A Hybrid Video Watermarking Technique based on DWT, SVD and SCHUR Decomposition

Ananya Adhikari
Department of Computer Science & Engineering
Maulana Abul Kalam Azad University of Technology, West Bengal
Salt Lake, Kolkata-700064, India

Mihir Sin
Department of Computer Science & Engineering
Maulana Abul Kalam Azad University of Technology, West Bengal
Salt Lake, Kolkata-700064, India

ABSTRACT
Video watermarking is a technique that used to protect the multimedia data, video in reference to authentication, proof of ownership, copy control etc. In this paper a robust video watermarking scheme is proposed that will protect videos from being unauthorized manipulation by the intruder. A hybrid transform technique consists of Discrete Wavelet Transform (DWT), Schur decomposition, Singular Value Decomposition (SVD) and a visual attention region determination scheme is used to determine the region of interest and region of non interest blocks of video frames. To make the scheme robust, watermark is embedded in the region of non interest blocks of HL band of the frame. The robustness against various attacks has been compared with high Normalized Cross Correlation (NCC) values and the imperceptibility of the watermark image with the original cover image has been compared with indicated high achievable Peak Signal to Noise Ratio (PSNR) values. The experimental results show that the proposed method has achieved a satisfactory performance against various notable attacks.

General Terms
Video watermarking

Keywords
Discrete wavelet Transform (DWT), Schur Decomposition, Singular Value Decomposition (SVD), Visual Attention Region.

1. INTRODUCTION
Digital media are nowadays being blow-out broadly throughout the Internet. Text, images, audio and videos are the digital media data that can be effortlessly replicated, fabricated or tampered by dishonest users. So it becomes necessary of ownership protection of these data in the digital world. Information hiding is a technique that is used to hide secret information into the original media to protect property rights, for authentication or to transmit secret information through the digital data like image, video etc. over internet. Watermarking is very effective scheme in information hiding field for ownership protection of these digital data [7].

Generally image watermarking is very popular among different watermarking scheme in watermarking research field but due to advancement of technology currently researchers are focusing on video watermarking. Nowadays researchers are using various methodology to make video watermarking robust against various attacks. Ling et al. [1] propose a video watermarking scheme which is mainly focused on geometrical distortion attack. They have used Harris-Affine interest point detector. The robustness of the scheme mainly depends on a precise detection of interest points. Preda et al. [2] propose a video watermarking methodology which is robust against some signal processing attacks. They have used Human Visual System (HVS) to adapt the watermarking scheme in DWT domain. A. Kirthika et al. [3] propose a real-time video watermarking robust against several attacks. They used Region of Interest (ROI) and Region of non Interest (RONI) to adapt the watermark embedding scheme in DWT domain. They have used two different watermarks to embed ROI and RONI zone. L. Agilandeeswari et al [7] propose a robust color video watermarking scheme which is based on hybrid embedding techniques such as DWT, SVD, and Contourlet Transform (CT). To achieve the robustness they used Arnold transformation to scramble the sliced color watermark image whereas slicing is done by using the bit plane slicing mechanism. By generating the eigen vector from color watermark image they embed the it in HL and LH band of host image. Agilandeeswari L et al [8] propose a highly efficient and robust video watermarking scheme which based on Hilbert transform and embedding is done in the Integer Wavelet Transform (IWT) domain. In this scheme the Hilbert coefficients of gray watermark image is embed to the selected blocks which is done using Principle Component Analysis (PCA) technique. Nasrin M. Makbol et al [9] propose a robust image watermarking scheme which based on Integer Wavelet Transform and Singular Value Decomposition (SVD). The embedding process is done by directly embedding the gray watermark pixel value into singular values of 1-level IWT decomposed sub bands. In this paper we have proposed a video watermarking scheme which is based on hybrid transform and visual attention region and it is robust against various image and video processing attacks. The paper is arranged in this way, section 2 will discuss about the methodologies we have used, section 3 will discuss about the proposed method, section 4 will discuss about experiment and the evolution and at the end section 5 will give the conclusion.

2. METHODOLOGIES
This section details the methodology I have used in my proposed work and also elaborates my proposed work in form of flow charts and algorithms. I have used three transformation techniques such as Discrete Wavelet Transform (DWT), Schur Decomposition, Singular Value Decomposition (SVD) and Visual Attention Region determination scheme.

2.1 Discrete Wavelet Transform (DWT)
The Discrete Wavelet Transform is one of the ways that can separate the smooth variation and details of an image. It is based on sub-band coding and it is the multi resolution description of an image [10]. It decomposes the input image
into four sub-bands like \( LL \) (low-low), \( HL \) (high-low), \( LH \) (low-high) and \( HH \) (high-high) are shown in Figure 1 (a).

The \( LL \) is achieved by low-pass filtering of both the rows and columns and it contains most of the information of the input image. The \( HH \) is achieved by high-pass filtering in both directions and it contains the edge and the texture of the input image. The \( HL \) and \( LH \) are achieved by low-pass filtering on one direction and high-pass filtering in the other direction. \( LH \) contains the vertical detail information to horizontal edges and \( HL \) contains the horizontal detail information to vertical edges. The \( LL \) can be decomposed again in the same way and the process can repeat up to any level and produce more sub-bands [11].

\[
d_{ij}(\delta, \sigma) = \frac{1}{\sigma_i} v_i^T u_j \delta_{ij} \leq 1
\]

\[
v_i^T v_j \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}
\]

And \( V_{x \times n} \) is an orthogonal matrix called right singular value of \( A \).

\[
V = [v_1, v_2, v_3, ..., v_r, v_{r+1}, ..., v_n] \quad \text{column vectors } v_i \quad \text{for } i = 1, 2, ..., n \text{ from an orthogonal set:}
\]

\[
v_i^T v_j \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}
\]

Here, \( S_{m \times n} \) is a diagonal matrix with singular values on the diagonal. The matrix \( S \) can be shown as following

\[
S = \text{diag}(\sigma_1, \sigma_2, ..., \sigma) \quad \text{for } i = 1, 2, ..., n \text{ where } \sigma_i \text{ are called Singular Value of matrix } A [10]. \text{ It can be proved that }
\]

\[
\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \sigma_4 \geq \sigma_r \geq 0 \quad \text{and} \quad \sigma_r = \sigma_{r+1} = \cdots = \sigma_n = 0
\]

\[2.3 \text{ Visual Attention Region}\]

Visual attention is a significant concern in some video/image applications, like video object discovery, video surveillance, and video retargeting. Visual attention analysis of a video excites human visual system (HVS) by spontaneously generating saliency maps of video frames and attentions on regions of interest (ROIs) in what human perceive.

In video watermarking strategies, the detection of the visual attention regions permits us to find out an acceptable watermarking energy that is resilient enough to persist transcoding attacks, when continuing a minimum perceptual distortion. The visual attention is based on Information Maximization (AIM) [15]. In Figure 3 visual attention region has shown which we achieved by employing the AIM-based method to the frame of “Container” video. After getting the visual attention regions, the remain portion of image will be designated as Region of No Interest (RONI) blocks [6].

\[2.4 \text{ SCHUR Decomposition}\]

The Schur decomposition is an important mathematical tool that is used in metrics analysis. Contrast to SVD the computational intricacy of Schur decomposition is not as much of and consequently it is functional for real time applications. Schur decomposition can be of use to several real matrixes. There are two kinds of this decomposition [16]:

(i) The complex Schur decomposition and

\[
\begin{align*}
I = & U S V^T \\
& m \times n \quad m \times n \quad m \times n \quad n \times n
\end{align*}
\]

Fig 1: Row-column computation of 2-D DWT (a) First level decomposition (b) Second level Decomposition

Fig 2: Illustration of Factoring A to USVT in Singular Value Decomposition

Fig 3: (a) Original video frame, (b) Visual attention region and RONI are defined as the opposite region to visual attention denoted by black portion.
(ii) The real Schur decomposition

In complex Schur decomposition: $f = UTU'$

where, $U$ is a unitary matrix, $U'$ is the conjugate transpose of $U$ and $T$ is an upper triangular matrix entitled the complex Schur type which has the eigen values off by the side of its diagonal.

In Real version of Schur Decomposition: $f = VRV'$

where $f$, $V$, $R$ and $V'$ are matrices that enclose just real information. Herein, $V$ is an orthogonal matrix, $V'$ is the transpose of $V$, and $R$ is a block upper triangular labelled the real Schur type. Here real Schur decomposition will be used because complex matrices needed double storage space. Schur decomposition leads to about $\frac{2}{3}N^3$ flops. SVD computations involve $11N^3$ flops. Eigen values in the Schur decomposition are as well extremely stable. If a square matrix $T$, with linearly independent Eigen vectors has full rank later there subsist an invertible matrix $Q$, such that $Q'TQ = \Lambda$, where $\Lambda$ is a diagonal matrix. However the entire square matrices are not diagonalizable. This shortage can be eased by Schur factorization or QR factorization. According to Schur each invertible matrix possibly will be expressed as a product of unitary matrix $U$ and upper triangle matrix $T$, $T = U_q \times I \times U$. In matrix $T$, the diagonal entries of $T$ are the Eigen values of $I$. This decomposition refers to the structure of the matrix $[4]$.

3. PROPOSED METHOD

The proposed method is based on hybrid transformation techniques consists of DWT, SVD and SCHUR techniques. To make the robustness against different attacks we have used visual attention region and embedding is done on $HL_2$ band on host image.

3.1 Embedding Process

Here, the proposed watermarking embedding process given in detail.

1. First divide the original video into frames.
2. Get the luminance space of the original video frame.
3. Now luminance space is segmented into 8x8 non-overlapping blocks and then applies 2-level 2D-DWT transform on $HL$ band to each block. This operation generates seven DWT sub-bands $[LL_1, LL_2, HL_2, LH_2, HH_2, LH_1$ and $HH_1]$.
4. Calculate the visual attention region by performing the process described above. Each k-th block of the n-th video frame is classified as Region of non Interest (RONI) or Region of Interest (ROI) block.
5. Apply the Schur operator on each 8x8 RONI block. It decomposes the sub-band's coefficient matrix into two independent matrices:

$$HL_2 = UHL_2SHL_2$$

6. Rescale watermark image such a way that the size of the watermark will equal to the size of the $HL_2$ sub-band of host image.
7. Apply 1-level DWT transform on $HL$ band to each 8x8 block. This operation generates four DWT sub-bands $[LL, HL, LH$ and $HH]$.
8. Apply SVD transform on $HL$ band to each 8x8 blocks of watermark and is decomposed into $U_w, V_w, \sigma_w$ then $\sigma$ (watermarked) is defined in the below equation.

$$\sigma = \sigma_v + (f \cdot \sigma_w)/S_{max}$$

where, $S_{max}$ is the largest singular value of watermark, $\sigma_v$ are the upper triangular eigen values obtained when schur transform is applied on the selected RONI blocks of each frame, $\sigma_w$ are the singular values obtained when svd is applied on watermark and $f$ is the scaling factor, chosen to increase the robustness of the method.
9. Inverse Schur operator has been applied on the modified $S HL_2$ matrix to get a modified coefficient matrix $HL_2'$ as follows:

$$UHL_2' \times S'HL_2' \times UHL_2'$$

10. Apply 2-level 2D IDWT on the modified coefficient matrix $HL_2'$. This operation produces the final watermarked video frame.
11. Convert the watermarked video frames from YUV to RGB color matrix.
12. Reconstruct the watermarked video frame into watermarked video.

3.2 Extraction Process

Here, the proposed watermarking extraction process given in detail.

1. Take watermarked video as input and divide the watermarked video into frames.
2. Get the luminance space of the watermarked video frame.
3. Now luminance space is segmented into non-overlapping blocks of size 8x8 pixels and then applies 2-level 2D-DWT transform on $HL$ band to each block. This operation generates seven DWT sub-bands $[wLL_1, wLL_2, wHL_2, wLH_2, wHH_2, wLH_1$ and $wHH_1]$.
4. Compute the visual attention region by executing the procedure defined above. Each k-th block of the n-th video frame is categorized as region of non interest (RONI) or region of interest (ROI) blocks.
5. Apply the Schur operator on each 8x8 RONI block. It decomposes the sub-band's coefficient matrix into two independent matrices:

$$WHL_2 = UwHL_2 S_{wHL_2}$$

the diagonal entries of $S_{wHL_2}$ are upper triangular eigen values $\sigma_v$
6. Extract singular values of watermark $\sigma_{ext}$ from the diagonal matrix $S_{wHL_2}$ as follows:

$$\sigma_{ext} = S_{max} \times (\sigma_v - \sigma_o)/f$$

7. To search out $HL$ coefficient matrix apply inverse Singular value Decomposition which is defined as

$$\text{watermark}_{HL} = U_w \times \sigma_{ext} \times V_w$$

8. Apply 1-level 2D IDWT on $\text{watermark}_{HL}$ to get the luminance of extracted watermark.
9. Convert the watermarked from YUV to RGB color matrix.

4. EXPERIMENTS AND EVALUATION

The implementation was carried out using Matlab R2015a on test videos and watermark of different sizes and formats collected from a Google database.
The performance of the proposed scheme has been measured in reference to imperceptibility and robustness. Experiment has been taken place using three sample cover videos named ‘container.avi’, ‘suzie.avi’, ‘foreman.avi’ and the a color watermark image which are depicted in Figure 4.

The watermarked videos and extracted watermark image are depicted in Figure 5.

\[
\text{NCC} = \frac{\sum_{i} \sum_{j} w(i,j) w'(i,j)}{\sqrt{\sum_{i} \sum_{j} [w(i,j)]^2}}
\]

BER: If the value of BER is 0 or close to then it will be called that there is no error in the extracted watermark and if the value is 1 or close to then there is error in the extracted watermark. The BER is calculated using the equation as,

\[
BER = \frac{\text{Wrongly Extracted watermark bits}}{\text{Total number of watermark bits embedded}}
\]

4.1 Visual perception or imperceptibility or transparency

The metrics Structural Similarity Index Measure (SSIM) and Peak Signal-to-Noise Ratio (PSNR) are used to calculate the degradation of the quality of frames that affected by several attacks. The minimum value of PSNR is 20dB considered as the acceptable value and the acceptable value of SSIM varies between 1 e.g. exact match and 0 e.g. no match.

\[
\text{PSNR} = 10 \log_{10} \left( \frac{225}{\text{MSE}} \right)
\]

Where, the Mean Square Error (MSE) is defined as

\[
\text{MSE} = \frac{1}{M \times N} \sum_{m=1}^{M} \sum_{n=1}^{N} (I(m,n) - I_w(m,n))^2
\]

Where \( I(m,n), I_w(m,n) \) are Original and watermarked images correspondingly.

\[
\text{SSIM}(w,w') = \frac{\sum_{i} \sum_{j} w(i,j) w'(i,j)}{\sqrt{\sum_{i} \sum_{j} [w(i,j)]^2} \sqrt{\sum_{i} \sum_{j} [w'(i,j)]^2}}
\]

w and \( w' \) is the embedded and extracted watermark correspondingly.

4.2 Robustness

The metrics Normalized Correlation Co-efficient (NCC) and Bit Error Rate (BER) are used to measure the robustness of the embedded watermark against several attacks. The standard value of these metrics varies between 0 and 1.

NCC: In this metric, if the value of NCC is close to 1 then both the watermarks embedded and the extracted will be called alike or correlated and will be called uncorrelated if value of NCC is close to 0. The NCC is calculated by using the equation as,

4.3 Attacks

The reliability and robustness of the implemented method is examined on the experimental videos considering the no attack situation (no attack have been applied), image processing attacks and geometrical attacks situations.

4.4 No attack

The effect of embedding watermark on the experimental cover video can be calculated effectively (in reference to transparency and robustness), considering no attack situation. The result is shown in Table 1.

Table 1 PSNR, NCC and BER values at no attack situation

<table>
<thead>
<tr>
<th>Videos</th>
<th>PSNR</th>
<th>NCC</th>
<th>BER</th>
</tr>
</thead>
<tbody>
<tr>
<td>container</td>
<td>43.88</td>
<td>0.9905</td>
<td>0.1933</td>
</tr>
<tr>
<td>suzie</td>
<td>44.31</td>
<td>0.9918</td>
<td>0.1889</td>
</tr>
<tr>
<td>foreman</td>
<td>46.12</td>
<td>0.9820</td>
<td>0.1857</td>
</tr>
</tbody>
</table>

4.5 Image processing and Geometric attacks

The experimental result of the implemented method has validated by applying various Image processing and Geometric attacks for example, Rotation attack (5°), Cropping attack, Gaussian noise (variance=0.01), Poisson noise, Salt and Pepper noise (density= 0.02), Median filtering of (3×3), Contrast adjustment, Average filtering, Motion Attack, Speckle Noise, Gaussian blur, Histogram attack, Median attack, frame averaging, Resize attack and Sharpening attack. Figure 6, 7, 8 has been depicted the attacked watermarked
video frames (considering the 6th frame of the sample videos and each video is consist of 150 numbers of frames) and the extracted watermark for mentioned attacks and we also have calculated the PSNR, SSIM, NCC and BER values.

<table>
<thead>
<tr>
<th>Video Name</th>
<th>Attack Name</th>
<th>Watermarked Frame</th>
<th>Extracted Watermarked Frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>container.avi</td>
<td>Rotation (5°)</td>
<td><img src="container_rotation.png" alt="Image" /></td>
<td><img src="container_rotation_extracted.png" alt="Image" /></td>
</tr>
<tr>
<td></td>
<td>PSNR: 30.615 SSIM: 0.5540 NCC: 0.9776 BER: 0.1900</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cropping</td>
<td></td>
<td><img src="container_cropping.png" alt="Image" /></td>
<td><img src="container_cropping_extracted.png" alt="Image" /></td>
</tr>
<tr>
<td></td>
<td>PSNR: 28.890 SSIM: 0.5445 NCC: 0.9753 BER: 0.1910</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frame Averaging</td>
<td></td>
<td><img src="container_frame_averaging.png" alt="Image" /></td>
<td><img src="container_frame_averaging_extracted.png" alt="Image" /></td>
</tr>
<tr>
<td></td>
<td>PSNR: 36.483 SSIM: 0.9004 NCC: 0.9779 BER: 0.1911</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salt &amp; Pepper noise (variance=0.01)</td>
<td><img src="container_salt_pepper.png" alt="Image" /></td>
<td><img src="container_salt_pepper_extracted.png" alt="Image" /></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PSNR: 46.875 SSIM: 0.7289 NCC: 0.9814 BER: 0.1887</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gaussian noise (variance=0.01)</td>
<td><img src="container_gaussian.png" alt="Image" /></td>
<td><img src="container_gaussian_extracted.png" alt="Image" /></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PSNR: 28.763 SSIM: 0.2158 NCC: 0.8946 BER: 0.2232</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median Attack</td>
<td></td>
<td><img src="container_median.png" alt="Image" /></td>
<td><img src="container_median_extracted.png" alt="Image" /></td>
</tr>
<tr>
<td></td>
<td>PSNR: 43.485 SSIM: 0.9639 NCC: 0.9834 BER: 0.1899</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contrast Adjustme nt</td>
<td><img src="container_contrast.png" alt="Image" /></td>
<td><img src="container_contrast_extracted.png" alt="Image" /></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PSNR: 36.2698 SSIM: 0.8653 NCC: 0.9694 BER: 0.1918</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poisson Noise</td>
<td></td>
<td><img src="container_poisson.png" alt="Image" /></td>
<td><img src="container_poisson_extracted.png" alt="Image" /></td>
</tr>
<tr>
<td></td>
<td>PSNR: 31.1639 SSIM: 0.4773 NCC: 0.9662 BER: 0.1894</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Video Name</th>
<th>Attack Name</th>
<th>Watermarked Frame</th>
<th>Extracted Watermarked Frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suzie.avi</td>
<td>Rotation (5°)</td>
<td><img src="suzie_rotation.png" alt="Image" /></td>
<td><img src="suzie_rotation_extracted.png" alt="Image" /></td>
</tr>
<tr>
<td></td>
<td>PSNR: 30.4564 SSIM: 0.6891 NCC: 0.9751 BER: 0.1856</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig 6: PSNR, NCC, SSIM and BER values for various image processing and geometric attacks on watermarked video (container.avi) and the extracted watermark.
Fig 7: PSNR, NCC, SSIM, and BER values for various image processing and geometric attacks on watermarked video (suzie.avi) and the extracted watermark.

<table>
<thead>
<tr>
<th>Attack Name</th>
<th>Video Name</th>
<th>Watermarked Frame</th>
<th>Extracted Watermarked</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cropping</td>
<td>Formen.avi</td>
<td><img src="image1.png" alt="image" /></td>
<td><img src="image2.png" alt="image" /></td>
</tr>
<tr>
<td>PSNR: 27.7143, SSIM: 0.6538, NCC: 0.9757, BER: 0.1828</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frame Averaging</td>
<td></td>
<td><img src="image3.png" alt="image" /></td>
<td><img src="image4.png" alt="image" /></td>
</tr>
<tr>
<td>PSNR: 38.4074, SSIM: 0.9189, NCC: 0.9758, BER: 0.1850</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salt &amp; Pepper noise (variance=0.01)</td>
<td></td>
<td><img src="image5.png" alt="image" /></td>
<td><img src="image6.png" alt="image" /></td>
</tr>
<tr>
<td>PSNR: 47.1618, SSIM: 0.7013, NCC: 0.9780, BER: 0.1845</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gaussian noise (variance=0.01)</td>
<td></td>
<td><img src="image7.png" alt="image" /></td>
<td><img src="image8.png" alt="image" /></td>
</tr>
<tr>
<td>PSNR: 28.7528, SSIM: 0.1440, NCC: 0.9151, BER: 0.2100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median Attack</td>
<td></td>
<td><img src="image9.png" alt="image" /></td>
<td><img src="image10.png" alt="image" /></td>
</tr>
<tr>
<td>PSNR: 44.9383, SSIM: 0.9718, NCC: 0.9831, BER: 0.1857</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contrast Adjustment</td>
<td></td>
<td><img src="image11.png" alt="image" /></td>
<td><img src="image12.png" alt="image" /></td>
</tr>
<tr>
<td>PSNR: 29.9882, SSIM: 0.8554, NCC: 0.9685, BER: 0.1815</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poisson Noise</td>
<td></td>
<td><img src="image13.png" alt="image" /></td>
<td><img src="image14.png" alt="image" /></td>
</tr>
<tr>
<td>PSNR: 31.8187, SSIM: 0.4713, NCC: 0.9776, BER: 0.1818</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Histogram Equalization</td>
<td></td>
<td><img src="image15.png" alt="image" /></td>
<td><img src="image16.png" alt="image" /></td>
</tr>
<tr>
<td>PSNR: 30.4969, SSIM: 0.7857, NCC: 0.9046, BER: 0.2126</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gaussian blur</td>
<td></td>
<td><img src="image17.png" alt="image" /></td>
<td><img src="image18.png" alt="image" /></td>
</tr>
<tr>
<td>PSNR: 42.7117, SSIM: 0.9056, NCC: 0.9765, BER: 0.1850</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speckle Noise</td>
<td></td>
<td><img src="image19.png" alt="image" /></td>
<td><img src="image20.png" alt="image" /></td>
</tr>
<tr>
<td>PSNR: 30.8472, SSIM: 0.4536, NCC: 0.9783, BER: 0.1790</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motion Attack</td>
<td></td>
<td><img src="image21.png" alt="image" /></td>
<td><img src="image22.png" alt="image" /></td>
</tr>
<tr>
<td>PSNR: 44.6106, SSIM: 0.9628, NCC: 0.9823, BER: 0.1843</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resize Attack</td>
<td></td>
<td><img src="image23.png" alt="image" /></td>
<td><img src="image24.png" alt="image" /></td>
</tr>
<tr>
<td>PSNR: 42.0104, SSIM: 0.9327, NCC: 0.9761, BER: 0.1856</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sharpening</td>
<td></td>
<td><img src="image25.png" alt="image" /></td>
<td><img src="image26.png" alt="image" /></td>
</tr>
<tr>
<td>PSNR: 47.7404, SSIM: 0.9843, NCC: 0.9794, BER: 0.1811</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Frame Averaging
PSNR :38.1108
SSIM : 0.9384
NCC : 0.9767
BER : 0.1858

Salt &Pepper noise (variance= 0.01)
PSNR :47.2817
SSIM : 0.7382
NCC : 0.9733
BER : 0.1829

Gaussian noise ( variance= 0.01)
PSNR :28.7501
SSIM : 0.1787
NCC : 0.9251
BER : 0.2024

Median Attack
PSNR :46.2812
SSIM : 0.9799
NCC : 0.9795
BER : 0.1844

Contrast Adjustment
PSNR :33.7234
SSIM : 0.8698
NCC : 0.9704
BER : 0.1871

Poisson Noise
PSNR :30.9348
SSIM : 0.4274
NCC : 0.9663
BER : 0.1824

Histogram Equalization
PSNR :28.0853
SSIM : 0.8226
NCC : 0.9290
BER : 0.2007

Gaussian blur
PSNR :43.3005
SSIM : 0.9217
NCC : 0.9770
BER : 0.1847

Speckle Noise
PSNR :29.4974
SSIM : 0.9305
NCC : 0.9493
BER : 0.1916

Motion Attack
PSNR :43.9834
SSIM : 0.9750
NCC : 0.9796
BER : 0.1847

Resize Attack
PSNR :41.3019
SSIM : 0.9460
NCC : 0.9772
BER : 0.1851

Sharpening
PSNR :47.6305
SSIM : 0.9885
NCC : 0.9741
BER : 0.1841

Fig 8: PSNR, NCC, SSIM and BER values for various image processing and geometric attack on watermarked video (foreman.avi) and the extracted watermark

4.6 Comparison between the implemented methods and the existing method
The effectiveness of the implemented method is vindicated by executing the existing methods (i) DWT, SVD, and Contourlet Transform (CT) [7], (ii) Hilbert transform and Integer Wavelet Transform (IWT) [8] and (iii) IWT-SVD [9] with the experimental dataset and the same is examine with the above introduced attacks. Figure 9 (a) and (b) depicts the comparison plots between the existing methods and the implemented method. We have experienced the robustness and imperceptibility of the implemented scheme using mainly image processing and geometric attacks, and also the same has been examined with the existing methods.

The outcomes and discussions presented here mainly focusing on the measurements of robustness of the existing methods and implemented method considering several attack situations. PSNR and NCC are used as the quality matrices here.

Table 2 Relative PSNR Values with Earlier Methodologies

<table>
<thead>
<tr>
<th>Attacks applied</th>
<th>Proposed scheme</th>
<th>[8]</th>
<th>[7]</th>
<th>[9]</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Attack</td>
<td>43.9804</td>
<td>47.78</td>
<td>65</td>
<td>43.6769</td>
</tr>
<tr>
<td>Rotation(5°)</td>
<td>30.6135</td>
<td>33.8165</td>
<td>54.0534</td>
<td>9.16300</td>
</tr>
<tr>
<td>Cropping</td>
<td>28.8960</td>
<td>39.3374</td>
<td>-</td>
<td>40.9067</td>
</tr>
<tr>
<td>Frame</td>
<td>36.4863</td>
<td>44.1762</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Table 3 Relative NCC Values with Earlier Methodologies

<table>
<thead>
<tr>
<th>Attacks applied</th>
<th>Proposed scheme</th>
<th>[8]</th>
<th>[7]</th>
<th>[9]</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Attack</td>
<td>0.9905</td>
<td>-</td>
<td>0.9999</td>
<td>-</td>
</tr>
<tr>
<td>Rotation (5°)</td>
<td>0.9748</td>
<td>0.9110</td>
<td>-</td>
<td>0.9061</td>
</tr>
<tr>
<td>Cropping</td>
<td>0.97</td>
<td>0.8610</td>
<td>-</td>
<td>0.9770</td>
</tr>
<tr>
<td>Frame Averaging</td>
<td>0.97</td>
<td>0.9912</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Salt &amp; Pepper noise (variance=0.01)</td>
<td>0.98</td>
<td>0.8750</td>
<td>0.98</td>
<td>0.8325</td>
</tr>
<tr>
<td>Gaussian noise</td>
<td>0.90</td>
<td>0.8677</td>
<td>0.89</td>
<td>0.8318</td>
</tr>
<tr>
<td>Median Attack</td>
<td>0.97</td>
<td>0.5117</td>
<td>0.75</td>
<td>0.9356</td>
</tr>
<tr>
<td>Contrast Adjustment</td>
<td>0.98</td>
<td>0.9898</td>
<td>0.92</td>
<td>0.9760</td>
</tr>
<tr>
<td>Poisson Noise</td>
<td>0.95</td>
<td>0.5125</td>
<td>0.72</td>
<td>-</td>
</tr>
<tr>
<td>Histogram Equalization</td>
<td>0.89</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gaussian blur</td>
<td>0.97</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Speckle Noise</td>
<td>0.97</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Motion Attack</td>
<td>0.97</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Resize Attack</td>
<td>0.97</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sharpening</td>
<td>0.98</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 3 depicts the NCC value of the experimental video container.avi with experimental watermark with having no attack and with different image processing and geometric attacks. The table depicts that at the presence of no attack almost 99% of the watermark extraction can be possible. From the Figure 9 (b) and Table 3, we conclude that the method is capable of extraction of 90% and above of watermark when the frames are attacked by Gaussian noise, Poisson noise, Salt and Pepper noise, Contrast Adjustment, Motion attack, and Histogram Equalization attacks respectively. The Figure 9 (a) and Table 2 depicts that the implemented scheme is greatly imperceptible at salt and pepper attack with PSNR value 46.87 after comparing with image processing attack. After observing the above results, we conclude that the implemented watermarking method is able to survive image processing attacks in an improved way except Frame averaging after comparing the existing watermarking methods in reference to imperceptibility level.

5. CONCLUSION

In this paper a hybrid transformation technique and visual attention region determination schemes are used in video watermarking. The detection of visual attention region helps to persist various attacks and possess a least perceptual distortion. The simulation result shows that the proposed scheme is robust against some image processing and geometric attacks compare to the existing algorithms such as Hilbert transform and Integer Wavelet Transform (IWT) [8], DWT, SVD, and Contourlet Transform (CT) [7] and IWT-SVD [9]. The watermarked frame quality is also good in terms of imperceptibility. In this work $H_L$ band is chosen to embed watermark which gives high robustness against geometric and filtering attacks. The efficiency of the scheme is improved by using Schur decomposition as least number of calculations required here.

Performance of the proposed scheme on transcoded video is not satisfactory. So it can be further improved.

6. REFERENCES


