Video watermarking using neural networks

S. Bhargavi Latha*

CSE Department,
Gokaraju Rangaraju Institute of Technology,
Hyderabad, India
Email: s.bhargavilatha@gmail.com
*Corresponding author

D. Venkata Reddy

ECE Department,
Mahatma Gandhi Institute of Technology,
Hyderabad, India
Email: dasarireddy@yahoo:com

A. Damodaram

CSE Department,
SIT,
Hyderabad, India
Email: damodarama@rediffmail.com

Abstract: Copyright protection for videos is important to prevent revenue loss for video generation companies by using video watermarking methods. Though many methods exists, still certain scope is noticed in robust video watermarking methods. Achieving features like trade-off between robustness and imperceptibility, speed, blind watermarking simultaneously is very challenging. The proposed work achieves the above said features using log-polar, DWT, and SVD techniques to embed watermark in a video and extract it when necessary. The objective is to protect the copyright and make the watermarking system blind, robust against frame drop attacks as well as achieving above features. This work also leverages scrambling, deep learning-based approach to generate secret sharing image from watermark to improve the speed compared to conventional tabular-based approach. We evaluated the method on our own dataset and proved that this method is outperforming compared to state-of-the-art methods in DWT and SVD domain.

Keywords: watermark; deep-neural network; DWT; SVD; scrambling; secret sharing.

Reference to this paper should be made as follows: Latha, S.B., Reddy, D.V. and Damodaram, A. (2021) 'Video watermarking using neural networks', *Int. J. Information and Computer Security*, Vol. 14, No. 1, pp.40–59.

Biographical notes: S. Bhargavi Latha is an Assistant Professor of Computer Science and Engineering at GRIET is presently engaged in pursuit of her PhD in Image Processing registered with JNTU, Hyderabad. Prior to PhD, she had earned her Master's of Technology in Computer Science and Engineering and

Bachelors of Technology in Electronics and Communication and Engineering. She has 12 years of experience in teaching. Her areas of interest include image processing, internet of things and network security. She has published eight research articles in leading journals and conference proceedings in the area of image watermarking and video watermarking.

D. Venkata Reddy received his Master's and PhD in Electronics and Communications from JNTUH. Presently, he is a Professor in Electronics and Communications Engineering and in-charge for campus placements in Mahatma Gandhi Institute of Technology, Hyderabad. His research interests include multi-valued logic, digital design, digital signal processing, and image processing. He has published more than 25 research articles in leading journals and conference proceedings in the area of digital design, digital signal processing, image and video watermarking. He is a member of IEEE Signal Processing Society and Computer Society, member of IETE and member of ISTE.

A. Damodaram received his BTech (CSE), MTech (CSE) and PhD degrees from Jawaharlal Nehru Technological University (JNTU). He has been serving JNTU since 1989 performed distinguished services for the university, as a Professor, the Head of the Department, Vice Principal, Director of UGC-Academic Staff College, Director School of Continuing and Distance Education, Director, University Academic Audit Cell, and Director, Academic and Planning. He Worked as a Professor and Vice Chancellor in Sri Venkateshwara University, Tirupati (Andhra Prahesh). His research interests include image processing, pattern recognition, network security, steganography and digital watermarking. He is a life member of various professional bodies.

1 Introduction

Protecting copyright of multimedia content is very mandate and important due to ubiquitous usage of data-sharing tools. This prevents or discourages illegal sharing or misuse of data. Multimedia content needed to be authenticated for copyright when necessary. So, there is a requirement of protection mechanism for it. One such approach is watermarking the content. Watermarking means insertion of unique information, might be customer unique code, i.e., logo, id, etc., into a multimedia and extraction of the same when necessary. To make watermark useful, it should obey few important properties like robustness against attacks: geometric (scaling, cropping, and rotation), frame dropping, filtering, compression, etc. and also watermark imperceptibility. Though many methods (Cox et al., 2008; Boland et al., 1995; van Schyndel et al., 1994; Hu et al., 2014; Boreiry and Keyvanpour, 2017; Kumar et al., 2016) exist towards this, still there is huge scope of research of robust watermarking methods. Hence, we are establishing a novel video watermarking approach, which uses both conventional methods as well as deep-learning-based method.

Though there are various methods present on watermarking, we discuss image and video watermarking methods on discrete wavelet transform (DWT), singular value decomposition (SVD), and neural network, also using these techniques in our approach. Averbuch et al. (2016) describes about DWT and explains important of DWT in watermarking. In SVD (Mohan and Kumar, 2008), few singular values compacts maximum energy of the signal, and also these values are robust against small perturbations on it. Tian and Ji (2010) inserted and extracted watermark bits into consecutive frames based on average pixel values in DWT domain. This method showed good performance against different filtering attacks, but fails when frame drop attack is applied. DWT-based approaches (Rathore et al., 2007; Liu et al., 2002) used DWT for watermarking, watermark is inserted into coefficients of LL band directly. Adding of watermark bits into low-frequency coefficients directly result into loss of watermark when any kind of filtering is applied on it or during compression. Thus, watermark bits will be modified.

Wang et al. (2010) applied temporal wavelet transform (TWT) on 32 consecutive frames, then selected resultant high frequency coefficients for inserting a watermark bit. As the method inserts watermark bit into multiple frames, retrieving of watermark bit becomes challenging when frame drop attack is presented and also finding the starting frame out of 32 frames during extraction process is very difficult when the clipping attack is applied. Kadu et al. (2016) employed key-based watermarking insertion along with DWT, where key is generated using watermark image and low frequency coefficients in DWT domain. These keys are also required during extraction of watermark bit though original video is not required. So, key should be preserved securely and required during extraction. 3-level DWT and scene change detector is used in Shukla and Sharma (2018) to insert a watermark bit into a video frame.

In few approaches, a combination of different transform domains had been explored to make watermarking system robust against different attacks. DWT and SVD were (Dey et al., 2012) applied on host video and watermark, but the method needs both U and V vector (out of SVD) matrices of watermark during extraction of a watermark, thus the method should store them. We implemented this approach and concluded that, using U and V components of a different watermark can also be used to extract it. This approach breaches security. The same approach was applied in Fung and Godoy (2011) and in-addition, the method also used DCT along with DWT, and SVD.

Along with DWT and SVD, few other methods like Takore et al. (2016) used genetic optimisation to achieve trade-off between robustness vs. perceptual quality. Nandi et al. used particle swarm optimisation technique (Nandi and Santhi, 2016) to achieve trade-off between quality vs. robustness. Sake used bee colony algorithm as an optimiser for embedding a watermark effectively in Sake and Tirumala (2016).

Though many methods exist on video watermarking, there is still a scope to build robust video watermarking solution, which achieves both robustness as well as perceptual quality. Having said as above, In this work, we develop a novel robust video watermarking solution, which achieves robust against rotation and scaling by applying log-polar mapping as a preprocessing technique before embedding a watermark into a singular values in DWT-SVD domain rather than directly computing singular values from the input video frame. Also, developed a deep neural network architecture to generate secret share images rather than conventional tabular-based method. In addition, we use firefly optimisation techniques to achieve trade-off between robustness vs quality. In this solution, watermark is embedded into a video frame in DWT and SVD domain due to their advantages, and neural network is chosen to eliminate time consuming process of generating secret shared image using tabular approach. Even in case if one watermark bit is lost, we can still successfully retrieve it using remaining bits as we have generated shared image using deep-neural network (DWT). Training of neural network is off-line process and generation of secret share is online. Besides this, the solution is blind watermarking method, which does not require original video during watermark extraction process.

This work is categorised into different sections. Section 2 introduces the techniques adopted for watermarking. In Section 3, we give detailed discussion on embedding method, secret sharing, and watermark extraction methodology. We empirically evaluated performance of the method against various kinds of videos in Section 4. We conclude the method by stating the methodology in the Section 5.

2 Background

In this section, we introduce basic techniques, before describing the watermark embedding and the extraction method, like log-polar transform, DWT, SVD, and deep neural networks and their applicability to video watermarking.

2.1 Log-polar mapping

Araujo and Dias (1996) proved log-polar mapping is akin to retino-cortical mapping, which transforms retinal image of our eyes into a form before it is examined by our brain. This transformation makes transformed form is invariant to rotation and scale. These properties enable the log-polar transform to be used in various image processing applications like visual object tracking (Long et al., 2017; Li et al., 2017; Bakhshande and Taghirad, 2015), image registration (Ohnishi et al., 2017; Zhang et al., 2016; Zeng et al., 2017) and image watermarking (Ouyang et al., 2015; Qu et al., 2017; Yang et al., 2018), etc.

Log-polar mapping indexes ordinary x, y spatial image coordinates by orthogonal axes (W, R), where R and W are ring number and wedge number in polar domain respectively.

$$r = \sqrt{\left((x - x_c)^2 + (y - y_c)^2\right)} \tag{1}$$

$$\theta = \tan^{-1}\left(\frac{y - y_c}{x - x_c}\right) \tag{2}$$

$$R = \frac{(n_r - 1)\left(\log^{\frac{r}{r_{\min}}}\right)}{\log^{\frac{r_{\max}}{r_{\min}}}} \tag{3}$$

$$W = \frac{n_w \theta}{2\pi} \tag{4}$$

where (r, θ) are polar coordinates for indexing purpose, point (x_c, y_c) is the position of the centre of the log-polar sampling pattern, nr is indices for rings n_w indices for wedges, and $r_{\min}$ is radii of smallest ring samples and $r_{\max}$ is largest rings of sample.

Image scaling, and rotation in spatial domain becomes shifts in *R* and *W* in log-polar domain. So, this property makes the log-polar mapping useful for watermarking.

2.2 Discrete wavelet transform

DWT has been performing well in various signal processing applications due to its simultaneous space and frequency analysis. The transform represents any signal in-terms of basic functions, also called as wavelets, using filter banks. DWT filters an image into multi-resolution such that signal can be analysed in spatial and frequency domain. 1D wavelet can be used to apply to 2D DWT, which results in 4 sub-bands, each sub band is having specific type of frequency coefficients. These sub-bands are used either for further splitting or applications like compression or watermarking (Ali, 2010). Simple one level DWT filter for decomposition in Figure 1 and example results of 3 level DWT is shown in Figure 2.

Figure 1 Represents DWT decomposition

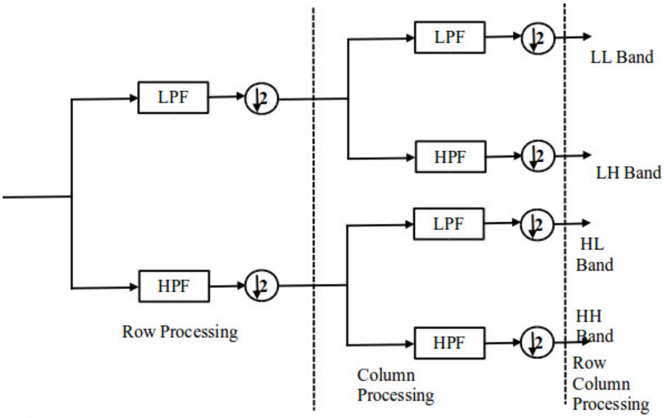


Figure 2 Sample 3-level DWT decomposition result (see online version for colours)

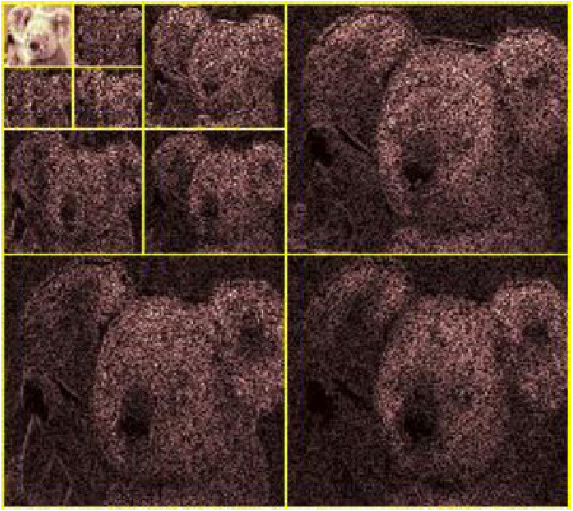


Figure 1 has set of filters: low pass filter (LPF), and high pass filter (HPF). Which are conjugate filters and termed as filter banks. There are various kinds of filter banks such as Daubechies, Haar, Coiflets, Symlets, etc. In this paper, we have chosen Haar Wavelets due to its simple operations and efficiency. Haar wavelets are developed by Hungarian mathematician Talukder and Harada (2007), which is akin to step function. The Haar wavelet operation is computed by averaging and differences the coefficient values. The reconstruction filter does reverse operation of the decomposition filter and is shown in Figure 1. The only difference between them is instead of down sampling in decomposition filter, here we up-sample in reconstruction filter. Due to these advantages, these filters are still being used in various applications. So, we also use this in present work.

2.3 Singular value decomposition

SVD is used for analysing signal energy by decomposing an image into three matrices and also is used for compressing a signal with the help of singular value matrix. It generates left and right singular vector matrices, diagonal matrix with singular values in diagonal direction. The diagonal matrix consists of singular value, which are in-turn represents its contribution in an energy of the signal. The diagonal matrix plays a key role in signal watermarking as well. The key property of singular values is that they are insensitive to small perturbation in a signal due to compression or noise. So, these can be used in inserting a watermark bit.

2.4 Secret sharing

We also employed scrambling for the watermark to eliminate burst errors, these are occurred when watermarked frame is cropped or highly attacked by using various signal processing methods. Single bit errors are corrected with ease with the help of existing error correcting codes than burst errors. In this paper, we employed a scrambling method (Rhine and Bhuvan, 2014) where row column transformations are used for scrambling. Keeping in mind the end goal to make the watermark more certain and secured, watermark can be shared among different specimens/samples or each bit of watermark is distributed or shared into different bits. This can be accomplished by utilising secret sharing method (Hsieh et al., 2007), in which image to be secreted will be represented with n-shares using a pre-determined code book. To retrieve the original image, at least few shares are enough. In our technique we utilised a strategy (Hsieh et al., 2007), where for producing the share image from the watermark, each piece of watermark alongside feature vector extricated from the cover image has been utilised to create share bits of relating watermark bit. For a typical 30 × 30 image, the share image will be 4× size due to the above said technique. This is randomly called as private share (p-share) image. This technique can also be used in video embedding. During the watermark extraction to verify the ownership or copyright, while extracting watermark for any validation procedures, the same process is employed and that extracted image will be called c-share. To extract the exact watermark, both c-share and p-share extracts are used. Sometimes due to errors, it is not possible to recover exact watermark, to overcome this problem morphological operation such as dilation or erosion will be employed.

3 Methodology

We develop a novel robust video watermarking solution, which achieves robust against rotation and scaling by applying log-polar mapping as a pre-processing technique before embedding a watermarking into a singular values in DWT-SVD domain rather than directly computing singular values from the input video frame. Also, we develop a deep neural network architecture to generate secret share images rather than conventional tabular-based method. In addition, we use firefly optimisation techniques to achieve trade-off between robustness vs. quality. We provide detailed description about secret sharing methodology, and both embedding and extraction of watermark process.

Secret share generation

We start by discussing about importance of secret sharing and then its methodology. Whenever we insert a watermark directly into a host video frame and apply attack on it, the watermark bits will be corrupted. One can use error correcting codes (Peterson et al., 1972; Pastawski et al., 2015; Mstafa and Elleithy, 2016; Procaccia et al., 2016) to correct erroneous bits, but those methods can help up to some extent. Other possibility of handling them is before inserting a watermark bit into a host video frame, watermark bit can be represented with multiple bits. Hence, though one or two bits are corrupted in this experiment, original bit is recovered with the help of remaining bits. Such methods are called as secret sharing approaches (Roy et al., 2015; Tuncer and Avci, 2016; Avci et al., 2016; Hsieh et al., 2007; Hsieh et al., 2008). In this work, we generate secret shared image based on concepts of Hsieh et al. (2007) and Hsieh et al. (2008), but we use deep-neural networks to generate a shared image instead of using tabular comparison, as it needs lot of computations and comparison. The advantage of the approach (Hsieh et al., 2007; Hsieh et al., 2008) is that, secret shared bits are formed based on content of input image as well as watermark bit, so robustness and imperceptibility will be maintained. First, we explain about tabular approach, then we move to DWT-based approach. In tabular approach, very first input values are computed along with watermark bit, input values would be singular values computed on low frequency band (LL) band of DWT image as given in Hsieh et al. (2007). This needs computation of multi-level DWT and SVD before performing tabular comparison and then generation of shared bits. Generally, this kind of approach is very expensive as it involves computation of DWT, SVD and tabular-based approach. So, we eliminated this by simply training the DWT approach though training takes time but during testing, it would have only forward pass to generate secret share image.

P-share generation

Recent advances in DWT (Hinton et al., 2015; Nguyen et al., 2015; Lee et al., 2018; Zeiler and Fergus, 2014; Simonyan and Zisserman, 2014) showed improved performance in image processing applications and encouraging to use deep-neural networks in every research field. The problem with them is they required lot of data to learn. Though it is taking time to converge during learning, it takes less time during testing or generating as it involves only forward pass. With this in mind, we construct two neural network-based architectures to generate two kinds of secret share images: p-share needed during embedding, and c-share needed during extraction. Figure 3 shows the neural network architecture during learning phase (p-share network). The p-share net is trained on multiple image patches of size 8 × 8 and watermark bits as input and secret shared bits, which are generated based on approach (Hsieh et al., 2007; Hsieh et al., 2008), are used as output. The learned network is as shown Figure 4 generates as secret share (pshare) bits before watermark embedding. The input dimension of the network is 65 i.e., 8 × 8 image and one watermark bit. It generates 4 output bits.

Figure 3 P-share generation training (see online version for colours)

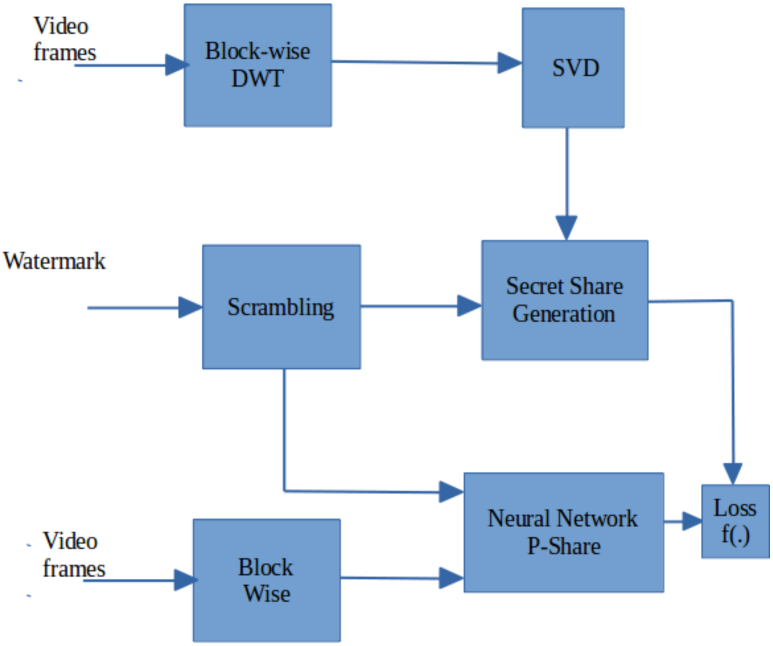
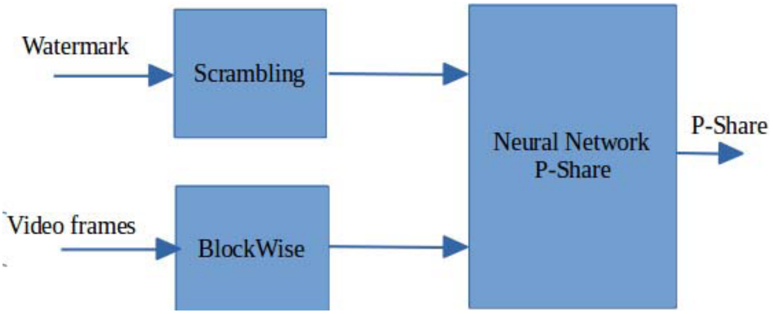


Figure 4 P-share generation (see online version for colours)



Embedding

Watermark is inserted into a host video in log-polar, DWT, and SVD domain to make it robust. A log-polar mapping is applied on each input video frame to make the method robust against rotation, and scaling as log-polar is invariant to this. Then, block of wise

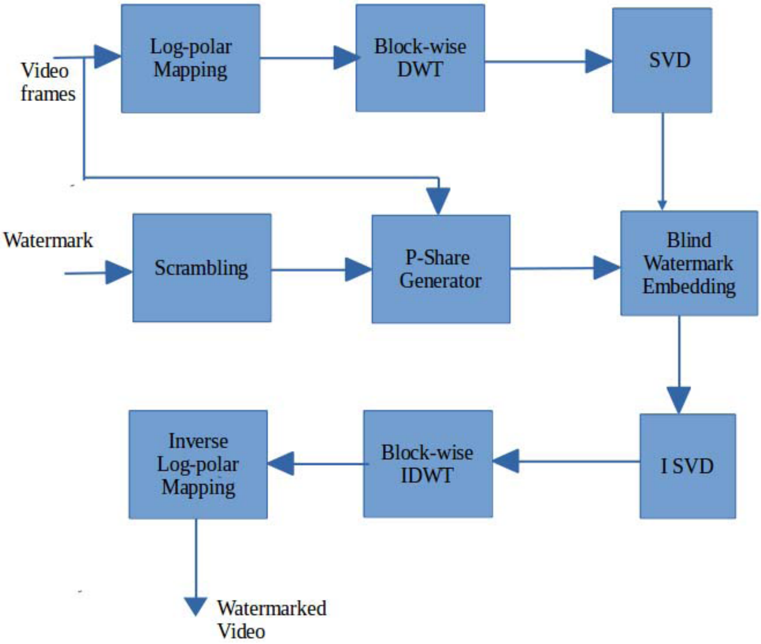
DWT is applied on each log-polar mapped frame to obtain four sub-bands, LL sub-band is chosen for applying SVD. The block size we fixed is 8 × 8 and applied normal DWT, this procedure is iterative. It results in multiple singular values. Top four singular values are chosen to embed a watermark bit as top values holds most of the energy of the input so that system will be robust against various kinds of filtering attacks. Watermark is scrambled to prevent burst errors and then expanded from $m \times m$ size to $2m \times 2m$ size using the secret sharing (p-share) approach that we discussed above. We embed the each bit of watermark into the second largest singular value as given in the equation below.

$$S_2 = \beta * (S_1 + S_3) + \alpha * w \tag{5}$$

where S_1, S_2, S_3 are the top three singular values in descending order. w watermark bit and scaling factor α. We set β is 0.95 as it is giving better results than other values. The same is discussed in results section as well.

This process will be continued till all the watermarked inserted into a frame. This whole process will be continued for each frame of a video. Figure 5 shows watermark insertion process.

Figure 5 Watermark insertion (see online version for colours)



C-share generation

Once watermark is embedded, this video may undergo various attacks like rotation, resize, and filter, etc. During which, few watermark bits will be corrupted. As we are correcting them using secret sharing approach, we need to construct a c-share image from the watermarked video frame using deep neural network. Before generating them, we design a deep neural network architecture (c-share network), which is generated on a separately created training data, where the input is a watermarked video and output is corresponding c-share image generated using tabular approach. The relevant diagram is as shown in Figure 6. The network is trained until network is converged. Training parameters are discussed in experimental section. While extracting a watermarking, we feed block of video frame into a neural network to generate corresponding share image. The trained network, which is used during testing is shown in Figure 7.

Figure 6 C-share image generation training (see online version for colours)

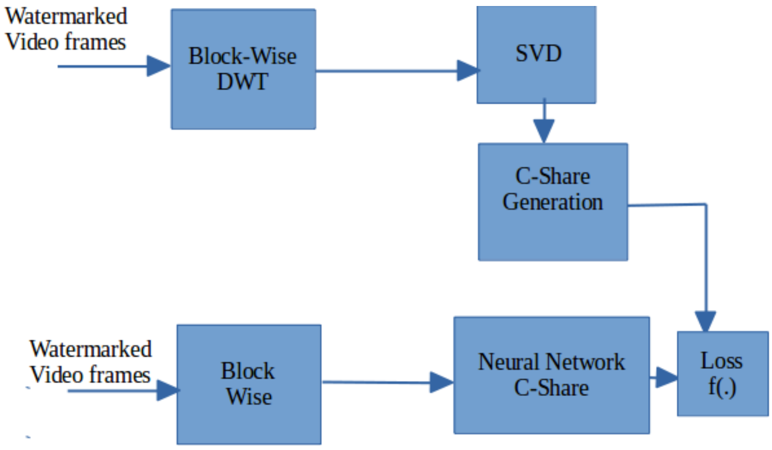
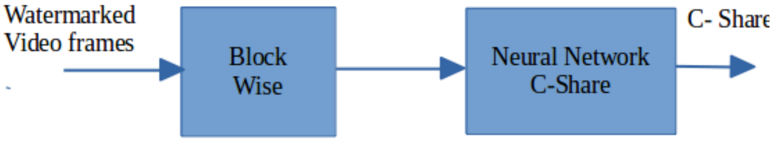


Figure 7 C-share image generation (see online version for colours)



Watermark extraction

We apply log-polar filtering over input watermarked video. Block-wise DWT is applied on the mapped video frame. Then SVD is applied over the LL band of DWT image. It results in a few singular values. Watermark is extracted, as this is a blind watermarking process, by comparing the singular values among them shelf. The comparison equation is as shown in following equations.

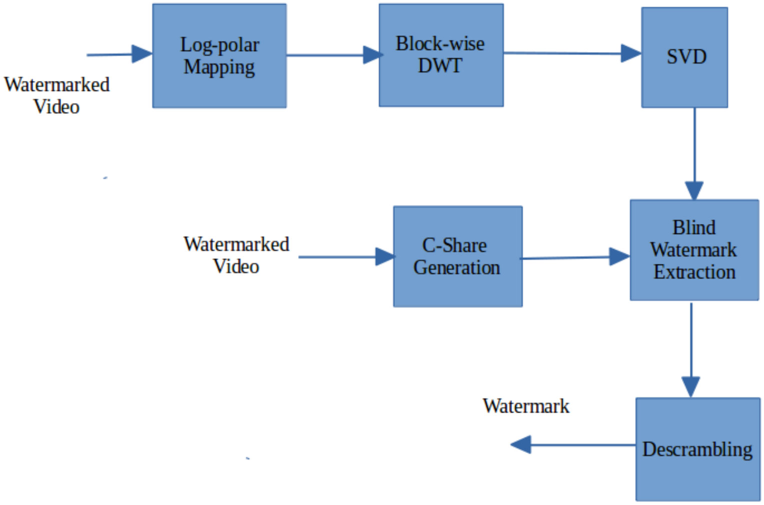
$$w = \begin{cases} 0, & \text{if } S_2 < \beta * (S_1 + S_3) \\ 1, & \text{otherwise} \end{cases} \tag{6}$$

where S_1, S_2, S_3 are the top three singular values in descending order. Watermark bit w and scaling factor α. β is set to 0.95.

On the other hand, c-share bits are extracted from a watermarked video frame. The c-share bits (image) is exclusive-ored with extracted bits. This results in a watermark bits of size $2m \times 2m$ and 2×2 block wise max pooling is performed to obtained watermark of size $m \times m$. Figure 8 describes about watermark extraction.

We also employed firefly optimisation algorithm (Latha et al., 2017) to obtain better scaling factor, which achieves trade-off between robustness vs. imperceptibility.

Figure 8 Watermark extraction (see online version for colours)



4 Experimental results

We discuss empirical results of the proposed method both quantitatively and qualitatively in this section. We evaluate the method on our own collected data-set consists of 100 videos with an average duration of 10 minutes. The data-set comprises of videos of various genre (e.g., sports, movies, cartoon, and news, etc.). We simulated the method in MATLAB. We select the best embedding scaling parameter using the firefly optimisation method by varying the scaling factor ranging from 1 to 35 with the multiples of 5. We considered all the videos are in avi format. We fixed the video size to 640 × 480 and the watermark size to 30 × 40, after secret sharing it would become 60 × 80 and the block size we fixed is 8 × 8. We constructed a fully connected p-share network to generate a p-share image with four layers with hidden neurons 32 in the first layer, 16 in the second layer, 8 in the third layer and 4 in the last layer. Also, we constructed c-share network to generate a c-share image, which is used at the extraction stage. The layers are fixed both in p-share and c-share image. The only difference is that, p-share net has 65 inputs and c-share net has 64 inputs.

Perceptual quality of watermarking process is measured by computing peak signal to noise ratio (PSNR) metric as shown in the equations (7) and (8) and structural similarity (SSIM) as in the equation (9) and robustness of the method is measured by calculating bit error rate (BER).

$$MSE = \frac{1}{mn} \sum_{i=1}^{m-1} \sum_{j=1}^{n-1} \left(I(i, j) - K(i, j) \right) \tag{7}$$

$$PSNR = 20 \log_{10}^{\left(\frac{MAX_t}{MSE} \right)} \tag{8}$$

$$SSIM(x,y)=\frac{\left(2\mu_x\mu_y+C_1\right)+\left(2\sigma_{xy}+C_2\right)}{\left(\mu_x^2+\mu_y^2+C_1\right)\left(\sigma_x^2+\sigma_y^2+C_2\right)} \tag{9}$$

We have selected a binary logo of size 30 × 40 as a watermark in our experimentation and is shown in Figure 9 and its scrambled output is shown in Figure 10.

Figure 9 Sample watermark

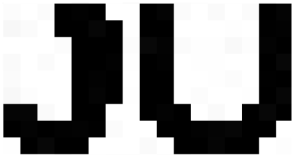
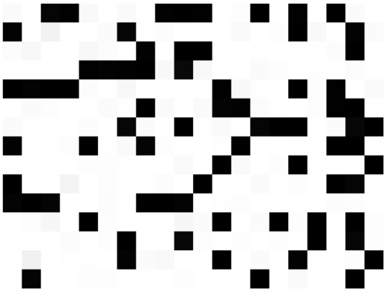
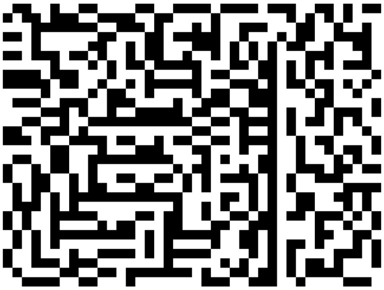


Figure 10 Scrambled output of the watermark



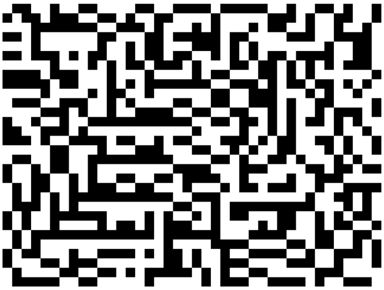
The logo along with a video frame is fed to p-share network to generate a p-share image of size 60 × 80. P-share image is not constant for every frame, it varies from frame to frame as it is being generated based on input video frame and the same is inserted into the corresponding frame. Example p-share image samples, which are generated using the watermark and the two sample video frames, are shown in Figures 11 and 12, respectively. Visually one can observe that these two are very different.

Figure 11 Share image generated from video frame 1



As we discussed about share generation, we feel that this is right place to discuss about network training procedure. We trained the p-share network on 10,000 input and output pairs using cross-entropy loss function and validated on 2,000 input and output pairs. We achieved an accuracy of 99.05% on validation set. The same training and validation set are used, except output data, for c-share net and achieved validation accuracy of 99.20%.

Figure 12 Share image generated from video frame 2



The imperceptibility of the video is measured by computing PSNR value and it is in the range of 43 to 54 depending on a video genre. SSIM is in the range of 0.91 to 0.998.

The range is enough to say the method is acceptable for watermarking a video to prove ownership.

Figure 13 Extracted watermark from one frame

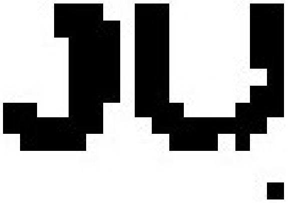
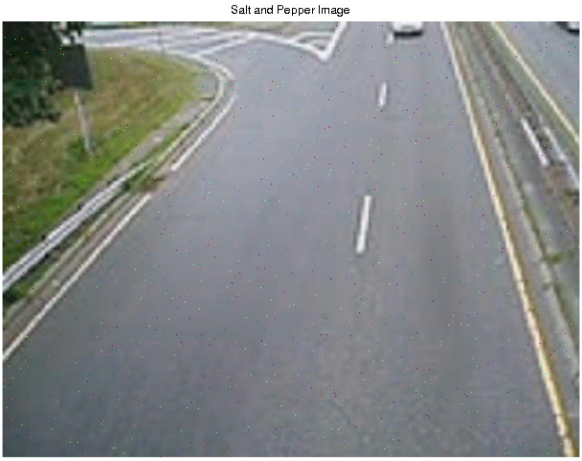


Figure 14 Salt and pepper noise with density of 0.02 (see online version for colours)



We simulated the most common attacks using MATLAB and FFMPEG for measuring robustness of the method. The attacks are filtering (Gaussian and median), scaling (50%

to 200%), rotation (–50 to +50), compression (mpg, mp4) with different bit rates (512 Kbps to 10 Mbps), frame rate conversion (22.34 to 30 fps), and frame drop attack.

We start analysing the result of watermark extraction from a signal without attack to different attacks. Figure 13 shows the watermark, which is extracted from signal with no attack. It contains almost negligible bits in error.

Figures 14 and 16 show an example video frame with salt and pepper noise is added with noise density 0.02 and 0.2 respectively. Figures 15 and 17 show corresponding extracted watermark with BER of 0.005% and 0.03%, respectively.

Figure 15 Retried watermark with the BER of 0.005%

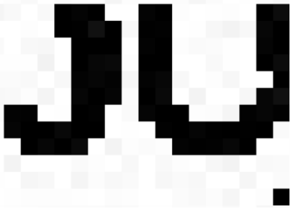


Figure 16 Salt and pepper noise with density of 0.2 (see online version for colours)

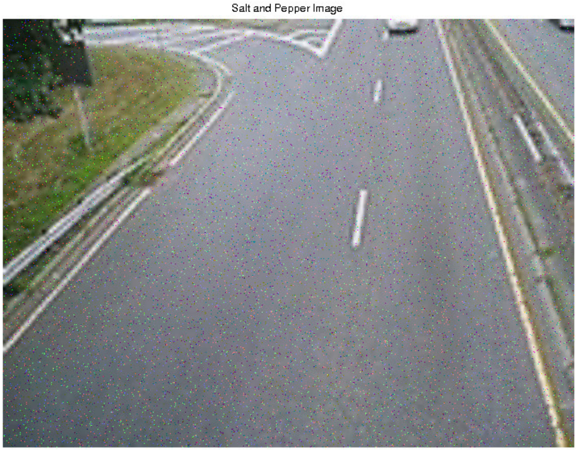
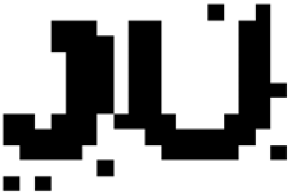


Figure 17 Retried watermark with the BER of 0.03%



Figures 18 and 19 show an example video frame with salt and pepper noise is added with noise density 0.02 and 0.2 respectively.

Table 1 shows BER performance comparison between different methods (Sake and Tirumala, 2016; Shukla and Sharma, 2018; Kadu et al., 2016) with different attacks. We implemented methods as given in the paper.

Table 2 shows BER variation against different attacks by keep β to 0.85, 0.95 and 0.99.

Figure 18 Sample frame under rotation-5 degrees attack (see online version for colours)

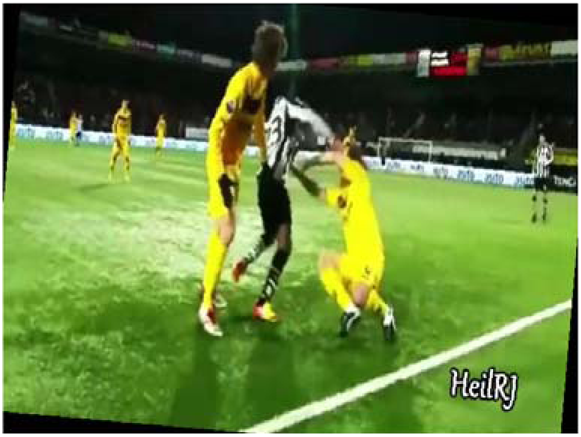


Figure 19 Retried watermark

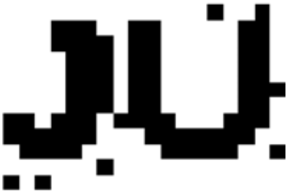


Table 1 Performance of the proposed method, where β is 0.95 and α is 25

Attack type	*BER (%)*			
	BDWTSVD (Sake and Tirumala, 2016)	*Level-3 DWT (Shukla and Sharma, 2018)*	*DWT (Kadu et al., 2016)*	*Proposed method*
Rotation 5 degrees	4.51	4.32	4.22	3.01
Rotation 5 degrees	4.53	4.32	4.23	3.02
MP4 compression (512 Kbp)	3.32	3.01	3.24	3.01
Scaling 200%	0.02	0.018	0.015	0.011
Scaling 50%	0.05	0.034	0.03	0.03
Cropping 30%	0.01	0.012	0.01	0.009
MPEG 2 (1,024Kbps)	2.93	2.53	2.98	2.50
Gaussian filtering	0.089	0.081	0.071	0.067

From Table 2, we conclude that when β is small then the updated S_2 value from the equation (6) is very close to S_3 than during extraction, it is not satisfying the comparison equation. When β is large then the updated S_2 which is close to S_1, even then it is not satisfying the extraction equation. So, we fixed β to 0.95. We also tried compression methods like avi, mkv jpeg and we achieved good performance and bit error rates are better than mp4 compression

Table 2 Performance of the proposed method, where α is 25

Attack type	*BER (%)*		
	$\beta = 0.85$	*$\beta = 0.95$*	*$\beta = 0.99$*
Rotation 5 degrees	5.10	4.32	4.90
Rotation 5 degrees	5.12	4.32	4.91
MP4 compression (512 Kbp)	4.60	3.01	4.20
Scaling 200%	0.10	0.018	0.09
Scaling 50%	0.23	0.034	0.1
Cropping 30%	0.9	0.012	0.08
MPEG 2 (1,024 Kbps)	5.63	2.53	3.81
Gaussian filtering	0.23	0.081	0.13

Table 3 Performance of the proposed method, where α is 25 and β is 0.95

Attack type	*BER (%)*		
	$s_2 = \beta(s_1 + s_3) + \alpha\ (wm)$	$s_2 = \beta(s_1 + s_4) + \alpha\ (wm)$	$s_2 = \alpha(wm)\ \beta(s_1) + \alpha(wm)$
Rotation 5 degrees	4.32	5.45	7.66
Rotation 5 degrees	4.32	5.46	7.61
MP4 compression (512 Kbp)	3.01	5.34	6.87
Scaling 200%	0.018	0.34	0.45
Scaling 50%	0.034	0.42	0.52
Cropping 30%	0.012	0.23	0.35
MPEG 2 (1,024 Kbps)	2.53	5.71	6.91
Gaussian filtering	0.081	2.01	4.13

Table 4 Performance of the proposed method with respect to PSNR and SSIM , where β is 0.95 and α is 25

Attack type	*BER (%)*			
	BDWTSVD (Sake and Tirumala, 2016)	*Level-3 DWT (Shukla and Sharma, 2018)*	*DWT (Kadu et al., 2016)*	*Proposed method*
PSNR	45.67	46.8	46.45	48.56
SSIM	0.93	0.945	0.936	0.956

Table 3 shows BER variation against different attacks by changing different kinds of embedding equation (5).

From Table 3 we can observe the present used equation is in 5 is giving better results than remaining other equations, though we changed extraction equation according to embedding equation.

We also compared performance evaluation with respect to PSNR and SSIM with the state-of-the art and given in 4.

We also computed time complexity of the algorithm for embedding and extraction process with and without Neural network-based approach. The system configuration:

- operating system: Linux
- software: Matlab, Python, FFMPEG
- CPU: Intel(R) Core(TM) i3-2120 CPU @ 3.30 GHz
- RAM:4GB.

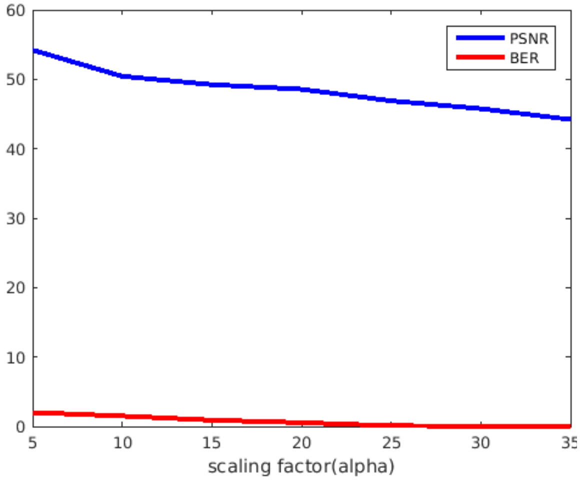
The watermark embedding speed is 40 frames per second (FPS) with neural network-based approach and 28 FPS with tabular-based approach. The MATLAB code is not optimised, this will be improved if we implement it in c and use multi threading. The watermark extraction process is much faster than the embedding as the extraction process not having computation of inverse SVD, DWT, and log-polar. The extraction process is almost double the embedding process.

Trade-off between the robustness and the imperceptibility is computed by varying the scaling factor ranging from 5 to 35 and measuring PSNR and BER. Figure 20 shows the scaling factor vs. BER and PSNR.

From Figure 20, we can notice, as the scaling factor is increased both PSNR and BER are decreased.

In addition, we also employed firefly optimisation algorithm[] to achieve trade-off between the robustness and the imperceptibility. The optimisation method has run PSNR and BER by simulating various attacks. The method results in best scaling factor which achieves said trade-off.

Figure 20 Scaling factor vs. BER, PSNR (see online version for colours)



5 Conclusions

We developed a robust blind video watermarking method, which inserts a watermark into a video frame based on log-polar, DWT, SVD and extracts it when necessary in the same domain. We also developed a neural network-based technique to generate secret share images in order to improve accuracy and computation speed. We successfully extracted the watermark with high accuracy from videos even when the signal undergone various signal processing attacks: filtering, rotation, scaling, compression, etc. The performance of the method is compared with the state of the art methods-based DWT and SVD techniques and proved that the method is outperforming compared to state of the art methods. The method also showed improved performance over other state of the art methods based on DCT, log-polar domain. As a future work, we extent and make the system more robust against other severe attacks that means increasing level of attacks, also want to compare with non DWT and SVD methods as well as atom kind of transforms and improve the speed.

References

Ali, M. (2010) *An Introduction to Wavelets and Haar Trans* [online] http://www.cs.ucf.edu/mali/haar.

Araujo, H. and Dias, J.M. (1996) 'An introduction to the log-polar mapping [image sampling]', *Proceedings II Workshop on Cybernetic Vision*, Sao Carlos, Brazil, pp.139–144.

Avci, E., Tuncer, T. and Avci, D. (2016) 'A novel reversible data hiding algorithm based on probabilistic XOR secret sharing in wavelet transform domain', *Arabian Journal for Science and Engineering*, Vol. 41, No. 8, pp.3153–3161.

Averbuch, A., Lazar, D. and Israeli, M. (1996) 'Image compression using wavelet transform and multiresolution decomposition', *IEEE Transactions on Image Processing*, January, Vol. 5, No. 1, pp.4–15.

Bakhshande, F. and Taghirad, H.D. (2015) 'Visual tracking using kernel projected measurement and log-polar transformation', *International Journal of Robotics*, Vol. 4, No. 1, pp.1–11.

Boland, F.M., O'Ruanaidh, J.J.K. and Dautzenberg, C. (1995) 'Watermarking digital images for copyright protection', *Fifth International Conference on Image Processing and its Applications*.

Boreiry, M. and Keyvanpour, M. (2017) 'Classification of watermarking methods based on watermarking approaches', *Artificial Intelligence and Robotics (IRANOPEN)*, Qazvin, pp.73–76.

Cox, I., Miller, M., Bloom, J., Fridrich, J. and Kalker, T. (2008) *Digital Watermarking and Steganography*, 2nd ed., Elsevier.

Dey, N., Das, P., Roy, A.B., Das, A. and Chaudhuri, S.S. (2012) 'DWT-DCT-SVD based intravascular ultrasound video watermarking', *World-Congress on Information and Communication Technologies (WICT)*, Trivandrum, pp.224–229.

Fung, C.W.H. and Godoy, W. (2011) 'A novel DWT-SVD video watermarking scheme using side view', *5th International Conference on Signal Processing and Communication Systems (ICSPCS)*, Honolulu, HI, 2011, pp.1–4.

Hinton, G., Vinyals, O. and Dean, J. (2015) 'Distilling the knowledge in a neural network', arXiv preprint arXiv:1503.02531.

Hsieh, S-L., Jian, J-J., Tsai, I-J. and Huang, B-Y. (2007) 'A color image watermarking scheme based on secret sharing and wavelet transform', *IEEE International Conference on Systems, Man and Cybernetics*, IEEE, Montreal, Que., pp.2143–2148.

Hsieh, S-L., Tsai, I-J., Huang, B-Y. and Jian, J-J. (2008) 'Protecting copyrights of color images using a watermarking scheme based on secret sharing and wavelet transform', *Journal of Multimedia*, October, Vol. 3, No. 4, pp.42–49.

Hu, Z., She, K., Wang, J. and Tang, J. (2014) 'Game theory based false negative probability of embedded watermark under unintentional and steganalysis attacks', *China Communications*, May, Vol. 11, No. 5, pp.114–123.

Kadu, S., Naveen, C., Satpute, V.R. and Keskar, A.G. (2016) 'Discrete wavelet transform based video watermarking technique', *International Conference on Microelectronics, Computing and Communications (MicroCom)*, Durgapur, pp.1–6.

Kumar, N.V., Sreelatha, K. and Kumar, C.S. (2016) 'Invisible watermarking in printed images', *1st India International Conference on Information Processing (IICIP)*, Delhi, pp.1–5.

Latha, S.B., Dasari, V.R. and Avula, D. (2017) 'Robust video watermarking using secret sharing, SVD, DWT and chaotic firefly algorithm", *International Journal of Intelligent Engineering and Systems*, Vol. 11, No. 1, pp.171–180.

Lee, S., Chen, T., Yu, L. and Lai, C. (2018) 'Image classification based on the boost convolutional neural network', *IEEE Access*, Vol. 6, pp.12755–12768.

Li, Y., Zhu, J., Song, W., Wang, Z., Liu, H. and Hoi, S.C.H. (2017) 'Robust estimation of similarity transformation for visual object tracking with correlation filters', arXiv:1712.05231.

Liu, H., Chen, N., Huang, J., Huang, X. and Shi, Y.Q. (2002) 'A robust DWT-based video watermarking algorithm', *IEEE International Symposium on Circuits and Systems, 2002. ISCAS 2002*, Phoenix-Scottsdale, AZ, pp.631–634.

Long, A.D., Narayanan, R.M., Kane, T.J., Rice, T.F. and Tauber, M.J. (2017) 'Foveal scale space generation with the log-polar transform', *Image Sensing Technologies: Materials, Devices, Systems, and Applications IV*, Vol. 10209, p.1020910, International Society for Optics and Photonics.

Mohan, B.C. and Kumar, S.S. (2008) 'A robust image watermarking scheme using singular value decomposition', *Journal of Multimedia*, Vol. 3, No. 1, p.715.

Mstafa, R.J. and Elleithy, K.M. (2016) 'A video steganography algorithm based on Kanade-Lucas-Tomasi tracking algorithm and error correcting codes', *Multimedia Tools and Applications*, Vol. 75, No. 17, pp.10311–10333.

Nandi, S. and Santhi, V. (2016) 'DWTSVD-based watermarking scheme using optimization technique, artificial intelligence and evolutionary computations in engineering systems', *Advances in Intelligent Systems and Computing*, Vol. 394, pp.69–77, Springer.

Nguyen, A., Yosinski, J. and Clune, J. (2015) 'Deep neural networks are easily fooled: high confidence predictions for unrecognizable images', *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.427–436.

Ohnishi, K., Kirk, D.R. and Techau, P.M. (2017) *Image Registration Using a Modified Log Polar Transformation*, U.S. Patent 9,547,884, issued January 17, 2017.

Ouyang, J., Coatrieux, G., Chen, B. and Shu, H. (2015) 'Color image watermarking based on quaternion Fourier transform and improved uniform log-polar mapping', *Computers Electrical Engineering*, Vol. 46, pp.419–432.

Pastawski, F., Yoshida, B., Harlow, D. and Preskill, J. (2015) 'Holographic quantum errorcorrecting codes: toy models for the bulk/boundary correspondence', *Journal of High Energy Physics*, Vol. 2015, No. 6, p.149.

Peterson, W.W., Wesley, W., Weldon Jr., E.J., Peterson, E.J.W. and Weldon, E.J. (1972) *Error-Correcting Codes*, MIT Press.

Procaccia, A.D., Shah, N. and Zick, Y. (2016) 'Voting rules as error-correcting codes', *Artificial Intelligence*, Vol. 231, pp.1–16.

Qu, Z., Cheng, Z., Luo, M. and Liu, W. (2017) 'A robust quantum watermark algorithm based on quantum log-polar images', *International Journal of Theoretical Physics*, Vol. 56, No. 11, pp.3460–3476.

Rathore, S.A., Gilani, S.A.M., Mumtaz, A., Jameel, T. and Sayyed, A. (2007) 'Enhancing invisibility and robustness of DWT based video watermarking scheme for copyright protection', *International Conference on Information and Emerging Technologies, 2007. ICIET 2007*, Karachi, pp.1–5.

Rhine, R. and Bhuvan, N.T. (2014) 'Image scrambling methods for image hiding: a survey', *International Journal of Computer Science and 48 Information Technologies (IJCSIT)*, Vol. 15, No. 2, pp.86–91.

Roy, R., Bandyopadhyay, S., Kandar, S. and Dhara, B.C. (2015) 'A novel 34 image secret sharing scheme', *International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, IEEE, pp.2072–2075.

Sake, A. and Tirumala, R. (2016) 'Bi-orthogonal wavelet transform based video watermarking using optimization techniques', *2nd International Conference on Solar Energy Photovoltaic*, ScienceDirect, 17–19 December.

Shukla, D. and Sharma, M. (2018) 'Robust scene-based digital video watermarking scheme using level-3 DWT: approach, evaluation, and experimentation', *Radioelectronics and Communications Systems*, January, Vol. 61, No. 1, p.112.

Simonyan, K. and Zisserman, A. (2014) 'Very deep convolutional networks for large-scale image recognition', arXiv preprint arXiv:1409.1556.

Takore, T.T., Kumar, P.R. and Devi, G.L. (2016) 'A modified blind image watermarking scheme based on DWT, DCT and SVD domain using GA to optimize robustness', *International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*, Chennai, pp.2725–2729.

Talukder, K.H. and Harada, K. (2007) 'HaarWavelet based approach for image compression and quality assessment of compressed image', *IAENG International Journal of Applied Mathematics*, Vol. 36, No. 1, p.1.

Tian, H. and Ji W. (2010) 'A digital video watermarking scheme based on 1D-DWT', *International Conference on Biomedical Engineering and Computer Science*, Wuhan, pp.1–3.

Tuncer, T. and Avci, E. (2016) 'A reversible data hiding algorithm based on probabilistic DNA-XOR secret sharing scheme for color images', *Displays*, Vol. 41, pp.1–8.

van Schyndel, R.G., Tirkel, A.Z. and Osborne, C.F. (1994) 'A digital watermark', *Proceedings of 1st International Conference on Image Processing*, Austin, TX, Vol. 2, , pp.86–90.

Wang, C., Zhang, C. and Hao, P. (2010) 'A blind video watermark detection method based on 3D-DWT transform', *IEEE International Conference on Image Processing*, Hong Kong, pp.3693–3696.

Yang, Z-F., Kuo, C-T. and Kuo, T-H. (2018) 'Authorization identification by watermarking in log-polar coordinate system', *The Computer Journal*, November, Vol. 61, No. 11, pp.1710–1723.

Zeiler, M.D. and Fergus, R. (2014) 'Visualizing and understanding convolutional networks', *European Conference on Computer Vision*, Springer, Cham, pp.818–833.

Zeng, L., Zhou, D., Liang, J. and Zhang, K. (2017) 'Polar scale-invariant feature transform for synthetic aperture radar image registration', *IEEE Geoscience and Remote Sensing Letters*, Vol. 14, No. 7, pp.1101–1105.

Zhang, J., Sohel, F., Bian, H., Bennamoun, M. and An, S. (2016) 'Forward-looking sonar image registration using polar transform', *OCEANS 2016 MTS/IEEE Monterey*, pp.1–6, IEEE.