BAYERO JOURNAL OF ENGINEERING AND TECHNOLOGY (BJET) VOL.16 No.2, MAY, 2021 PP 89-95 ISSN: 2449 – 0539

DEVELOPMENT OF AN IMPROVED AUDIO WATERMARKING SCHEME USING STATIONARY WAVELET TRANSFORM AND SINGULAR VALUE DECOMPOSITION TECHNIQUE

Muzakkar S[#1] ,S. M. Sani[#2] and Tekanyi A. M. S. [#3]

[#]*Department of Electronics and Telecommunications Engineering, Ahmadu Bello University Zaria, Nigeria*
[#1]muzakkars@yahoo.com[#2] smsani@abu.edu.ng [#3] amtekanyi@abu.edu.ng

ABSTRACT

In this era, the widespread use of internet and wireless communication has made replication and modification of multimedia data easier than ever, due the fact that these multimedia data are now stored in digital form. This necessitates the needs for the protection of these multimedia data against illegal piracy. This paper developed the scheme of SWT-SVD technique to improve robustness and inaudibly of audio watermark. The host audio signal was partitioned into non-overlapping frames of equal size and number of frames was equivalent to the number of watermarking bits. The scheme was subjected to various attacks with a view of evaluating its performance on different audios namely: classical, country, folk, jazz, metal, pop and rock music respectively. The simulation was done on MATLAB 2018a using Signal Processing and Communications Toolbox. The results obtained from simulations showed that the developed watermarking scheme when compared to existing scheme was found to be better in terms of robustness and inaudibility. Classical music recorded an improvement in terms of inaudibility with Signal to Noise Ratio (SNR) value of 9.43% over existing watermarking scheme. Additionally, developed watermarking scheme achieved improvement in terms of robustness with Normalized Cross Correlation (NCC) value of 1.52% and reduction in Bit Error Rate (BER) with average value of 7.89% relative to existing watermarking scheme.

Keywords: Inaudibility, Robustness, Watermark, BER, SNR

1. INTRODUCTION

The widespread use of Internet and wireless networking has made the distribution of multimedia data much easier than ever. The fact that multimedia data are now stored in digital form, thus making replication and modification of such media so easy by users (Hu and Hsu, 2015). This necessitates the needs for digital multimedia data protection against illegal copying, distribution, and misuse of digital media. Digital watermarking is the technology to embed or extract independent source of signal in the watermarked carrier (Qian and Xianghong 2016). Watermarking provides copyright protection, monitoring copies of the content, tamper resistance, authentication, and annotation (Jelena *et al.,* 2014). A good watermarking scheme should have some important traits such as imperceptibility, robustness, security, and embedding capacity (Kulkarni *et al.,* 2014). Digital watermarking aims to conceal proprietary data into the media object. When needs arises, owner can extract these data and declare copyright. (Qian and Xianghong 2016).

1.1 Review of Related Literature

The following presents a review of similar research works undertaken by some researchers that are relevant to this research area. Sakriti, (2015). implemented digital audio watermarking using Discrete Wavelet Transform (DWT) and Discrete HAAR Transform (DHT) on music signal, although the work satisfactorily improved robustness, capacity, and imperceptibility when compared to Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT). But this algorithm can only be applied on audio files with wav formats; therefore, it cannot work on audio files with Advance Audio coding (AAC) or mp3 formats. Hazem *et al.,* (2016) reported on a new method for digital watermarking based on two dimensional cellular automata, in this work, cellular automata were used to generate key and embedded the watermark. The experimental result demonstrated that the watermark was imperceptible and achieved greater similarity between watermarked and original audio. But the algorithm cannot withstand compression, resampling and cropping attacks when audio files with multi-channel is employed. Devasis, (2017) implemented an invisible digital audio watermarking using DWT-DCT. In the work, gray scale image was scrambled by Arnold Scrambling algorithm and used it as watermark on the host audio. Cyclic codes were then watermark image was little bit degraded after being removed from the watermarked audio, hence the algorithm did not satisfy the condition of perfect reconstruction. Shijun *et al.,* (2018) presented a reversible audio watermarking algorithm by modifying the statistical feature in time domain. In the work, the histogram of the statistical features of the algorithm was shifted for data hiding to embed watermark bits in to the frames. This algorithm was robust and reversible to common signal processing operations. But the advert effect of frames length on the robustness of the technique against MP3 compression was significant. Arashdeep and Malay, (2018) worked on High Embedding Capacity and Robust Audio Watermarking for Secure Transmission using Tamper Detection, the algorithm was robust against intentional and unintentional attacks. However, there is need for further improvement in the aspect of reducing Bit Error Rate in order to achieve better robustness and inaudibility. Chen *et al.,* (2019) reported on digital audio watermarking scheme based on quantization embedding system in the wavelet domain, In the work, amplitude quantization embedding rules along with synchronization mechanism was used to embed and extract the watermark. This algorithm achieved improved robustness in terms of resampling, low pass filtering and MP3 compression attacks. However, when fixed parameters are employed, it was difficult to combine sound quality with resistance to compression and attacks. In view of the reviewed literature, From the work of Arashdeep and Malay, (2018) there is need for further improvement in the aspect of reducing Bit Error Rate in order to achieve better robustness and inaudibility.

2. METHODOLGY

Lena digital image (form Standard Image database) was used as watermark on seven different audio samples of different genre were used as host audio, they are: Classical, Country, Folk, Jazz, Metal, Pop and Rock music's respectively.

2.1 Stationary Wavelet Transform (SWT)

This technique was first introduced to reestablish the time (or translation) invariance property that was lost in the DWT (Morsi *et al.,* 2008). The relationship of the SWT approximation and detail coefficient, $c_{j,k}$,$d_{j,k}$ at any decomposition level, j and sample index, k of the original signal, $x(k)$ through a low and high pass filter with arbitrary integer, m (Morsi *et al.,* 2008) are represented by the following expressions:

$$c_{j,k} = \sum_m c_{j-1,k+2^{j-1}m}\ g(m) \quad (1)$$

$$d_{j,k} = \sum_m c_{j-1,k+2^{j-1}m}\ h(m) \quad (2)$$

2.2 Singular Value Decomposition (SVD)

SVD is a numerical analysis tool that effectively decomposes a matrix in to three matrices of the same size as original matrix. Both one way and non-symmetric properties are preserved in this algorithm, which are not obtainable in other transformation techniques such as DFT, DWT and DCT. An arbitrary matrix, A of size $m \times n$ U and $V$are unitary matrices of size $m \times m$ and $n \times n$ respectively, and S is a diagonal matrix of size $m \times n$, which has non zero singular values as its diagonal components of SVD are shown in equation 3 (Wei C. *et al.,* 2009).

$$A = USV^T \quad (3)$$

The eigenvalue decomposition of $A\,A^T$ and A^TA can be used to evaluate the unitary matrices U and V. additionally, S can be evaluated by taking the square root of the eigenvalues of either $A\,A^T$ or A^TA (Min-jae, 2016) as expressed in the following equations:

$$A\,A^T = US^2U^T \quad (4)$$

$$A^TA = VS^2V^T \quad (5)$$

2.3 Implementation Principle

This algorithm was implemented by hybridization of stationary wavelet transform and singular value decomposition. All the embedding and extraction methods are implemented on MATLAB R2018a.

a. Embedding process

i. Framing and Segmenting input host Audio signal into non-overlapping frames.

ii. Conversion and splitting watermark image in to row and column

iii. Decomposing the Frames using SWT

iv. Embedding the watermark into the transformed frame SVD embedding technique.

v. Perform inverse 3rd- level SWT to obtain the watermarked Audio.

b. Extraction process

i. Framing and Segmenting the watermarked audio into non-overlapping frames

ii. Decomposing the watermarked frames using SWT

iii. 3D matrix conversion on the frames using SVD

iv. Apply SWT invers to the cover frames

v. Decrypt extracted bits and recover the watermark.

The original watermark in Figure 1 was scrambled into two-dimensional binary sequence of size 64 × 64 as depicted in Figure 2.

Figure 1: Watermark image

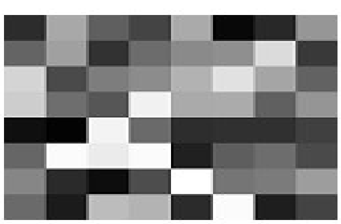


Figure 2: $n \times n$ Binary Image

Host Audio signal was segmented and then converted using equation (1) and (2) into 168 non-overlapping frames. The input audio used consisted of 44100 samples per frames, the quantity of frames of an audio signal are equal to the number of bits to be watermarked. Equation 4 and 5 were used to further decomposed the frames. In order to obtain the watermarked Audio, a 3-level inverse SWT was applied to the frames. The Inverse SWT cover frame was achieved using the mathematical expression in equation 6.

$$c_{j-1} = \bar{R}_{[j]}\left(\sum_m c_{j-1,k+2^{j-1}m}\, g(m) + \sum_m c_{j-1,k+2^{j-1}m}\, h(m)\right) \qquad (6)$$

Figure 3 shows the host audio signal before embedding the watermark into it, while Figure 4 illustrates the watermarked audio signal after the watermark was embedding, this shows that the developed algorithm achieved high similarity between the host and watermarked audios respectively.

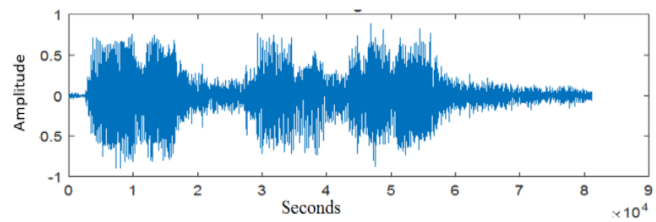


Figure 3: Host Audio

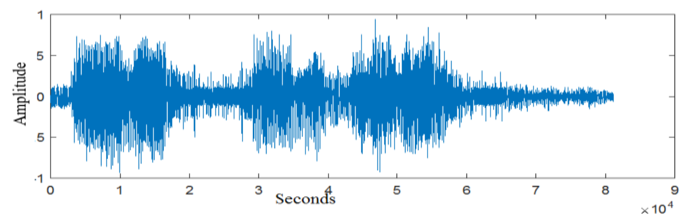


Figure 4: Watermarked audio

2.4 Application of Attacks

In this research work, different values were assigned for the implementation of those i.e 1, 2, 3 4 and 6 for AWGN, re-sampling re-quantization and cropping1 attack respectively, when the assigned value is switched in the algorithm, the particular attack is applied as shown in Table 1.

Table 1: Attack Description

S/N	Attack	Description	Label
1	AGWN	SNR 30dB	1
2	Resampling	44.1->22.05>44.1	2
3	Re-quantization	16->8->16	3
4	Cropping 1	25%	3

2.5 Performance Metrics

2.5.1Signal-to-Noise Ratio (SNR)

SNR refers to the logarithmic ratio between original audio signal and distorted watermark signal, usually measured in decibel. SNR value is computed according to equation 7 (Al-Haj, 2014).

$$SNR\ (dB) = 10log_{10}\ \frac{\sum_n A_n^2}{\sum_n (A_n - A_n')^2} \qquad (7)$$

where:

A corresponds to the original signal.
A' corresponds to the watermarked signal.
n is an arbitrary number.
$\sum$ denotes the summing operator.

2.5.2 Bit Error Rate (BER)

The BER is the error bits occurring per second between the embedded watermark and extracted watermark (Iqbal *et al.*, 2014).

$$BER = \frac{\sum_{i=1}^{M} \sum_{j=1}^{M} A_O(i,j) \oplus A_E\,(i,j)}{M \times M} \qquad (8)$$

where:

A_W and A_E are the original and extracted the watermarked bits.
$M \times M$ is the watermark size in bits.
i and j are the indexes of the binary.
$\oplus$ is an exclusive OR operator.

2.5.3Normalized Cross Correlation (NCC)

NCC is used to evaluate the similarity between the extracted and the original watermark. NCC close to 1 represents very high correlation and NCC close to 0 represents very low correlation (Tsai, (2015).

$$\text{NCC}\ (\mathbf{I}, \mathbf{I'}) = \frac{\sum_{i=1}^{X} \sum_{j=1}^{Y} i(i\,,j)\, i'(i\,,j)}{\sqrt{\sum_{i=1}^{X} \sum_{j=1}^{Y} i^2(i\,,j)}\ \sqrt{\sum_{i=1}^{X} \sum_{j=1}^{Y} i'^2\,(i\,,j)}} \qquad (9)$$

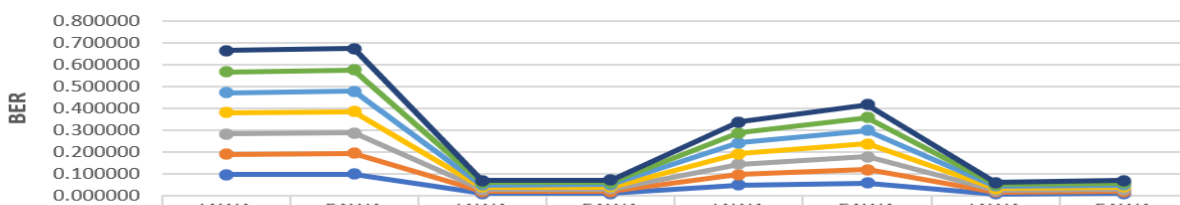
where:

X and Y is size of binary image.
I and I' denote the original watermark and the extracted watermark respectively.
$i$and j are the indexes of the binary.

3 RESULTS AND DISCUSSIONS

3.1 Results of the Application of Attacks on the Watermarked Audios

This subsection discussed the results and analysis obtained from improved audio watermarking algorithm (IAWA) and that of existing audio watermarking algorithm (EAWA) when subjected to different attacks. For the purpose of this study, seven audio samples: Classical, Country, Folk, Jazz, Metal, Pop and Rock music's are used as host audio for the existing and improved image watermarking algorithm. The performance of the IAWA was evaluated analytically and compared with that of EAWA using Bit Error Rate (BER), Signal to Noise Ratio (SNR) and Normalized Cross Correlation (NCC) as performance metrics.



	AGWN		Resampling		Re-quantisation		CROPPING1	
	IAWA	EAWA	IAWA	EAWA	IAWA	EAWA	IAWA	EAWA
Rock	0.097104	0.098301	0.010012	0.010204	0.049801	0.059804	0.009216	0.009951
Pop	0.096501	0.097707	0.010021	0.010312	0.046702	0.059702	0.008867	0.009976
Metal	0.090901	0.091813	0.010011	0.010211	0.048801	0.059801	0.009412	0.009947
Jazz	0.097802	0.098104	0.010101	0.010301	0.047803	0.059802	0.009101	0.009983
Folk	0.091509	0.092711	0.010001	0.010102	0.048801	0.059804	0.008516	0.009961
Country	0.094711	0.096803	0.010001	0.010101	0.047701	0.059702	0.007964	0.009967
Classical	0.096801	0.098912	0.010101	0.010201	0.048702	0.059701	0.008989	0.009992

Figure 5: BER Results Under Different Attacks

BAYERO JOURNAL OF ENGINEERING AND TECHNOLOGY (BJET) VOL.16 No.2, MAY, 2021 PP 89-95 ISSN: 2449 – 0539

3.1.1 Analysis of BER of the Watermarked Audios Under AWGN Attack

As depicted in Figure 5, the BER values of the seven watermarked audios produced by IAWA are all lower as compared to that of the EAWA when subjected to AWGN attack, with reduction of BER by 2.13%, 2.16%, 1.29%, 0.30%, 0.99%, 1.23% and 1.21% for Classical, Country, Folk, Jazz, Metal, Pop and Rock music's respectively over EAWA. The percentage reduction was obtained from the following mathematical expression.

$$\text{Percentage Reduction} = \frac{\sum \text{EAWA} - \text{IAWA}}{\sum \text{EAWA}} \text{x } 100 \qquad (10)$$

3.1.2 Analysis of BER of the Watermarked Audios Under Resampling Attack

As shown in Figure 5, the BER values of the seven watermarked audios produced by IAWA are all lower as compared to that of the EAWA when subjected to resampling attack with reduction of BER by 0.98%, 0.99%, 0.99%, 1.94%, 1.95%, 2.82% and 1.88%for Classical, Country, Folk, Jazz, Metal, Pop and Rock music's respectively over EAWA, and the percentage reduction were obtained from Equation 10.

3.1.3 Analysis of BER of the Watermarked Audios Under Requantization Attack

as illustrated in figure 5, the BER values of the seven watermarked audios produced by IAWA are all lower as compared to that of the EAWA when subjected to re-quantization attack with reduction of BER by 18.42%, 20.10%, 18.39%, 20.06%, 18.39%, 21.77 and 16.72% for Classical, Country, Folk, Jazz, Metal Pop and Rock music's respectively over EAWA and the percentage reduction was obtained from Equation 10.

3.1.4 Analysis of BER of the Watermarked Audios Under Cropping1 Attack

As shown in Figure 5, the BER values of the seven watermarked audios produced by IAWA are all lower when compared to that of the EAWA when subjected to cropping 1 attack, with reduction in BER by 10.03%, 20.09%, 14.50%, 8.83%, 5.37%, 11.11% and 7.38% for Classical, Country, Folk, Jazz, Metal, Pop and Rock music's respectively, over EAWA, and the percentage reduction was obtained from Equation 10. According digital watermarking standards, Low BER signifies high robustness (Akhaee and Al-Haddad 2015). Therefore, reduction in BER indicates improvement in terms of robustness achieved by the developed algorithm.

3.2 NCC Results After Application of Attack on the Watermarked Audios

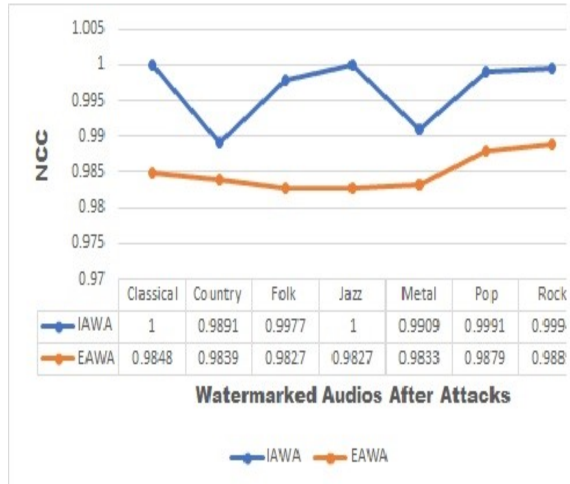


Figure 6: NCC Result on Watermarked Audios.

3.2.1 Analysis of NCC of the Watermarked Audios After Attacks

As illustrated from Figure 6 the NCC values of the seven watermarked audios produced by IAWA are all higher as compared to that of the EAWA after the attacks were subjected, with NCC improvement of 1.52%, 0.52%, 1.50%, 1.73%, 0.76%, 1.12% and 1.05% for Classical, Country, Folk, Jazz, Metal, Pop and Rock music's respectively over EAWA, thus, indicating that the developed IAWA achieved higher robustness when compared to the EAWA. The percentage improvement was obtained using equation 11.

$$\text{Percentage Improvement} = \frac{\sum \text{IAWA} - \sum \text{EAWA}}{\sum \text{IAWA}} \text{x } 100 \quad (11)$$

3.3 SNR Results After Application of Attack on the Watermarked Audios

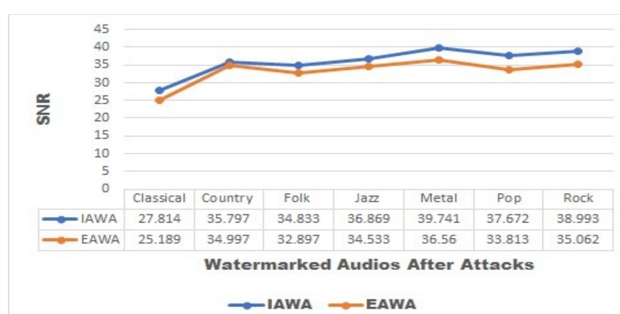


Figure 7: SNR Results After the Attacks

3.3.1 Analysis of SNR of the Watermarked Audios After Attacks

As illustrated in Figure 7, the SNR values of the seven watermarked Audios produced by IAWA are all higher as compared to that of the EAWA when subjected to all the attacks, with SNR improvement of 9.43%, 2.23%, 5.55%, 6.33%, 8.00%, 10.24% and 10.08% for Classical, Country, Folk, Jazz, Metal, Pop and Rock music's respectively over EAWA. An algorithm manifesting an SNR value above 20db signifies sufficient perceptual transparency [9]. Thus, indicating that the developed IAWA achieved higher inaudibility when compared to the EAWA. This is as a result of transformation technique with respect to the SWT i.e. translation invariance. and the percentage improvement was obtained using equation (11).

4 CONCLUSION

This paper presented an improved audio watermarking algorithm, the technique applied the use of SWT and SVD to make the watermarking algorithm more robust and more inaudible against various attacks. The performance of the developed algorithm was evaluated using BER, SNR and NCC as performance metrics. SNR was calculated in order to measure the inaudibility whilst the BER and NCC were calculated to measure the robustness of the technique when subjected to signal processing attacks. Classical music recorded an improvement in terms of inaudibility with Signal to Noise Ratio (SNR) value of 9.43% over existing watermarking scheme. Additionally, developed watermarking scheme achieved improvement in terms of robustness with Normalized Cross Correlation (NCC) value of 1.52% and reduction in Bit Error Rate (BER) when subjected to AWGN, Resampling, Re-quantization and cropping1 attacks.

5 REFERENCES

Al Haj. (2014). An imperceptible and robust audio watermarking algorithm. *EURASIP Journal on Audio, Speech, and Music Processing*, 37(1), 100-112.

Akhaee M. A. and Al-Haddad S. A. R. (2015) "Robust Digital Speech for online Speaker Recognition". *Hindaw*i *Mathematical Problems in Engineering*, 1-12.

Arashdeep K. and Malay K.D. (2018). High Embedding Capacity and Robust Audio Watermarking for Secure Transmission Using Tamper Detection. *ETRI Journal,* 40(1), 133-145.

Chen C.J., Chen S.T., Zhao M. and Lin M.J. (2019). Advances in Intelligent Information hiding and Multimedia Signal Processing. *Smart Innovation, Systems and Technologies* 156 133-141.

Devasis P. (2017). Invisible Digital Audio Watermarking using DWT-DCT based Transform. *IOSR Journal of Electronics and Communication Engineering (IOSR JECE),*12(5), 14-19.

Hazem H., AbdelLatif A. D., Alia M., Alfonso O. and Manuel A. (2016). Blind Audio Watermarking Technique Based on Two Dimensional Cellular Automata. *International Journal of Security and Its Applications,* 10(9), 175 – 184

Hu H.T. and Hsu L.Y. (2015). Robust, transparent and high-capacity audio watermarking in DCT domain. *ELSEVIER Signal Processing,* 109, 226–235.

Iqbal S.M.A., Chowdhury A.K., Akhter A. and Dhar K., (2014). A Novel Audio Watermarking Algorithm Based on Neighboring Samples. *16th Int'l Conference Computer and Information Technology*, 16(1), 473-477.

Jelena M., Ivan K. and Edis F. (2014). Wavelet based Watermarking approach in the Compressive Sensing Scenario. *4th Mediterranean Conference on Embedded Computing MECO Budva, Montenegro,* 24(3), 118- 123

Kulkarni E., Metkar S.P., Harshad C. K. and Kaustubh M. (2014). Audio Watermarking by using Image Comparison. *IEEE International*

Conference on Advances in Engineering & Technology Research (ICAETR - 2014), 10(6), 31-36.

Min-jae H., JeeSok L., MiSuk L. and Hong-GooK. (2016). SVD Based Adaptive QIM Watermarking on Stereo Audio Signals. *IEEE Transaction on Multimedia*, 1-10.

Morsi G. W. and El-Hawary. M. E. (2008). A New Perspective for the IEEE Standard 1459-2000 Via Stationary Wavelet Transform in the Presence of Nonstationary Power Quality Disturbance. *IEEE Transactions on Power Delivery*, 23(4), 2356-2365.

Sakriti D. (2015). Digital Audio-Watermarking using DWT-HAAR Transform. *International Journal of Advanced Research in Computer Science and Software Engineering*, 5(6), 804-808

Shijun X., Le Y., and Yi W. (2017). Robust and Reversible Audio Watermarking by Modifying Statistical Features in Time Domain. *Hindawi Advances in Multimedia*, 1-10.

Qian, C. and Xianghong, T. (2016). A digital Audio Watermarking Algorithm Based on Independent Component Analysis. *Biomedical Engineering and informatics journal,* 1053-1058.

Tsai S. (2015). A Robust Zero-Watermarking Scheme for Digital Audio. *International Journal of Information and Electronics Engineering, 5(2), 117-121.*

Wei C., Yixing Y. and Shengming L. (2009). Robust Image watermarking based on Singular Value Decomposition in DT-CWT Domain. *IEEE International workshop on imaging system and Techniques.* 28(9), 303-309.