A Robust Video Watermarking Algorithm Based on SVD and VC

A.Radha
Associate professor
ECE dept
DVR & Dr. HS MIC
College of Technology

R.Harikishan
DECS
DVR & Dr. HS MIC
College of Technology
harikishan2010@gmail.com

R.Priyakanth
Associate professor
ECE dept
DVR & Dr. HS MIC
College of Technology

Abstract

When the digital world was at its infant stage, it was quite complicated and required a high level of expertise to duplicate digital contents so that they look like the original. However, in the present digital world, this is not true. Today, it is made so easy that almost anyone can duplicate, reproduce or manipulate digital data, with no degradation in data quality. As a consequence there was a search for a mechanism for providing the copyright protection scheme for providing the authentication of digital data. Digital watermarking is one of the most widely used technologies to perform this task. The basic idea of watermarking is to hide some information (the watermark) into a digital media, so that it can be later extracted or detected for a variety of purposes like identification and authentication. In this paper we propose a robust video watermarking algorithm that makes use of Fractional Fourier Transform (FrFt), Singular Value Decomposition (SVD) and Visual Cryptography (VC) and generates two shares namely the Master share and the Ownership share, for each selected frame of the video that is to be watermarked. The two shares of a particular frame will give no information individually about the embedded watermark. Visual cryptography is used to reveal the embedded watermark by combining the two shares of a frame. The properties of FrFt and SVD make the watermarking more robust to common signal processing attacks. The probability of proving the ownership is also more in case of attacks like frame insertion and frame deletion.

Key words – Digital watermarking, Visual secret sharing schemes, Singular value decomposition.

1. Introduction

Watermarking is the process of embedding information into a digital signal which may be used to verify its authenticity or ownership identity. The digital signal can be an image, an audio or a video signal. The watermarking schemes can be classified into three categories: non-oblivious watermarking, semi-oblivious watermarking, and oblivious watermarking, based on the process of extracting the watermark. A watermarking scheme is said to be non-oblivious, if to extract the embedded watermark it requires the host frame (of the host video). Semi-oblivious watermarking scheme requires the watermark and/ or some side information, instead of the host frame to extract the embedded watermark. Where as in case of oblivious watermarking scheme neither extra-information nor host frame are required to extract the watermark. Depending on the work domain in which the watermark is embedded, the

watermarking schemes can be further classified into two categories: spatial-domain watermarking schemes and transform- domain watermarking schemes. The spatial domain watermarking schemes will directly modify the intensity values of the pixels of the host frame and hence they are more robust to geometrical attacks, where as in transform-domain watermarking, first a transformation or a combination of transformations are applied on the host frame and then the watermark is embedded in the transformed coefficients of the host frame. The different transformation techniques that are most often used are: The discrete Fourier transform (DFT), the discrete cosine transform (DCT), the discrete wavelet transform (DWT).

In this paper a video watermarking scheme based on Singular Value Decomposition (SVD) and Visual Cryptography (VC) is proposed for providing an authentication to a video. The watermark is embedded in the transform domain by using the Fractional Fourier transform (FrFt).Since, the video can be regarded as the sequence of the correlated images, digital video watermarking can be achieved by either applying still image technology to each film frame or using dedicated methods that exploit inherent features of the video sequence. Here we use the first method. The host video is divided into frames and then three frames are randomly selected using pseudo random number generator. Each selected frame is then divided into 4x4 non-overlapping blocks and a sub frame is formed by selecting some blocks, using pseudo random number generator (PRNG) seeded with a secret key K. Features of the sub frame are extracted by applying FrFt and SVD on the sub frame. A binary map is constructed based on the extracted features. A master share is then constructed using this binary map. An ownership share is constructed by using the master share and the watermark to be embedded (secret image). The ownership can be proved by revealing the secret image through stacking the master share and ownership share of one of the selected frame of the concerned video.

A good watermarking scheme should possess the following basic characteristics.

Robustness- Robustness is the major requirement of the watermarking, without which we cannot recover the embedded watermark successfully when the video is subjected to attacks ranging from normal signal processing attacks to attacks like frame insertion, frame deletion and frame swapping which are more prominent in video attacks.

Imperceptibility- It means that the watermarked video and the original video should have no difference to human perception or the watermark should be invisible.

Security- It refers to the point that the embedded watermark should not be revealed by unauthorized persons and should be able to be revealed by the authorized person

2. BACKGROUND

2.1 Fractional Fourier Transform (FRFT)

The Fractional Fourier Transform is a linear transformation and is a generalized form of classical Fourier transform. We can think of it as the Fourier transform with **n-th** order where **n** need not be an integer and hence it can convert a signal to any intermediate domain between time and frequency.

Definition:

The FRFT is defined with the help of a transformation kernel K_α which is defined as follows [1]

$$K_\alpha(x,u) = \begin{cases} \delta(x-u) if\ \alpha\ is\ a\ multiple\ of\ 2\pi \\ \delta(x+u) if\ \alpha+\pi\ is\ a\ multiple\ of\ 2\pi \\ \sqrt{\frac{1-j\ cot\alpha}{2\pi}} \times \exp\ j\left(\frac{u^2+x^2}{2}\cot\alpha - xu\csc\alpha\right) \\ if\ \alpha\ is\ not\ a\ multiple\ of\ \alpha \end{cases} \quad (1)$$

The Fractional Fourier transform of order α for a function f(x) is denoted and defined then as follows

$$F_\alpha(f(x)) = \int_{-\infty}^{\infty} K_\alpha(x,u) f(x) dx \quad 0 \le |\alpha| \le 2$$
(2)

The Fractional Fourier transform possess the following properties [2]

Identity operator - The fractional Fourier transform with order 0 acts as an identity operator because of the fact that F_0 of a signal is the signal itself. Similarly the FRFT of order 2π is identical to successive application of the ordinary Fourier transform 4 times and results in the signal itself. Hence we have

$$F_0 = F_{2\pi} = I$$

Fourier transform operator- The fractional Fourier transform with order π/2 i.e., $F_{\pi/2}$ is the Fourier transform operator and gives the output as the Fourier transform of input signal.

Successive application- Successive application of Fractional Fourier transform to a signal is equivalent to applying a single transformation of order equal to sum of individual transforms.

$$F_\alpha\left(F_\beta\left(f(x)\right)\right) = F_{\alpha+\beta}(f(x))$$

Inverse transform - The Fractional Fourier transform of order –α is the inverse of transform of order α.

$$F_{-\alpha}(F_\alpha(f(x)) = F0\ (f(x)) = f(x)$$

The 2-dimensional discrete fractional Fourier transform of order (α, β) on a frame signal is calculated as follows [3]

$$F_{\alpha,\beta}(m,n) = \sum_{p=0}^{M-1}\sum_{q=0}^{N-1} f(p,q) K_{\alpha,\beta}(p,q,m,n)$$
(3)

And the inverse DFRFT of the signal is given by

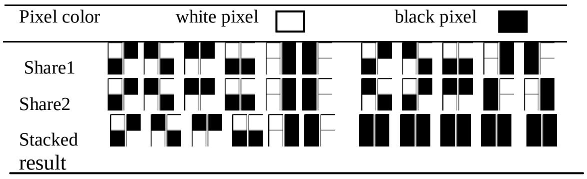
$$f(p,q) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} F(m,n)\ K_{-\alpha,-\beta}(p,q,m,n)$$
(4)

Where (α, β) is the order of 2D-FrFt and (K_α,K_β) are the 1D-DFrFt kernels.

2.2 Visual Cryptography (VC)

Visual Cryptography is the cryptography technology, which uses the human visual characteristics to decrypt the encrypted image. This technology was first proposed by Naor and Shamir in 1994 [4]. The attracting feature of this technology is that the secret image can be recovered without any computation. This is achieved by a technique called as K out of N shares threshold (K≤N). In this technique N transparencies (shares) are drawn from the secret image and for revealing the secret image we need to overlap at least K shares of the N shares. Any combination of shares less than K does not reveal the secret image. Another technique used is N out of N shares, where all the N shares must be stacked to reveal the secret image. These techniques are called as Visual Secret Sharing (VSS) schemes. In this paper we use a (2, 2) visual secret sharing scheme, where two shares called Master share and Ownership share are generated based on secret image and the secret image can be extracted by stacking the two shares. A typical 2 out of 2 VSS will be as shown below [5].

Table 1An example of (2 out of 2) visual secret sharing scheme



2.3 Singular Value Decomposition (SVD)

Any real m x n matrix X can be uniquely decomposed as

$$X = U D V^T$$

Where U is m x n orthogonal matrix, whose columns are the eigen vectors of X X^T.

V is n x n orthogonal matrix, whose columns are the eigen vectors of X^T X.

D is n x n diagonal matrix whose diagonal elements are non-negative real values called the singular values. The singular values are the square roots of the eigen values of XX^T.

The features of an image can be obtained from these three matrices of SVD by applying it to the image as a matrix. The singular values in D represent the luminance (energy) of the image while the corresponding singular vectors of the two diagonal matrices represent the horizontal and vertical details (edge) of the image. Slight variation of these singular values does not result in a large change in visual perception of human eye.

Some rigorous features of SVD:

Transport invariance- The matrix X and its transpose X^T have the same non-zero singular values.

Flip invariance – The matrix X its row flip X_{rf} and its column flip X_{cf} all will have the same non-zero singular values.

Rotation invariance – The matrix X and X_r (X rotated by an arbitrary angle) will have the same singular values.

Scale invariance- If we scale up every row in the matrix X by L1 and every column in the matrix by L2 then the resulting matrix will have each singular value $\sqrt{L1L2}$ times the corresponding eigen values of the original matrix and the original matrix and the scaled up matrix will have the same no. of singular values.

The above features of SVD make it useful to withstand several geometrical attacks [6].

3. PROPOSED SCHEME

3.1 Ownership registration procedure

Let us consider, H is the frame of size M × N that is extracted from the host video and the secret image S is a binary image of size m × n.

3.1.1 Master share construction algorithm

1. Divide the host video into frames
2. Select three frames randomly from the video and apply the following steps on each of the three selected frames
3. Divide the frame into 4x4 non-overlapping blocks
4. Select m x n blocks using Pseudo random number generator seeded with a secret key k
5. Apply Fractional Fourier transform of order α, β on all the m x n blocks
6. Apply SVD on all transformed blocks
7. Form a matrix X by collecting the first singular value of each matrix
8. Calculate the binary map B of the matrix X as

$$B_{ij} = \begin{cases} 0 \text{ } if \text{ } X_{ij} < X_{avg} \\ 1 \text{ } if \text{ } X_{ij} \geq X_{avg} \end{cases}$$

Where $X_{avg}$is the average value of all pixels in X. Here 1 denotes the white pixel and 0 denotes the black pixel.

6. Assume that M is a master share of size 2m x 2n pixels. Divide the master share into non-overlapping 2x2 blocks and the content of each block is determined according to the following master share generation rule:

If B_i is a white pixel then

$$M_{ij} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

If Bi is a black pixel then

$$M_{ij} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

3.1.2 Ownership share construction algorithm

After generating master share M, we generate owner-ship share O with the help of master share M and binary secret image S. Assume that O is the ownership share of size 2m x 2n. Divide O into non-overlapping 2x2 blocks. Now ownership share is constructed according to the following rule:

If $S_i = 1$ and $M_{ij} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ then $O_i = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

If $S_i = 1$ and $M_{ij} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ then $O_i = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

If $S_i = 0$ and $M_{ij} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ then $O_i = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

If $S_i = 0$ and $M_{ij} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ then $O_i = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

The ownership share O should be registered with a certified authority (CA) for further authentication.

3.2Ownership identification process

3.2.1 Identifying the type of attack

1. Divide the suspected video into frames
2. Select those three frames from the video that were used during the ownership registration process and compare them with the three stored frames. If at least one of them matches with the stored frames, then the attack is a normal signal processing attack and apply the algorithm for normal attacked frame on the matched frame. Otherwise the attack may be one of the following: frame insertion, frame deletion or frame swapping. In this case apply the algorithm for special attacks

3.2.2 Algorithm for normal frame attack

1. Divide the selected frame into 4x4 non-overlapping blocks
2. Select m x n blocks using Pseudo random number generator seeded with a secret key k
3. Apply Fractional Fourier transform of order (α, β) on all the m x n blocks
4. Apply SVD on all transformed blocks
5. Form a matrix X' by collecting the first singular value of each matrix
6. Generate the master share M' according to the master share generation rule.
7. Retrieve the secret image S' by stacking the master share M' and the ownership share O of the corresponding frame kept by the certified authority CA.
8. Divide the secret image S' into non-overlapping2x2 blocks. Let us denote these blocks byS''.
9. Get the reduced secret image S'' as
10. $S''_{ij} = \begin{cases} 0 \; if \; \sum_i \sum_j S'_{ij} < 2, \\ 1 \; if \; \sum_i \sum_j S'_{ij} > 2 \end{cases}$

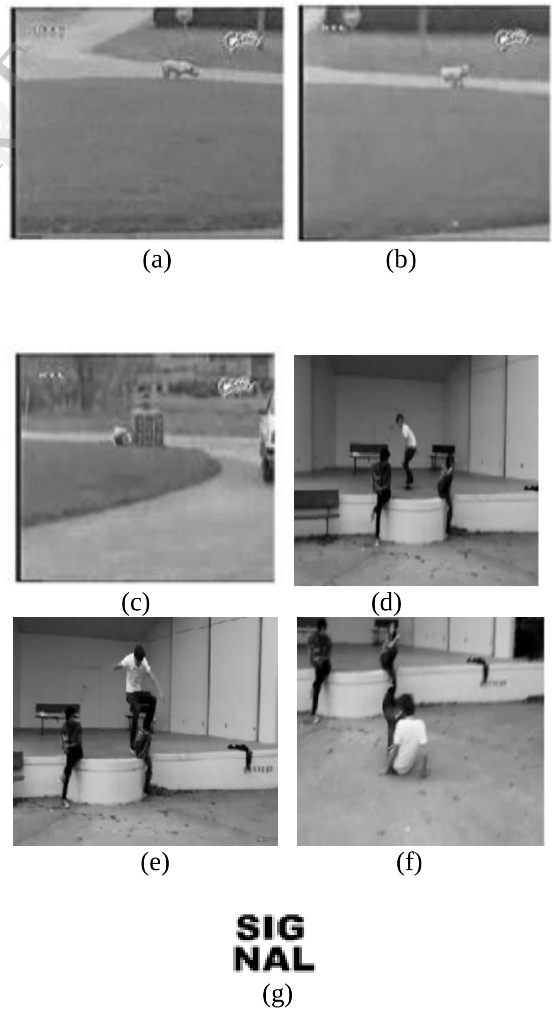
3.2.3 Algorithm for special frame attacks

1. Compare the frames of the suspected video that are next to the stored frame number of the original video with the stored frame and select the matching frame for Frame insertion attack.
2. In case of frame deletion attack, Check for the frame from the suspected video that is similar to one of the stored frames and select that frame for ownership identification.
3. In case of frame swapping attack, the stored frame may be moved to some other number in the attacked video and keep on comparing the stored frame with the frames in the suspected video until a match is found.

After selecting the right frame apply the algorithm for normal frame attacks on that frame.

4. Results

We have performed various experiments on the test frames of the video that is to be watermarked, to check the performance of the proposed algorithm. As a part of the experiment we have selected three frames randomly from the test video the “frame 2”, “frame 30” and “frame 90” with each frame size being 240 × 320 pixels. A binary image of size 28 x 40 is selected as a secret image that is to be embedded. The discrete fractional Fourier transform of order is taken as α= 2 and β= 2. The test frames and the secret image we consider are shown below in fig.1. Figures 2a and 2b show the Master share and the corresponding Ownership share of the frame 90 and figure 2c shows the stacked result of the two shares and the reduced secret image is shown in figure 2d.


(a) (b)

(c) (d)

(e) (f)

(g)

Fig 1 Test frames and secret image

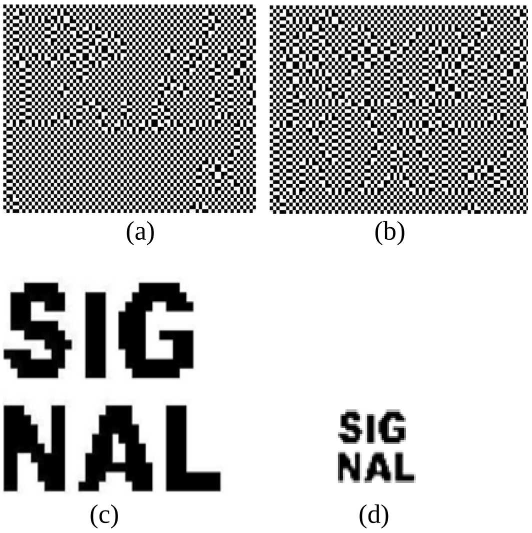


Fig 2 a. Master share b. Ownership share c. Stacked Result d. Reduced secret image

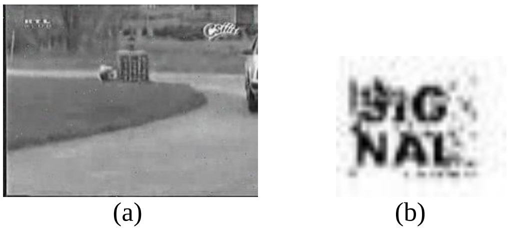


Fig 3 a.Frame 90 after adding noise (PSNR= 21.03 dB) b. Retrieved secret image(NC=0.9063)

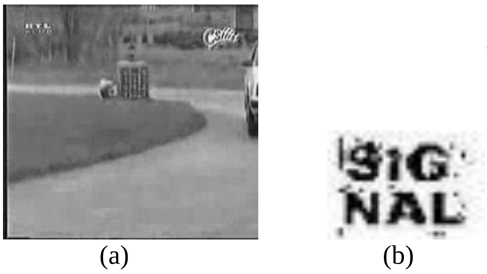


Fig 4 a.Frame 90 jpeg compressed PSNR= (25.47 dB) b. Retrieved secret image(NC= 0.9000)

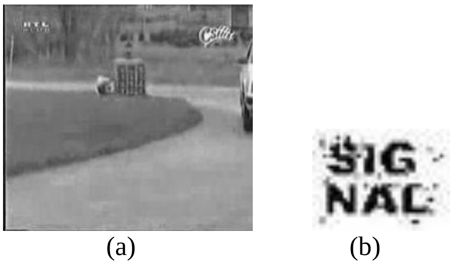


Fig 5 a. Frame 90 resized (PSNR=27.04 dB) b. Retrieved secret image(NC= 0.9009)

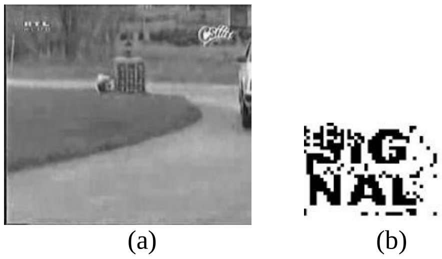


Fig 6 a. Frame 90 average filtered (PSNR=25.18dB) b. Retrieved secret image (NC= 0.8920)

The Normalized correlation (NC) is used to measure the similarity between the extracted image and the original secret image. It is defined as follows

$$NC = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} s_{ij}\oplus s''_{ij}}{m\times n} \qquad (5)$$

Where m x n represents the size of the secret image and S_{ij} , S''_{ij} are the original and retrieved secret images respectively, $\oplus$ denotes the exclusive-or (XOR) operation.

We have used another parameter called PSNR (peak signal to noise ratio) to measure the degradation of the attacked frame with the actual frame. The more the PSNR, the better is the frame quality. The PSNR is defined as follows

$$PSNR = 10\ \log_{10}\frac{255^2}{MSE}\ dB \qquad (6)$$

Where MSE is the mean squared error between the original frame H and the attacked frame H′ and is given by

$$MSE = \frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}[H(i,j) - H'(i,j)]^2 \qquad (7)$$

Figures 3 to 6 represent the results of the experiment when the test frame is subjected to various attacks like jpeg compression, noise addition, resize and average filtering and these results are used to estimate the robustness of the proposed algorithm. All the attack simulations are carried out by using Matlab platform. All the above attacks are shown for the 90th frame of the test video. The attacks are explained as follows:

Jpeg compression: The host frame is subjected to jpeg compression with a quality factor of 70%. The PSNR value of the compressed frame is 25.09(14.06 dB).The NC of the retrieved secret image is 0.9000

Noise addition: We have produced a noisy frame by adding salt & pepper noise to the original frame. The PSNR of the noisy frame is 21.06(13.23 dB) and the NC value of the retrieved secret image is 0.9063

Resize: The host frame is resized to 120 × 160 pixels and then scaled up to the actual frame size. The PSNR value of the resulting frame is 27.06(14.32 dB) and the NC value of the retrieved secret image is 0.9009

Average filtering: The host frame is applied to an average filter with window of size 3 x 3.The PSNR value of the resulting frame is 25.19(14.01 dB) and the NC value of the retrieved secret image is 0.8920.

Table2. Results of various frame attacks

Frame attack	test frame1 of video1		test frame2 of video1	
	PSNR	NC	PSNR	NC
Jpeg compression	27.7459	0.9143	28.8210	0.9080
Noise addition	32.1718	0.9071	32.2419	0.8991
Resize	33.4290	0.8884	31.1875	0.9161
Average filtering	27.7639	0.8714	27.1120	0.8946

Table3. Results for various frame attacks

Frame attack	test frame3 of video1		test frame1 of video2	
	PSNR	NC	PSNR	NC
Jpeg compression	25.4700	0.9000	24.0228	0.9098
Noise addition	21.0300	0.9063	35.0142	0.9036
Resize	27.0400	0.9009	40.1868	0.9248
Average filtering	25.1800	0.8920	32.9745	0.8929

Table4. Results for various frame attacks

Frame attack	test frame2 of video2		test frame3 of video2	
	PSNR	NC	PSNR	NC
Jpeg compression	33.3093	0.9013	32.0173	0.9008
Noise addition	29.7107	0.9001	30.0248	0.8995
Resize	39.0366	0.8972	40.0159	0.9189
Average filtering	32.8032	0.8995	30.5689	0.9006

5. Discussions:

The strengths of the proposed algorithm are security and robustness. The robustness can be observed based on the experimental results from section 4. In most of the attacks the retrieved secret image is close to the original secret image (NC≥0.90) and the least NC value is 0.8920 still which is a close approximation. Security is achieved through the use of pseudo random number generator usage and fractional Fourier transform. The frames which we select for embedding the watermark are randomly selected from the host video frame sequence using a pseudo random number generator seeded with a secret key K and hence it is difficult to identify the exact frames for the attacker in which the watermark is present. The second level of security is achieved through the order of fractional Fourier transform. The coefficients of the FrFt are more sensitive to the order and even for a small change in the order; there will be a large difference in the coefficient values. Therefore without knowing the correct order the attacker cannot get the embedded secret image. The following set of figures show the frame 30 of the test video when fractional Fourier transform is applied on it with different orders. Even in case of frame deletion attack, the probability is less that all the selected frames are deleted and similarly in case of frame insertion and frame swapping the selected frames are just shifted from their actual position in the original video. Hence in case of frame insertion, deletion and swapping also the ownership can be proved.

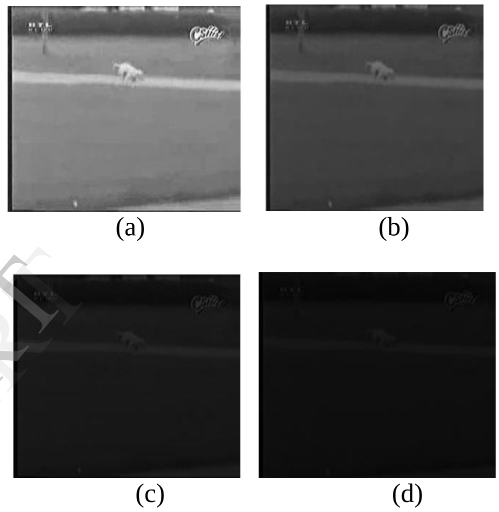


Fig 8 a. Original test frame b. After FrFt with order (-1/2,1/2)
c. After FrFt with order (3/2,-5/2) d. After FrFt with order (-7/4,-7/2)

6. Conclusion:

We conclude that our proposed algorithm is robust against most of the normal signal processing attacks and also against attacks like frame insertion, frame deletion and frame swapping. Also the advantage of the proposed scheme is that the video can be watermarked without processing the whole frame series of the video and with the limitation that the video is not of very long duration and the adjacent frames are highly correlated, with a very less number of frames from the video we can provide the authentication to the video which simplifies the watermark embedding and retrieving process.

References:

[1] M. Naor, A. Shamir, Visual cryptography, in :proceedings of the advances in cryptology, EUROCRYPT'94,Lecture notes in computer science, vol. 950, Springer-verlag,1995,pp. 1-12

[2] H.M. Ozaktas, O.Arikan, Digital computation of the fractional Fourier transform, IEEE transactions on signal processing 9 (1996) 2141-2149.

[3] V.A. Narayanan, K.M.M Prabhu, The fractional Fourier transform: theory, implementation and error analysis, Microprocessors and Microsystems 27 (2003) 511-521.

[4] S.C. Pei, M.H. Yeh, Two dimensional discrete fractional Fourier transform, signal processing 67(1998) 99-108.

[5] Shyong Jian Shyu , Efficient visual secret sharing scheme for color images, Pattern recognition, Volume 39, Issue 5, May 2006, Pages 866-880.

[6] B. Zhou, J. Chen, A geometric distortion resilient image watermarking algorithm based on SVD, a Chinese journal of image and graphics 9 (2004) 506-512.

AUTHOR'S BIOGRAPHY

Mrs. Radha Abburi received B. Tech degree in Electronics & Communication Engineering from JNTU, Anantapur, A.P, India and M. Tech degree in Embedded Systems from JNTU, Kakinada, A.P, India in 2003 and 2010 respectively. From 2005 till date she is working as Associate Professor in the department of Electronics and Communication Engineering in DVR & Dr HS MIC College of Technology, Kanchikacherla, A.P, India. Presently she is pursuing Ph.D. from KL University, Vaddeswaram, A.P, India, in the area of System Development.

Harikishan Repala received his B.Tech degree in Electronics and Communication Engineering from Sri Sarathi Institute of Engineering and Technology, Nuzvid, A.P, India in 2009.He is currently pursuing M.tech in Digital Electronics and Communication Systems (DECS) in Devineni Venkata Ramana & Dr.HimaSekhar MIC College of Technology, Kanchikacherla ,A.P, India.

R.Priyakanth received B.Tech degree in Electronics & Control Engineering from PVP Siddhartha Institute of Technology, Vijayawada, A.P, India and M.Tech degree in Communication and Radar Systems from KLUniversity, Vaddeswaram, Guntur Dist., A.P, India in 2002 and 2005 respectively. From 2005 till date he is working in the department of Electronics and Communication Engineering in Devineni Venkata Ramana & Dr.HimaSekhar MIC College of Technology, Kanchikacherla, A.P, India. Presently he is pursuing Ph.D from JNTUK, Kakinada, A.P, India. His research interests include Multimodal Signal Processing, Biomedical Image Processing.