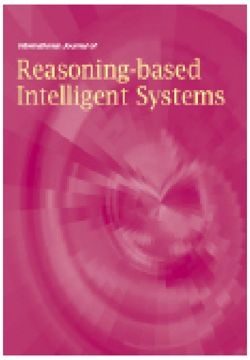


International Journal of Reasoning-based Intelligent Systems

ISSN online: 1755-0564 - ISSN print: 1755-0556
https://www.inderscience.com/ijris

A block-based fragile watermarking scheme for digital image authentication and tamper recovery

Monalisa Swain, Debabala Swain

DOI: 10.1504/IJRIS.2023.10052941

Article History:

Received:	03 August 2022
Accepted:	23 November 2022
Published online:	19 March 2024

A block-based fragile watermarking scheme for digital image authentication and tamper recovery

Monalisa Swain* and Debabala Swain

Department of Computer Science,
Rama Devi Women's University,
Bhubaneswar, India
Email: ys.monalisa@gmail.com
Email: debabala@rdwu.ac.in
*Corresponding author

Abstract: Unauthorised access and modification of multimedia content are on the rise with the huge increase in digital communication and multimedia information exchange over the web. For the transmitted images to remain protected and authentic, fraud identification and restoration procedures are required. In light of the aforementioned difficulties, a unique self-embedding block-based fragile watermarking method is introduced with enhanced tamper identification and restoration abilities. In this presented method of watermarking, the cover image is split into non-overlapping 4 × 4 block segments. Seven MSBs of each pixel in the block are used to create the watermark data. The mapping number for each block, where the recovery information is embedded, is created using a key value for that block. The various tampering rates and the number of altered images are used to evaluate the proposed methodology. The SSIM and PSNR values of the recovered image illustrate the uniqueness and efficacy of the method.

Keywords: spatial domain; fragile watermarking; least significant bit; LSB; singular value decomposition; authentication; image recovery.

Reference to this paper should be made as follows: Swain, M. and Swain, D. (2024) 'A block-based fragile watermarking scheme for digital image authentication and tamper recovery', *Int. J. Reasoning-based Intelligent Systems*, Vol. 16, No. 1, pp.1–7.

Biographical notes: Monalisa Swain is currently pursuing her PhD in the Department of Computer Science, Rama Devi Women's University, Bhubaneswar, Odisha. Her area of research interest includes information security, machine learning, and IoT.

Debabala Swain is working as an Associate Professor in the Department of Computer Science, Rama Devi Women's University, Bhubaneswar, Odisha. Her area of research interest includes high-performance computing, information security, machine learning, and IoT.

1 Introduction

People are transferring information through the Internet more frequently as a result of the network's quick development. The extent and scope of information transmission through digital media have surpassed all previous records. Due to advanced image processing technology and the widespread accessibility of editing tools, digital image content is susceptible to unauthorised changes and alterations. Identification, localisation, and recovery of modified images have emerged as major issues (Shi et al., 2016; Liu and Yuan, 2021). To safeguard information, several techniques are employed in cryptography. The perception of data itself is altered by encryption methods. Only a valid key may be used to decrypt the information. Digital watermarking is an approach for embedding a secret message or information into an image, where the secret message or information is referred to as the watermark (Gutub et al., 2010; Barni et al., 2006). The image on which the hidden message is placed is referred to as the host or cover image. The watermark comprises authentication bits for identity verification as well as recovery bits for recovering the altered data of the cover image.

In this study, we propose a unique self-embedding watermarking approach for recovering images with high tampering rates. We were able to retrieve 75% of the tampered-with photos. Singular value decomposition and the LSB method are used in the watermarking approach. In this proposed technique, the cover image is split into 4 × 4 blocks. The singular value decomposition of the block is used to produce authentication data, and the block average is used for recovery data. The watermarks are applied using the LSB technique to boost the watermarked image's imperceptibility. The testing results show that the suggested approach delivers a higher quality recovered picture with acceptable visual quality by adding less distortion to watermarked images.

The remaining section is organised as follows. Section 2 provides an overview of related work. Section 3 describes the proposed scheme's algorithm for watermark embedding and extraction. Section 4 contains the experimental results and analysis. Finally, Section 5 brings the overall observations to a conclusion.

2 Related works

Digital watermarking approaches are most commonly employed in industrial, remote sensing (Khosravi et al., 2018; Jindal et al., 2018), and healthcare applications for forgery detection, monitoring, fingerprinting, copyright protection, and ownership identification (Singh and Agarwal, 2016). For tamper identification and recovery, many researchers have suggested self-embedding fragile watermarking methods.

In 2018, Singh and Singh (2016) proposed a watermarking method with improved quality of the recovered image. For altered location identification, positioning, and restoration, a block-wise technique is utilised. This technique reduces the false tampered detection error by embedding authentication bits in the block itself rather than the mapped block. This scheme's usage of two-level tampered detection techniques adds to its effectiveness. Altered portions will be identified at level 2 of the tamper detection methods if they are not found at level 1 of the analysis. This technique uses block size (2 × 2) and is further protected by the use of predetermined user keys. The recovery was accomplished using 3 × 3 neighbouring blocks, hence eliminating the tampering coincidence problem, if the recovery bits are unable to extract from the mapped block.

In 2018, Wang et al. presented a fragile watermarking approach with dynamic embedding capacity. This approach removed three least significant bits (LSBs) from the original image and separated them into blocks of size 2 × 2. Following that, SVD was conducted block-wise to get the eigenvector. The eigenvector trace was then computed and translated into 9-bit sequences. Each block's authentication data set was built using this 9-bit sequence. Different processes for different block textures were used to generate recovery data. If the block was smooth, the recovery data was calculated by taking 5 MSBs from each block's mean value. Alternatively, if the block appeared rough, the recovery code was created by extracting the AC and DC coefficients from the DCT of the block. Block-wise, the recovery watermark is incorporated into its mapping block after it has been encrypted using a binary pseudo-random sequence. The three-level detection technique is used on the detecting side to detect and identify tampered locations.

In 2020, Gul and Ozturk presented a watermarking approach for effective image tamper detection and recovery. The suggested method uses merely two LSBs of the pixels to effectively incorporate a watermark without degrading the image's visual appearance. The suggested approach split the source image into 16 non-overlapping major blocks. Thereafter, four partner blocks were chosen from the main blocks using a look-up table and then subdivided into four sub-blocks. These sub-blocks average values are considered recovery information. The triple recovery information for each partner block is then produced by combining the recovery data from the three additional partner blocks. Each partner block is separated into 16 × 16 blocks and again subdivided into four sub-blocks. The recovery bits were inserted in the three sub-blocks. In the fourth sub-block, the two LSBs of the pixels were set to zero to create the authentication information. The MD5 hash technique is employed to generate hash code for the four sub-blocks that used a block number and a secret key. Finally, the LSB technique is used to insert 128-bit authentication data into the two LSBs of the fourth sub-block.

In 2020, Hemida and He proposed an image watermarking method based on a quantum chaos map and BTC. The host image is separated into blocks of size 2 × 2 in the suggested approach. To limit watermark capacity, BTC is employed block-wise to create recovery and authentication data. Using the LSB approach, the recovery and authentication watermarks are integrated into the pixel bits. To overcome the limitation of watermarking, such as low security and limited key space, a quantum chaos map is employed to produce the embedding position of recovery information and extend the key space. To increase the imperceptibility of the watermarked image, the LSBs are chosen to incorporate the watermark data. The testing findings reveal that the suggested approach adds less distortion to watermarked images.

3 Proposed method

In this section, an improved strategy against tamper identification and recovery is discussed. The proposed watermarking scheme uses a single LSB of the pixels for embedding the watermark efficiently maintaining perceptibility. First, the host image is divided into 4 × 4 non-overlapping blocks. An 8-bit block authentication data is created block-wise using the singular value decomposition technique. The recovery watermark data is generated by taking the block average, which helped in the restoration of the tampered block. The proposed method includes two main processes: watermark generation and embedding tamper identification and recovery.

3.1 Watermark generation and embedding

The embedding procedure uses the spatial domain for watermarking to achieve maximum fragility against the majority of attacks. For each block, the watermark is created by using the seven MSB of the pixels in that block. The watermark bits replace the first LSB layer of the cover image. Each block of size 4 × 4 generates two different sorts of watermarks: the first one, termed authentication bits (8 bits), is used to identify and locate the tampered blocks, and the second one, termed recovery bits (8 bits), is used to restore the content of the modified image. To create block dependency, recovery bits are incorporated into the mapped block and authentication bits are embedded into the block itself.

The first step is to partition the source image C, a 512 × 512 greyscale image, into 4 × 4 non-overlapping blocks. The watermark is then created block-by-block. When the cover image C is a colour image, it is first processed for dimension reduction, which involves splitting a three-dimensional image into three two-dimensional images. From this, the three layers of the carrier image K_i (R, G, and B), where i = 1, 2, and 3, respectively, are obtained. Each layer image becomes a greyscale image after dimension reduction, and watermarking is performed as a greyscale image. The watermarked image K' is then created by combining the three layers of the watermarked image K_i*.

The watermark embedding processes are shown in Figure 1. The following details the steps involved in creating and embedding a watermark.

Step 1 The cover image C was first split into non-overlapping blocks of size 4 × 4,

$$C = \{C_i \mid i = 1, 2, ..., Z\}$$

where $Z = N \times N / 16$ is the total number of blocks.

Step 2 For each block C_i, a mapping block C_j is derived as follows.

$$j = (key \times i) \bmod Z + 1 \quad (1)$$

where $key[1, Z]$ is a prime number chosen by the embedder and is the secret key.

Step 3 Eight authentication bits are created for each block in the following manner.

Each block is divided into three matrices U, S, and V by applying SVD.

$$C_i = USV^T \quad (2)$$

After applying SVD, the trace of the singular matrix S is generated by taking the summation of diagonal elements of S. This trace value is then converted into 8 bits in binary format. X-OR an operation are applied between 7 MSBs of each pixel of the block to generate 16 bits and again X-OR is applied between the first 8 bits and the other 8 bits of generated 16 bits to get 8 bits. Again X-OR operations between these 8 bits and trace 8 bits are applied to get the authentication watermark.

Step 4 By calculating the average intensity of each block, recovery information is produced for each block. It is calculated as follows:

$$Avg = \left(\sum_{i=1}^{16} Pixel\ Value_i\right) \Big/ 16 \quad Avg \in [0, 255] \quad (3)$$

Step 5 Watermark embedding: The LSB replacement mechanism is used to embed the watermark. Seven MSBs of the cover image remain unaltered in this watermark embedding approach, while the watermarks are substituted for one LSB. The recovery bits are embedded block-wise in their mapped block. The watermarked image (C_w) is generated when all of the cover image's blocks have been embedded.

3.2 *Tamper identification and recovery*

The received watermarked image's authenticity is checked block-by-block on the receiving end. The encoded watermark bits will let the recipient identify the altered blocks after identification. The receiver will use the recovery bits from each mapped block of the received image to recover the tampered blocks. The tamper identification and recovery processes are shown in Figure 2. In the case of a colour image, watermarked carrier image K' is divided into three watermarked layered images K_i' (R, G, and B) by reducing dimension processing. Each layer is processed as a greyscale image.

During the tampering identification process, the received image is processed block-wise of the size of 4 × 4 pixels. Authentication bits for each received block are generated using the same method used during watermarking and matched with authentication information embedded during watermarking to mark the block as authentic or tampered.

During the recovery process, each block's block mapping number is produced using the same embedding key. When the mapping block is authentic, the embedded recovery bits are extracted from LSB and are used as block recovery values. Whereas for tampered mapping block, authentic blocks from four-neighbourhood are used and their average intensity value replaces all of the pixel values in the tampered block. The tampered block is now remarked as authentic.

4 Experimental results and analysis

In this section, tests are carried out on various widely used 512 × 512 greyscale and colour images in order to assess the effectiveness of the suggested approach. Only one LSB of each pixel is utilised for watermark insertion in the suggested approach. The suggested scheme's invisibility and structural similarity index (SSIM) is shown using nine images. The host images and its watermarked images are illustrated in Figures 3(a) and 3(b) and Figures 4(a) and 4(b), respectively. The graphic makes it obvious that human eyes are unable to discriminate between the original and watermarked images' visual variations. The efficiency of the algorithm has been evaluated using the peak signal-to-noise ratio (PSNR) and SSIM metrics.

Figure 1 Watermark generation and embedding

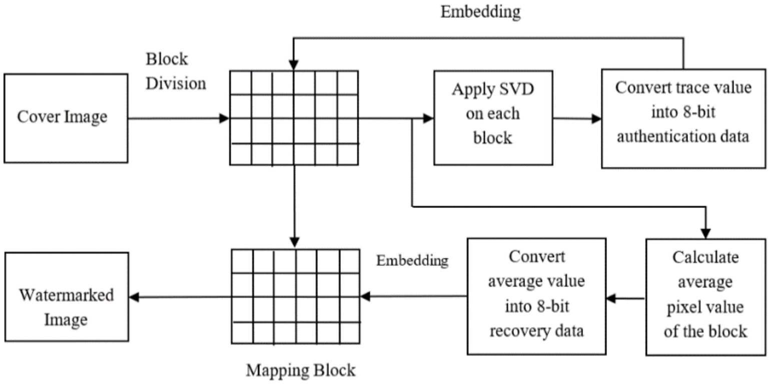
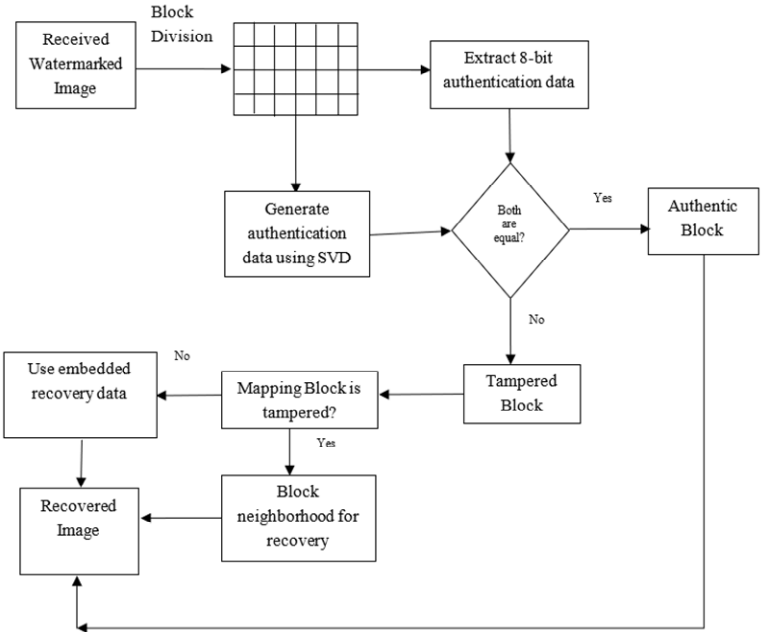

Embedding
Cover Image
Block Division
Apply SVD on each block
Convert trace value into 8-bit authentication data
Watermarked Image
Embedding
Convert average value into 8-bit recovery data
Calculate average pixel value of the block
Mapping Block

Figure 2 Tamper identification and recovery


Received Watermarked Image
Block Division
Extract 8-bit authentication data
Generate authentication data using SVD
Both are equal?
Yes
Authentic Block
No
Use embedded recovery data
No
Mapping Block is tampered?
Tampered Block
Yes
Recovered Image
Block neighborhood for recovery

In Table 1, the statistical findings are displayed. The proposed scheme's PSNR ranges from 52.38 dB to 57.18 dB, as indicated in Table 1. The outcomes show that the watermarked images' visual quality has been successfully preserved.

Table 1 Watermarked image

Cover image	*PSNR*	*SSIM*	*TIME (sec)*
Lena	55.45 dB	0.9969	21.31
Clown	57.18 dB	0.9965	17.12
Lake	55.38 dB	0.9978	7.83
Truck	55.14 dB	0.9979	7.65
Jetplane_colour	52.38dB	0.9973	22.98
House_colour	52.38 dB	0.9993	20.87
Mandril_colour	52.39 dB	0.9998	17.55
Peppers_colour	52.44 dB	0.9998	17.44
Pirate	55.48 dB	0.9978	14.32

The effectiveness of the suggested technique is analysed with respect to general cropping and tampering attacks. In general, tampering involves changing the watermarked image by including contents that were taken from that image. Figures 3 and 4 shows the results of the proposed algorithm against general tampering.

The tamper detection result is shown in Figures 3(c) and 3(d) and Figures 4(c) and 4(d), respectively. From the results, it can be observed that the proposed method can locate the tampered region. Table 2 displays the recovered images' PSNR and SSIM results. A high PSNR and SSIM value is attained, showing that the suggested approach is effective in locating tamper areas and recovering tampered images. The performance of the proposed method under different percentages of attacks are shown in Table 3, Figures 5 and 6. The truck image has the highest PSNR and SSIM values of any recovered image against 75% cropping attacks, with values of 32.18 dB and 0.6453, respectively. As shown in Table 3, the clown image has the highest SSIM and PSNR results for 50% of attacks, with values of 35.73 dB and 0.8142, respectively.

Figure 3 Experimental outputs for grayscale images, (a) watermarked (b) tampered (c) tamper localised (d) recovered image

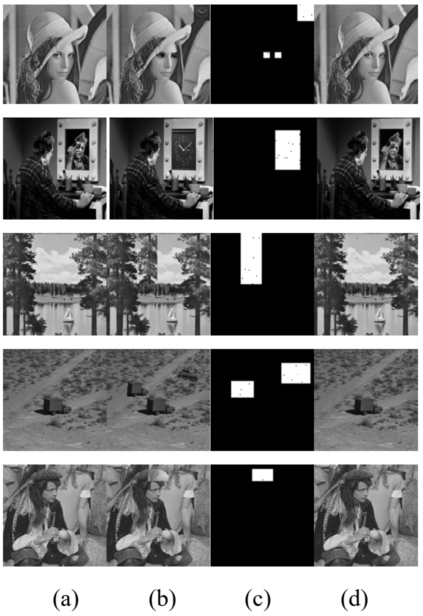
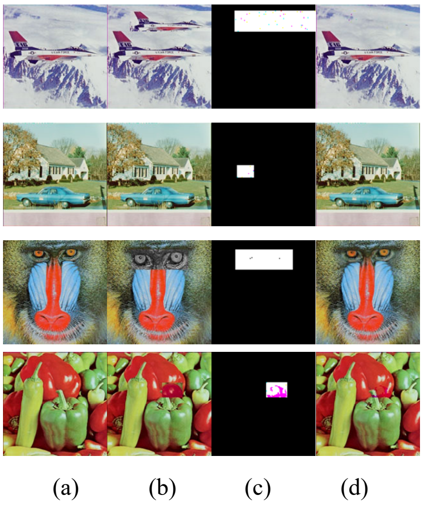


Table 2 Quality of recovered image using proposed scheme against general tampering

Cover image (512 × 512)	*PSNR of recovered image*	*SSIM*
Lena	47.80 dB	0.9907
Clown	41.41 dB	0.9679
Lake	45.29 dB	0.9810
Truck	43.08 dB	0.9689
Jetplane_colour	40.64dB	0.9137
House	37.69 dB	0.9817
Mandril_colour	28.60 dB	0.9000
Peppers	33.31 dB	0.9929
Pirate	50.07 dB	0.9928

Figure 4 Experimental outputs for colour images, (a) watermarked (b) tampered (c) tamper localised (d) recovered image (see online version for colours)



Images of Lena, clown, lake, truck, and pirate are used to assess the time complexity of the suggested approach. Watermark embedding time and tamper detection time are measured in comparison to 20%, 25%, 30%, 50%, and 75% cropping attacks, and shown in Table 4. The tests are carried out using a machine with an Intel Core i3, 1.20 GHz CPU, and 4 GB RAM.

Table 3 Recovered image against different percentages of attack

Cover image		*Tamper rate*				
		20%	*25%*	*30%*	*50%*	*75%*
Lena	PSNR	42.69 dB	38.36 dB	37.49 dB	34.61 dB	31.64 dB
	SSIM	0.9649	0.9206	0.9017	0.8164	0.6710
Clown	PSNR	38.20 dB	37.27 dB	38.29 dB	35.73 dB	31.72 dB
	SSIM	0.9237	0.8922	0.9034	0.8142	0.5984
Lake	PSNR	39.31 dB	36.92 dB	37.17 dB	34.63 dB	31.67 dB
	SSIM	0.9288	0.8818	0.8874	0.7947	0.6216
Truck	PSNR	40.50 dB	37.55 dB	38.37 dB	34.70 dB	32.18 dB
	SSIM	0.9222	0.8928	0.9001	0.7912	0.6453
Pirate	PSNR	39.40 dB	37.69 dB	37.62 dB	33.96 dB	31.73 dB
	SSIM	0.9247	0.8974	0.8957	0.7627	0.6027

Figure 5 (a) Tampered Lena image (30%–75%) (b) Tamper localised image (c) Recovered image

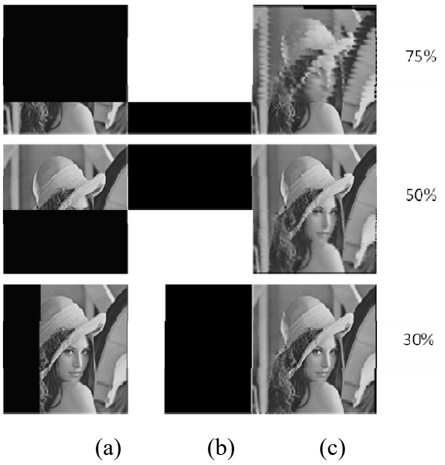


Table 5 shows the comparison of the performance of the proposed method to that of other approaches in terms of PSNR value with a Lena image. The table shows that the watermarked image got using the suggested scheme has a higher PSNR value than the Gul and Ozturk (2020), Di Martino and Sessa (2012), Qian et al. (2011), Qin et al. (2017), Yang and Shen (2010), and Zhang et al. (2009) methods. Additionally, the table makes it clearly evident that the suggested method's recovered image PSNR value is higher than the Gul and Ozturk (2020), Qin et al. (2017), Yang et al. (2014) and Zhang et al. (2009) methods. The method (Qian et al., 2011) yields the high quality restored image when the tampering rate is under 35%. Furthermore, the approaches of Di Martino and Sessa (2012) and Yang and Shen (2010) have a high quality of recovered images against 50% tampering attacks. Our suggested approach has a higher recovery of up to 75% modification of content, which is higher than the other approaches (Gul and Ozturk, 2020; Di Martino and Sessa, 2012; Qian et al., 2011; Qin et al., 2017; Yang and Shen, 2010; Yang et al., 2014; Zhang et al., 2009).

Figure 6 (a) Tampered Lena image (20%–25%) (b) Tamper localised image (c) Recovered image

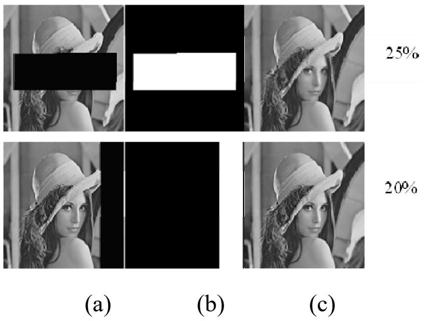


Table 4 Proposed scheme's time complexity

Cover image	*Watermark embedding time (sec)*	*Tamper detection and recovery time*				
		20%	*25%*	*30%*	*50%*	*75%*
Lena	21.31	7.59	6.70	6.95	7.57	10.29
Clown	17.12	6.42	6.79	6.94	11.01	9.59
Lake	7.83	6.73	7.31	6.96	7.51	9.95
Truck	7.65	6.46	6.81	7.23	8.24	8.91
Pirate	14.32	6.50	6.64	7.21	7.69	9.57

Table 5 Performance comparisons

Methods	*Watermarked image PSNR*	*Recovered image PSNR*	*Condition of recovery*
Gul and Ozturk (2020)	44.14	30.49	75%
Di Martino and Sessa (2012)	43.70	33.6	50%
Qian et al. (2011)	37.90	35	35%
Qin et al. (2017)	44.27	29.41	45%
Yang and Shen (2010)	40.70	32	50%
Yang et al. (2014)	51.30	24.36	50%
Zhang et al. (2009)	37.90	29.9	59%
Proposed	54.24	31.64	75%

5 Conclusions

This proposed scheme proposed a novel watermarking scheme that recovers up to 75% of tampered images. The authentication information is generated block-wise using the SVD of each block and recovery of tampered region is achieved by taking block average and block neighbourhood approach. By applying various attack percentages to various sections of the watermarked photos, the efficiency of the proposed method has been demonstrated. According to experimental findings, the proposed technique successfully restores up to 75% of tamper-rated images. The proposed technique excels at invisibility, spotting alter regions, and recovering modified content. The experiment findings demonstrate that the proposed approach delivers high quality recovery. In future work, frequency domain techniques can be used for watermarking. Dual watermarking can likewise be accomplished using the suggested approach in combination with robust watermarking methods.

References

Barni, M., Doërr, G. and Cox, I.J. (2006) 'Editorial: steganography and digital watermarking', *IEE Proceedings – Information Security*, Vol. 153, No. 3, p.75, https://doi.org/10.1049/ip-ifs:20069026.

Di Martino, F. and Sessa, S. (2012) 'Fragile watermarking tamper detection with images compressed by fuzzy transform', *Information Sciences*, Vol. 195, pp.62–90, https://doi.org/10.1016/j.ins.2012.01.014.

Gul, E. and Ozturk, S. (2020) 'A novel triple recovery information embedding approach for self-embedded digital image watermarking', *Multimedia Tools and Applications*, Vol. 79, Nos. 41–42, pp.31239–31264, https://doi.org/10.1007/s11042-020-09548-4.

Gutub, A.A-A. (2010) 'Pixel indicator technique for RGB image steganography', *Journal of Emerging Technologies in Web Intelligence*, Vol. 2, No. 1, https://doi.org/10.4304/jetwi.2.1.56-64.

Hemida, O. and He, H. (2020) 'A self-recovery watermarking scheme based on block truncation coding and quantum chaos map', *Multimedia Tools and Applications*, Vol. 79, Nos. 25–26, pp.18695–18725, https://doi.org/10.1007/s11042-020-08727-7.

Jindal, H., Kasana, S.S. and Saxena, S. (2018) 'Underwater pipelines panoramic image transmission and refinement using acoustic sensors', *International Journal of Wavelets, Multiresolution and Information Processing*, Vol. 16, No. 3, p.1850013, https://doi.org/10.1142/s0219691318500133.

Khosravi, M.R., Rostami, H. and Samadi, S. (2018) 'Enhancing the binary watermark-based data hiding scheme using an interpolation-based approach for optical remote sensing images', *International Journal of Agricultural and Environmental Information Systems*, Vol. 9, No. 2, pp.53–71, https://doi.org/10.4018/ijaeis.2018040104.

Liu, T. and Yuan, X. (2021) 'A dual-tamper-detection method for digital image authentication and content self-recovery', *Multimedia Tools and Applications*, Vol. 80, No. 19, pp.29805–29826, https://doi.org/10.1007/s11042-021-11179-2.

Qian, Z., Feng, G., Zhang, X. and Wang, S. (2011) 'Image self-embedding with high-quality restoration capability', *Digital Signal Processing*, Vol. 21, No. 2, pp.278–286, https://doi.org/10.1016/j.dsp.2010.04.006.

Qin, C., Ji, P., Zhang, X., Dong, J. and Wang, J. (2017) 'Fragile image watermarking with pixel-wise recovery based on overlapping embedding strategy', *Signal Processing*, Vol. 138, pp.280–293, https://doi.org/10.1016/j.sigpro.2017.03.033.

Shi, H., Wang, X., Li, M., Bai, J. and Feng, B. (2016) 'Secure variable-capacity self-recovery watermarking scheme', *Multimedia Tools and Applications*, Vol. 76, No. 5, pp.6941–6972, https://doi.org/10.1007/s11042-016-3328-z.

Singh, D. and Singh, S.K. (2016) 'Effective self-embedding watermarking scheme for image tampered detection and localization with recovery capability', *Journal of Visual Communication and Image Representation*, Vol. 38, pp.775–789, https://doi.org/10.1016/j.jvcir.2016.04.023.

Singh, P. and Agarwal, S. (2016) 'A self recoverable dual watermarking scheme for copyright protection and integrity verification', *Multimedia Tools and Applications*, Vol. 76, No. 5, pp.6389–6428, https://doi.org/10.1007/s11042-015-3198-9.

Wang, C., Zhang, H. and Zhou, X. (2018) 'A self-recovery fragile image watermarking with variable watermark capacity', *Applied Sciences*, Vol. 8, No. 4, p.548, https://doi.org/10.3390/app8040548.

Yang, C-W. and Shen, J-J. (2010) 'Recover the tampered image based on VQ indexing', *Signal Processing*, Vol. 90, No. 1, pp.331–343, https://doi.org/10.1016/j.sigpro.2009.07.007.

Yang, S., Qin, C., Qian, Z. and Xu, B. (2014) 'Tampering detection and content recovery for digital images using halftone mechanism', *IEEE Xplore*, https://doi.org/10.1109/IIH-MSP.2014.39.

Zhang, X., Wang, S. and Feng, G. (2009) 'Fragile watermarking scheme with extensive content restoration capability', *Digital Watermarking*, pp.268–278, https://doi.org/10.1007/978-3-642-03688-0_24.