



Hybrid domain watermarking technique for copyright protection of images using speech watermarks

Rohit M. Thanki¹ · Ashish M. Kothari²

Received: 17 August 2018 / Accepted: 9 April 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

When digital images are shared over an open access network such as the internet, facebook, WhatsApp, and other social media, then the security of these images are required. The digital watermarking is one approach for the security of images (e.g. copyright protection, ownership authentication). In most of the watermarking approaches, secret information such as owner binary logos and texts are used for protection of images. These days, biometric watermarks such as human speech signals are preferred for protection of images. In this paper, a watermarking technique based on various signal processing transforms is proposed and implemented for the security of image using human speech signal. In this technique, the first discrete cosine transform (DCT) and then singular value decomposition (SVD) are applied on the watermark speech signal to get its hybrid coefficients which are inserted into hybrid coefficients of the cover image to get a watermarked image. These hybrid coefficients of cover image are first generated using discrete wavelet transform (DWT) and then fast discrete curvelet transform (FDCuT) is applied on it. The performance of techniques is tested for standard speech database such as TIMIT in terms of imperceptibility, robustness and payload capacity. The experimental results and comparison show that the proposed watermarking technique performs better than the existing watermarking techniques available in the literature. This technique may also be used for security of speech signal against spoof attack.

Keywords Biometric · Color image · Curvelet transform · Speech signal · Security · TIMIT database · Watermarking

1 Introduction

In modern society, digital images are easily shared using various platforms such as Facebook, Instagram, WhatsApp, and other social media. When these images are transmitted over the internet then they are easily copied or attacked by imposters. This issue creates serious problems like copyright protection. Thus, copyright protection of images is necessary when it is distributed over an open network (Langehaar et al. 2000; Thanki and Kothari 2017; Agarwal et al. 2019). The digital watermarking is one of many techniques used for security of images (e.g., copyright protection and authentication) (Borra and Swamy 2013). In this technique,

the watermark is hidden into the cover information to get secure watermarked information where watermark may be binary logos or texts and which have an identity of the owner or copyright holder (Thanki et al. 2017b, 2018a, b; Borra et al. 2017). But nowadays, many similar logos or texts are available in the market which creates confusion in ownership authentication. Thus, researchers are finding solutions for this issue by using biometric modalities of individual as secret watermark information. These biometric modalities can be physical characteristics, behavioral characteristics and are unique for every individual. But these characteristics have limitation against spoof attacks. Therefore, the security of biometric characteristics is also required against this attack (Jain and Kumar 2012; Jain et al. 2004; Jain and Uludag 2003a, b; Ratha et al. 2001). The watermarking is also one of the possible solutions for this issue (Rege 2012; Jain and Uludag 2002, Jain and Uludag 2003a, b; Jain et al. 2002). When biometric data is incorporated with the watermarking technique then it is referred to as 'biometric watermarking technique' (Rege 2012).

✉ Rohit M. Thanki
rohitthanki9@gmail.com

Ashish M. Kothari
amkothari@aits.edu.in

¹ C. U. Shah University, Wadhwan, Gujarat, India

² Atmiya University, Rajkot Gujarat, India

A. K. Jain and his research team proposed various watermarking techniques using various transforms for the security of fingerprint data and face data (Jain and Uludag 2002, 2003a, b; Jain et al. 2002). Vatsa et al. proposed various watermarking techniques based on various transforms for the security of biometric characterizes such as fingerprint and face (Vatsa et al. 2004, 2009; Noore et al. 2006, 2007). Rege (2012) proposed watermarking techniques using various transforms for the security of various biometric characterizes such as the face, speech, fingerprint, and sign (Inamdar and Rege 2014; Inamdar et al. 2010). The various watermarking techniques were also available in the literature for the security of iris features (Edward et al. 2011; Feng and Lin 2007; Park et al. 2007; Ko 2005). Recently, A. K. Singh and his research team proposed various watermarking techniques for security of medical images in telemedicine applications (Singh 2019; Thakur et al. 2018a, b; Kumar et al. 2018).

The rest of this paper is organized as follows. In Sect. 2, a summary of related works is given. The novel contributions of proposed work are highlighted in Sect. 3. The information of various transform used for designing of the technique is given in Sect. 4. The description on the proposed technique in the hybrid domain is given in Sect. 5. The comparative results and analysis of the proposed technique with existing techniques are described in Sect. 6. The possible application of the proposed technique is described in Sect. 7. In lastly, conclusions on this proposed technique is given in Sect. 8.

2 Related work

Speech watermarking technique is one of the applications of the biometric watermarking technique. Here, the speech signal is taken either as a cover medium or secret watermark information. Various existing techniques are related to the proposed technique and are summarized below.

El-Gazar et al. (2018) proposed a singular value decomposition (SVD) and data encryption standard (DES) based speech watermarking technique. In this technique, the DES algorithm was applied to the watermark image to generate an encrypted watermark image. Then, this encrypted watermark image was embedded into the singular value of speech signal to get watermarked speech signal. Revathi et al. (2018) proposed a discrete wavelet transform (DWT) based speech watermarking technique. In this technique, information of cover and watermark are speech signals. Authors were also giving person identification method based on watermarked speech signal using clustering algorithms. Merrad and Saadi (2018) and Merrad et al. (2018) proposed a hybrid domain speech watermarking technique using discrete cosine transform (DCT), DWT and sub-sampling. Here, first, sub-sampling is applied to a speech signal to get segment speech signal. The 1st DWT is applied to each

segment of the signal and then, DCT is applied to wavelet coefficients on it to get hybrid coefficients of each segment of the speech signal. These hybrid coefficients of speech segments are modified by watermark image and the key to getting its modified hybrid coefficients. The inverse DCT and inverse DWT are applied to this modified hybrid coefficients to get watermarked segmented speech signal. Finally, inverse sub-sampling is applied to a watermarked segmented speech signal to get watermarked speech signal. Ali et al. (2018) proposed a Hurst exponent and zero-crossing based speech watermarking technique. Thanki et al. (2018b) proposed a hybrid speech watermarking technique. Here, watermark information modifies hybrid coefficients of the speech signal using a scaling factor. Tsai and Yang (2018) proposed DCT and error correcting codes (ECC) based audio watermarking technique. The ECC is used for securing the watermark information.

Thanki et al. (Thanki and Borisagar 2017) proposed a curvelet transform and compressive sensing (CS) encryption-based audio watermarking technique. In this technique, the encrypted speech signal is used for ownership identification of audio signals. Li et al. (2017) proposed QIM and LSB based speech watermarking technique. Here, quantized bits of cover speech frames are modified by watermark bits using LSB substitution approach. Nematollahi et al. (2017a) proposed Least Significant Substitution (LSB) based speech watermarking technique. Here, Watermark logo is inserted into LSB of line spectral frequencies of the cover speech signal.

Nematollahi et al. (2017b) proposed a speech watermarking technique using quantization of the L_p -norm. Here, each speech signal frame is divided into two vectors based on the odd and even index values. The QIM is used to insert watermark information into the ratio of L_p -norm between these two indices. Finally, Lagrange optimization technique is used to reduce embedding distortion in the modified signal. Talbi et al. (2017) proposed a DCT and amplitude modulation based watermarking technique for speech signal. Here, frequency coefficients (which are DCT in nature) of the cover image are modified by amplitude modulated speech signal. Renza et al. (2016) proposed a quantization index modulation (QIM) and DWT based watermarking technique for copyright protection of color image.

Nematollahi et al. (2015) proposed a speech watermarking technique using hybridization of DWT + SVD. This existing technique was the application of Bhat technique (Bhat et al. 2010) for the security of speech signal. Inamdar and Rege (2014) proposed multiple watermarks based watermarking technique for protection of facial features and speech signal. Mani et al. (Mani and Lakshmi 2013) proposed a watermarking technique using DCT for the security of speech signal. Patel et al. (2011) proposed frequency masking and Fast Fourier Transform (FFT) based

speech watermarking technique. Jundale and Patil (2010) proposed the DWT based speech watermarking technique. Here, approximation coefficients of cover image are modified by watermark speech signal. Kaur et al. (2010) designed signature and speech based multimodal biometric system.

The watermarking techniques in the literature are mainly robust and applied for protection of speech signal or digital image. First, four existing watermarking techniques: Talbi technique (Talbi et al. 2017), Renza technique (Renza et al. 2016), Inamdar technique (Inamdar and Rege 2014), and Jundale technique (Jundale and Patil 2010) are chosen for copyright protection of images using speech watermarks. These existing techniques are designed using DCT and DWT. But the results of these techniques are not good for imperceptibility to cover image and perceptual quality of the extracted watermark speech signal. Also, these techniques directly insert watermark speech signal into a cover image. Thus, these techniques have less security to the watermark speech signal.

3 Major contributions of the proposed work

In this paper, a hybrid watermarking technique using speech watermark is developed and proposed to overcome some limitations of existing techniques (Talbi et al. 2017; Renza et al. 2016; Inamdar and Rege 2014; Jundale and Patil 2010). The usage of the DWT along with Fast Discrete Curvelet Transform (FDCuT) provides better transparency to watermarked image compared to existing techniques. The novel features of this paper are that (1) this paper uses haar wavelet matrix to compute wavelet coefficients of the cover image instead of conventional wavelet computational method such as `dwt2` in MATLAB. (2) this paper uses hybridization of DCT+SVD for the security of watermark information instead of the encryption method. The special features of the proposed technique are the following:

- *Robustness and imperceptibility* in the proposed technique, the hybrid coefficients of the watermark speech signal are inserted into the hybrid coefficients of the cover image which increase the imperceptibility and robustness of the proposed technique as compared to many related existing techniques (Talbi et al. 2017; Renza et al. 2016; Inamdar and Rege 2014; Jundale and Patil 2010).
- *Security* in the existing techniques (Talbi et al. 2017; Renza et al. 2016; Inamdar and Rege 2014; Jundale and Patil 2010), the watermark speech signal is directly inserted into the cover image. Thus, these techniques provide less security for the speech signal. In the proposed technique, the hybrid coefficients of the speech signal are

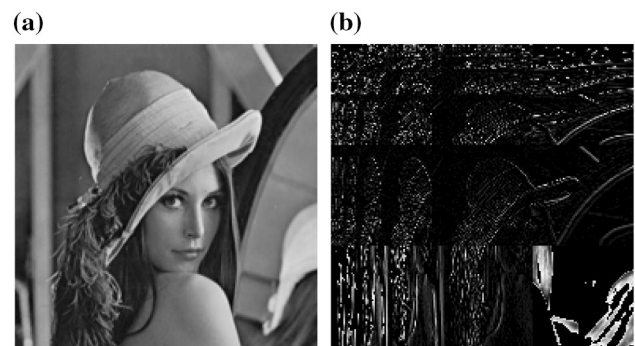


Fig. 1 a Test lena image. b Its wavelet coefficients

inserted into the cover image instead of the actual value of the signal.

4 Mathematical preliminaries

In this paper, hybridization of various transforms is utilized for designing of the proposed technique. The information of these transforms is given as per follows:

4.1 Discrete Wavelet Transform (DWT)

In the proposed technique, the watermark speech signal is modified the hybrid coefficients of the cover image. These coefficients are curvelet coefficients (high frequency in nature) of all wavelet subbands of the cover image. Here, Haar wavelet matrix (Saxena et al. 2017; Yan 2009; Vidakovic 1999) and its inverse version is used to get wavelet coefficients of the cover image. The wavelet coefficients of the cover image are shown in Fig. 1; here the representation of coefficients differs from the conventional representation of wavelet coefficients. This happens due to the use of the Haar wavelet matrix instead of using the conventional `dwt2` function of MATLAB.

The other reason behind choosing the Haar wavelet matrix in the proposed method (Saxena et al. 2017) is that (1) this wavelet transform has sparse data compared to another transform such as Walsh. (2) The input and output of this wavelet matrix is the same size and it is a power of 2. (3) The wavelet analyses the local feature of the image due to its orthogonality property.

4.2 Discrete curvelet transform (DCuT)

The DCuT (Candes et al. 2006; Candes and Donoho 2004) is an advanced image processing transform and represents an image in its edges. The reason behind DCuT used in watermarking is that due to its sparsity properties, it overcomes the limitation of conventional image transformation such as

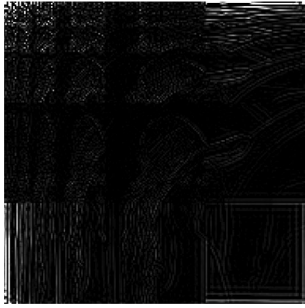


Fig. 2 Hybrid coefficients (in high frequency nature) lena image

FFT, DCT, and DWT. The limitation of missing directional selectivity of DWT overcomes by DCuT. *Sparsity means that “the matrix entries decay nearly exponentially fast, and they are well organized in the sense that very few nonnegligible entries occur near a few shifted diagonals”* (Candes et al. 2006). In this paper, 1st level DCuT is applied to all wavelet coefficients of the cover image to get its hybrid coefficients. Then, choose hybrid coefficients (in high frequency nature) for embedding of the watermark signal. The hybrid coefficients of the cover image are shown in Fig. 2. The coefficients show that the value of many coefficients is near zero and less visual information of cover image appears in

it. Thus, sparsity property of DCuT is utilized in this proposed technique for obtaining hybrid coefficients of the cover image.

4.3 Discrete cosine transform (DCT)

In this proposed technique, hybrid coefficients of the watermark speech signal are used to achieve security before embedding into the cover image. These coefficients are generated using hybridization of DCT + SVD. The DCT (Jain 1989) converts the signal into its frequency coefficients. The frequency coefficients of the speech signal are shown in Fig. 3.

4.4 Singular value decomposition (SVD)

The signal is reshaped into the matrix for proper application of SVD (Golub and Reinsch 1970) on it. The SVD is applied to the signal matrix and decomposes it into three matrices: U, S, and V. The S matrix values have real numbers and are placed diagonally in the matrix. As the S matrix values are stable, it makes SVD more popular in watermarking. In the proposed technique, SVD is used to generate stable hybrid coefficients (singular values of DCT coefficients) of the watermark speech signal. The hybrid coefficients of

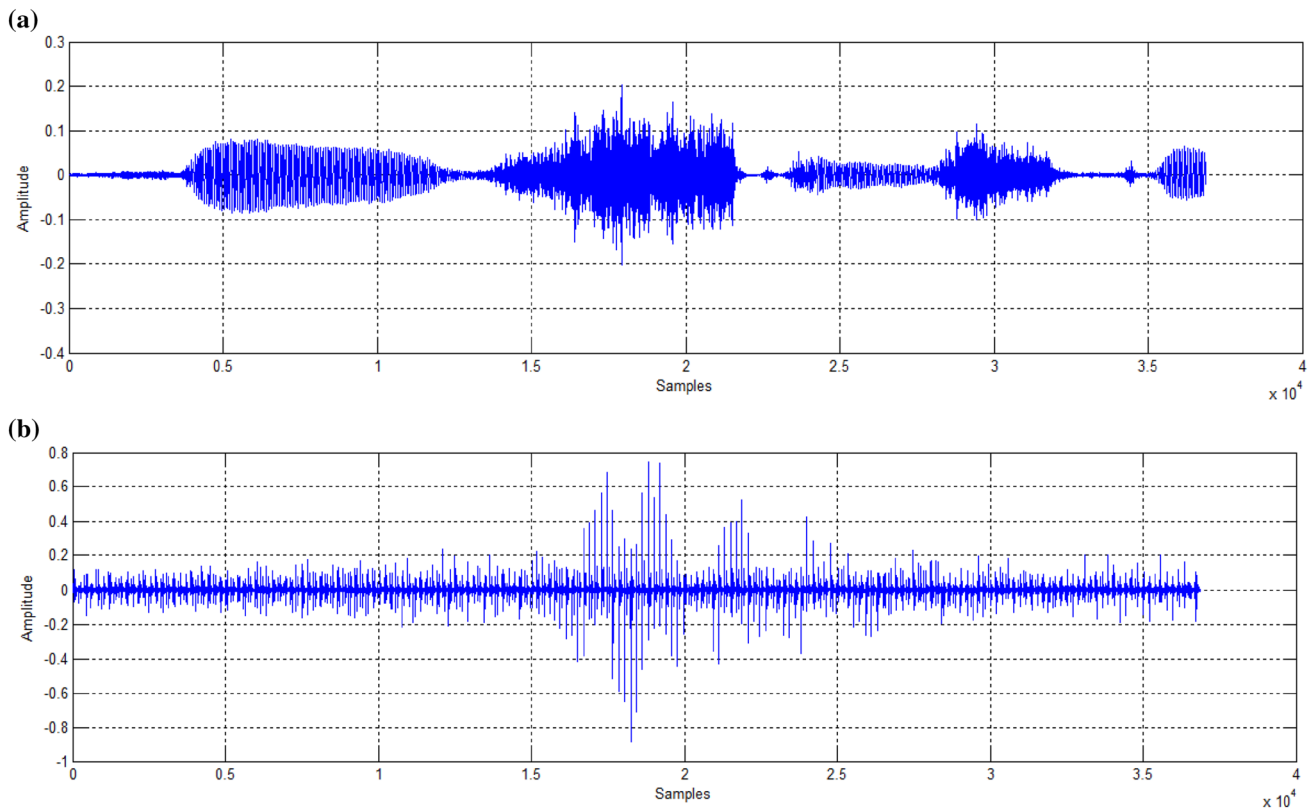


Fig. 3 **a** Speech signal. **b** Its DCT coefficients

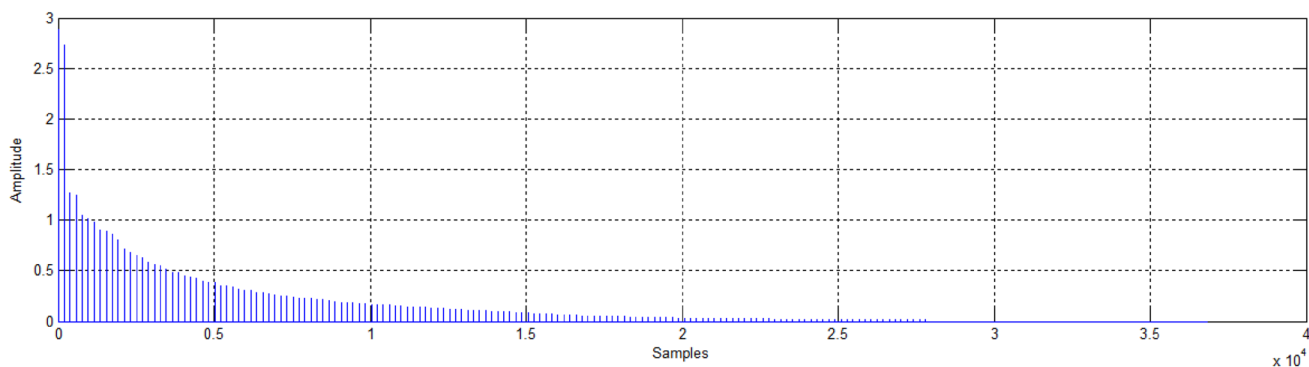


Fig. 4 Hybrid coefficients of speech signal

the watermark speech signal are shown in Fig. 4. The main significance of Fig. 4 is that it shows sparse information of watermark speech signal which is difficult to predict by an attacker when illegally extracted it from the watermarked image.

5 Proposed technique

In this paper, a non-blind, high imperceptible and robust watermarking technique based on DWT-FDCuT is proposed. The steps for watermark speech signal embedding and watermark speech signal extraction are given in the following subsections. In the proposed technique, the curvelet transform is applied to wavelet coefficients of the cover image to get hybrid coefficients of the cover image. Then high frequency hybrid coefficients of cover image are chosen for embedding of hybrid coefficients of the watermark speech signal. The hybrid coefficients of the watermark speech signal are obtained using DCT and SVD.

For extraction the watermark speech signal from the watermarked image, there are various security keys are required. These keys such as embedding factor, original hybrid coefficients of host image and information of orthogonal matrices U , V are required for this purpose. These keys are provided security to images in various stages in technique. The requirement of these keys is making proposed technique little be complicated but provide excellent security to images. Because attacker or imposter is required three keys to extract watermark speech signal which is very difficult to get compared to getting one cipher key. These all keys are numeric values depending on characteristics of the watermark speech signal and required a trade-off in technique. The keys are generated using various signal transforms such as DWT, DCuT, and SVD. The values of these keys are represented in bytes. The overall size of all keys is approximately around 100 kB which can be stored in any computer system, server or cloud. This is a very small

amount data consideration of modern times of computer and server which have very high storage capacity for data storage.

5.1 Process for embedding of watermark speech signal in cover image

The process for embedding of a watermark speech signal is shown in Fig. 5. The steps for embedding of the watermark speech signal is given as follow. The watermark speech signal w is first reshaped in the matrix. The reason behind signal conversion into the matrix is that it is easy to be processed in MATLAB.

Step 1 The watermark speech signal w is reshaped into watermark matrix M . The DCT coefficients (D) of matrix M is obtained by applying DCT on it

$$D = DCT(M). \quad (1)$$

Step 2 Apply SVD on DCT coefficients of matrix M to obtained matrices of U , S , and V . The values of the singular (S_w) matrix or watermark matrix M is taken as hybrid coefficients of the watermark speech signal

$$[U_w, S_w, V_w] = SVD(D). \quad (2)$$

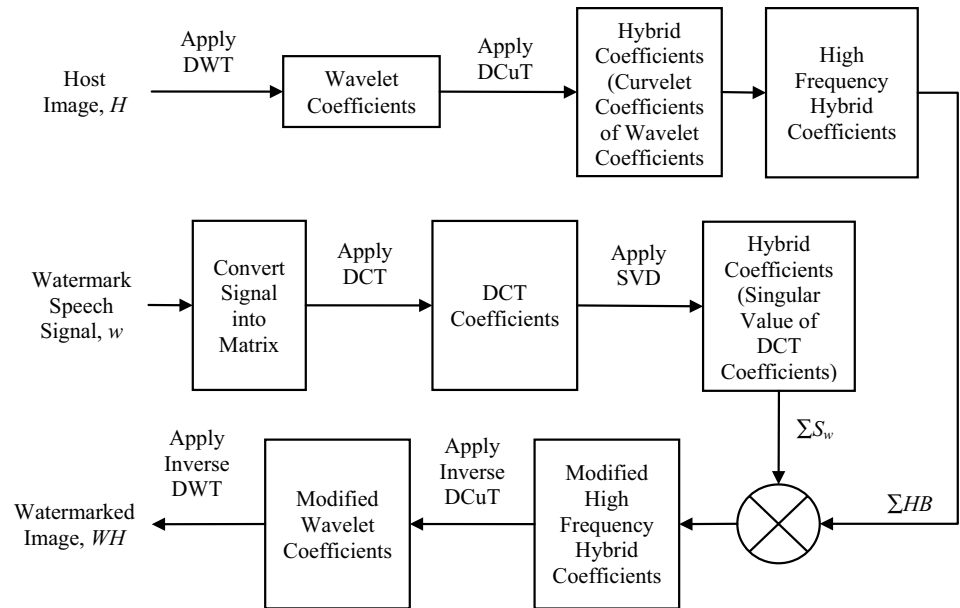
Step 3 Take cover image H . The wavelet coefficients (WA) of cover image H are obtained using below Eq. (3):

$$WA = \Psi \times H \times \Psi', \quad (3)$$

where Ψ and Ψ' is the Haar wavelet matrix and its inverse version, respectively.

Step 4 The 1st level DCuT is applied to wavelet coefficients of the cover image to obtained its hybrid coefficients (HB). The high frequency hybrid

Fig. 5 Proposed embedding process



coefficients are chosen for embedding of hybrid coefficients of the watermark matrix

$$HB = DCuT(WA) \tag{4}$$

Step 5 The hybrid coefficients of the watermark matrix modify the high frequency hybrid coefficients of cover image with help of embedding factor (k) and using multiplicative watermarking equation

$$HB^*_{High_Coeff} = HB_{High_Coeff} \times (1 + k \times S_w), \tag{5}$$

where HB^* is modified high frequency hybrid coefficients of the cover image. The k is an embedding factor which is set by the user. It scales the pixels of the watermark while it is embedded into the cover image. While small values of k , leads to high imperceptibility, large values lead to high robustness. Hence its value is chosen so as to reduce the trade-off between imperceptibility and robustness.

Step 6 The modified wavelet coefficient (WA^*) of the cover image is obtained by applying inverse DCuT on modified hybrid curvelet coefficients with unmodified curvelet coefficients of the cover image

$$WA^* = IDCuT(HB^*). \tag{6}$$

Step 7 Finally, the watermarked image WH is obtained using below Eq. (7)

$$WH = \Psi' \times WA^* \times \Psi, \tag{7}$$

where WH is the watermarked image.

5.2 Process for extraction of watermark speech signal from watermarked image

For extraction of the watermark speech signal from watermark image, information regarding hybrid coefficients of the original cover image, embedding factor and orthonormal matrices (U matrix, V matrix) is required. This information is used three secret keys in the proposed technique. The block diagram of the process of extraction of the watermark speech signal from the watermarked image is shown in Fig. 6.

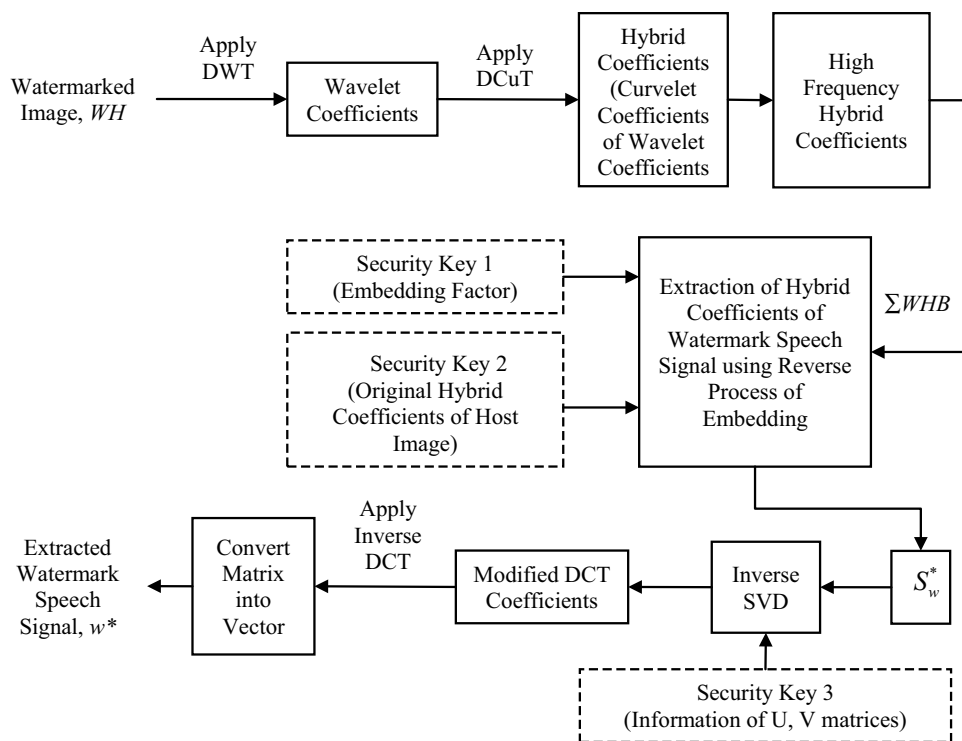
The steps for extraction of a watermark speech signal is given as follow:

- Step 1 Take watermarked image WH . The wavelet coefficients of watermarked image WH are obtained using Eq. 3
- Step 2 The 1st level DCuT is applied to wavelet coefficients of the watermarked image to obtain its hybrid coefficients (WHB). The high frequency hybrid coefficients are chosen for extraction of hybrid coefficients of the watermark speech matrix
- Step 3 Extract hybrid coefficients of the watermark speech matrix using the embedding factor (security key 1) and original hybrid coefficients of cover image (security key 2) as follow:

$$S_w^* = \left\{ \frac{HB^*_{High_Coeff}}{HB_{High_Coeff}} - 1 \right\} / k, \tag{8}$$

where S_w^* is the extracted hybrid coefficients of the watermark speech matrix.

Fig. 6 Proposed extraction process



Step 4 Apply inverse SVD on hybrid coefficients of the watermark speech matrix using orthonormal matrices U, V (Security key 3) to extract its DCT coefficients

$$D^* = U_w \times S_w^* \times V_w', \tag{9}$$

where D^* is extracted DCT coefficients of the watermark speech matrix.

Step 5 Apply inverse DCT on extracted DCT coefficients of the watermark speech matrix to get actual pixel values of the watermark speech matrix

$$M^* = IDCT(D^*), \tag{10}$$

where M^* is extracted watermark speech matrix.

Step 5 Finally, watermark speech matrix reshapes into a signal to get the extracted watermark speech signal w^* on the detector side

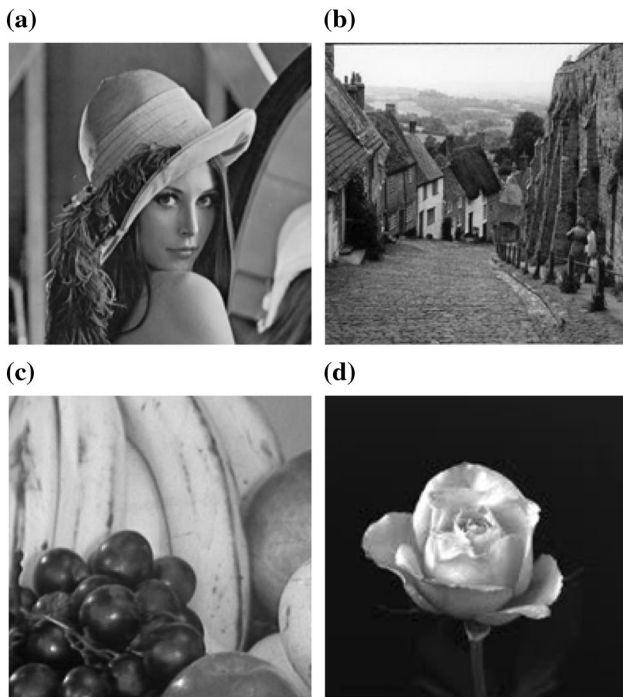
$$w^* = reshape(M^*), \tag{11}$$

where w^* is extracted watermark speech signal.

6 Experimental results

The comparative analysis of the proposed technique is tested by various grayscale cover images (192×192 pixels) which are taken from SIPI database (University of South Carolina 2017) (shown in Fig. 7). The 8-bit mono man speech signal and woman speech signal are used as watermark speech signal (shown in Fig. 8). The duration of the speech signal is 2 s with 36,864 samples. The word of man speech used is “just long enough to make you feel important”. The word of woman speech used is “This is a”. These speeches are obtained from the freesound website (Speech Signal Database 2017). The performance analysis of technique is done using various values of embedding factor k .

A cover image of size 192×192 is multiplied with Haar wavelet matrix and its inverse wavelet matrix, to obtain wavelet coefficients of size 192×192 of the cover signal. The 1st level DCuT is applied to wavelet coefficients to obtain its curvelet coefficients (high and low). For the embedding of the watermark speech signal, curvelet coefficients of size 192×192 with high frequency are selected. The watermark speech signal with 36,864 samples is reshaped into matrix M of the size of 192×192 . The DCT coefficients of the matrix are obtained using application of 1D DCT on it. The



forward SVD is then applied on DCT coefficients to obtain U matrix, S matrix and V matrix with each size 192×192 . The values of S matrix are embedded into high frequency curvelet coefficients of the cover image to get watermarked image.

The payload capacity (PC) of watermark technique is defined the number of watermark bits inserted into the cover image and can be calculated using below equation (Kutter and Petitcolas 1999):

$$PC = \frac{Watermark_Bits}{Host_Bits} bpb, \tag{12}$$

where PC is a payload capacity, $Watermark_Bits$ is a total number of bits for watermark data, $Host_Bits$ is a total number of bits for cover data, and bpb is the bit per bit.

In this proposed technique, the total numbers of bits for watermark speech signal are 589,776 and total numbers of bits for the cover image are 303,536. Thus, payload capacity (PC) of the proposed technique is calculated as:

$$PC(bpb) = \frac{589776}{303536} = 1.94. \tag{13}$$

Fig. 7 Test cover images a Lena. b Goldhill. c Fruit. d Rose

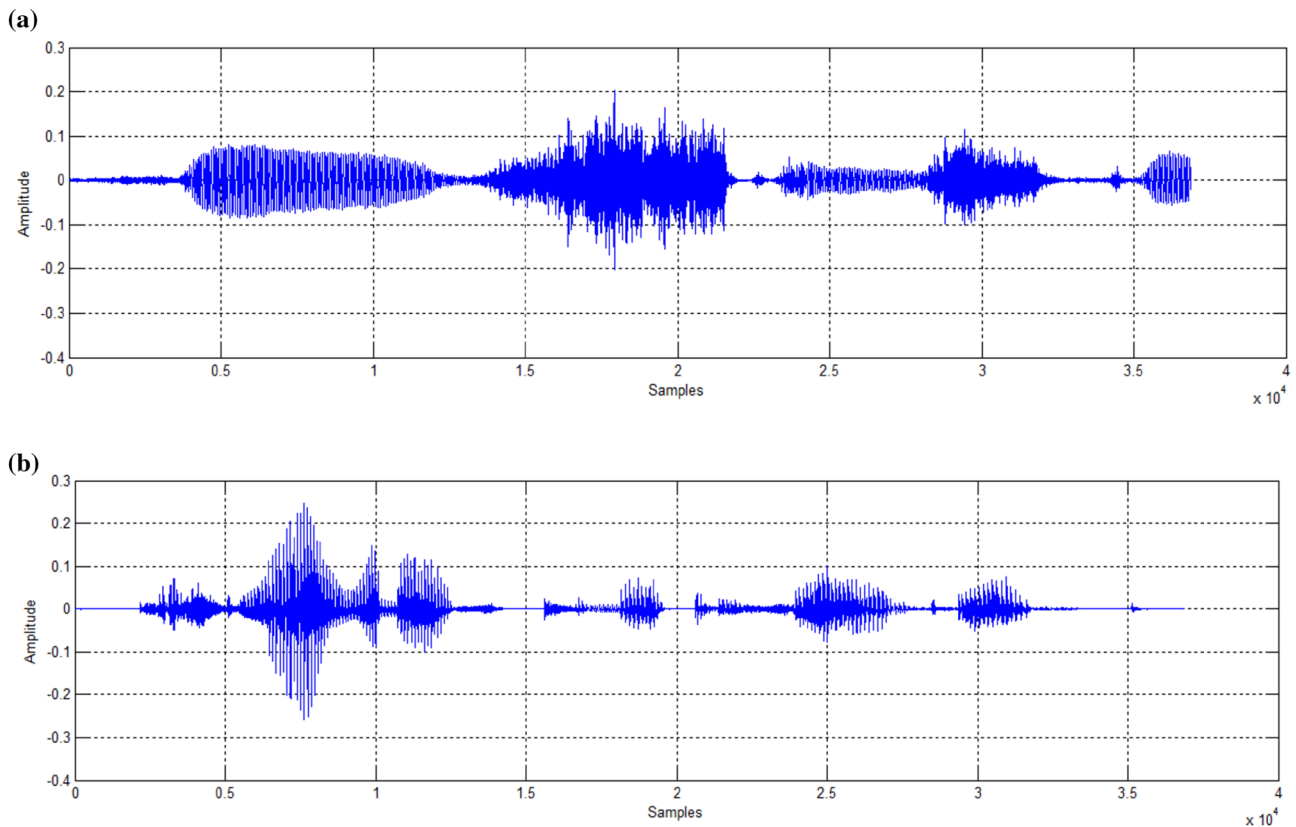


Fig. 8 Watermark signals a Woman speech. b Man speech

Table 1 Perceptual quality decision about extracted watermark speech signal based on SF value

SF values	A decision about perceptual quality of extracted watermark speech signal
$0 < SF < 0.4$	Extraction of signal is not possible and not audible
$0.4 < SF < 0.5$	Extraction of signal is possible and audible
$0.5 < SF < 0.6$	Extraction of signal is possible with average audible quality
$0.6 < SF < 1.0$	Extraction of signal is possible with good audible quality

6.1 Evaluation parameters for proposed technique

The peak signal to noise ratio (PSNR) (Kutter and Petitcolas 1999) is used to measure the imperceptibility of the proposed technique. The robustness of the proposed technique is measured by Similarity Factor (SF) (Jundale and Patil 2010), Bit Error Rate (BER) (Kutter and Petitcolas 1999), and Watermark to Noise Ratio (WNR) (Zhong 2007). The SF, BER, and WNR are calculated using the following equations. The WNR is measured in decibel values while SF and BER are measured in numerical values.

$$SF(w, w^*) = \frac{w \times w^*}{\sqrt{w \times w^*}}, \quad (14)$$

$$BER = \frac{\text{No. of Wrong bits between } w^* \text{ and } w}{\text{No. of bits of } w}, \quad (15)$$

where w is the watermark speech signal and w^* is the extracted watermark speech signal, respectively.

$$D_w = \frac{1}{M \times N} \sum_{x=1}^M \sum_{y=1}^N |WH(x, y) - H(x, y)|^2,$$

$$D_a = \frac{1}{M \times N} \sum_{x=1}^M \sum_{y=1}^N |CWH(x, y) - WH(x, y)|^2, \quad (16)$$

$$WNR = 10 \log_{10} \left(\frac{D_w}{D_a} \right),$$

where WH is a watermark image, H is a cover image, D_w is a distortion occurred during embedding process, CWH is a corrupted watermark image by attacks, D_a is a distortion occurred due to attacks applied on the watermarked image, and WNR is a watermark to noise ratio.

For robustness checking of watermarking technique, the acceptable value of BER is near to zero (Kutter and Petitcolas 1999) and value of WNR should be -20 dB to 20 dB (Zhong 2007). Based on the SF values for the watermark speech signal, the decision of the perceptual quality of the extracted watermark speech signal is decided (Jundale and

Patil 2010). If the SF value is less than 0.4 , then extracted watermark speech signal is not audible. If the SF value is in between 0.4 to 0.6 then extracted watermark speech signal is audible with a lot of noise. The quality of the signal is average. If the SF value is greater than 0.6 , then extracted watermark speech signal is audible with less noise and the quality of the signal is good. Table 1 summarizes perceptual quality decision for extracted watermark speech signal based on SF value.

6.2 Performance analysis of proposed technique

The resultant images using the proposed technique are shown in Fig. 9. The resultant images are obtained using the Lena image as a cover image and woman speech as watermark signal. By comparing the obtained results, this is indicated that this technique is invisible in nature.

The PSNR, SF, and BER values of the proposed technique for various embedding factors are given in Table 2. Referring to Table 2, it indicates that the quality of the watermarked image varies with embedding factors, while the quality of extracted watermark speech is stable with varying embedding factors. It is observed that the PSNR values start decreasing after embedding factor value (100). It is also observed that the SF and BER values are stable for all embedding factor for all tested speech signal. The reason behind getting the same value of SF and BER for all embedding factor is that in the technique, hybrid coefficients (singular value of DCT coefficients) of the watermark speech signal are used to modified high frequency hybrid coefficients of the cover image. The hybrid coefficients of the watermark speech signal have sparse diagonal values which are less modified during embedding and extraction process. Also, these coefficients of the watermark speech signal have less information regarding actual information of speech. That way, the same value of SF and BER achieved in this technique. It is shown that this technique has a good trade-off between watermarking requirements.

Furthermore, for the stability of the proposed watermarking technique, watermarked image and extracted watermark woman speech signal for various embedding factor values are shown in Figs. 10 and 11, respectively. The visual results in Figs. 10 and 11 shows that this proposed technique performs equally well for various embedding factors.

For robustness checking of the proposed technique, the performance of the technique is analyzed against various watermarking attacks. The SF and WNR values of extracted watermark woman's speech signal under watermarking attacks are given in Table 3. Figure 12 shows the SF values of extracted watermark woman's speech signals against various watermarking attacks. The robustness results are generated using embedding factor $k=400$. The SF values being

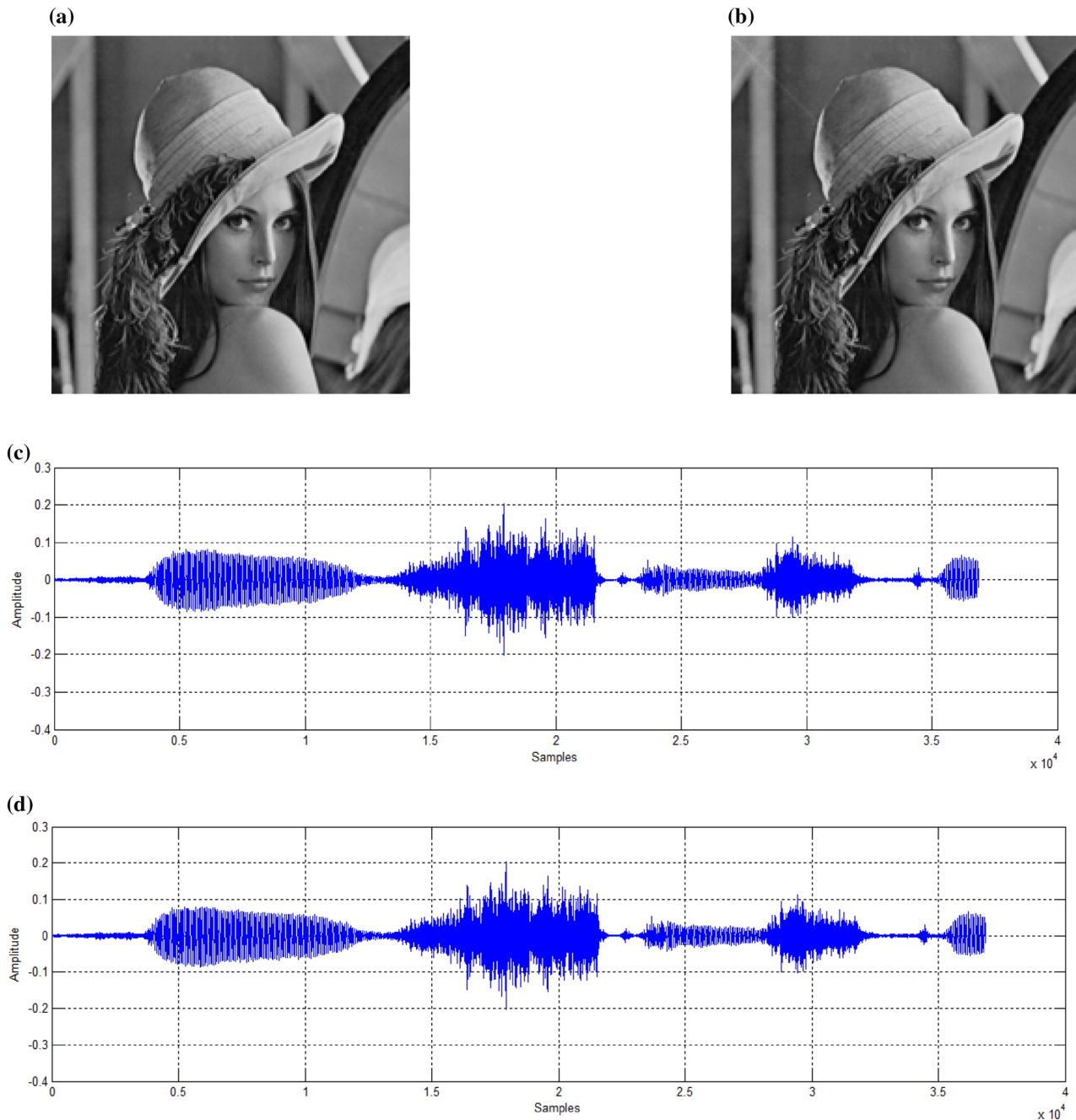


Fig. 9 **a** Original cover image, **b** watermarked image, **c** original watermark woman speech signal, **d** extracted watermark woman speech signal

greater than 0.4 shows that the extraction of a watermark speech signal can be performed when attacks are applied to the watermarked image. For Table 3, it is observed that for various attacks, the WNR values is varied from -21 to 15 dB. The results in Table 3 indicate that this technique is robust in nature.

6.3 Computational complexity of proposed technique

The computational complexity of the proposed technique of any watermarking technique can be measured by the computational time when it is used for copyright protection of the image. Here, four image transforms are used for designing and implementation of the proposed watermarking technique. Here, the computational time is divided into two

Table 2 PSNR, SF and BER values of proposed technique (a) For watermark woman speech signal (b) For watermark man speech signal

Embedding factor	PSNR (dB)	SF	BER	PSNR (dB)	SF	BER
(a)	Lena image			Goldhill image		
100	40.69	0.9843	0.0035	41.32	0.9843	0.0035
200	34.67	0.9843	0.0036	35.31	0.9843	0.0035
300	31.15	0.9843	0.0035	31.78	0.9843	0.0036
400	28.65	0.9843	0.0036	29.28	0.9843	0.0035
	Fruit image			Rose image		
100	41.19	0.9843	0.0036	41.32	0.9843	0.0036
200	35.17	0.9843	0.0035	35.30	0.9843	0.0036
300	31.64	0.9843	0.0035	31.78	0.9843	0.0035
400	29.15	0.9843	0.0035	29.28	0.9843	0.0036
(b)	Lena image			Goldhill image		
100	44.56	0.9636	0.0141	45.19	0.9636	0.0143
200	38.54	0.9636	0.0142	39.17	0.9636	0.0142
300	35.02	0.9636	0.0143	35.66	0.9636	0.0142
400	32.51	0.9636	0.0143	33.15	0.9636	0.0142
	Fruit image			Rose image		
100	45.06	0.9636	0.0142	45.19	0.9636	0.0144
200	39.03	0.9636	0.0142	39.17	0.9636	0.0142
300	35.51	0.9636	0.0142	35.65	0.9636	0.0142
400	33.01	0.9636	0.0143	33.15	0.9636	0.0142

**Fig. 10** a Original cover image, b–e watermarked image with embedding factor value = 100, 200, 300 and 400

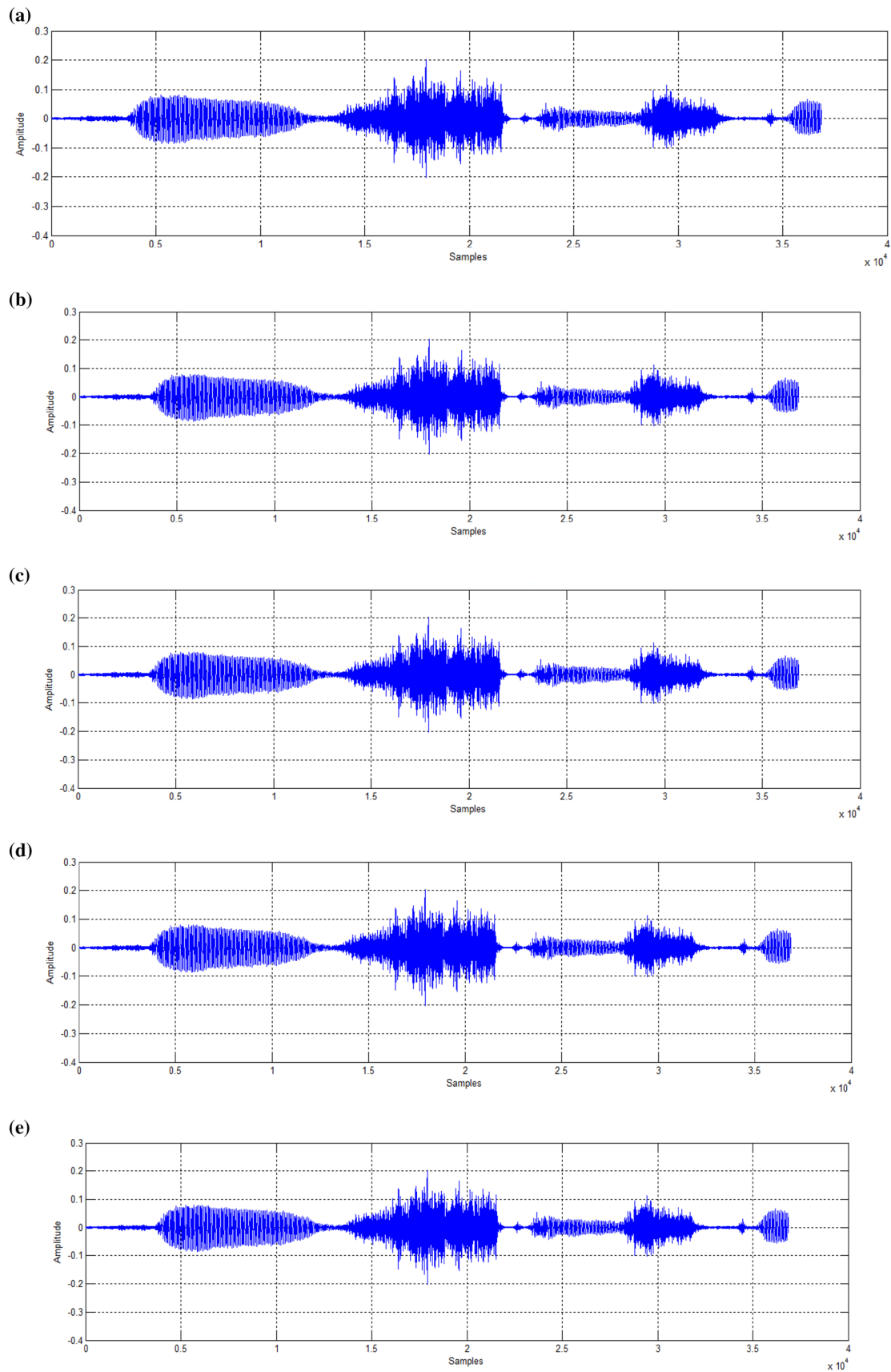


Fig. 11 a Original watermark woman speech signal. b–e Extracted watermark woman speech signal with embedding factor value = 100, 200, 300 and 400

Table 3 SF and WNR values of Proposed Technique under Various Watermarking Attacks

Attacks	Lena image		Goldhill image		Fruit image		Rose image	
	SF	WNR (dB)	SF	WNR (dB)	SF	WNR (dB)	SF	WNR (dB)
JPEG (Q=90)	0.7192	11.2105	0.6663	7.7967	0.7200	12.5444	0.7195	14.5848
JPEG (Q=70)	0.6677	5.6095	0.5635	1.7177	0.6899	7.9381	0.6921	9.1450
JPEG (Q=50)	0.6285	3.4159	0.5102	- 0.1643	0.6685	6.0644	0.6744	7.3117
Gaussian noise ($\sigma=0.0001$)	0.7171	10.7431	0.6835	10.0075	0.7119	10.4072	0.7043	10.4512
Salt and pepper noise ($\sigma=0.0005$)	0.7146	8.9963	0.6693	6.0446	0.7152	10.4162	0.6781	10.6380
Speckle noise ($\sigma=0.0004$)	0.7218	11.6642	0.6831	10.3056	0.7027	9.1392	0.7183	14.3355
Median filter (3×3)	0.6822	7.0522	0.5169	- 0.2362	0.7138	11.2382	0.7059	11.1705
Weiner filter (3×3)	0.6343	3.9346	0.4523	- 1.8521	0.6603	5.8312	0.6860	8.9116
Gaussian LPF (3×3)	0.6921	7.2695	0.6129	3.8680	0.7050	8.8863	0.7074	12.0788
Sharpening	0.7405	0.0000	0.7122	0.0000	0.7367	0.0000	0.7297	0.0000
Histogram Equalization	0.4215	- 12.3327	0.4707	- 11.3079	0.5828	- 4.3949	0.4254	- 20.7131
Cropping	0.4693	- 5.9952	0.4062	- 8.3640	0.5034	- 2.6901	0.5946	0.9177

Fig. 12 Similarity Factor (SF) values of extracted watermark woman speech signal against various watermarking attacks

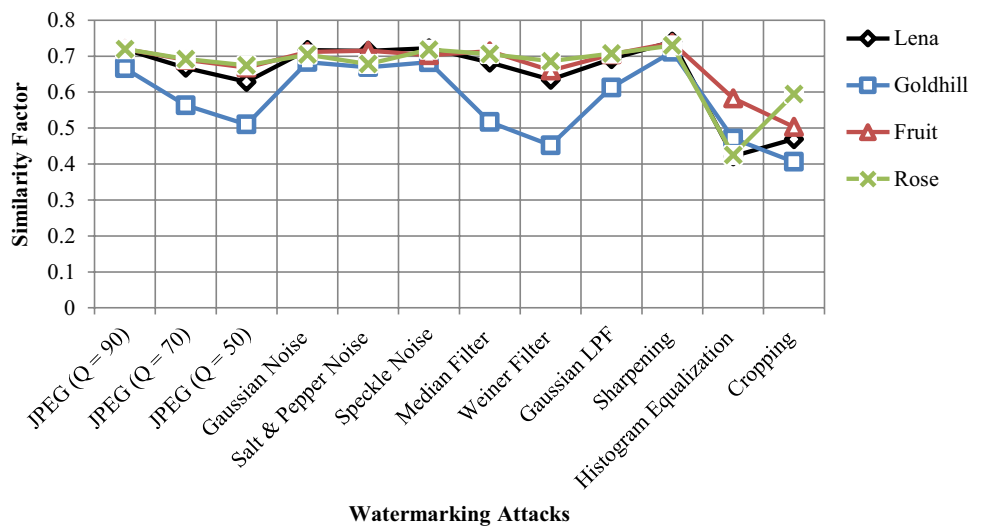


Table 4 Computational Time (s) of Proposed Technique for Different Embedding Factor Values

Test image	Embedding factor k=100		Embedding factor k=200		Embedding factor k=300		Embedding factor k=400	
	T _{Embed} (s)	T _{Extract} (s)	T _{Embed} (s)	T _{Extract} (s)	T _{Embed} (s)	T _{Extract} (s)	T _{Embed} (s)	T _{Extract} (s)
Lena	1.7924	0.3622	1.7393	0.3556	1.7764	0.3499	1.7753	0.4148
Goldhill	1.7972	0.3569	1.7251	0.3488	1.9439	0.3550	1.7641	0.3513
Fruit	2.0627	0.4171	1.8209	0.3560	1.8158	0.3557	1.7508	0.3579
Rose	1.8424	0.3789	1.7937	0.3459	1.7807	0.3564	1.7707	0.3614

parts such as embedding time and extraction time. These two times have importance in this technique because they are part of a continuous procedure. In this paper, the computation time is calculated by summation of embedding time and extraction time. The embedding time is a time taken for embedding watermark speech signal into the cover image to get watermarked image. While the extraction time is a time

taken for extraction of the watermark speech signal from the watermarked image. Table 4 shows the computational time for watermark speech signal embedding and extraction for varying embedding factor values.

The proposed technique is implemented using a machine with 2 GHz processor and 8 GB RAM using MATLAB 2016b software. The average computational time of

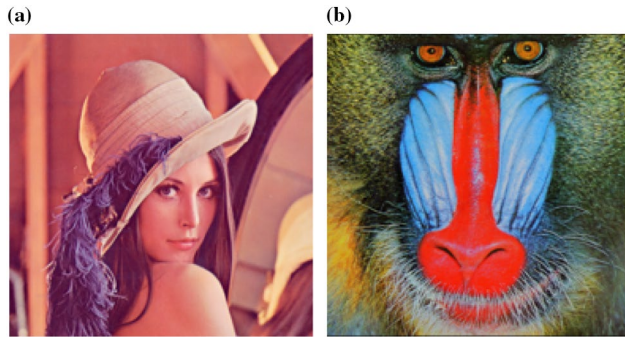


Fig. 13 a color lena image, b color baboon image

watermark speech signal embedding process is 1.8095 s and watermark speech signal extraction process is 0.3640 s. Thus, the average total computational time of the proposed technique is 2.1735 s. This indicates that the usage of four image transforms in the proposed technique does not affect the computational complexity of the watermarking technique. The computational time of the proposed technique is compared with fast existing watermarking techniques (Thakkar and Srivastava 2017; Nguyen et al. 2006) which were used for copyright protection of the image. The computational time of Nguyen et al. (2006) is 11.026 s and computational time of Thakkar and Srivastava (2017) is 5.19 s. This indicates that the proposed technique computes faster than the existing techniques (Thakkar and Srivastava 2017; Nguyen et al. 2006).

Fig. 14 a Watermarked color lena image, b extracted man speech signal

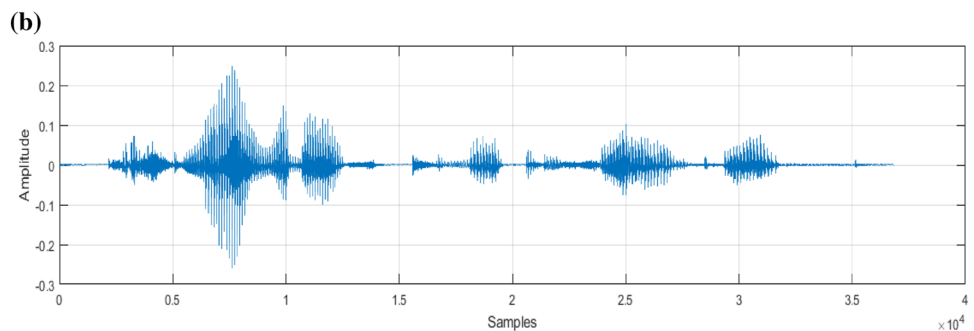


Table 5 PSNR, SF and BER value of Proposed Technique for Color Image

Embedding factor	PSNR (dB)	SF	BER	T_{Embed} (S)	T_{Extract} (S)
Color lena image					
100	45.19	0.9636	0.0143	1.9844	0.3904
200	39.17	0.9636	0.0142	2.0477	0.4289
300	35.65	0.9636	0.0141	2.0242	0.4121
400	33.15	0.9636	0.0143	2.0125	0.4319
Color baboon image					
100	44.92	0.9636	0.0141	2.0569	0.3873
200	38.90	0.9636	0.0141	2.0334	0.4197
300	35.37	0.9636	0.0143	2.0686	0.3907
400	32.88	0.9636	0.0143	2.1047	0.4200

6.4 Performance analysis of proposed technique for color images

The experimental results indicated that this technique works effectively for grayscale images. In addition to this, the performance of the technique is tested for color standard images such as Lena, baboon shown in Fig. 13. A man speech signal is used as watermark information.

Here, the cover color image is broken into three color channels (R-G-B) and the values of the red channel of cover image modifies by watermark speech signal and embedding factor. The remaining process for embedding and extraction of the watermark speech signal in cover

Table 6 SF and WNR values of proposed technique for color images under various watermarking attacks

Attacks	Color lena image		Color baboon image	
	SF	WNR (dB)	SF	WNR (dB)
JPEG (Q=90)	0.5291	4.1092	0.5274	- 0.0592
JPEG (Q=70)	0.4843	1.6189	0.4391	- 2.5917
JPEG (Q=50)	0.5003	0.4294	0.4049	- 3.6210
Gaussian noise ($\sigma=0.0001$)	0.4877	8.8993	0.7149	9.9043
Salt and pepper noise ($\sigma=0.0005$)	0.5256	7.0973	0.7003	13.0941
Speckle noise ($\sigma=0.0004$)	0.5005	6.3490	0.7132	9.7188
Median filter (3×3)	0.4863	7.0455	0.5216	- 1.0000
Weiner Filter (3×3)	0.4800	3.4252	0.5290	- 0.3932
Gaussian LPF (3×3)	0.4830	5.8949	0.6583	4.8739
Sharpening	0.4860	0.4335	0.5092	- 2.6795
Histogram equalization	0.5964	- 16.9379	0.4970	- 8.9952
Cropping	0.4912	- 1.8994	0.6840	6.2186

image is the same as in Sect. 5. Figure 14 shows a watermarked color image and extracted watermark speech for embedding factor value = 100. The robustness of the proposed technique for color images is verified under various watermarking attacks. Table 5 shows the quality measures of the proposed technique for color images. The robustness results of the proposed technique for color images are given in Table 6.

The robustness results of Table 6 show that the proposed technique performs equally well for the color images. The results also show that the imperceptibility of the proposed technique is also good for color images.

6.5 Performance analysis of proposed technique for TIMIT database

The experimental results indicated that this technique works effectively for grayscale images and color images. In addition to this, the performance of the technique is tested for one of standard speech databased such as TIMIT database (Zue et al. 1990). The database contains speeches of 630 American English speakers. This database was generated and designed by the Massachusetts Institute of Technology (MIT), SRI International and Texas Instruments Inc. (TI) in 1990. Here, for testing of this proposed technique, four speech signals from the database are taken as watermark speech signal which is given in Fig. 15.

The remaining process for embedding and extraction of a watermark speech signal in cover image is the same as in Sect. 5. Figure 16 shows a watermarked Lena image and extracted watermark speech for embedding factor value = 100. The robustness of the proposed technique for TIMIT database speech signals is verified under various watermarking attacks. Table 7 shows the quality measures of the proposed technique for various TIMIT database speech

signal. Here, grayscale Lena image is taken as a cover image while values of embedding factor k are taken as 100 and 400, respectively. The robustness results of the proposed technique for TIMIT database speech signals are given in Table 8. The robustness results are generated using embedding factor $k=400$.

The results of Table 7 shows that the proposed technique is equally performed for standard speech database like the TIMIT database. The robustness results of Table 8 show that the proposed technique provides robustness to TIMIT database speech signals. The results of Table 8 also show that the proposed technique provides partially robustness to TIMIT database speech signals against the histogram equalization attack, Weiner filter, and cropping attack.

6.6 Performance analysis of proposed technique for StirMark benchmark

In this section, performance analysis of proposed techniques for StirMark benchmark is given. Here, StirMark benchmark attacks were described by Steinebach et al. (2001) which are applied to the watermark speech signal and measured distortion has taken place in an extracted watermark speech signal using the proposed technique. This benchmark is tested embedding strength and watermark audibility strength of the given algorithm. The StirMark benchmark performance of the proposed technique is compared with standard benchmark reported in Steinebach et al. (2001) and is summarized in Table 9. The comparison results show that the proposed technique fulfilled requirements of StirMark benchmark test.

6.7 False positive test (FPT) of proposed technique

The results of this test are useful in evaluating the security requirements of the proposed technique. A false positive is a

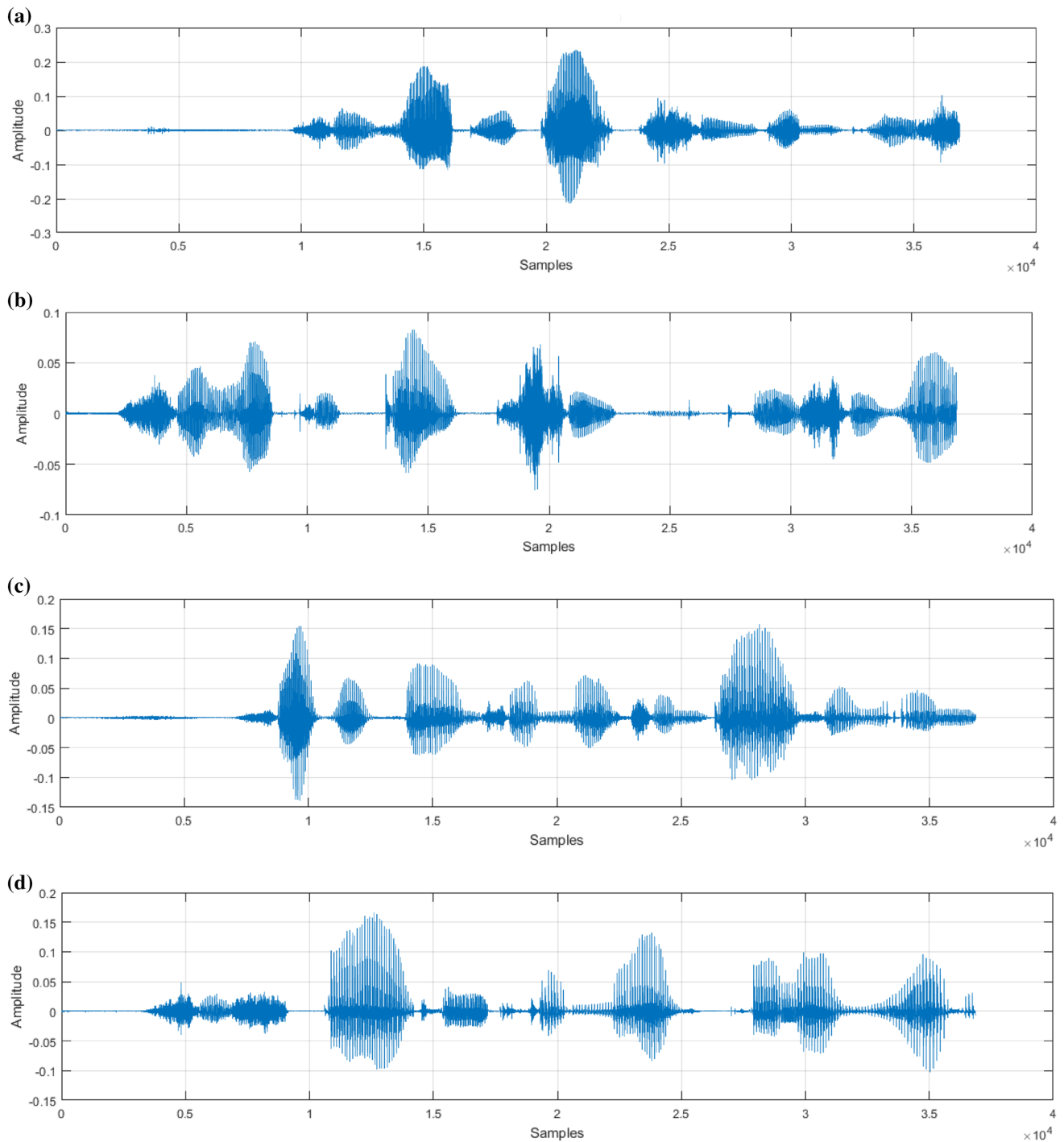


Fig. 15 Sample speech signals from TIMIT database **a** SP1, **b** SP2, **c** SP3, **d** SP4

result of the extraction of a watermark from an unauthorized image, which doesn't actually belong to the owner (Thanki et al. 2017a). Since false positives encourage malicious owners in claiming other unauthorized images, this problem should be avoided. Given multiple non-watermarked images from SIPI database and a watermark woman speech signal, the proposed watermark extraction algorithm is applied on

each one of them using the same security keys to extract the watermark speech signal from non-watermarked images. The sample subjective and objective results for 4 non-watermarked images are shown in Table 10. For testing on standard image database, by assuming that a false positive arises, if the SF value of such extracted watermark speech is above than 0.4 and BER value, is near to zero; a false positive

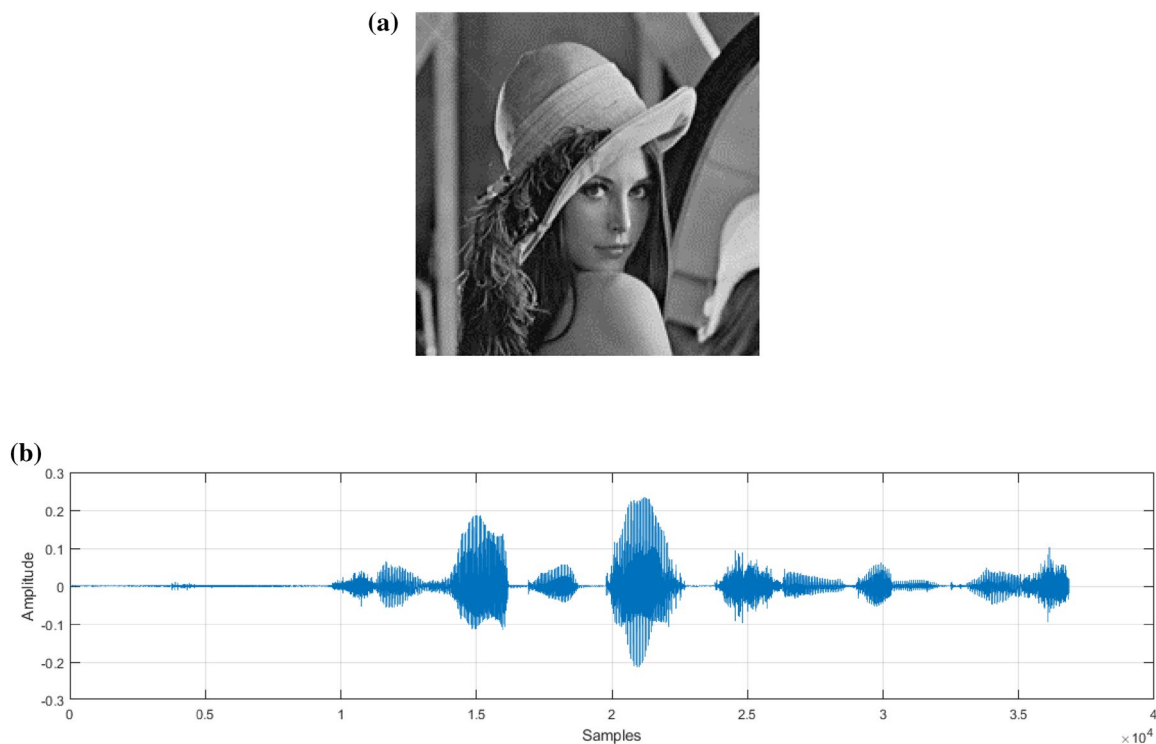


Fig. 16 a Watermarked lena image, b extracted SP1 speech signal

Table 7 PSNR, SF and BER value of proposed technique for TIMIT database speech signals (a) For embedding factor $k=100$. (b) For embedding factor $k=400$

Speech signals	PSNR (dB)	SF	BER	T_{Embed} (s)	$T_{Extract}$ (s)
(a)					
SP1	41.8527	0.9796	0.0130	1.2840	0.3044
SP2	48.9021	0.8998	0.0306	1.5271	0.2901
SP3	45.4420	0.9537	0.0133	0.9026	0.2639
SP4	45.8191	0.9509	0.0215	1.1248	0.3331
(b)					
SP1	42.6319	0.9796	0.0129	0.9530	0.2652
SP2	49.5416	0.8998	0.0310	0.9049	0.2488
SP3	46.1183	0.9537	0.0133	0.9660	0.2565
SP4	46.5012	0.9509	0.0215	0.8667	0.2281

ration of zero is obtained for the proposed technique, when tested on around 100 images of SIPI database. According to Cox et al. (2000), “A false positive rate of 10^{-6} can meet the security requirements” and hence this proposed technique can meet the security requirements of watermarking.

6.8 Security analysis of proposed technique

In this paper, in order to provide security of watermark speech signal, hybridization of discrete cosine transform

(DCT) and singular value decomposition (SVD) is used with embedding factor k . When a watermark speech signal is extracted from the watermarked image, imposter can't get watermark speech signal without the security keys. Hence, the proposed technique has high security. Figure 17 shows the hybrid coefficients of the watermark speech signal using DCT and SVD.

The reason behind using hybrid coefficients of watermark instead of encrypted watermark for the security of watermark speech signal in this technique is that hybrid coefficients of each watermark speech signal differs and depends on behavioral characteristics of a person. While in encryption, fix type of encrypted key is applied to watermark speech signal to get an encrypted signal. Once encrypted key becomes public then the security of technique will be compromised. That ways, dynamic nature of hybrid coefficients of the watermark speech signal is better than the static encrypted key for generation secured watermark speech signal.

Further, for security analysis of proposed technique, testing of technique is done using the U and V matrix of other watermark speech signal for extraction of original watermark speech signal at extraction side. For this purpose, the woman speech signal image is inserted into cover image to generate watermarked image using proposed scheme. At extraction, first, extract singular value of watermark speech signal using this scheme. After that, U and V matrix of other

Table 8 SF and WNR values of proposed technique for TIMIT database speech signals

Attacks	SP1 Signal		SP2 Signal		SP3 Signal		SP4 Signal	
	SF	WNR (dB)	SF	WNR (dB)	SF	WNR (dB)	SF	WNR (dB)
JPEG (Q=90)	0.7063	10.0370	0.5485	2.9255	0.6489	6.4893	0.6398	6.1773
JPEG (Q=70)	0.6442	4.4382	0.4166	- 2.5993	0.5493	0.9529	0.5353	0.6348
JPEG (Q=50)	0.5992	2.2493	0.3599	- 4.8168	0.4863	- 1.2529	0.4777	- 1.5599
Gaussian noise ($\sigma=0.0001$)	0.7040	9.4435	0.5410	2.1777	0.6443	5.7510	0.6335	5.5018
Salt & pepper noise ($\sigma=0.0005$)	0.7087	7.5654	0.5217	- 0.3976	0.6020	3.8611	0.6038	3.1519
Speckle Noise ($\sigma=0.0004$)	0.7098	10.4897	0.5578	3.2862	0.6550	6.8964	0.6462	6.5645
Median filter (3×3)	0.6619	5.9444	0.4553	- 0.8838	0.5825	2.6379	0.5698	2.3299
Weiner filter (3×3)	0.6050	2.8357	0.3668	- 3.8681	0.5044	- 0.4084	0.4883	- 0.7184
Gaussian LPF (3×3)	0.6716	6.0773	0.4688	- 1.0372	0.5927	2.5492	0.5804	2.2350
Sharpening	0.7330	0.0000	0.6380	0.0000	0.7017	0.0000	0.6955	0.0000
Histogram equalization	0.3695	- 13.6203	0.1976	- 21.0586	0.2797	- 17.3997	0.2705	- 17.7093
Cropping	0.4171	- 7.4366	0.2625	- 14.9724	0.3369	- 11.2410	0.3305	- 11.5512

Table 9 Performance analysis of StirMark benchmark of proposed technique

Attacks	Percentage of destroyed watermark audio defined by StirMark Benchmark (Steinebach et al. 2001) (%)	Percentage destroyed watermark audio obtained by proposed technique (%)
Noise (Max. value = 50)	33	30
Noise (Max. value = 300)	66	60
High pass filter	18	25
Low pass filter	66	60
Delay	17	22
L/R split	50	55
Chorus	33	38
Flanger	66	70
Enhance	33	30
Compressor	33	30
Pitch shift	83	80
Resampling	100	95
Time stretch	83	78
Zero-cross-inserts	100	96
Copy samples	83	80
Flip samples	50	45
Cut samples	83	80

speech signal such as men speech signal is used to extract actual woman speech signal. But the result in Fig. 18 shows that the extraction of woman speech signal is not possible by using U and V matrix values of men speech signal. That's situation indicated that this technique is also provide security against impostor manipulation such as wrong presentation of U and V matrix.


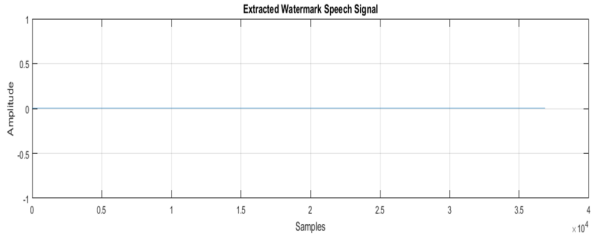

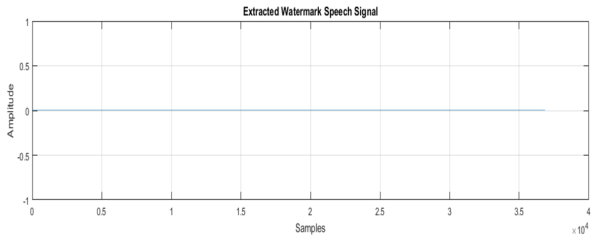

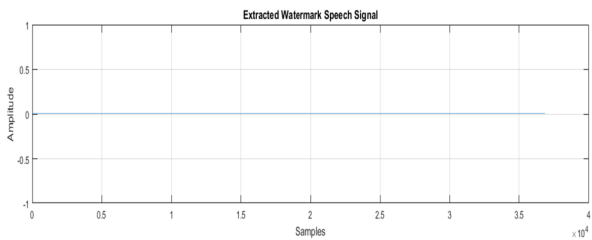
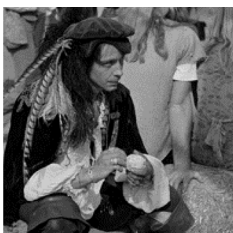
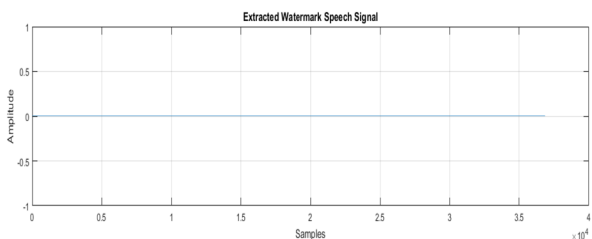
6.9 Comparison of proposed technique with existing techniques

The SF values comparison between the proposed technique and Jundale technique (Jundale and Patil 2010) which was the first paper to used human speech signal as watermark information for grayscale Lena image under attacks are given in Table 11. Referring the Table 11, the proposed technique has been provided better robustness against noise attacks and compression attack compared to Jundale technique (Jundale and Patil 2010).

The comparative comparison of proposed technique with existing techniques (Talbi et al. 2017; Inamdar et al. 2010; Jundale and Patil 2010) by different parameters are given in Table 12. Here, the complexity of the watermark algorithms is compared based on the used transform for speech signal embedding and the provision of security to a watermark speech signal. Based on these two parameters, the proposed technique is more complex compared to Talbi technique (Talbi et al. 2017), Inamdar technique (Inamdar and Rege 2014) and Jundale technique (Jundale and Patil 2010) as all these techniques used either one or two transforms only for watermark embedding. The proposed technique is more secure than the existing techniques (Talbi et al. 2017; Inamdar et al. 2010; Jundale and Patil 2010) as it embeds hybrid coefficients of the watermark speech signal instead of its actual value. The quality measure values of proposed technique indicate that this technique performed better than existing watermarking techniques.

The comparison of the proposed technique with the Renza technique (Renza et al. 2016) for the security of color image using speech watermark signal is given in Table 13. The comparison shows that this technique performs better than the Renza technique (Renza et al. 2016).

Table 10 Results of false positive test (FPT) for proposed technique

Sl. No	Non-watermarked Image	Extracted Watermark Speech Signal using Proposed Technique	SF	BER
1			0.0000	0.9856
2			0.0000	0.9856
3			0.0000	1.0000
4			0.0000	0.9856

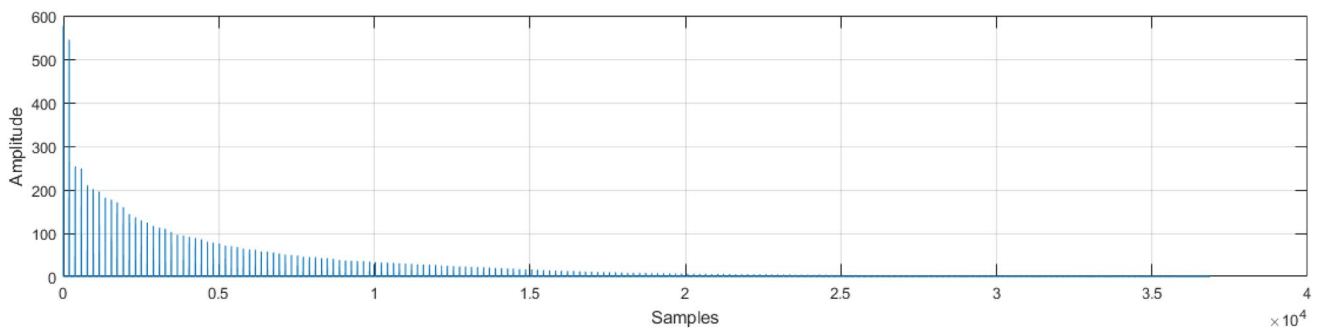


Fig. 17 Secured hybrid coefficients of watermark speech signal using security keys

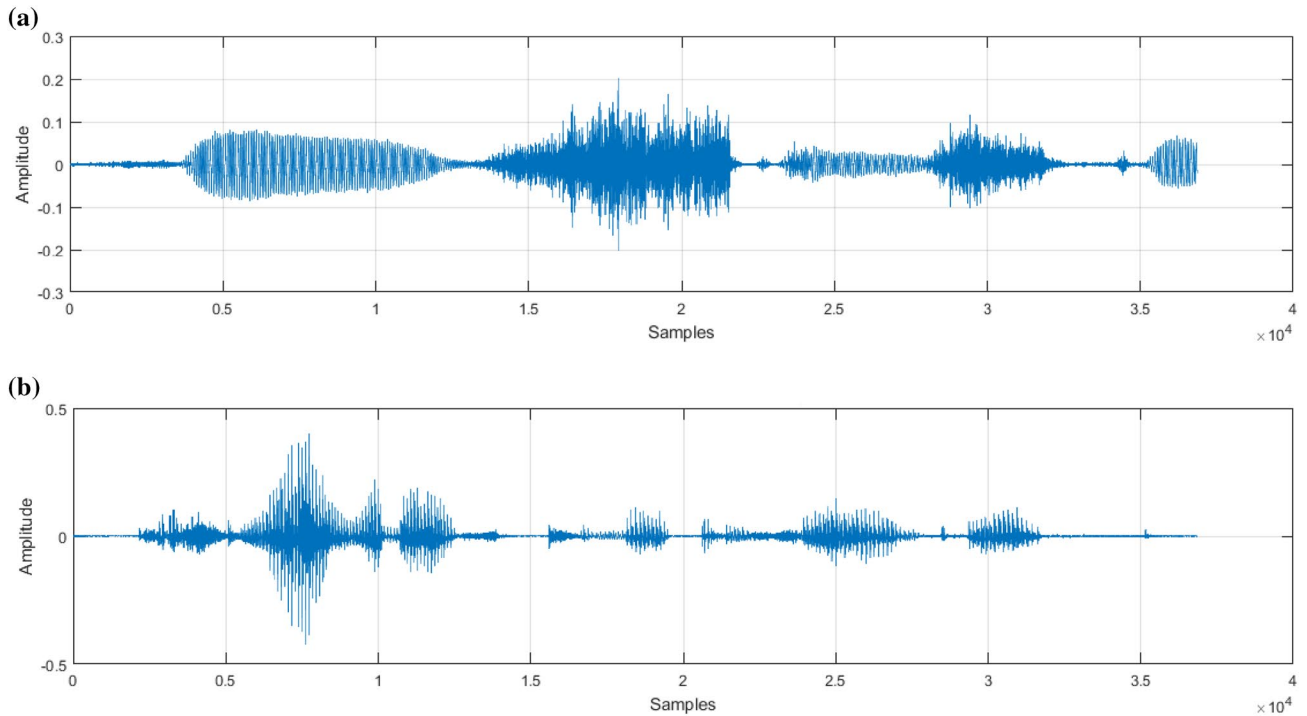


Fig. 18 Security analysis of proposed scheme, **a** actual embed watermark woman speech signal, **b** extracted watermark woman speech signal with SF=0.0016 and BER=0.9964

Table 11 SF values of the proposed technique and jundale technique [43]

Attacks	SF for watermark speech signal	
	Jundale technique (Jundale and Patil 2010)	Proposed technique
Cropping	0.76	0.47
Salt and pepper noise	0.51	0.71
Histogram equalization	0.85	0.42
Gaussian noise	0.40	0.72
Median filtering	0.67	0.68
Sharping	0.83	0.74
JPEG (Q=50)	0.45	0.63

7 Security of speech signal at storage of biometric system: possible application of proposed technique

In this paper, a watermarking technique is proposed for the security of images using the speech signal. This technique also provides security to watermark a speech signal using different keys. One possible application of the proposed technique is shown in Fig. 19. This scenario shows the application of the proposed technique for the storage of individual speech signal in the biometric image database.

In this application, when an individual presents his/her speech signal to the system sensor, then the query speech signal is compared with the enrolled speech signal. The

Table 12 Comparison of proposed technique with existing techniques

Features	Jundale technique (Jundale and Patil 2010)	Inamdar technique (Inamdar and Rege 2014)	Talbi technique (Talbi et al. 2017)	Proposed technique
Watermarking domain	Wavelet	Hybrid	Frequency	Hybrid
Security provided to watermark speech signal	Not reported	Not reported	Not reported	SVD and DCT
Maximum PSNR (dB)	41.63	34.028	24.819	45.19
Maximum SF	0.8713	0.9775	0.7843	0.9843

Table 13 Comparison of proposed technique with Renza technique

Features	Renza technique (Renza et al. 2016)	Proposed technique
Processing domain	Frequency	Hybrid
Security provided to watermark speech signal	QIM	SVD and DCT
Similarity between color watermarked image	0.8240	0.9991
Similarity between extracted watermark Speech Signal	95%	96.36%

watermarked image with embedded watermark speech signal is given to watermark extractor to get his/her enrolled speech signal. The extracted enrolled speech signal is given to the matcher module for speech authentication of his/her speech signal. In Fig. 19, suppose user 1 is the employer of x company and have all keys for extraction of his/her speech signal from the watermarked image. The user 2 and user 3 have different keys and try to enter into the company as an imposter using these keys. When user 1 sends a request to system database about watermarked image with his/her speech signal embedded then the system database sends watermarked biometric image to user 1, where the user 1 tries to extract his/her speech signal using all three keys. Then a comparison of the extracted speech signal with his/her query speech signal is performed using the matcher module. If the result of the comparison is greater than the

predefined system threshold value then the user 1 is allowed to enter the company. This scenario is shown in Fig. 19 as dotted blue portion.

When imposter sends a fake request about watermarked image, then system database sends the watermarked image to an imposter. Then, imposter tries to extract enrolled speech signal of someone from a watermarked image without proper knowledge of the secret keys. There are various scenarios shown in Fig. 19 as dotted red portion where user 2 has knowledge of key 1, key 2 but not key 3. The user 3 has knowledge of key 1 and key 3 but not key 2. The user 2 can get the singular value of DCT coefficients of a watermark speech signal using key 1, key 2 but cannot get actual watermark speech signal from watermarked image. Without the knowledge of enrolled speech signal, speech signal matching is not performed and user 2 is not allowed to enter into the

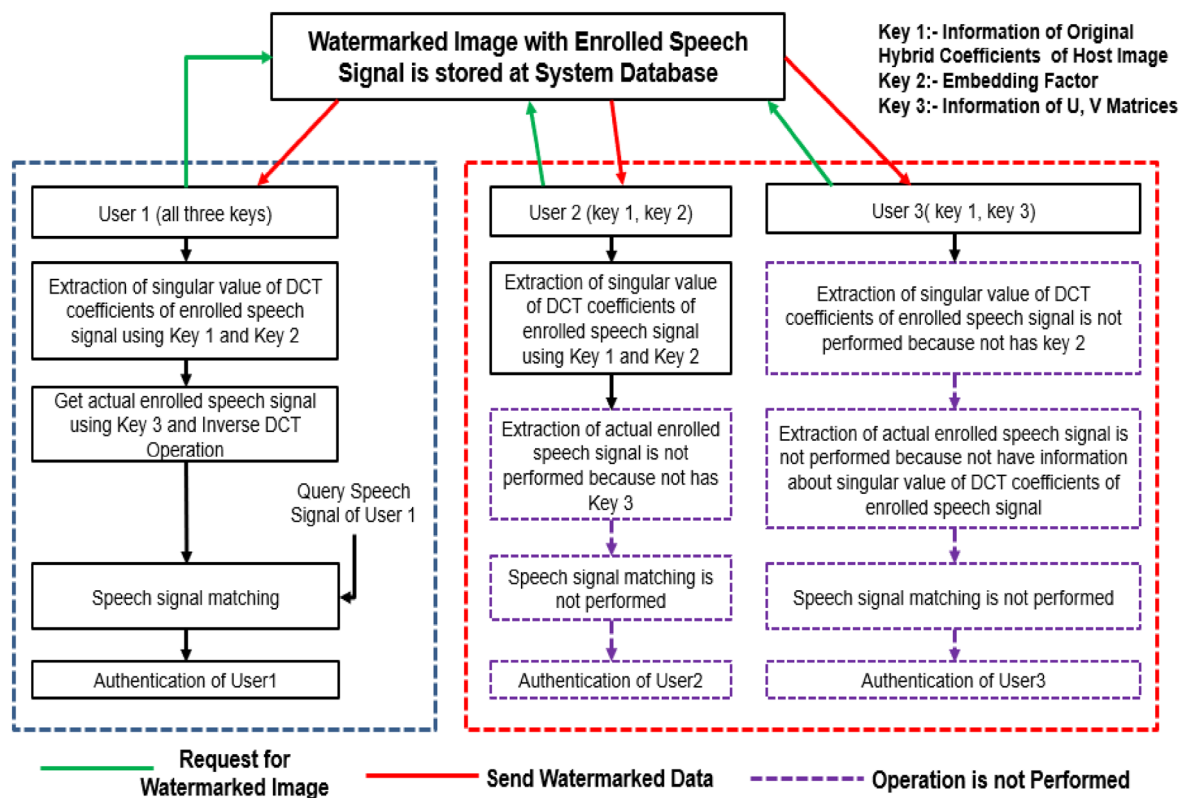


Fig. 19 Application scenario of proposed technique

company. Similarly, the user 3 cannot get the singular value of DCT coefficients of the watermark speech signal as he has no information of hybrid coefficients of the original cover image. The user 3 is also not allowed to enter the company because he does not have enrolled speech signal for speech signal matching. This indicates that the proposed technique may provide security to human speech signal against spoof attack at system database.

Other possible practical application of this proposed technique may be in biometric as a service (BAAS) in the cloud. In the cloud service, various biometric data are stored in the cloud server. Therefore, this proposed technique can be providing a solution for the security of biometric data on the web server.

8 Conclusions and future direction

This paper proposed a novel hybrid domain image watermarking technique using watermark speech signal. For non-blind extraction of the watermark speech signal, original hybrid coefficients of cover image and information of U, V matrices of the watermark speech signal are required as secret keys. The comparative analysis shows that this proposed technique performs better than existing techniques for all requirement of watermarking techniques. This proposed technique may be used for copyright protection of images as well as speech signal security at the storage of the biometric system. The performance of the proposed technique is tested for various types of multimedia data such as grayscale images, color images, TIMIT database speech signals. The experimental results of the proposed technique show that this technique can be applied to the security of various types of multimedia data. The limitation of the proposed technique is that when embedding factor value increases, the noise distortion appears in the watermarked image. In the future, the design of hardware architectures for hardware implementation of the proposed technique can be considered.

References

- Agarwal N, Singh AK, Singh PK (2019) Survey of robust and imperceptible watermarking. *Multimedia Tools Appl*. <https://doi.org/10.1007/s11042-018-7128-5>
- Ali Z, Hossain MS, Muhammad GM, Aslam M (2018) New zero-watermarking algorithm using hurst exponent for protection of privacy in telemedicine. *IEEE Access* 6:7930–7940
- Bhat V, Sengupta I, Das A (2010) An adaptive audio watermarking based on the singular value decomposition in the wavelet domain. *Digit Signal Process* 20(6):1547–1558
- Borra S, Swamy GN (2013) Sensitive digital image watermarking for copyright protection. *Int J Netw Secur (IJNS)* 15(2):95–103
- Borra S, Lakshmi HR, Dey N, Ashour A, Shi F (2017) Digital image watermarking tools: state-of-the-art. *Front Artif Intell Appl* 296:450–459
- Candes E, Donoho D (2004) New tight frames of curvelets and optimal representations of objects with piecewise-C2 singularities. *Commun Pure Appl Math* 57:219–226
- Candes E, Demanet L, Donoho D (2006) Fast discrete curvelet transforms. *Multiscale Modell Simul* 5(3):861–899
- Cox IJ, Miller ML, Bloom JA (2000) Watermarking applications and their properties. In: *Proceedings international conference on Information technology: coding and computing, 2000*. IEEE. pp. 6–10
- Edward S, Sumanthi S, Ranihemamalani R (2011) Person authentication using multimodal biometrics with watermarking. In: *2011 International conference on signal processing, communication, computing, and networking technologies*, pp 100–104
- El-Gazar S, Abbas AM, El-Dolil S, El-Dokany IM, Dessouky MI, El-Rabaie ESM, El-Samie FEA (2018) Efficient SVD speech watermarking with encrypted images. *Int J Speech Technol* 21(4):953–965
- Feng G, Lin Q (2007) Iris feature based watermarking algorithm for personal identification. *Int Symp Multispectral Image Process Pattern Recognit*. <https://doi.org/10.1117/12.748180>
- Golub GH, Reinsch C (1970) Singular value decomposition and least squares solutions. *Numer Math* 14(5):403–420
- Inamdar V, Rege P (2014) Dual watermarking technique with multiple biometric watermarks. *Sadhana* 39(1):3–26
- Inamdar V, Rege P, Arya M (2010) Offline handwritten signature based blind biometric watermarking and authentication technique using biorthogonal wavelet transform. *Int J Comput Appl* 11(1):19–27
- Jain A (1989) *Fundamentals of digital image processing*. Prentice Hall Inc., New Jersey, pp 150–153
- Jain A, Kumar A (2012) Second generation biometrics, the ethical, legal and social context. In: Mordini E, Tzovaras D (eds) *Biometric recognition: an overview*. Springer, Berlin, pp 49–79. https://doi.org/10.1007/978-94-007-3892-8_3
- Jain A, Uludag U (2002) Hiding fingerprint minutiae in images. In: *Proceedings of 3rd workshop on automatic identification advanced technologies*, pp 97–102
- Jain A, Uludag U (2003) Multimedia content protection via biometrics-based encryption. In: *Proceedings of IEEE ICME'03*, pp III-237
- Jain A, Uludag U (2003b) Hiding biometric data. *IEEE Trans Pattern Anal Mach Intell* 25(11):1494–1498
- Jain A, Uludag U, Hsu R (2002) Hiding a face in a fingerprint image. In: *Proceedings of IEEE 16th international conference on pattern recognition*, vol 3, pp 756–759
- Jain A, Ross A, Prabhakar S (2004) An introduction to biometric recognition. *IEEE Trans Circ Syst Video Technol* 14(1):4–20
- Jundale V, Patil S (2010) Biometric speech watermarking technique in images using wavelet transform. *IOSR J Electr Commun Eng* 33–39
- Kaur M, Girdhar A, Kaur M (2010) Multimodal biometric system using speech and signature modalities. *Int J Comput Appl* 5(12):13–16
- Ko T (2005) Multimodal biometric identification for large user population using fingerprint, face and iris recognition. In: *Proceedings of the 34th workshop on applied imagery and pattern recognition*. <https://doi.org/10.1109/aipr.2005.35>
- Kumar C, Singh AK, Kumar P (2018) Improved wavelet-based image watermarking through SPIHT. *Multimedia Tools Appl* 15:1–14. <https://doi.org/10.1007/s11042-018-6177-0>
- Kutter M, Petitcolas F (1999) A fair benchmark for image watermarking systems *Electronic imaging '99. Secur Watermarking Multimedia Contents* 3657:1–14
- Langelar G, Setyawan I, Lagendijk R (2000) Watermarking of digital image and video data—a state of art review. *IEEE Signal Process Mag* 20–46

- Li S, Song S, Lu W, Sun D, Wei J (2017) Parameterization of LSB in self-recovery speech watermarking framework in big data mining. *Secur Commun Netw* 2017:1–12. <https://doi.org/10.1155/2017/3847092>
- Mani M, Lakshmi T (2013) Speech watermarking using logarithmic approach. *Int J Adv Res Electr Instrum Eng* 2(10):5007–5011
- Merrad A, Saadi S (2018) Blind speech watermarking using hybrid scheme based on DWT/DCT and sub-sampling. *Multimedia Tools Appl* 77(20):27589–27615
- Merrad A, Benziane A, Saadi S, Hafaifa A (2018a) Robust blind approach for digital speech watermarking. In: 2018 2nd international conference on natural language and speech processing (ICNLSP). IEEE, pp 1–5
- Nematollahi M, Al-Haddad S, Zarafshan F (2015) Blind digital speech watermarking based on eigen-value quantization in DWT. *J King Saud Univ Comput Inf Sci* 27(1):58–67
- Nematollahi M, Vorakulpipat C, Gamboa Rosales H (2017a) Semi fragile Speech watermarking based on least significant bit replacement of line spectral frequencies. *Math Probl Eng* 2017:1–9. <https://doi.org/10.1155/2017/3597695>
- Nematollahi MA, Vorakulpipat C, Gamboa Rosales H (2017b) Optimization of a blind speech watermarking technique against amplitude scaling. *Secur Commun Netw* 2017:1–13. <https://doi.org/10.1155/2017/5454768>
- Nguyen C, Tay D, Deng G (2006) A fast watermarking system for H.264/AVC video. In: Asia specific IEEE conference on circuits and systems, pp 81–84
- Noore A, Singh R, Vatsa M, Houck M, Morris K (2006) Robust biometric image watermarking for fingerprint and face template protection. *IEICE Electron Express* 3(2):23–28
- Noore A, Singh R, Vatsa M, Houck M (2007) Enhancing security of fingerprints through contextual biometric watermarking. *Forensic Sci Int* 169(2):188–194
- Speech Signal database: <http://freesound.org>. Accessed 2017
- Park K, Jeong D, Kang B, Lee E (2007) A study on iris feature watermarking on face data. In: Adaptive and natural computing algorithms, pp. 415–423. https://doi.org/10.1007/978-3-540-71629-7_47
- Patel R, Sharwankar U, Thakare V (2011) Secure transmission of password using speech watermarking. *Int J Comput Sci Technol* 2(3):315–318
- Ratha N, Connell J, Bolle R (2001) Enhancing security and privacy in biometric based authentication systems. *IBM Syst J* 40(3):614–634
- Rege P (2012) Biometric watermarking. National Seminar on Computer Vision and Image Processing, Rajkot
- Renza D, Dora L, Sanchez J (2016) Highly transparent steganography scheme of speech signals into color images using quantization index modulation. In: Mexican conference on pattern recognition. Springer, Champ, pp 241–250
- Revathi A, Sasikaladevi N, Jeyalakshmi C (2018) Digital speech watermarking to enhance the security using speech as a biometric for person authentication. *Int J Speech Technol* 21(4):1021–1031
- Saxena P, Khandelwal Y, Khandelwal R (2017) Haar transform for the numerical solutions of ordinary differential equations and boundary value problem with maple. *Int J Eng Manag Sci* 4(4):8–12
- Singh AK (2019) Robust and distortion control dual watermarking in LWT domain using DCT and error correction code for color medical image. *Multimedia Tools Appl*. <https://doi.org/10.1007/s11042-018-7115-x>
- Steinebach M, Petitcolas FA, Raynal F, Dittmann J, Fontaine C, Seibel S, Ferri LC (2001) StirMark benchmark: audio watermarking attacks. In: Proceedings international conference on information technology: coding and computing, 2001. IEEE, pp. 49–54
- Talbi M, Fatima SB, Cherif A (2017) Speech modulation for image watermarking. In: 2017 International conference on control, automation and diagnosis (ICCAD). IEEE, pp 522–527
- Thakkar F, Srivastava V (2017) A fast watermarking algorithm with enhanced security using compressive sensing and principle components and its performance analysis against a set of standard attacks. *Multimedia Tools Appl* 76(14):15191–15219
- Thakur S, Singh AK, Ghrera SP, Elhoseny M (2018a) Multi-layer security of medical data through watermarking and chaotic encryption for tele-health applications. *Multimedia Tools Appl* 78(3):3457–3470
- Thakur S, Singh AK, Ghrera SP (2018b) NSCT domain-based secure multiple-watermarking technique through lightweight encryption for medical images. *Concurr Comput Pract Exp* 15:5108. <https://doi.org/10.1002/cpe.5108>
- Thanki R, Borisagar K (2017) Watermarking scheme with CS encryption for security and piracy of digital audio signals. *Int J Inf Syst Model Des (IJISMD)* 8(4):38–60
- Thanki RM, Kothari AM (2017) Digital watermarking: technical art of hiding a message. In: Intelligent analysis of multimedia information. IGI Global, pp 431–466
- Thanki R, Borra S, Dwivedi V, Borisagar K (2017a) An efficient medical image watermarking scheme based on FDCuT–DCT. *Eng Sci Technol Int J* 20(4):1366–1379
- Thanki R, Dwivedi V, Borisagar K, Borra S (2017b) A Watermarking algorithm for multiple watermarks protection using RDWT-SVD and compressive sensing. *Inform Int J Comput Inf* 41(4):479–493
- Thanki R, Borra S, Brisagar K (2018a) A Hybrid watermarking technique for copyright protection of medical signals in teleradiology. handbook of research on information security in biomedical signal processing, pp. 320–349
- Thanki R, Borisagar K, Borra S (2018b) Speech watermarking technique using the finite ridgelet transform, discrete wavelet transform, and singular value decomposition. In: Advance compression and watermarking techniques for speech signals. Springer, Cham, pp 27–45
- Tsai SE, Yang SM (2018) An effective watermarking method based on energy averaging in audio signals. *Math Probl Eng* 2018:1–8. <https://doi.org/10.1155/2018/6420314>
- University of South Carolina SIPI Image Database: <http://sipi.usc.edu/database/database.php>. Accessed 2017
- Vatsa M, Singh R, Mitra R, Noore A (2004) Digital watermarking based secure multimodal biometric system. In: Proceedings of 2004 IEEE international conference on systems, man and cybernetics, pp. 2983–2987
- Vatsa M, Singh R, Noore A (2009) Feature based RDWT watermarking for multimodal biometric system. *Image Vis Comput* 27(3):293–304
- Vidakovic B (1999) Statistical modelling by wavelets. Wiley, Hoboken, pp 115–116
- Yan J (2009) Wavelet matrix. Department of Electrical and Computer Engineering, University of Victoria, Victoria, BC, Canada, November, 1–2
- Zhong J (2007) Watermark embedding and detection. arXiv preprint [arXiv:0706.0427](https://arxiv.org/abs/0706.0427)
- Zue V, Seneff S, Glass J (1990) Speech database development at MIT: tIMIT and beyond. *Speech Commun* 9(4):351–356