# DESIGNING OF A RESILIENT WATERMARKING SCHEME UTILIZING LIFTING WAVELET TRANSFORM AND SIFT FOR VIDEOS

Himanshu Agarwal[1], Shweta Agarwal[2], Rakesh Ahuja[3]

[1]Assistant Professor, Department of Computer Science and Engineering, Moradabad Institute of Technology, Moradabad (U.P.) – India
Himanshu.agg2000@gmail.com
[2]Assistant Professor, Department of Mathematics, Ramabai Ambedkar Govt. Degree College, Amroha-Gajraula (U.P.) – India
Shweta.agg2000@gmail.com
[3]Professor (Research), Institute of Engineering and Technology, Chitkara University, (Punjab) – India
Ahuja2305@gmail.com

***ABSTRACT***

*With the advent of advanced software tools for editing of videos, it is possible to manipulate digital content, leading to substantial financial losses for content creators. Various industries, including entertainment, suffer greatly from the detrimental effects of digital piracy. Platforms such as WhatsApp, Facebook, Instagram, and Twitter have an immense user base, with millions of individuals actively utilizing these platforms. Unfortunately, many users share digital content without appropriately acknowledging or crediting the original creators. To safeguard the rights and lawful reparation of data owners, the implementation of digital watermarking emerges as a highly effective technique. However, existing watermarking methods often fail to adequately preserve the intrinsic relationships within the host data due to the suboptimal selection of locations to embed the watermark. As a result, they prove to be ineffective in combating digital piracy. In light of these challenges, this research paper presents a digital video watermarking technique that leverages the 2D-Lifting Wavelet Transform (2D-LWT) and Scale Invariant Feature Transform (SIFT) to analyze, extract, and process essential patterns or features. Specifically, the SIFT method is employed to extract ten invariant key points from the approximation part of 2D-LWT, exploiting the luminance part of every video frame. The proposed technique is rigorously evaluated through popular objective metrics, considering various attacks such as geometrical distortions, temporal desynchronization, video compression, and the insertion of noise. The evaluation results demonstrate a significantly strong correlation value, close to one, and excellent imperceptibility, establishing the superiority and practical viability of the proposed scheme.*

***KEYWORDS***—*Digital Piracy, 2D-LWT, SIFT, HD Videos, Key Points.*

## 1. INTRODUCTION

The widespread utilization of multimedia materials in online education, social media entertainment, and the unauthorized distribution of videos such as, trendy TV shows and highly rated movies via the use of internet have led to significant financial losses for the rightful owners of these media contents. Engaging in activities such as illicit recording, alteration, reproduction, or sharing and dissemination of digital multimedia materials are classified as digital piracy. In order to safeguard the intellectual property rights of content creators, the practice of watermarking has emerged as a useful strategy over the past two decades. Digital watermarking involves the embedding of identifying

DOI: 10.5281/zenodo.10441424

information, such as a logo or pattern, into digital content in a manner that allows it to be extracted later. safeguards the identity of the owner by preventing illegal manipulation and sharing, while also providing evidence that can be utilized in legal proceedings.

Several research papers [1, 2, 3] have presented numerous digital watermarking techniques by categorized them into spatial and transform domain based on how the extra information (watermark) is inserted into the digital content. In spatial domain techniques [4, 5] the protection against unauthorized processing is achieved by directly modifying the pixel values. However, transform domain techniques [6, 7] involve transforming the data into different frequency components and choosing the appropriate component for inserting the additional information.

However, all the methods proposed by the researchers mentioned in the papers [8, 9, 10, 11] and discussed above failed to retain the embedded watermark value due to inadequate selection of watermarking locations. To address this limitation, we utilize data features in the watermarking process by making slight modifications to embed the watermark [12, 13]. Moreover almost all the researchers focuses their work considering SD quality videos even though most of videos produced in modern time having HD quality[14].

The introduced approach introduces an innovative video watermarking technique that combines 2D-LWT (2D Lifting Wavelet Transform) with semantically meaningful features obtained through the SIFT (Scale-Invariant Feature Transform) technique. This method effectively addresses unauthorized processing while simultaneously preserving the video's visual fidelity.

### 1.1 Key Points and Offerings of this Study

This study proposes a HD video watermarking scheme utilizing 2D-LWT and SIFT. The preference for the 2D-LWT above the other popular transforms like DWT and DCT stems from its ideal reconstruction capability and efficient in-place computations during the wavelet transform. Invariant key points detected via SIFT are used as for watermark embedding.

This research introduces the contributions as outlined below:

- An efficient and reversible technique called 2D-LWT is employed for wavelet transforms. The main advantage of using 2D-LWT is that it reduces computation by half compared to similar method of DWT and it also minimizes rounding errors from floating-point arithmetic, making it highly suitable for the applications that process the videos.
- In order to enhance the resilience of the watermark, the proposed scheme employs the scale and rotation invariant method called SIFT which express robustness to occlusion.
- In order to enhance practical applicability, a grayscale image is utilized as watermark instead of binary image. Grayscale watermark can accommodate a larger amount of content and exhibit improved identification characteristics compared to binary watermark.
- The efficacy of the proposed scheme is demonstrated through evaluation results, showing correlation coefficients (CC) and Mean Perceptual Similarity Index (MPSI) close to 1 and high values of Average PSNR between watermarks and video frames in practical applications.

The ensuing sections of this document are structured as follows: In Section 2, we delve into the fundamental principles of LWT, SIFT methods. Following that, Section 3 will present the innovative approach we have developed for attack-resistant watermarking techniques. We will then move on to Section 4, where we will outline the experimental framework and showcase the compelling outcomes achieved through our method. Finally, in Section 5, we will draw insightful conclusions from our study.

## 2. Research Background

This section provides a concise overview of the fundamental concept and mathematical underpinnings necessary for the proposed video watermarking scheme, including LWT (Lifting Wavelet Transform) and the application of SIFT (Scale-Invariant Feature Transform) for feature point detection.

DOI: 10.5281/zenodo.10441424

## 2.1 Lifting Wavelet Transform (LWT)

The Lifting Wavelet Transformation (LWT) is an advanced technique employed in signal and image processing, providing a flexible and optimized framework for wavelet decomposition and reconstruction. Its innovative features offer significant advantages over conventional methods like the Discrete Wavelet Transform (DWT), including enhanced accuracy, reduced computational complexity, and improved efficiency. The LWT utilizes a set of progressive steps to transform the input signal or image, incorporating updating and predicting operations that are iteratively applied to acquire wavelet coefficients at multiple scales. The process begins by partitioning the input signal into two distinct sample groups: even samples (approximation coefficients) and odd samples (detail coefficients). While the even samples capture the low-frequency content, the odd samples encompass the high-frequency details. The LWT then engages in alternating updating and predicting steps to refine these coefficients [15].
During the updating step, the LWT adjusts the even samples by considering the odd samples and previously computed approximation coefficients. This process effectively captures the interdependencies between adjacent samples while minimizing redundancy in the signal. Conversely, the predicting step employs the updated even samples and the original odd samples to estimate new odd samples. This procedure efficiently enhances the high-frequency components while preserving essential details. Through the iterative application of these updating and predicting steps, the LWT performs a comprehensive multi-scale decomposition, generating wavelet coefficients at varying scales or levels. This decomposition yields a comprehensive representation of the input signal, delineating its frequency content across multiple resolutions.
Additionally, the LWT facilitates seamless and efficient reconstruction of the original signal or image from the wavelet coefficients. The reconstruction process entails the application of inverse lifting steps in reverse order, ensuring minimal information loss during restoration. A notable advantage of the LWT lies in its ability to perform in-place calculations, minimizing the need for additional memory for intermediate results. This intrinsic feature renders the LWT particularly well-suited for real-time applications and memory-constrained systems [16]. In summary, the Enhanced Wavelet Transformation represents a versatile and optimized approach to wavelet decomposition and reconstruction. Its applications span diverse domains such as image and signal processing, compression, denoising, feature extraction, and watermarking, among others, providing significant advancements in accuracy and computational efficiency.

## 2.2 Identification of Feature Points through SIFT

The Scale-Invariant Feature Transform (SIFT) is a powerful computer vision technique used for feature detection and extraction in digital images. It provides robustness to variations in scale, rotation, translation, and changes in viewpoint, making it highly effective in various applications such as object recognition, image stitching, and 3D reconstruction. At its core, SIFT identifies distinctive local features within an image that are invariant to changes in scale. These features, known as keypoints or interest points, are selected based on their stability and uniqueness across different scales [17]. SIFT achieves this by employing a multi-scale approach, where the image is analyzed at different levels of detail using a series of Gaussian filters. The SIFT algorithm starts by constructing a scale space representation of the image, where each scale level is generated by convolving the original image with Gaussian filters of different sizes. This process allows for the detection of features at varying scales, from fine details to larger structures.
The SIFT applies a Difference of Gaussian (DoG) operator to detect potential keypoints. The DoG operator is obtained by subtracting adjacent scales of the Gaussian-blurred images, highlighting areas with significant intensity changes. These regions correspond to keypoints that are distinctive and stable under scale variations. After identifying potential keypoints, SIFT applies a process called keypoint localization. This step involves precise localization of the keypoints by refining their positions based on the local image gradients and eliminating unstable keypoints that are likely to be noise or artifacts. To ensure scale invariance, SIFT computes a descriptor for each key point that captures its local appearance. This descriptor is computed based on the local image gradient orientations and magnitudes within a region surrounding the key point. By considering the orientation

DOI: 10.5281/zenodo.10441424

and intensity variations, the descriptor is designed to be robust to changes in scale, rotation, and illumination. The final step of the SIFT algorithm is feature matching, where keypoints from different images are compared to establish correspondences. This enables tasks such as image stitching, object recognition, and tracking by matching similar keypoints across multiple images.
One of the key advantages of SIFT is its robustness to various transformations and its ability to handle cluttered and occluded scenes [18]. It has proven to be highly effective even in challenging scenarios with partial object occlusion or significant changes in viewpoint. In summary, the Scale-Invariant Feature Transform (SIFT) is a robust computer vision technique that detects and extracts distinctive features from digital images. By employing a multi-scale approach and computing descriptors that are invariant to scale, rotation, and translation, SIFT enables reliable feature matching and robust performance in various applications requiring accurate and invariant image analysis.

## 3. PROPOSED VIDEO WATERMARKING METHOD

In this section, we present the comprehensive details of the watermark embedding and extraction stages employed by our proposed method. For enhanced identification characteristics and a stronger contextual relationship effect, we utilize grayscale image as watermark instead of binary watermark [19]. Grayscale watermarks offer a greater capacity for carrying a significant amount of watermark information, making them more suitable for practical applications. To ensure resilience against high lossy compression, our proposed method focuses on utilizing the luminance part of video rather than the chrominance components [20]. Subsection 4.1 elaborates on the meticulous selection process of embedding locations within the video frames. The block diagram in Figure 4.1 illustrates the process of watermark embedding, while Figure 4.2 showcases the corresponding watermark extraction method. The extraction method recognize the keypoints in a manner similar to the one used for embedding the watermark. By subtracting the intensity values of the identified locations from the intensity values obtained through the same procedure applied to the original video, the corresponding watermark values are extracted.

### 3.1 FINDING EMBEDDING SITES

In this subsection, we outline the process of selecting watermark locations using Algorithm 1. Our method initiates by converting every frame into its corresponding YCbCr components and subsequently transforming the luminance part 'Y' into frequency coefficients through LWT. The Algorithm 1 provides a detailed explanation of how our proposed approach calculates ten key feature points. These feature points are obtained by applying the SIFT algorithm to the approximation component of the LWT transformed frame, ultimately serving as the designated locations for our watermark embedding process.

| **Algorithm 1: Finding the Watermarking Locations** |
|---|
| Input: HD RGB video (HDV)<br>Output: Watermark Locations (WL)<br>Begin:<br>1. for j=1 to number of frames, convert the frame Fj in YCbCr and decompose each component<br>(Fj_Y, Fj_Cb, Fj_Cr) →rgbtoycbcr(Fj)<br>2. Apply 2D-LWT on FjY to separate the frame into its constituent components. The approximation component is AC and the detailed components are represented as DH, DV and DD.<br>[AC, DH, DV, DD] = 2D-LWT(Fj_Y)<br>3. Finding ten key feature as watermark locations (WLs) on AC using SIFT.<br>WLs = SIFT_WLs [AC,10]<br>4. End of for loop<br>End |

### 3.2 WATERMARK EMBEDDING AND EXTRACTION PROCESS

DOI: 10.5281/zenodo.10441424

In order to illustrate the processes of watermark embedding and extraction, we refer to Figure 3.1 and Figure 3.2. The embedding process begins by acquiring the watermark and subsequently scaling it using a factor 'α' to ensure both robustness and perceptibility. The watermark is then molded onto the Watermarking Locations (WLs) obtained through the utilization of Algorithm 1.On the other hand, the watermark extraction process follows a reverse course compared to the embedding process. Initially, we extract the frames from both the watermarked and the original videos. Subsequently, we calculate the PSNR value between the watermarked video frame and the corresponding frame from the original video. If the calculated value of PSNR remains below a predefined threshold, denoted as '*TS*' (in the developed algorithm, we consider '*TS*' as 10), we discard the current frame and repeat the procedure until we locate a frame with a PSNR value equal to or higher than '*TS*'. The selected frames for extraction are then transformed into YCbCr components using standard transformation equations. By applying 2D-LWT on the Y component and employing SIFT on its approximation component (AC), we precisely determine the original insertion locations of the watermark. The values of the extracted watermark are then divided by the factor of scale denoted by 'α' to obtain the actual values of intensity.

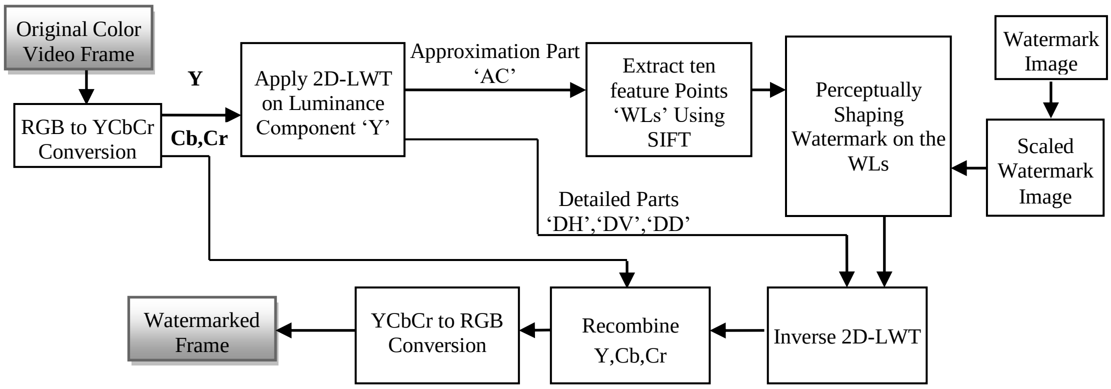



**Fig. 3.1.** Video Watermark Embedding Process using 2D-LWT and SIFT

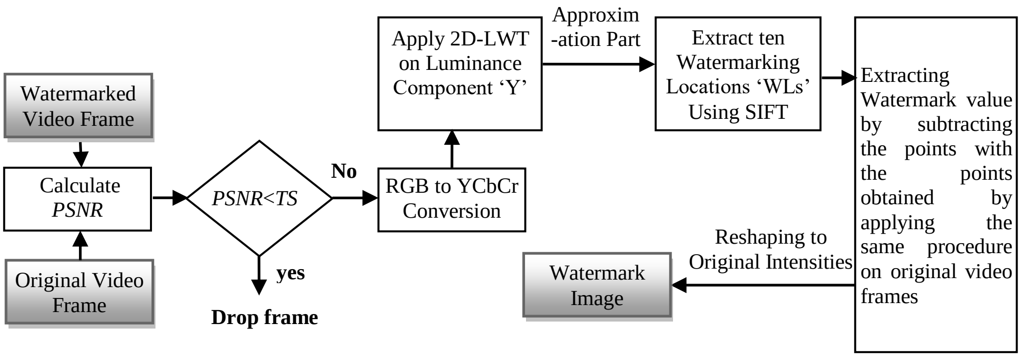



**Fig. 3.2.** Video Watermark Extraction Process using 2D-LWT and SIFT

## 4. Setup and Findings of the Experiment

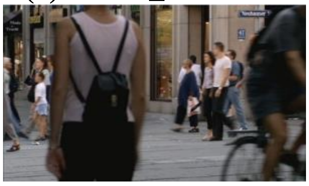
The proposed scheme is simulated on a PC comprises of a tenth generation i3 processor and 12 GB of RAM, utilizing MATLAB R2016a software. To assess the effectiveness of the method, we conducted performance evaluations using four different high-definition (HD) color videos as test content. These videos, namely 'Speed_Bag.avi,' 'Snow_Mountain.avi,' 'Tractor.avi,' and 'Pedestrian.avi,' each possess dimensions of 1920 x 1080 pixels and consist of 300 frames. To assess the impact different attacks, all videos were initially transformed into uncompressed AVI format for comprehensive investigation. For reference, Fig. 4.1 showcases the frames of the four high-definition videos.

DOI: 10.5281/zenodo.10441424


(a) ‘Speed_Bag.avi’

(b) ‘Snow_Mountain.avi’

(c) ‘Tractor.avi’



(d) ‘Pedestrian.avi’

**Fig. 4.1** High Definition Videos (a, b, c, d)

To validate the efficacy of the proposed methodology, all four videos were individually subjected to watermarking using the grayscale image "cameraman.tiff," a standard image encompassing a wide range of gray shades. Figuring as a benchmark, Fig. 4.2 showcases both the original frame and the corresponding watermarked frame of 'Snow_Mountain.avi' obtained through the implementation of the proposed method. In order to disrupt the synchronization of crucial information within the video frames, a range of attack categories, including temporal synchronization, compression, and insertion of various types of noise, were employed. The performance of the proposed method was evaluated by testing the videos after these attacks, focusing on maintaining the imperceptibility and resilience of the watermark. Prior to conducting the tests, careful consideration was given to selecting appropriate intensities for the distortions to ensure practical usefulness of the proposed method in video processing applications.

(a) Video Frame **:**Original

(b) Video Frame: Watermarked

**Fig. 4.2** Comparative Frames - Original vs. Watermarked

### 4.1 Evaluation of Imperceptibility

Evaluating the imperceptibility of the watermarked video represents a pivotal and challenging aspect in assessing the impact of the embedded watermark on video quality. In order to develop an effective watermarking scheme, it is crucial to employ reliable and valid testing procedures. Given the easy accessibility of flawless video editing software, it is imperative that any distortions in the resulting watermarked video remain undetectable by exploiting the Human Visual System (HVS). To ensure the dependability of the proposed scheme, two widely recognized objective metrics, namely Average Peak Signal-to-Noise Ratio (APSNR) and Mean Perceptual Similarity Index (MPSI), are employed to evaluate the imperceptibility of the videos after watermarking. These metrics serve as reliable indicators in assessing the extent to which the watermarked videos align with perceptual expectations, enabling the determination of the overall effectiveness of the proposed method.

Average Peak Signal-to-Noise Ratio (APSNR): The APSNR is determined by applying Equation (4.1), which involves calculating the PSNR between the watermarked video frame (WVF) and the corresponding original video frame (OVF) of dimensions H and W. The Mean Square Error (MSE) is utilized in Equation (4.2) to quantitatively measure the disparity between WVF and OVF.

DOI: 10.5281/zenodo.10441424

$$\text{PSNR} = 10\log_{10}\frac{(Max\_In)^2}{\text{MSE}} \tag{4.1}$$

$$\text{MSE} = \frac{\sum_{p=1}^{H}\sum_{q=1}^{W}\left[OVF(p,q) - WVF(p,q)\right]^2}{\text{HxW}} \tag{4.2}$$

Here, Max_In represents the maximum intensity value of the frame under examination, set to 255. The APSNR is derived by computing the average value of the calculated PSNR across all frames.

**Mean Perceptual Similarity Index (MPSI):** The Mean Perceptual Similarity Index (MPSI) utilizes the Structural Similarity Index Metric (SSIM) to assess the perceptual similarity between individual frames of the original and watermarked videos in a block-wise manner. The SSIM value ranges from -1 to +1, where a value of +1 indicates perfect similarity between the compared video frames. By calculating the SSIM for each frame and subsequently obtaining the mean value, we derive the MPSI. The SSIM can be computed using Equation (4.3) as follows:

$$\text{SSIM}(OVF, WVF) = \left[\text{Lc}(OVF, WVF)^{\delta}.\text{Cc}(OVF, WVF)^{\eta}.\text{Sc}(OVF, WVF)^{\xi}\right] \tag{4.3}$$

The luminance comparison, contrast comparison, and structure comparison functions are denoted as Lc, Cc, and Sc, respectively. These functions enable the evaluation of the relative importance of luminance, contrast, and structure components. To appropriately adjust the significance of each component, positive parameters $\delta$, $\eta$, and $\xi$ are commonly employed [21].

Table 4.1 presents the outcomes of evaluating the video's quality and similarity prior to any attack by examining the average values of APSNR and MPSI. Remarkably high values of APSNR and MPSI were obtained across all videos, indicating the exceptional capability of the proposed scheme in achieving superior visual quality following the watermarking process.

**Table 4.1** *APSNR* and *MPSI Values* Before Undergoing to Any Attack.

| Type of Attack | Speed_Bag | | Snow_Mountain | | Tractor | | Pedestrian | |
|---|---|---|---|---|---|---|---|---|
| | APSNR | MPSI | APSNR | MPSI | APSNR | MPSI | APSNR | MPSI |
| **No Attack** | 72.61 | 0.99 | 72.92 | 1 | 73.23 | 1 | 70.08 | 0.99 |

## 4.2 Evaluation of Robustness

The correlation between the two watermarks the embedded one and its extracted counterpart is evaluated using the Coefficient of the Correlation (CC), a widely recognized metric for assessing robustness. When the CC value between the two watermarks the embedded and the extracted approaches to one, it signifies that the extracted watermark is nearly identical to the embedded watermark, thereby demonstrating the resilience of the proposed scheme against signal modifications. Conversely, a CC value close to zero indicates dissimilarity between the embedded and extracted watermarks. The calculation of CC is performed using Equation (4.4).

$$\text{CC} = \frac{\sum_m\sum_n\left(OW_{mn} - \overline{OW}\right)\left(EW_{mn} - \overline{EW}\right)}{\sqrt{\left(\sum_m\sum_n\left(OW_{mn} - \overline{OW}\right)^2\right)\left(\sum_m\sum_n\left(EW_{mn} - \overline{EW}\right)^2\right)}} \tag{4.4}$$

Here, $OW$ and $EW$ represent the watermark matrices of identical dimensions (m x n), with $\overline{OW}$ representing the mean value of the original watermark signal, and $\overline{EW}$ denoting the mean value of the extracted watermark signal. To assess the performance of the proposed method, the grayscale image, namely 'cameraman.tiff,' was embedded as watermark signal in the test videos. Fig. 4.3 depicts the original watermark and the corresponding extracted watermark from the watermarked video 'Snow_Mountain.avi' when subjected to a frame insertion attack.

DOI: 10.5281/zenodo.10441424

| | |
|---|---|
|  |  |
| (a)<br>Original Watermark<br>“cameraman.tiff” | (b)<br>Recovered Watermark<br>“cameraman.tiff” after<br>frame insertion attack |

**Fig. 4.3.** Comparison of Original and Extracted Watermarks in the Presence of Frame Insertion Attack

The resilience of the watermark in all the HD watermarked videos, subjected to a range of attacks, is quantified by calculating the average Correlation Coefficient (CC) value. The watermark used for this evaluation is 'cameraman.tiff'. The average value of CC considering all videos are presented in Table 4.2. To ensure optimal robustness, a minimum acceptable CC value of 0.7 is considered as the threshold for determining copyright, as suggested by Agarwal and Husain [14]. Remarkably, in all cases, CC values exceeding 0.7 were achieved, confirming the proposed method's capacity to deliver a satisfactory level of robustness when confronted with various attacks.

**Table 4.2** Average Value of CC, APSNR and MPSI for Watermarked Videos.

| Type of Attack | CC | APSNR | MPSI |
|---|---|---|---|
| **No Attack** | 0.96 | 72.21 | 1 |
| **Frame Swapping 5%** | 0.92 | 71.29 | 0.99 |
| **Frame Swapping 10%** | 0.80 | 69.61 | 0.99 |
| **Frame Averaging 5%** | 0.87 | 71.21 | 0.99 |
| **Frame Averaging 10%** | 0.81 | 68.15 | 0.99 |
| **Frame Insertion (In Any Number)** | 0.96 | 72.21 | 1 |
| **Frame Rate Change 30 to 15 fps** | 0.96 | 72.21 | 1 |
| **Frame Rate Change 30 to 45 fps** | 0.96 | 72.21 | 1 |
| **Lossy Compression Motion JPEG 2000 with CR = 5** | 0.93 | 51.03 | 0.99 |
| **Lossy Compression Motion JPEG 2000 with CR = 10** | 0.92 | 48.01 | 0.99 |
| **MPEG-4 AVC (.avi to .mp4 conversion)** | 0.71 | 37.99 | 0.97 |
| **Salt & Pepper Noise Density 0.001** | 0.87 | 35.01 | 0.98 |
| **Salt & Pepper Noise Density 0.0005** | 0.89 | 38.11 | 0.98 |
| **Salt & Pepper Noise Density 0.0003** | 0.91 | 40.75 | 0.99 |

The efficacy of the scheme regarding robustness and the resulting imperceptibility of the watermark is presented in Table 4.2. The table showcases the results obtained after subjecting the videos to different temporal desynchronization, video compression, and noise insertion attacks. Notably, it is evident from Table 4.2 that the proposed method successfully maintains visual quality and watermark robustness despite variations in frame rates, whether they decrease or increase. The Average Peak Signal-to-Noise Ratio (APSNR) consistently exceeds 60 dB, indicating high fidelity in video quality. Additionally, the Mean Perceptual Similarity Index (MPSI) consistently reaches a value of one across various cases, demonstrating the preservation of perceptual similarity between original and watermarked videos. The Correlation Coefficient (CC) values generally exceed or approach 0.9 in all cases, except for the MPEG-4 AVC attack that significantly reduces video size. However, it is evident

DOI: 10.5281/zenodo.10441424

from the table that the CC values consistently remain above the threshold of 0.7 for all types of attacks, underscoring the effectiveness of the proposed method in safeguarding against such attacks.

**Table 4.3** Robustness based on Average CC Values following Geometric Synchronization Attacks

| Type of Attack | CC |
|---|---|
| **Cropping 5%** | 0.97 |
| **Cropping 10%** | 0.94 |
| **Cropping 20%** | 0.91 |
| **Uniform Scaling 0.80** | 0.85 |
| **Uniform Scaling 0.95** | 0.91 |
| **Uniform Scaling 1.05** | 0.90 |
| **Uniform Scaling 1.20** | 0.86 |

Furthermore, to evaluate the robustness of the watermark, the quantitative results in terms of Correlation Coefficient (CC) values for two well-known geometric synchronization attacks are provided in Table 4.3. The first attack involves uniform cropping of video frames, with cropping percentages of 5%, 10%, and 20%. The second attack involves scaling the videos using various scaling parameters. The results demonstrate the effectiveness of the proposed scheme in maintaining high robustness against both cropping and scaling attacks.

In conclusion, based on the simulation results of the proposed scheme, it can be observed that the imperceptibility level remains excellent, as indicated by the high values of APSNR and MPSI, even when the videos are subjected to the aforementioned video attacks.

### 4.3 Comparative Analysis of the Proposed Methodology versus Cutting-Edge Schemes

In order to thoroughly assess the robustness of the present approach in safeguarding digital watermark against various threats, the achieved outcomes of the scheme is contrasted with the findings from cutting-edge methodologies [12, 27] that demonstrate resilience in the face of diverse attacks. Additionally, certain elements within the comparative Table 4.4 were deliberately annotated with dashed lines, indicating that the authors of said methodologies neglected to evaluate their schemes against the aforementioned attacks.

**Table 4.4** Comparative Analysis of Robustness between Proposed Vs. Cutting Edge Schemes

| S. No. | Type of Attack | Agarwal et al. [12] | Shukla and Sharma [22] | Proposed |
|---|---|---|---|---|
| 1 | Salt & Pepper Noise (Density = 0.001) | 0.84 | 0.74 (Density= 0.05) | 0.87 |
| 2 | Swapping of 5% Frames | 0.87 | 0.49 (Three frames swapped) | 0.92 |
| 3 | Frame Insertion | 0.92 (Insertion of Ten frames) | - | 0.96 (Any Number of frames) |
| 4 | Cropping 20% | 0.96 (10% Cropping) | - | 0.91 |

## 5. Conclusion

This paper presents an innovative approach aimed at safeguarding the intellectual property rights and ensuring fair compensation for content creators. Drawing inspiration from the remarkable achievements in diverse areas of computer science, we have developed an advanced video watermarking technique that is both imperceptible and robust. Our proposed methodology leverages the SIFT (Scale-Invariant Feature Transform) technique to identify invariant feature points for precise watermark embedding. To enhance resilience against compression attacks, we strategically embed the watermark in the luminance component of the video frame. Furthermore, we utilize 2D-LWT (2D-

DOI: 10.5281/zenodo.10441424

Discrete Wavelet Transform) for its renowned capabilities in flawless reconstruction and streamlined computations when judge against Discrete Wavelet Transform.

To evaluate the effectiveness of our approach, we subject it to rigorous testing in the face of diverse categories of temporal manipulation, noise addition, and video compression attacks. A grayscale image is employed as the watermark, while commonly used metrics such as APSNR (Average Peak Signal-to-Noise Ratio) and modern MPSI (Mean Perceptual Similarity Index) are utilized to assess imperceptibility. Additionally, the strength of the embedded watermark is evaluated using the correlation coefficient metric. The simulation results confidently demonstrate that our proposed method excels in both resilience against attacks and preservation of perceptual fidelity among the evaluated videos. This breakthrough offers a promising solution for protecting digital content and ensuring the rightful compensation of content owners.

## REFERENCES

[1]. Agarwal, Himanshu, Rakesh Ahuja, and S. S. Bedi. "Highly robust and imperceptible luminance based hybrid digital video watermarking scheme for ownership protection." International Journal of Image, Graphics and Signal Processing 4, no. 11 (2012), pp. 47-51.
[2]. Mishra, Priyanka, and Rakesh Ahuja. "Review of Secured, Robust and Imperceptible Digital Watermarking Scheme." In 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO), pp. 1-5. IEEE, 2021.
[3]. Yu, Xiaoyan, Chengyou Wang, and Xiao Zhou. "A survey on robust video watermarking algorithms for copyright protection." Applied Sciences 8, no. 10 (2018) pp. 1891.
[4]. Parah, Shabir A., Javaid A. Sheikh, Umer I. Assad, and Ghulam M. Bhat. "Realisation and robustness evaluation of a blind spatial domain watermarking technique." International journal of electronics 104, no. 4 (2017): 659-672.
[5]. Singh, Amit Kumar, et al. "A novel technique for digital image watermarking in spatial domain." Parallel Distributed and Grid Computing (PDGC), 2012 2nd IEEE International Conference on. IEEE, 2012, pp. 424-429.
[6]. Ahuja, Rakesh, Amandeep Kaur, Vikas Lamba, Vinay Kukreja, Ambuj Agarwal, and Manish Sharma. "Securing copyright of digital video by exploiting quantized dc coefficients." In 2021 6th International Conference on Signal Processing, Computing and Control (ISPCC), pp. 423-427. IEEE, 2021.
[7]. Kaarna, Arto, and Pekka Toivanen. "Digital watermarking of spectral images in PCA/wavelet-transform domain." IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No. 03CH37477). Vol. 6. IEEE, 2003.
[8]. Nin, Jordi, and Sergio Ricciardi. "Digital watermarking techniques and security issues in the information and communication society." In 2013 27th International Conference on Advanced Information Networking and Applications Workshops, pp. 1553-1558. IEEE, 2013.
[9]. Yavuz, Erkan, and Ziya Telatar. "Improved SVD-DWT based digital image watermarking against watermark ambiguity." In Proceedings of the 2007 ACM symposium on Applied computing, pp. 1051-1055. 2007.
[10]. Abraham, Jobin, and Varghese Paul. "An imperceptible spatial domain color image watermarking scheme." Journal of King Saud University-Computer and Information Sciences 31, no. 1 (2019),pp.125-133.
[11]. Ahuja, Rakesh, and S. S. Bedi. "Robust Video Watermarking Scheme Based on Intra-Coding Process in MPEG-2 Style." International Journal of Electrical & Computer Engineering 7, no. 6 (2017), pp. 2088-8708.
[12]. Agarwal, Himanshu, Farooq Husain, and Praveen Saini. "Next generation noise and affine invariant video watermarking scheme using Harris feature extraction." In Advances in Computing and Data Sciences: Third International Conference, ICACDS 2019, Ghaziabad, India, April 12–13, 2019, Revised Selected Papers, Part II 3, pp. 655-665. Springer Singapore, 2019.
[13]. Agarwal, Himanshu, and Farooq Husain. "Protecting Ownership Rights of Videos Against Digital Piracy: An Efficient Digital Watermarking Scheme." International Journal of Communication Networks and Information Security 13, no. 2 (2021): 290-301.
[14]. Agarwal, Himanshu, and Farooq Husain. "Development of payload capacity enhanced robust video watermarking scheme based on symmetry of circle using lifting wavelet transform and SURF." Journal of Information Security and Applications 59 (2021): 102846.
[15]. Dang, Shuwen, Weifeng Tian, and Feng Qian. "EMD-and LWT-based stochastic noise eliminating method for fiber optic gyro." Measurement 44, no. 10 (2011): 2190-2193.

DOI: 10.5281/zenodo.10441424

[16]. Singh, Amit Kumar. "Robust and distortion control dual watermarking in LWT domain using DCT and error correction code for color medical image." Multimedia Tools and Applications 78 (2019): 30523-30533.

[17]. Burger, Wilhelm, and Mark J. Burge. "Scale-invariant feature transform (SIFT)." In Digital Image Processing: An Algorithmic Introduction, pp. 709-763. Cham: Springer International Publishing, 2022.

[18]. Li, Yanshan, Weiming Liu, Xiaotang Li, Qinghua Huang, and Xuelong Li. "GA-SIFT: A new scale invariant feature transform for multispectral image using geometric algebra." Information Sciences 281 (2014): 559-572.

[19]. Gaurav Bhatnagar, Q. M. Jonathan Wu, Balasubramanian Raman, ―Robust gray-scale logo watermarking in wavelet domain‖, Computers and Electrical Engineering 3 (2012): pp. 1164-1176.

[20]. Mettripun, Narong, Thumrongrat Amornraksa, and Edward J. Delp. "Robust image watermarking based on luminance modification." Journal of Electronic Imaging 22, no. 3 (2013): 033009

[21]. Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). "Image quality assessment: from error visibility to structural similarity." IEEE transactions on image processing, 13(4), pp. 600-612.

[22]. Shukla, D., & Sharma, M. "A novel scene-based video watermarking scheme for Copyright Protection." Journal of Intelligent Systems, 27, no.1, (2018), pp.47-66.

## AUTHORS

**Dr. Himanshu Agarwal** did his B.Tech in Computer Engineering & Information Technology from UPTU, Lucknow in 2004. He received his M. Tech degree in Computer Science and Engineering from Teerthanker Mahaveer University (Uttar Pradesh) in 2012 and received his Ph.D degree from Dr. A P J Abdul Kalam Technical university, Lucknow in 2023. Presently he is working as Assistant Professor in Moradabad Institute of Technology, Moradabad from March-2005 onwards. His areas of interest are Digital Image processing, Information Security and Digital Watermarking. He has published more than fifteen papers in international journals and conferences including publications from M.Tech and Ph.D. research work.

**Dr. Shweta Agarwal** Ph.D in Mathematics from MJPRU Bareilly. Presently she is working as an Assistant Professor in the Department of Mathematics, Ramabai Ambedkar Govt. Degree College, Amroha-Gajraula (U.P.). Her areas of interest are Digital Image processing, Information Security and Fluid Dynamics. She has published more than twelve research papers in international journals and conferences including publications.

**Dr. Rakesh Ahuja** Ph.D in Computer Science and Engineering from IFTM University, Moradabad. Presently he is working as Professor in Institute of Engineering and Technology, Chitkara University, (Punjab). His areas of interest are Digital Image processing and Information Security. He has published more than twenty research papers in international journals and conferences including publications.

DOI: 10.5281/zenodo.10441424