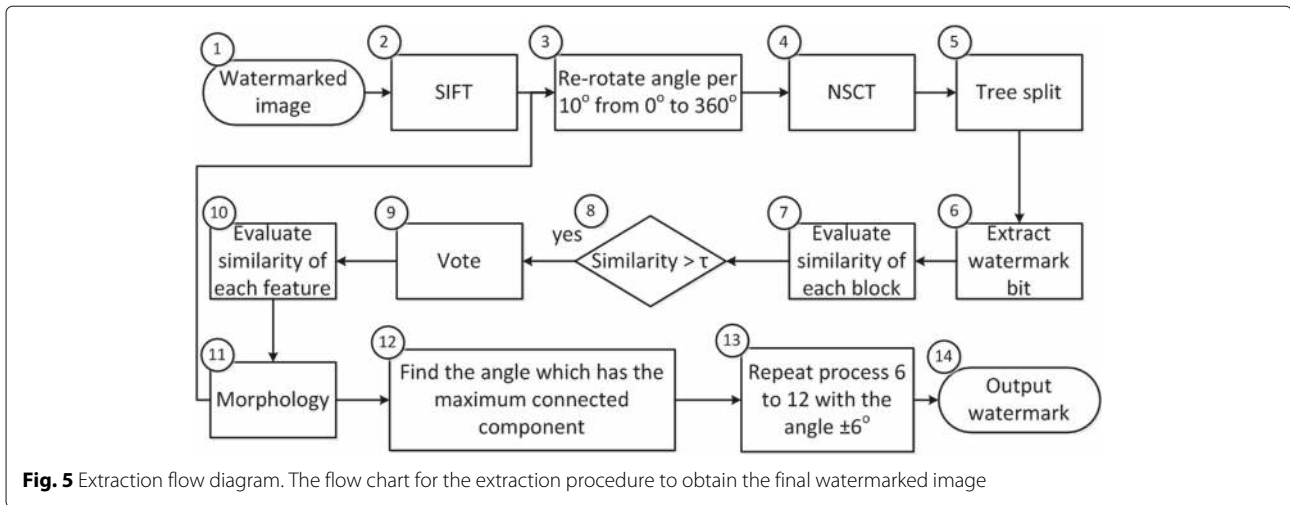
3.2 Extraction phase

Analogous to the embedding process, primarily, the orientation of the feature region is computed and then the tree-split approach is applied to the post-NSCT region. In the tree split process, assuming that if the region is segregated into n blocks, the formula (3) is applied to every block for obtaining the corresponding watermark bit information. Moreover, this process would result in the capturing of n watermark images. Besides, the segregated blocks containing no high-frequency information are utilized for voting the matching bit; and lastly, the block corresponding to the matching bit is found. The standard of voting determines the captured bit, if 0 is voted, so the matching bit would be 0; otherwise, it would be 1. If each block encompasses high-frequency information, so vote for bits is based on no high-frequency information with its matching position. Once the watermark image is obtained, it is scaled to four times larger in size and voting and scaling are repeated as aforementioned, until it reaches the top of the tree. The watermark is obtained, after accomplishing the abovementioned process. Further, the watermark is rotated anticlockwise, depending on the orientation of that region, which is termed as discrete rotation.

Nevertheless, when the image is tampered due to rotation, so the obtained square region shall not be the region assimilated during the embedding process and it will result in capturing distortion. Figure 4b depicts the variation in the embedded information, when the image is rotated by 50°. It is apparent from the figure that the captured image of each block is diverse from the segregated portions during the embedding process. In order to resolve this issue, the image rotation is used to ascertain the best angle. Each unit is assigned with a value of 10°. Figure 4c, d represents the captured position of each rotated angle. It can be noticed that after the rotation, the
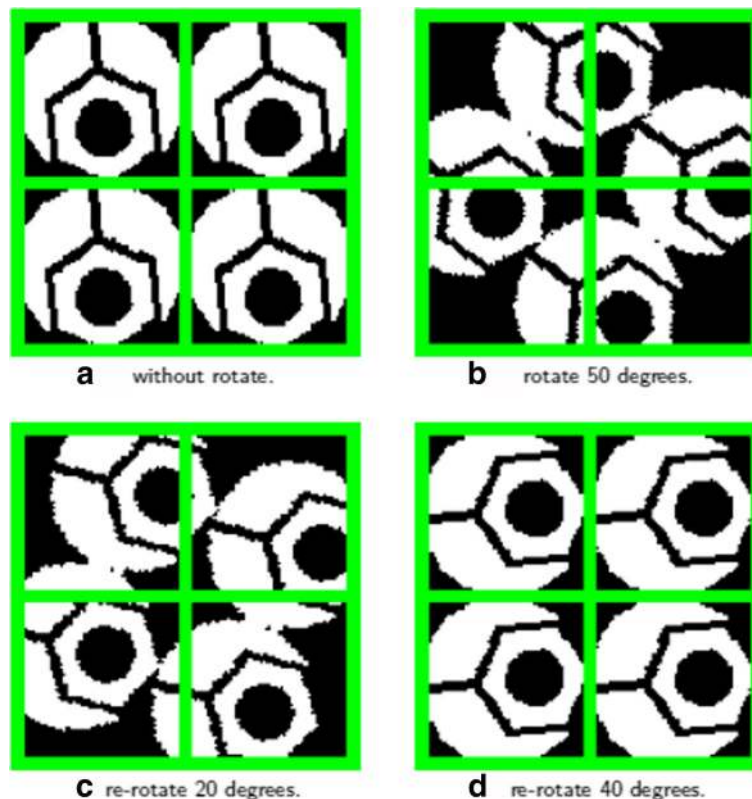
Fig. 4 The example of rotation search. **a** is the feature region without rotation and the location of embedded watermark. **b** is the captured feature region with 50° rotation, in comparison with the feature region without rotation. They do not match with each other. **c** is the captured feature region after being re-rotated at 20°, in comparison with the feature region without rotation. They do not match with each other. **d** is the captured feature region after being re-rotated at 40°, in comparison with the feature region without rotation. They have a great matching

Hua *et al. EURASIP Journal on Image and Video Processing* (2017) 2017:10

Page 6 of 17

Fig. 5 Extraction flow diagram. The flow chart for the extraction procedure to obtain the final watermarked image

capture region and the embedded region look dissimilar. Moreover, while comparing two captured blocks at a time, it is apparent that the consistency should be greater than τ for proceeding with the voting process, which is determined by a trial-and-error method on a training image.

When the voting process is accomplished, once again the watermark images are compared. Furthermore, if the angle is found to be perfect, then the result is collected from diverse regions. The extraction flow diagram presenting the procedure to obtain the final watermarked image is shown in Fig. 5. The error angle is segregated as shown in Fig. 6b, e. The black portion indicates the matching and white signifies the unmatched parts. Based on these characteristics, the morphology closing for discarding the segregated black portions that are matched. (Closing indicates the process of dilation (Eq. 7) which is done first and then process of erosion (Eq. 8) subsequently.) *A* represents the region for dilation,
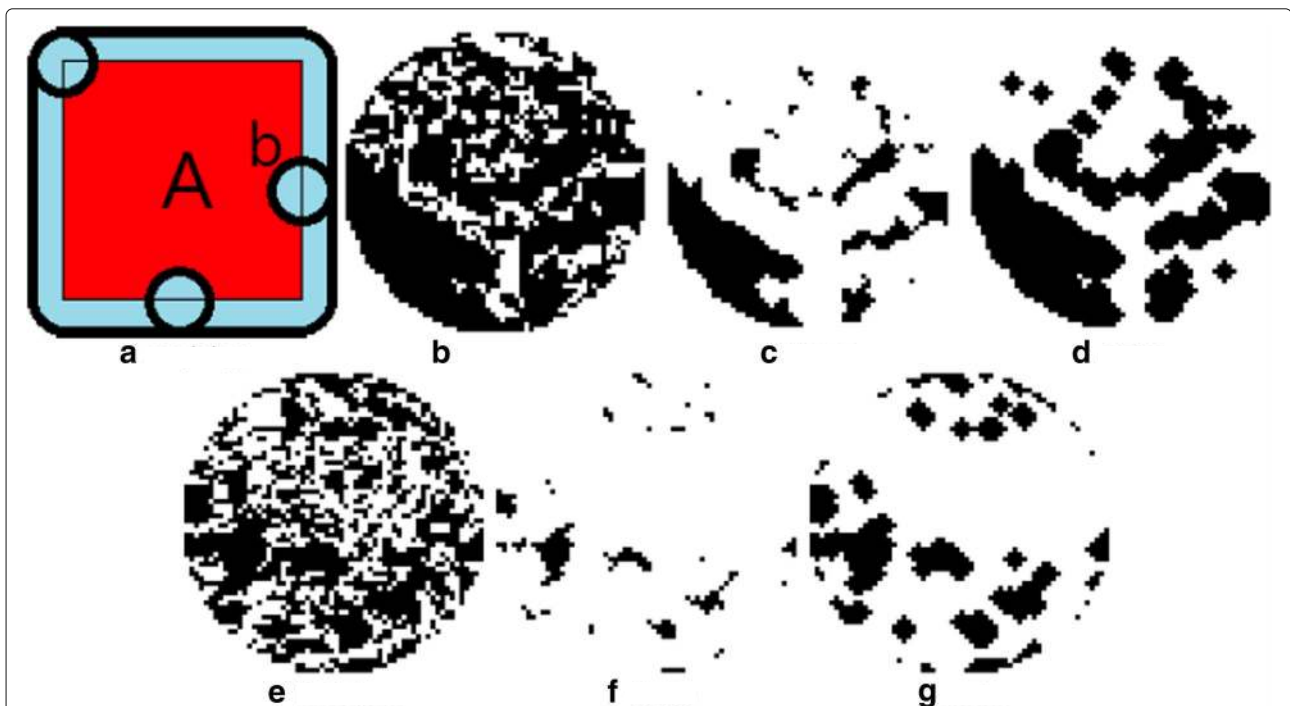
Fig. 6 The examples of using morphology to determine the maximum connected component. **a** is the morphology diagram. **b**, **c**, and **d** are not tampered by rotation and are re-rotated at 0°. They are the result after doing morphology closing on the perfect angle of watermark. **e**, **f**, and **g** are not tampered by rotation but are re-rotated at 10°. They are the result after doing morphology closing on the error angle of watermark. The *black part* is the connected component. **b** Exclusive-or. **c** dilation. **d** erosion. **e** Exclusive-or. **f** dilation. **g** erosion

Hua *et al. EURASIP Journal on Image and Video Processing* (2017) 2017:10

Page 7 of 17

Fig. 7 Nine test images and a watermark binary image. Test images images with a 512 × 512 resolution, and the watermark embedded as a 64 × 64 binary image. **a** Lenna. **b** aerial. **c** airfield. **d** lake. **e** Goldhill. **f** Barbara. **g** Baboon. **h** elaine. **i** peppers. **j** ntrust
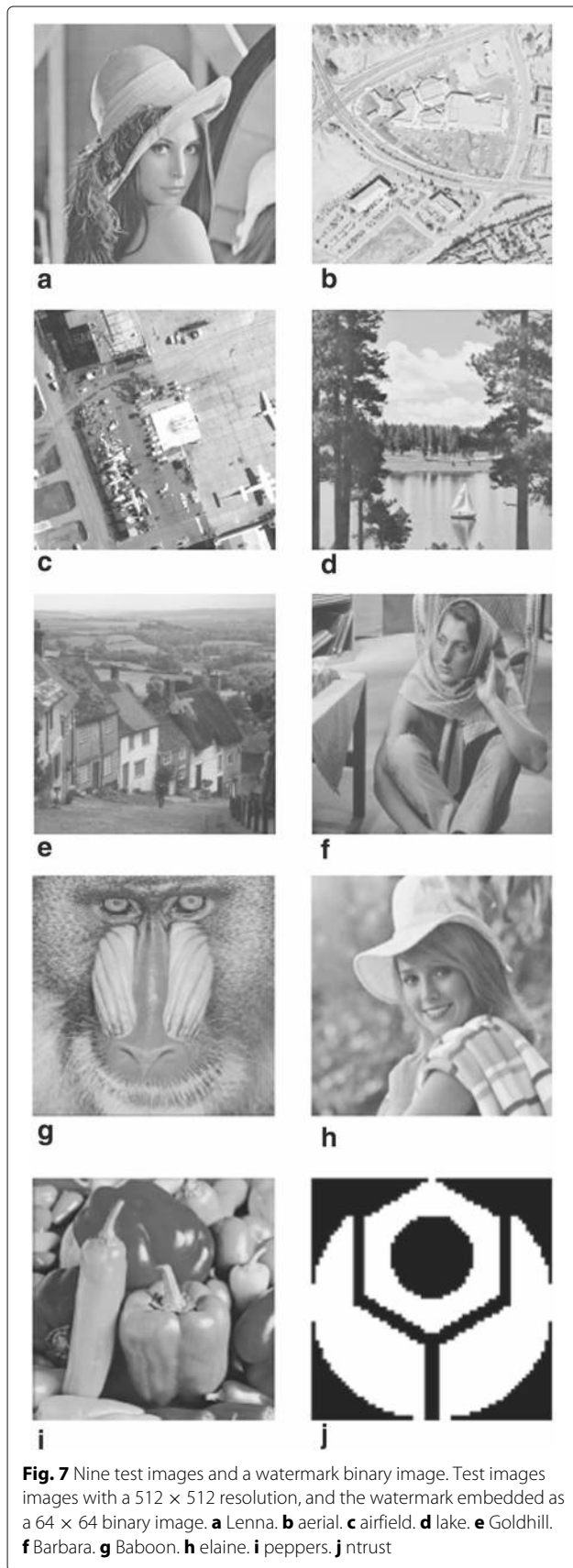
and *B* signifies the structuring elements. Dilation means that *B* is a circle on the pixels around *A* as continuum, to make *A* larger. Erosion means the region in which *A* deducts *B*, to make *B* smaller, as illustrated in Fig. 6. Figure 6c, d, f, g depicts the result of performing the closing operation on two different angles. The biggest black region is known as the connected component and the angle of biggest connected component is the accuracy angle. The region is searched based on the accuracy angle and the subsequent step is to search the accuracy angle with a deviation of ± 6° for obtaining the best rotated angle. If the deviation is under 1°, when compared with the confirmed angle, the region between ± 6° is fixed as the accuracy angle.

$$D(A, B) = \bigcup_{b \in B} A + b \tag{7}$$

$$E(A, B) = \bigcap_{b \in -B} A + b \tag{8}$$

4 Experimental results and discussion

In our experiments, nine images from the USC-SIPI database (http://sipi.usc.edu/database) are deployed (Fig. 7). These images have a 512 × 512 resolution, and the watermark embedded on it is a 64 × 64 binary image, as shown in Fig. 7j. The *N* feature regions for embedding are obtained, once the SIFT process gets over. In our experiments, the value of *N* is set as five, which is determined based on a trial-and-error method on a training image. Therefore, five feature regions are essential to proceed with the further process. In case of an overlap between these regions, then the process is continued until five disjoint feature regions are obtained. The value of *N* is fixed based on the requirements of the model. Moreover, if the value of *N* is larger, then more feature regions would be essential for information embedding, consequently leading to a superior chance of capturing watermarks with enhanced quality. Nevertheless, tampering effect on the original image would be highly intense, if there are more regions with embedded information. Additionally, computed peak signal-to-noise ratio (PSNR) dB values would be lower.

The proposed method is compared with Li [27] and Patra [7]. Moreover, based on the following three major criteria, the performance of the proposed algorithm can be found: (1) the perfection of watermark under diverse tampering conditions (i.e., the robustness); (2) the changeability of the information embedded in the image (PSNR); (3) the amount of capacity offered by an image. After performing the robust experiments, the normalized Hamming similarity (NHS) is computed. The NHS

Hua *et al. EURASIP Journal on Image and Video Processing* (2017) 2017:10

Page 8 of 17

formula is depicted in (9). $w(i)$ and $\overline{w}(i)$ correspondingly illustrate the original watermark and post-captured watermark; M represents the total amount of bits in the watermark image; $\oplus$ signifies the exclusive-or operation.

$$\text{NHS} = 1 - \frac{1}{M} \left\{ \sum_{i=0}^{M-1} [w(i) \oplus \overline{w}(i)] \right\} * 100 \qquad (9)$$

In order to tamper the image, the following values are set: JPEG quality factor $10 \sim 90\%$, rotation angle $10° \sim 90°$, scaling $\times 0.65$ to $\times 1.75$, median filter 3×3 mask $\sim 9 \times 9$ mask, shearing X and Y are $1 \sim 10\%$. Therefore, for capturing the watermark from the N feature region and also for accomplishing the NHS computation, select the largest NHS value to be W, as illustrated in (10).

$$W = \max(\text{NHS}_1, \text{NHS}_2, ..., \text{NHS}_N) \qquad (10)$$

For instance, considering Li's [27] approach, it can be observed that the lowpass sub-bands of NSCT are embedded directly. Hence, the captured watermark would produce blurred texture caused by the relatively high-frequency components, as a result of diverse texture details. In Fig. 8, the subplots 8a and b portrays the Lenna image's extracted watermark of Li's [27] approach and our proposed method, respectively. Furthermore, in our proposed method, the tree split algorithm is employed to find out the regions with high-frequency information and such regions are not embedded. Additionally, the voting mechanism is deployed, in order to preserve the perfect nature of the watermark, thereby resolving the high-frequency problem.

With the purpose of resisting the image tampering, such as rotation and deformation, the SIFT-produced orientations are not unique. The proposed method provides a unique orientation depending on their content, thereby it can be positioned perfectly. Moreover, the notion of rotation searching and morphology is employed to determine the exact capture angle. From the computations done so far, it is apparent that our method is superior in terms of capture quality than Li's [27] approach, and it can resist all forms of tampering. It can be noticed from Fig. 9 that the blue line specifies our W under diverse tampering; red line signifies the W of Li's [27] approach; black line depicts the W of Patra's [7] model. Furthermore, it can be witnessed that, for the tampering with diverse JPEG quality factor, the proposed method provides superior results than that of Li [27], and the
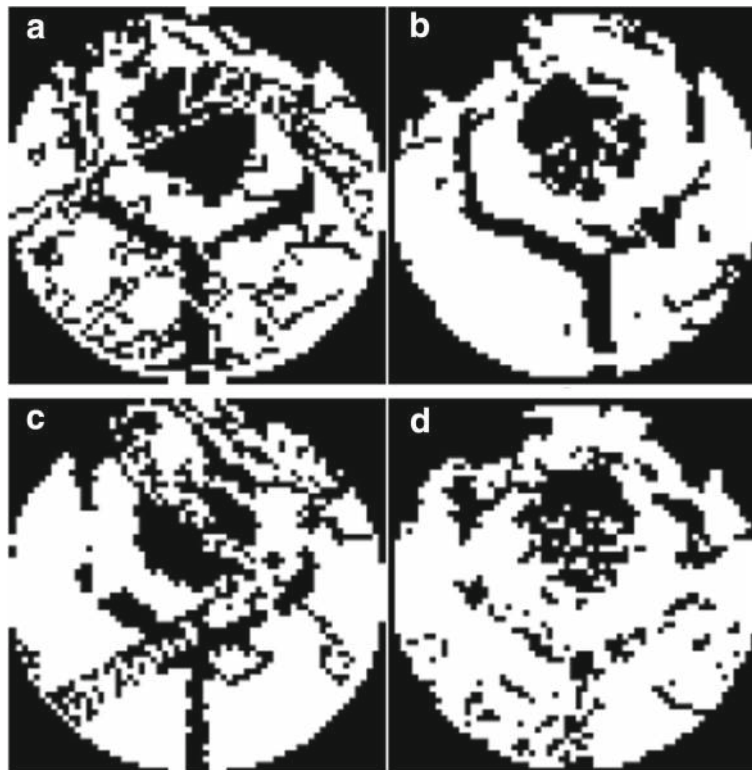
Fig. 8 The results of watermarking. The results of watermark extracted by **a** Li's [27] method (Lenna). **b** Proposed method (Lenna). **c** Li's [27] method (peppers). **d** Proposed method (peppers)

Hua *et al. EURASIP Journal on Image and Video Processing*   (2017) 2017:10

Page 9 of 17
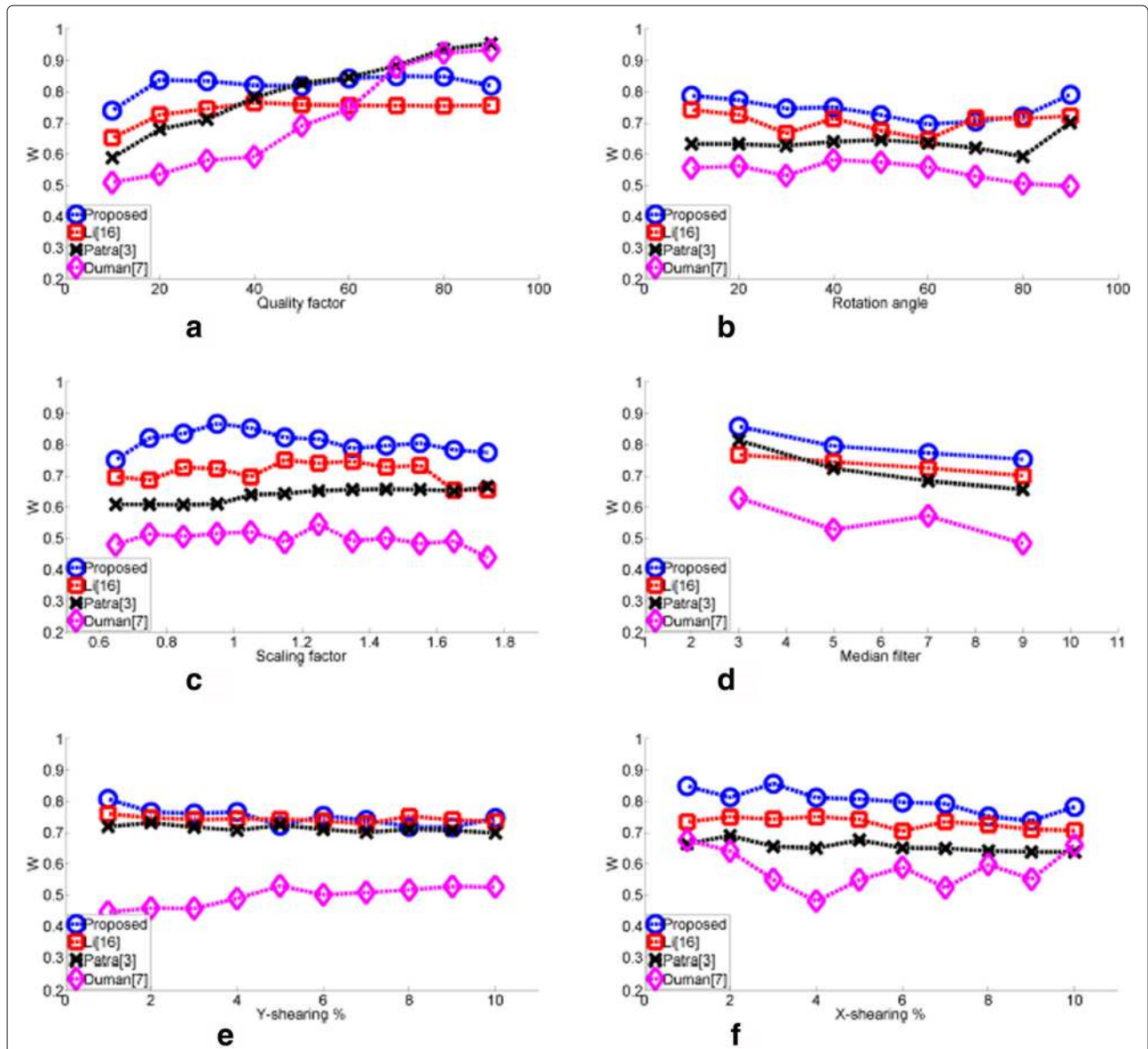
Fig. 9 The maximum NHS comparison among various attacks for Lenna image. Results for JPEG compression, rotation attacks, scaling attacks, median filter, and *Y* and *X* shearing percentages of the proposed approach in comparison with Li [27], Patra [7], and Duman [11]. **a** JPEG compressiion. **b** rotation attacks. **c** scaling attacks. **d** median filter. **e** shearing Y %. **f** shearing X %

maximum *W* is greater than 0.885. When compared with Patra's [7] model, predominantly, while processing the JPEG with quality factor greater than 50, the proposed method underperforms. On the other hand, when the quality factor is lesser than 50, the proposed method is comparatively superior than that of Patra's [7] model. Further, since the notion of SIFT feature searching is deployed in the proposed method, it is superior than that of Patra's [7], while resisting rotation and deformation.

Considering the rotation experiment, even though certain angles of the proposed method are lower than that of Li's [27] approach, the average angle *W* is fairly superior

to that of Li's [27] model. While taking into account about the scaling experiment, as the feature scale of Li's [27] model is chosen from the fixating regions, if the scaling factor is zoomed, it is dubious to capture feature regions obtained from productive embedding. In the median filter experiment, due to the watermark capture failure, *W* of the proposed method is superior to Li's [27] model for diverse masks. While comparing our proposed method for the Lenna experimental map with Patra's [7] model, it can be seen that the proposed method yields better results. Moreover, apart from comparing the value of *W*, the change of capacity and embedding of Patra [7] segregate the image into blocks and one block was made as a

Hua *et al. EURASIP Journal on Image and Video Processing* (2017) 2017:10

Page 10 of 17

unit for embedding the watermark. The size of the block was 8×8; therefore, Patra [7] was able to embed 4096 bits into a 512×512 image. However, more fixing and modifications of the image happen in Patra's [7] approach. In our proposed method, an amalgamation of the SIFT and NSCT approaches are utilized, in which most of the image details are unchanged and only a certain region of the image is modified. Furthermore, all coefficients in a region are permitted to be embedded. Additionally, as portrayed in Table 1, after the information embedding, the proposed method's watermarked image has almost 1.8 dB greater PSNR than Patra's [7], and as displayed in Table 2, the proposed method's capacity can reach 17,689 bits (in the maximum feature region with a size of 133×133). The detailed results of the other experiments are illustrated in Tables 3, 4, 5, 6, 7, and 8.

It can be clearly seen from the JPEG experiment in Table 3 that the value of W for the pepper image, with JPEG quality 40, is lower than Li's [27] approach. Besides, there is no disruption due to high-frequency noises, and the watermark captured by the proposed method is more pleasant for the human vision. Furthermore, as displayed in Fig. 8c, d these are the captured watermark images of JPEG quality 40 tampered peppers, obtained by employing Li's [27] and the proposed method, respectively. It is apparent from the rotation experiment in Table 4 that all the output images of the proposed method have better angle averages W than the other approaches. Nevertheless, it can be noticed from the scaling experiment in Table 5, when there is an progressive increase in the image scaling, the value of W computed by Li's [27] model, would drop down steadily. On the other hand, the proposed method is very robust to scaling. While considering the case of the image Elaine, even though the W average of the proposed method is mediocre under diverse scaling conditions, the W average of Li's [27] becomes slowly declining, when there is $\times 1.5$ scaling, and the deterioration is faster for larger scaling values. Henceforth, in general, the proposed method is comparatively superior than Li's [27] model for the scaling experiment.

Table 1 Image quality evaluation based on PSNR

Image	Proposed	Li [27]	Patra [7]	Duman [11]
Lenna	42.8048	41.0184	41.0102	43.2878
Aerial	48.3773	42.6563	39.7875	45.2514
Airfield	44.5894	42.0151	40.3889	44.3458
Lake	42.5213	41.5496	40.8763	45.2097
Goldhill	45.9514	41.8544	42.5871	44.9961
Barbara	42.9446	40.8478	41.3454	44.3548
Baboon	44.9735	43.5218	41.4154	44.3998
Elaine	45.9933	41.6803	41.0367	44.1249
Peppers	41.7276	40.4359	40.1475	44.2377

Table 2 Image capacity evaluation

Image	Proposed	Li [27]	Patra [7]	Duman [11]
Lenna	17,689	18,769	4096	1024
Aerial	6561	14,161	4096	1024
Airfield	18,769	19,881	4096	1024
Lake	18,769	17,689	4096	1024
Goldhill	15,625	16,641	4096	1024
Barbara	16,641	21,609	4096	1024
Baboon	17,689	17,689	4096	1024
Elaine	10,201	13,225	4096	1024
Peppers	18,769	19,321	4096	1024

It is apparent from the results of the median experiment shown in Table 6, although the W values are inferior for Goldhill, Barbara, and pepper images, but the results are extremely superior for the other six images. Moreover, the shearing results of the proposed method are identical to the other approaches. Generally, in case of non-deforming tampering processes like zipping and blurring, the SIFT approach offers better accuracy in capturing the feature points and the proposed model provides superior capture quality. Even though there are few losses in the certain feature points due tampering effects like deforming, the watermark can be captured from the other feature regions that are stable. Finally, in the case of majority of the image tampering processes, the proposed approach preserves a great robust capability.

5 Conclusions

Earlier, the concept of deformation was a very big issue for watermarking technologies. Nevertheless, this issue was addressed after the assimilation of feature searching into the watermarking models. Moreover, the notion of feature searching is inadequate to sustain deformation; enhancing the capacity of the watermarking scheme is the other fact that requires consideration. Consequently, NSCT approach offers greater capacity for further progression in the watermarking technologies. The experimental results have proven that the NSCT approach yields superior capacity, in comparison with other DCT-based watermarking schemes, even though the amalgamation of both SIFT and NSCT was explored by Li [27], which offers a high capacity and robust results. However, while considering the quality of the captured information, there are further challenges like blurred texture caused by the relatively high-frequency components. Therefore, in this work, the proposed method includes the concepts of tree split, voting, rotation searching, and morphology, and this resolves the issues caused by high-frequency noises in NSCT computing, thereby greatly improving the image quality.

Hua *et al. EURASIP Journal on Image and Video Processing*   (2017) 2017:10

Page 11 of 17

Table 3 Normalized Hamming similarity results for JPEG attack

Image	Scheme	JPEG 10	JPEG 20	JPEG 30	JPEG 40	JPEG 50	JPEG 60	JPEG 70	JPEG 80	JPEG 90
Lenna	Proposed	0.7400	0.8379	0.8340	0.8204	0.8185	0.8421	0.8492	0.8482	0.8194
	Li [27]	0.6528	0.7261	0.7455	0.7652	0.7594	0.7561	0.7555	0.7548	0.7565
	Patra [7]	0.5886	0.6792	0.7114	0.7813	0.8274	0.8455	0.8828	0.9360	0.9539
	Duman [11]	0.5098	0.5361	0.5811	0.5918	0.6904	0.7441	0.8779	0.9238	0.9336
Aerial	Proposed	0.6447	0.6663	0.6185	0.7006	0.7103	0.6944	0.7151	0.6844	0.7161
	Li [27]	0.5941	0.6812	0.6295	0.6615	0.6360	0.6437	0.6673	0.6263	0.6370
	Patra [7]	0.6003	0.7866	0.9050	0.8057	0.8420	0.9224	0.8799	0.9465	0.9763
	Duman [11]	0.4854	0.5713	0.6182	0.6230	0.7188	0.7012	0.7021	0.7285	0.7881
Airfield	Proposed	0.6389	0.7316	0.7623	0.7590	0.7413	0.7762	0.7603	0.7716	0.7442
	Li [27]	0.5791	0.6489	0.6693	0.6818	0.6605	0.6944	0.6647	0.6799	0.6793
	Patra [7]	0.5835	0.7019	0.7153	0.7937	0.8406	0.8577	0.8948	0.9297	0.9431
	Duman [11]	0.5518	0.5605	0.6543	0.6885	0.6797	0.7175	0.7324	0.7832	0.8779
Lake	Proposed	0.7545	0.7891	0.7839	0.8204	0.8282	0.7791	0.8353	0.8482	0.8330
	Li [27]	0.6615	0.6922	0.7384	0.7384	0.7542	0.7729	0.7261	0.7529	0.7700
	Patra [7]	0.6033	0.6746	0.6970	0.7742	0.8274	0.8479	0.8914	0.9299	0.9534
	Duman [11]	0.5049	0.5400	0.6104	0.6348	0.6787	0.7109	0.7930	0.8379	0.8896
Goldhill	Proposed	0.6460	0.7649	0.7519	0.7536	0.7781	0.7917	0.8026	0.8101	0.7910
	Li [27]	0.6269	0.6999	0.6970	0.6977	0.7087	0.6935	0.7090	0.6977	0.7019
	Patra [7]	0.6296	0.6929	0.6750	0.8022	0.8293	0.8540	0.8901	0.9351	0.9509
	Duman [11]	0.5635	0.5801	0.6465	0.6563	0.7178	0.7295	0.7754	0.8203	0.8760
Barbara	Proposed	0.6980	0.7313	0.7829	0.7855	0.7584	0.7051	0.7581	0.7972	0.7910
	Li [27]	0.6967	0.7322	0.7487	0.7490	0.7558	0.7578	0.7561	0.7652	0.7584
	Patra [7]	0.5591	0.6497	0.6841	0.7551	0.8174	0.8457	0.8926	0.9316	0.9543
	Duman [11]	0.4961	0.5195	0.5996	0.6289	0.6006	0.6621	0.7100	0.8223	0.9111
Baboon	Proposed	0.5995	0.7426	0.7468	0.7674	0.7539	0.7578	0.7716	0.7571	0.7820
	Li [27]	0.5778	0.6263	0.6492	0.6579	0.6495	0.6531	0.6512	0.6537	0.6460
	Patra [7]	0.6038	0.7222	0.7200	0.8159	0.8643	0.8840	0.9146	0.9463	0.9585
	Duman [11]	0.5830	0.6123	0.6533	0.6631	0.6514	0.6934	0.7402	0.7549	0.7432
Elaine	Proposed	0.5478	0.7235	0.7293	0.6767	0.7245	0.7261	0.7003	0.7151	0.7161
	Li [27]	0.6660	0.6851	0.7364	0.7329	0.7506	0.7590	0.7516	0.7578	0.7565
	Patra [7]	0.5957	0.6770	0.6711	0.7800	0.8440	0.8594	0.9041	0.9414	0.9580
	Duman [11]	0.4922	0.5283	0.5996	0.6816	0.7676	0.7920	0.8701	0.9189	0.9424
Peppers	Proposed	0.7016	0.7464	0.7435	0.7800	0.7710	0.7972	0.7943	0.7978	0.8017
	Li [27]	0.6996	0.7768	0.7548	0.7875	0.7623	0.7891	0.7920	0.7920	0.8010
	Patra [7]	0.5886	0.6792	0.7114	0.7813	0.8274	0.8455	0.8828	0.9360	0.9539
	Duman [11]	0.4541	0.4932	0.5566	0.5791	0.6914	0.7432	0.8027	0.8896	0.9453

Hua *et al. EURASIP Journal on Image and Video Processing* (2017) 2017:10

Page 12 of 17

Table 4 Normalized Hamming similarity results for rotation attack

Image	Scheme	10°	20°	30°	0°	50°	60°	70°	80°	90°
Lenna	Proposed	0.7875	0.7733	0.7464	0.7503	0.7258	0.6954	0.7051	0.7213	0.7910
	Li [27]	0.7422	0.7261	0.6663	0.7145	0.6757	0.6466	0.7158	0.7132	0.7225
	Patra [7]	0.6328	0.6331	0.6270	0.6399	0.6458	0.6360	0.6199	0.5925	0.7012
	Duman [11]	0.5557	0.5625	0.5313	0.5820	0.5742	0.5596	0.5293	0.5059	0.4980
Aerial	Proposed	0.6776	0.6218	0.7083	0.7158	0.6240	0.6970	0.6305	0.7245	0.7090
	Li [27]	0.6205	0.6076	0.6366	0.6308	0.5969	0.6189	0.6350	0.6373	0.6205
	Patra [7]	0.6287	0.6460	0.6682	0.6763	0.6812	0.6736	0.6438	0.6108	0.6980
	Duman [11]	0.5332	0.5635	0.5752	0.5635	0.5908	0.5576	0.5283	0.5176	0.4951
Airfield	Proposed	0.7703	0.6037	0.7474	0.7891	0.7258	0.5975	0.7290	0.7758	0.7235
	Li [27]	0.6683	0.6424	0.6360	0.6240	0.6324	0.6689	0.7238	0.7028	0.6405
	Patra [7]	0.6328	0.6443	0.6677	0.6692	0.6675	0.6545	0.6484	0.6013	0.6965
	Duman [11]	0.5137	0.5410	0.5410	0.5615	0.5586	0.5918	0.5381	0.5127	0.4883
Lake	Proposed	0.7345	0.7287	0.6818	0.6815	0.6818	0.6460	0.7187	0.7271	0.7678
	Li [27]	0.7361	0.7106	0.7261	0.7277	0.6967	0.7070	0.6912	0.7006	0.6970
	Patra [7]	0.6375	0.6641	0.6782	0.6733	0.6626	0.6575	0.6401	0.6138	0.7014
	Duman [11]	0.5088	0.5488	0.5820	0.5439	0.5684	0.5469	0.5293	0.5020	0.4639
Goldhill	Proposed	0.6460	0.7649	0.7519	0.7536	0.7781	0.7917	0.8026	0.8101	0.7910
	Li [27]	0.6809	0.6757	0.6764	0.6731	0.7087	0.6802	0.7109	0.7154	0.6783
	Patra [7]	0.6257	0.6340	0.6438	0.6484	0.6558	0.6387	0.6245	0.5850	0.7017
	Duman [11]	0.5635	0.5801	0.6465	0.6563	0.7178	0.7295	0.7754	0.8203	0.8760
Barbara	Proposed	0.7839	0.7513	0.7897	0.7939	0.7494	0.7235	0.7093	0.6786	0.7342
	Li [27]	0.7332	0.7216	0.7216	0.7006	0.7099	0.7432	0.7380	0.7222	0.7177
	Patra [7]	0.6211	0.6279	0.6462	0.6482	0.6470	0.6436	0.6206	0.5940	0.7014
	Duman [11]	0.5273	0.5664	0.5635	0.5625	0.5684	0.5596	0.5244	0.5342	0.5049
Baboon	Proposed	0.7497	0.7164	0.7422	0.6948	0.7345	0.7048	0.7303	0.7429	0.6818
	Li [27]	0.6550	0.7090	0.6305	0.6463	0.6428	0.6599	0.6344	0.6353	0.6483
	Patra [7]	0.5984	0.6316	0.6370	0.6431	0.6436	0.6287	0.6152	0.5879	0.7014
	Duman [11]	0.5195	0.5498	0.5488	0.5576	0.5947	0.5811	0.5684	0.5156	0.5078
Elaine	Proposed	0.6909	0.6873	0.6974	0.6043	0.6970	0.6983	0.6308	0.7196	0.6954
	Li [27]	0.7380	0.7429	0.7300	0.7397	0.7251	0.7319	0.7290	0.7364	0.7300
	Patra [7]	0.6272	0.6313	0.6567	0.6655	0.6621	0.6533	0.6428	0.6069	0.7017
	Duman [11]	0.5439	0.5537	0.5400	0.5488	0.5889	0.5557	0.5508	0.5488	0.4883
Peppers	Proposed	0.8714	0.6944	0.6893	0.7410	0.7442	0.7300	0.6402	0.8227	0.8411
	Li [27]	0.7248	0.7248	0.7222	0.7326	0.7410	0.7074	0.7645	0.7232	0.7468
	Patra [7]	0.6321	0.6460	0.6509	0.6448	0.6487	0.6536	0.6416	0.6130	0.6995
	Duman [11]	0.5078	0.5293	0.5850	0.5410	0.5859	0.5811	0.5430	0.5186	0.4775

Hua *et al. EURASIP Journal on Image and Video Processing*   (2017) 2017:10

Page 13 of 17

Table 5 Normalized Hamming similarity results for scaling attack

Image	Scheme	×0.65	×0.75	×0.85	×0.95	×1.05	×1.15	×1.25	×1.35	×1.45	×1.55	×1.65	×1.75
Lenna	Proposed	0.7506	0.8214	0.8353	0.8672	0.8530	0.8230	0.8175	0.7881	0.7962	0.8046	0.7833	0.7752
	Li [27]	0.6970	0.6864	0.7264	0.7232	0.6970	0.7510	0.7406	0.7468	0.7280	0.7339	0.6541	0.6557
	Patra [7]	0.6089	0.6089	0.6072	0.6104	0.6399	0.6433	0.6523	0.6563	0.6572	0.6570	0.6519	0.6653
	Duman [11]	0.4795	0.5137	0.5059	0.5156	0.5205	0.4883	0.5449	0.4912	0.5010	0.4834	0.4912	0.4404
Aerial	Proposed	0.6967	0.7135	0.6860	0.6621	0.6877	0.7216	0.6902	0.6673	0.6663	0.6828	0.6657	0.6990
	Li [27]	0.6059	0.6567	0.6890	0.6302	0.6557	0.6405	0.6525	0.6066	0.5940	0.6289	0.6043	0.5992
	Patra [7]	0.5828	0.5820	0.5674	0.6028	0.5991	0.6008	0.6033	0.5959	0.5896	0.5891	0.5889	0.5928
	Duman [11]	0.4932	0.5137	0.4912	0.4922	0.4766	0.4912	0.4727	0.4902	0.4883	0.4863	0.4883	0.4932
Airfield	Proposed	0.6776	0.6970	0.6370	0.7661	0.7309	0.7829	0.7862	0.7116	0.6815	0.8033	0.6586	0.7539
	Li [27]	0.6043	0.6286	0.6470	0.6609	0.7390	0.7129	0.6663	0.6873	0.6512	0.6583	0.6537	0.6531
	Patra [7]	0.5872	0.6001	0.5996	0.6064	0.6084	0.6072	0.6055	0.6052	0.5977	0.6025	0.6035	0.6067
	Duman [11]	0.5254	0.5039	0.5068	0.4912	0.5010	0.4834	0.5127	0.4971	0.5313	0.5234	0.4980	0.5029
Lake	Proposed	0.7200	0.7794	0.7694	0.8068	0.8524	0.8359	0.8049	0.7991	0.8230	0.8072	0.7807	0.8395
	Li [27]	0.7158	0.7248	0.6928	0.7487	0.7468	0.6731	0.7222	0.6893	0.6906	0.6289	0.6124	0.6079
	Patra [7]	0.6094	0.6287	0.6218	0.6211	0.6338	0.6353	0.6523	0.6492	0.6257	0.6416	0.6475	0.6433
	Duman [11]	0.5273	0.5117	0.5039	0.5078	0.4854	0.5020	0.5068	0.5146	0.5000	0.5205	0.4951	0.5166
Goldhill	Proposed	0.6712	0.7271	0.7387	0.7846	0.7862	0.7481	0.7697	0.7784	0.6851	0.6941	0.7122	0.7196
	Li [27]	0.6660	0.6702	0.7090	0.7190	0.7154	0.6993	0.6886	0.6848	0.6796	0.6660	0.6457	0.5966
	Patra [7]	0.5667	0.5891	0.5991	0.5876	0.6177	0.6162	0.6252	0.6179	0.6116	0.6172	0.6162	0.6165
	Duman [11]	0.5137	0.5146	0.5225	0.5352	0.4902	0.4805	0.5127	0.5186	0.4746	0.4971	0.5020	0.5166
Barbara	Proposed	0.7099	0.7222	0.8088	0.7332	0.8288	0.8023	0.7536	0.7742	0.7661	0.7629	0.7358	0.7981
	Li [27]	0.7232	0.7106	0.7251	0.7355	0.7264	0.7339	0.7322	0.7274	0.7019	0.6838	0.6618	0.6831
	Patra [7]	0.5935	0.5806	0.5959	0.6030	0.6125	0.6206	0.6350	0.6206	0.6157	0.6233	0.6326	0.6294
	Duman [11]	0.4727	0.4922	0.5146	0.4902	0.4863	0.4854	0.4971	0.5127	0.4756	0.4844	0.5088	0.4941
Baboon	Proposed	0.7074	0.7464	0.7784	0.6634	0.7865	0.7742	0.7745	0.7952	0.7639	0.7494	0.7500	0.7458
	Li [27]	0.5817	0.6224	0.6008	0.6495	0.6253	0.5940	0.6001	0.6376	0.6147	0.5904	0.5662	0.5833
	Patra [7]	0.5444	0.5405	0.5627	0.5676	0.5808	0.5669	0.5803	0.5757	0.5623	0.5769	0.5596	0.5754
	Duman [11]	0.5098	0.4990	0.5127	0.5068	0.5098	0.4922	0.5449	0.4756	0.4912	0.5137	0.4912	0.4717
Elaine	Proposed	0.6793	0.7051	0.7484	0.6738	0.6838	0.6437	0.7691	0.7064	0.7006	0.6928	0.7448	0.7019
	Li [27]	0.6654	0.6812	0.7332	0.7523	0.7477	0.7345	0.7455	0.7645	0.7474	0.7513	0.7506	0.5921
	Patra [7]	0.5903	0.5891	0.6045	0.6074	0.6208	0.6211	0.6289	0.6331	0.6216	0.6362	0.6321	0.6338
	Duman [11]	0.4775	0.4824	0.4893	0.5078	0.4834	0.5244	0.56225	0.5186	0.4961	0.4824	0.4854	0.5000
Peppers	Proposed	0.8224	0.7926	0.8630	0.8876	0.7145	0.8472	0.8708	0.8669	0.8456	0.8718	0.8792	0.7455
	Li [27]	0.6809	0.7054	0.7119	0.7720	0.7500	0.7461	0.7167	0.7148	0.7161	0.7174	0.7135	0.7099
	Patra [7]	0.6101	0.6011	0.6030	0.6260	0.6270	0.6169	0.6365	0.6372	0.6255	0.6296	0.6189	0.6367
	Duman [11]	0.4961	0.4756	0.5029	0.4932	0.5225	0.5391	0.5332	0.5254	0.4961	0.5205	0.5059	0.5391

Hua *et al. EURASIP Journal on Image and Video Processing* (2017) 2017:10

Page 14 of 17

Table 6 Normalized Hamming similarity results for median filter attack

Image	Scheme	Mask 3 × 3	Mask 5 × 5	Mask 7 × 7	Mask 9 × 9
Lenna	Proposed	0.8579	0.7959	0.7729	0.7529
	Li [27]	0.7681	0.7448	0.7245	0.6999
	Patra [7]	0.8137	0.7236	0.6836	0.6563
	Duman [11]	0.6299	0.5283	0.5723	0.4834
Aerial	Proposed	0.7332	0.7196	0.7122	0.6696
	Li [27]	0.6560	0.6105	0.5950	0.6244
	Patra [7]	0.7339	0.6494	0.6104	0.5967
	Duman [11]	0.6084	0.5557	0.5146	0.5000
Airfield	Proposed	0.7141	0.7393	0.6986	0.6395
	Li [27]	0.6725	0.6418	0.6334	0.5982
	Patra [7]	0.7451	0.6702	0.6323	0.6240
	Duman [11]	0.5322	0.5039	0.5088	0.4883
Lake	Proposed	0.8269	0.8440	0.7620	0.7258
	Li [27]	0.7607	0.7487	0.7048	0.6977
	Patra [7]	0.7651	0.6763	0.6460	0.6296
	Duman [11]	0.4883	0.4590	0.5195	0.5449
Goldhill	Proposed	0.7817	0.7455	0.6983	0.6531
	Li [27]	0.7041	0.6838	0.6977	0.6983
	Patra [7]	0.7815	0.6934	0.6565	0.6313
	Duman [11]	0.5967	0.5449	0.5518	0.5010
Barbara	Proposed	0.7813	0.7710	0.7329	0.6860
	Li [27]	0.7552	0.7455	0.7284	0.7109
	Patra [7]	0.7695	0.6875	0.6477	0.6284
	Duman [11]	0.6152	0.5205	0.5068	0.5254
Baboon	Proposed	0.7590	0.7758	0.7116	0.6130
	Li [27]	0.6457	0.6250	0.6118	0.6098
	Patra [7]	0.6765	0.6094	0.5894	0.5691
	Duman [11]	0.5234	0.5215	0.5205	0.5469
Elaine	Proposed	0.6718	0.6570	0.6557	0.7209
	Li [27]	0.7464	0.7016	0.6815	0.6463
	Patra [7]	0.7917	0.7200	0.6736	0.6406
	Duman [11]	0.5029	0.5000	0.5547	0.5039
Peppers	Proposed	0.8023	0.8207	0.8401	0.8130
	Li [27]	0.7904	0.7859	0.7642	0.7452
	Patra [7]	0.8025	0.7141	0.6726	0.6416
	Duman [11]	0.6143	0.4219	0.4697	0.4951

Hua *et al. EURASIP Journal on Image and Video Processing*   (2017) 2017:10

Page 15 of 17

Table 7 Normalized Hamming similarity results for *Y* shearing % attack

Image	Scheme	1%	2%	3%	4%	5%	6%	7%	%	9%	10%
Lenna	Proposed	0.8075	0.7658	0.7613	0.7661	0.7225	0.7539	0.7406	0.7187	0.7171	0.7464
	Li [27]	0.7603	0.7477	0.7426	0.7435	0.7429	0.7380	0.7284	0.7523	0.7416	0.7358
	Patra [7]	0.7197	0.7314	0.7173	0.7073	0.7244	0.7104	0.7007	0.7109	0.7073	0.6985
	Duman [11]	0.4453	0.4580	0.4570	0.4893	0.5293	0.5010	0.5078	0.5166	0.5273	0.5264
Aerial	Proposed	0.6441	0.5940	0.6612	0.6660	0.6525	0.6502	0.6521	0.6589	0.6034	0.6896
	Li [27]	0.6696	0.6715	0.6738	0.6760	0.6841	0.6818	0.6515	0.6618	0.6715	0.6657
	Patra [7]	0.6382	0.6399	0.6477	0.6335	0.6511	0.6494	0.6445	0.6543	0.6523	0.6482
	Duman [11]	0.4609	0.5049	0.4932	0.4697	0.5234	0.5430	0.4912	0.5029	0.5186	0.5313
Airfield	Proposed	0.6751	0.6967	0.7403	0.6037	0.7093	0.6970	0.6143	0.6082	0.7329	0.6899
	Li [27]	0.6718	0.6873	0.6899	0.6541	0.6699	0.6596	0.6709	0.6667	0.6651	0.6686
	Patra [7]	0.6416	0.6658	0.6516	0.6511	0.6614	0.6516	0.6663	0.6714	0.6558	0.6611
	Duman [11]	0.4932	0.5039	0.4775	0.5225	0.5420	0.5000	0.5195	0.5010	0.5391	0.5059
Lake	Proposed	0.7620	0.7322	0.7668	0.7405	0.7087	0.7374	0.6915	0.6864	0.5849	0.6932
	Li [27]	0.7432	0.7519	0.7390	0.7322	0.7358	0.7319	0.7374	0.7229	0.7342	0.7264
	Patra [7]	0.6853	0.6929	0.6841	0.6804	0.6882	0.6726	0.6704	0.6853	0.6777	0.6780
	Duman [11]	0.5020	0.4639	0.4756	0.5176	0.4766	0.4990	0.5088	0.5205	0.5000	0.5186
Goldhill	Proposed	0.7251	0.6754	0.6944	0.6537	0.6899	0.6518	0.6495	0.6818	0.6802	0.6143
	Li [27]	0.6767	0.7099	0.7022	0.6893	0.6822	0.7035	0.7057	0.7016	0.6660	0.6983
	Patra [7]	0.6448	0.6580	0.6428	0.6467	0.6328	0.6406	0.6372	0.6587	0.6340	0.6445
	Duman [11]	0.4668	0.4814	0.4531	0.4941	0.5303	0.5156	0.5156	0.5146	0.5059	0.5156
Barbara	Proposed	0.7930	0.8065	0.7852	0.7297	0.7368	0.6877	0.7600	0.7164	0.7064	0.7339
	Li [27]	0.7671	0.7626	0.7655	0.7545	0.7523	0.7513	0.7552	0.7445	0.7419	0.7397
	Patra [7]	0.6748	0.6777	0.6709	0.6750	0.6772	0.6750	0.6738	0.6650	0.6685	0.6614
	Duman [11]	0.4824	0.4795	0.4688	0.4746	0.4980	0.5049	0.5127	0.5332	0.4922	0.5400
Baboon	Proposed	0.7484	0.7500	0.7464	0.7158	0.6909	0.7442	0.6738	0.7374	0.7087	0.6024
	Li [27]	0.6392	0.6450	0.6499	0.6499	0.6463	0.6525	0.6486	0.6508	0.6489	0.6528
	Patra [7]	0.6274	0.6482	0.6406	0.6372	0.6331	0.6497	0.6252	0.6587	0.6338	0.6431
	Duman [11]	0.4961	0.4609	0.4883	0.5010	0.5166	0.4902	0.5000	0.5098	0.5088	0.5313
Elaine	Proposed	0.7348	0.7235	0.7203	0.6776	0.7048	0.7332	0.6586	0.6783	0.6576	0.6764
	Li [27]	0.7506	0.7552	0.7545	0.7536	0.7558	0.7561	0.7536	0.7377	0.7432	0.7749
	Patra [7]	0.6675	0.6829	0.6655	0.6685	0.6721	0.6804	0.6689	0.6721	0.6577	0.6670
	Duman [11]	0.5020	0.4727	0.4482	0.4600	0.4785	0.4824	0.5068	0.4746	0.5391	0.5322
Peppers	Proposed	0.9028	0.8589	0.7807	0.8253	0.8075	0.7571	0.7623	0.7410	0.7158	0.7274
	Li [27]	0.8030	0.7787	0.7713	0.7258	0.7248	0.7235	0.7122	0.7193	0.7826	0.7668
	Patra [7]	0.6726	0.6980	0.6763	0.6755	0.6853	0.6838	0.6716	0.6792	0.6824	0.6699
	Duman [11]	0.4941	0.4932	0.4912	0.4902	0.5088	0.5000	0.4912	0.5283	0.5049	0.5000

Hua *et al. EURASIP Journal on Image and Video Processing* (2017) 2017:10

Page 16 of 17

Table 8 Normalized Hamming similarity results for *X* shearing % attack

Image	Scheme	1%	2%	3%	4%	5%	6%	7%	8%	9%	10%
Lenna	Proposed	0.8482	0.8127	0.8566	0.8117	0.8075	0.7968	0.7923	0.7519	0.7380	0.7820
	Li [27]	0.7348	0.7503	0.7432	0.7510	0.7422	0.7045	0.7355	0.7251	0.7103	0.7067
	Patra [7]	0.6641	0.6904	0.6553	0.6504	0.6765	0.6511	0.6494	0.6414	0.6382	0.6379
	Duman [11]	0.6758	0.6406	0.5498	0.4805	0.5488	0.5889	0.5234	0.5977	0.5508	0.6602
Aerial	Proposed	0.6670	0.6902	0.6851	0.6615	0.6605	0.6647	0.6760	0.6799	0.6912	0.6757
	Li [27]	0.7167	0.6321	0.6318	0.6783	0.6854	0.6292	0.6925	0.6857	0.6718	0.6728
	Patra [7]	0.6492	0.6660	0.6467	0.6494	0.6572	0.6387	0.6455	0.6501	0.6411	0.6345
	Duman [11]	0.5840	0.5488	0.4863	0.4863	0.5303	0.5020	0.5508	0.5029	0.4990	0.5498
Airfield	Proposed	0.7542	0.7584	0.7229	0.7057	0.6457	0.6347	0.5727	0.6809	0.7154	0.6402
	Li [27]	0.6718	0.6873	0.6899	0.6541	0.6699	0.6596	0.6709	0.6667	0.6651	0.6686
	Patra [7]	0.6406	0.6599	0.6470	0.6418	0.6636	0.6501	0.6431	0.6509	0.6414	0.6379
	Duman [11]	0.6318	0.5225	0.4863	0.5791	0.5576	0.5244	0.5430	0.5078	0.5215	0.5508
Lake	Proposed	0.8004	0.8388	0.7555	0.8437	0.7852	0.8194	0.7339	0.7736	0.7287	0.7161
	Li [27]	0.7513	0.7558	0.7603	0.7494	0.7293	0.7313	0.7274	0.7293	0.7258	0.7145
	Patra [7]	0.6846	0.6953	0.6838	0.6821	0.6851	0.6765	0.6643	0.6755	0.6660	0.6545
	Duman [11]	0.6279	0.5684	0.5020	0.5693	0.5977	0.4814	0.5938	0.5283	0.4932	0.5762
Goldhill	Proposed	0.7703	0.8043	0.7988	0.7804	0.7875	0.6815	0.6047	0.5998	0.6315	0.6298
	Li [27]	0.6938	0.7067	0.7064	0.7057	0.6764	0.6986	0.6812	0.7070	0.6880	0.7051
	Patra [7]	0.6931	0.7161	0.6729	0.6853	0.6841	0.6841	0.6804	0.6692	0.6777	0.6633
	Duman [11]	0.6123	0.6768	0.5205	0.4795	0.5410	0.5264	0.5234	0.5928	0.5039	0.5801
Barbara	Proposed	0.7636	0.7426	0.6961	0.7096	0.7445	0.7723	0.5756	0.6340	0.5730	0.5601
	Li [27]	0.7368	0.7497	0.7335	0.7358	0.7371	0.7277	0.7406	0.7355	0.7309	0.7422
	Patra [7]	0.6538	0.6597	0.6453	0.6460	0.6628	0.6326	0.6418	0.6389	0.6411	0.6160
	Duman [11]	0.5957	0.6055	0.5361	0.4971	0.5830	0.4912	0.6250	0.5664	0.5781	0.5586
Baboon	Proposed	0.7277	0.7613	0.7623	0.7028	0.6683	0.7067	0.7765	0.6260	0.6505	0.6744
	Li [27]	0.6557	0.6541	0.6521	0.6486	0.6421	0.6402	0.6402	0.6466	0.6399	0.6318
	Patra [7]	0.6057	0.6365	0.6060	0.6074	0.6326	0.6089	0.6116	0.6118	0.6152	0.5999
	Duman [11]	0.5752	0.5713	0.5176	0.5254	0.5215	0.4922	0.5146	0.5127	0.5322	0.5664
Elaine	Proposed	0.7125	0.6638	0.6883	0.6389	0.7258	0.6634	0.6676	0.6741	0.6434	0.6621
	Li [27]	0.7477	0.7474	0.7500	0.7452	0.7345	0.7345	0.7400	0.7406	0.7284	0.7393
	Patra [7]	0.6487	0.6567	0.6406	0.6482	0.6677	0.6306	0.6448	0.6458	0.6321	0.6274
	Duman [11]	0.6797	0.4629	0.5781	0.5518	0.5039	0.5762	0.5342	0.5469	0.5137	0.5762
Peppers	Proposed	0.8773	0.7836	0.7665	0.7571	0.7881	0.8353	0.8207	0.8198	0.7552	0.7539
	Li [27]	0.7988	0.7691	0.7500	0.7316	0.7232	0.7778	0.7910	0.7833	0.7574	0.7510
	Patra [7]	0.6602	0.6804	0.6511	0.6589	0.6665	0.6426	0.6458	0.6494	0.6475	0.6394
	Duman [11]	0.6084	0.5332	0.5332	0.5850	0.5879	0.5811	0.5576	0.6035	0.5684	0.5273

Hua *et al. EURASIP Journal on Image and Video Processing* (2017) 2017:10

Page 17 of 17

Authors' contributions
KH and KS proposed the framework of this work and drafted the manuscript; BD and YH carried out the algorithm studies, participated in the simulation, and helped draft the manuscript. VS participated in the discussion, corrected the English errors, and helped polish the manuscript. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan (ROC). [2]Department of Computer Science and Information Engineering, National Ilan University, Yilan County, Taiwan (ROC). [3]Computer Science and Engineering Department, Thapar University, Patiala, Punjab, India.

References
1. J Sanchez-Riera, K-L Hua, Y-S Hsiao, T Lim, SC Hidayati, W-H Cheng, A comparative study of data fusion for RGB-D based visual recognition. Pattern Recogn. Lett. **73**, 1–6 (2016)
2. W-H Cheng, C-W Wang, J-L Wu, Video adaptation for small display based on content recomposition. IEEE Trans. Circ. Syst. Video Technol. **17**(1), 43–58 (2007)
3. T Lim, K-L Hua, H-C Wang, K-W Zhao, M-C Hu, W-H Cheng, in *Multimedia Signal Processing (MMSP), 2015 IEEE 17th International Workshop On 2015 Oct 19*. Vrank: voting system on ranking model for human age estimation (IEEE, 2015), pp. 1–6
4. K-L Hua, C-H Hsu, SC Hidayati, W-H Cheng, Y-J Chen, Computer-aided classification of lung nodules on computed tomography images via deep learning technique. OncoTargets Ther. **8**, 2015–2022 (2015)
5. GC Langelaar, RL Lagendijk, Optimal differential energy watermarking of DCT encoded images and video. IEEE Trans. Image Process. **10**(1), 148–158 (2001)
6. TK Das, S Maitra, J Mitra, Cryptanalysis of optimal differential energy watermarking (DEW) and a modified robust scheme. IEEE Trans. Image Process. **53**(2), 768–775 (2005)
7. JC Patra, JE Phua, D Rajan, DCT domain watermarking scheme using Chinese remainder theorem for image authentication. IEEE Int. Conf. Multimed. Expo., 111–116 (2010)
8. H Ling, Z Lu, F Zou, Improved differential energy watermarking (IDEW) algorithm for DCT encoded image and video. IEEE Int. Conf. Signal Process. **3**, 2326–2329 (2004)
9. Z Wei, K N.Ngan, Spatial just noticeable distortion profile for image in DCT domain. IEEE Int. Conf. Multimed. Expo., 925–928 (2008)
10. Y Niu, J Liu, S Krishnan, Q Zhang, Combined just noticeable difference model guided image watermarking. IEEE Int. Conf. Multimed. Expo., 1679–1684 (2010)
11. O Duman, O Akay, A new method of wavelet domain watermark embedding and extraction using fractional Fourier transform. IEEE Int. Conf. Electr. Electron. Eng. (ELECO), 187–191 (2011)
12. S Ababneh, R Ansari, A Khokhar, Compressed-image authentication using compensated signature embedding. IEEE Int. Conf. Electro/Information Tech., 476–481 (2007)
13. S Ababneh, A Khokhar, R Ansari, Compensated signature embedding based multimedia content authentication system. IEEE Int. Conf. Image Process., 393–396 (2007)
14. S Ababneh, R Ansari, A Khokhar, Improved image authentication using closed-form compensation and spread-spectrum watermarking. IEEE Int. Conf. Acoust. Speech and Signal Process., 1781–1784 (2008)
15. H Yuan, X-P Zhang, A multiscale fragile watermark based on the Gaussian mixture model in the wavelet domain. IEEE Int. Conf. Acoust. Speech Signal Process., 413–416 (2004)
16. H Yuan, X-P Zhang, A secret key based multiscale fragile watermark in the wavelet domain. IEEE Int. Conf. Multimedia Expo., 1333–1336 (2006)
17. X Wang, J Wang, H Peng, A semi-fragile image watermarking resisting to JPEG compression. IEEE Int. Conf. Manag. e-Commerce e-Government., 498–502 (2009)
18. T Huynh-The, O Banos, S Lee, Y Yoon, T Le-Tien, Improving digital image watermarking by means of optimal channel selection. Expert Syst. Appl. **62**, 177–189 (2016). doi:10.1016/j.eswa.2016.06.015
19. X-y Wang, P-p Niu, H-y Yang, C-p Wang, A-l Wang, A new robust color image watermarking using local quaternion exponent moments. Inf. Sci. **277**, 731–754 (2014)
20. AM Abdelhakim, HI Saleh, AM Nassar, Quality metric-based fitness function for robust watermarking optimisation with bees algorithm. IET Image Process. **10**(3), 247–252 (2016). doi:10.1049/iet-ipr.2015.0379
21. KC Choi, CM Pun, in *2016 13th International Conference on Computer Graphics, Imaging and Visualization (CGiV)*. Difference expansion based robust reversible watermarking with region filtering (IEEE, Morocco, 2016), pp. 278–282. doi:10.1109/CGiV.2016.61
22. M Moosazadeh, A Andalib, in *2016 Second International Conference on Web Research (ICWR)*. A new robust color digital image watermarking algorithm in DCT domain using genetic algorithm and coefficients exchange approach (University of Science and Culture, Tehran, 2016), pp. 19–24. doi:10.1109/ICWR.2016.7498441
23. S Kadu, C Naveen, VR Satpute, AG Keskar, in *2016 IEEE Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*. A blind video watermarking technique for indoor video content protection using discrete wavelet transform, (Bhopal, 2016), pp. 1–6. doi:10.1109/SCEECS.2016.7509264
24. T-H Tsai, W-H Cheng, C-W You, M-C Hu, AW Tsui, H-Y Chi, Learning and recognition of on-premise signs from weakly labeled street view images. IEEE Trans. Image Process. **23**(3), 1047–1059 (2014)
25. DG Lowe, Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. **60**(2), 91–110 (2004)
26. AL Cunha, J Zhou, MN Do, The nonsubsampled contourlet transform: theory, design, and application. IEEE Trans. Image Process. **15**(10), 3089–3101 (2006)
27. L Li, J Qian, J-S Pan, High capacity watermark embedding based on local invariant features. IEEE Int. Conf. Multimedia Expo., 1311–1314 (2010)
28. John Canny, A computational approach to edge detection. IEEE Trans. Pattern Anal. Mach. Intell. **PAMI-8**(6), 679–698 (1986)