Received: November 25, 2023. Revised: January 19, 2024.

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Invisible Video Watermarking Based on Frame Selection and LSB Embedding

Inas Jawad Kadhim[1]* **Prashan Premaratne[2]** **Zena Ahmed Alwan[1]**

[1]*Electrical Engineering Technical College, Middle Technical University, Baghdad, Iraq*
[2]*School of Electrical and Computer and Telecommunications Engineering,*
University of Wollongong, North Wollongong, NSW, 2522, Australia
* Corresponding author's Email: inasjk@mtu.edu.iq

Abstract: Watermarking digital media content could undoubtedly discourage and prevent such illegal sharing activities, safeguarding the copyrights of the original creators. Such content could make use of invisible watermarking, not just for robustness and security, but also for imperceptibility. The proposed research presents a frame selection strategy based on the difference between the absolute histograms of consecutive frames. To increase security and robustness and avoid embedding the watermark frame by frame, ten frames with the maximum absolute histogram value are selected for watermark embedding. The size of the watermark would determine the number of non-overlapping blocks that a key frame is divided into in order to embed the watermark. The watermarks are embedded into the LSB of the frame-blocks, which have the most matched location where the embedding error will be minimum with high imperceptibility. Security can be further enhanced by generating a secret key based on non-sequential video frames, block location, and watermark size. The proposed method surpasses other comparable methods in terms of imperceptibility and robustness to image processing and geometrical attacks, and it is also computationally inexpensive. The results of the proposed method outperformed the best state-of-the-art methods in MPSNR and MNCC by 16 dB and 0.018, respectively.

Keywords: Video watermarking, Least significant bit, Absolute histogram difference, Frame selection.

1. Introduction

Information technology has brought about massive improvements in every aspect of life, but it has also led to significant issues of copyright infringement, including unauthorized copying, distribution, and sales of digital content.

Copyright protection is becoming increasingly significant as a means of enhancing the efficient use of network information [1]. To safeguard the image, video, and audio data, watermarking is a popular information embedding technique. Video watermarking is the type of watermarking technique that is most frequently utilized, due to the increasing frequency of copyright infringement and abuse in video media content [2]. The goal of utilizing a video watermark is to prevent copyright infringement of the real owner of the video.

Basic watermark systems consisting of encoders and decoders face several significant challenges, such as imperceptibility, security, robustness, and blind detection capabilities [2, 3]. The procedure of embedding a watermark in a video can negatively affect its viewability, as additional data is added to the original content. Thus, preserving the perceived quality of the video after watermarking is an important part to consider [4].

Watermarks if not visible to the human eye are said to be imperceptible and difficult to detect using detection tools [2, 4, 5]. The insertion of watermarks should not have a negative impact on the quality of a video. Robustness is defined as the ability of a watermark to withstand attacks, such as common signal processing attacks or malicious activities, without the embedded watermark being removed. The most important assurance that a digital item is safe from hackers is that security-integrated data cannot be altered [6, 7]. The capacity of watermarking is another factor considered when evaluating it. Capacity is less important for watermarking videos due to the sheer number of frames present in a video clip.

Watermarking is used in steganographic techniques for information embedding. Data is concealed via steganography in both science and the arts. There are two techniques for embedding images or videos in steganography: spatial and transform domain embedding. Data are directly embedded in the least significant bits (LSBs) of the spatial domain in the pixels of the video frame as watermarks. On the other hand, transform domain techniques (such as Fourier, discrete cosine, or wavelet transforms) modify the frequency coefficients of the video frames and encode the information into them. An example of spatial domain watermarking is the LSB technique, which inserts the watermark into the least significant bits of the original video. The watermark can be located anywhere on the video frame or only at a specific location.

The proposed method objective is to achieve the highest imperceptibility and security of the video watermarked while maintaining a good embedding capability and robustness. This paper presents a simple and efficient algorithm for watermarking videos invisibly that makes use of the LSB principle and optimum block selection.

The main contributions of this work include: (1) The proposed invisible video watermarking method satisfies the different positions of watermarks and achieves a good performance trade-off between imperceptibility and security; (2) frame selection strategy based on the absolute histogram difference between neighbouring frames is proposed to avoid embedding the watermark frame by frame and to solve the contradiction between the invisibility and security, (3) The use of a frame-watermark matching technique for determining the location to embed the watermark helps to reduce the error during the embedding process and improves the imperceptibility, (4) addition of further level of security to the methodology via secret key based non-sequencing video frames, blocks location, and watermark size. (5) Research evidence shows that the suggested approach surpasses current leading methods.

The rest of this work is structured into four sections: the first deals with a literature review, the second outlines the proposed approach, the third analyses the results and provides discussion, and the fourth summarizes the work.

2. Literature review

As explained in the introduction, information technology has sparked a significant transformation in the area of digital communication. In addition to its numerous advantages, it makes it easier to distribute, use, and replicate digital material, endangering the security of digital media ownership. Digital watermarking is a technique created specifically to address these problems. Watermarking a video that serves as a media cover is a difficult and complex process.

In the past decade, along with the development of an overwhelming amount of media content, many different watermarking methods have been created and documented in academic research. Depending on the embedding domains, these techniques are primarily categorized in the spatial and transform domains [2, 8].

Applications that previously only supported grayscale videos have now been expanded to include colour content. Inserting a Watermark into colour media can be accomplished by inserting it in the RGB colour space or by separately embedding a watermark into each colour channel [3, 9, 10] or just one particular channel, such as blue [11, 12]. With the exception of the blue channel, the HVS is less responsive. On the other hand, the majority of methods [13] use the Y channel of the YUV colour space as the watermark embedding channel. The watermark may be embedded directly in a frame after selecting an embedding channel, or it may first be translated into another domain.

In the spatial domain, there are four different types of watermarking systems [14]: LSB-based [5, 15], block-based [16], statistical [17, 18] and feature point-based [19]. In LSB-based, the watermark is embedded by changing the LSB of each pixel in the host image or video. Watermarking in the spatial domain is dominated by the LSB method as the most basic technique. Modifying the LSBs has no impact on perceptual quality since they contain less pertinent information [15]. Block-based systems partition the host image into various blocks before embedding the watermark in each of them. After that, the watermark is used to modify each block's pixel's intensity. These techniques are straightforward and effectively utilize computation. In the feature point-based approach, the watermark is embedded by changing the image's invariant characteristics at the encoder. The decoder then verifies whether the embedded watermark is present. These methods are relatively straightforward and computationally effective because the watermark is simply placed in the frame without any kind of alteration. However, they are not very strong and have a low information-hiding capacity [20].

In the system-based transform domain watermarking approach, a host frame in a video sequence is transformed into a new domain before the watermark is embedded [20]. The most widely used transforms include fast Fourier transform (FFT) [21], discrete cosine transform (DCT) [22], discrete wavelet transform (DWT), dual-tree complex wavelet transform (DTCWT) and singular value decomposition (SVD) [23, 24]. Each of these transformations, which are used to simulate video frames in multiple ways, have unique characteristics. To achieve better performance, numerous algorithms can be used for choosing embedded regions, frame selection, and interesting regions, as described in previous research such as [4, 6, 7, 25-29]

In [25], the authors proposed a robust video watermarking technique based on the spatial domain. Luminance adaptation and an edge mask were used to choose the suitable embedding region and adjust the watermark intensity.

In [6], the authors introduced a frame selection technique that takes into account the statistical relationship between frames in video, block size, and the capacity for inserting the watermark. The lifting wavelet transform (LWT) determines the process of selecting frames and the watermark is then inserted into the selected keyframes based on the quantization of the difference in coefficients between the sub-bands of the LWT. The proposed Chronological-MS method is used by [7] to identify the interesting regions for better embedding. The secret message was extracted and embedded into the video frames using a wavelet transform. Here, the binary image for embedding and extraction is created using the bit plane approach. A robust video watermarking-based nonblind colour method was presented by [26]. The method started by separating the frames into non-moving and moving components. Then, each colour channel's non-moving components were treated separately using a block-oriented watermarking scheme. A different procedure for embedding a watermark in an image was applied to blocks with lower entropy, and the result was the creation of the watermarked frame. The ideal frame prediction utilizing the deep belief network framework was provided by [27]. First, random frames from each video were sent to the genetic algorithm model as input, which then selected the frames in the best way possible so that the PSNR would be high. Each frame was given a label of one or zero, with one designating the frame with a higher PSNR and zero designating the frame with a lower PSNR. The authors in [28] proposed frame selection based on the quantity of scene changes made in the video. For effective watermark embedding, a blend of graph-based transform, singular valued decomposition, and hyperchaotic encryption is used. The proposed method was discovered to be resistant to a variety of attacks. In [29], embedding is carried out on a chosen frame. To strengthen the security of the watermark, frames are chosen at random. A linear congruential generator processes two integer keys to determine which frames to use. The chosen frame is then changed from RGB to YCbCr in terms of colour space. Based on the block, the Y channel was chosen to undergo Tchebichef transformation, and the coefficient for each block of the transformed results was chosen and compiled into a matrix. A singular matrix is then chosen for the watermark embedding after this matrix has undergone SVD transformation. The study [4] uses Tchebichef transformation and singular value decomposition to insert a watermark on a chosen frame with the most edge where a two-step Arnold transformation was evaluated to distribute the edge region. Watermarked video imperceptibility is improved by 1dB PSNR average compared to previous methods.

Despite the significant progress made in video watermarking research, several limitations remain. Many existing methods are computationally expensive, especially those that operate in the transform domain. This can make them impractical for real-time applications. Additionally, some methods are not robust to common video processing operations, such as compression, noise addition, and rotation. This can make it difficult to extract the watermark from the video after it has been processed. Additionally, it has been explained how well various suggested strategies may operate with colour images. Watermark robustness can be increased by converting the colour space from RGB to YCbCr and embedding it on the Y channel [13, 24, 29]. Nonetheless boosting robustness runs the risk of lowering imperceptibility. Finally, many methods do not achieve a good balance between robustness, imperceptibility, and security. These limitations and drawbacks highlight the need for further research on video watermarking algorithms that are computationally efficient, robust to a wide range of attacks, imperceptible to human viewers, secure and capable of embedding a significant amount of data.

3. Proposed algorithm

Digital watermarking, or simply watermarking, refers to the hiding of information within multimedia data in host data. In recent years, several types of host data have been studied for information hiding. Our

Received: November 25, 2023. Revised: January 19, 2024.

proposed strategy is based on a frame selection approach that utilizes the absolute difference in histograms between adjacent frames. In order to avoid embedding the watermark frame by frame and to increase security and robustness, ten frames are selected based on the maximum absolute histogram to include the watermark. The key frames get divided into non-overlapping blocks based on the size of the watermark. The watermarks are embedded into the LSB of the frame-blocks, with a more consistent position where the embedding error will be minimum. Detailed descriptions are provided in Sections 3.1, 3.2, and 3.3, and the embedding procedure is described in detail in 3.3, while the watermark extraction procedure is outlined in Section 3.4.

3.1 Frames selection key

The proposed method for identifying appropriate frames for watermarking is crucial to improve the performance and security of the video. Selecting a certain number of frames makes it more difficult for the attacker to detect and remove the watermark. This method protects the copyright of the video and helps reduce the file size of the video.

This technique takes advantage of the high sensitivity of the human visual system to changes in video content. The system aims to place watermarks in areas where they can difficulty be detected by viewers. This is generated by selecting frames with high absolute histogram difference (AH). The goal is to strike a balance between watermark invisibility and robustness, as frames with noticeable changes are more forgiving in terms of introducing slight modifications for watermark embedding.

Overall, this technique aligns with the principles of invisible video watermarking, where the goal is to embed information while minimizing visual degradation and maximizing resistance to various video processing operations.

$$\text{AH} = \sum_{K=1}^{L} \left| \left(H_{\boldsymbol{k}}(\boldsymbol{j}) - H_{\boldsymbol{k+1}}(\boldsymbol{j}) \right) \right| \quad (1)$$

where $H_k(i,j)$ and $H_{k+1}(i,j)$ represent the absolute histogram value of the k^{th} frame image at level j. L denotes the total number of levels for the absolute histogram. After that, the next step compares these values in order to choose 10 maximum absolute histogram value frames, then these frames will be the key frames of the video. This can be used, for example, to assess the degree of change or motion in consecutive frames of a video, where the histogram represents the pixel value distribution within each frame.

Fig. 1 shows the different motion energy of some frames in xylophone, superfalcon, situp and Foreman. In the case of the Forman video, the 10 highest motion frames are (184, 183, 185, 182, 181, 186, 180, 179, 187 and 188).

3.2 Preprocessing of video frame

The video frames are expected to be in RGB image format. Convert the selected RGB frames to grayscale frames by using the luminosity method.

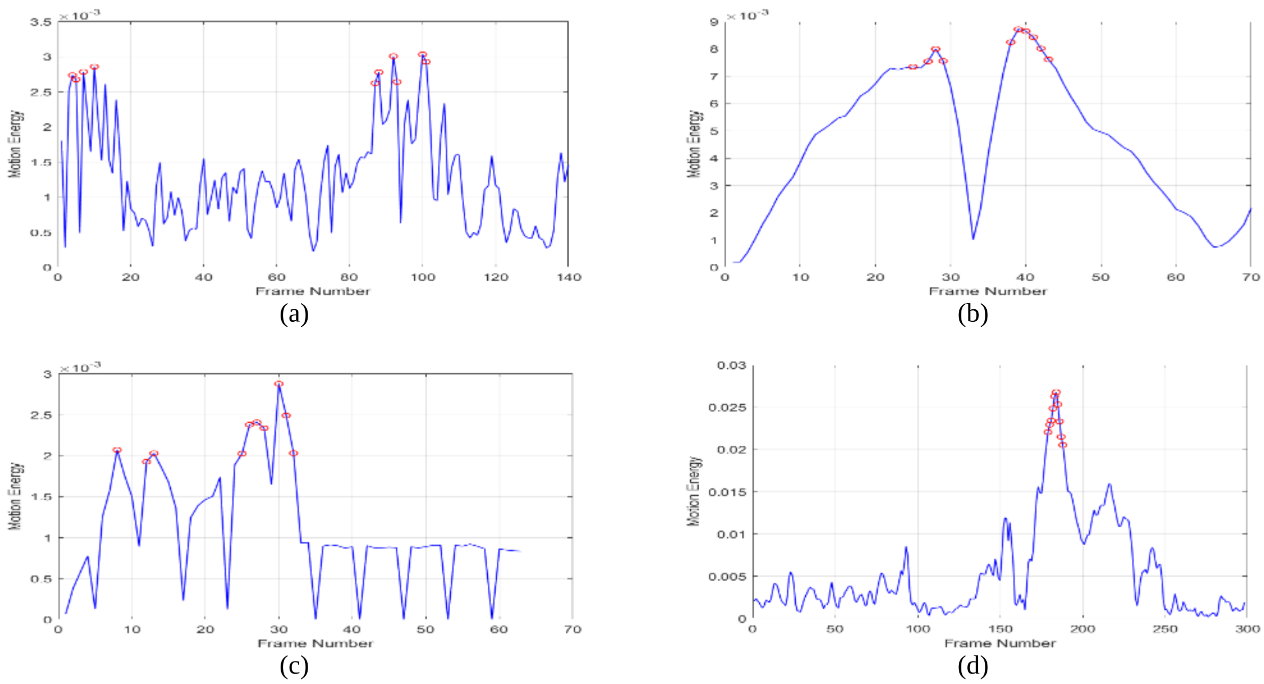

Figure. 1 Distribution of motion energy in: (a) Xylophone, (b) Superfalcon, (c) Situp, and (d) Foreman

 DOI: 10.22266/ijies2024.0630.04

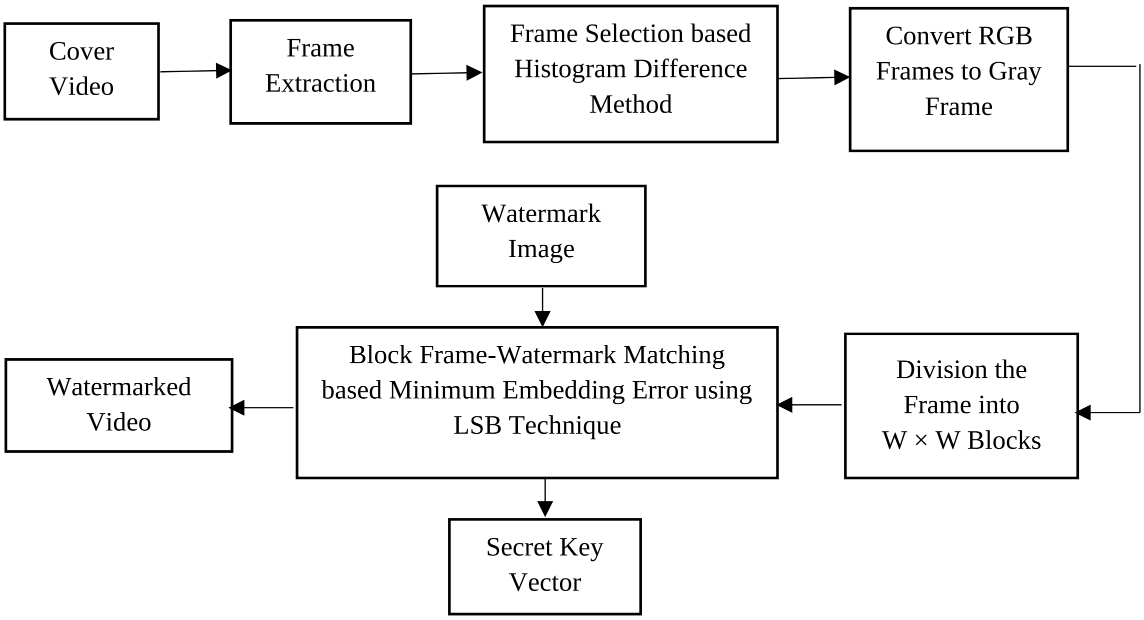

Figure. 2 The general framework of the proposed video watermark embedding procedure

Grayscale frames contain only luminance information and are often used in image and video processing tasks.

3.3 Block division algorithm

The next step after the selection of the best grayscale frames is to select the best block for embedding the watermark image. Dividing each frame F into the non-overlapping blocks of size W × W pixels, where W is the size of the watermarking image and assuming F_R , F_C are to be the number of rows and number of columns to that frame, the total number of blocks found in N number of blocks will be:

$$N = \lfloor F_R / \mathrm{W} \rfloor \times \lfloor F_C / \mathrm{W} \rfloor \quad (2)$$

3.4 Watermark embedding process

The proposed watermarking method involves embedding the watermark image into selected host frames by template matching over each block and embedding the watermark at the position with the minimum embedding error block.

$$E(W, FB) = \sum_{r=o,c=0}^{m=1,n=1} |W(r,c) - FB(r,c)| \quad (3)$$

A minimum error technique is used to choose the embedding position.

$$(r', c') = argmin_{r,c} \, | \, W(r,c) - FB(r,c)| \quad (4)$$

Where $W(r,c)$ is the original watermark, $FB(r,c)$ is the (r,c) th block in the selected frame, *m* and *n* are the dimensions of the block frame, where (r', c') denotes the optimum value of $FB(r,c)$ where watermark image has the best match with block of the frame.

Here, the LSB technique is used for watermark embedding, it is a very popular algorithm used in spatial domain watermarking. The watermark is inserted into the LSBs of the host frame, and the same process is used to extract the watermark. One main advantage of this technique is that it is straightforward to implement while offering high visual transparency with a minimal effect on the quality of the original image.

$$FB'(r,c) = FB(r,c) \oplus W(r,c) \quad (5)$$

where $FB'(r,c)$ is the watermarked frame.
The detailed steps of the proposed video watermarking process are illustrated in Fig. 2. The watermarked locations are then stored as a secret key vector, which contains information about non-sequential frames, block locations, and watermark size. Each watermark placed in the selected frame block will be assigned the secret key is updated. This secret key vector is used to retrieve the client-side watermark and verify the video's authenticity. This method increases the security of the watermarking process and ensures that the copyright of the video is protected.

Repeat through all selected frames, blocks, and watermarks using the watermarking process and evaluate the quality metrics. Finally, combine all the watermarked frames to create a watermarked video. This video now has a watermark embedded in it, and the secret key vector is preserved for future reference.

(a) (b) (c)

Figure. 3: (a)Original video frame, (b)Watermark image and (c)Watermarked frame

Table 1. Tested video

	Video Name	No. of Frame	Duration	Frame Rate
1	xylophone.mp4	141	4.7000	30
2	superfalcon.mp4	64	2.1340	30
3	situp.mp4	71	2.4150	29.9700
4	Foreman.mp4	299	9.9433	30.0707

3.5 Extraction algorithm module

In this process, the same techniques used in the embedding process, such as the LSB method, can be used in reverse to extract the verify the authenticity of the video. The secret key vector created during the embedding process is then used to properly align and extract the watermark image from the host frames.
The steps for extracting the watermark are given below:

1. The watermarked video is divided into individual frames, which are the main elements of the video.
2. Each RGB frame is converted to a grayscale frame. This conversion simplifies the watermark extraction process because it deals with luminance information only.
3. Detect key frames using the secret key vector.
4. Extract the location (x, y) and size of blocks (W×W) from the secret key vector. The secret key vector provides information about the location and size of the blocks within each key frame where the watermark was embedded. This information helps in accurately locating the watermark within the frames.
5. Divide the selected frames into non-overlapping blocks (W×W). Using the provided block size *(W×W)*, the chosen frames are divided into non-overlapping blocks. These blocks correspond to the regions where the watermark was embedded.
6. Find the location of watermark: the Eq. (6) is used to determine the location of the watermark within the selected frame. This equation likely defines the upper-left and lower-right corners of the watermark region.

$$(x_1, y_1) = (x + W, y + W) \tag{6}$$

7. Extract watermark bits from LSB of blocks: In the specified block defined by the upper-left *(x, y)* and lower-right (x_1, y_1) corners, the watermark bits are extracted from the LSB of the pixel values. This is where the watermark information was embedded during the watermarking process.
8. Display watermark image: The extracted watermark bits are reconstructed into the watermark image. This image represents the watermark that was embedded in the original video. Displaying the watermark image provides a visual representation of the extracted watermark.

4. Results and discussion

The proposed work has been implemented using MATLAB R2022a as the platform and has been evaluated on four standard videos. The dataset utilized is free to download at https://media.xiph.org/video/derf. The videos that were used were: xylophone, superfalcon, situp, and foreman. The evaluation was done in terms of imperceptibility and robustness. The video frame size of the xylophone is 464KB, superfalcon is 2.01 MB and situp is 1.56 MB and the number of frame and frame rates is shown in Table 1. The watermark used in the experiments was a binary image logo of 512 × 512. Fig. 3 shows a sample video frame, a watermarked image, and a corresponding watermarked frame.

4.1 Imperceptibility evaluation

The imperceptibility of the watermarked video frames is evaluated to assess the quality of the watermarked frames. The imperceptibility is measured by comparing the selected cover video frames (FS) with the watermarked video frames (FW). Three evaluation metrics could be utilized to assess the effectiveness of the proposed method: mean squared error (MSE), mean peak signal-to-noise ratio (MPSNR), and mean structural similarity index measure (MSSIM) which are defined by Eqs. (8), (10) and (12), respectively and are given below with their respective expressions.

The MSE measures the degree of similarity between the original video frame and the watermarked frame by computing the average energy of the error signal, which is the difference between the watermarked frame and the original frame. The MSE is given by the equation:

$$MSE = \sum_{I=0}^{F_R-1}\sum_{J=0}^{F_C-1}\frac{1}{F_R \times F_C}([FS(i,j) - FW(i,j)])^2 \quad (7)$$

$$MMSE = \frac{1}{F_t}\sum_{i=1}^{F_t} MSE_i \quad (8)$$

Where FS (i, j) stands for the pixel values of a selected frame, while FW (i, j) signifies the pixel values of the corresponding watermarked frame. i and j denote the location of the pixels in the frame $F_R \times F_C$, where F_t denotes the total number of frames in the video.

PSNR plays an important role in distinguishing between original and watermarked frames through mean squared error. average PSNR values are obtained by summing the PSNR values of all frames selected and then dividing by the total number of frames. A higher PSNR values indicate the efficiency of the method. The equation for the MPSNR calculation is given:

$$PSNR = \frac{10\, log_{10}\,(255)^2}{MSE} \quad (9)$$

$$MPSNR = \frac{1}{F_t}\sum_{i=1}^{F_t} PSNR_i \quad (10)$$

SSIM is an important metric for dissimilarity detection by comparing the structure and perceived quality of two images or video frames The average MSSIM is calculated as follows:

$$SSIM = \frac{(2\mu_{FS}\mu_{FW} + a_1)(2\sigma_{FSFW} + a_2)}{(\mu_{FS}^2 + \mu_{FW}^2 + a_1)(\sigma_{FS}^2 + \sigma_{FW}^2 + a_2)}$$

$$a_1 = (b_1 L)^2$$

$$a_2 = (b_2 L)^2 \quad (11)$$

$$MSSIM = \frac{1}{F_t}\sum_{i=1}^{F_t} SSIM_i \quad (12)$$

where μ_{FS} and μ_{FW} are the mean intensity values of frames FS and FW respectively. σ_{FS}^2 and σ_{FW}^2 are the variance of frames FS and FW respectively. σ_{FSFW} is the covariance of FS and FW. a_1 and a_2 are the two stabilizing parameters, L is the dynamic range of pixel values ($2^{\#bits\ per\ pixel}$- 1). b_1 and b_2 are also stabilizing parameters, with the commonly used values being $b_1 = 0.01$ and b_2 = 0.03.

Table 2. Perceptual quality measure for all the frames xylophone video

Watermarked Frame	Frame 1	Frame 2	Frame 3	Frame 4	Frame 5	Frame 6	Frame 7	Frame 8	Frame 9	Frame 10
Random Frame Number	100	92	101	10	7	88	4	5	93	87
MSE	0.0207	0.0204	0.0206	0.0211	0.0206	0.0211	0.0205	0.0205	0.0211	0.02112
PSNR	68.833	68.913	68.855	69.149	68.861	69.108	69.889	69.875	69.044	69.7193
SSIM	0.9891	0.9992	0.9833	0.9802	0.9897	0.9799	0.9884	0.9889	0.9890	0.9995

Table 3. Perceptual quality measure for all cover videos

Cover Video	MMSE	MPSNR	MSSIM
Xylophone	0.02077	69.2247	0.98872
Superfalcon	0.02124	68.3721	0.98869
Situp	0.02041	69.3894	0.99895
Foreman	0.02011	69.4753	0.99972
Average	0.02063	69.1154	0.99402

Table 4. Robustness measure for all the frames xylophone video under no attacks

Watermarked Frame	Frame 1	Frame 2	Frame 3	Frame 4	Frame 5	Frame 6	Frame 7	Frame 8	Frame 9	Frame 10
Random Frame Number	100	92	101	10	7	88	4	5	93	87
BER	0.00149	0.00149	0.00149	0.00153	0.00149	0.00153	0.00149	0.00149	0.00151	0.001498
NC	0.99914	0.99983	0.99725	0.99821	0.99904	0.99912	0.99944	0.99985	0.99901	0.99976

Table 5. Robustness measure for all cover videos under no attacks

Cover Video	MBER	MNC
Xylophone	0.0015	0.99907
Superfalcon	0.0013	0.99917
Situp	0.0012	0.99920
Foreman	0.00132	0.99918
Average	0.00133	0.999155

Table 6. Robustness measure of the suggested method's resilience against different attacks

Attack Index	Video Attack	Xylophone	Superfalcon	Situp	Foreman
		MNC	MNC	MNC	MNC
MF	Median filtering (3 × 3)	0.9999	0.9998	0.9996	0.9997
WF	Wiener filtering (3 × 3)	0.9989	0.9988	0.9985	0.9976
GF	Gaussian filtering (3 × 3)	0.9987	0.9978	0.9977	0.9976
RO	Rotation (45°)	0.9945	0.9957	0.9975	0.9955
TR	Translation (30, 30)	0.9983	0.9989	0.9992	0.9989
HE	Histogram equalization	0.9979	0.9984	0.9985	0.9985
SH	Sharpening	0.9982	0.9980	0.9980	0.9994
GN	Gaussian noise (0, 10%)	0.9883	0.9881	0.9878	0.9878
SP	Salt and pepper noise (10%)	0.9875	0.9872	0.9869	0.9880

The analysis of the proposed method based on PSNR, MSE and SSIM values by varying the number of frames is depicted in Tables 2 and 3 shows the analysis of the proposed method based on MMSE, MPSNR and MSSIM corresponding to each video. From the analysis, it is evaluated that the proposed method acquires high imperceptibility with an average value of MMSE as 0.02063, MPSNR as 69.1154and MSSIM as 0.99402. The results indicate the possibility of using the proposed video watermark copyright protection.

4.2 Robustness evaluation

The robustness of a watermarking technique is determined by comparing the original watermark logo to the extracted watermark logo. The parameters used for evaluation are mean normalized correlation (MNC) and mean bit error rate (MBER). A high MNC value close to or equal to one, and a low MBER value indicate a successful and robust watermarking technique, as it shows that the watermark signal has been accurately extracted and has not been affected by any attacks. The formulas for calculating MNC and MBER are given:

$$NC = \frac{\sum_i \sum_j W(i,j) W'(i,j)}{\sum_i \sum_j [W(i,j)]^2} \tag{13}$$

$$MNC = \frac{1}{F_t} \sum_{i=1}^{F_t} NC_i \tag{14}$$

$$BER\ (in\ \%) = \frac{N_W - N_{W'}}{N_W} \times 100 \tag{15}$$

$$MBER = \frac{1}{F_t} \sum_{i=1}^{F_t} BER_i \tag{16}$$

where $W(i,j)$ represents the original watermark, while $W'(i,j)$ is extracted watermark. The number of watermark bits embedded is represented by N_W , while $N_{W'}$ represents the number of bits correctly retrieved.

Table 4 displays the results of watermarking the xylophone video by changing the number of frames without any external interference. Table 5 shows the analysis of the proposed method based on MBER and MNC when applied to each video without attacks. The results indicate that the proposed method has a low MBER and high MNC. The proposed method's ability to withstand various image and video attacks is also tested by evaluating the quality of the extracted watermark logo. Watermarked video has undergone nine different types of attacks to assess the proposed method's robustness. These attacks include: median filtering (MF) (3 × 3), wiener filtering (WF) (3 × 3), Gaussian filtering (GF) (3 × 3), rotation (RO) (45◦), translation (TR) (30, 30), histogram equalization (HE), sharpening (SH), Gaussian noise (GN) (0, 10%), salt and pepper noise (SP) (10%). Table 6 summarizes the MNC values for the four videos under consideration for each of the nine attacks. The table demonstrates that the proposed method achieves superior MNC values.

Received: November 25, 2023. Revised: January 19, 2024.

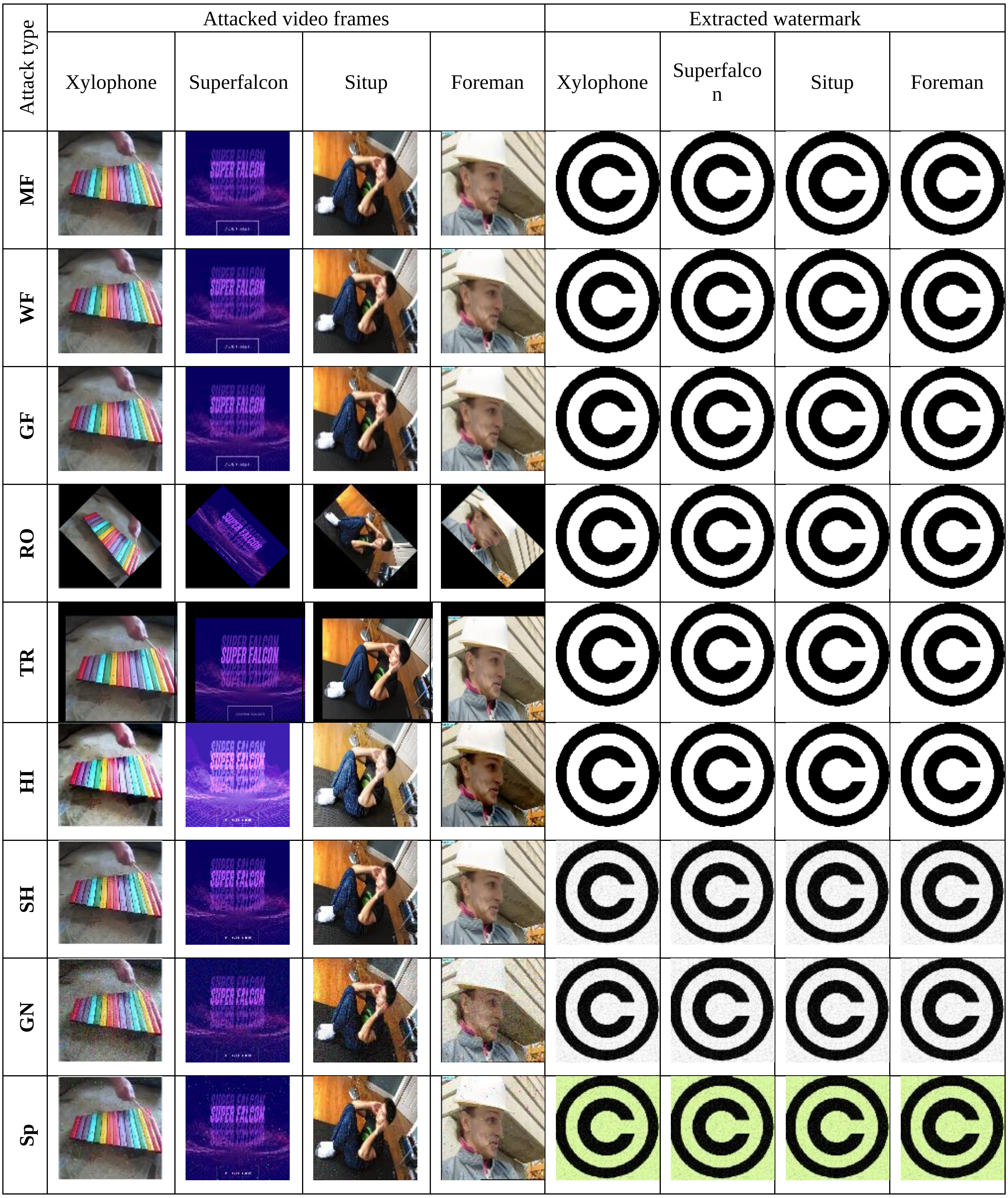

Figure. 4 The visual quality of watermarked video frames and the resulting extracted watermark logo image when subjected to different types of attack

The quality of the extracted watermark logos for each of the nine attacks for all videos is shown in Fig. 4. It can be seen from the figure that the proposed method consistently produces a high-quality watermark, even under attack.

4.3 Performance analysis against existing techniques

To further evaluate the imperceptibility and robustness of the proposed method, it has been compared to four recent video watermarking

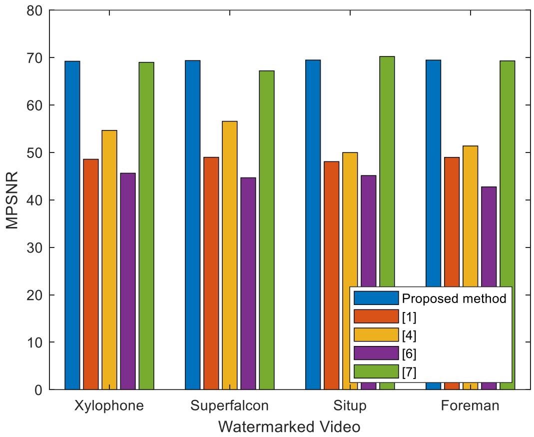

Figure. 5 Comparative analysis of the proposed method against considered techniques in terms of MPSNR under no attacks

techniques in terms of MPSNR and MNC both under attack and no attack conditions. The four state-of-the-art-based video watermarking techniques namely by [1, 4, 6, 7][1], [4], [6], and [7] are implemented from scratch as described in the papers using the same four videos. Fig. 5 shows MPSNR values, which are plotted and compared against other methods for all four videos without any attacks. When using ten frames, the MPSNR values of Xylophone video measured by the proposed method, [1], [4], [6], and [7] methods are 69.23 dB, 48.57 dB, 54.65 dB, 45.63 dB and 69.00 dB respectively. Likewise, when using Superfalcon video, the MPSNR values measured by the proposed method and existing methods [1], [4], [6], and [7] are 69.37 dB, 48.98 dB, 56.57 dB, 44.67 dB and 67.21 dB respectively. The MPSNR values of the Situp video for the proposed method and the existing methods were measured to be 69.48 dB, 48.07 dB, 49.98 dB, 45.13 dB, and 70.23 dB, respectively. Likewise, the MPSNR of the fourth video measured by the proposed method and above existing methods are 69.47 dB, 48.97 dB, 51.37 dB, 42.75 dB and 69.31 dB respectively. From the analysis of Fig. 5, it can be shown that the proposed method has a higher MPSNR value than the existing methods.

Table 7. Comparative analysis of the proposed method against considered techniques in terms of MNC under various attacks for Xylophone video

Video Attack	Proposed Method	[1]	[4]	[6]	[7]
MF	0.9999	0.9984	0.9663	0.9981	0.9477
WF	0.9989	0.9975	0.9681	0.9964	0.9620
GF	0.9987	0.9979	0.9841	0.9972	0.9554
RO	0.9945	**0.9970**	0.9852	0.9943	0.9379
TR	0.9983	0.9977	0.9859	0.9973	0.9832
HE	0.9979	0.9978	0.9709	0.9971	0.9733
SH	0.9982	0.9970	0.9721	0.9964	0.9329
GN	0.9883	0.9838	0.9672	0.9816	0.9571
SP	0.9875	0.9822	0.9569	0.9817	0.9403
Average	**0.9958**	0.9943	0.9729	0.9933	0.9544

Table 8. Comparative analysis of the proposed method against considered techniques in terms of MNC under various attacks for Superfalcon video

Video Attack	Proposed Method	[1]	[4]	[6]	[7]
MF	0.9998	0.9987	0.9650	0.9977	0.9470
WF	0.9988	0.9980	0.9663	0.9961	0.9609
GF	0.9978	**0.9981**	0.9841	0.9972	0.9541
RO	0.9957	**0.9975**	0.9819	0.9935	0.9369
TR	0.9989	0.9982	0.9844	0.9968	0.9810
HE	0.9984	0.9983	0.9671	0.9958	0.9722
SH	0.9980	0.9977	0.9688	0.9962	0.9312
GN	0.9881	0.9841	0.9664	0.9816	0.9549
SP	0.9872	0.9827	0.9552	0.9816	0.9391
Average	**0.9959**	0.9948	0.9710	0.9929	0.9530

Table 9. Comparative analysis of the proposed method against considered techniques in terms of MNC under various attacks for Situp video

Video Attack	Proposed Method	[1]	[4]	[6]	[7]
MF	0.9996	0.9985	0.9651	0.9981	0.9473
WF	0.9985	0.9981	0.9665	0.9964	0.9612
GF	0.9977	**0.9980**	0.9842	0.9973	0.9543
RO	0.9975	**0.9975**	0.9821	0.9945	0.9374
TR	0.9992	0.9985	0.9845	0.9973	0.9811
HE	0.9985	0.9982	0.9673	0.9960	0.9724
SH	0.9980	0.9975	0.9691	0.9962	0.9319
GN	0.9878	0.9839	0.9667	0.9814	0.9559
SP	0.9869	0.9820	0.9553	0.9817	0.9401
Average	**0.9960**	0.9947	0.9712	0.9932	0.9535

Table 10. Comparative analysis of the proposed method against considered techniques in terms of MNC under various attacks for Foreman video

Video Attack	Proposed Method	[1]	[4]	[6]	[7]
MF	0.9997	0.9986	0.9661	0.9980	0.9471
WF	0.9976	0.9980	0.9679	0.9962	0.9611
GF	0.9976	**0.9984**	0.9840	0.9971	0.9544
RO	0.9955	**0.9976**	0.9851	0.9941	0.9370
TR	0.9989	0.9986	0.9857	0.9971	0.9812
HE	0.9985	0.9980	0.9701	0.9963	0.9723
SH	0.9994	0.9974	0.9711	0.9960	0.9317
GN	0.9878	0.9843	0.9666	0.9813	0.9552
SP	0.9880	0.9821	0.9565	0.9807	0.9393
Average	**0.9959**	0.9948	0.9726	0.9929	0.9533

The proposed method has been designed to withstand various image processing attacks. Tables 7-10 compare the proposed method with four recent techniques found in literature in terms of MNC values under the considered attacks. The results indicate that the proposed method offers higher resistance to various attacks compared to state-of-the-art video watermarking techniques with high-density embedding. Although one existing method [1] performs better against Gaussian filtering and rotation attacks, the proposed method still outperforms it in overall analysis.

5. Conclusion

The main contribution of this research is to enhance the imperceptibility quality of watermarked videos and achieve a better trade-off between security and robustness performance. The proposed video watermarking technique is a robust method for embedding watermarks in selected frames based on the absolute difference in histograms between neighbouring frames. The watermarks are embedded into the least significant bits of the frame-blocks with the closest match. The block's location, non-sequence of frames, and watermark size are disguised from potential attackers by incorporating different secret keys. The proposed method is computationally efficient. Testing on benchmark videos for standard image processing operations and video attacks have highlighted the robustness and the security of the proposed method. The results of the proposed method outperformed the best state-of-the-art methods in terms of MPSNR and MNCC by 16 dB and 0.018 respectively. We envisage that the method could be improved by incorporating optimization algorithms that allow for the embedding of multiple watermarks, thus providing stronger protection against compression attacks.

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

The paper conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft preparation, and visualization, have been done by 1st and 3rd authors. 2nd author has done the review and editing.

References

[1] R. Singh, H. Mittal, and R. Pal, "Optimal keyframe selection-based lossless video-watermarking technique using IGSA in LWT domain for copyright protection", *Complex & Intelligent Systems*, Vol. 8, No. 2, pp. 1047-1070, 2022.

[2] I. Joseph and J. Mandala, "Comprehensive Review on Video Watermarking Security Threats, Challenges, and Its Applications", *ECS Transactions*, Vol. 107, No. 1, p. 13833, 2022.

[3] M. Patil, P. Bamane, and S. Hasarmani, "Efficient Watermarking in Color Video Using DWT-DFT and SVD Technique", In: *Proc. of Progress in Advanced Computing and Intelligent Engineering*, pp. 75-87, 2021.

[4] N. Rijati, PNADR Ignatius, and M. Setiadi, "Imperceptible Improvement using Edge Area Selection for Robust Video Watermarking Using Tchebichef-Singular Value Decomposition", *International Journal of Intelligent Engineering and Systems*, Vol. 15, No. 2, pp. 298-306, 2022, doi: 10.22266/ijies2022.0430.27.

[5] G. Gupta, V.K. Gupta, and M. Chandra, "Review on video watermarking techniques in spatial and transform domain", In: *Proc. of Advances in Intelligent Systems and Computing*, doi10.1007/978-81-322-2752-6_67, 2016.

[6] A. Bhardwaj, V.S. Verma, and R.K. Jha, "Robust video watermarking using significant frame selection based on coefficient difference of lifting wavelet transform", *Multimedia Tools and Applications*, Vol. 77, No. 15, pp. 19659-19678, 2018.

[7] A.M.U. Wagdarikar and RK Senapati, "Optimization based interesting region identification for video watermarking", *Journal of Information Security and Applications*, Vol. 49, p. 102393, 2019.

[8] B. Meryem and M. Samira, "A short survey on image zero-watermarking techniques based on visual cryptography", In: *Proc. of 2018 9th International Symposium on Signal, Image, Video and Communications*, pp. 157-162, 2018.

[9] H. Mirza, H. Thai, and Z. Nakao, "Digital video watermarking based on RGB color channels and principal component analysis", In: *Proc. of International conference on knowledge-based and intelligent information and engineering systems*, pp. 125-132, 2008.

[10] R. Alkanhel and H.A. Abdallah, "Securing color video when transmitting through communication channels using DT-CWT-Based watermarking", *Electronics*, Vol. 11, No. 12, p. 1849, 2022.

[11] J.W. Salih, S.H. Abid, and T.M. Hasan, "Imperceptible 3d video watermarking technique based on scene change detection", *International Journal of Advanced Science and Technology*, Vol. 82, pp. 11-22, 2015.

[12] A. Mishra and P. Chuttani, "Uncompressed Video Watermarking Using Motion Vectors and Back Propagation Network", *International Journal of Engineering Research & Technology*, Vol. 3, pp. 1023-1030, 2014.

[13] A. Karmakar, A. Phadikar, and A. Mukherjee, "Video watermarking scheme resistant to rotation and collusion attacks", In: *Proc. of Emerging Trends in Computing and Communication*, pp. 95-101, 2014.

[14] R. V Mahule and C.A. Dhawale, "Analysis of image security techniques using digital image watermarking in spatial domain", *International Journal of Computer Applications*, Vol. 975, p. 8887, 2015.

[15] V. Sharma, M. Gangarde, and S. Oza, "A Spatial Domain Based Secure and Robust Video Watermarking Technique Using Modified LSB and Secret Image Sharing", *ICTACT Journal on Image and Video Processing*, Vol. 10, No. 1, p. 2019. doi10.21917/ijivp.2019.0293.

[16] F. Arab and M. Zamani, "VW16F: a robust video watermarking using simulated block based spatial domain technique", In: *Proc. of Proceedings of the International Conference on Security and Management (SAM), The Steering Committee of The World Congress in Computer Science, Computer*, pp. 202-207, 2018.

[17] L. Li, X. Li, T. Qiao, X. Xu, S. Zhang, and C.-C. Chang, "A novel framework of robust video watermarking based on statistical model", In: *Proc. of International Conference on Cloud Computing and Security*, pp. 160-172, 2018.

[18] K. Su, D. Kundur, and D. Hatzinakos, "Statistical invisibility for collusion-resistant digital video watermarking", *IEEE Transactions on Multimedia*, Vol. 7, No. 1, pp. 43-51, 2005.

[19] Z. Lv, H. Guan, Y. Huang, S. Zhang, and Y. Zheng, "Enhancing Feature Point-Based Video

Watermarking against Geometric Attacks with Template", *Wireless Communications and Mobile Computing*, Vol. 2021, p. 2021.

[20] M. Asikuzzaman and M.R. Pickering, "An overview of digital video watermarking", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 28, No. 9, pp. 2131-2153, 2017.

[21] X. Jin and J. Kim, "Robust Digital Watermarking for High-definition Video using Steerable Pyramid Transform, Two Dimensional Fast Fourier Transform and Ensemble Position-based Error Correcting", *KSII Transactions on Internet and Information Systems (TIIS)*, Vol. 12, No. 7, pp. 3438-3454, 2018.

[22] M.G. Busiri and R. Munir, "Mobile application of video watermarking using discrete cosine transform on Android platform", In: *Proc. of 2017 5th International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering (ICICI-BME)*, pp. 207-211, 2017.

[23] A. Fadhilah, B. Hidayat, and R.D. Atmaja, "Analisis dan Simulasi Video Watermarking Menggunakan Metode Dual Tree Complex Wavelet Transform (DT-CWT) dan Singular Value Decomposition (SVD)", *Jurnal Rekayasa Elektrika*, Vol. 12, No. 2, pp. 41-47, 2016.

[24] R. Esfahani, M.A. Akhaee, and Z. Norouzi, "A fast video watermarking algorithm using dual tree complex wavelet transform", *Multimedia Tools and Applications*, Vol. 78, No. 12, pp. 16159-16175, 2019.

[25] G. Chen, C. Kang, D. shun Wang, X. Zhao, and Y. Huang, "A Robust Video Watermarking Algorithm Based on Spatial Domain", In: *Proc. of 2018 7th International Conference on Energy and Environmental Protection*, pp. 412-419. Atlantis Press, 2018.

[26] P. Rasti, S. Samiei, M. Agoyi, S. Escalera, and G. Anbarjafari, "Robust non-blind color video watermarking using QR decomposition and entropy analysis", *Journal of Visual Communication and Image Representation*, Vol. 38, pp. 838-847, 2016.

[27] S.S. Alotaibi, "Video Watermarking for Copyright Protection with Genetic-Based Frame Selection", *Journal of King Abdulaziz University*, Vol. 8, No. 2, pp. 33-45, 2019.

[28] C. Sharma, B. Amandeep, R. Sobti, T.K. Lohani, and M. Shabaz, "A secured frame selection based video watermarking technique to address quality loss of data: combining graph based transform, singular valued decomposition, and hyperchaotic encryption", *Security and Communication Networks*, Vol. 2021, p. 2021.

[29] A. Setyono and D. Setiadi, "Robust Video Watermarking using Tchebichef Transform and Singular Value Decomposition on the Selected Frame Based YCbCr Color Space", *International Journal of Intelligent Engineering and Systems*, Vol. 13, No. 6, pp. 432-441, 2020, doi: 10.22266/ijies2020.1231.38.