Audio Watermarking using DWT-SVD-BFO

Nancy Tayal
M. Tech Student
Department of Computer Engineering
Indus Institute of Engineering & Technology, Kinana

Nikita Sagar
Head of the Department
Department of Computer Engineering
Indus Institute of Engineering & Technology, Kinana

Abstract

The main aim of this work is to develop a new watermarking algorithm within an existing discrete wavelet Transform (DWT) and singular value decomposition (SVD) framework. This resulted in the development of a combination of DWT-SVD-BFO (bacterial foraging optimization) watermarking algorithm. In this new implementation, the embedding depth was generated dynamically thereby rendering it more difficult for an attacker to remove, and watermark information was embedded by manipulation of the spectral components in the spatial domain thereby reducing any audible distortion. Further improvements were attained when the embedding criteria was based on bin location comparison instead of magnitude, thereby rendering it more robust against those attacks that interfere with the spectral magnitudes. The further aim of this thesis is to analyze the algorithm from a different perspective.
Keywords: discrete wavelet Transform (DWT), singular value decomposition (SVD), VHS

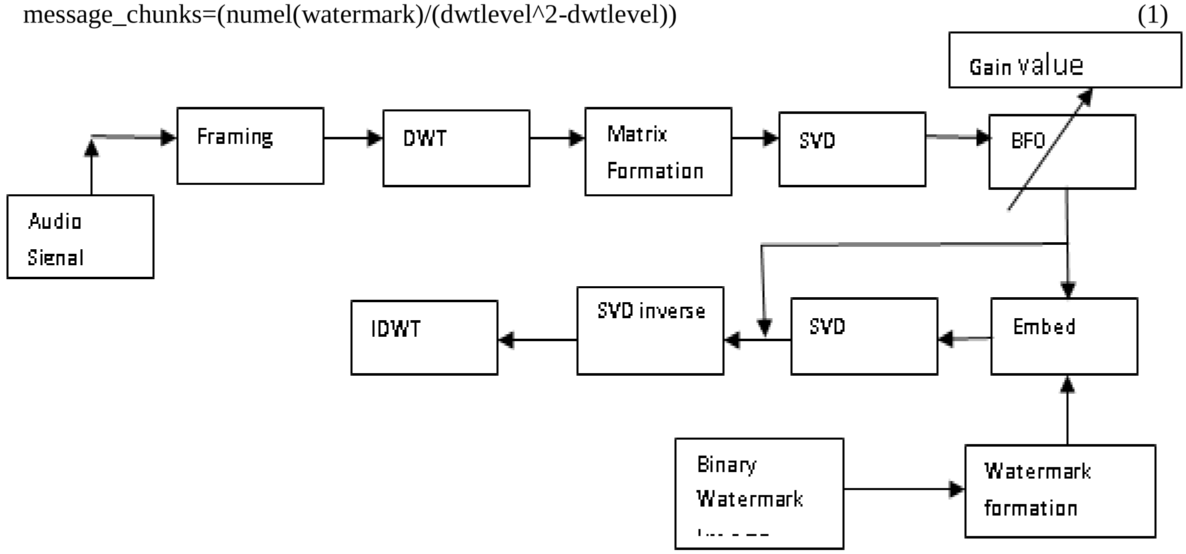
I. INTRODUCTION

Technological advances in computing, communications, consumer electronics and their convergence have resulted in phenomenal increases in the amount of digital content that is being generated, stored, distributed, and consumed. The term "content" broadly refers to any digital information, such as digital audio, video, graphics, animation, images, text, or any combinations of these types. This digital content can be easily accessed, perfectly copied, rapidly disseminated and massively shared without it losing quality, as opposed to the situation with earlier analogue media, such as audio cassettes and Video Home System (VHS) tapes.

However, these advantages of digital media formats over analogue transform into disadvantages with respect to copyright management, because the possibility of unlimited copying without a loss of fidelity has led to a considerable financial loss for copyright holders. In our work audio watermarking is the target area because of high digitally spread content of audio, music etc. The algorithm process must possess some characteristics like imperceptibility, exact detection etc. To match these, our work will be evaluated on the basis of peak signal to noise ratio (PSNR) and normalized cross correlation (NCC).

Till now many researchers have used optimization algorithms to tune the gain value but signal value for whole audio signal was used. This increases the chance of perception for gain value by third user. In our work, we have divided the audio signal into chunks and gain value is calculated by BFO as per the number of chunks. The number of chunks depends upon the watermark level used. We have developed the dynamic MATLAB script which can work for desired DWT levels. The mathematical formulation of that is given in equation 1.

message_chunks=(numel(watermark)/(dwtlevel^2-dwtlevel)) (1)



II. METHODOLOGY

The audio watermarking is necessary to avoid the original sound form theft and tempering. To protect and watermark the audio signal the algorithm should be such that it audio signal shouldn't be changed in any aspect and it should be robust and message hidden in the audio signal should not be retrieved very easily by others. To fulfill these we proposed the combination of Discrete Wavelet transform (DWT) and Singular Value Decomposition (SVD) is used which is further modified by bacterial Foraging optimization (BFO). The embedding formula used in our case is

$$S_w = S + \alpha . W \qquad (2)$$

Where S_w is the watermarked audio signal, S is the original audio signal, α is the gain factor which is the measure of robustness and imperceptibility of message and W is the watermark message bit.

The gain factor is the deciding factor for imperceptibility and robustness of watermarked audio signal and it should be optimum so that retrieval of watermark message is easier as high gain factor can lead to distorted extraction of message but robustness increases with high gain value. So we have used BFO optimization algorithm (discussed in previous chapter) to tune the gain factor value for tradeoff between PSNR (peak signal to noise ratio) and MSE (mean square error). The objective function chosen for this purpose considers the three evaluation parameters: PSNR, NCC (normalized cross correlation), MSE. The block diagram for watermark embedding is shown in figure 1. Four level DWT decomposition is used in this work for more robust and imperceptibility. For four-level DWT decomposition, this is done by forming a matrix of the detail sub-bands (D1, D2, D3 and D4) as shown in Figure 2. The resultant DWT matrix is processed by the SVD transform to embed the watermark bits, as will be explained in the next subsection.

Table - 1
Decomposition of input signal into 4 levels DWT

DWT1							
DWT2				*DWT2*			
DWT3		*DWT3*		*DWT3*		*DWT3*	
DWT4	*DWT4*	*DWT4*	*DWT4*	*DWT4*	*DWT4*	*DWT4*	*DWT4*

A. Algorithm for Watermark Embedding:

1) Convert the binary image watermark into a one-dimensional vector b of length M×N. A watermark bit bi may take one of two values: 0 or 1.

$$b_i = \{0,1\} \qquad 1 \le i \le M \times N$$

2) Sample the original audio signal at a sampling rate of 44,100 samples per second and partition the sampled file into N frames. The optimal frame length will be determined experimentally in such a way to increase data payload.
3) Perform a four-level DWT transformation on each frame. This operation produces five multi-resolution sub-bands: D1, D2, D3, D4, and A4. The D sub-bands are called 'detail sub-bands' and the A4 sub-band is called 'approximation sub-band'. The five sub-bands are arranged in the vector.
4) Arrange the four detail sub-bands D1, D2, D3, and D4 in a matrix D as shown in Figure 4.2. The matrix formation is done this way to distribute the watermark bits throughout the multi-resolution sub-bands D1, D2, D3, and D4. Forming the matrix with the Ds, rather than using A alone, is done to allow for matrix formation and subsequent application of the matrix-based SVD operator. The size of matrix D is 4 × (L/2), where L refers to the length of the frame.
5) Decompose matrix D using the SVD operator. This operation produces the three orthonormal matrices Σ, U, and V^T as follows:

$$D = U * \Sigma * V^T \qquad (3)$$

where the diagonal matrix Σ has the same size of the D matrix. The diagonal σii entries correspond to the singular values of the D matrix. However, for embedding purposes, only a 4 × 4 subset of matrix Σ, assigned the name S hereafter, is used as shown below. This is a trade-off between imperceptibility (inaudibility) and payload (embedding capacity). That is, using the whole Σ matrix for embedding will increase embedding capacity but will lead to severe distortion in imperceptibility (inaudibility) of the watermarked audio signal.

$$S = \begin{bmatrix} S_{11} & 0 & 0 \\ 0 & S_{22} & 0 \\ . & . & . \\ . & . & S_{nn} \end{bmatrix}$$

6) Arrange 12 bits of the original watermark bit vector b into a scaled 4 × 4 watermark matrix W. The watermark bits must be located in the non-diagonal positions within the matrix, as shown below.

$$W = \begin{bmatrix} 0 & bit1 & bit2 & bit3 \\ bit4 & 0 & bit5 & bit6 \\ bit7 & bit8 & 0 & bit9 \\ bit10 & bit11 & bit12 & 0 \end{bmatrix}$$

7) BFO algorithm starts from this step. The searching space dimension of bacteria is equal to the number of chunks as it will be final tuned gain value. The table 1 correlates the bio term in BFO algorithm with our technical counterpart.

Table - 1
Technical counterpart of bio inspired variables

	Variable in Bio Inspired Algorithm	Terms in our technical concept
1	Position of bacteria/swarms	Gain factor values
2	Number of dimension of searching space	Number of gain factors to be tuned for embedding
3	Update in positions	Change in the gain factor's value

The fitness function used for our proposed scheme in BFO takes PSNR, MSE and NCC in consideration which is:

$$\text{fitness value} = \left(\frac{1}{\text{PSNR}+\text{NCC}}\right) + \text{MSE} \qquad (4)$$

8) Since fitness value should be minimized so inverse of PSNR and NCC is considered here. For each collection of gain values (number of gain values is equal to number of audio signal chunks) the embedding algorithm is executed and fitness value is calculated by above equation. Once all iterations are finished, the gain factor values for minimum fitness function is picked as the final gain values, which are used further for embedding of watermark message. Embed watermark matrix W bits into matrix S according to the additive-embedding formula of equation 2.
9) Decompose the new watermarked matrix Sw using the SVD operator. This operation produces three new orthonormal matrices as follows:

$$S_w = U_1 * S_1 * {V_1}^T$$

The matrices U_1 and ${V_1}^T$ are stored for later use in the extraction process. This makes the proposed watermarking algorithm semi-blind, as the whole original audio frame is not required in the extraction process.
10) Apply the inverse SVD operation using the U and VT matrices, which were unchanged, and the S1 matrix, which has been modified according to Equation below. The Dw matrix given below is the watermarked D matrix.

$$D_w = U * \Sigma' * V^T$$

Where matrix Σ′ is the original Σ matrix with the S sub-matrix replaced by the S1 sub-matrix.
11) Apply the inverse DWT operation on the Dw matrix to obtain the watermarked audio frame.
12) Repeat all previous steps on each frame. The overall watermarked audio signal is obtained by concatenating the watermarked frames obtained in the previous steps.

B. Retrieval of Watermark:

Given the watermarked audio signal and the corresponding U1 and V1 matrices that were computed previously and stored for each frame, the embedded watermark can be extracted according to the procedure outlined in Figure 3 and described in detail in the following steps:

1) Step 1: Obtain the matrix S1′ from each frame of the watermarked audio signal following the general steps presented in Figure 3.

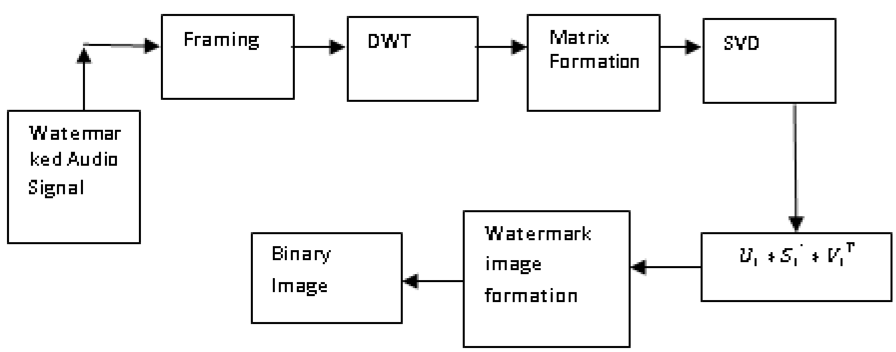


Fig. 3: Extraction block diagram of watermark

2) Step 2: Multiply matrix S1′ by U1 and V1 which were computed in the watermark embedding procedure and stored for use in the extraction process. This results in the following matrix.

$$S'_w = U_1 * {S_1}' * {V_1}^T$$

3) Step 3: Extract the 12 watermark bits from each frame by examining the non-diagonal values of matrix Sw'. It has been experimentally noticed that there are two groups of non-diagonal values that are extremely distinct. The values at the positions where a 0 bit has been embedded tend to be much smaller than those values at the positions where a 1 bit has been embedded. Thus, to determine the watermark bit W(n), the average of non-diagonal values is first computed, name it avg, then for each non-diagonal value Sw'ij, W(n) is extracted according to the following formula:

$$W(n) = \begin{cases} 0 & S_{wij} \le avg \\ 1 & \text{otherwise} \end{cases}$$

4) Step 4: Construct the original watermark image by assembling the bits extracted from all frames.

III. RESULTS

As discussed in previous chapter this thesis work is to suggest a new method for audio watermarking and analysis is done in MATLAB. We have used MATLAB 2013a's signal processing toolbox to test our proposed algorithm and a comparison is done with the already existing DWT-SVD algorithm. Results of reference paper are not quoted here, as test conditions are different along with sample watermark images and input audio samples. MATLAB's signal processing toolbox provided may functions ready to use which reduces our hassle to write script for those and we were able to concentrate on our proposed work's implementation. The results has been tested for a recorded signal at 44100 Hz frequency as well as at live recording of audio signal at the same frequency. Different types of watermark messages with varying sizes are used for analysis purpose. As described PSNR, NCC and MSE will be defining parameters in our work. The recorded input signal is shown in subplot 1 in figure 4.

The watermark message is embedded into audio signal and depth of embedding is based on the gain factor used. The gain factor should be optimal so that there is a tradeoff between PSNR and MSE as discussed in previous chapter. For this purpose a bio optimized algorithm named bacterial foraging optimization (BFO) is used which gives the tuned gain factor value which results in high PSNR and low MSE.

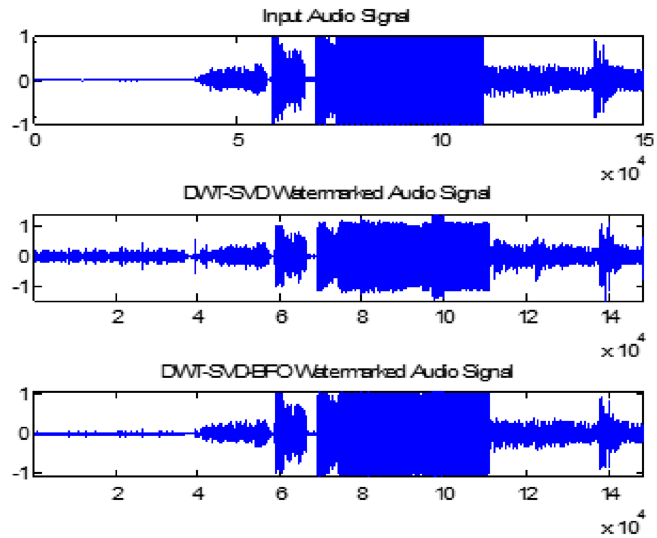


Fig. 4: (a) Recorded input signal (b) DWT-SVD watermarked signal (c) DWT-SVD-BFO watermark signal

The BFO parameters used for the purpose are tabulated in table2. The step size for the movement of bacteria and searching space dimension are important for a fine tuning and it can be observed by plotting the objective function value plot for whole iterations. With increase in iteration number the value of objective function should decrease and if it settles to a minimum value after certain iterations and doesn't change further then that would be a case of best optimization of BFO.

Table – 2
BFO initialization parameters

No of bacteria	*4*
No of chemotactic steps	*2*
Searching space dimension	*50*
Length of swim	*2*
Probability of elimination-dispersal	*0.25*

The searching space dimension of bacteria in BFO depends upon the number of tuning variables in the application used. In our case the input audio signal is divided into 50 chunks and for each chunk a different gain factor is used which gives optimal results for embedding. Thus we have 50 gain values for each chunk instead of single common value as in the case of DWT-SVD audio signal embedding process. For DWT-SVD embedding process the gain factor has been fixed at 10 and in proposed it is 50 in numbers having different values with mean of 2.0637. Results have been checked for different watermark message of varying size with same recorded audio signal. Results are shown in table 3 in appendix. The size of watermark message used for iteration isn't chosen so large as BFO is an iterative process and large message size will increase the execution time of algorithm. A bar plot for table 3 is shown in figure 5 below. The above comparison plot shows that proposed work is better than DWT-SVD embedding process as PSNR, NCC are higher and MSE is less than DWT-SVD process. The higher value of normalized cross correlation of the output watermarked audio signal with the input signal confirms the higher similarity between two as watermarked signal shouldn't be disturbed by message embedding and it should retain the originality. Similar is the case with PSNR. High PSNR indicates the more robustness of watermarked signal and less prone to attacks. This value is also high for our proposed case embedding process. The error should be less for embedding signal and since NCC is higher in our case so MSE will be lesser. Another analysis drawn from the above table is that the robustness of embedding process depends upon the size of image watermark message. A survey over results states that message with large number of columns than rows produce most robust watermarked signal and best in rest evaluation parameters also. To confirm it another test for the live recorded audio signal in noisy environment is recorded and tested for the same watermark messages. The audio signal is recorded for 5 seconds at frequency of 14500 Hz. The results for this are shown in table 4 in appendix. Again the analysis of bar graph and above table proves the point that the watermark message with more number of columns than rows provides more security and no change in the original audio signal. Almost 53% of increase in PSNR for the 2nd message is achieved by our proposed scheme.

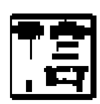
IV. CONCLUSION

This work proposes a new algorithm considering the tuning of gain factor for embedding of watermark message in the audio signal. Proposed algorithm is based on quantization in DWT domains with SVD while considering the more active components of the signal.

The performance of the algorithm is provided by evaluating the performance parameters such as peak signal to noise ratio, normalized correlation, and mean square error. From the results it is inferred that proposed algorithm is more robust than the DWT-SVD. The performance of the algorithm is improved by using the tuning of gain factor depending upon the number of chuncks of audio signal in the embedding process. Choosing proper gain value and wavelet filters have considerable effect on the performance of the algorithm. Wavelet filters' decomposition level are considered to evaluate the algorithm performance completely. From studies level 4 produces better results bringing a tradeoff between PSNR and MSE.

APPENDIX

Table - 3
Evaluation of proposed scheme for different watermark message

	Input Audio	*Watermark Message*	*DWT-SVD*			*DWT-SVD-BFO*		
			PSNR	*NCC*	*MSE*	*PSNR*	*NCC*	*MSE*
1	*Recorded Input signal of frequency 44100 Hz*	 *Size=27*22*	*7.2973*	*0.17689*	*0.18631*	*7.3959*	*0.18631*	*0.18213*
2		 *Size=20*50*	*20.0076*	*0.9219*	*0.0099818*	*34.1841*	*0.99844*	*0.00038156*
3		 *Size=36*39*	*5.2703*	*0.25985*	*0.29713*	*5.4447*	*0.29179*	*0.28543*

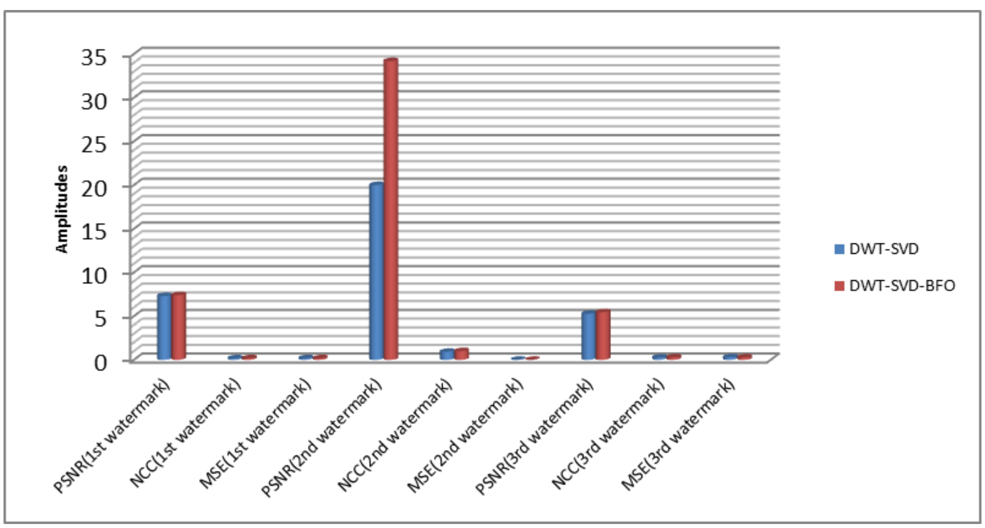
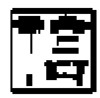


Fig. 5: Bar plot comparison of table 3 values

Table – 4
Evaluation of proposed scheme for different watermark message

	Input Audio	Watermark Message	DWT-SVD			DWT-SVD-BFO		
			PSNR	NCC	MSE	PSNR	NCC	MSE
1	Live noisy Recorded Input signal of frequency 44100 Hz	 Size=27*22	14.8713	0.3062	0.031564	16.287	0.437	0.022784
2		 Size=20*50	17.6615	0.54947	0.016601	26.597	0.90414	0.0021213
3		 Size=36*39	15.3364	0.405	0.028357	20.4728	0.74922	0.0086899

References

[1] Komal V. Goenka, Pallavi K. Patil,” Overview of Audio Watermarking Techniques” International Journal of Emerging Technology and Advanced Engineering, Volume 2, Issue 2, February 2012
[2] Ali Al-Haj,” An imperceptible and robust audio watermarking algorithm” EURASIP Journal on Audio, Speech, and Music Processing 2014.
[3] Darabkh, K.,”Imperceptible and Robust DWT-SVD-Based Digital Audio Watermarking Algorithm”, Journal of Software Engineering and Applications, 2014.
[4] Yekta Said Can, Fatih Alagoz, Melih Evren Burus,” A Novel Spread Spectrum Digital Audio Watermarking Technique” Journal of Advances in Computer Networks, Vol. 2, No. 1, March 2014
[5] Hwai-Tsu Hu, Hsien-Hsin Chou, Chu Yu and Ling-Yuan Hsu,” Incorporation of perceptually adaptive QIM with singular value decomposition for blind audio watermarking” EURASIP Journal on Advances in Signal Processing 2014
[6] Prayoth Kumsawat,” A Genetic Algorithm Optimization Technique for Multiwavelet-Based Digital Audio Watermarking” EURASIP Journal on Advances in Signal Processing Volume 2010
[7] Mehdi Sadeghzadeh, and Mahsa Taherbaghal,”A New Method for Watermarking using Genetic Algorithms” International Conference on Machine Learning, Electrical and Mechanical Engineering (ICMLEME'2014) Jan. 8-9, 2014 Dubai (UAE).
[8] Hong Peng, Zulin Zhang, Jun Wang And Peng Shi,” Audio Watermarking Framework Using Multi-Objective Particle Swarm Optimization” International Journal Of Innovative Computing, Information And Control ICIC International Control Volume **9**, Number **7**, July **2013**
[9] Hsiang-Cheh Huang, Feng-Cheng Chang,” Survey of Bio-inspired Computing for Information Hiding” Journal of Information Hiding and Multimedia Signal Processing, Volume 6, Number 3, May 2015
[10] Hongqin Shi,” DWT and SVD based Watermarking Scheme with Circulation” JOURNAL OF SOFTWARE, VOL. 9, NO. 3, MARCH 2014
[11] Huan Zhao1,Fei Wang1, Zuo Chen1, Jun Liu,” A Robust AudioWatermarking Algorithm Based on SVD-DWT”, ELEKTRONIKA IR ELEKTROTECHNIKA, ISSN 1392-1215, VOL. 20, NO. 1, 2014.
[12] Sartid Vongpraphip and Mahasak Ketcham,” An Intelligence Audio Watermarking Based on DWT- SVD Using ATS”, IEEE 2009
[13] Krishna Rao Kakkirala and Srinivasa Rao Chalamala,” Digital Audio Watermarking Using DWT-SVD and Secret Sharing”, International Journal of Signal Processing Systems Vol. 1, No. 1 June 2013
[14] N.V.Lalitha, G.Suresh, Dr.V.Sailaja,” Improved Audio Watermarking Using DWT-SVD”, International Journal of Scientific & Engineering Research Volume 2, Issue 6, June-2011
[15] Jyotirmayee Mishra, M.V.Patil,” An Effective Audio Watermarking using DWT-SVD”, International Journal of Computer Applications (0975 – 8887) Volume 70– No.8, May 2013
[16] AKSHYA KUMAR GUPTA, MEHUL S RAVAL,” A robust and secure watermarking scheme based on singular values replacement”, *Sa¯dhana¯* Vol. 37, Part 4, August 2012

