



# Optimal keyframe selection-based lossless video-watermarking technique using IGSA in LWT domain for copyright protection

Roop Singh<sup>1</sup> · Himanshu Mittal<sup>2</sup> · Raju Pal<sup>2</sup>

Received: 19 May 2021 / Accepted: 19 October 2021 / Published online: 13 November 2021  
© The Author(s) 2021

## Abstract

Video piracy is a challenging issue in the modern world. Approximately 90% of newly released films were illegally distributed around the world via the Internet. To overcome this issue, video watermarking is an effective process that integrates a logo in video frames as a watermark. Therefore, this paper presents an efficient lossless video-watermarking scheme based on optimal keyframe selection using an intelligent gravitational search algorithm in linear wavelet transform. This technique obtains color motion and motionless frames from the cover video by the histogram difference method. One-level linear wavelet transform is performed on the chrominance channel of motion frames and a low-frequency sub-band LL opts for watermark embedding. The performance of the proposed technique has been evaluated against 12 video processing attacks in terms of imperceptibility and robustness. Experiments demonstrate that the proposed technique outperforms five state-of-the-art schemes on the considered attacks.

**Keywords** Video watermarking · Linear wavelet transform · Intelligent gravitational search algorithm

## Introduction

In recent years, the unauthorized users can easily access the digital media content (image, audio, and video). These content are illegally copied, manipulated, and distributed [1,2] across the globe over the internet. Around 90% newly released movies are illegally recorded by the camcorder device and distributed via the internet globally [3]. The Dark Knight movie is one example of video piracy whose 7 million copies were delivered illegally in 6 months after its release [4]. As a result, copyright and content security are the most prevalent issues in the modern world. To address the same, video watermarking is one of the promising solutions which

embeds information into video frames for copyright protection [5].

Many watermarking algorithms have been proposed over the last decade and have been classified based on specific characteristics. In general, video-watermarking methods operate in two domains, namely spatial and frequency. In the spatial domain, a watermark is directly embedded by modifying the pixels. Approaches of this domain have low computational complexity. However, these methods have poor data-hiding ability and low robustness [6,7]. On the contrary, frequency domain methods are pretty much efficient in this respect. In the frequency domain, the watermark is integrated with the wavelet coefficients of video frame. Some of the popular frequency domain methods are discrete Fourier transform (DFT), singular value decomposition (SVD), discrete cosine transform (DCT), and discrete wavelet transform (DWT). However, DWT is often employed due to its multi-resolution capabilities. This domain methods accomplish high payload capacity, imperceptibility, robustness, and more secure [6]. Further, This domain methods are computationally expensive as compared to spatial domain methods.

Generally, researchers focus on the three parameters of a video watermarking method, i.e., imperceptibility, robustness, and payload capacity [8]. In video watermarking, imperceptibility corresponds to the effectiveness of a water-

✉ Himanshu Mittal  
himanshu.mittal224@gmail.com

Roop Singh  
roopsolanki@gmail.com

Raju Pal  
raju3131.pal@gmail.com

<sup>1</sup> Department of Electronics and Communication Engineering, Uttarakhand Technical University, Dehradun, Uttarakhand, India

<sup>2</sup> Department of Computer Science and Engineering, Jaypee Institute of Information Technology, Noida 201304, India

marking method in concealing a watermark into the cover video frames, while robustness measures the efficiency of a watermarking process in recovering the watermark from video frames after video attacks. Moreover, payload capacity refers to the number of bits contained within each frame. The complexity of a watermarking method often increases with the increased payload capacity. In the literature, it has been witnessed that these three parameters are contradictory and confined with each other [6]. Therefore, it is a difficult task to develop an effective watermarking system that is efficient in terms of these parameters along with low computational complexity.

In the literature, many video-watermarking schemes are presented to achieve the objectives mentioned above. Table 1 lists some of the popular video-watermarking schemes. It can be observed that most schemes perform watermark embedding in DWT, DCT, and SVD domains [9–15]. Generally, such schemes are computationally complex [28] and unable to recover the lossless watermark image due to shift-variant property [29]. Moreover, few schemes are based on linear wavelet transform (LWT) [17–19]. The LWT-based video-watermarking schemes are resistant to both image processing attacks and temporal video attacks such as impulse noise, Gaussian noise, filtering, compression attacks, but fails against geometric attacks [25–28,30]. Further, it can be envisioned in Table 1 that existing video-watermarking schemes incorporate watermark in the non-motion frame of luminance components which results in poor imperceptibility [31]. Moreover, most video watermarking schemes employ single scaling factors (SEF) approach that significantly affects the balance between imperceptibility and robustness. To attain a better equilibrium between imperceptibility and robustness, the integration of video-watermarking schemes with multiple scaling factors (MSF) is a promising solution [20–22,24,32]. However, the selection of the optimum value of MSF is an NP-complete problem [33] which can be solved by employing meta-heuristic algorithms.

Meta-heuristic algorithms are optimization algorithms that imitate the optimization behavior of natural phenomena [34,35]. Gravitational search algorithm (GSA) [36] is one of such meta-heuristic algorithms inspired by the Newtonian law of gravity. In GSA, the optimal solution is obtained through a collection of objects which co-ordinates with each other according to the law of gravity and law of motion [37]. In comparison to different meta-heuristic algorithms, GSA has a low computational cost and high convergence rate [38]. In addition, GSA has been broadly acknowledged in the literature for multimodal challenges, notably for clustering applications [39]. Moreover, GSA advantages in finding the best solution using the current positions only, and therefore, it is considered as a memory-less algorithm [36,40]. However, due to the lack of demographic diversity and an inappropriate balance between exploration and exploitation, it often

stagnates into local optima [41]. In the literature, researchers have proposed several variants of GSA. Liu et al. [42] suggested dynamically adapting inertia factors for improving position updation. Further, Olivas et al. [43] introduced an interval type-2 fuzzy system-based modified variant of GSA which improves the exploitation and exploration of the search space. The adaptive GSA (AGSA) presented by Mirjalili et al. [44] allows the exploitation of the GSA to be modified based on the current situation. A variation of GSA called an exponential kbest gravitation search algorithm (eKGSA) had been introduced by Mittal et al. [45] for optimal thresholds for multi-level image segmentation. The hierarchical gravitation search algorithm proposed by Wang et al. [46] deals with premature convergence and low search capacity. Rawal et al. [47] presented fast convergent GSA which utilizes a sigmoidal function and an exponential step size to accelerate convergence and exploitation. Recently, Mittal et al. [48] presented a new variant of GSA, intelligent gravitational search algorithm (IGSA), which outperformed GSA regarding convergence rate and solution precision.

Therefore, the key contributions of the paper are twofold, (1) a new video-watermarking technique has been proposed, termed as a lossless video-watermarking technique using intelligent gravitational search algorithm and Hessenberg transform in linear wavelet transform (IGSA-LH) and (2) to attain equilibrium between imperceptibility and robustness, intelligent gravitational search algorithm (IGSA) has been leveraged to acquire an optimal set of multiple scaling factors. For experimental analysis, the proposed technique has been evaluated on four standard benchmark videos against 12 image and video attacks in terms of imperceptibility parameters, namely mean peak signal to noise ratio (MPSNR) and mean structural similarity index (MSSIM) and robustness parameter, i.e., mean normalized correlation (MNC). Further, the obtained results are compared with five existing video watermarking techniques, namely Karmakar et al. [2], Bhardwaj et al. [49], Farri et al. [17], Kuraparthi et al. [24], and Agilandeeswari et al. [50].

The remaining paper is arranged in the following order. The preliminaries for the proposed technique are presented in the next section. The third section describes the proposed technique followed by the experimental findings presented in the fourth section. Finally, the last section draws the conclusion.

## Preliminaries

### Color space conversion from RGB to YUV

The RGB color space's red (R), green (G), and blue (B) components are transformed to the YUV color space's luminance (Y) and chrominance (U, V) components. The luminance

**Table 1** Comparative description of video-watermarking schemes in frequency domain

Author/article	Methodology	Author objectives	Performance evaluation parameter	Application
Choi et al. [9]	(1) DCT (1) DWT (2) PRN (3) Secret key	(1) High robustness (2) High video quality (1) Transparency (2) Robustness	(1) PSNR (2) BER (1) PSNR (2) BER	Copyright protection Copyright protection
Preda et al. [10]				
Yassin et al. [11]	(1) DWT (2) PCA	(1) High imperceptibility (2) Robustness	(1) PSNR (2) NC	–
Faragallah et al. [12]	(1) DWT + SVD	(1) An efficient (2) Robust	(1) PSNR (2) NC	Copyright protection
Masoumi et al. [13]	(1) 3D-DWT (2) Spread spectrum	(3) Imperceptible (1) Transparency (2) Robustness	(3) BER (1) PSNR (2) NC	Copyright protection
Thanh et al. [14]	(1) DCT (2) KAZE feature	(1) High robustness	(1) PSNR (2) SSIM	–
Masoumi et al. [15]	(1) 3D-DWT (2) CDMA (3) PRN	(1) High transparency (2) Robustness	(1) MSE (2) PSNR (3) SSIM (4) MS-SSIM	Ownership verification
Rast et al. [16]	(1) DWT (2) SVD (3) QR decomposition	(1) Imperceptibility (2) Robustness	(1) PSNR (2) CC	Copyright protection
Fari et al. [17]	(1) LWT (2) SVD (3) Arnold transform	(1) Imperceptibility (2) Robustness	(1) PSNR (2) SSIM (3) NCC	Copyright protection
Suresh et al. [18]	(1) LWT (2) GWO (3) Chaotic key	(1) Imperceptibility (2) Robustness	(4) BER (1) PSNR (2) SSIM (3) NC	Copyright protection Copyright protection

**Table 1** continued

Author/article	Methodology	Author objectives	Performance evaluation parameter	Application
Ayubi et al. [19]	(1) DWT (2) LWT (3) SVD	(1) Imperceptibility (2) Robustness	(1) PSNR (2) SSIM (3) NC	—
Aditya et al. [20]	(1) DCT (2) DWT	(1) Imperceptibility (2) Robustness	(1) PSNR (2) NC	—
Agrawal et al. [21]	(1) DWT (2) GA-PSO	(1) Imperceptibility (2) Robustness	(1) PSNR (2) NC	—
Gupta et al. [22]	(1) DWT (2) GSO	(1) Imperceptibility (2) Robustness	(1) PSNR (2) NC	—
Kumar et al. [23]	(1) RDWT-SVD (2) MOO	(1) Imperceptibility (2) Robustness	(1) PSNR (2) SSIM (3) NC	—
Kuraparthi et al. [24]	(1) DWT-SVD (2) ABC	(1) Imperceptibility (2) Robustness	(1) PSNR (2) NC	Copyright protection
Sahu et al. [25]	(1) SIFT (2) DCT	(1) Robustness	(1) PSNR	—
Dutta et al. [26]	(3) Feature points (1) SIFT (2) DCT	(1) Robustness	(1) PSNR	—
Singh et al. [27]	(3) Feature points (1) SIFT (2) DWT (3) Arnold transform	(1) Strong imperceptibility (2) Strong robustness	(1) MPSNR (2) MSSIM (3) NC (4) BER	Copyright protection

components (Y) reflect the majority of the frame information and overall strength, while the chrominance components (U, V) indicate the color information of the frame [31]. The mathematical color conversion formula is given by Eq. (1):

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0 \\ 127 \\ 127 \end{bmatrix} \begin{bmatrix} 0.2989 & 0.5866 & 0.1145 \\ -0.1688 & -0.3312 & 0.5000 \\ 0.5000 & -0.4184 & -0.0816 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

## Linear wavelet transform

DWT is a first-generation wavelet that yields floating-point coefficients. At any time, these coefficients may alter during subsequent processing. As a result, information is lost during watermark embedding due to the truncation of floating-point pixel values. LWT replaces the up and downsampling of DWT with split and merges into each level, resulting in a nonlinear wavelet transform. Because of the split and merge procedure, the computational complexity is nearly half of DWT [51]. To alleviate the same, the linear wavelet transform (LWT) is an extended version of DWT based on second-generation wavelets [52]. LWT splits an image into low-frequency (LL) and high-frequency (LH, HL, HH) subband. In contrast to DWT, LWT transforms the value of an image pixel from integers to integers, resulting in lossless, computationally faster, and reliable execution. There are three key steps, namely split, predict, and update which are discussed below. The complete procedure of LWT is illustrated in Fig. 1.

1. **Splitting:** Consider an image  $F(m, n)$  that is split into two sections, i.e., even ( $F_e(m, n)$ ) and odd ( $F_o(m, n)$ ) which are defined by Eqs. (2) and (3), respectively:

$$F_e(m, n) = F(m, 2n) \quad (2)$$

$$F_o(m, n) = F(m, 2n + 1) \quad (3)$$

2. **Dual lifting (Predict):** Generally, the prediction operator ( $P^*$ ) merges many even parts and applies the resulting value to the actual odd part. The odd part is evaluated using a prediction operator from the local neighborhood's even coefficients. The error in odd part prediction defined

by Eq. (4) yields high-frequency coefficients ( $h(m, n)$ ) and updated value of odd part is given by Eq. (5):

$$h(m, n) = F_o(m, n) - P^*[F_e(m, n)] \quad (4)$$

$$F_o(m, n) = h(m, n) + P^*[F_e(m, n)] \quad (5)$$

3. **Primal lifting (update):** Similarly, the low-frequency coefficients ( $h^*(m, n)$ ) are produced by modifying the even part along with the updating value ( $U_h(m, n)$ ) which is given by Eq. (6). Further,  $U_h(m, n)$  is updated through  $F(m, n)$  using Eq. (7):

$$h^*(m, n) = F_e(m, n) + U_h(m, n) \quad (6)$$

$$U_h(m, n) = U[F(m, n)] \quad (7)$$

## Arnold transform

Arnold transform (AT) is a technique for scrambling images that improve protection and identifies their true owner [52]. The AT is used to scramble a watermark logo before embedding. It shuffles the pixel positions to create the new chaotic image iteratively, resulting in a scrambled image. An unauthorized user cannot retrieve the watermark logo from the watermarked image without knowing the security key. Figure 2 depicts the watermark logo and its scrambled version at different iterations. The figure shows that as the number of iterations increase, the watermark logo scrambles efficiently. The AT is equated to a square watermark logo of dimension  $R_2 \times R_2$  as depicted in Eq. (8):

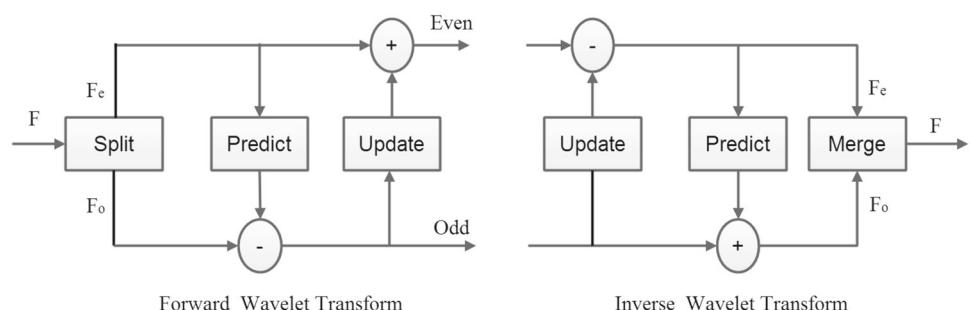
$$\begin{bmatrix} a_0 \\ b_0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \bmod R_2, \quad (8)$$

where  $(a, b)$  represents the watermark pixel values and  $(a_0, b_0)$  represents the scrambled image pixel values at  $i$ th iteration.

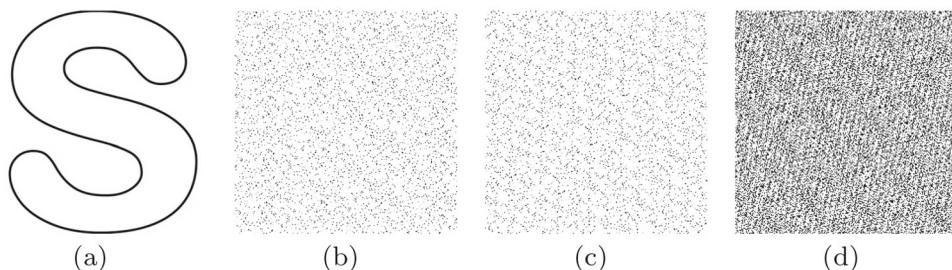
The inverse Arnold transform (IAT) recovers the watermark logo through Eq. (9):

$$\begin{bmatrix} a \\ b \end{bmatrix} = \left( \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + \begin{bmatrix} R_2 \\ R_2 \end{bmatrix} \right) \bmod R_2 \quad (9)$$

**Fig. 1** Linear wavelet transform



**Fig. 2** Arnold transform at different iterations: **a** Watermark logo, **b** itr = 50, **c** itr = 75, **d** itr = 100



## Hessenberg transform

The Hessenberg transform (HT) is a procedure for factorizing a general matrix ( $S$ ) using orthogonal similarity transformations [53,54]. The HT of a matrix ( $S$ ) is given by Eq. (10):

$$\text{HT}[S] = Q \times H \times Q^T, \quad (10)$$

where  $Q$  and  $H$  are orthogonal and Hessenberg matrix, respectively, such that  $h_{ij} = 0$  and  $i > j + 1$ . Usually, Hessenberg transform is performed on household matrices. The Householder matrix ( $P$ ) is an orthogonal matrix defined by Eq. (11):

$$P = (I_r - 2uu^T)/u^Tu, \quad (11)$$

where  $I_r$  and  $u$  are  $m \times m$  identity matrix and nonzero vector in  $R^m$ , respectively. The overall procedure consists of  $m - 2$  steps for a matrix ( $S$ ) of size  $m \times m$ . Therefore,  $H$  is computed by Eq. (12):

$$H = (P_1 P_2 \cdots P_{m-3} P_{m-2})^T \times S (P_1 P_2 \cdots P_{m-3} P_{m-2}) \quad (12)$$

$$\implies H = Q^T \times S \times Q \quad (13)$$

$$\implies S = Q \times H \times Q^T, \quad (14)$$

where  $Q = P_1 P_2 \cdots P_{m-3} P_{m-2}$ .

For example, HT of a matrix ( $S$ ) of size  $4 \times 4$  is computed as follows:

$$S = \begin{bmatrix} 228 & 39 & 208 & 51 \\ 245 & 66 & 63 & 65 \\ 140 & 215 & 237 & 158 \\ 36 & 65 & 90 & 121 \end{bmatrix} \quad (15)$$

After performing HT on a matrix ( $S$ ), matrix ( $Q$ ) and matrix ( $H$ ) are given as follows:

$$Q = \begin{bmatrix} 1.0000 & 0 & 0 & 0 \\ 0 & -0.8613 & 0.5039 & 0.0656 \\ 0 & -0.4921 & -0.7950 & -0.3546 \\ 0 & -0.1266 & -0.3377 & 0.9327 \end{bmatrix} \quad (16)$$

$$H = \begin{bmatrix} 228.0000 & -142.4106 & -162.9320 & -23.6301 \\ -284.4662 & 255.7505 & 109.1450 & -85.5923 \\ 0 & 246.4589 & 113.4365 & -68.8549 \\ 0 & 0 & -29.5251 & 54.8130 \end{bmatrix} \quad (17)$$

## Intelligent gravitational search algorithm (IGSA)

Mittal et al. [48] proposed a variant of the gravitational search algorithm, intelligent gravitational search algorithm (IGSA), to improve the solution precision. IGSA focuses on enhancing the exploitation ability among the objects. To do so, the position equation is modified by including the global best solution ( $gBest$ ) and global worst solution ( $gWorst$ ). In IGSA, the attraction of objects towards  $gBest$  is proportionated with  $gWorst$ . Mathematically, the modified position equation of IGSA for an  $i$ th object in  $d$ th dimension at  $t$ th iteration is depicted in Eq. (18):

$$x_i^d(t+1) = x_i^d(t) + v_i^d(t+1) + r(t) \times \underbrace{\frac{(gBest^d(t) - x_i^d(t))}{|(\rho \times gWorst^d(t)) - x_i^d(t)|}}_{\text{Intelligent component}}, \quad (18)$$

where  $r(t) \in [0, 1]$  and  $\rho$  is a constant whose value is taken as 0.9.

## Proposed technique

The proposed technique, termed as a lossless video-watermarking technique using intelligent gravitational search algorithm and Hessenberg transform in linear wavelet transform (IGSA-LH), is explained in the following four sections: identification

of motion frames and keyframes (“Identification of motion frames and keyframes”), embedding process (“Embedding process”), extraction process (“Extraction process”), and finding optimal scaling factors through IGSA (“Selection of multiple scaling factors using IGSA algorithm”). The overall procedure of the proposed technique is illustrated in Fig. 3.

### Identification of motion frames and keyframes

Motion frames are identified by performing the histogram difference method on cover video [55]. The complete keyframe selection procedure is illustrated in Fig. 4. In this method, the absolute histogram difference of two video frames back to back is calculated. If the absolute histogram difference between consecutive video frames is greater than a predefined threshold, a frame is identified as a motion frame. Similarly, keyframes are detected from the selected motion frames followed by entropy. Suppose, the single motion frame entropy is greater than the average entropy of motion frames then that motion frame is identified as a keyframe. Figure 5 depicts the identified keyframes corresponding to each cover video. A detailed workflow for identifying keyframes are referred to motion frames which is given below:

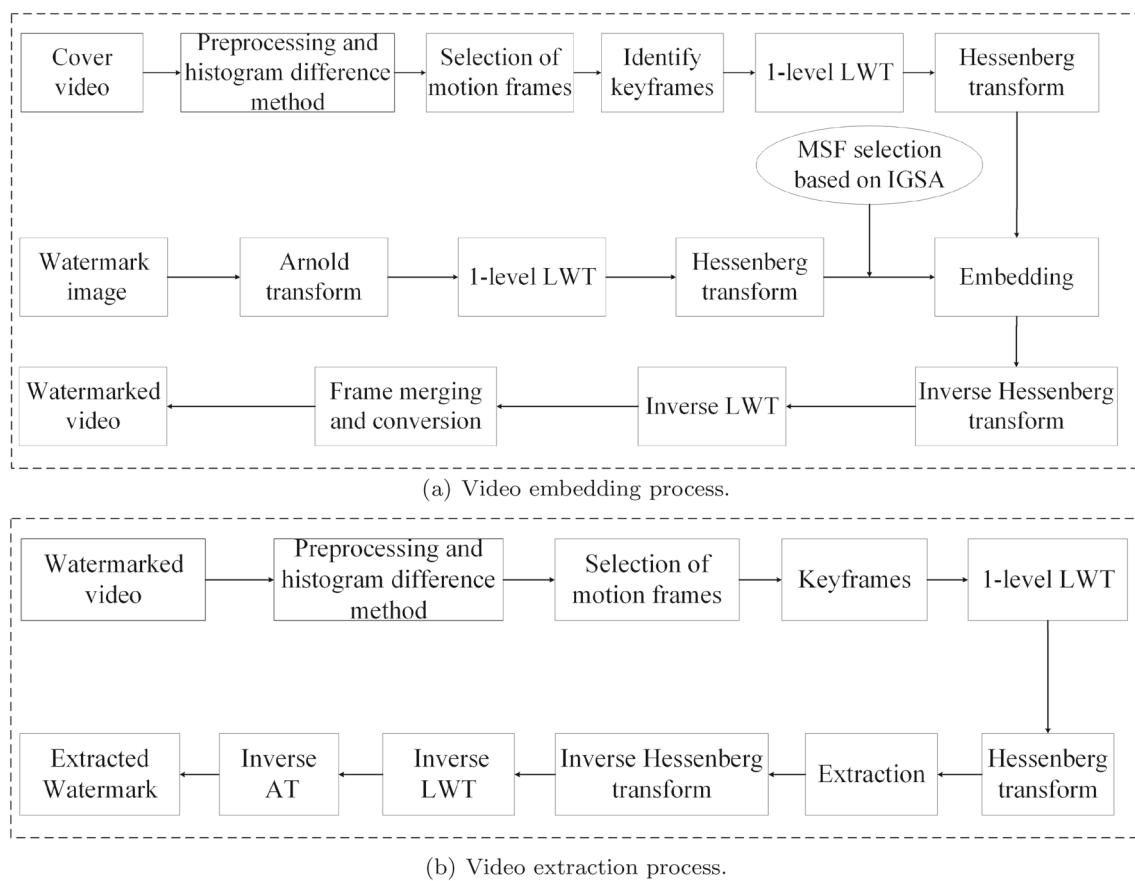
1. The cover video ( $V$ ) is divided into  $k$  frames of size  $R_1 \times R_1$ , namely  $f_1, f_2, \dots, f_k$ .
2. Compute the sum of the absolute histogram differences ( $h_1 - h_2$ ) between two adjacent frames.
3. The sum of the absolute differences is compared to a pre-defined threshold [55]. If the sum exceeds the threshold, the frame is referred as the motion frame (MF).
4. Continue until last frame of video.
5. Calculate the entropy of all the selected motion frames.
6. If the single frame entropy is greater than the compared average entropy of motion frames then that frame is selected as keyframe ( $k_f$ ).

### Embedding process

In this section, a scrambled watermark is incorporated into the keyframes. The following steps are considered while embedding watermark:

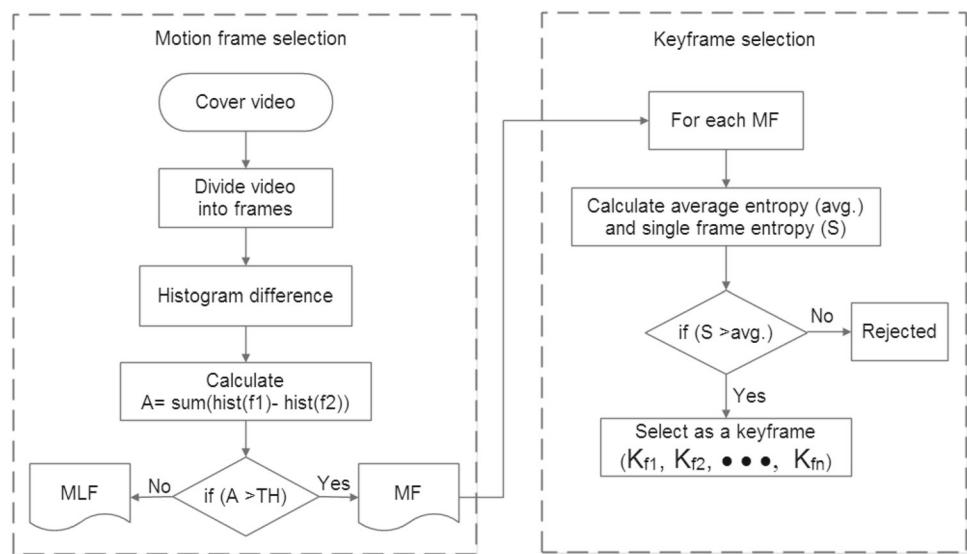
1. Convert keyframe ( $K_f$ ) into YUV color space and select chrominance component ( $K_v$ ) by Eq. (19):

$$[K_y \ K_u \ K_v] = \text{rgb2yuv}(K_f) \quad (19)$$



**Fig. 3** The proposed video-watermark embedding and extraction process

**Fig. 4** The workflow of identification of motion frames and keyframes



2. The chrominance component ( $K_v$ ) is decomposed into four sub-bands ( $LL_v, LH_v, HL_v, HH_v$ ) by performing 1-level LWT defined by Eq. (20) and consider sub-band ( $LL_v$ ) for embedding purpose:
$$[LL_v, HL_v, LH_v, HH_v] = \text{LWT}(M_v) \quad (20)$$
3. Perform HT on sub-band  $LL_v$  using Eq. (21):
$$\text{HT}[LL_v] = Q_v \times H_v \times Q_v^T \quad (21)$$
4. Select watermark logo ( $w$ ) of size ( $R_2 \times R_2$ ) and perform AT to obtain scrambled watermark ( $s_w$ ) for security enhancement purpose.
5. Scrambled watermark ( $s_w$ ) is decomposed into four sub-bands by performing 1-level LWT using Eq. (22) and sub-band  $LL_w$  is selected for further processing:
$$[LL_w, HL_w, LH_w, HH_w] = \text{LWT}(s_w) \quad (22)$$
6. Perform HT on sub-band  $LL_w$  using Eq. (23):
$$\text{HT}[LL_w] = Q_w \times H_w \times Q_w^T \quad (23)$$
7. Apply IGSA to obtain the set of MSF factors ( $\alpha$ ) according to the procedure detailed in “Selection of multiple scaling factors using IGSA algorithm”.
8. Embed the component  $H_w$  into  $H_v$  using Eq. (24):
$$\tilde{H}_v = H_v + \alpha \times H_w \quad (24)$$
8. Apply inverse HT (IHT) on modified matrix ( $\tilde{H}_v$ ) to obtain a Hessenberg embedded matrix ( $\tilde{L}L_v$ ) by Eq. (25):
$$\tilde{L}L_v = Q_v \times \tilde{H}_v \times Q_v^T \quad (25)$$
9. Apply inverse LWT (ILWT) on matrix ( $\tilde{L}L_v$ ) to obtain the watermarked video frames ( $W_v$ ).
$$W_v = \text{ILWT}[\tilde{L}L_v, LH_v, HL_v, HH_v] \quad (26)$$
10. Finally, convert all YUV watermarked video keyframes into RGB watermarked video keyframes and merges with rest video frames to obtain watermarked video ( $V^*$ ).

## Extraction process

In extraction, the reverse operation of the embedding process is performed to recover the watermark logo and is presented below:

1. Divide the RGB watermarked video ( $V^*$ ) into  $k$  frames, namely  $f_1^*, f_2^*, \dots, f_k^*$ , and extract keyframes ( $k_f^*$ ).
2. Convert the watermarked video keyframes ( $k_f^*$ ) into YUV color space and select chrominance component ( $k_v^*$ ).
3. The chrominance component ( $K_v^*$ ) is decomposed into four sub-bands ( $LL_v^*, LH_v^*, HL_v^*, HH_v^*$ ) by performing 1-level LWT defined by Eq. (27) and consider sub-band ( $LL_v^*$ ) for extraction purpose

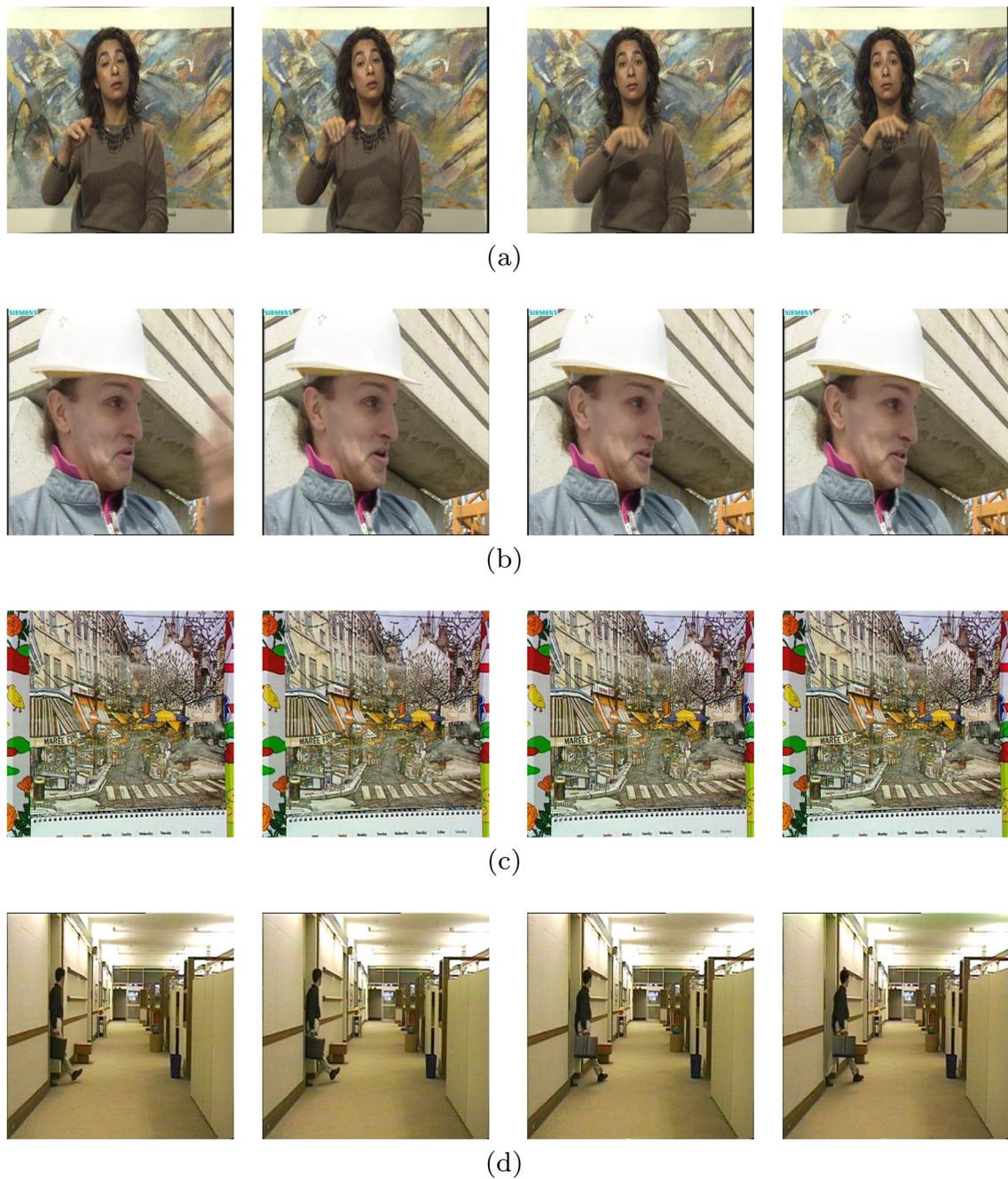
$$[LL_v^*, LH_v^*, HL_v^*, HH_v^*] = \text{LWT}(k_v^*) \quad (27)$$

4. Apply HT on sub-band  $LL_v^*$  of watermarked video keyframes by Eq. (28):

$$\text{HT}[LL_v^*] = Q_v^* \times H_v^* \times Q_v^{*T} \quad (28)$$

5. Extract watermark logo using Eq. (29):

$$H_w^* = \frac{(H_v^* - H_v)}{\alpha} \quad (29)$$



**Fig. 5** The extracted keyframes corresponding to **a** Silent, **b** Foreman, **c** Mobile, **d** Hall monitor

6. Perform IHT on matrix ( $H_w^*$ ) using Eq. (30):

$$\text{LL}_w^* = Q_w \times H_w^* \times Q_w^T \quad (30)$$

7. Apply ILWT on matrix ( $\text{LL}_w^*$ ) to obtain the extracted and scrambled watermark ( $\text{LL}_{ew}$ ) using Eq. (31):

$$\text{LL}_{ew} = \text{ILWT}[\text{LL}_w^*, \text{LH}_w, \text{HL}_w, \text{HH}_w] \quad (31)$$

8. Again, Apply IAT to recover the watermark logo ( $w^*$ ).

#### Selection of multiple scaling factors using IGSA algorithm

The scaling factor is a critical parameter in watermark embedding. It regulates imperceptibility and robustness simultaneously. However, to maintain the trade-off between the above two parameters, a single scaling factor completely fails. The IGSA algorithm is employed to identify the optimal set of MSF through an objective function given in Eq. (32). The

**Fig. 6** Watermarked video frames and their corresponding extracted watermark logos **a** Silent, **b** Foreman, **c** Mobile, **d** Hall monitor



**Table 2** The performance of the proposed technique under various attacks for considered videos in terms of robustness

Attack index	Video attack	Silent	Foreman	Mobile	Hall monitor
MF	Median filtering ( $3 \times 3$ )	0.9998	0.9993	0.9994	0.9999
WF	Wiener filtering ( $3 \times 3$ )	0.9997	0.9998	0.9992	0.9995
GF	Gaussian filtering ( $3 \times 3$ )	0.9997	0.9994	0.9996	0.9992
RO	Rotation ( $45^\circ$ )	0.9998	0.9986	0.9994	0.9997
TR	Translation (30, 30)	0.9999	0.9993	0.9997	0.9998
CR	Cropping CR (center)	0.9995	0.9996	0.9992	0.9997
SH	Sharpening	0.9996	0.9997	0.9996	0.9995
GC	Gamma correction ( $\gamma = 0.6$ )	0.9993	0.9997	0.9997	0.9994
HI	Histogram equalization	0.9998	0.9993	0.9994	0.9996
GN	Gaussian noise (0, 10%)	0.9880	0.9878	0.9872	0.9877
SP	Salt and pepper noise (10%)	0.9882	0.9885	0.9884	0.9889
PN	Poisson noise	0.9953	0.9942	0.9921	0.9947

complete procedure for determining optimal scaling factors by IGSA is given below:

1. IGSA initializes a random population where, each individual consists of ' $d$ ' scaling factors.
2. The objective function described by Eq. (32) is used to calculate the fitness of each individual:

$$\text{OF}(i, j)_{\max} = \left( \frac{\text{MSSIM}(f, \tilde{f}) + \text{MNC}(f, \tilde{f})}{2} + \frac{\sum_{x=1}^K \text{MNC}(w, \tilde{w})}{K} \right) \quad (32)$$

where  $\text{MSSIM}(f, \tilde{f})$  and  $\text{MNC}(f, \tilde{f})$  evaluate mean structural similarity index and mean normalized correlation between cover video frames ( $f$ ) and watermarked

video frames ( $\tilde{f}$ ).  $\text{MNC}(w, \tilde{w})$  measures the mean normalized correlation between the watermark logo ( $w$ ) and extracted watermark logo ( $\tilde{w}$ ), and  $K$  denotes the total number of performed attacks.

3. Each individual of IGSA is updated according to Eq. (18).
4. This process is continued until the maximum iterations are reached.

## Results and discussions

Experiments are simulated on MATLAB 2020a on a 2.5 GHz, I5 processor, and 8 GB RAM system. The efficacy of the proposed technique (IGSA-LH) has been evaluated on four standard benchmark videos: Silent, Foreman, Mobile, and Hall monitor in terms of imperceptibility and robustness. Imperceptibility is examined between cover video and water-

Attack type	Attacked video frames				Extracted watermark			
	Silent	Foreman	Mobile	Hall monitor	Silent	Foreman	Mobile	Hall monitor
1. MF								
2. WF								
3. GF								
4. RO								
5. TR								
6. CR								
7. SH								
8. GC								
9. HI								
10. GN								
11. SN								
12. PN								

**Fig. 7** The quality of watermarked video frames and extracted watermark logos after various attacks

marked video using MPSNR and MSSIM, while robustness is evaluated between watermark logo and extracted watermark logo using MNC. The cover videos and watermark logo are taken from the online available database [56,57]. The parameter settings of the considered techniques are taken from the respective literature. The experimental results are studied in the following sections: “Imperceptibility analy-

sis of the proposed technique” analyzes the imperceptibility while robustness is examined against supposed attacks in “Robustness analysis of the proposed technique”. “Performance analysis against existing techniques” presents the comparative analysis of the proposed technique (IGSA-LH) with five recent video-watermarking schemes under considered attacks in terms of MPSNR, MSSIM, and MNC values.

In addition, statistical validation of the proposed technique is discussed in “Statistical Analysis of the proposed technique”. Finally, “Comparative analysis of time complexity” provides the comparative analysis of the considered techniques in terms of time complexity.

### Imperceptibility analysis of the proposed technique

The imperceptibility evaluates the quality of watermarked video frames. The imperceptibility is assessed between cover video frames ( $f$ ) and watermarked video frames ( $\tilde{f}$ ). The performance of the proposed technique (IGSA-HT) has been examined on all considered four videos in terms of MPSNR and MSSIM which are defined by Eqs. (33) and (34), respectively. Figure 6 shows the quality of watermarked video frames which are quite similar to keyframes depicted in Fig. 5). The proposed technique attains higher MPSNR values as **47.65, 48.97, 47.98, 48.07, 48.52** and MSSIM values as **0.9997, 0.9998, 0.9997, 0.9993, 0.9998** against each considered video. The figure depicts that the proposed technique performs effectively in the embedding phase.

$$\text{MPSNR}(f, \tilde{f}) = \frac{1}{F} \sum_{i=1}^F \text{PSNR}_i \quad (33)$$

$$\text{MSSIM}(f, \tilde{f}) = \frac{1}{F} \sum_{i=1}^F \text{SSIM}_i, \quad (34)$$

where  $F$  denotes the total number of video frames.

### Robustness analysis of the proposed technique

The quality of extracted watermark logo under the various image and video attacks is measured by robustness. The robustness is assessed between the original watermark logo and extracted watermark logo in terms of MNC which is defined by Eq. 35. A total of 12 attacks are confronted on the watermarked video to evaluate the robustness of the proposed technique. These attacks includes (i) median filtering ( $3 \times 3$ ), (ii) wiener filtering ( $3 \times 3$ ), (iii) Gaussian filtering ( $3 \times 3$ ), (iv) rotation ( $45^\circ$ ), (v) translation (30, 30), (vi) cropping (center), (vii) sharpening, (viii) gamma correction ( $\gamma = 0.6$ ), (ix) histogram equalization, (x) Gaussian noise (0, 10%), (xi) salt and pepper noise (10%), and (xii) Poisson noise. Table 2 tabulates the MNC values for all considered four videos against each attack. It can be visualized from the table that the proposed technique attains superior MNC values. The quality of extracted watermark logos corresponding to each attack for all videos are illustrated in Fig. 7. It can be observed from the figure that the proposed technique recovers a superior quality watermark against each attack except

Cover video	IGSA-LH			Karmakar et al. [2]			Bhardwaj et al. [49]			Farri et al. [17]			Kuraparthi et al. [24]			Agilandeswari et al. [50]		
	MPSNR	MSSIM	MPSNR	MSSIM	MPSNR	MSSIM	MPSNR	MSSIM	MPSNR	MSSIM	MPSNR	MSSIM	MPSNR	MSSIM	MPSNR	MSSIM		
Silent	<b>48.97</b>	<b>0.9995</b>	42.65	0.9897	46.63	0.9989	42.67	0.9397	46.78	0.9876	45.82	0.9892						
Foreman	<b>47.98</b>	<b>0.9997</b>	42.57	0.9889	44.07	0.9875	46.71	0.9373	46.72	0.9874	45.76	0.9890						
Mobile	<b>48.07</b>	<b>0.9993</b>	41.98	0.9843	44.13	0.9821	41.23	0.9844	46.83	0.9879	45.90	0.9895						
Hall monitor	<b>48.52</b>	<b>0.9999</b>	42.37	0.9863	43.75	0.9983	42.31	0.9973	46.70	0.9881	45.67	0.9897						
Average	<b>48.39</b>	<b>0.9996</b>	42.39	0.9873	44.65	0.9917	43.23	0.9647	46.76	0.9878	45.79	0.9894						

Gaussian noise, salt and pepper noise, and Poisson noise:

$$\text{MNC}(w, \tilde{w}) = \frac{1}{F} \sum_{i=1}^F NC_i \quad (35)$$

## Performance analysis against existing techniques

The proposed technique (IGSA-LH) has been compared with five recent video-watermarking techniques, namely Karimakar et al. [2], Bhardwaj et al. [49], Farri et al. [17], Kuraparthi et al. [24], and Agilandeswari et al. [50]. “Comparative analysis of imperceptibility” discusses the imperceptibility performance of the proposed technique against existing techniques in terms of MPSNR and MSSIM. While “Comparative analysis of robustness” studies the robustness of the considered techniques in terms of MNC parameters on 12 video attacks.

## Comparative analysis of imperceptibility

Table 3 highlights the MPSNR and MSSIM values corresponding to each video against compared schemes. It can be noticed from the table that the proposed technique attains an average value of MPSNR as **48.39** and MSSIM value as **0.9996**. These superior values confirm that the proposed technique outperforms the compared schemes. For the quantitative analysis, the MPSNR values are plotted against considered methods under all considered four videos which are shown in Fig. 8.

## Comparative analysis of robustness

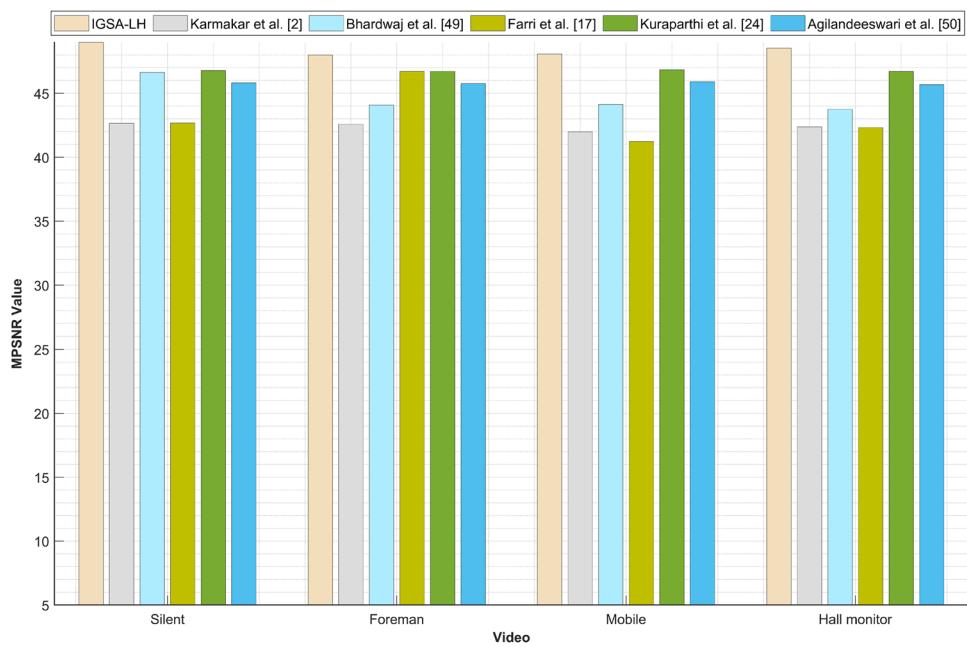
Tables 4, 5, 6, 7, 8, 9, 10, 11, 12 and 15 illustrate the comparative analysis of robustness of the considered techniques against 12 different video attacks.

- (i) **Median filtering attack:** The median attack with kernel size  $3 \times 3$  is performed on watermarked video frames. Table 4 compiles the individual and average values of four videos for all considered schemes. The proposed technique achieves individual MNC values as (**0.9990**, **0.9984**, **0.9991**, **0.9986**) and average MNC value as **0.9988** against this attack for Silent, Foreman, Mobile, and Hall monitor videos. These results confirm that the IGSA-LH technique outperforms compared schemes against median filtering attack.
- (ii) **Wiener filtering attack:** To evaluate the robustness, the watermarked video frames are attacked by wiener filtering. Table 5 compiles the MNC individual values and averaged MNC values for the considered four videos against each considered scheme. The averaged MNC values attained by the considered schemes are

as (**0.9993**, **0.9830**, **0.9956**, **0.9986**, **0.9880**, **0.9869**). These results confirm that the average MNC value of the IGSA-LH technique is superior to the compared schemes that show the IGSA-LH technique’s robustness under wiener filtering attack.

- (iii) **Gaussian filtering attack:** Gaussian filtering with kernel size  $3 \times 3$  is encountered on watermarked video frames. Table 6 presents the average MNC values as (**0.9992**, **0.9837**, **0.9965**, **0.9990**, **0.9910**, **0.9876**) corresponding to each considered schemes for all considered four videos. From the table, it can be confirmed that the proposed technique attains a higher average MNC value as compared schemes. Therefore, the proposed technique performs effectively under Gaussian filtering attacks.
- (iv) **Rotation attack:** A watermarked video is rotated clockwise by  $45^\circ$  angle to prove the robustness. For this attack, the average MNC values are as (**0.9997**, **0.9835**, **0.9951**, **0.9960**, **0.9921**, **0.9884**) which are tabulated in Table 7. This table visualizes that the proposed technique’s average MNC value is superior to the compared schemes. Hence, the proposed technique outperforms under this attack.
- (v) **Translation attack:** To evaluate the effectiveness of the proposed technique, a translation attack (30, 30) shifts the pixels value of the watermarked video. Under this attack, Table 8 highlights the average MNC values corresponds to each considered scheme. These values are as (**0.9996**, **0.9834**, **0.9950**, **0.9958**, **0.9920**, **0.9890**) which confirm that the proposed technique has the highest average MNC value against compared schemes. Therefore, the proposed technique performance is outstanding under this attack.
- (vi) **Cropping attack:** In cropping attack, a watermarked video frame is cropped 25% from the center. The average MNC values obtained by each scheme under this attack are as (**0.9993**, **0.9833**, **0.9947**, **0.9957**, **0.9904**, **0.9868**) which are visualized in Table 9. It can be observed from the table that the proposed technique attains a higher average MNC value. This table concludes that the IGSA-LH technique is quite effective under cropping attack.
- (vii) **Sharpening attack:** A sharpening attack is performed on watermarked video frames. The average MNC values correspond to sharpening attack against considered schemes are as (**0.9988**, **0.9828**, **0.9951**, **0.9956**, **0.9932**, **0.9934**) and visualized in Table 10. It can be envisioned from the table that the proposed technique attains a higher average MNC value which is reflected in bold. Hence, the proposed technique is quite resilient to sharpening attack.

**Fig. 8** Comparative analysis of IGSA-LH technique against considered techniques in terms of imperceptibility



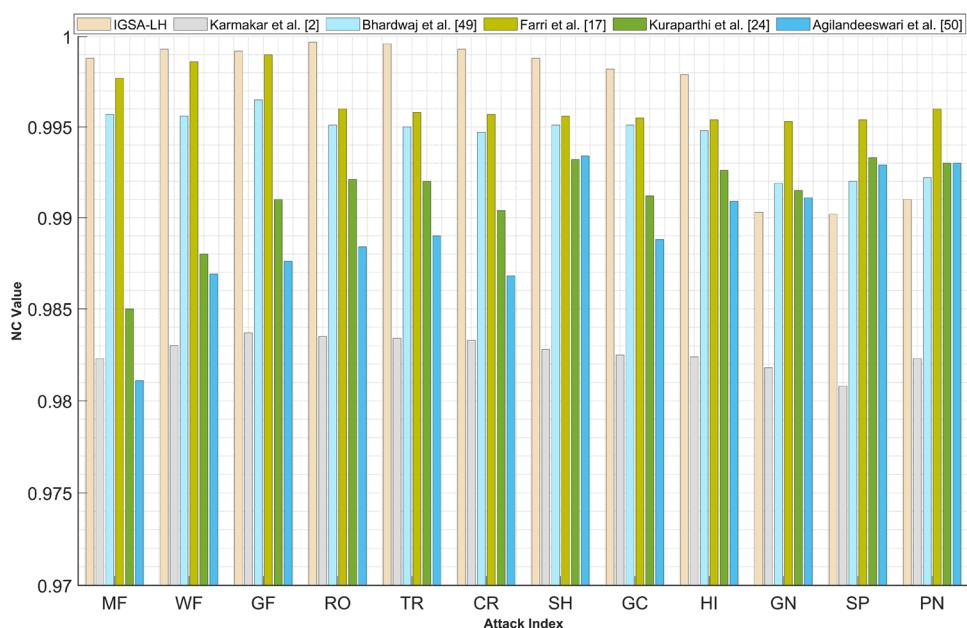
- (viii) **Histogram equalization attack:** Histogram attack is confronted on watermarked video frames to compare the robustness. In histogram attack, the considered schemes attains average MNC values as **(0.9979, 0.9824, 0.9948, 0.9954, 0.9926, 0.9909)**, respectively, and depicted in Table 12. Histogram attack gets maximum average MNC value **(0.9979)** corresponding to the proposed technique. This value confirms that the performance of the proposed technique is quite effective under this attack.
- (ix) **Gamma correction attack:** The contrast of the watermarked video frames is enhanced by gamma correction (0.6) and average MNC values under this attack are presented in Table 11. Table 11 tabulates the average MNC values as **(0.9982, 0.9825, 0.9951, 0.9955, 0.9912, 0.9888)** corresponding to each considered schemes. these results confirm that the average MNC value is higher for proposed technique.
- (x) **Salt and pepper noise attack:** Watermarked video frames are attacked by salt and pepper noise (1%) and average MNC values corresponding to each scheme are depicted in Table 14. Farri et al. [17] gets maximum average MNC value as **0.9954**. Therefore, the proposed technique is performing better than Karmakar et al. [2]. However, Farri et al. [17] outperforms.

xi. **Gaussian noise attack:** In Gaussian noise, the watermarked video frames are destroyed by 10% noise density. Table 13 depicts the average MNC values as **(0.9903, 0.9818, 0.9919, 0.9953, 0.9915, 0.9911)**. It can be visualized from the table that the IGSA-LH technique performs better than Karmakar et al. [2]. However, Farri et al. [17] attains a higher average MNC value **0.9953** which conforms that this scheme outperforms under the Gaussian noise attack.

xii. **Poisson noise attack:** Poisson noise attack is performed on video-watermarked frames to compare the robustness. The average NC values against each scheme are envisioned in Table 15. The average MNC values are as **(0.9910, 0.9823, 0.9922, 0.9960, 0.9925, 0.9930)** for considered schemes, respectively. It can be observed from the table that Farri et al. [17] is also outperforming under this attack.

Moreover, the average MNC values of all the considered videos corresponding to each attack against considered schemes are also presented graphically for better visualization, shown in Fig. 9. It can be illustrated from the figure that the proposed technique outperforms all considered attacks except noise attacks. Therefore, the proposed technique is resilient to considered attacks.

**Fig. 9** Comparative analysis of IGSA-LH technique against considered techniques in terms of robustness



### Statistical analysis of the proposed technique

To statistically validate the performance of the proposed technique (IGSA-LH), a non-parametric Friedman's test is performed for each considered performance parameters, i.e., MPSNR, MSSIM, and MNC. The test contains two hypothesis, i.e., the null hypothesis ( $H_0$ ) and the alternative hypothesis ( $H_1$ ). As per the  $H_0$ , all parameter values generated by comparative methods are significantly equal, while  $H_1$  says that comparative methods are significantly different. The  $p$  value return by test is 0.01 for 30 different executions which is less than considered significant level ( $\alpha = 0.05$ ). Therefore, it can be stated that  $H_0$  fails and the obtained results are significantly different. Further, Table 15 depicts the ranking of the considered techniques for each parameter, respectively, where the proposed technique ranked first. It can be confirmed from the tables that the proposed technique is statistically better than the considered techniques in terms of each parameter.

### Comparative analysis of time complexity

The proposed technique has also been compared with five recent state-of-the-art schemes in terms of embedding and extraction time for 300 video frames of each considered video. Table 17 depicts the embedding and extraction time in seconds. It can be confirmed from the table that the proposed technique spans few seconds only as compared techniques.

These results show that the proposed technique has low time complexity, making the proposed technique suitable for video watermarking in real-time. Hence, the proposed technique is computationally efficient.

### Conclusion

This paper presents a lossless efficient video-watermarking technique based on an optimal keyframes selection using IGSA and HT in the LWT domain. In this scheme, a scrambled watermark logo is incorporated into the keyframes followed by a one-level LWT. IGSA algorithm acquires a set of MSF which maintains the equilibrium between imperceptibility and robustness. The security of the IGSA-LH technique has been improved by performing Arnold transform on the watermark logo prior to embedding. The experimental results were validated against 12 video attacks and compared five recent state-of-the-art schemes. The comparative analysis of imperceptibility and robustness validate that the IGSA-LH technique is quite resilient to attacks. In future work, the performance of the proposed technique can be evaluated on color watermark logos over different video attacks along with various parameters. Further, the elliptic curve cryptography (ECC) technique can be applied to a proposed technique for security enhancement. Moreover, the applicability of the proposed technique can be extended to big-data applications.

**Table 4** The robustness (MNC) of the considered techniques under median filtering attack for considered videos

Cover video	IGSA-LH	Karmakar et al. [2]	Bhardwaj et al. [49]	Farr et al. [17]	Kuraparthi et al. [24]	Agilandeeswari et al. [50]
Silent	<b>0.9990</b>	0.9827	0.9987	0.9968	0.9889	0.9795
Foreman	<b>0.9984</b>	0.9783	0.9914	0.9976	0.9829	0.9807
Mobile	<b>0.9991</b>	0.9897	0.9956	0.9982	0.9827	0.9801
Hall monitor	<b>0.9986</b>	0.9783	0.9972	0.9980	0.9855	0.9842
Average	<b>0.9988</b>	0.9823	0.9957	0.9977	0.9850	0.9811

**Table 5** The robustness (MNC) of the considered techniques under wiener filtering attack for considered videos

Cover video	IGSA-LH	Karmakar et al. [2]	Bhardwaj et al. [49]	Farr et al. [17]	Kuraparthi et al. [24]	Agilandeeswari et al. [50]
Silent	<b>0.9998</b>	0.9834	0.9973	0.9990	0.9825	0.9875
Foreman	<b>0.9993</b>	0.9789	0.9934	0.9995	0.9901	0.9865
Mobile	<b>0.9990</b>	0.9893	0.9951	0.9969	0.9890	0.9862
Hall monitor	<b>0.9992</b>	0.9805	0.9966	0.9989	0.9902	0.9872
Average	<b>0.9993</b>	0.9830	0.9956	0.9986	0.9880	0.9869

**Table 6** The robustness (MNC) of the considered techniques under Gaussian filtering attack for considered videos

Cover video	IGSA-LH	Karmakar et al. [2]	Bhardwaj et al. [49]	Farr et al. [17]	Kuraparthi et al. [24]	Agilandeswari et al. [50]
Silent	<b>0.9994</b>	0.9840	0.9975	0.9992	0.9901	0.9868
Foreman	<b>0.9992</b>	0.9801	0.9968	0.9997	0.9905	0.9873
Mobile	<b>0.9990</b>	0.9890	0.9954	0.9980	0.9912	0.9879
Hall monitor	<b>0.9991</b>	0.9817	0.9962	0.9991	0.9923	0.9885
Average	<b>0.9992</b>	0.9837	0.9965	0.9990	0.9910	0.9876

**Table 7** The robustness (MNC) of the considered techniques under rotation attack for considered videos

Cover video	IGSA-LH	Karmakar et al. [2]	Bhardwaj et al. [49]	Farr et al. [17]	Kuraparthi et al. [24]	Agilandeswari et al. [50]
Silent	<b>0.9997</b>	0.9897	0.9979	0.9951	0.9909	0.9856
Foreman	<b>0.9998</b>	0.9780	0.9904	0.9972	0.9916	0.9890
Mobile	<b>0.9997</b>	0.9898	0.9954	0.9962	0.9923	0.9893
Hall monitor	<b>0.9996</b>	0.9763	0.9968	0.9956	0.9937	0.9896
Average	<b>0.9997</b>	0.9835	0.9951	0.9960	0.9921	0.9884

**Table 8** The robustness (MNC) of the considered techniques under translation attack for considered videos

Cover video	IGSA-LH	Karmakar et al. [2]	Bhardwaj et al. [49]	Farr et al. [17]	Kuraparthi et al. [24]	Agilandeswari et al. [50]
Silent	<b>0.9997</b>	0.9894	0.9981	0.9955	0.9914	0.9871
Foreman	<b>0.9995</b>	0.9782	0.9905	0.9967	0.9922	0.9901
Mobile	<b>0.9998</b>	0.9893	0.9950	0.9960	0.9925	0.9897
Hall monitor	<b>0.9994</b>	0.9766	0.9963	0.9952	0.9919	0.9890
Average	<b>0.9996</b>	0.9834	0.9950	0.9958	0.9920	0.9890

**Table 9** The robustness (MNC) of the considered techniques under cropping attack for considered videos

Cover video	IGSA-LH	Karmakar et al. [2]	Bhardwaj et al. [49]	Farr et al. [17]	Kuraparthi et al. [24]	Agilandeswari et al. [50]
Silent	<b>0.9990</b>	0.9899	0.9978	0.9952	0.9906	0.9867
Foreman	<b>0.9992</b>	0.9780	0.9900	0.9970	0.9896	0.9872
Mobile	<b>0.9994</b>	0.9890	0.9951	0.9958	0.9902	0.9854
Hall monitor	<b>0.9995</b>	0.9765	0.9960	0.9949	0.9911	0.9878
Average	<b>0.9993</b>	0.9833	0.9947	0.9957	0.9904	0.9868

**Table 10** The robustness (MNC) of the considered techniques under sharpening attack for considered videos

Cover video	IGSA-LH	Karmakar et al. [2]	Bhardwaj et al. [49]	Farr et al. [17]	Kuraparthi et al. [24]	Agilandeswari et al. [50]
Silent	<b>0.9989</b>	0.9879	0.9981	0.9953	0.9917	0.9925
Foreman	<b>0.9982</b>	0.9781	0.9903	0.9970	0.9927	0.9932
Mobile	<b>0.9989</b>	0.9894	0.9956	0.9942	0.9936	0.9929
Hall monitor	<b>0.9990</b>	0.9756	0.9964	0.9958	0.9947	0.9949
Average	<b>0.9988</b>	0.9828	0.9951	0.9956	0.9932	0.9934

**Table 11** The robustness (MNC) of the considered techniques under gamma correction attack for considered videos

Cover video	IGSA-LH	Karmakar et al. [2]	Bhardwaj et al. [49]	Farr et al. [17]	Kuraparthi et al. [24]	Agilandeswari et al. [50]
Silent	<b>0.9983</b>	0.9875	0.9980	0.9950	0.9905	0.9913
Foreman	<b>0.9980</b>	0.9778	0.9901	0.9971	0.9897	0.9817
Mobile	<b>0.9985</b>	0.9890	0.9955	0.9940	0.9915	0.9924
Hall monitor	<b>0.9981</b>	0.9757	0.9967	0.9957	0.9929	0.9899
Average	<b>0.9982</b>	0.9825	0.9951	0.9955	0.9912	0.9888

**Table 12** The robustness (MNC) of the considered techniques under histogram equalization attack for considered videos

Cover video	IGSA-LH	Karmakar et al. [2]	Bhardwaj et al. [49]	Farr et al. [17]	Kuraparthi et al. [24]	Agilandeeswari et al. [50]
Silent	<b>0.9976</b>	0.9873	0.9978	0.9951	0.9919	0.9893
Foreman	<b>0.9981</b>	0.9780	0.9900	0.9970	0.9933	0.9902
Mobile	<b>0.9982</b>	0.9891	0.9952	0.9941	0.9937	0.9943
Hall monitor	<b>0.9977</b>	0.9752	0.9963	0.9953	0.9916	0.9897
Average	<b>0.9979</b>	0.9824	0.9948	0.9954	0.9926	0.9909

**Table 13** The robustness (MNC) of the considered techniques under Gaussian noise attack for considered videos

Cover video	IGSA-LH	Karmakar et al. [2]	Bhardwaj et al. [49]	Farr et al. [17]	Kuraparthi et al. [24]	Agilandeeswari et al. [50]
Silent	0.9881	0.9863	<b>0.9970</b>	0.9943	0.9889	0.9903
Foreman	0.9871	0.9776	0.9886	<b>0.9980</b>	0.9943	0.9912
Mobile	0.9980	0.9883	0.9856	<b>0.9938</b>	0.9895	0.9916
Hall monitor	0.9879	0.9750	<b>0.9964</b>	0.9949	0.9932	0.9913
Average	0.9903	0.9818	0.9919	<b>0.9953</b>	0.9915	0.9911

**Table 14** The robustness (MNC) of the considered techniques under salt and pepper noise attack for considered videos

Cover video	IGSA-LH	Karmakar et al. [2]	Bhardwaj et al. [49]	Farr et al. [17]	Kuraparthi et al. [24]	Agilandeswari et al. [50]
Silent	0.9878	0.9865	<b>0.9968</b>	0.9940	0.9951	0.9911
Foreman	0.9873	0.9773	0.9884	<b>0.9981</b>	0.9934	0.9918
Mobile	0.9975	0.9843	0.9858	<b>0.9943</b>	0.9930	0.9952
Hall monitor	0.9880	0.9751	<b>0.9968</b>	0.9951	0.9917	0.9935
Average	0.9902	0.9808	0.9920	<b>0.9954</b>	0.9933	0.9929

**Table 15** The robustness (MNC) of the considered techniques under Poisson noise attack for considered videos

Cover video	IGSA-LH	Karmakar et al. [2]	Bhardwaj et al. [49]	Farr et al. [17]	Kuraparthi et al. [24]	Agilandeswari et al. [50]
Silent	0.9885	0.9867	<b>0.9968</b>	0.9953	0.9914	0.9917
Foreman	0.9891	0.9786	0.9876	<b>0.9986</b>	0.9929	0.9919
Mobile	0.9982	0.9887	0.9865	<b>0.9941</b>	0.9922	0.9948
Hall monitor	0.9880	0.9753	<b>0.9978</b>	0.9959	0.9951	0.9937
Average	0.9910	0.9823	0.9922	<b>0.9960</b>	0.9925	0.9930

**Table 16** Friedman's test of the proposed technique and considered techniques over MPSNR, MSSIM, and MNC

Rank	Mean PSNR (MPSNR)		Mean SSIM (MSSIM)		Mean PSNR (MNC)	
	Method	Mean rank value	Method	Mean rank value	Method	Mean rank value
1	IGSA-LH	<b>2.67</b>	IGSA-LH	<b>2.04</b>	IGSA-LH	<b>1.89</b>
2	Kuraparthi et al. [24]	3.10	Farri et al. [17]	2.79	Farri et al. [17]	2.56
3	Agilandeswari et al. [50]	3.56	Kuraparthi et al. [24]	3.17	Bhardwaj et al. [49]	3.27
4	Farri et al. [17]	4.01	Bhardwaj et al. [49]	3.96	Kuraparthi et al. [24]	3.51
5	Bhardwaj et al. [49]	4.32	Karmakar et al. [2]	4.12	Agilandeswari et al. [50]	4.11
6	Karmakar et al. [2]	4.56	Agilandeswari et al. [50]	4.63	Karmakar et al. [2]	4.24

**Table 17** Time complexity with considered video-watermarking techniques (in s)

Video	Proposed		Existing technique				Kuraparthi et al. [24]				Agilandeswari et al. [50]			
	IGSA-LH		Karmakar et al. [2]		Bhardwaj et al. [49]		Farri et al. [17]		Kuraparthi et al. [24]		Agilandeswari et al. [50]			
	EmT	ExT	EmT	ExT	EmT	ExT	EmT	ExT	EmT	ExT	EmT	ExT		
Silent	<b>23.6</b>	<b>7.43</b>	65.05	8.10	43.56	9.93	45.06	11.63	30.12	8.27	38.14	9.11		
Foreman	<b>19.12</b>	<b>6.28</b>	64.21	7.90	41.21	7.61	42.71	9.31	29.56	8.15	37.87	9.17		
Mobile	<b>14.89</b>	<b>5.15</b>	60.19	7.01	39.51	7.17	41.01	8.87	29.43	8.21	37.89	8.78		
Hall monitor	<b>20.41</b>	<b>5.89</b>	59.18	6.13	41.39	8.57	42.89	10.27	29.74	8.25	37.52	9.53		

EmT embedding time, ExT extraction time

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Asghar MN, Ghanbari M (2012) Mikey for keys management of h. 264 scalable video coded layers. *J King Saud Univ Comput Inf Sci* 24(2):107–116
- Karmakar A, Phadikar A, Phadikar BS, Maity GK (2016) A blind video watermarking scheme resistant to rotation and collusion attacks. *J King Saud Univ Comput Inf Sci* 28(2):199–210
- Balamurugan N, Adimoolam M, John A, et al (2021) A novel efficient algorithm for duplicate video comparison in surveillance video storage systems. *J Ambient Intell Humaniz Comput* 1–12
- Stelter B, Stone B (2009) Digital pirates winning battle with studios. *New York Times*, New York, p 5
- Asikuzzaman M, Pickering MR (2017) An overview of digital video watermarking. *IEEE Trans Circuits Syst Video Technol* 28(9):2131–2153
- Singh R, Ashok A (2021) An optimized robust watermarking technique using CKGSA in frequency domain. *J Inf Secur Appl* 58(102):734
- Singh R, Ashok A, Saraswat M (2020) Optimised robust watermarking technique using CKGSA in DCT-SVD domain. *IET Image Process* 14(10):2052–2063
- Phadikar A (2013) Multibit quantization index modulation: a high-rate robust data-hiding method. *J King Saud Univ Comput Inf Sci* 25(2):163–171
- Choi D, Do H, Choi H, Kim T (2010) A blind mpeg-2 video watermarking robust to camcorder recording. *Signal Process* 90(4):1327–1332
- Preda R, Vizireanu DN (2011) Robust wavelet-based video watermarking scheme for copyright protection using the human visual system. *J Electron Imaging* 20(1):013022
- Yassin NI, Salem NM, El Adawy MI (2012) Block based video watermarking scheme using wavelet transform and principle component analysis. *Int J Comput Sci Issues (IJCSI)* 9(1):296
- Faragallah OS (2013) Efficient video watermarking based on singular value decomposition in the discrete wavelet transform domain. *AEU-Int J Electron Commun* 67(3):189–196
- Masoumi M, Amiri S (2013) A blind scene-based watermarking for video copyright protection. *AEU-Int J Electron Commun* 67(6):528–535
- Thanh TM, Hiep PT, Tam TM, Tanaka K (2014) Robust semi-blind video watermarking based on frame-patch matching. *AEU-Int J Electron Commun* 68(10):1007–1015
- Masoumi M, Rezaei M, Hamza AB (2015) A blind spatio-temporal data hiding for video ownership verification in frequency domain. *AEU-Int J Electron Commun* 69(12):1868–1879
- Rasti P, Samiei S, Agoyi M, Escalera S, Anbarjafari G (2016) Robust non-blind color video watermarking using qr decomposition and entropy analysis. *J Vis Commun Image Represent* 38:838–847
- Farri E, Ayubi P (2018) A blind and robust video watermarking based on IWT and new 3D generalized chaotic sine map. *Nonlinear Dyn* 93(4):1875–1897
- Suresh M, Sam IS (2020) Optimized interesting region identification for video steganography using fractional grey wolf optimization along with multi-objective cost function. *J King Saud Univ Comput Inf Sci* :1–8
- Ayubi P, Barani MJ, Valandar MY, Irani BY, Sadigh RSM (2021) A new chaotic complex map for robust video watermarking. *Artif Intell Rev* 54(2):1237–1280
- Aditya K, Choudhary A, Raj A, Sing M (2017) Video watermarking based on cuckoo search and scene change detection. In: 2017 international conference on intelligent sustainable systems (ICISS). IEEE, pp 420–424
- Agrawal P, Khurshid A (2014) Dwt and ga-psp based novel watermarking for videos using audio watermark. In: International conference in swarm intelligence. Springer, pp 212–220
- Gupta G, Gupta V, Chandra M (2018) An efficient video watermarking based security model. *Microsyst Technol* 24(6):2539–2548
- Kumar PM, Srinivas K, Prabhakar B (2021) A new video watermarking using redundant discrete wavelet in singular value decomposition domain with multi-objective optimization. *Concurr Comput Pract Exp* 24:e6217
- Kuraparthi S, Kollati M, Kora P (2019) Robust optimized discrete wavelet transform-singular value decomposition based video watermarking. *Traitement du Signal* 36(6):565–573
- Sahu N, Sur A (2017) Sift based video watermarking resistant to temporal scaling. *J Vis Commun Image Represent* 45:77–86
- Dutta T, Gupta HP (2017) An efficient framework for compressed domain watermarking in p frames of high-efficiency video coding (hevc)-encoded video. *ACM Tran Multimed Comput Commun Appl (TOMM)* 13(1):1–24
- Singh KM (2018) A robust rotation resilient video watermarking scheme based on the sift. *Multimed Tools Appl* 77(13):16419–16444
- Rajab L, Al-Khatib T, Al-Haj A (2015) A blind DWT-SCHUR based digital video watermarking technique. *J Softw Eng Appl* 8(04):224
- Sharma S, Sharma H, Sharma JB (2019) An adaptive color image watermarking using RDWT-SVD and artificial bee colony based quality metric strength factor optimization. *Appl Soft Comput* 84(105):696
- Lin YT, Huang CY, Lee GC (2011) Rotation, scaling, and translation resilient watermarking for images. *IET Image Process* 5(4):328–340
- Asikuzzaman M, Alam MJ, Lambert AJ, Pickering MR (2014) Imperceptible and robust blind video watermarking using chrominance embedding: a set of approaches in the DT CWT domain. *IEEE Trans Inf Forensics Secur* 9(9):1502–1517
- Mittal H, Saraswat M (2020) A new fuzzy cluster validity index for hyper-ellipsoid or hyper-spherical shape close clusters with distant centroids. *IEEE Trans Fuzzy Syst* 29:3249–3258
- Pal R, Saraswat M (2017) Data clustering using enhanced biogeography-based optimization. In: 2017 tenth international conference on contemporary computing (IC3). IEEE
- Raju P, Subash Y, Rishabh K et al (2020) EEWG: energy-efficient weighted clustering method based on genetic algorithm for HWSNS. *Complex Intell Syst* 6(2):391–400
- Tripathi AK, Mittal H, Saxena P, Gupta S (2021) A new recommendation system using map-reduce-based tournament empowered whale optimization algorithm. *Complex Intell Syst* 7(1):297–309

36. Rashedi E, Nezamabadi-Pour H, Saryazdi S (2009) GSA: a gravitational search algorithm. *Inf Sci* 179(13):2232–2248
37. Mittal H, Pandey AC, Saraswat M, Kumar S, Pal R, Modwel G (2021) A comprehensive survey of image segmentation: clustering methods, performance parameters, and benchmark datasets. *Multimed Tools Appl* 1–26
38. Mittal H, Pandey AC, Pal R, Tripathi A (2021) A new clustering method for the diagnosis of covid19 using medical images. *Appl Intell* 51(5):2988–3011
39. Anari B, Torkestani JA, Rahmani AM (2017) Automatic data clustering using continuous action-set learning automata and its application in segmentation of images. *Appl Soft Comput* 51:253–265
40. Mittal H, Saraswat M (2018) cKGSA based fuzzy clustering method for image segmentation of RGB-D images. In: 2018 eleventh international conference on contemporary computing (IC3). IEEE, pp 1–6
41. Mittal H, Pal R, Kulhari A, Saraswat M (2016) Chaotic kbest gravitational search algorithm (ckgsa). In: 2016 ninth international conference on contemporary computing (IC3). IEEE, pp 1–6
42. Liu J, Xing Y, Ma Y, Li Y (2020) Gravitational search algorithm based on multiple adaptive constraint strategy. *Computing* 102(10):2117–2157
43. Olivas F, Valdez F, Melin P, Sombra A, Castillo O (2019) Interval type-2 fuzzy logic for dynamic parameter adaptation in a modified gravitational search algorithm. *Inf Sci* 476:159–175
44. Mirjalili S, Lewis A (2014) Adaptive gbest-guided gravitational search algorithm. *Neural Comput Appl* 25(7):1569–1584
45. Mittal H, Saraswat M (2018) An optimum multi-level image thresholding segmentation using non-local means 2d histogram and exponential kbest gravitational search algorithm. *Eng Appl Artif Intell* 71:226–235
46. Wang Y, Yu Y, Gao S, Pan H, Yang G (2019) A hierarchical gravitational search algorithm with an effective gravitational constant. *Swarm Evolut Comput* 46:118–139
47. Rawal P, Sharma H, Sharma N (2020) Fast convergent gravitational search algorithm. In: Recent trends in communication and intelligent systems. Springer, pp 1–12
48. Mittal H, Saraswat M (2019) An automatic nuclei segmentation method using intelligent gravitational search algorithm based superpixel clustering. *Swarm Evolut Comput* 45:15–32
49. Bhardwaj A, Verma VS, Jha RK (2018) Robust video watermarking using significant frame selection based on coefficient difference of lifting wavelet transform. *Multimed Tools Appl* 77(15):19659–19678
50. Agilandeswari L, Ganesan K (2018) RST invariant robust video watermarking algorithm using quaternion curvelet transform. *Multimed Tools Appl* 77(19):25431–25474
51. Zear A, Singh PK (2021) Secure and robust color image dual watermarking based on LWT-DCT-SVD. *Multimed Tools Appl* 1–18
52. Koley S (2019) A feature adaptive image watermarking framework based on phase congruency and symmetric key cryptography. *J King Saud Univ Comput Inf Sci* :1–10
53. Bhatnagar G, Wu QJ (2013) Biometrics inspired watermarking based on a fractional dual tree complex wavelet transform. *Future Gener Comput Syst* 29(1):182–195
54. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process* 13(4):600–612
55. Thounaojam DM, Khelchandra T, Singh K, Roy S et al (2016) A genetic algorithm and fuzzy logic approach for video shot boundary detection. *Comput Intell Neurosci* 2016:1–12
56. Sipi (2020) <http://www.sipi.usc.edu/database/>. Accessed 2 Mar 2020
57. Xiph (2020) <https://media.xiph.org/video/derf/>. (Accessed on 01/31/2020)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.