

Automatika

Journal for Control, Measurement, Electronics, Computing and Communications



ISSN: 0005-1144 (Print) 1848-3380 (Online) Journal homepage: <https://www.tandfonline.com/loi/taut20>

A DCT domain smart vicinity reliant fragile watermarking technique for DIBR 3D-TV

Tariq Bashir, Imran Usman, Abdulaziz Abdullah Albeshar, Samer Hassan Atawneh & Syed Saud Naqvi

To cite this article: Tariq Bashir, Imran Usman, Abdulaziz Abdullah Albeshar, Samer Hassan Atawneh & Syed Saud Naqvi (2020) A DCT domain smart vicinity reliant fragile watermarking technique for DIBR 3D-TV, *Automatika*, 61:1, 58-65, DOI: [10.1080/00051144.2019.1670992](https://doi.org/10.1080/00051144.2019.1670992)

To link to this article: <https://doi.org/10.1080/00051144.2019.1670992>



© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 07 Nov 2019.



Submit your article to this journal [↗](#)



Article views: 416



View related articles [↗](#)








View Crossmark data [↗](#)



Citing articles: 2 View citing articles [↗](#)



A DCT domain smart vicinity reliant fragile watermarking technique for DIBR 3D-TV

Tariq Bashir ^a, Imran Usman ^b, Abdulaziz Abdullah Albeshir ^b, Samer Hassan Atawneh ^b and Syed Saud Naqvi ^a

^aDepartment of Electrical and Computer Engineering, COMSATS University Islamabad, Islamabad, Pakistan; ^bCollege of Computing and Informatics, Saudi Electronic University, Riyadh, Kingdom of Saudi Arabia

ABSTRACT

This work presents a vicinity reliant intelligent fragile watermarking scheme for depth image-based rendering technique used for three-dimensional television. Depth map of a centre image is implicitly inserted in the block-based discrete cosine transform (DCT) of the same using an aggregate, which also takes into account the presence of its neighbourhood blocks. Based upon the parity of a Boolean operation on the aggregate, parity is modulated which implicitly embeds the watermark. Genetic algorithm is then utilized to select the appropriate frequency bands in the DCT domain to become eligible for watermark embedding based on imperceptibility requirements. Experimental results demonstrate the usefulness of the proposed scheme in terms of its resistance against a set of fragile watermarking attacks and its ability to detect and localize tempering attempts.

ARTICLE HISTORY

Received 12 November 2018
Accepted 18 September 2019

KEYWORDS

Vicinity reliant fragile watermarking (VRFW); discrete cosine transform (DCT); depth image-based rendering (DIBR); three-dimensional television (3D-TV); genetic algorithm

1. Introduction

With the advent of Internet and high-speed connected networks, multimedia applications have protruded extensively in the last decade. Among such applications, three-dimensional television (3D-TV) is one of the most evolving and sought out technology derived by the desire of viewing prerecorded content closest to the reality. Depth image-based rendering (DIBR) is one of the most promising techniques to realize the virtual view synthesis for depth perception [1]. In DIBR, a centre image and its associated depth map is utilized to generate corresponding left and right images. These are then used to produce anaglyph image that yields depth perception with suitable viewing aids [2]. Among many challenges, content authentication is one of the primary concerns for DIBR 3D-TV content producers. Amidst many solutions, fragile digital watermarking is one of the promising techniques. In such a watermarking, slightest modification in the cover work would destroy the watermark such as to raise alarms against any tempering attempts [3].

A fragile watermark, when embedded, is likely to be destroyed by the slightest modification of the watermarked content. This enables and ensures a mechanism where an extracted watermark can be compared with the inserted bits to indicate an attempt of tempering. More advanced fragile watermarking systems are also able to localize tempered regions of the cover work. Temper localization not only helps to identify the areas

of an image where an attack was attempted but also provides an insight into the attacker's region of interest and intention. In addition, it also helps to rectify and correct the changes made in the cover work. More recent watermarking schemes are also able to automatically recover the changed bits [4].

To be adequate, a fragile watermarking technique must localize any tempering, geometric transformations, and be able to sustain any legitimate manipulations, such as image compression. It must not contain any security gaps for attacks such as cut and paste [5,6] and birthday attacks [5,7]. Researchers have shown that a very well established necessity to encounter such attacks is block-wise dependence. Barreto et al. [5] has indicated that contextually deterministic dependence is susceptible to transplantation attacks [6,8]. To deal with the transplantation attacks, Barreto et al. [5] proposed a scheme based on the hash function calculation. However, this scheme is relatively time-consuming and the size of the blocks limits accuracy of temper localization.

Several schemes have been proposed in literature including Wu and Liu's scheme [9] of inserting a binary watermark sequence in DCT coefficients in a lookup table, and Wong's public key scheme [10]. However, these schemes, along with some others [11–13] are block-wise independent and therefore are exposed to cover up, transplantation and vector quantization attacks [6,14].

CONTACT Tariq Bashir  tariq.haque81@gmail.com

This article has been republished with minor changes. These changes do not impact the academic content of the article.

Recognizing the importance of block-wise dependence, several schemes have been proposed including Edupuganti et al.'s [15] and Nyeem et al.'s [16] schemes based upon cyclic redundancy checks to authenticate the features of a block stored in pairs of mapping blocks. Li et al. [17] proposed a scheme that uses a binary feature-map extracted from an underlying image as a watermark. This scheme is simple to implement and takes into consideration of many of the challenges in fragile watermarking. However, it was not developed for DIBR and thus does not take into consideration the depth map for embedding purpose. In addition, appropriate frequency bands are not selected for watermark embedding. Yu et al. [14] utilized the concept of stereo matching and block categorization and classification to propose a fragile watermarking scheme. Their scheme classifies the non-overlapping blocks of the left image as matchable or non-matchable based on stereo matching of left and right images. For tamper localization, they employed a detection technique with alterable-length watermark to increase accuracy of tamper localization at the receiver side. Nevertheless, compared with the proposed technique, their technique lags in imperceptibility.

All these schemes were proposed for digital images and do not incorporate ways to be used for 3D-TV. Recently, Ali et al. [18] proposed a transform-based watermarking algorithm for copyright of 3D images by using discrete wavelet transform and singular value decomposition transforms to utilize their individual properties. Nevertheless, their method is not developed for content authentication and is suitable only for copyright protection applications. Rana et al. [19] developed an Sacle Invariant Feature Transform (SIFT) feature-based coefficient selection scheme for robust watermarking of 3D DIBR images which are view invariant. This scheme [19] exploits the shift-invariance and directional property of the DIBR-based view synthesis process to embed the watermark.

In this work, we propose a fragile watermarking scheme for DIBR-based 3D-TV. The proposed scheme takes into consideration the neighbourhood, or the vicinity, of the embedding location while calculating a secret sum associated with the implicit embedding of the watermark bit. It enhances Li's [17] DCT-based watermarking scheme by selecting an appropriate frequency band using the genetic algorithm (GA). In addition, the depth map of the image itself is used as a source for generating block-wise dependence. The watermark is created using the depth map and a pseudorandom sequence using a predefined secret key. The same key is used at the receiver side to authenticate the embedded watermark. In the process of watermark embedding, any of the images from the stereoscopic image pairs can be used for the embedding of the watermark. For the purpose of this study, the proposed scheme is

named as vicinity reliant fragile watermarking (VRFW) technique.

The rest of the paper is organized as follows. Section 2 explains the proposed technique with subsections on important related topics such as general architecture, GA-based training and testing, the watermark embedding and extraction processes. Section 3 presents the experimental results and discussion, whereas the conclusion is presented in Section 4.

2. VRFW technique

2.1. General architecture

In this section, the proposed VRFW watermarking scheme is presented. We first describe the general architecture of VRFW and then elaborate the individual modules in later subsections. The rendering operation for the left-eye and the right-eye image synthesis is explained separately in Section 2.4. Figure 1 shows its general architecture, whereby, a stereoscopic left or right image is block DCT transformed to produce frequency-domain coefficients. GA is then utilized to select the appropriate frequency band in block DCT for embedding of the watermark based on imperceptibility property of the watermark. On the other hand, the watermark is generated using the depth map, a secret key and a pseudorandom sequence. The frequency-domain coefficients and the watermark bit are then used to implicitly embed the watermark using a secret sum and block-wise neighbourhood dependence. Conjointly, these two provide an implicit mechanism for resistance against a number of attacks a fragile watermarking application is susceptible to. Based on the performance of the particular candidate frequency band selected by GA in terms of imperceptibility characteristics, the algorithm decides to continue or terminate its generational evolution. In the next section, the proposed embedding scheme is discussed in detail.

2.2. Watermark embedding algorithm

Figure 2 shows the watermark embedding process of the proposed VRFW scheme. First, a stereo left or right image X is 8×8 block DCT transformed to generate Y .

$$Y = DCT_{8 \times 8}(X), \quad (1)$$

and

$$Y = \bigcup_{i=1}^r Y_i(k), \quad (2)$$

where $r = (p \times q)/(8 \times 8)$ is the number of DCT blocks, p and q are the number of rows and columns of image matrix, respectively, and $k = 0, 1, 2, \dots, 63$ represent the zigzag coefficient index of each 8×8 DCT

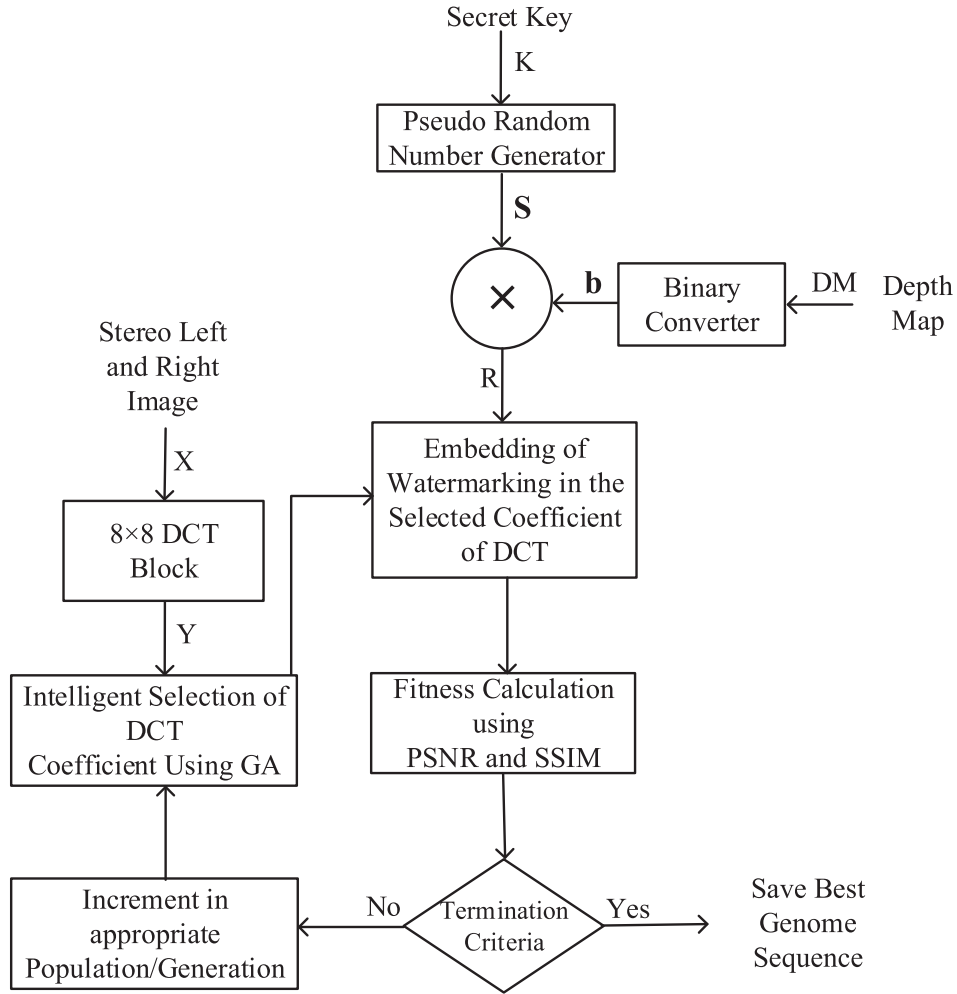


Figure 1. General architecture of VRFW scheme.

block. Using Y , a binary image M of the same size as Y is generated using the following equation:

$$M_i(j) = \begin{cases} 1, & \text{if } Y_i(j) \neq 0 \\ 0, & \text{otherwise} \end{cases}, \quad (3)$$

On the other hand, the depth map DM is converted into a binary image using a predefined threshold $0 \leq \alpha \leq 255$ to generate b . All pixel values in depth map lesser than the defined threshold are converted to 0 while the larger values are converted to 1. It is then multiplied by a pseudorandom sequence S which is generated using a secret key K . We call it R .

$$R = b.S, \quad (4)$$

The final watermark message (W) is then generated by taking exclusive OR between M and R using:

$$W = M \oplus R \quad (5)$$

In our GA simulation, we define our genome sequence as a binary sequence with length equal to the DCT block size, i.e. 64. Initially, some ‘ON’ bits’ locations in the middle-frequency bands are considered as watermarkable and the zero bits’ locations are ignored. These

watermarkable frequency locations are then optimized in the GA simulation based on the imperceptibility of the watermark which is taken as fitness of every genome and is calculated using Equation (9).

Corresponding to each watermarkable coefficient selected by GA, we compute a secret sum according to the following equation:

$$\text{SUM}_i(j) = \sum_{m=1}^9 \sum_{n \in h} (W_m(n) \oplus W_i(j)) \cdot Y_m(n), \quad (6)$$

where $n = 0, 1, 2, \dots, 63$, and m is the dependence neighbourhood including the block itself as show in Figure 3. The h in Equation (6) represents the set of indices of the watermarkable coefficients. Let $\Psi_i(j)$ be the concatenation of $\text{SUM}_i(j)$ and $Y_i(j)$ in one’s complement format. Then, let $\Psi'_i(j)$ be the concatenation of $\Psi_i(j)$ with $\Psi_i(j)$. Finally, the watermarked image Y' is produced by modulating the watermarkable coefficients of each block in a way that,

$$\text{Alarm}(\Psi'_i(j)) = W_i(j), \quad (7)$$

where $\text{Alarm}()$ is a function that results in 1 if the count of ‘1’ bits is odd and 0 if it is even. Finally, inverse DCT

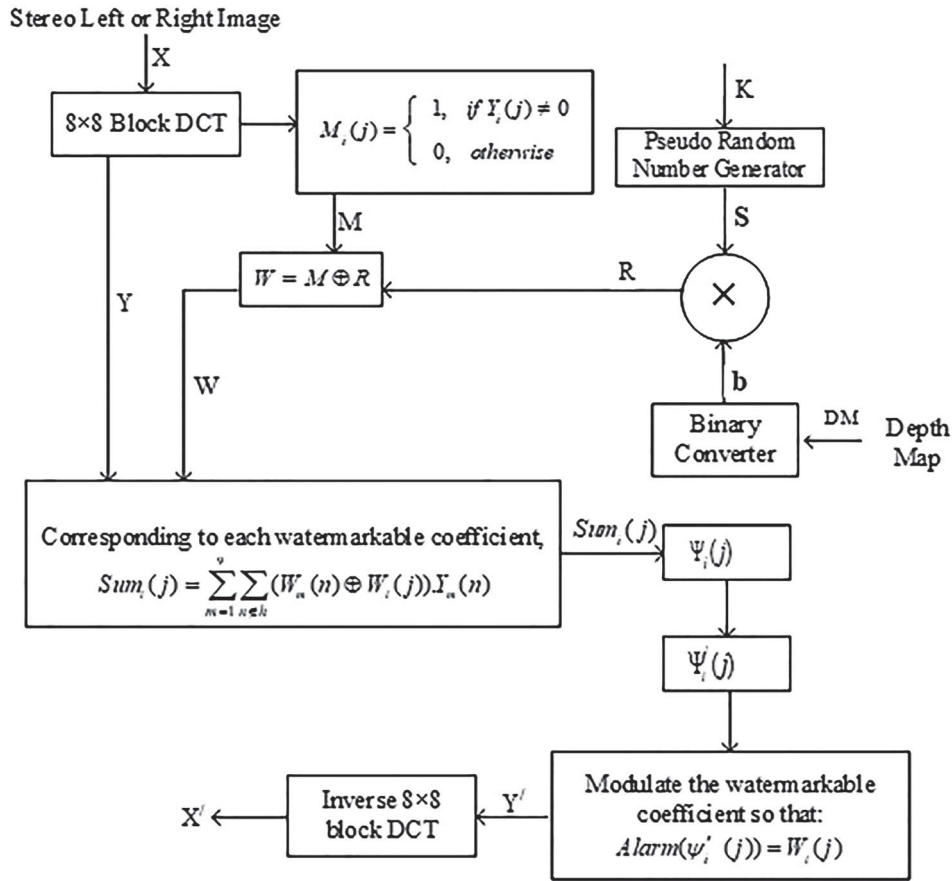


Figure 2. Block diagram of watermarking embedding phase.

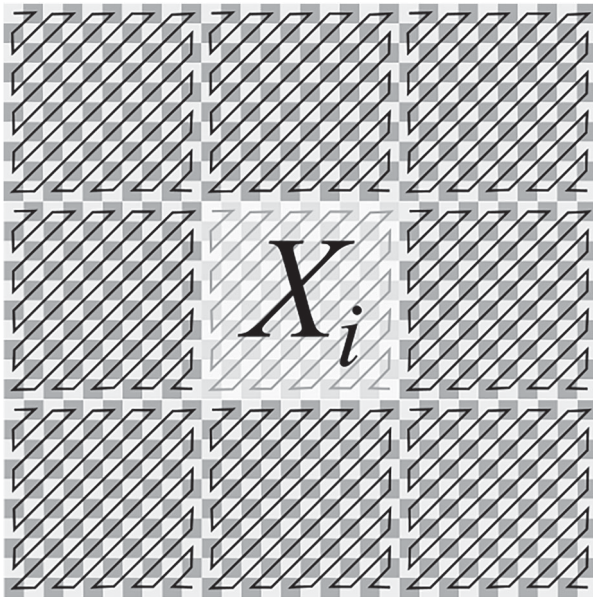


Figure 3. Nine-neighbourhood system of block X_i within 8×8 block DCT.

is performed on Y' to generate the watermarked image in the spatial domain.

$$X' = \text{IDCT}(Y'), \quad (8)$$

From a given population in GA simulation, the individuals satisfying the fitness criteria are selected for

the reproduction. For this purpose, a fitness value for each individual in the population is calculated as given below:

$$f_c = \left(\frac{\text{PSNR}}{50} + \text{SSIM} \right) / 2, \quad (9)$$

where f_c is the fitness value of the current chromosome, whereas PSNR and SSIM denote Peak Signal to Noise Ratio and Structural Similarity Index Measure, respectively. PSNR is divided by 50 to scale its value in accordance with SSIM value. They are the quality metrics that give better visual image analysis in any given scenario and are calculated between original image and watermarked image. In any population, best individuals are selected based on f_c to make the next generation using replication, mutation and crossover to achieve diversity as well as convergence in the solution space. The algorithm iteratively enhances the fitness, generation by generation, until the stopping criteria are met. The convergence or stopping criteria is mainly based on the optimal choice of DCT coefficients for the embedding of the watermark. The best DCT locations are saved as K_0 . The keys K and K_0 are transmitted over a secure network, or hardcoded in the decoder, to be used in watermark extraction.

2.3. VRFW watermark extraction algorithm

The proposed scheme provides blind watermark extraction, i.e. the original image is not required on the extraction side. First, the watermarked image X' is 8×8 DCT transformed to generate Y' . The same embedding procedure is adopted to generate the binary sequence M , R and W . The same key K_0 is used to identify the watermarked coefficients, and their indices are saved in h . For the corresponding marked coefficients, $SUM_i(j)$ is calculated using Equation (6). The concatenation of $SUM_i(j)$ with $Y_i(j)$ in *one's* compliment to produce $\Psi_i(j)$ and $\Psi'_i(j)$ as explained in the previous section. The watermarked image is then authenticated using Equation (7). If Equation (7) does not hold, the image is tempered and the block which raises the alarm is the tempering location.

2.4. Rendering operation

The synthesis of left-eye and/or the right-eye images is done through pixel-wise warping of the centre image [20]. For any depth image, let the depth values range from ζ_{near} to ζ_{far} . Where ζ_{near} is the nearest object plane and ζ_{far} is the farthest object plane in the 3D scene. For an 8-bit depth image, the depth values map the intensity values $[0, 255]$ in the range $[\zeta_{far}, \zeta_{near}]$. The proposed rendering operation warps the pixel location according to the depth value using the following equations:

$$x_L = x_C + \left(\frac{\tau_x}{2} \times \frac{f}{\zeta} \right), \quad (10)$$

and

$$x_R = x_C - \left(\frac{\tau_x}{2} \times \frac{f}{\zeta} \right). \quad (11)$$

where x_L and x_R are the corresponding x -coordinate of the pixels in the left and right eye-images, respectively. x_C is the x -coordinate corresponding pixel in the centre image. ζ characterizes the depth value of the current pixel corresponding to the centre image. τ_x represents the camera distance and f is the focal length of the camera.

3. Experimental results and discussion

The proposed VRFW scheme is implemented in Matlab using Intel Core i7 2.4 GHz processor with an 8 GB onboard RAM. Middlebury Stereo Vision Lab 2006 image datasets [21–23] are used for the evaluation of our proposed technique. The size of each image as well as the corresponding depth map used in experiments are 256×256 . In our GA simulation, we have used Matlab GA Toolbox with a population size of 100 and maximum generation size of 20. The fitness function used is presented in Equation (9) and comprises of both PSNR and SSIM performance metrics.

Figure 4 shows the strength of the watermark embedded in different 8×8 blocks of DCT for Bowling2 image. It can be seen that the proposed VRFW scheme is intelligent enough to learn different frequency bands and their response to additive watermarking. A high strength watermark is embedded in areas where it is easily concealed, and a low strength watermark or none is embedded in areas where it

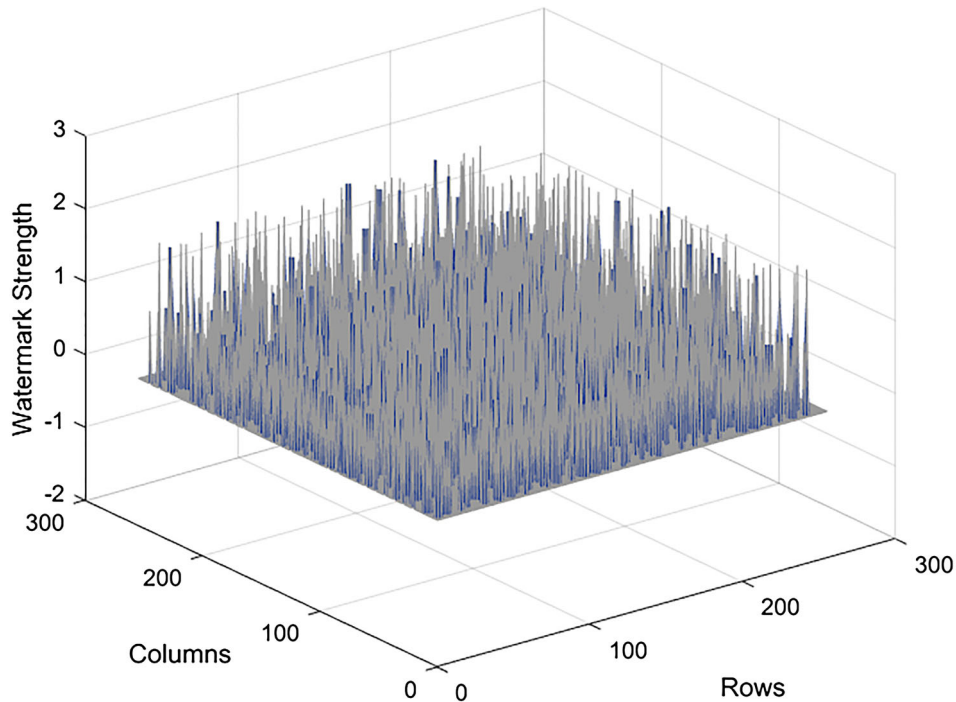


Figure 4. Watermark strength distribution in block DCT frequency bands.

would be more perceptible and increase distortion. At the same time, the proposed scheme caters for the imperceptibility constraints of the fragile watermarking domain. Also, most of the watermark is embedded in the middle frequency bands which are highly desired in watermarking applications. If the watermark is embedded in higher frequencies, it is susceptible to a low-pass filter attack and can be removed easily. On the other hand, adding a watermark to lower frequency regions

will increase the noise in smooth areas of the image and will be highly perceptible.

Table 1 demonstrates the performance of the proposed VRFW scheme in terms of watermark imperceptibility. Different values of PSNR and SSIM corresponding to different images are shown. The higher values of PSNR (which are around 50 dB) and SSIM (which are almost 1) depict that even after the insertion of the watermark, the proposed technique is able to

Table 1. PSNR and SSIM values for watermarked Middlebury datasets.

Images	Aloe	Baby1	Bowling1	Cloth1	Flowerpots	Lampshade1	Midd1
PSNR	51.10644	49.45362	51.10644	50.16846	49.93881	50.86423	51.14057
SSIM	0.999695	0.999338	0.999171	0.999687	0.999215	0.999186	0.999268

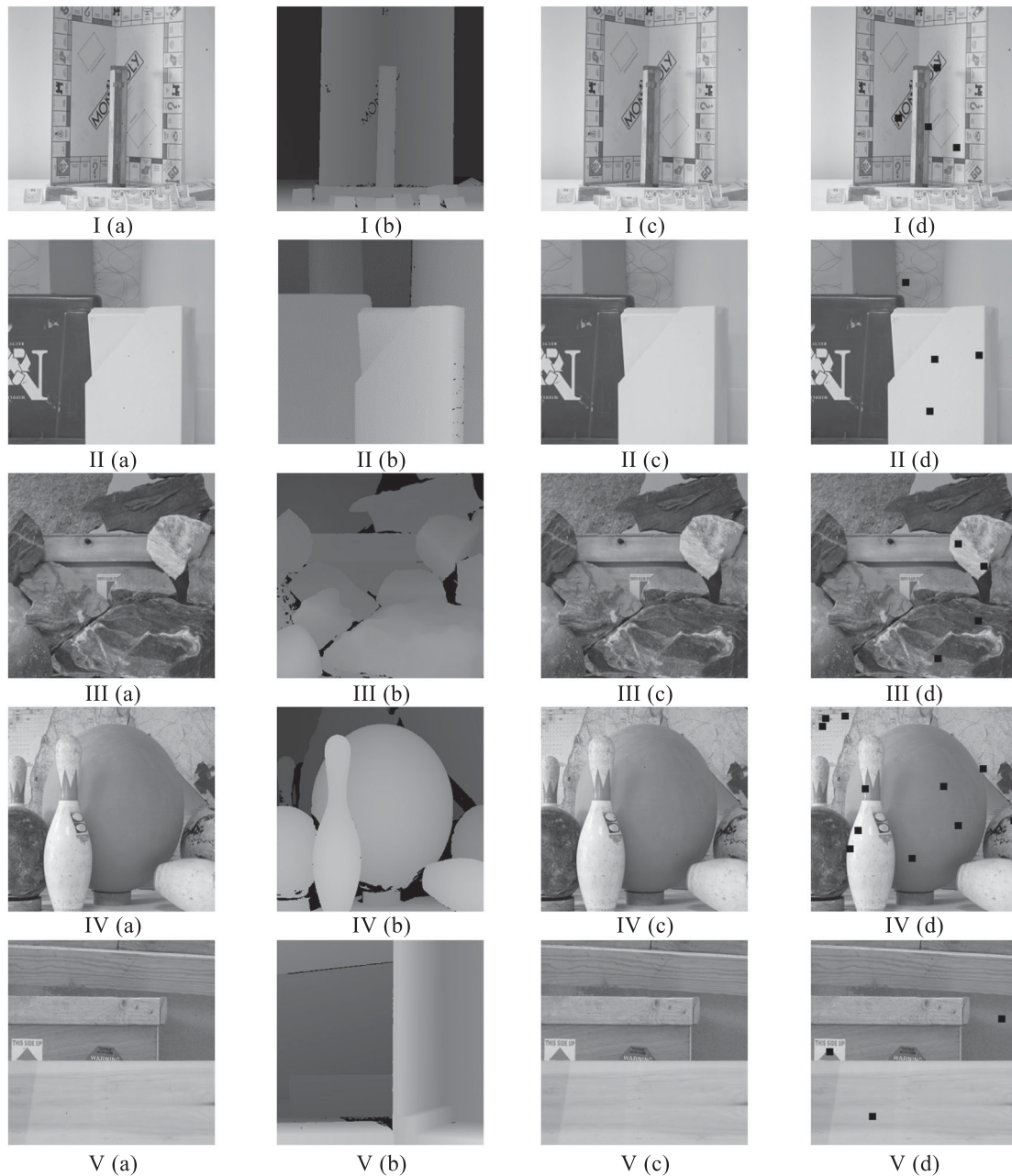


Figure 5. Performance of the proposed technique on Middlebury 2006 datasets [21] in terms of watermark embedding, its authentication and tamper localization. (a) Original left-eye image, (b) its associated depth map, (c) watermarked image and (d) its temper detected and temper-localized image for (I) Monopoly, (II) Plastic, (III) Rocks1, (IV) Bowling2 and (V) Wood2 images.

Table 2. Performance comparison of the proposed technique and Yu et al.'s technique [14] in terms of PSNR values.

Images	Aloe	Baby1	Bowling1	Cloth1	Flowerpots	Lampshade1
Proposed scheme	51.11	49.45	51.11	50.17	49.94	50.86
Yu et al.'s technique [14]	49.95	48.68	50.74	49.92	49.73	49.04

keep the watermark imperceptible and subsides visual attenuation of the images.

Figure 5 shows the performance of the VRFW in terms of watermark embedding in the left-eye image, its authentication and tamper localization. The results shown in Figure 5 demonstrate the subjective outcomes of the proposed scheme for Monopoly, Plastic, Rocks1, Bowling2 and Wood2 images through (I)–(V) of the Middlebury [21] database. Figure 5(a–d) represents the original image, its associated depth map, watermarked image and the temper-localized image, respectively. Addition and removal of artefacts attacks were mounted on the watermarked images to produce temper-localized images. It can be seen that there is no visual difference in the original and watermarked images. The black squared regions in the last column demonstrate the areas where slightest modifications are made in the watermarked image. Tempering attempts are first detected and then localized, which is very much desirable in real-world applications.

Table 2 presents the performance comparison of the proposed technique with the technique presented by Yu et al. [14] in terms of watermark imperceptibility. For this purpose, a number of images are selected from the Middlebury Stereo Vision Lab 2006 image datasets [21] and the PSNR values are computed between the original and the watermarked images generated by the two techniques. As shown from the table, the proposed technique performs better as compared with Yu et al.'s technique in terms of PSNR values. This is due to the fact that appropriate frequency bands are selected intelligently in the proposed technique utilizing a minimum number of coefficients for watermark embedding. Another reason for this enhanced performance is that the proposed technique finds the most suitable bits for watermark embedding. In addition, it uses only four bits to alter (and most of the time it does not need to alter) while embedding the watermark. This provides very little distortion even if the watermark is embedded. It is also visible from the SSIM and PSNR values in Table 1, which are very near to the original image and completely imperceptible to the human visual system. Through our experimentation in the laboratory, we found that there is no significant change in the quality of the watermarked 3D view.

One of the major drawbacks of the proposed scheme is its time complexity during the training phase. Since the proposed scheme is developed on simulation-based optimization using the GA, the time complexity is high during the training process. However, once a generalized suitable genome is generated, it takes a

fraction of a second to watermark the image. This is primarily due to the fact that the appropriate frequency bands are provided by the best-evolved genome for all the block locations in the DCT coefficients. The only computational cost then remains is following the rest of the steps in the algorithm for watermarking the image.

4. Conclusion

In this paper, a fragile watermarking technique is proposed for DIBR 3D-TV images with improved visual quality. The proposed VRFW scheme is based on GA and block DCT transform. Block-wise neighbourhood reliant watermarking scheme is based on neighbourhood aggregate and parity modulation. This creates strong block-wise dependence to enforce security against different types of attacks. As a result, the watermarked images obtained are not only better in visual quality but also the scheme is able to detect and localize tempering attempts for 3D-TV media.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Tariq Bashir  <http://orcid.org/0000-0002-8014-7816>

Syed Saud Naqvi  <http://orcid.org/0000-0002-6335-3538>

References

- [1] Sun W, Xu L, Au OC, et al. An overview of free view-point depth-image-based rendering. Proceedings of the second APSIPA annual summit and conference; 2010 Dec 14–17; Singapore: Biopolis. p. 1023–1030.
- [2] Zadokar SR, Raskar VB, Shinde SV. A digital watermarking for anaglyph 3D images. IEEE 2013 international conference on advances in computing, communications and informatics (ICACCI), Karnataka, India; 2013 Aug 22–25.
- [3] Singh D, Singh SK. DCT based efficient fragile watermarking scheme for image authentication and restoration. *Multimed Tools Appl.* 2017;76(1):953–977.
- [4] Usman I, Khan A, Chamlawi R, et al. Image authenticity and perceptual optimization via genetic algorithm and a dependence neighborhood. *Int J Appl Math Comput Sci.* 2007;4:615–620.
- [5] Barreto PS, Kim HY, Rijmen V. Toward secure public-key blockwise fragile authentication watermarking. *IEE Proc Vis Image Signal Process.* 2002;149(2):57–62.
- [6] Li CT, Si H. Wavelet-based fragile watermarking scheme for image authentication. *J Electron Imaging.* 2007;16:013009.

- [7] Dhole VS, Patil NN. Self-embedding fragile watermarking for image tampering detection and image recovery using self-recovery blocks. International conference on computing communication control and automation (ICCUBEA), Pune; 2015. p. 752–757.
- [8] Yeun MM, Mintzer F. An invisible watermarking technique for image verification. International conference on image processing, California, USA; 1997. p. 680–683.
- [9] Wu M, Liu B. Watermarking for image authentication. In: International conference on image processing. Chicago, Illinois, 1998; Vol. 2. p. 437–441.
- [10] Wong PW, Memon N. Secret and public key image watermarking schemes for image authentication and ownership verification. IEEE Trans Image Process. 2001;10:1593–1601.
- [11] Lee TY, Lin SD. Dual watermark for image tamper detection and recovery. Pattern Recognit. 2008;41:3497–3506.
- [12] Li C, Wang Y, Ma B, et al. A novel self-recovery fragile watermarking scheme based on dual-redundant-ring structure. Comput Electr Eng. 2011;37:927–940.
- [13] Qin C, Chang CC, Chen PY. Self-embedding fragile watermarking with restoration capability based on adaptive bit allocation mechanism. Signal Process. 2012;92:1137–1150.
- [14] Yu M, Wang J, Jiang G, et al. New fragile watermarking method for stereo image authentication with localization and recovery. AEU – Int J Electron Commun. 2015;69:361–370.
- [15] Edupuganti VG, Shih FY, Chang I-C. An efficient block-based fragile watermarking system for tamper localization and recovery. Intell Automa Soft Comput. 2011;17:257–267.
- [16] Nyeem H, Boles W, Boyd C. Counterfeiting attacks on block-wise dependent fragile watermarking schemes. Proceedings of the 6th international conference on security of information and networks, Aksaray, Turkey; 2013. p. 86–93.
- [17] Li CT. Digital fragile watermarking scheme for authentication of JPEG images. IEE Proc, Vis Image Signal Process. 2004;151:460–466.
- [18] Al-Haj A, Farfoura ME, Mohammad A. Transform-based watermarking of 3D depth-image-based-rendering images. Measurement. 2017;95:405–417.
- [19] Rana S, Mattu V, Sur A. SIFT based view invariant watermarking for DIBR 3D image. Fifth national conference on computer vision, pattern recognition, image processing and graphics (NCVPRIPG), Patna, India; 2015.
- [20] Lin YH, Wu JL. A digital blind watermarking for depth-image-based rendering 3D images. IEEE Trans Broadcast. 2011;57:602–611.
- [21] Scharstein D, Szeliski R. Middlebury stereo vision lab datasets. 3D Standard datasets. [cited 2017 Jan 12]. Available from: <http://vision.middlebury.edu/stereo/>.
- [22] Scharstein D, Pal C. Learning conditional random fields for stereo. IEEE computer society conference on computer vision and pattern recognition (CVPR 2007); 2007 June; Minneapolis, MN.
- [23] Hirschmüller H, Scharstein D. Evaluation of cost functions for stereo matching. IEEE computer society conference on computer vision and pattern recognition (CVPR 2007); 2007 June; Minneapolis, MN.