

# Key-frame extraction based video watermarking using speeded up robust features and discrete cosine transform

Kapre Bhagyashri S., Rajurkar Archana M.

Department of Computer Science and Engineering, MGM's College of Engineering, Nanded, India

## Article Info

### Article history:

Received Sep 28, 2022

Revised Nov 28, 2022

Accepted Jan 17, 2023

### Keywords:

Discrete cosine transform

speed-up robust features

Embedding

Entropy

Extraction

Pearson correlation coefficient

Watermark

## ABSTRACT

Due to advancements in the internet and multimedia technologies, unauthorised users can easily modify video content. As a result, video authentication has been established as a viable solution for ensuring multimedia security. We propose a key-frame based video watermarking scheme based on discrete cosine transform (DCT). First, the pearson correlation coefficient (PCC) is used to detect the shot boundaries of the input video. To reduce the difficulties created by traditional video watermarking systems; an entropy measure is employed to detect key-frames from input video. Traditional schemes entail embedding the entire watermark into all frames of the video, which is inefficient and time-consuming. To improve the security, robustness, and imperceptibility of the proposed video watermarking scheme, speeded up robust feature points are extracted from each key-frame of the shot and used as reference points for embedding and detection of watermark. The embedded watermark is extracted blindly without using the original data during the extraction process. The results of the experiments reveal that the proposed technique effectively detects shot boundaries under a variety of camera operations and outperforms in terms of imperceptibility and resilience.

*This is an open access article under the [CC BY-SA](#) license.*



## Corresponding Author:

Kapre Bhagyashri S.

Department of Computer Science and Engineering, MGM's College of Engineering

Nanded, India

Email: kapre\_bs@mgmce.ac.in

## 1. INTRODUCTION

Recent developments in multimedia technologies and the transmission of digital video over a network show how significant digital video is becoming to be as a broadcasting, communication, and entertainment medium. It gets more difficult to transmit and store raw video as storage capacity grows and video quality improves. Due to the development of video compression, almost all digital videos are now transferred via networks in a compressed format. Different codec, including H.261, H.262, H.263, and H.264, are available for the videos. These codec are used to encode raw videos. New generations of H.264/AVC video codec are widely used due to their great compression efficiency and strong network compatibility.

Initially, watermarks were incorporated into all frames of the video using the same image watermarking approach. However, such types of technique es require more time for embedding and extraction of watermarks, and they fail to address the challenges caused by the temporal dimension of video sequence. It has been observed that the video quality degrades if image watermarking techniques are directly applied to it. Therefore, the key-frame based approach comes into the picture. In this approach, key-frames are extracted from video for embedding and extraction of watermark. This approach requires less time for the watermarking process. This efficiently reduces the time complexity and helps to improve the visual quality of

watermarked video. The work presented in this paper is motivated by the requirement of a practical video watermarking scheme that provides authentication to H.264/AVC-based compressed video with high imperceptibility and robustness that protects copyright property efficiently. The proposed video watermarking scheme provides several advantages over existing techniques. In this paper, pearson correlation coefficient (PCC) based shot detection technique is introduced and key-frame from each shot is extracted using statistical measure called entropy.

The organization of paper is: section 2 reviews the existing work. In section 3, the proposed video watermarking scheme is presented. Section 4 presents experimental study on quality and efficiency of the proposed video watermarking scheme. Section 5 describes conclusions.

## 2. RELATED WORK

The various video watermarking schemes have been developed in the recent area, in which the main idea is to insert secret information (watermark) into the selected key-frames of the video by introducing some changes that are acceptable in terms of accuracy and usually invisible to the legitimate user. Video watermarking techniques are divided into two different schemes based on embedding domain like frequency domain (FD) and spatial domain (SD). In SD, the watermark is inserted into the intensity (pixel) values of the frame, whereas in the FD the watermark information is inserted into the frequency coefficients of the video frame instead of modifying direct pixel values. As a result this type of approach provides high robustness and less visible distortion than SD approach. Most research in the field of video watermarking focuses on the FD approach. However, it requires more computational complexity compared to SD approach.

Generally in image watermarking schemes, watermark information is inserted into the image itself either in blocks or region of interest, whereas, in video watermarking schemes watermark is inserted in different ways like: i) frame by frame [1], [2], ii) region based [3]–[5], iii) key-frame based [2], [6] approach. In the frame by frame approach, watermark information was inserted in every frame of the video. This type of approach is very effective and robust against frame attacks like frame dropping, inserting, and swapping. But it is impractical, requires more time for embedding and extraction of watermark information and also increases the size of video. To overcome these challenges, many researchers have used region-based approach for embedding and extraction of watermark. In this type of approach, robust regions or moving blocks are detected within a host frame for embedding watermark. This approach was revealed that these algorithms provide security, high imperceptibility and robustness against common attacks. The main disadvantage is that the watermark's accuracy is dependent on the locations of motion segments or regions retrieved during the extraction process. To improve the correctness and to reduce time complexity the third approach was introduced wherein, the representative frames are selected from each shot or scene of video sequence for embedding and extraction of watermark. This methodology reduces a huge amount of time required for watermarking process. Moreover, it avoids frame redundancy and improves the stability and robustness of the watermarking technique.

In [7], Li *et al.* have presented a semi-fragile video watermarking scheme for compressed domain videos. The numerical relationship among discrete cosine transform (DCT) non-zero coefficients was considered as an authentication code. In this scheme, first the frame number was converted into an 18-bit watermark sequence, and then the generated code was embedded into a 4×4 sub-block which contains at least three DCT non-zero coefficients. This watermarking technique shows good transparency and tamper detection. Himeur *et al.* have developed a chaotic encryption-based video watermarking scheme [8]. In which, the key-frames were extracted using the gradient magnitude similarity deviation (GMSD) technique for embedding and extraction of watermarks. A blind and secure watermark embedding and extraction technique was adopted using discrete wavelet transform (DWT) and singular value decomposition (SVD). The drawback of this system was that, if the watermarked key frame is lost from the video sequence, it fails to recover the watermark and provides weak resistance to geometric distortion. In [9], a semi-blind video watermarking scheme using speed-up robust features (SURF) and visual cryptography was introduced. A shot detection technique that uses the histogram difference of consecutive frames was introduced. This technique provides robustness against different signal processing and geometric distortion. However, this scheme needs to store generated shares securely. Sethuraman *et al.* [10] have introduced a key-frame based watermarking technique wherein the structural similarity index metric–absolute difference metric (SSIM-AMD) techniques were adopted for the identification of non-redundant frames. Then, the entropy–AMD method was used to select a key-frame. Furthermore, DWT is applied to decompose the key-frame into sub-bands. To avoid false-positive attacks, the principal component of the watermark image block was computed and embedded into the middle band of DWT. The strength of the watermark was decided by calculating the scaling factor using the ant colony optimization (ACO) technique. It was observed that this scheme was

robust against video processing and false-positive attacks. It provides high performance in terms of imperceptibility and robustness.

In [11], the authors designed a semi fragile video authentication technique using DWT and DCT transform wherein, the robust features were extracted and used to generate a content-based authentication code (CBAC). That code is scrambled using Arnold's transform to generate a quick response code. Thereafter, a quick response code is embedded into the middle frequency sub-band of DWT and extracted blindly without using original video information. This technique outperforms well in terms of watermarked video perceptual quality and discriminating between intentional and unintentional manipulations. To detect and localise tampered area locations in [12], the authors have developed a chromatic DCT-based video watermarking approach. In this scheme, tamper detection was done by using different features of the H.264/AVC coding standards. An experimental result shows that the developed technique was used to detect spatial attacks as well as help localise tampered regions.

A hyper-chaotic Lorentz based video watermarking approach has been developed in [13]. In which case, watermark embedding and detection were performed by extracting specific frames from the host video's non-motion frames. Then, watermark embedding was done using the 3D-DWT transform. The developed approach fails to resist temporal attacks. Therefore, this approach is to detect tampered areas but not be able to localise them. In [14], the authors have introduced a fragile video watermarking approach wherein, content-based authentication code is generated using the Arnold transforms. This approach was able to detect and localise tampered area locations. In [15], authors have developed a video authentication technique using audio and video features. The authentication code was generated by using both video and audio content from an MP4 clip. Then, the generated code was embedded in the subtitles. This approach allows frame addition and removal to be detected. One more approach was designed for authentication of the MP4 format in [16], wherein an encrypted hash value of audio data was inserted into the synchronisation content of the MP4 file. The designed approach was robust against compression and able to detect tampered area locations.

A new video authentication technique based on the generation of watermark images has been developed in [17]. All watermark images were embedded into all video frames using the DWT transform. During the extraction process, the embedded watermark was extracted and analysed to detect spatial attacks. Thereafter, a binary sequence was generated from all the extracted watermark images that is used to determine the type of temporal attack such as frame removal, addition, re-ordering, and localising tampered frames. In [18], authors have developed a semi-fragile video watermarking approach. This approach was used for tampered area detection and localization. The watermark embedding was done in the P and B frames of the video in low frequency components. The developed approach was robust and imperceptible, but requires original information during the extraction process. Aditya *et al.* [19] have designed a video watermarking scheme for tamper detection. In this approach, triple transformations like DWT, DCT, and SVD were used to improve the robustness of the designed approach. Both host and watermark videos were transformed by using the DWT and DCT successively. Then, SVD was employed on the original video, and in the same way, SVD was applied to the watermark to obtain singular values. The singular values of watermark were embedded into the singular values of host video with some embedding strength. However, the designed approach was robust, but there was a possibility of false detection.

The literature review on video watermarking systems revealed that, i) embedding watermarks in all frames of the video consumes more time; ii) the watermark block embedded in the key-frame can withstand frame-dropping attacks, iii) these strategies fail to achieve a good balance of resilience and imperceptibility.

In this paper, we have developed a video watermarking technique that provides the tread off between robustness and imperceptibility. Video shot boundaries are detected using PCC and the key-frames are extracted using a statistical measure called entropy value. A maximum entropy valued frame is extracted, which provides maximum information compared to other frames of the same shot. Then, SURF feature-based square regions are extracted for embedding watermark. The SURF feature points are commonly invariant to rotation, scaling, and translating, so they naturally fit into the requirements of geometrically robust image watermarking. A DCT-based embedding technique has been introduced to achieve high robustness and imperceptibility.

### 3. PROPOSED SCHEME

A H.264 video watermarking system is explained in this section. First the input video is divided into frames, and then shot boundaries are detected using the PCC technique. Then, the entropy value of each frame of a shot is calculated, and the maximum entropy-valued frame of a shot is selected as a key-frame. The proposed shot boundary and key-frame selection algorithms are used for reducing watermark embedding time and providing robustness against distortion, noise, illumination changes, object motions, and camera operations such as zoom-in and zoom-out. A resilient SURF feature-based watermarking approach in the DCT domain is applied to each key-frame of a video. During the watermark embedding process, SURF

feature points are detected from each key-frame. These detected key-points are robust to various geometric and image transformations, such as scaling and rotation, blurring, and JPEG compression. The detected feature points are utilised for the generation of non-overlapped square regions for embedding and extraction of watermarks. The detailed watermarking process is explained.

### 3.1. Shot boundary detection and key-frame extraction

The process of proposed shot boundary detection using PCC is explained in this section. Initially, host video is decomposed into frames and then each frame is divided into R, G and B channels. In the proposed shot boundary detection, first frame of each channel of the host video is considered as a first frame (FF) for the first shot. The PCC is measured using (1) between first frame FF and successive frames ( $F_i$ ) for each red, green and blue channel of respective shot. If measured PCC value of FF and  $F_i$  is greater than threshold for each channel then that frame is added into the current shot, otherwise, it is considered as first frame of next shot. Same process is applied on each frame of the host video. The proposed shot boundary detection algorithm is described in Algorithm 1. In order to achieve high accuracy in shot boundary detection, we need to decide appropriate threshold values. For the PCC of red, green and blue channels, the threshold values are measured using the mean ( $\mu_R, \mu_G$  and  $\mu_B$ ), and variance ( $\sigma_R^2, \sigma_G^2$  and  $\sigma_B^2$ ) of  $PC_R, PC_G$ , and  $PC_B$  using (2) to (4).

PCC between two images P and Q is calculating in terms of covariance is given as in (1).

$$\rho_{cc}(P, Q) = \frac{1}{N-1} \times \frac{1}{M-1} \sum_{i=1}^N \sum_{j=1}^M \left( \frac{P(i,j) - \mu_P}{\sigma_P} \right) \left( \frac{Q(i,j) - \mu_Q}{\sigma_Q} \right) \quad (1)$$

Where,  $\mu_P$  and  $\sigma_P$  are the mean and standard deviation of P, respectively, and  $\mu_Q$  and  $\sigma_Q$  are the mean and standard deviation of Q, respectively.

$$Th_R = \mu_R + a \times \sigma_R^2 \quad (2)$$

$$Th_G = \mu_G + a \times \sigma_G^2 \quad (3)$$

$$Th_B = \mu_B + a \times \sigma_B^2 \quad (4)$$

The selected values of  $a$  is considered in this research is 1.2.

#### Algorithm 1: Shot boundary detection

```
Initially, the host video is preprocessed into number of frames ( $F_1, F_2, F_3, \dots, F_n$ )
FF =  $F_1$  // First frame of host video and first shot
For i=2 to n
     $PC_j = PCC(FF, F_i)$  // where j is R, G and B channel
    If  $PC_j > Th_j$  then // where  $Th_j$  is the threshold value of R, G and B channel.
         $F_i$  is added into current FF shot
    Else
         $F_i$  is considered as first frame of next shot
         $FF = F_i$ 
    End If
End For
```

Then, a key-frame is selected from each shot using a statistical entropy measure. The randomness that can be used to classify an image is measured using entropy [8]. Entropy value of each frame of a shot is measured, and the frame with the highest entropy value is chosen as a key-frame for the shot. The same procedure is repeated for each shot, to obtain all the key-frames. The following (5) is used to determine the entropy value.

$$Entropy_{F_j} = \sum p(F_{ij}) \times \log_2(p(F_{ij})) \quad (5)$$

p encloses the histogram counts of image and  $F_{ij}$  is the  $i^{th}$  frame of  $j^{th}$  shot.

$$KeyFrame_j = \max(Entropy_{F_j}) \quad (6)$$

Where,  $KeyFrame_j$  is the key-frame of the  $j^{th}$  shot.

### 3.2. Generation of invariant regions using SURF

In the proposed video watermarking approach, a SURF feature based invariant regions are detected for embedding and extraction of watermark. Initially, SURF feature points are extracted and those feature points are considered as centre of circular region with radius  $r$ . Each circular region is converted into square region using circumscribed square. It is a square surrounding a circle such that the circumference of the circle touches the midpoints of the four sides of the square. The diameter of the circle is equal to the side length of the square. Figure 1 shows the generation of square regions using Figures 1(a) to 1(d). Figure 1(a) shows the circumscribed square of a circle, Figure 1(b) shows the extracted SURF feature points; Figure 1(c) indicates the generated circular regions using feature points and radius. And Figure 1(d) shows the generated square regions using circular regions.

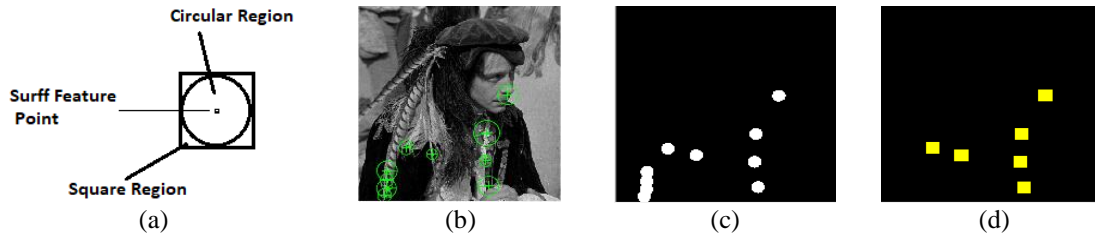


Figure 1. SURF feature based region extraction (a) circumscribed square of a circle (b) SURF feature points (c) Circular region with radius  $r$  (d) Non overlapped circumscribed square of a circles

### 3.3. Watermark embedding

In this section, the detailed watermark embedding procedure is explained:

- Step 1: Initially, shot boundaries are detected using PCC from input video and then key-frame for each shot is detected using entropy measure.
- Step 2: Each key-frame is converted from RGB color space to YCbCr color space and Y component is used for further processing.
- Step 3: SURF feature points are detected for Y component of each key-frame of the video.
- Step 4: Detected feature points are used to generate non-overlapped circumscribed square regions of size  $32 \times 32$ .
- Step 5: Each non-overlapped square region of a key-frame is further divided into non-overlapped sub-blocks of size  $8 \times 8$ .
- Step 6: DCT is applied on each sub-block ( $8 \times 8$ ) and get mid-frequency coefficients.
- Step 7: Un-correlated pseudorandom (PN) sequences such as Psudo0 and Psudo1 are generated using a secret key. Psudo0-sequence is used to embed watermark bit 0 and Psudo1-sequence is obtained to embed watermark bit 1. Size of each of the two PN-Sequence must be equal to the number of mid-frequency elements of  $8 \times 8$  DCT transformed sub-block.
- Step 8: The generated Psudo0 and Psudo1 are embedded with decimal sequence  $s$ , which is generated using secret prime number  $p$  into the mid-frequency coefficients of DCT using following (7).

$$IW_{mid} = \begin{cases} \text{if } W = 0 & I_{mid} + (\alpha \times Psudo_0 \times s) \\ \text{else} & I_{mid} + (\alpha \times Psudo_1 \times s) \end{cases} \quad (7)$$

Where,  $I_{mid}$  indicate mid-frequency coefficient,  $Psudo_0$  and  $Psudo_1$  are pseudorandom sequences,  $W$  is the watermark and  $\alpha$  indicates the scaling factor.

- Step 9: Inverse DCT is applied and then, the original square region is replaced with the watermarked one.
- Step 10: The above embedding operation is done repeatedly until all the invariant regions are watermarked.
- Step 11: From step-2 to step-9 all steps are applied on each key-frame of a shot and finally, watermarked video is generated.

### 3.4. Watermark extraction

We employ a correlation-based watermark detection approach to recover the watermark. Following steps are applied on watermarked video to extract watermark.

- Step 1: Shot boundaries are detected using PCC for watermarked video and then entropy measure is used to select watermarked key-frame.

- Step 2: Each watermarked key-frame is converted from RGB color space to YCbCr color space and Y component is used for further processing.
- Step 3: SURF feature points are extracted from Y component.
- Step 4: Detected feature points are used to generate non-overlapped circumscribed square regions of size  $32 \times 32$ .
- Step 5: Each non-overlapped square region ( $32 \times 32$ ) of a key-frame is further divided into non-overlapped sub-blocks of size of  $8 \times 8$ .
- Step 6: DCT is applied on each sub-block  $8 \times 8$  to get watermarked mid-frequency coefficient.
- Step 7: The s-sequence formed using the same prime number used in watermark embedding and it is correlated with the DCT watermarked mid coefficients values using (8). If the Corr is greater than Threshold T then watermark bit is 0 otherwise it is 1. Equation (9) is used to detect watermark.

$$Corr = \frac{1}{N} (IW'_{mid}, s) \quad (8)$$

$$W'_n = \begin{cases} 0 & \text{if } Corr > T \\ 1 & \text{else} \end{cases} \quad (9)$$

Where,  $IW'$  DCT coefficient of watermarked frame

s- Decimal sequence generated using prime number

T- Threshold (which is decided based on the need to minimize error.

Corr – is the correlation value

#### 4. EXPERIMENTAL RESULTS

The experimentation of proposed video watermarking scheme is estimated in MATLAB 2021a with I3 processor. The efficiency of the proposed scheme has been calculated on standard six videos: Bowling, Coastguard, Silent, Salesman, News and Foreman in terms of imperceptibility and robustness. The perceptual quality of watermarked video is measured using peak signal to noise ratio (PSNR). In terms of imperceptibility, a greater PSNR suggests better performance. PSNR of watermarked image is calculated using (11).

$$PSNR = 10 \log_{10} \frac{255^2}{\frac{1}{N \times M} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} (Of_{x,y} - Wf_{x,y})^2} \quad (10)$$

The robustness of the watermark is determined by its resistance to attempts to remove the watermark content using various types of image or signal processing attacks. The robustness of watermark is measured using normalized cross correlation (NCC). It is evaluated using (17).

$$NCC = \frac{\sum_{i=1}^N \sum_{j=1}^M W(i,j) W'(i,j)}{\sum_{i=1}^N \sum_{j=1}^M W^2(i,j) W'^2(i,j)} \quad (11)$$

##### 4.1. Imperceptibility and robustness analysis

To evaluate the invisibility of the proposed video watermarking approach, we have used PSNR as an evaluation parameter, by comparing the original video and the watermarked video. Table 1 shows the average PSNR values of watermarked videos. We present the average PSNR values for all frames in the host video to eliminate the effect of randomness. From Table 1, it is found that the average PSNR of all the watermarked videos across all videos is above 62.85 dB, which proves the good imperceptibility of the watermarked videos.

Table 1. The average PSNR values of the watermarked videos

Video	PSNR
Silent	63.21
Foreman	63.13
News	63.35

The NCC values of a content-based watermark extracted from a watermarked video frame following various attacks are shown in Figure 2. The proposed approach is shown to be resistant to image processing

attacks such as Gaussian noise, salt and pepper noise, Poisson noise, blurring, brightening, frame averaging, and swapping. The NCC value of the detected watermark is found to be between 0.9 and 1.

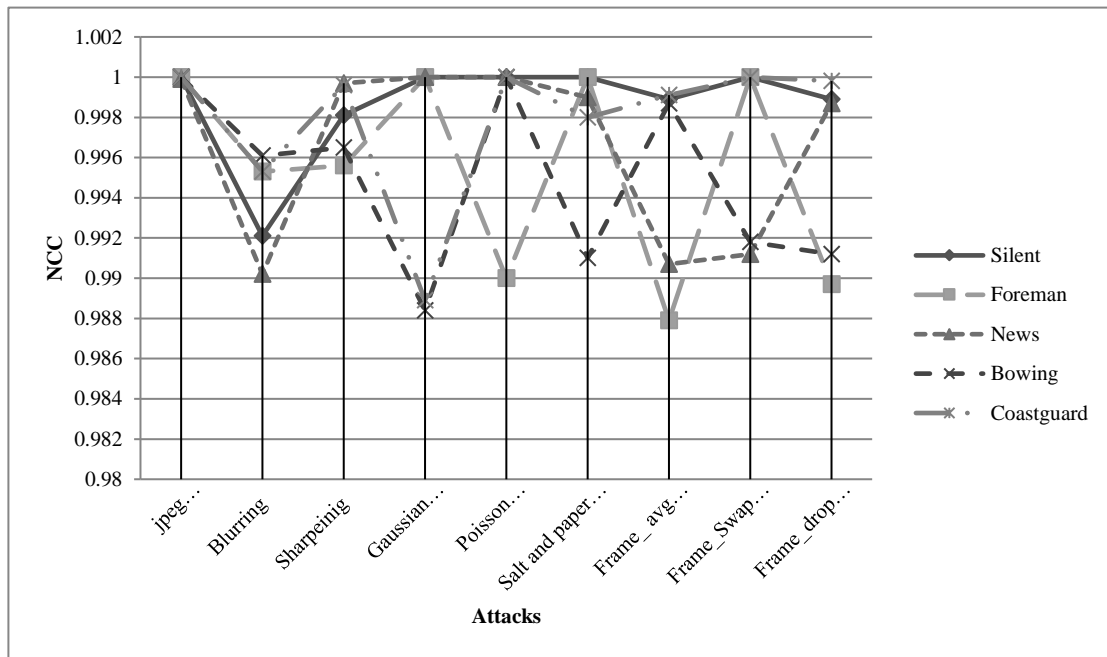


Figure 2. NCC values of watermark after applying various attacks

The proposed entropy-based key-frame extraction scheme is evaluated using precision and recall measure. In this research the detected key-frames are compared with the ground truth, which is generated manually by five human observers after watching the videos. The similarity between detected key-frames using proposed scheme and the ground truth is then measured. The recall and precision are computed using following (12) and (13) respectively:

$$Recall = \frac{TP}{TP+FN} \times 100 \% \quad (12)$$

$$Precision = \frac{TP}{TP+FP} \times 100 \% \quad (13)$$

Where, true positive (TP) indicates true positive means that the extracted frame by proposed scheme and by human observers are same. If the key-frame extracted by proposed scheme is not observed by human observed then that is called false positive (FP) and false negative (FN) means if the key-frame is observed by human but not by proposed scheme. The results of key-frame detection by proposed scheme for 'silent' video sequence are given in Figure 3 wherein, two different shots of a 'silent' video are shown; one is indicated using green color box and other by using red color box. Figure 4 shows accurately and semantically extracted key-frame of each shot shown in Figure 3 of video 'silent' using entropy measure.

## 4.2. Comparative analysis

### 4.2.1. Comparative analysis of proposed key-frame extraction

In this section, the results of proposed scheme are compared to existing video watermarking approaches in the literature. We have compared our results of SVD based key-frame with GSMD [8] scheme. Table 2 shows average values of recall and precision for all the above schemes. Recall and precision values obtained by the proposed method are 95%. So it is proved that the proposed entropy based key-frame extraction scheme outperforms the performances GSMD [8]. From the experimentation, it is revealed that the proposed key-frame extraction strategy is less difficult, segments shots precisely, and extracts key-frames from each shot with strong robustness.

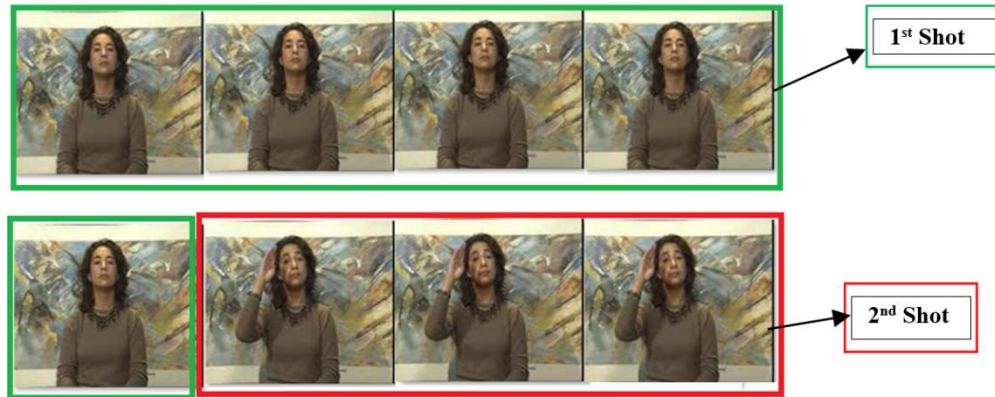


Figure 3. Example of shot boundaries detection from 'silent' video sequence



Figure 4. Extracted key-frames using entropy from each shot of Figure 3

Table 2. Comparative analysis of Recall and precision values of GSMD [8] and proposed scheme for 'silent video'

	GSMD [8]	Proposed
Recall	95	95.5
Precision	93	95

#### 4.2.2. Comparative analysis of proposed video watermarking scheme in terms of imperceptibility and robustness

In order to evaluate the performance of the proposed scheme, the results of the proposed video watermarking scheme is compared with the results of related video watermarking schemes given in [8]–[10], which we introduced and discussed in the related work. Table 3 shows the results of comparing the proposed scheme with other recent schemes under different kinds of attacks like blur, brighten, Gaussian noise attack, Salt and Pepper, median filtering, frame dropping and averaging. It is observed that the proposed scheme outperforms all the above mentioned existing schemes. It observed from the Table 3, the average NCC value of proposed scheme is higher than 99% for almost all the attacks.

Table 3. NCC Watermark Variation in Several Attacks vs. Existing Schemes [8]–[10]

Attacks	[8]	[9]	[10]	Proposed
Compression	0.997	0.9883		0.9991
Gaussian Noise	1	0.9844	0.92	1
Salt and Pepper	1	0.9844	0.87	1
Blurring	0.978		0.95	0.9911
Median Filtering	1	0.9965	0.93	0.9999
Frame averaging	0.99	0.99	0.9	0.9993
Brighten			0.94	1
Frame Dropping			0.95	1
Frame Swapping			0.92	1



The PSNR values obtained from the proposed algorithm is compared with the existing schemes which are shown in Figure 5. In human visual perception, PSNR value of above 50 dB is considered as better visual quality of watermarked video [9]. Figure 5 shows the average PSNR values of proposed scheme and existing schemes [2], [8]–[10]. It is observed that the proposed scheme shows PSNR value is above 62.8dB, which is better than the existing schemes [2], [8], [10] except [9]. As in [9], authors have considered only boundary blocks of selected key-frames for embedding and extraction of watermark.

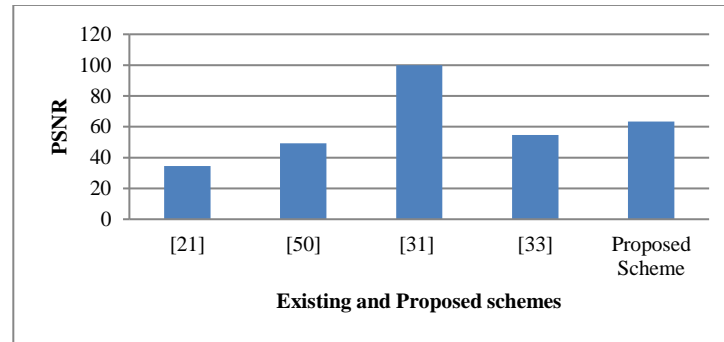


Figure 5. PSNR of watermarked video of proposed scheme compared with exiting methodology

## 5. CONCLUSION

In this research, we provide a blind, efficient, and secure video watermarking scheme for H.264/AVC videos. Initially, we have developed a novel and efficient shot detection technique using PCC. The reason behind the use of PCC is that it is robust against object motion, camera operation, and illumination changes. As embedding watermark in every frame is time-consuming, we have proposed an entropy based key-frame extraction algorithm, which is used for selecting key-frames from each shot of the video. From each key-frame, SURF feature point-based square regions are extracted for embedding and watermark extraction. These extracted regions are robust to various geometric and photographic transformations, such as scaling and rotation, blurring, and JPEG compression. Further, a DCT-based watermark embedding algorithm is used to improve the imperceptibility and robustness of the proposed watermarking scheme. Moreover, the embedded watermark is extracted blindly using the extraction function. In terms of imperceptibility, security, and resilience, promising results have been achieved using the proposed approach.

## ACKNOWLEDGEMENTS

The authors would like to thank Chairman, Kamal kishor kadam and Director, Dr. Geeta S. Lathkar, HOD of CSE department for their constant encouragement, valuable suggestions, and support for carrying out this work as a part of my Ph.D. work.





## REFERENCES

- [1] C. H. Wu, Y. Zheng, W. H. Ip, C. Y. Chan, K. L. Yung, and Z. M. Lu, "A flexible H.264/AVC compressed video watermarking scheme using particle swarm optimization based dither modulation," *AEU - International Journal of Electronics and Communications*, vol. 65, no. 1, pp. 27–36, 2011, doi: 10.1016/j.aeue.2010.02.003.
- [2] X. Jiang, Q. Liu, and Q. Wu, "A new video watermarking algorithm based on shot segmentation and block classification," *Multimedia Tools and Applications*, vol. 62, no. 3, pp. 545–560, 2013, doi: 10.1007/s11042-011-0857-3.
- [3] T. M. Thanh, P. T. Hiep, T. M. Tam, and K. Tanaka, "Robust semi-blind video watermarking based on frame-patch matching," *AEU - International Journal of Electronics and Communications*, vol. 68, no. 10, pp. 1007–1015, 2014, doi: 10.1016/j.aeue.2014.05.004.
- [4] N. I. Yassin, N. M. Salem, and M. I. E. Adawy, "Entropy based video watermarking scheme using wavelet transform and Principle Component Analysis," *International Conference on Engineering and Technology, ICET 2012 - Conference Booklet*, vol. 9, no. 1, pp. 296–301, 2012, doi: 10.1109/ICEngTechnol.2012.6396128.
- [5] C. Hsu and Y. Hou, "A Visual Cryptography and Statistics Based Method for Ownership Identification of Digital Images," *Proceedings of World Academy of Science, Engineering and Technology*, vol. 1, no. 2, pp. 665–668, 2007.
- [6] M. Masoumi and S. Amiri, "A blind scene-based watermarking for video copyright protection," *AEU - International Journal of Electronics and Communications*, vol. 67, no. 6, pp. 528–535, 2013, doi: 10.1016/j.aeue.2012.11.009.
- [7] C. Li, Y. Yang, K. Liu, and L. Tian, "A Semi-Fragile Video Watermarking Algorithm Based on H.264/AVC," *Wireless Communications and Mobile Computing*, vol. 2020, 2020, doi: 10.1155/2020/8848553.
- [8] Y. Himeur and A. Boukabou, "A robust and secure key-frames based video watermarking system using chaotic encryption,"





- Multimedia Tools and Applications*, vol. 77, no. 7, pp. 8603–8627, 2018, doi: 10.1007/s11042-017-4754-2.
- [9] Z. Bahrami and F. Akhlaghian Tab, “A new robust video watermarking algorithm based on SURF features and block classification,” *Multimedia Tools and Applications*, vol. 77, no. 1, pp. 327–345, 2018, doi: 10.1007/s11042-016-4226-0.
- [10] S. P. A. Sathya and S. Ramakrishnan, “Non-redundant frame identification and keyframe selection in DWT-PCA domain for authentication of video,” *IET Image Processing*, vol. 14, no. 2, pp. 366–375, 2020, doi: 10.1049/iet-ipr.2019.0341.
- [11] A. Hammami, A. Ben Hamida, and C. Ben Amar, “Blind semi-fragile watermarking scheme for video authentication in video surveillance context,” *Multimedia Tools and Applications*, vol. 80, no. 5, pp. 7479–7513, 2021, doi: 10.1007/s11042-020-09982-4.
- [12] L. Tian, H. Dai, and C. Li, “A Semi-fragile Video Watermarking Algorithm Based On Chromatic Residual DCT,” *Multimedia Tools and Applications*, vol. 79, no. 3–4, pp. 1759–1779, 2020, doi: 10.1007/s11042-019-08256-y.
- [13] Z. Cao and L. Wang, “A secure video watermarking technique based on hyperchaotic Lorentz system,” *Multimedia Tools and Applications*, vol. 78, no. 18, pp. 26089–26109, 2019, doi: 10.1007/s11042-019-07809-5.
- [14] R. Munir and Harlili, “A Secure Fragile Video Watermarking Algorithm for Content Authentication Based on Arnold Cat Map,” *Proceedings of 2019 4th International Conference on Information Technology: Encompassing Intelligent Technology and Innovation Towards the New Era of Human Life, InCIT 2019*, pp. 32–37, 2019, doi: 10.1109/INCIT.2019.8912074.
- [15] K. S. Wong, C. S. Chan, and A. P. Maungmaung, “Lightweight authentication for MP4 format container using subtitle track,” *IEICE Transactions on Information and Systems*, vol. E103D, no. 1, pp. 2–10, 2020, doi: 10.1587/transinf.2019MUI0001.
- [16] A. P. M. Maung, Y. Tew, and K. S. Wong, “Authentication of MP4 file by perceptual hash and data hiding,” *Malaysian Journal of Computer Science*, vol. 32, no. 4, pp. 304–314, 2019, doi: 10.22452/MJCS.VOL32NO4.4.
- [17] Y. Vybornova, “A New Watermarking Method for Video Authentication with Tamper Localization,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12334 LNCS, pp. 201–213, 2020, doi: 10.1007/978-3-030-59006-2\_18.
- [18] C. N. Sujatha and P. Sathyanarayana, “DWT-based blind video watermarking using image scrambling technique,” *Smart Innovation, Systems and Technologies*, vol. 106, pp. 621–628, 2019, doi: 10.1007/978-981-13-1742-2\_62.
- [19] B. P. Aditya, U. G. K. Avaneesh, K. Adithya, A. Murthy, R. Sandeep, and B. Kavyashree, “Invisible Semi-Fragile Watermarking and Steganography of Digital Videos for Content Authentication and Data Hiding,” *International Journal of Image and Graphics*, vol. 19, no. 3, 2019, doi: 10.1142/S0219467819500153.

## BIOGRAPHIES OF AUTHORS



**Kapre Bhagyashri S.**     is a master's holder in Computer Science and Engineering from the Swami ramanand Tirth University, Nanded, Maharashtra. She has worked as a Assistant professor in Computer Science and Engineering Department. She can be contacted at email: kapre\_bs@mgmcen.ac.in.



**Rajurkar Archana M.**     is working in MGM's College of Engineering, Nanded, Maharashtra as Professor, Head of the Department (Associate). She has 31 years experience. She has published 46 international and 13 national conference papers. Above 24 journals are published in international journals. Her academic qualification is Ph.D from IIT Roorkee. Her research area includes image and video retrieval. She can be contacted at email: rajurkar\_am@mgmcen.ac.in.