Journal of Theoretical and Applied Information Technology
30th April 2018. Vol.96. No 8

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

IMAGE TAMPER DETECTION AND RECOVERY USING LIFTING SCHEME-BASED FRAGILE WATERMARKING

[1]TAHA BASHEER TAHA, [2]RUZELITA NGADIRAN, [3]PHAKLEN EHKAN, , [4]MOHAMAD T. SULTAN

[1,2,3]School of Computer and Communication Engineering, University Malaysia Perlis, Perlis, Malaysia

[4]Department of Computer Science, Cihan University-Erbil, Erbil, Iraq

ABSTRACT

High prevalence of digital images imposed a great interest in the process of authority and integrity investigation. In many cases, high accuracy image tamper detection techniques involves with high computational overhead in addition to the complexity produced from recovering original image process. In this work, blind image tamper detection and self-recovery is presented using Lifting Scheme which characterized by simplicity and integer based calculations and LSB modification. Experimental results reveal that proposed model performs well in terms of detection and recovery for different types of tampering as removing and cloning. Furthermore, produced watermarked images have very accepted perceptual quality in terms of subjective and objective evaluations.

Keywords: *Lifting Scheme, LWT, Image Tamper Detection, Image Recovery, Fragile Watermarking.*

1. INTRODUCTION

Due to the wide spread of digital images in today's digital age, sometimes seeing is no longer believing, since tampering these digital images is easy and simple by using available image editing tools. Many tampered images emerge in news items, scientific experiments and even legal evidences as criminal investigation or road accident imaging. Therefore, authenticities of images can not be granted until verifying the integrity of these image from tampering [1]. So, distinguishing the original images from faked ones and establishing the authenticity of digital photographs have become a serious challenges in present time [2]. A fragile watermark embedding is a feasible method for solving the problem of image authentication and tamper detection. A fragile watermark that may be a logo or meta data, is embedded into a cover image, and by comparing the extracted watermark with original one tampering can be detected [3]. In addition to tampering detection, it is usually important to restore original data from the tampered one to determine the intent of alteration.

It is recommended that tamper detection algorithm can satisfy all or most of the following requirements:

(1) Tampering detection: The scheme should raise an alarm if the image has been altered.

(2) Imperceptibility: Any modification on the image caused by watermark embedding should be imperceptible to human vision system (HVS) or affect perceptual quality of original image.

(3) Locating tempered areas: The scheme should locate the tempered areas or identify the type of attack.

(4) Self recovery: The original data is being recovered (Image can restore it self).

(5) Blind detection: Original image is not required in tamper detection process.

(6) Efficiency: The algorithm processing complexity should be reduced to minimum.

The proposed system is designed to meet mentioned requirements by using a low computational fragile watermarking algorithm that depends on Lifting Wavelet Transform (LWT) presented by Sweldens [4]. Some LWT coefficients in second approximation decomposition are modified by changing the value of the Least Significant Bit (LSB) by watermark pixel to embed a fragile watermark that can be blindly detected for tampering detection and locating forged areas in digital images. Original pixels can be reconstructed after tampering as they were re-embedded from approximation band into one of the LSBs of middle frequency bands. To evaluate the proposed method,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

watermarked images have been tampered by covering blocks, cloning attack, and removing attack. Tampered areas were successfully detected and original pixels were mostly recovered. The usage of LWT instead of DWT or DCT simplified the processing duty for its simple calculation and integer to integer transformation. As a result, better computational efficiency is produced that make the proposed system suitable for embedded systems.

For perceptual quality evaluation, two metric are used, Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index (SSIM) [5]. Watermarked images has PSNR values ranging from 28 to 36 dB and SSIM values ranging from 0.96 – 0.99. As objective evaluation, watermarked images have very accepted perceptual quality.

Rest of the paper is arranged as follow, a literature is presented in section two, section three is an introduction to lifting wavelet transform, section four is the proposed method, and experimental results is shown in section five. Finally, the work is concluded in section six.

2. RELATED LITERATURE

Image processing schemes that aim to detect image variation can be implemented in time (spatial) domain or frequency (transform) domain.

A low complexity fragile watermarking scheme was proposed by Lin et al. [6] using a multi-level authentication scheme. The operations which are used were simple as parity check and comparison between average intensities. The hierarchical structure of the design increase the accuracy of tamper localization, that is, once a tampered sub-block is detected, the other three sub-blocks within the same block are recognized as being tampered as well. Seungwu Han et.al [7] developed a spatial domain method where the original image is divided into blocks, and the features of each block is hidden in the same block. In extraction, the hidden features are extracted and compared to detect the existence of tampering.

Chang et al. [8] proposed a fragile watermarking scheme in which the authentication data of an image, as the eight pixel values surrounding the center of the block, the block number, the user secret key, and user ID is inserted into adaptive least significant bits of the embedded pixels. And the number of the least significant bits was determined by corresponding block type. When authenticating image integrity, first authenticated information is retrieved from the least significant bits of each center block. Then, a new authenticated message of the test image is generated in the same manner. By comparing the embedded authentication message and the newly generated message, tampered blocks can be identified.

Another attempt is presented by Ching-Sheng Hsu a, Shu-Fen Tu [3]. This study aims to integrate probability to the proposed scheme to improve image tampering detection accuracy and precision. The scheme includes two processes: the embedding of an image authentication message and tampering detection.

As frequency domain based tamper detection attempt, Qi et.al utilized DWT to embed the image features in the approximation subband to high frequencies bands, however, recovery process is not included[9]. Sarika et.al employed both wavelet and singular value decomposition (SVD) to detect image tampering. In wavelet, both transmitted and extracted are compared to check whether the image is altered or not, while the SVD is used to recover the original image's content[10]. SVD is also used by Tafti and Hassannia [11] where the statistical information from original image's Lower Upper and SVD is computed with the addition of cellular automata. This combination is used to construct a cipher key that is unique for the host image and will be changed if any tampering occurs.

Watermark existing is considered as a crucial factor to investigate authentication [12] and integrity, as a consequence, it is involved in many tamper detection attempts in literature where different algorithms are proposed for detecting image forgery and self-recovery. A number of surveys are presented to summarize such attempts as [13] and [14].

Generally, mentioned time domain based tampering attempts lack to self-recovery process, however, attempts alike are characterized by simplicity and low computation overhead. While detecting forgery and recover original data in transform domain used to have complicated operations as combining both DWT and SVD in [10]. In this paper, we presents a fragile watermarking scheme for image tamper detection and recovery. The feature of proposed scheme is the ability to detect and recover the image by depending on basic and simple calculations that consist of LWT and LSB embedding. As a result, the simplicity of time domain attempts and high performance of frequency domain attempts in detecting alteration and self-recovery are obtained.

Journal of Theoretical and Applied Information Technology
30th April 2018. Vol.96. No 8

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3. THE LIFTING SCHEME

Lifting scheme or Lifting Wavelet Transform (LWT) is introduced by Swedens [4] to construct wavelet coefficient without using Fourier transform. It utilizes the spatial domain to produce wavelets which provide low computational process in compare with filter banks used in regular DWT. Furthermore, the Lifting scheme has integer to integer computation which makes it suitable to be mapped to hardware design where the floating points are not often recommended.

The core idea beyond lifting scheme is to divide the original signal into approximation and details coefficients. This can be done via three major steps [15] include:

i. Split

It also called the lazy wavelet, where the signal is being separated into odd and even samples.

ii. Predict

In this step the difference between the prediction of a signal and the real signal is stored. In other words, even samples are used to predict the odd ones, and the difference between them is stored in the location of odd samples. This result produces the details signal (edges of the image), because by default, image values are tended to have linear changing, and the predicted odd value is the average of its two adjacent even samples. The difference tends to be zero in linear intensity change but it will have a higher value where an abrupt transition in pixels' values occurred as the border of black jacket in front of white background. Predict step can be written as (1)

$$Y2n+1=Xo- P(Xe) \quad (1)$$

iii. Update

When the average of each two pixels in a signal are calculated, that means the overall structure of that signal is obtained with half energy [16]. However, due to the non- linearity changing in image pixels, the average (even samples interposed the odd ones) cannot be taken directly and it needs to be "Updated" with the differences computed in predict steps. Update step can be written as (2).

$$Y2n=Xe + U (Y2n+1) \quad (2)$$

In proposed method the simplest form of Lifting scheme is used, which is cdf (2,2), a member of Cohen-Daubachies and Feauveau wavelets with natural coefficient that can be implemented using basic shift operators[15]. LWT, as discrete wavelet transform can be applied in more than one level, by repeating the same steps on approximation coefficients produced by update step to get more compressed version of original signal. The inverse of lifting scheme can easily be implemented by reversing the order and exchanging the sign of the predict and update steps. Figs. 1 and 2 show the forward and reverse LWT respectively.

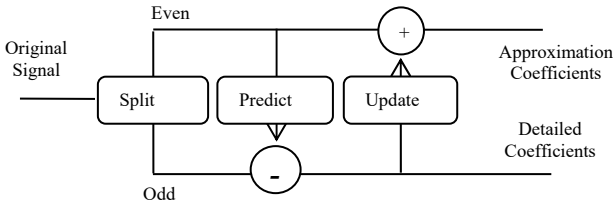

Figure 1: Forward LWT

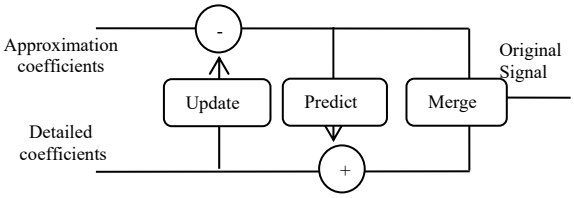

Figure 2: Inverse LWT

By applying one LWT decomposition on digital image, four bands will be created, one for approximation coefficients as low frequency band, two mid frequency bands and other for details as high frequency band. In proposed method, another LWT decomposition is applied on the approximation band to get more abstracted approximation coefficients that will be employed in proposed method (Figure 3).

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4. PROPOSED METHOD

Proposed method is described using several steps, first is approximation band distribution, which is used to prepare the image for self-recovery in case of tampering, second is watermark embedding process, third is watermark extraction, and fourth is the recovery process. In addition, an analysis had been made to find the suitable band for watermark embedding in last sub section.

A. Approximation band distribution:

Two level of LWT is applied on host image of size 512 × 512 to produce seven frequency bands as shown in Figure 3. Approximation band (SS2) of size 128 × 128 has the most important features of the image. So, each bit of SS2 coefficients will be embedded in one of D1 and S2 LSBs, hence, each band will carry 4 bits from SS2. These data will be used as reference to restore original data in case of tampering.

The usage of D1 and S2 have been chosen since middle frequency bands have the suitable robustness to carry the original information as well as high imperceptibility that has no large impact on image quality.

B. Embedding Process

Figure 4 depicts the embedding process that is summarized in following steps:

1- Watermark is automatically generated as set of zeroes and ones with the size that equal to band size after second decomposition (128 * 128 in this case). The reason beyond using auto generated watermark is to control the size of the watermark according to host image size, furthermore, there is no need for extra storage to save external watermark.

2-After LWT decomposition, Each coefficient of approximation band DD3 is modified by changing one of its LSBs by the value of equivalent binary watermark bit.

3- Finally, Inverse LWT is applied to reconstruct the watermarked image that carries the fragile watermark.

C. Watermark Extraction:

Receiver side aims to verify watermarked image authentication. LWT is applied on the watermarked image and LSBs are extracted from DD3 coefficients. Watermark bits are re-generated in the same way of embedding process and compared with the extracted watermark. In case of mismatching, each location of tampered pixels is stored and marked to be recovered in recover phase (Figure 5).

D. Recovery Process:

Approximation band coefficients that are embedded inside middle frequency bands will be restored in case of mismatching between generated watermark and extracted one. The recovery process is done by collecting the LSBs from all D1 and S2 coefficients, since they have the data of original SS2. And altered SS2 coefficients will be replaced by the collected values. After applying ILWT, tampered location will be restored.

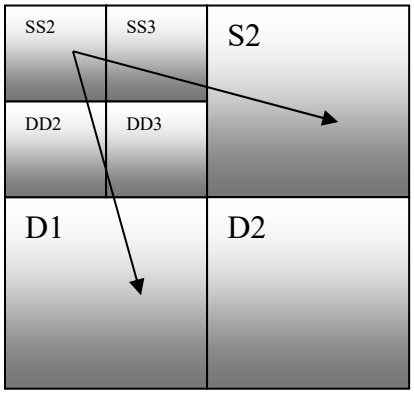

Figure 3: Two LWT decomposition bands

Journal of Theoretical and Applied Information Technology
30th April 2018. Vol.96. No 8

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

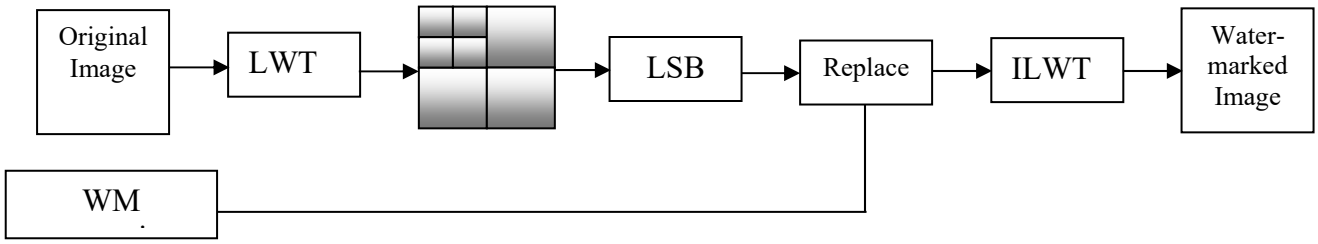

Figure 4: Embedding Process

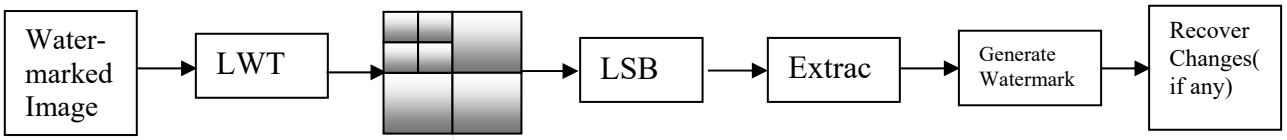

Figure 5: Extracting Process

E. Watermark embedding Bands

While embedding in LWT coefficients, it is important to point that the watermark must still exist after applying the reverse process of LWT (ILWT), since LWT is an integer to integer transform, "floor" operation is used [17], which may drop the external watermark bit. Although the LWT is lossless algorithm, original signals coefficients are modified by embedded watermark bits. So in our experiment, watermark is embedded in LSB of the four band of second decomposition and the false positive (possibly misjudging the untampered pixels as tampered ones [3]) is measured.

Five images where used for testing, thumbnails are shown in figure 6. Table 1 shows the number of false positive errors in each image for bands (SS2,SS3,DD2,DD3). We find that the best band to embed the watermark is in SS2, however SS2 is the approximation band that is distributed to other band for recovery process, hence DD3 is chosen for watermark embedding. 7th LSB is used for embedding since it has better false positive ratio than 8th one, and has approximately the same impact on image in compare with 8th bit.

Figure 6: Thumbnail of tested images, Images are numbered as I1-I5 left to right.

Journal of Theoretical and Applied Information Technology
30th April 2018. Vol.96. No 8

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 1: Number of false positive bits in different bands

	Number of False Positive bits			
IMAGE	SS2	SS3	DD2	DD3
I1	8	71	75	27
I2	0	0	0	0
I3	0	0	0	0
I4	0	5	3	2
I5	0	4	8	1
Total	8	80	86	30

5. EXPERIMAENTAL RESULTS

Watermarked image quality and ability to detect and recover tampering in different situations is presented in this section along with a summary of features of the proposed method.

A. *Watermarked Image Quality*

Peak signal to noise ratio (PSNR) is considered as a simple objective pixel-based methods in image quality evaluation. PSNR is logarithmic transformation of MSE. PSNR equation is given as:

$$\mathrm{PSNR} = 20\log_{10}\left(\frac{\mathrm{MAX}}{\mathrm{MSE}}\right) \quad (3)$$

And

$$\mathrm{MSE} = \frac{1}{m*n} + \sum_{x=0}^{m-1}\sum_{y=0}^{n-1} \| I(x,y) - I'(x,y) \|^2 \quad (4)$$

Where m, n are image dimensions and I and I' are original and target image respectively.

In addition to PSNR, structural similarity index (SSIM) is used as more reliable measurements [5]. The equation of SSIM is given by:

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2+\mu_y^2 + C_1)(\sigma_x^2+\sigma_y^2 + C_2)} \quad (5)$$

Where x and y are two non-negative image signals. μx, μy, are the mean intensity, σx, σy are the standard deviations for the original and distorted images respectively, C1, and C2 are constants.

Quality of watermarked images is measured using these metrics. As we observe from Table 2 SSIM and PSNR values are high. For objective evaluation, watermarked images has no visual distortion or any sign of watermark artifact. As shown in Figure 7.

Table 2: PSNR AND SSIM for different images

IMAGE	PSNR	SSIM
I1	33.8607	0.9871
I2	28.3141	0.9649
I3	35.1795	0.9886
I4	36.0161	0.9914
I5	29.6412	0.9803

Figure 7: Thumbnail of some watermarked images

B. *Tamper Detection*

Original watermark that is generated is set of zeroes and ones appears like tiny check board as shown in Figure 8. After two levels of decomposition, the generated watermark is compared with the extracted watermark from DD3

Journal of Theoretical and Applied Information Technology
30th April 2018. Vol.96. No 8

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

band. In case of mismatching, the location of corresponding pixel is located to indicate the tampered regions that need to be restored. Different tampering attacks have been tested and detected as listed in following sub sections.

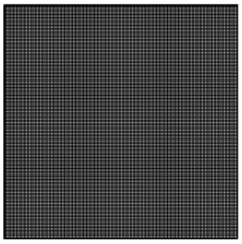

Figure 8: Generated Watermark

i. Clone detection:

Proposed algorithm is tested against clone attack, Cameraman image is watermarked and altered by cloning the two buildings. Extracted watermark is affected where the buildings are cloned, cloning areas are successfully recovered on the restored image. (Figure 9).

a

Figure 9 Results after cloning objects. a :Tampered Cameraman image. b: Recovered Image.

ii. Remove detection:

Two buildings in watermarked Cameraman image are removed, after applying proposed detection and recovery algorithm, the most of buildings pixels were successfully restored (Figure 10).

a

b

Figure 10 Results after removing objects. A :Tampered Cameraman image. b: Recovered Image.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

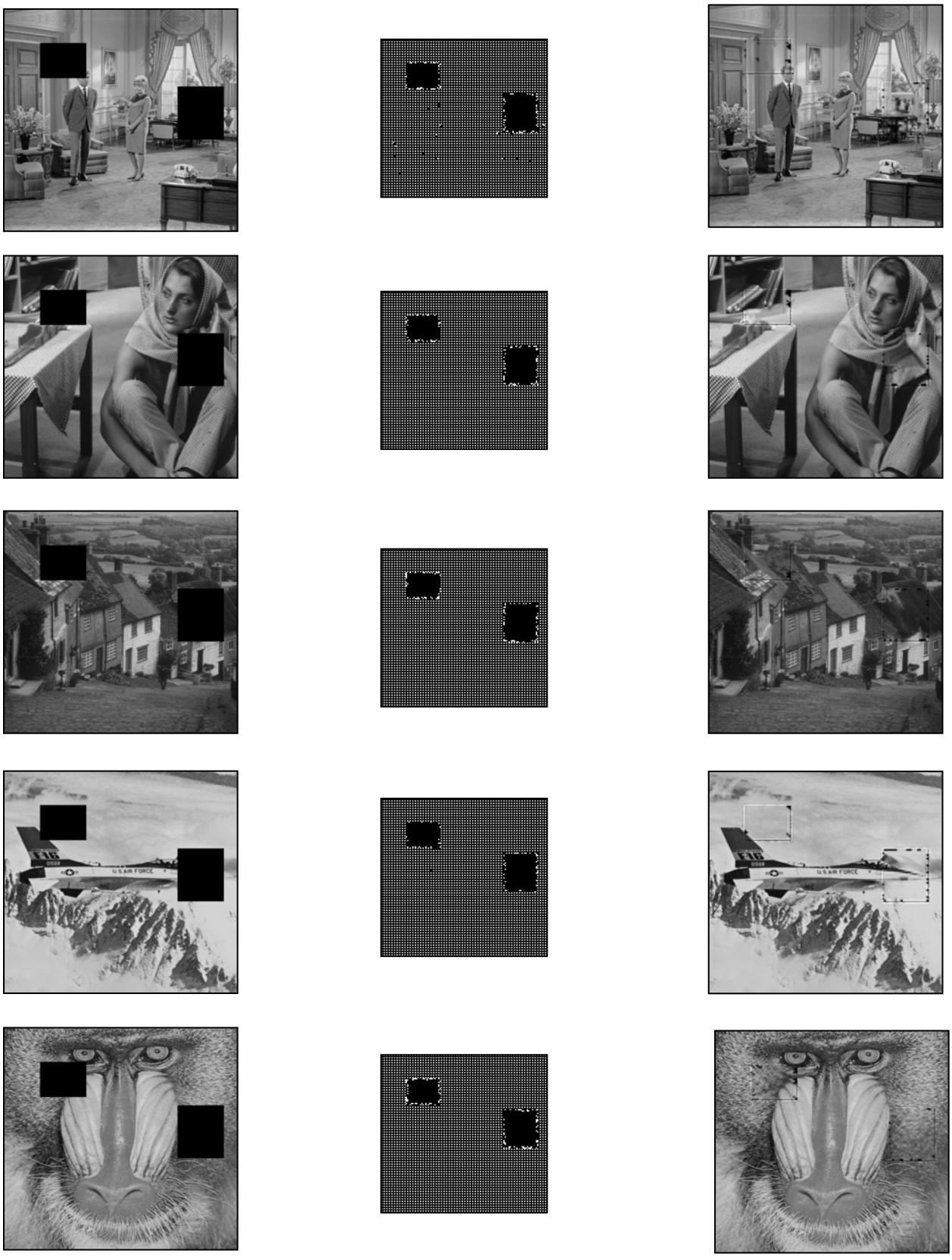

Figure 11: Different tampered images (Left), Extracted watermark (Middle), Restored images (Right)

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

iv. **Block tampering:**

Tested images are partially covered by blocks. As shown in Figure 11, proposed algorithm was able to reconstruct original data after detecting the tampered places on the watermark. Results shows that the algorithm detects the place of embedding, and most of the altered pixels are restored correctly although the blocking was completely hiding the covered areas. Which indicated that proposed algorithm has high performance in tamper detection and recovery in addition to its simplicity.

C. General Results Discussion

Performance of the proposed system is analyzed in terms of memory utilization, simplicity, visual quality, and accuracy.

- **Memory Utilization:**

Since the detecting algorithm has blind extraction feature no original image is needed for extracting the watermark or recovery. which reduce the required memory. Also, the watermark is automatically generated in the system so no watermark image to be stored. As a result the proposed watermark has efficient memory utilization.

- **Simplicity:**

Proposed algorithm depended on LWT, which is integer to integer calculation process and its equations for embedding and extracting are simple and depends on dividing by numbers two and four which can be implemented by logical shift operators. Also, proposed algorithm based on LSB changing which make the algorithm applicable to be implemented on embedded systems.

- **Visual Quality:**

Watermarked images have high perceptual quality in terms of PSNR or SSIM and for different utilized images. As objective evaluation, no watermark artifacts or image degradation can be noticed.

- **Accuracy:**

Proposed algorith compares between generated watermark and extracted one. Since the watermark is a binary number, the value might not be chaged although there is an alteration. However, creating a dotted watermark as adjacent zeros and ones reduces this issue. Generally speaking, proposed algorithm characterized by simplicity in implementation and good accuracy of recovering the original image using new method of LWT approximate band distribution. However, false positive errors initiating by ILWT may be enhanced in future.

6. CONCLUSION

A new fragile watermarking algorithm for blind image taper detection and recovery is presented in this paper. The algorithm designed to have simple computations and integer-based operations and combines the feature of time domain based simplicity and transform domain performance. that suitable to be implemented in limited processor applications. Using a combination of LSB modification and LWT produced a watermarking system that can detect and locate tampering . And by using the proposed LWT approximation band distribution method, tampered areas were successfully restored after cloning, removing and blocking attacks. Simplicity in implementation and high performance combines the features of time domain and frequency domain tampering detection techniques together within one system. In addition, produced watermarked images has high perceptual quality in terms of subjective and objective evaluations.

ACKNOWLEDGMENT:

The author would like to acknowledge the support from the Fundamental Research Grant Scheme (FRGS) under a grant number of **FRGS/1/2017/ICT05/UNIMAP/02/2** from the **Ministry of Higher Education Malaysia**.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

REFERENCES:

[1] Wang, W., Dong, J., & Tan, T. (2009, August). A Survey of Passive Image Tampering Detection. In *IWDW* (Vol. 9, pp. 308-322).

[2] Mishra, M., & Adhikary, F. (2013). Digital image tamper detection techniques-a comprehensive study. *arXiv preprint arXiv:1306.6737*.

[3] Hsu, C. S., & Tu, S. F. (2010). Probability-based tampering detection scheme for digital images. *Optics Communications*, *283*(9), 1737-1743.

[4] Sweldens, W. (1995, September). Lifting scheme: a new philosophy in biorthogonal wavelet constructions. In SPIE's 1995 International Symposium on Optical Science, Engineering, and Instrumentation (pp. 68-79). International Society for Optics and Photonics.

[5] Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. Image Processing, IEEE Transactions on, 13(4), 600-612.

[6] Lin, P. L., Hsieh, C. K., & Huang, P. W. (2005). A hierarchical digital watermarking method for image tamper detection and recovery. *Pattern recognition*, *38*(12), 2519-2529.

[7] Han, S., Jin, H. L., Fujiyoshi, M., & Kiya, H. (2006, December). Lossless data hiding in the spatial domain for image tamper detection. In Intelligent Signal Processing and Communications, 2006. ISPACS'06. International Symposium on (pp. 760-763). IEEE.

[8] Chang, C. C., Hu, Y. S., & Lu, T. C. (2006). A watermarking-based image ownership and tampering authentication scheme. *Pattern Recognition Letters*, *27*(5), 439-446.

[9] Qi, X., Xin, X., & Chang, R. (2009, November). Image authentication and tamper detection using two complementary watermarks. In *Image Processing (ICIP), 2009 16th IEEE International Conference on* (pp. 4257-4260). IEEE.

[10] Bhosale, S., Thube, G., Jangam, P., & Borse, R. (2012, August). Employing SVD and Wavelets for Digital Image Forensics and Tampering Detection. In Advances in Mobile Network, Communication and its Applications (MNCAPPS), 2012 International Conference on (pp. 135-138). IEEE.

[11] Tafti, A. P., & Hassannia, H. (2014). Active Image Forgery Detection Using Cellular Automata. In Cellular Automata in Image Processing and Geometry (pp. 127-145). Springer International Publishing.

[12] Liew, S. C., & Zain, J. M. (2011). Tamper localization and lossless recovery watermarking scheme. In Software Engineering and Computer Systems (pp. 555-566). Springer Berlin Heidelberg.

[13] Arun Anoop, M. (2015, March). Image forgery and its detection: A survey. In Innovations in Information, Embedded and Communication Systems (ICIIECS), 2015 International Conference on (pp. 1-9). IEEE.

[14] Qazi, T., Hayat, K., Khan, S. U., Madani, S. A., Khan, I. A., Kołodziej, J., ... & Xu, C. Z. (2013). Survey on blind image forgery detection. Image Processing, IET, 7(7), 660-670.

[15] Gholipour, M. (2011). Design and implementation of lifting based integer wavelet transform for image compression applications. In Digital Information and Communication Technology and Its Applications (pp. 161-172). Springer Berlin Heidelberg.

[16] Jensen, A., & la Cour-Harbo, A. (2001). The Discrete Wavelet Transform via Lifting. In Ripples in Mathematics (pp. 11-24). Springer Berlin Heidelberg.

[17] Angelopoulou, M. E., Masselos, K., Cheung, P. Y., & Andreopoulos, Y. (2008). Implementation and comparison of the 5/3 lifting 2D discrete wavelet transform computation schedules on FPGAs. Journal of signal processing systems, 51(1), 3-21.