HYBRID AND BLIND WATERMARKING FRAMEWORK FOR PRIVACY PROTECTION AND CONTENT AUTHENTICATION OF DIGITAL MULTIMEDIA

Swati J. Patel[1]
Dr. Mehul C. Parikh[2]

•

[1]Ph.D. Scholar, Information Technology Department, Gujarat Technological University, Ahmedabad, India. swatijpatel@ldce.ac.in
[2]Associate Professor, Information Technology Department, L.D.College of Engineering, Ahmedabad, India. mehulcparikh@ldce.ac.in

Abstract

Nowadays, due to inexpensive and conveniently available internet access at the fingertip, the illegitimate sharing of digital multimedia i.e. image, audio, and video is becoming a universal and significant threat. Illegal transmission of digital multimedia through the internet creates an issue of authentication and copyright protection; hence, piracy protection is a vital need for protecting digital media. Digital watermarking is a method of preventing digital theft in which additional information, known as a watermark, is inserted into digital multimedia. This technology was originally designed for still photos, but it has subsequently been expanded to include additional multimedia artifacts such as audio and video, due to its countless use in today's era. Digital watermarking is an effective method of limiting piracy and providing authenticity and copyright ownership to digital content. Watermarking can be performed either in the spatial or in the transform domain. In this paper, a hybrid digital video watermarking technique for copyright protection, data security, and content authentication of multimedia, based on Discrete Wavelet Transform, Discrete Cosine Transform, and Singular Value Decomposition is presented. The authenticity of the content has been ensured by embedding a watermark in the transform domain, while copyright protection has been provided by a strong watermark. The experimental results show that the proposed schemes achieve a PSNR greater than 51 dB on average, which illustrates that the proposed method gives excellent performance for robustness, authentication, and security. A comparison of the proposed framework to various cutting-edge techniques illustrates its effectiveness and superiority.

Keywords: Digital Video Watermarking, Robustness, Imperceptible, Authentication, Copyright Protection, Discrete Wavelet Transform, Discrete Cosine Transform, Singular Value Decomposition

1. Introduction

Almost everyone's smartphone is connected to one of the several widely available and fairly priced high-speed internet networks in today's world. People's conceptions of information transmission have been radically altered by the power of digitalization, which now allows them to send and receive text, data, images, and even films using their mobile devices. They are inadvertently exposing themselves to a serious technological risk of having their personal information misused. Anyone can easily copy and distort someone else's work, resulting in a slew of piracy-related concerns as well as a detrimental influence on the owner's financial benefits and intellectual property rights [1][2]. The challenges of video watermarking are greater than those of image watermarking. The data in the video is inherently more scattered between frames. Uneven distribution of moving and stationary areas. The complete video watermarking process can be divided into three stages, as indicated in figure 1: watermark attachment or embedding, watermark sending or distribution over the channel, and watermark extraction or detection [3].

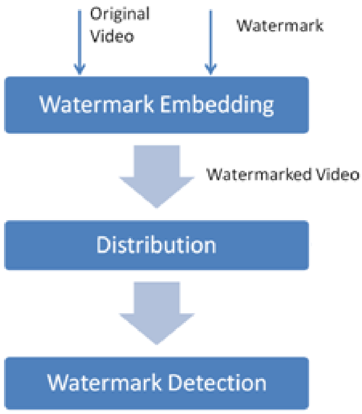


Figure 1: *Watermarking Process*

Section 2 represents the literature review. Section 3 introduces the three main concepts of the paper i.e., Discrete Wavelet Transform (DWT), Discrete Cosine Transform (DCT), and Singular Value Decomposition (SVD). Section 4 introduces the proposed method of video watermarking. Section 5 represents experimental results and discussion. Section 6 gives the concluding remarks.

2. Literature Review

Digital watermarking has three parameters: data payload, fidelity, and robustness. Digital watermarking has long been used to protect still photos, but it is now employed to protect audio and video files[4]. A comparison of different works is illustrated in Table 1.

Table 1: *Summary of earlier work done*

Ref.	Technique used	Purpose	Scheme Type	Results	Attacks	Remark
[5]	PCA, RDWT	Robustness, Imperceptibility	Blind	PSNR = 67.87, MSE = 0.0106, NCC = 0.995, SSIM = 0.9997	GN, Rot, Sc, Compression, RS	It can be improved to work for color images.
[6]	DWT, SVD, Entropy, Pixel position shuffling	Robustness, Imperceptibility, Security	Non-Blind	PSNR= 42.6369, NC = 1	JPEG, Rot, GN, S&P	It is not time Efficient
[7]	DCT, DWT, Arnold	Robustness, Imperceptibility	Non-Blind	PSNR = 47.1836, NC = 0.1936	Rot, Noise, RS, JPEG	No balance between PSNR and NC value is achieved
[8]	2D –DWT,	Robustness,	Non-	PSNR = 54.96,	S&P, Rot,	Not good for

	Encryption	Imperceptibility, Security	Blind	MSE = 0.2047, CC = 0.9749	JPEG	geometric attacks and have low capacity
[9]	DWT, SVD, Block selection scheme	Robustness, Security	Non-Blind	PSNR = 61.7524, SSIM = 0.9999, Cor = 1	GN, S&P, Rot, JPEG	It is not good for active attacks
[10]	DCT, DWT, SVD, Arnold	Robustness, Imperceptibility, Security	Non-Blind	PSNR = 43.88, NC = 0.9888, BER = 0.2174	JPEG, S&P, GN, Rot, Cr, RS	For the high gain factor, the quality of the watermarked image is poor

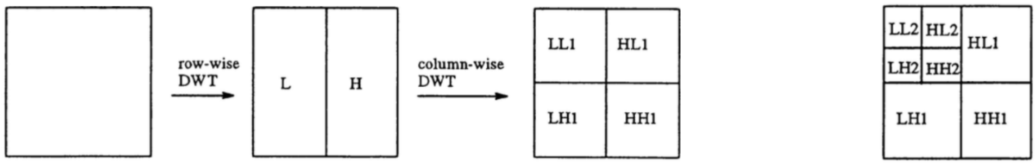
Abbreviations used for the various attack:
GN: Gaussian Noise
Rot: Rotation
Sc: Scaling
RS: Resize
Cr: Cropping
S&P: Salt and Pepper

3. Transform Domain Methods

The video frame of a host video sequence is translated into the new domain, and then the inverse frequency domain is applied after embedding a watermark. Data is converted using transforms like the Discrete Wavelet Transform (DWT), Discrete Cosine Transform (DCT), and Singular Value Decomposition (SVD). Each transform has its properties for video frame representation. This approach offers the advantages of being more resistant to malicious attacks, more stable, and more imperceptible than a spatial domain [11] [12].

3.1 Discrete Wavelet Transform (DWT):

In discrete wavelet transform, video frame pixels are converted into wavelets, which can subsequently be used for wavelet-based compression and coding [13]. A video frame is divided into four bands using a mathematical technique called lower resolution approximation (LL), vertical (LH), horizontal (HL), and diagonal (D) (HH). The low-frequency component is designated as LL, while the high-frequency sections are designated as LH, HH, and HL. Multiple-scale wavelet decompositions can be obtained by repeating the process as shown in figure2(b). The low-frequency district information is a frame that is quite similar to the original frame. This frequency district contains the majority of the original frame's signal information. The level detail, upright detail, and diagonal detail of the original image are represented by the frequency districts LH, HL, and HH, respectively. The watermark can be inserted in these three subbands to retain higher image quality because human eyes are significantly more sensitive to the low-frequency component (the LL subband). They are sensitive to changes in the smooth district of a frame, but not too small changes in the edge, profile, or streak, according to the HVS character [14]. The watermark's robustness is increased by embedding it in higher-level sub-bands. However, the visual integrity may be compromised, as evaluated by PSNR. In the high-frequency region, the edges and texture can be easily distinguished using the DWT. As a result, it's difficult to see that inserting the watermarking signal into the frame's DWT-converted high-frequency band has a large amplitude coefficient.



(a) *(b)*

Figure 2: *(a) First level of decomposition (b) Second level of decomposition*

3.2 Discrete Cosine Transform (DCT):

Watermarking based on DCT can be divided into two types: global and block-based. A full DCT is conducted on the entire frame in the former, whereas a frame is divided into non-overlapping blocks of a particular size and the DCT is performed on each block [1] [15]. The DCT technique separates the coefficients into three frequency bands: low, middle, and high, with the mid-frequency band being the most commonly used for watermark embedding. Because the low-frequency band and DC component contain the greatest signal energy per frame, they are the most critical portions of a video signal, and any change to this band has a significant impact on how the human eye perceives distortion both are vulnerable to watermarking attacks [16]. Watermarking with DCT is based on two facts. The first is that much of the signal energy is concentrated in the low-frequency sub-band, which contains the frame's most significant visual elements. The second fact is that high-frequency frame components are frequently deleted due to compression and noise attacks [14]. DCT can be conducted in a variety of dimensions, including 1D, 2D, and 3D. In the context of images or frames, 2D DCT is used to convert them into a cosine series [15].

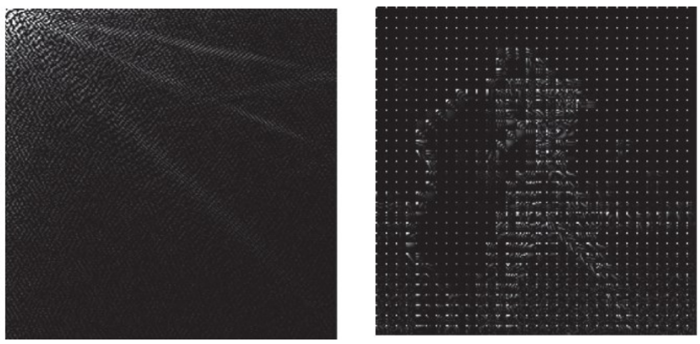


Figure 3: *(a) The whole frame (Global) (b) Block-wise process*

1D-DCT is given as:

$$f(i) = F(u) \tag{1}$$

2D-DCT is given as:

$$F(u,v) = \alpha(u)\alpha(v) \sum_{i=0}^{M-1}\sum_{j=0}^{M-1} f(i,j) cos\left[\frac{(2\mathrm{i}\ + 1)\mathrm{u}\pi}{2\mathrm{M}}\right] cos\left[\frac{(2\mathrm{i}\ + 1)\mathrm{v}\pi}{2\mathrm{M}}\right]$$

Where u,v = 0,1,2,...M-1, M is size of sequence; f(i,j) is image in spatial domain and F(u,v) is in frequency domain

(2)

$$\alpha(u) = \left(\frac{1}{\sqrt{M}}\ if\ u = 0\right) or\ (\sqrt{\frac{2}{M}}\ if\ \ u \neq 0)$$

4. Singular Value Decomposition (SVD):

Singular Value Decomposition is a technique used in a variety of applications, including image compression, image concealment, and image watermarking. A frame is an array of nonnegative scalar entries that can be treated as a matrix from the viewpoint of linear algebra. A matrix can be decomposed into three matrices of the same size as the original matrix using the SVD transformation [14]. Considering a frame of the video sequence is a square matrix of size M×M, the formula for SVD is defined as,

$$f = \mathrm{USV}^T \tag{3}$$

Where U and V are orthogonal (or unitary) matrices and S is a diagonal matrix, with the diagonal elements in the descending order of S being called the singular values of the frame (f) and V^T is the transpose of a MxM matrix containing the orthonormal eigenvectors.

Watermarking strategies based on SVD embed the watermark by changing U, V, or S. Due to the strong stability of the singular values, SVD techniques are commonly used in video watermarking. This means that when a little perturbation is applied to a frame, these values do not fluctuate much. Although this characteristic of the SVD provides robustness to attacks, a limitation is that performing it on an image is computationally expensive[17].

5. Proposed Method

We proposed a DWT, DCT, and SVD-based hybrid watermarking technique using the salient properties of DWT, DCT, and the SVD in our suggested watermarking scheme. The algorithm for watermark embedding and extraction processes are described in algorithm 1 and algorithm 2 respectively, which shows how the watermark is embedded with the host frame and how the embedded watermark is extracted from the attacked watermarked frame. First of all Host video is taken and converted into a sequence of frames. Among them all video frames, the keyframes are selected for further processing. The proposed approach for watermarking is depicted in figure 4.

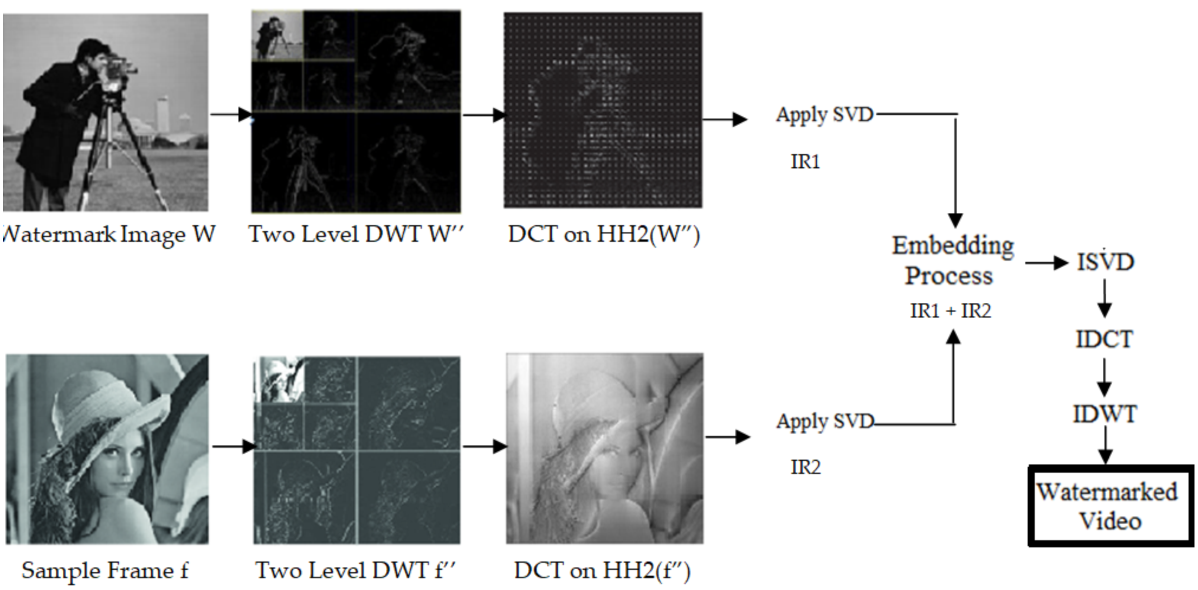


Figure 4: Schematic Representation of the Proposed System

Algorithm 1: Watermark Embedding Process

```
Host Video V, f = Frames of (V)
Watermark Image W
W' = DWT (W)
W" = DWT (W')  // Two Level DWT
Result1 = DCT on HH2(W") as per eqn. (2)
IR1 = SVD (Result1) as per eqn. (3)   // Intermediate Result1
fRow = All frames of Video V
for f = 1 to n where f = All keyframes of Video V
{
f' = DWT(f)
f" = DWT (f')
Result 2 = DCT on HH2(f") as per eqn. (2)
IR2 = SVD (Result 2) as per eqn. (3)                // Intermediate Result 2
IR3 = IR1 + IR2    // Embedding Process
RR1 = ISVD (IR3)
RR2 = IDCT (RR1)
Ri = IDWT (RR2)
Final Result FR = ∑_{i=1}^{n} Ri
}
```

Algorithm 2: Watermark Extraction Process

```
Watermarked Video FR
Rawf = frames (FR)
F = keyframe (Rawf)
For f=1 to n where f = All keyframes of Video FR
{
f' = DWT(f)
f" = DWT(f')
R1 = DCT on HH2(f") as per equ (2)
IR1 = SVD (R1) as per equ (3)
R2 = IDCT (S (IR1)) on singular values only.
R = IDWT (R2)
Final Extracted Result FER = ∑_{i=1}^{n} Ri
}
```

6. Experimental Results and Discussion

In this proposed watermarking algorithm various mp4 videos of different backgrounds and different lengths are used. Videos are converted into a sequence of frames, frame size is 1280×720. Figure 5 shows the cameraman's image taken as a watermark of size 256×256. Figure 6 shows a frame of videos taken as input videos. Imperceptibility and robustness are the attributes that are assessed for the proposed scheme based on PSNR and MSE values. The term imperceptibility refers to the fact that the video's quality should not be compromised once the watermark is applied. After embedding the watermark, Peak Signal to Noise Ratio (PSNR) is used to calculate the imperceptibility, the degradation in the watermarked video, compared to the host video. PSNR is expressed as a decibel scale. Higher the value of PSNR higher the quality of the video. PSNR is represented as shown in equation (4). Watermark robustness refers to the fact that the watermark is not destroyed after intentional or unintentional attacks and may still be utilized to offer certification. It is calculated "after the attack." Mean Square Error (MSE) measures the mean of the square of the original watermark and the extracted watermark from the attacked frame for robust capabilities. The lower the value of MSE lower will be the error. MSE is represented as shown in equation (5).

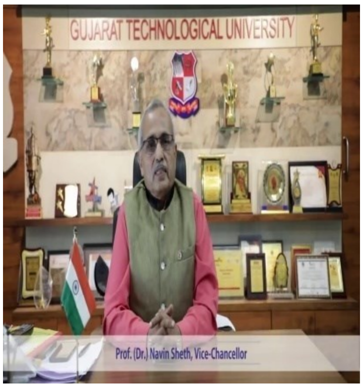
$$\text{PSNR} = 10log_{10}[\frac{P^2}{MSE}] \tag{4}$$

Where P is the peak signal value. P is equal to 255 for frames having a channel depth of 8-bit.

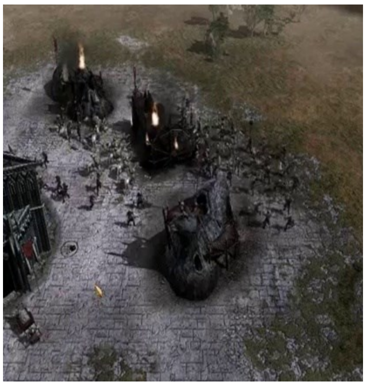
$$\text{MSE} = \frac{1}{ab}\sum_{x=1}^{a}\sum_{y=1}^{b}[\text{c(x, y)} - \text{e(x, y)}]^2 \tag{5}$$

Where a and b are the height of the original frame and distorted frame, respectively. c(x,y) is the pixel value of the host video frame and e(x,y) is the pixel value of the embedded video frame.

Figure 5: *Watermark*

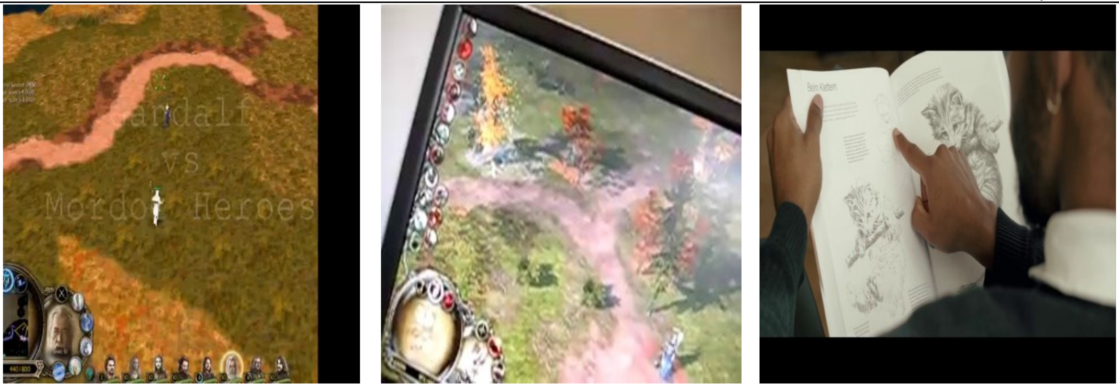


Test Video 1
(HonourableVCSirMessage)



Test Video 2
(Battle for Middle Earth 2 Goblins Fighting)

Test Video 3
(Filmstro & Film Riot One Minute Short Film)



Test Video 4 (Gandalf vs Mordor Heroes)

Test Video 5 (Lord Of The Rings Battle For Middle Earth 2)

Test Video 6 (Perfection - 1 Minute Short Film)

Figure 6: *Frame of various host videos*

The PSNR values of the watermarked videos are shown in Table 2. These values demonstrate the scheme's undetectable property, as the PSNR values are high, implying that there is very little quality distortion after embedding the watermark. The graphical representations of the same are shown in figure 7.

Table 2: *Performance results in terms of PSNR (in dB)*

Test Data	DCT+SVD	DCT+DWT+ SVD	DWT+DCT+SVD
Test Video 1	54.9353	51.9685	55.1118
Test Video 2	53.8887	51.4956	54.1171
Test Video 3	53.9193	53.6994	54.0628
Test Video 4	51.8861	51.2743	52.5427
Test Video 5	50.8698	50.0266	51.5588
Test Video 6	53.8909	52.9037	54.5608

The MSE values of the watermark are shown in Table 3. These MSE values show the average term difference between the original video and the output, watermarked video. The lower the MSE values show the lower the level of degradation. The graphical representations of the same are shown in figure 8.

Table 3: *Values of MSE of the watermark embedded in various host videos*

Test Data	DCT+SVD	DCT+DWT+ SVD	DWT+DCT+SVD
Test Video 1	3.2097287e-06	6.3554344e-06	3.0818992e-06
Test Video 2	3.2443077e-06	4.086472e-06	3.0875129e-06
Test Video 3	3.2215482e-06	4.266353e-06	3.0920397e-06
Test Video 4	7.2462806e-06	7.457069e-06	7.0825214e-06
Test Video 5	7.2585012e-06	9.938821e-06	6.984163e-06
Test Video 6	4.2426826e-06	5.1242314e-06	3.498743e-06

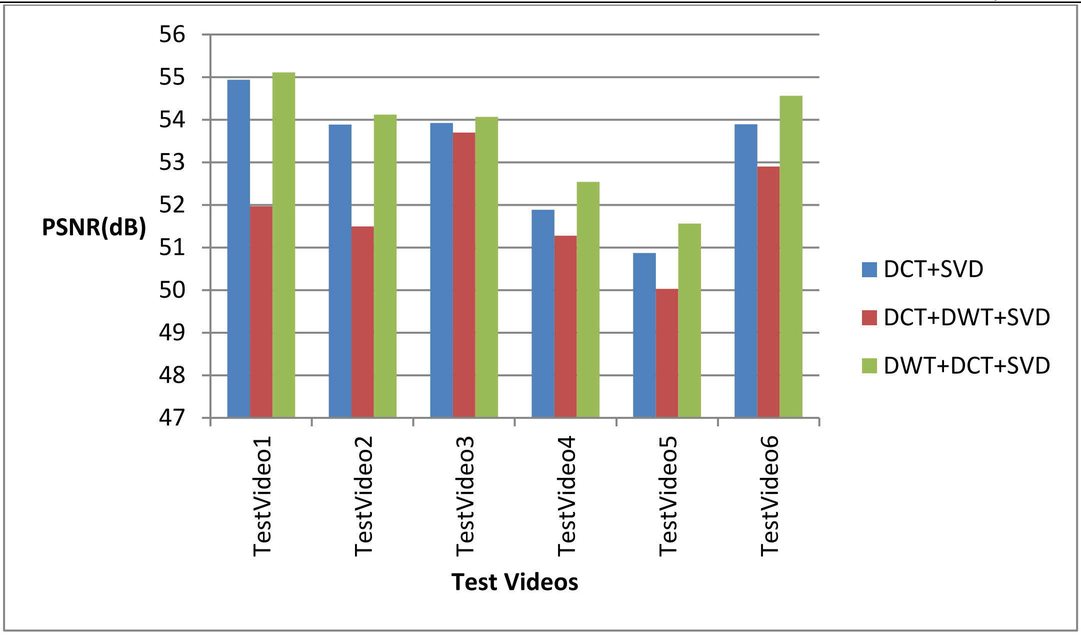


Figure 7: *Comparision between various combinations for imperceptibility*

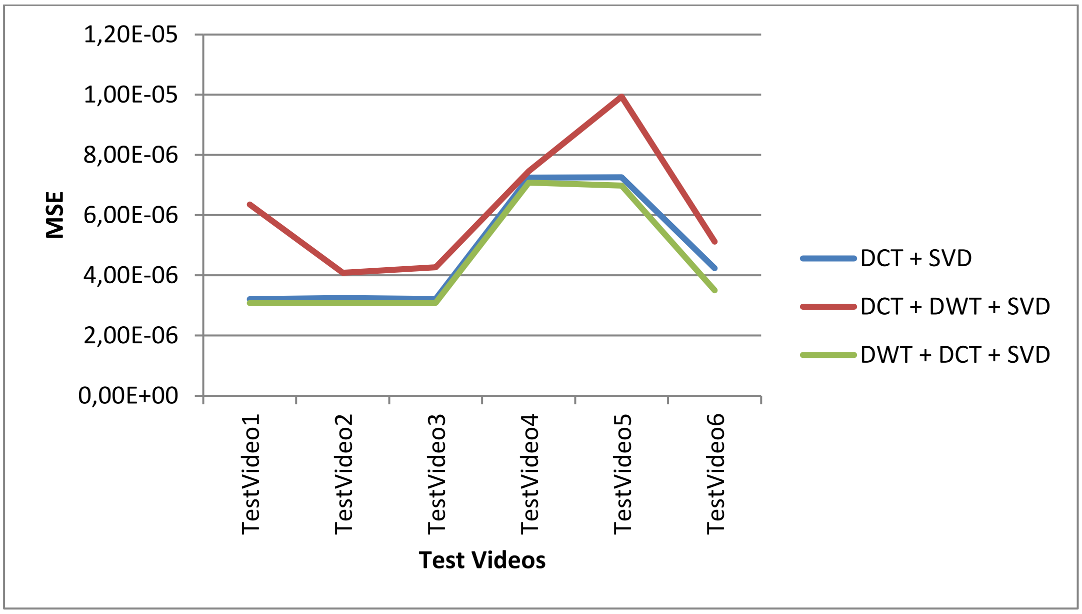


Figure 8: *Comparision between various combinations for robustness*

The current state of the art as discussed in the literature review section mainly focuses on the few numbers of bits or blocks within an image or a few numbers of images which make them more prone to various attacks such as geometric attacks, active attacks, compression attack, JPEG attack and many more. While the proposed approach is designed for watermarking the entire video with a special focus on keyframes thereby safeguarding the system from various attacks.

7. Conclusion and future work

The proposed DWT+DCT+SVD digital video watermarking technique outperforms the state-of-the-art methods while being more robust and safe. The DWT takes care of the finer variations of the images at various scales and it also helps in the localization of information. The DCT plays an important role in capturing a finite sequence of data points and the SVD is pivotal for preserving important geometrical insights. In the pursuit of the added advantage of DWT, DCT, and SVD combination the computational time of the proposed approach can be considered as a small trade-off against a significant improvement in watermarking performance. As a part of future work, the same approach may be deployed and analyzed for analog systems such as audio signals. Also, the experimental analysis to discretize the scale parameter will be an interesting study to pursue.

References

[1] M. Asikuzzaman and M. R. Pickering, "An Overview of Digital Video Watermarking," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 9, pp. 2131–2153, 2018.

[2] G. Kaur, S. S. Kasana, and M. K. Sharma, "An efficient watermarking scheme for enhanced high-efficiency video coding/h.265," *Multimed. Tools Appl.*, vol. 78, no. 9, pp. 12537–12559, 2019.

[3] M. R. Keyvanpour, N. Khanbani, and M. Boreiry, "A secure method in digital video watermarking with transform domain algorithms," *Multimed. Tools Appl.*, vol. 80, no. 13, pp. 20449–20476, 2021.

[4] G. Doërr and J. L. Dugelay, "A guide tour of video watermarking," *Signal Process. Image Commun.*, vol. 18, no. 4, pp. 263–282, 2003.

[5] D. Rajani and P. R. Kumar, "An optimized blind watermarking scheme based on principal component analysis in redundant discrete wavelet domain," *Signal Processing*, vol. 172, p. 107556, 2020.

[6] P. Garg and R. Rama Kishore, "Secured and multi optimized image watermarking using SVD and entropy and prearranged embedding locations in transform domain," *J. Discret. Math. Sci. Cryptogr.*, vol. 23, no. 1, pp. 73–82, 2020.

[7] A. K. Abdulrahman and S. Ozturk, "A novel hybrid DCT and DWT based robust watermarking algorithm for color images," *Multimed. Tools Appl.*, vol. 78, no. 12, pp. 17027–17049, 2019.

[8] S. P. Ambadekar, J. Jain, and J. Khanapuri, *Digital image watermarking through encryption and DWT for copyright protection*, vol. 727. Springer Singapore, 2019.

[9] D. G. Savakar and A. Ghuli, "Robust Invisible Digital Image Watermarking Using Hybrid Scheme," *Arab. J. Sci. Eng.*, vol. 44, no. 4, pp. 3995–4008, 2019.

[10] W. Zheng *et al.*, "Robust and high capacity watermarking for image based on DWT-SVD and CNN," *Proc. 13th IEEE Conf. Ind. Electron. Appl. ICIEA 2018*, pp. 1233–1237, 2018.

[11] F. Arab, M. Zamani, S. Poger, C. Manigault, and S. Yu, "A framework to evaluate the performance of video watermarking techniques," *2019 IEEE 2nd Int. Conf. Inf. Comput. Technol. ICICT 2019*, pp. 114–117, 2019.

[12] S. Das, M. Banerjee, and A. Chaudhuri, "An improved video key-frame extraction algorithm leads to video watermarking," *Int. J. Inf. Technol.*, vol. 10, no. 1, pp. 21–34, 2018.

[13] A. M. U. Wagdarikar, R. K. Senapati, and S. Ekkeli, *A secure video watermarking approach using CRT theorem in DCT domain*, vol. 521. Springer Singapore, 2019.

[14] M. I. Khan, M. Rahman, and I. H. Sarker, "Digital Watermarking for Image Authentication Based on Combined DCT , DWT and SVD Transformation."

[15] P. Garg and R. R. Kishore, "Performance comparison of various watermarking techniques," *Multimed. Tools Appl.*, 2020.

[16] H. Fang, H. Zhou, Z. Ma, W. Zhang, and N. Yu, "A robust image watermarking scheme in DCT domain based on adaptive texture direction quantization," *Multimed. Tools Appl.*, vol. 78, no. 7, pp. 8075–8089, 2019.

[17] G. Singh and N. Goel, "Entropy based image watermarking using discrete wavelet transform and singular value decomposition," *Proc. 10th INDIACom; 2016 3rd Int. Conf. Comput. Sustain. Glob. Dev. INDIACom 2016*, vol. 59, no. 11, pp. 2700–2704, 2016.