Research Journal of Applied Sciences, Engineering and Technology 5(14): 3678-3683, 2013
ISSN: 2040-7459; e-ISSN: 2040-7467

Submitted: June 23, 2012 Accepted: July 28, 2012 Published: April 20, 2013

Video Compression-Encryption using Three Dimensional Discrete Fractional Transforms

Neeru Jindal and Kulbir Singh
Department of ECE, Thapar University, Patiala - 147004, Punjab (India)

Abstract: The temporal correlation between consecutive frames is exploited. The proposed algorithm is based on video compression-encryption using 3 Dimensional (3-D) discrete fractional transforms, which makes full use of the additional degree of freedom provided by the fractional orders to achieve an optimum domain in the compression and encryption. We first used Discrete Fractional Fourier Transform (DFrFT) to compress the adjacent difference frame of video and thereafter the obtained compressed difference frame was encrypted using Discrete Fractional cosine Transform (DFrCT) with fractional keys. The worth of the reconstructed video was measured with Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR). Noteworthy improvement in results using our method from that of current existing methods is the efficacy of this study.

Keywords: Compression-encryption, CR, discrete fractional transforms, DFrCT, MSE, PSNR

INTRODUCTION

The advancement of digital communication technologies persuades the development of multimedia applications such as video compression, audio compression, streaming, object recognition, content analysis. There are several methods available in literature for digital video compression and encryption (Schonberg *et al*., 2007; Wang and Ma, 2010; Cho *et al*., 2004; Singh and Manimegalai, 2011; Roy and Pradhan, 2011). There are 3 kinds of unswerving areas suitable for video encryption in literature. The first one encrypts video stream before the compressed code (Lian *et al*., 2004). A study pointed out that such a method drastically amends the source structure and syntax (Bao *et al*., 2007) and the follow-up coding efficiency is exaggerated. The second one encrypts the compressed video stream after being compressed and coded (Cheng and Li, 2000). The early video encryption technique availed of this method, which usually uses the soaring security strength of the traditional cryptographic algorithm such as DES, IDEA and RSA to meet the high security requirements. However, it has controlled some serious disadvantages such as high computation complexity and changed video formats. The third one combines encryption with compression. It partially encrypts video data (Winkler, 2005). In the proposed algorithm, the third approach is implemented and the temporal difference between adjacent frames is calculated. Thereafter, 3 dimensional (3-D) DFrFT is implemented for frame difference compression. The DFrCT encrypts this compressed difference frame. The comparative mean square error results prove that the proposed method endows with better video quality at the receiving end. In general, these algorithms encrypt the key data of video sequence that has great significance for the video reconstruction such as intra-prediction mode, motion vector difference, adjacent frame difference, quantization coefficients, block-matching motion compensation (Gorpuni, 2009; Lian *et al*., 2005) using transforms (Ishwar *et al*., 2008; Yeung and Zhu, 2009; Servais and Jager, 1997). The 2 dimensional (2-D) DFT has been implemented for image compression in recent years (Singh and Sinha, 2008; Candan *et al*., 2000).

The Fractional Fourier Transforms (FrFT) has found many applications in the area of signal processing, for solution of differential equations, swept frequency filters, pattern recognition and study of time-frequency distributions. The simplest generalization of the FrFT for 2 dimensions is given by:

$$X_{\alpha,\beta}(u,s)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}K_{\alpha,\beta}(u,s;t,r)x(t,r)\,dtdr \tag{1}$$

where,

$$K_{\alpha,\beta}(u,s;t,r)=K_{\alpha}(u,t)K_{\beta}(s,r) \tag{2}$$

After the continuous FrFT were derived, many researchers attempted their best to develop discrete counterpart of it. Santhanam and McClellan (1995) first reported the work. 2 dimensional (2-D) forward and

Corresponding Author: Neeru Jindal, Research Scholar, Department of ECE, Thapar University, Patiala - 147004, Punjab (India), Tel.: 91-9465060700

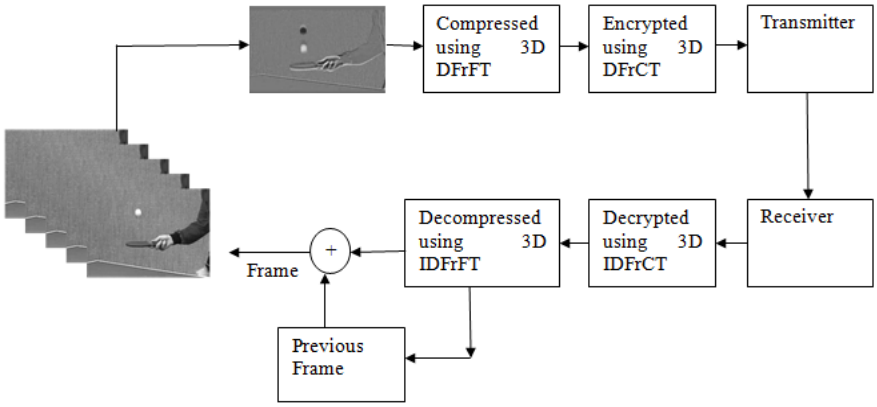

Fig. 1: Encoder and decoder using discrete fractional transforms

inverse DFrFT for (m, n) and (p, q) points are defined with separable forms as Pei and Yeh (1998):

Forward 2-D DFrFT:

$$X_{(\alpha,\beta)}(m,n)=\sum_{p=0}^{M-1}\sum_{q=0}^{N-1}x(p,q)\mathrm{K}_{(\alpha,\beta)}(p,q,m,n) \tag{3}$$

Inverse 2-D DFrFT:

$$x(p,q)=\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}X_{(\alpha,\beta)}(m,n)\mathrm{K}_{(-\alpha,-\beta)}(p,q,m,n) \tag{4}$$

For 2-D DFrFT, 2 individual angles of rotation $\alpha=\alpha\pi/2$ and $\beta=b\pi/2$ in 2 dimensions are taken with '*a*' and '*b*' as fractional orders and can be implemented by row-column computation in case of 2-D separable kernel *K (p, q, m, n)*. The signal can be recovered back by 2-D FrFT operation with inverted angles (-α,-β).

The DFrCT uses the Eigen decomposition of the Discrete Cosine Transform (DCT) kernel. The exclusive eigenvectors are obtained from the even Hermite-Gaussian eigen-vectors of the Fourier matrix in the cosine case. The kernel matrix of the N point DFrCT is defined as Pei and Yeh (2001):

$$C_{N,\alpha}=V_N D_N^{2\alpha/\pi}V_N^T=V_N\begin{bmatrix}1 & & & 0\\ & e^{-2j\alpha} & & \\ & & \ddots & \\ 0 & & & e^{-j2(N-1)\alpha}\end{bmatrix}V_N^T \tag{5}$$

where, $V_N = [v_0 \mid v_1 \mid ...\mid V_{2N-2}\mid v_k$ is the eigenvector derived from k order DFT Hermite eigenvector. Steps for constructing the *N* point DFrCT kernel with angular parameter α are summarized (Pei and Yeh, 2001) as follow:

Step 1: Compute M_c point DFT Hermite even eigenvectors (where $M_c = 2(N -1)$).
Step 2: Use Step 1 to calculate the DCT-I eigenvectors from the DFT Hermite even eigenvectors.
Step 3: Find the DFrCT transform kernel by the following equation:

$$C_{N,\alpha}=V_N D_N^{2\alpha/\pi}V_N^T$$

DFrCT has mathematical properties of unitarity, additivity (of rotations), periodicity and reality. In this study, we present an algorithm for the video compression and encryption using 3-D discrete fractional transforms by exploiting temporal correlation between adjacent video frames.

MATERIALS AND METHODS

Let in addition to the 2 spatial dimensions, one considers the dimension of time, the 3 dimensional separable DFrFT can be defined as Eq. (6) and Eq. (7):

Forward 3-D DFrFT:

$$X_{(\alpha,\beta,\gamma)}(m,n,j)=\sum_{p=0}^{M-1}\sum_{q=0}^{N-1}\sum_{k=0}^{j-1}x(p,q,k)K_{(\alpha,\beta,\gamma)}(p,q,m,n,j,k) \tag{6}$$

Inverse 3-D DFrFT:

$$x(p,q,k)=\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\sum_{j=0}^{j-1}X_{(\alpha,\beta,\gamma)}(m,n,j)K_{(-\alpha,-\beta,-\gamma)}(p,q,m,n,j,k) \tag{7}$$

- If $\alpha=\beta=\pi/2$ any $\gamma=\alpha\pi/2$ the 3-D DFrFT performs the conventional discrete Fourier Transform (DFT)

Table 1: Comparison of mean square error

Claire				Trevor			
Compression	Reference (Winkler, 2005; Maniccam and Bourbakis, 2004) (MSE)	Proposed algorithm (MSE)	Improvement	Compression	Reference (Winkler, 2005; Maniccam and Bourbakis, 2004) (MSE)	Proposed algorithm (MSE)	Improvement
52.26%	0.6363	0.000958	0.635342	31.73%	0.5041	0.000957	0.503143
73.09%	1.0413	0.000951	1.040349	48.20%	0.9023	0.000927	0.901373
82.65%	1.3416	0.000915	1.340685	59.99%	1.1752	0.00092	1.17428
87.24%	1.5500	0.000899	1.549101	65.31%	1.3684	0.000984	1.367416
89.60%	1.7065	0.000883	1.705617	67.79%	1.5469	0.000909	1.545991
90.96%	1.8374	0.000882	1.836518	69.57%	1.7417	0.000922	1.740778
91.89%	1.9648	0.000884	1.963916	71.12%	1.9534	0.000945	1.952455
92.61%	2.0912	0.000880	2.09032	72.52%	2.1785	0.000963	2.177537
93.24%	2.2191	0.000881	2.218219	73.83%	2.4140	0.000857	2.413143
93.77%	2.3532	0.000884	2.352316	75.07%	2.6637	0.000935	2.662765

- If $\alpha = \beta = \pi/2$ any $\gamma = 0$, the 3-D DFrFT perform 2-D DFrFT i.e., only row-column Transforms
- If $\alpha = \beta = 0$ any $\gamma = \alpha\pi/2$, the 3-D DFrFT performs only time-axis transforms
- If $\alpha = \beta = 0$ any $\gamma = 0$, the 3-D DFrFT performs the identity transform

Since 2-D DFrFT maintains desired properties of Fractional Fourier Transform (FrFT) (Santhanam and McClellan, 1995) so 3-D DFrFT also maintains the properties of unitary, angle additivity, time reversal and DFT rotation. However, the DFrFT is a separable transform. This implies that a multi-dimensional DFrFT may be implemented as a series of one-dimensional discrete fractional Fourier transforms. Several fast algorithms for implementing the DFrFT in both one and 2 dimensions exist. Thus, for example, a fast 2-D DFrFT could be implemented on the rows and columns of each of the video frames. This could be followed by a fast one dimensional (1-D) DFrFT along the time axis (i.e., corresponding pixels in each of the frames.)

In this study, video compression-encryption using 3 discrete fractional transforms was implemented. The video was partitioned into video frames. In the proposed algorithm, the redundancy of adjacent frames was exploited and difference frame was obtained. The encoder and decoder for discrete fractional transforms are shown in Fig. 1.

The difference frame was compressed using 3-D discrete fractional Fourier transform. The frame could be divided into sub-blocks of size 32 × 32, 16 × 16, 8 × 8 etc. The smallest sub-block size increased compression and computational complexity. The proposed algorithm had used 8 × 8 sub-block size. The degree of data diminution as a result of the compression process is known as Compression Ratio (CR). The CR is equal to the size of the original image divided by the size of the compressed image. The various compression ratios were implemented on algorithm given in Table 1. Then, 3-D DFrFT for 3 fractional orders α, β, γ were applied at particular CR. The 3 fractional orders were kept same in the proposed work. Optimization of fractional orders was decided by varying fractional order value between 0 and 1. The value at which PSNR reached maximum was selected as optimum value. The optimum value depends upon the video sequence and the CR. The final step in compression process was to quantize the transformed coefficients and run length coding. The decoding end used divergent processes with same inverse fractional order keys. The block diagram for compression using DFrFT is shown in Fig. 2.

The compressed frame was encrypted using DFrCT. The frame was discrete fractional cosine transformed 3 times using fractional orders δ_1, δ_2 and δ_3 respectively. In the intermediate stages, we put 2 Random Phase Masks (RPM), $R_1(m_1, n_1) = \exp[i2\pi\psi(m_1, n_1)]$ and $R_2(m_2, n_2) = \exp[-i2\pi\psi(m_2, n_2)]$ respectively, serving as phase filters, where functions $\psi_1(m_1, n_1)$ and $\psi_2(m_2, n_2)$ were randomly generated homogeneously distributed functions with values (0, 1). Thus the resultant transformed function Ψ(m, n) can be written as:

$$\psi(m,n) = F_c^{\alpha_3}\left\{\psi_2(m_2,n_2)R(m_2,n_2)\right\} \quad (8)$$

With

$$\psi_2(m_2,n_2) = F_c^{\alpha_2}\left\{\psi_1(m_1,n_1)R(m_1,n_1)\right\}$$

and

$$\psi_1(m_1,n_1) = F_c^{\alpha_1}\left\{f(m_0,n_0)\right\}$$

The resultant function Ψ (m, n) can be regarded as the encrypted image (Liu *et al*., 2001). The decryption process was the reverse operation with respect to the encryption with inverse fractional keys and complex conjugate of RPM's. The test videos sequences Claire,

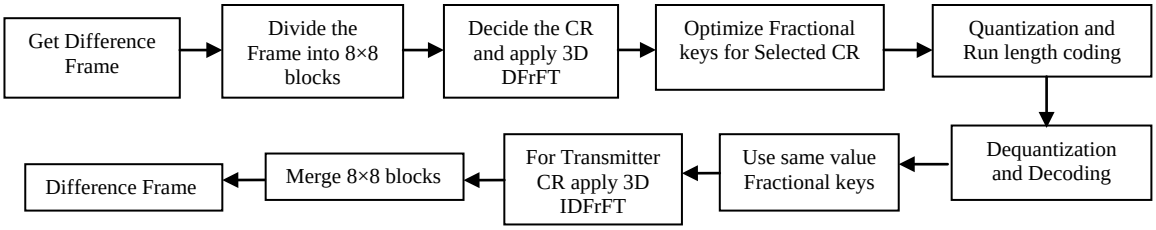

Fig. 2: Compression using DFrFT

Claire (352×288) Tennis (320×240) Foreman (176×144) Trevor (352×288) Heart

Fig. 3: Test videos

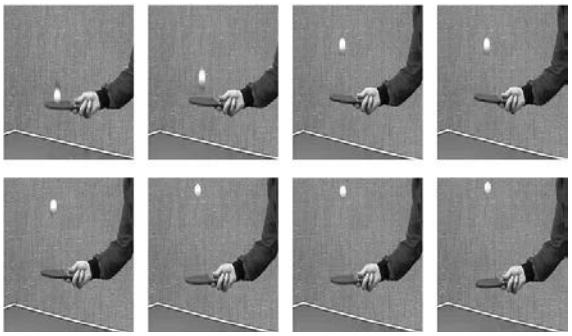

Fig. 4: Frame number 1, 4, 8, 13, 17, 20, 23 and 25: Note the motion of tennis ball

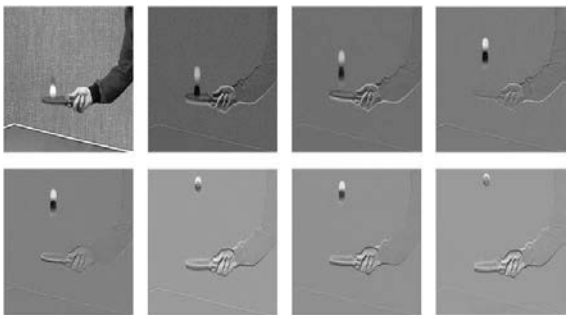

Fig. 5: Adjacent frame differences

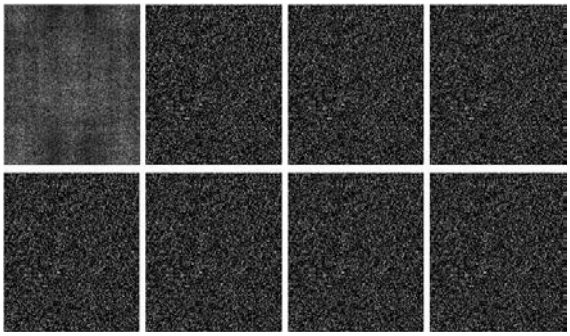

Fig. 6: Compressed and encrypted frames

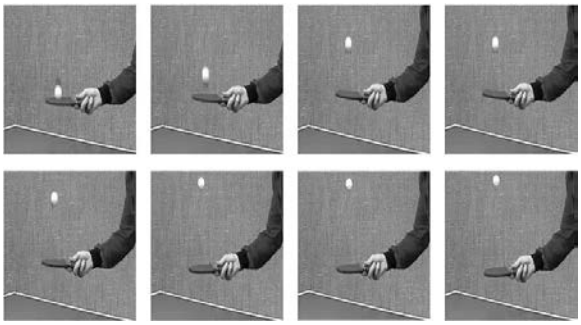

Fig. 7: Recuperate video frame number 1, 4, 8, 13, 17, 20, 23 and 25

tennis, Trevor, Foreman and heart were considered for experiment purpose shown in Fig. 3. The adjacent frames of test video (Tennis) and their difference frames are shown in Fig. 4 and 5. The Fig. 6 shows compressed and encrypted frames from experiment results of proposed algorithm. Recuperate video frames at the receiving end are shown in Fig. 7.

RESULTS AND DISCUSSION

The subjective and objective methods are usually adopted to evaluate the performance of video algorithms. The best way to assess the worth of a video is to subjectively evaluate it because in most of the cases, human eyes act decisive receivers. The subjective quality measurement Mean Opinion Score (MOS) has been used for many years (ITU-R BT.500-11). While these tests are the best way to measure "true" perceived quality, they are complex, time-consuming and expensive (ITU-R BT.500-11). Hence, they are often impractical, when real-time online quality monitoring of several video channels is desired. Looking for faster alternatives, signals could turn to simple error measures such as the Mean Squared Error (MSE) or the Peak Signal-To-Noise Ratio (PSNR). These criteria are considered to be objective due to the fact that they rely on the pixel luma and chroma values of the input and output video frames and do not include any subjective human intervention in the quality assessment process. The MSE is defined as:

$$MSE = \frac{1}{TXY}\sum_{t}\sum_{x}\sum_{y}\left(o(t,x,y)-r(t,x,y)\right)^2 \quad (9)$$

where, X, Y and T are height, width and time axis of video frame.

The *o (t, x, y)* is original video frame and *r (t, x, y)* is the reconstructed frame. The lower value of MSE demonstrates a high quality of video at receiver's end. The worth of the video at receiving end was measured with MSE and PSNR. The proposed algorithm MSE was compared with Winkler (2005) and Maniccam and Bourbakis (2004) as shown in Table 1. The PSNR was used as a worth measure of reconstructed video in proposed algorithm. It is defined as: PSNR = $10\log M^2/MSE$, where M is the maximum value that a pixel can take (e.g., 255 for 8-bit images). Table 2 shows results for test video sequences. The reconstructed video must be secure from attacks. The performance of this algorithm has been also discussed.

Security: The video sequence was partitioned into frames in the proposed algorithm. Video encryption of each compressed difference frame was realized separately by multi-keys using DFrCT. Multi-keys were changed chaotically for encryption of each frame. The security for proposed algorithm is provided by 2-ways. First, statistical attackers break the cryptosystem by utilizing the temporal correlation between consecutive video frames. The adversary has to decrypt each frame separately for same video in the proposed algorithm. Second, it makes brute-force attack infeasible. Brute force attack is based on exhaustive key search and is feasible only for the cryptosystems with relatively small key space. Proposed algorithm uses 3 multi-keys each of 32 bit with a key space size 2^{96} = 79228162514264337593543950336 for each frame.

Because of individual frame encryption in proposed algorithm, if a frame is corrupted or lost during transmission, it would not affect the decryption of other frames. Therefore, time will be saved by avoiding iteration transmission of remaining frames due to single frame lost.

Performance: Yeung *et al*. (2011) have also proposed a method for video encryption using 8 × 8 transforms in H.264 and MPEG-4. PSNR between original frame and the decrypted frame calculated by Yeung *et al*. (2011) is 54 dB and 49 dB using H.264 and MPEG-4, respectively. The proposed algorithm Table 2 shows that the MSE and PSNR between original image and encrypted image is very small 0.00122 and is very high 76.17 dB, respectively at different compression ratios for Foreman video. The high PSNR 22.17 dB from H.264 and 27.17 dB from MPEG-4 of proposed algorithm show the successful retrieval of the input frame. The PSNR of proposed algorithm was also compared with the PSNR of Servias *et al*. (1997).

Table 2: PSNR of video sequence

Compression	Tennis (in dB)	Foreman (in dB)	Claire (in dB)	Trevor (in dB)	Heart (in dB)
15%	76.95	76.17	77.33	77.45	76.93
25%	77.30	76.08	77.48	77.30	76.94
35%	77.08	75.80	77.21	77.33	76.78
50%	76.00	75.74	77.31	77.25	77.00
60%	77.25	76.00	77.43	77.49	76.82
70%	76.86	75.60	77.45	77.48	76.89
80%	77.12	76.15	77.50	77.30	77.11
90%	76.83	75.94	77.52	77.53	76.63

(Servais *et al*., 1997) video sequence of 256 × 256 was taken and PSNR vs. frame serial number was calculated using 3-D discrete Cosine transform (DCT). The maximum value of PSNR was 35 dB in Servais *et al*. (1997), which is smaller as compared to that of proposed work, i.e., 77 dB (PSNR) for Heart sequence of 256 × 256. So an improvement of 42 dB in PSNR was obtained with proposed algorithm.

CONCLUSION

In the proposed study, the temporal correlation between consecutive frames is exploited. The difference frame was extracted from the adjacent frames and compressed using 3-D DFrFT at different compression ratios. The compressed difference frame was encrypted using DFrCT. The worth of reconstructed frame was evaluated with MSE and PSNR. An improvement of 2.32 in MSE (Maniccam and Bourbakis, 2004; Maniccam, 2001) and PSNR of 22.17 dB (Yeung *et al*., 2011) has been attained in the proposed algorithm. The compressed-encrypted video was secured due to large key space size of 2^{96} and infeasible to brute force attack.

REFERENCES

Bao, X., J. Jiang, W. Yuan and Y. Li, 2007. Study of CABAC-based digital video encryption in the H.264/AVC standard. J. Commun., 28(6): 24-29.

Candan, C., M.A. Kutay and H.M. Ozaktas, 2000. The discrete fractional Fourier transform. IEEE Trans. on Signal Process., 48(5): 1329-1337.

Cheng, H. and X.B. Li, 2000. Partial encryption of compressed images and videos. IEEE Trans. Signal Proc., 48(8): 24-39.

Cho, S., D. Kim and W.A. Pearlman, 2004. Lossless compression of volumetric medical images with improved 3-D SPIHT algorithm. J. Dig. Imaging, 17(1): 57-63.

Gorpuni, P., 2009. Development of fast motion algorithms for video compression. M.Tech Thesis, National Institute of Technology, Rourkela.

Ishwar, S., P.K. Meher and M.N.S. Swamy, 2008. Discrete tchebichef transform-A fast 4×4 algorithm and its application in image/video compression. IEEE International Symposium on Circuits and Systems (ISCAS), pp: 260-263.

Lian, S., J. Sun and Z. Wang, 2004. Quality analysis of several typical MPEG video encryption algorithms. J. Image Graphics, 9(4): 483-490.

Lian, S., Z. Liu, Z. Ren and Z. Wang, 2005. Selective video encryption based on advanced video coding. Advances in Multimedia Information Processing-PCM, LNCS 3767, pp: 281-290.

Liu, S., L. Yu and B. Zhu, 2001. Optical image encryption by cascaded fractional Fourier transform with random phase filtering. J. Opt. Commun., 187(1-3): 57-63.

Maniccam, S.S., 2001. Image-Video Compression. Encryption and Information Hiding, State University of New York at Binghamton, State University of New York, pp: 104.

Maniccam, S.S. and N.G. Bourbakis, 2004. Image and video encryption using SCAN patterns. Pattern Recognit., 37(4): 725-737.

Pei, S.C. and M.H. Yeh, 1998. Two dimensional discrete fractional Fourier transform. Signal Proc., 67(1): 99-108.

Pei, S.C. and M.H. Yeh, 2001. The discrete fractional cosine and sine transforms. IEEE Trans. Signal Proces., 49(6): 1198-1207.

Roy, M. and C. Pradhan, 2011. Secured selective encryption algorithm for MPEG-2 video. 3rd International Conference on Electronics Computer Technology, pp: 420-423.

Santhanam, B. and J.H. McClellan, 1995. The DRFT-a rotation in time-frequency space. International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp: 921-925.

Schonberg, D., C. Yeo, S.C. Draper and K. Ramchandran, 2007. On compression of encrypted video. Data Compression Conference, pp: 173-182.

Servais, M. and G.D. Jager, 1997. Video compression using the 3 Dimensional Discrete Cosine Transform (3D-DCT). Proceedings of the South African Symposium on Communications and Signal Processing (COMSIG), pp: 27-32.

Singh, N. and A. Sinha, 2008. Optical image encryption using fractional Fourier transform and chaos. Opt. Las. Eng., 46(2): 117-123.

Singh, K.J. and R. Manimegalai, 2011. Fast random bit encryption technique for video data. Eur. J. Sci. Res., 64(3): 437-445.

Wang, S. and H. Ma, 2010. Improved H.264 video compression based on cubic macro-blocks. International Conference on Audio Language and Image Processing, pp: 90-93.

Winkler, S., 2005. Digital Video Quality: Vision Models and Metrics. John Wiley & Sons, Chichester, pp: 175, ISBN: 0470024046.

Yeung, S.K.A. and S. Zhu, 2009. Partial video encryption based on alternating transforms. IEEE Signal Proc. Lett., 16(10): 893-896.

Yeung, S.K.A., S. Zhu and B. Zeng, 2011. Perceptual video encryption using multiple 8 × 8 transform in H.264 and MPEG-4. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp: 2436-2439.