Video steganography: recent advances and challenges

Jayakanth Kunhoth[1] · Nandhini Subramanian[1] · Somaya Al-Maadeed[1] ·
Ahmed Bouridane[2]

Abstract

Video steganography approach enables hiding chunks of secret information inside video sequences. The features of video sequences including high capacity as well as complex structure make them more preferable for choosing as cover media over other media such as image, text, or audio. Video steganography is a prominent as well as the evolving field in the information security domain and significant number of video steganography methods are proposed in recent years. This article provides a comprehensive review of video steganography methods proposed in the literature. This article initially reviews various raw domain-based video steganography methods. In particular, the raw domain-based methods include spatial domain approaches such as least significant bits (LSB), transform domain-based methods such as discrete wavelet transform, discrete cosine transform, etc. Furthermore, the article looks into various compressed domain steganography methods. A critical comparative analysis is included in the article to analyze and contrast the steganography methods proposed in the literature. A brief description of various evaluation matrices for video steganography methods is provided in this article. Moreover, a brief introduction to steganalysis and video steganalysis is provided. The article concludes with a discussion focused on the limitations and challenges of the video steganography methods. Further, a brief insight into future directions in video steganography systems is provided.

Keywords Video steganography · Raw and compressed videos · Data hiding · Motion vectors · Discrete Wavelet Transform (DWT) · Discrete Cosine Transform (DCT)

✉ Jayakanth Kunhoth
  j.kunhoth@qu.edu.qa

  Nandhini Subramanian
  nandhini.reborn@gmail.com

  Somaya Al-Maadeed
  S_alali@qu.edu.qa

  Ahmed Bouridane
  abouridane@sharjah.ac.ae

1   Computer Science and Engineering, Qatar University, Al Jamia street, Doha, Qatar

2   Cybersecurity and Data Analytics Research Center, University of Sharjah,
    Sharjah, United Arab Emirates

1 Introduction

Technology and its development are inevitable in our daily lives. Recent developments in technology have led to the production of cheaper storage, better quality cameras, and inexpensive sensors. Combining this all together has given rise to the internet of things paradigm. This in turn has increased the total amount of data that is being transferred and stored, even sensitive information like personal details, medical information, medical images, banking details, and so on. Information security has been a topic of research ever since the beginning of digital communication. Information hiding techniques are used to counter the attacks on data and provide security, privacy, confidentiality, and integrity to the data [3].

Information hiding techniques are majorly categorized into cryptography, digital watermarking, and steganography.Cryptography [50, 108] is the popular information hiding method used to encrypt the plain data into cipher data. The encrypted cipher data is decrypted back to plain data. Digital watermarking is the basic technique used for hiding watermarks like company logos, and trademarks to claim authorship and ownership. Even though Cryptography and digital watermarking is unbreakable, the encrypted message is visible to the Human Visual System (HVS). On the other hand, steganography aids in hiding the secret information inside the carrier without any traces to HVS [9].

Steganography is not new and has been in existence ever since the BCs. Before the digital era, the information transfer happened by shaving the slave's head, invisible inks, waxes, and silks. As digital media developed, the steganography method has also evolved. Based on the digital media used as the carrier, technical steganography can be divided into image, audio, text and video steganography [48]. Steganography is reduced to the prisoners' problem, where Alice and Bob are inside the prison [10]. Eve is the warden who oversees all the communication between Alice and Bob. Now, Alice and Bob are planning to escape the prison, and to pursue the plan, they have to communicate in a way where Eve does not get any suspicions. Using cryptography and digital watermarking in this situation only protects the content of the information, however, Eve will realize about their secret communication. Steganography is the only option for Alice and Bob to communicate without creating any suspicion. The escape plan is hidden inside a normal-looking image and communicated between Alice and Bob. Eve can see only the normal-looking cover image. Figure 1 explains the overall workflow of the steganography and steganalysis from Alice, Bob, and Eve's perspectives.

Video steganography is the process of hiding secret information inside videos. The secret information can be any media like text, audio, images, video, and binary file and the carrier video can be raw/compressed in any format. A detailed classification of the video steganography methods based on different criteria is given in Fig. 2. The first level of classification is based on the format of the cover video. The cover video considered are either in the raw domain or compressed domain. Raw domain videos are further classified into spatial domain and transform domain. Least Significant Bits (LSB) substitution and other significant methods are included in the spatial domain. Discrete Wavelet Transform (DWT) and Discrete Cosine Transform (DCT) are the extensively used transformation methods to convert the cover videos into the transform domain. After converting the video into the transform domain, embedding of the secret information is undertaken. In the compressed domain, video steganography uses compressed cover videos. Compressed videos have less storage space compared to raw videos and the embedding happens during or after the compression of the videos. Motion Vectors, intra-prediction modes, entropy coding modules,
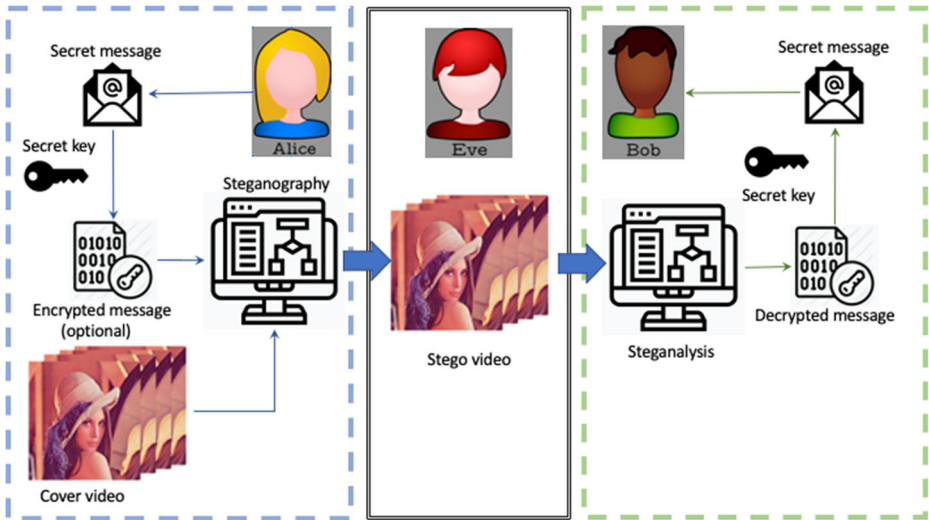
Fig. 1 Steganography reduced to the prisoners' problem. Alice and Bob communicate with each other using steganography methods. Eve does not suspect their secret communication since the secret information is not visible to HVS

and DCT/DST techniques are the extensively used video steganography methods in the compressed domain.

Video steganography has its application in different domains/fields where covert communication is often used. The popular fields where video steganography is utilized are
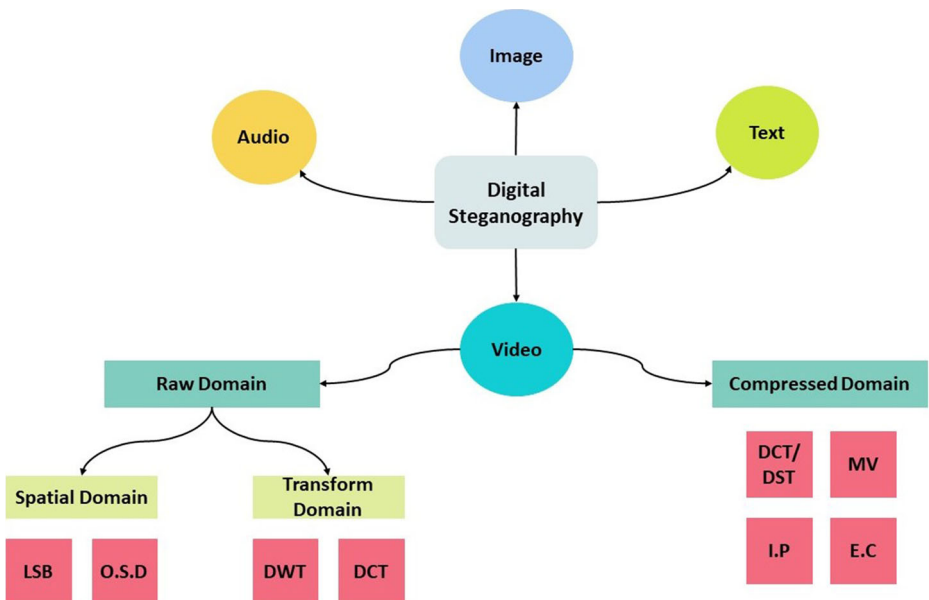

Fig. 2 Hierarchical classification of video steganography methods

intelligence agencies, the military, the medical sector, and multimedia. Intelligence agencies always prefer covert communication when they communicate inside as well as outside the agency. Video steganography is widely used in this case where they can hide the very existence of the secret message from the attacker. Similar to intelligence agencies, military organizations are also widely using steganography techniques to cover up their communication. Because unauthorized disclosure of the secret data can result in national security issues. The medical sector has also benefitted from the application of video steganography. The current advancement in the health sector has made the storage of the patient's information in digital form. Further, this information is stored in the cloud and can be transferred to respective patients or authorized health care providers with the help of internet connectivity. Transferring medical data over the internet is a critical problem since any data loss that happened due to cyberattacks can negatively affect the patient health. The medical sector is using video steganography techniques to conceal their private information from unauthorized entities when it is transferred via communication channels. Apart from that video steganography is used to preserve the privacy of authorized individuals detected in the video sequences captured by the surveillance camera. The data of the individuals are embedded inside the video sequences from the surveillance camera.

The main characteristics of any steganography method are imperceptibility, security, robustness, and hiding capacity. There is always a compromise between the security, robustness, and hiding capacity of the steganography methods. The most popular method for steganography is the Least Significant Bits (LSB) substitution method for image, audio, and video steganography. Traditional video steganography methods are simple, effective, and quick. However, overloading the carrier image may increase the hiding capacity but will compromise the security and robustness in return [69, 103]. Compressed domain has better security and robustness compared to raw domain methods. The storage space is less for compressed video when compared to raw videos. However, during compression, some redundant data that are useful may be lost. Recently, deep learning methods are another perspective applied in the steganography field and have produced exceptional results also. Deep learning methods have increased hiding capacity, security, and imperceptibility but are time-consuming and complex.

In the past, few number of attractive review articles about video steganography are published [7, 69, 86, 96, 103, 105] . Most of these surveys are very brief and mainly concentrated on data-hiding in either raw or compressed domains. In addition, they considered only video steganography and not discussed video steganalysis which is somewhat equally important. In this paper, a collocation of all the methods available for video steganography dating over the past two decades is reviewed. A comprehensive review is done and the methods are grouped and summarized. Not only the methodology but also, the existing research gaps available, the challenges faced are delineated. Further analysis is done to point out the future direction. Further, a brief introduction to video steganalysis is also provided.

2 Video steganography in raw domain

The raw domain-based video steganography methods consider the cover video as a sequence of frames and the data embedding operation is applied to each frame separately. The general data embedding procedure in the raw domain is shown in the Fig. 3.

Initially, the cover video sequence is transformed into multiple frames. Then the secret data is hidden inside the frames using various methods. In the raw domain, the secret data
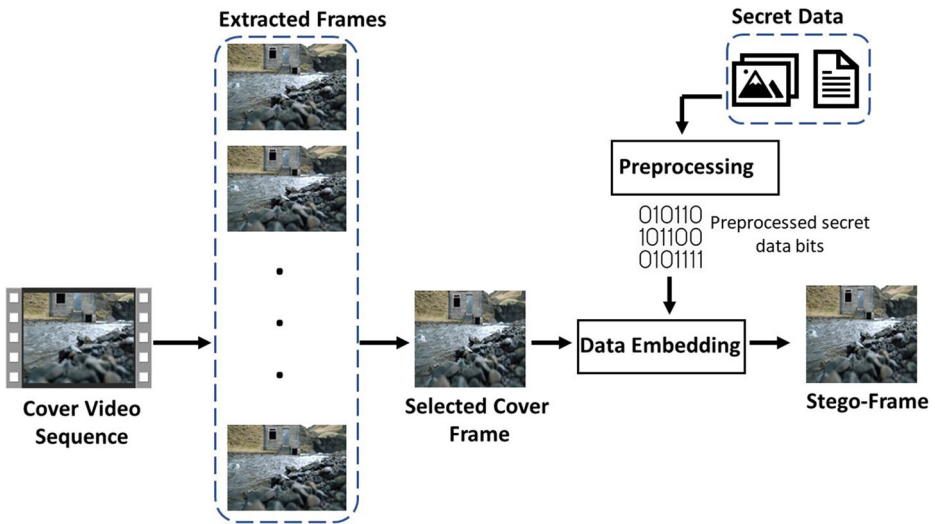
Fig. 3 Steganography procedure in raw video domain

is directly embedded in the spatial domain of the cover frame, or the cover frame is transformed into the frequency domain and secret data is embedded in the frequency domain. Before embedding, the secret data is subjected to preprocessing. Many of the methods have applied encryption techniques, error-correcting codes, etc to preprocess the secret data. The preprocessing of secret data is implemented to ensure the security of the secret data even if the cover video suffers any attacks or frame drops during transmission. The row domain-based method can be classified into two types: data hiding in the spatial domain and data hiding in transform domain methods.

2.1 Data hiding in spatial domain

Data hiding in spatial domain techniques utilize the pixel values of the cover frame to hide the secret data. It means the secret data bits are embedded directly into the pixel intensity values. The least significant bits (LSB) or LSB substitution is a prevalent method where secret data bits are embedded into the least significant bits of cover pixels intensities. utilized for data hiding in the spatial domain. This section is focused on the discussion of various LSB and other spatial domain methods proposed in the literature for data hiding in videos.

2.1.1 Least significant bits methods

The Least Significant Bits based methods are the commonly used algorithm for image, audio, and video steganography. LSB methods are simple and effective. Usually, LSB methods are described as k-LSB substitution methods where k stands for the number of secret bits that can be hidden. Based on the embedding algorithm, the value of k is changed. The hiding capacity of the embedding algorithm depends on the number of bits (k) that can be manipulated in the cover video. Increasing the number of bits for hiding can increase the hiding capacity but will overburden the carrier leading to exposure of the secret information.
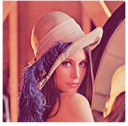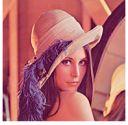
A single pixel in the frame of the cover video has 8 to 24 bits based on the format of the video. Grayscale frames have 8 bits whereas a color frame has 24 bits. A color frame consists of 3 channels (RGB) and each channel consists of 8 bits per pixel. Similarly, an RGB secret frame of the video consists of 8 bits per pixel for all the 3 channels. The least significant bits of the cover frame is replaced with the most significant bits of the secret frame. During extraction, the LSB bits of the stego is taken and 0s are padded to get an approximation of the intended secret information. Along with different secret media, different levels of hiding capacity are considered, and MSE, PSNR, and SSIM values of different combinations are given in Table 1. From the Table 1, it can be noted that the PSNR value decreases with increasing hiding capacity. The imperceptibility, security, and robustness are also low with increased hiding capacity. However, the computational time taken for performing all the combinations, even with the increased hiding capacity is similar. It is considered that there is always a compromise between the hiding capacity and the imperceptibility, security, and robustness of the proposed system.

In the last two decades, different variants of the LSB-based substitution technique [8, 21, 22, 29, 60, 72, 115, 124] are used for hiding data in video streams. Ramalingam [99] proposed a software tool named "stego machine" for hiding the secret text files inside the video. The proposed approach utilized the traditional 1 LSB approach. Ie, only 1 least significant bit of the pixel values in the cover frame is used for embedding the data. The four least significant bits (4 LSB method) of the cover frame are utilized for hiding the secret data in [34]. Before embedding, the secret data (frame) is subjected to partition using the non-uniform rectangular partition algorithm and the resulted grids of the secret data are embedded in the cover frame. , In the last decade, most of the LSB-based approaches included the cover pixel selection technique, secret data encoding as well encryption techniques to improve the security and robustness of the data hiding algorithm. "HASH LSB" [17], an extended version of the LSB approach integrated hash function along with the LSB substitution method. The proposed LSB scheme followed a '3-3-2' embedding pattern to hide the secret data in the cover image. For a cover pixel, the '3-3-2 LSB pattern' utilizes 3 LSBs of the red component, 3 LSBs of the green component, and 2 LSBs of the blue component for hiding the data. The hashing function is employed for selecting the suitable bits in LSBs of the pixels to embed the data. The "HASH LSB" technique along with the RSA encryption scheme is introduced in [40]. The proposed work utilizes the RSA algorithm for encrypting the secret data before data hiding using the 'HASH LSB' embedding scheme.

Significant number of methods have introduced the application of encryption and preprocessing methods (including encoding the secret message with error-correcting codes) for handling the secret data before embedding it inside the cover frame. Encryption and computer forensics have been employed with the 4 LSB method for data embedding in [80]. Although encryption techniques can provide additional security for data, the proposed method is more prone to attack. Because the 4 LSB substitution can cause significant visual degradation to the cover video after embedding the secret data. Yadav et al. [131] focused on improving the security of the secret data in the LSB-based data hiding approach by introducing an encryption scheme. The encryption scheme involves the XOR operation between the secret data and the secret key to encrypt the secret data before embedding.

Mstafa et al. [81] introduced error-correcting codes along with the encryption technique in the LSB-based data hiding scheme. The proposed approach converts the cover frame into the YCbCr color space and utilized Y, U, and V components for hiding the data. Initially, the pixel positions of Y, U, and V components in the cover frames are altered using a specific key. The secret data is encoded using hamming code and the resulted encoded data is further

Table 1 Comparative analysis of different hiding capacity using the LSB method with Lena as the cover object

Cover video	Secret	Stego frame	Capacity	MSE	PSNR	SSIM
	'a'		0. 25	1.4e-05	96.67	0.99
	'ab'		0.25	2.16e-05	94.78	0.99
	'this is a secret message'		0.5	0.00013	86.83	0.99
			0.125	0.499	51.14	0.99
			0.25	2.086	44.93	0.98
			0.375	8.47	38.85	0.96
			0.5	34.22	32.79	0.87

encrypted with another key by applying the XOR operation. In a selected pixel of the cover frame, 7 bits of encoded and encrypted secret data are embedded in a manner where 3 bits are in the Y component and the remaining every 2 bits in the U component as well as in the V component. Instead of hamming code, BCH codes are used for encoding secret data in Mustafa et al. [84]. The proposed work followed the same color space conversion and pixel

position alteration scheme introduced in [81]. After that, the bits position of the secret data is changed using a private key. Then, the secret data was encoded using BCH codes. In each cover pixel, 8 bits of secret data are embedded by following a 3-3-2 LSB pattern (3 bits in Y, 3 bits in U, and 2 bits in V). Integration of error-correcting code with encryption technique significantly improved the security of the proposed data hiding methods.

LSB substitution-based method [1] employed a metaheuristic optimization algorithm namely 'cuckoo search' for preprocessing the secret data. The cuckoo search algorithm is implemented to process the secret message byte by byte. It arranges the bits of each byte (of the secret data) in distinct five forms before embedding them inside the cover frame. In the cover frame, the Euclidean distance method is utilized for finding the appropriate pixel for embedding the secret data. Furthermore, Leavy flight random walk methodology is employed for traversing from the current cover pixel to the next cover pixel. The secret data is hidden in the cover pixel by following the 3-3-2 LSB embedding pattern. Jha et al. [37] proposed an extended version of LSB for hiding the video sequence inside the video sequence. Before embedding the secret frame inside the cover frame, the pixels of the selected cover frame is subjected to scrambling using the prime factorization technique. Then the bits of the secret frame is embedded into the cover frame using the spiral LSB technique. Khan et al. [46] focused on hiding the secret data using the LSB approach in the keyframes of the cover video sequence. Various statistical features such as standard deviation, skewness, and kurtosis are utilized for extracting the keyframes from the cover video. Moreover, the AES algorithm is used to encrypt the secret message before embedding.

To improve the robustness and security of the LSB-based data hiding techniques, various adaptive steganography approaches for the LSB method are proposed in the literature. The adaptive steganography approach hides the secret data in a specific predefined region of interest (moving objects, skin regions, etc.) in the cover frame. Edges of the objects in the cover frame are chosen for predominantly hiding the secret data in [43]. The Canny edge detection technique is utilized for detecting the edges. The 4-LSB method is used for embedding the secret data in the detected edge pixels of the cover frame. Non-edge pixels are also utilized for hiding the data by using the 2-LSB method. Moreover, the RSA algorithm is employed for providing additional security to the secret data by encryption mechanism. Background objects and foreground objects (except their face) in the cover frame are selected as the region of interest for hiding the medical images [6]. The human vision region of interest is utilized to classify the background and foreground objects in the cover frame. Motion attention index value and variation range are used to determine the human vision region of interest. Moreover, the AES mechanism is implemented to encrypt the medical images. . Mstafa et al. [88] proposed a 4-LSB method to embed the secret data in corner points available in the cover frame. The shi-Tomasi algorithm is used to detect regions of corner points within the cover video frames. The secret data is encrypted using Arnold's cat map technique. The obtained result shows that the proposed approach is robust against artificial noises.

In the reference [78], the cover video is treated as a sequence of frames, and each cover frame is partitioned into four quadrants. Each byte in the secret data is divided into 4 pairs of bits, where each pair of bits are embedded in each quadrant of the cover frame. LSB substitution-based approach proposed in [39] introduced a cover frame selection mechanism based on DNA alphabets. To improve the security of the embedding data, all pixels of the selected cover frame are not chosen for embedding the secrete data. Instead, suitable pixels are selected by the construction of a burger chaotic map for pixel plotting over the cover frame followed by randomization of the pixel points obtained in the chaotic map

using a random number generator. The linear congruential generator is used for generating the random numbers. Younus et al. [134] utilized the knight tour algorithm for selecting the suitable random pixels in the cover frame for hiding the secret data. The knight tour algorithm is based on the movement of the knight on the chessboard. Once the suitable pixels are selected, the secret data is hidden in the last 2 LSBs of the cover pixel.

LSB substitution method is integrated with patch-wise code formation technique for hiding a video inside another video [97]. Initially, the cover frame is preprocessed using the fuzzy adaptive median filtering method to remove the impulse noises. Further, the redundancies in the frame are removed using the pixel clustering technique. The secret data is embedded in the least significant bits of the cover pixels. After embedding the cover video is transformed into an encoded format using the patch-wise code formation technique. The patch-wise code formation technique is included to improve security and reduce the transmission time.

2.1.2 Other spatial domain methods

Most of the video steganography techniques for raw videos in the spatial domain have relied on LSB substitution to embed the secret data. Besides that, a couple of works proposed in the last decade have utilized non-LSB methods to hide the secret data in the spatial domain of raw videos. Jangid et al. [36] utilized K-means clustering and LBP features to embed the secret data. The cover frames are converted into Lab color space and the K-means clustering algorithm is implemented to group the cover frames into different clusters. Only selected clusters of the cover frames are chosen for embedding the secret data. LBP methodology is utilized for hiding the secret data in the selected cluster of the cover frame. The obtained evaluation results show that the proposed method achieved better imperceptibility than the method which embedded the secret data in the transform domain (IWT) using the LSB approach. An adaptive steganography approach [47] utilized the Cb component in the YCbCr color space of cover frames for embedding the secret data. Firstly, the skin regions (face) in the cover frames are detected. The skin region detection methodology involves converting the RGB frames to HSV color space followed by applying morphological dilation as well as filling operation. The frames containing skin regions are transformed to YCbCr color space and the frame with the least MSE value is chosen for embedding the secret data.

Blocks of the cover frame that have nonuniform colors are exploited for embedding the secret data [13]. The regional histogram optimization technique is implemented to find the appropriate cover pixels (the blocks with nonuniform colors). The regional histogram optimization method divides a cover frame into multiple blocks and the histogram dispersion of each block is plotted to find the blocks that have uniform colors. The blocks with uniform colors are excluded and the rest of the blocks are utilized to embed the secret data. Although the proposed approach is simple, it delivered acceptable imperceptibility. A reversible lossless data hiding technique based on a histogram distribution constrained scheme is proposed in [4]. The proposed approach utilizes the luminance component of raw video frames for embedding the data. Firstly, the luminance component of the video frame is extracted and separated into multiple non-overlapping blocks. After generating multiple non-overlapping blocks, the arithmetic difference of each block is calculated and secret data is placed into the blocks by shifting the arithmetic difference values of the block. The experimental results show that the proposed method is robust against the h.264/AVC compression.

Kelash et al. [45] utilized the average histogram values of the cover frames to determine the suitable frames from cover video sequences for embedding the secret data. Initially, the

histogram variation of each frame is computed and the frames having variation greater than the histogram constant value are selected for hiding the data. Each selected frame is broken down into blocks and suitable pixels are selected by comparing the consecutive blocks. Each suitable pixel is divided into two parts. The data is embedded in the right part of the pixel and the count of the bits is altered while embedding is encoded in the left part of the pixel. Ramalingam and Isa [101] proposed a data hiding scheme to embed the secret data in random RGB components of the cover frame. The pixels on the cover frame are randomly permuted using a random key (seed) and a pseudo-random number generator. Every 8 bits of the secret message are embedded in the random pixel by following the specific order "RGBBGRGG". It means the first and fifth bit of the secret message is embedded in the red component of the pixel, the third and fourth bit in the blue component of the pixel, and the rest of the bits are embedded in the green component of the pixel.

Most of the LSB and other spatial domain methods discussed in this section achieved acceptable imperceptibility as well as data hiding capacity. However, the robustness of the proposed methods is a concern and many of the methods have not conducted any quantitative analysis to evaluate their robustness. Most of the spatial domain-based methods are prone to steganalysis attacks and are not robust against compression as well as noise attacks. A critical analysis of different methods in spatial domain steganography is given in Table 2.

2.2 Data hiding in transform domain

Unlike the spatial domain-based method which directly embeds the secret data in raw pixel intensities of the cover pixel, the transform domain-based method converts the blocks of cover frames in the spatial domain to the transform domain. After that, the secret data is embedded in the least significant bits of transform coefficients. The general workflow of data hiding in the transform domain is shown in the Fig. 4. Discrete wavelet transform (DWT) and discrete cosine transform (DCT) are two predominantly used transform function in video steganography [53, 83, 94, 117]. The general description of both DWT, and DCT functions and discussions of methods using DWT, and DCT functions for video steganography are provided in this section.

2.2.1 Discrete wavelet transform

A signal can be represented in either the time domain or frequency domain and each domain capture interesting features in their domain. In a stationary signal, the frequency components won't change with time, whereas, in non-stationary signals, the frequency changes over time. Wavelet transformation represents the signals in time-frequency and so is effective with non-stationary signals. Fourier transform is common with stationary signals. Generally, the time domain information is passed through low pass and high pass filters at different levels to decompose the information. Similarly, the frequency domain information is captured by decomposing the depth of the signals. Wavelet transform can be either continuous or discrete. Discrete Wavelet transformation is of interest in image processing tasks as it is simple, operational, and effective.

Discrete Wavelet Transform (DWT) decomposes the signal into sets with significant and insignificant information. The significant information is related to general appearance and is called low-frequency DWT coefficients. Similarly, the insignificant information represents the behavior of the signals and is called the high-frequency coefficients. A single signal is passed through a set of filters and decomposed into two parts - approximation and details.

Table 2 Critical analysis of video steganography methods in spatial domain

Method	Secret	Preprocessing	Psnr	Imperceptibility	Robustness	Capacity	Remarks	Ref.
1-LSB	Text	Encryption	N/A	↑	→	N/A	(-) Security is not tested	[99]
4-LSB	Video	Non uniform rectangular partition	29.15	↑	→	1.5 bpp	(+) No obvious visual distortion	[34]
Hash-LSB	Text	Nil	44.2	↑	→	2.6 bpp	(-) Hash based LSB alone is not secure enough under steganalysis attack	[17]
Hash-LSB	Text	Encryption	74.18	↑	→	100 %	(-) Security is not tested	[40]
4-LSB	Text	Encryption	N/A	→	→	12.5 %	(-)Significant visual distortion	[80]
LSB	Video	Encryption	35.4	↑	→	N/A	(-) Prone to steganalysis	[131]
3-2-2 LSB in YUV	Image	Encryption and ECC	52	↑	↑	90 KB	(+) Acceptable security against steganalysis attacks	[81]
3-3-2 LSB in YUV	Text	Encryption and ECC	55.3	↑	↑	246 KB	(+) Increased security and capacity	[84]
3-3-2	Image	Cukoo search algorithm	51.19	↑	→	N/A	(+) High embedding efficiency with less visual distortion	[1]
Spiral LSB	Video	Nil	42.51	↑	→	25%	(-) Low hiding capacity	[37]
LSB	Text	Encryption	55.60	↑	→	N/A	(-) Security is not tested	[46]
Adaptive 4 2 -LSB	Text	Encryption	53.55	↑	→	128 KB	(-) Supports AVI carrier files only	[43]
Adaptive 1,2 3 LSB	Text	Encryption	67.17	↑	→	6300 KB	(+) Highly secure method	[6]
Adaptive LSB	Text	Encryption	52.24	↑	→	N/A	(+)Only minimal degradation in cover video	[39]
2-LSB	Text	Encryption	67.36	↑	→	N/A	(-) Security not tested	[134]
Adaptive 4-LSB	Image	Encryption	61.44	↑	↑	0.069 bpp	(-) Low embedding capacity	[88]
Adaptive (YCbCr components)	Text	Nil	85.18	↑	↑	N/A	(-) Security not tested	[47]
Adaptive (Histogram analysis)	Text	Nil	30 - 50	↑	→	N/A	(-) Security not tested	[13]
Adaptive (Histogram analysis)	Text	Nil	48.84	↑	→	1.1 %	(-) low embedding capacity	[45]

Secret: type of secret data used for hiding, Preprocessing : Preprocessing applied on secret data before embedding for improving its security, ↑ : High, ↓ : Low
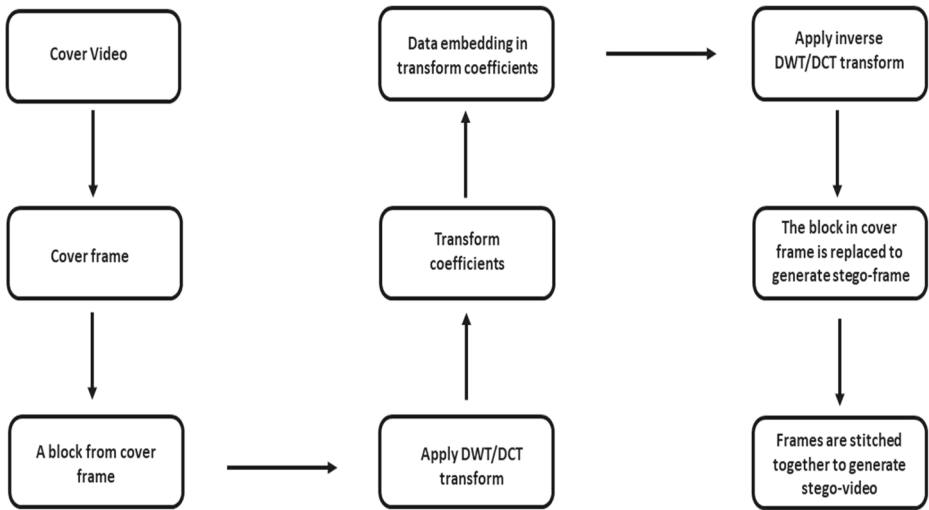
Fig. 4 Data hiding in transform domain: A general workflow

The rows and columns of an r × c image are passed and processed independently. The formula used for decomposing the rows and columns are given in (1) and (2) respectively.

$$i(x, y) = \begin{cases} \sum_0^{n-1} I(r, m).h_L(m-r) &, r \equiv 0 (mod 2) \\ \sum_0^{n-1} I(r, m).h_H(m-r) &, r \equiv 1 (mod 2) \end{cases} \qquad (1)$$

where I(r,c) is the image, h_L is the low pass filter and h_H is the high pass filter.

$$i'(x, y) = \begin{cases} \sum_0^{n-1} i(m, c).h_L(m-c) &, c \equiv 0 (mod 2) \\ \sum_0^{n-1} i(m, c).h_H(m-c) &, c \equiv 1 (mod 2) \end{cases} \qquad (2)$$

Finally, the low pass components are arranged in the top half while the high pass components are arranged in the bottom half. The same steps are repeated for several iterations based on the application. Every time the transformation is applied to the low-frequency components.

The data hiding technique based on the LSB substitution approach in the wavelet domain is implemented in [95]. The cover frames of the video sequence are transformed to the wavelet subbands by using the lazy lifting wavelet transform technique. The three least significant bits of each transform coefficient are used to embed the secret data. Moreover, the meta-information about the hiding scheme is embedded in the LSBs of the audio component. The meta-information is required for the receiver to extract the secret data. A hybrid data hiding method [2] used RSA encryption, three-level 2D-DWT operation, DCT operation, and LSB substitution method to embed the secret data in the cover video. For Each cover frame, the red component is extracted and subjected to 3 level 2D- DWT operation. Only the HH band of the red component is decomposed into three levels. Later, the DCT operation is performed on the resulted HH band (The HH band obtained after three-level decomposition). Among the obtained DCT coefficients, middle-frequency subbands are selected for embedding the secret data. To provide additional security for the secret data, it is encrypted using the RSA algorithm. DWT-based method [49] utilized only the red channel of the RGB cover frame for hiding the secret data. The red channel of the cover frame

is extracted and the DWT operation is performed to decompose it into frequency subbands. The secret data is embedded in the HH subbands using the LSB approach. Sushmitha et al. [114] proposed an approach for hiding the secret video inside a cover video in the wavelet domain. DWT operation is applied to the cover frames and the LSB substitution approach is used to hide the secret frames in the HL, HH, and LH subbands of the cover frame. Moreover, the proposed approach is extended to hide two secret videos in a single cover video. The cover video is split into two parts, one secret video is embedded in the first part and another secret video in the second part using the same DWT technique and LSB method.

Owing to the fact that the adaptive steganography technique can provide additional security and robustness, few adaptive steganography approaches based on the wavelet domain are proposed in the literature. Lu et al. [74] proposed an adaptive technique to hide the biometric data in the frequency subbands of the video frames. The suitable frame and regions of interest inside the frame for embedding the data are selected by implementing a motion analysis technique. Initially, a watermarking mechanism is employed to embed the sequence number for each cover frame. The watermarking mechanism allows the receiver to extract the information from the cover frame without information loss. One level DWT operation is performed on each watermarked frame to divide it into subbands. The motion analysis is performed initially on each frame and the frame with higher motion activity is selected. In each selected frame, motion analysis is again performed for each block to find the blocks with higher motion activity. The secret biometric data are embedded in the blocks with higher motion activity. Multiple object tracking algorithm is employed in [87] for finding the suitable region of interest to hide the secret data. In the cover frame, only pixels of moving objects are chosen for embedding the secret data. The multiple objects tracking algorithm based on background subtraction and Kalman filtering is applied to the cover frames to detect the moving objects. The secret data is embedded in the DWT or DCT coefficients of the regions that have moving objects. Moreover, the secret data is encoded using error-correcting codes before embedding it in the region of interest.

Skin tone areas present in the cover frames are considered as the region of interest for hiding covert information. Embedding inside the skin tone areas is based on the fact that the skin tone areas in the frame have better immunity against noises. Kumar et al. [57] utilized the skin tone areas in the cover frames for embedding the secret data. A skin detection algorithm is implemented to detect the frames and regions containing the skin. The selected cover frames are decomposed into frequency subbands using a three-level DWT operation. The secret data is embedded in the LSBs of the transform coefficients using the 1-LSB algorithm. Obtained results show that the proposed method is robust against MPEG-4 compression attacks. Sadek et al. [104] utilized an adaptive skin detection algorithm to generate a skin map for each cover frame. Further, the skin maps are converted to a skin-block-map to eliminate the error-prone skin pixels and choose good pixels for hiding the secret data. The red and blue channels of the skin regions are used for embedding the data. The red and blue channels are transformed into wavelet domains by applying three-level 2D-DWT and the secret data is hidden in the frequency coefficients. The introduction of the skin block map to eliminate the error-prone pixel has improved the robustness of the proposed method against the MPEG-4 compression attack. But the process is computationally expensive and eliminating error-prone pixels resulted in less hiding capacity. In [82] facial regions present in the video frames are exploited for hiding the secret data. The proposed work employed the Viola-Jones algorithm and KLT tracking algorithm to detect and track the facial regions in the cover frame. The detected regions in the cover frame are decomposed to frequency subbands by applying DWT. The secret message is encoded using BCH codes and the encoded

secret is embedded in the transform coefficients. Further, the key used for encoding and information about the region of interest is embedded in the non-facial regions.

A trained artificial neural network and LSB algorithm are used for data hiding in the DWT domain [42]. Initially, the extracted cover frames are transformed to the wavelet domain using DWT operation. The trained artificial neural network classifier is utilized to select the suitable regions in the frequency subbands for hiding the secret data. Once the suitable regions are identified, the secret data is embedded using the LSB algorithm. Suresh et al. [112] integrated oppositional grey wolf optimization algorithm and DCT-based keyframe extraction mechanism for hiding the secret data in the DWT domain. The oppositional grey wolf optimization algorithm is employed for enhancing the visual quality, minimizing the distortion in the cover video after embedding, and improving the security of the secret data. DCT operation is used to detect the scene changes and extract the keyframes. Once the keyframes are obtained, the optimal regions in the keyframes for hiding secret data are selected using the oppositional grey wolf optimization algorithm. The optimal regions are decomposed into frequency subbands by applying two-level DWT. Only LL and HH bands are used for embedding the secret data.

Wahab et al. [121] proposed a hybrid approach for video steganography based on discrete wavelet transform (DWT) and histogram shifting. After transforming the cover frame to the wavelet domain by performing the DWT operation, the histogram shifting technique is implemented to embed the secret data. Unlike traditional DWT methods that directly embed the secret data in subbands, the proposed work selects the histogram of subbands with higher frequency values and a part of the selected histograms are subjected to shifting operation. The shifting operation is carried out to create the space for embedding the secret data. Dalal and Juneja [15] utilized the frequency subbands of the luminance component for hiding the secret data. The proposed work converts the RGB cover frames to YUV color space and extracts the luminance component (Y). Second level 2D-DWT is applied to the Y component and 16 subbands are generated. Among 16 generated subbands, 8 middle subbands are used for embedding the secret data. Evaluation results show that the proposed method is robust to noise attacks and compression attacks to an extent.

Dalal and Juneja [16] conducted a study to compare the performance of Orthogonal and bi-orthogonal filters used for frame decomposition in DWT domain-based data hiding methods. During implementation, the cover frame is decomposed by applying one level 2D-DWT with orthogonal filters and bi-orthogonal filters separately. The secret image is hidden in the LH and HL sub-bands of the cover frame. The obtained results show that Bi-orthogonal wavelet filters outperformed the orthogonal filters for DWT domain-based video steganography applications. A comparative analysis study is conducted in [109] to examine the performance of DCT, DWT, and CvT (curvelet transform) for hiding the secret data inside video in the transform domain. LSB substitution method is implemented to hide the data in the transform domain after applying the transform function. Obtained evaluation results show that CvT based method delivered better imperceptibility in both conditions (with and without the presence of noises). Moreover, the DCT and DWT approaches are more computationally expensive than CvT.

Few works in the literature have utilized the integer wavelet transform for hiding the secret data inside the video. An integer wavelet transform-based fusion approach is applied for minimizing the distortion associated with hiding the secret data inside the video [90]. Both the cover frame and secret frame are decomposed by applying the IWT. The decomposed cover frame and secret frame are fused by adding the wavelet coefficients of

respective sub-bands of both the cover frame and secret frame. Then inverse integer transform was applied to the fused matrix to generate the stego-video. Evaluation results show that the proposed method generated the stego-video with less distortion and acceptable robustness. Ramalingam and Isa [100] proposed a method based on Haar IWT and LSB for embedding the text messages inside AVI video files. The RGB frames of the cover video are decomposed to frequency sub-bands by applying one-dimensional Haar IWT. Before decomposing the RGB frame into frequency sub-bands, the video frames are subjected to normalization. Normalization is implemented to prevent the overflow or underflow that may happen while altering the transform coefficients of the cover video. The LSBs of higher frequency bands HH, HL, and LH are used for embedding the secret text data.

2.2.2 Discrete cosine transform

DCT is also a transform function like DWT that divides the image into spectral sub-bands. The major difference between DCT and DWT is the earlier one generates more frequency bands and provides higher frequency resolution. Nevertheless, DWT generates few frequency bands and provide high spatial resolution. Significant amount of works in the literature used the DWT domain to embed the secret data in raw videos. Unlike DWT, the DCT domain is not frequently used in the literature to hide the secret data inside the raw videos . On the other hand, video steganography methods proposed in the compressed domain have utilized the DCT domain extensively for hiding the secret data. This section discusses the DCT based method in the raw video domain only.

In the raw video domain, the two-dimensional DCT is applied to each frame of the video separately and transforms the frame into low, middle, and high-frequency bands. The secret data is hidden in the transform coefficients of either one or multiple bands.

Consider an arbitrary frame I of resolution $J \times K$. And T is the transformed frame generated by applying DCT on I. The DCT coefficients of T are calculated using the equation,

$$T_{xy} = \alpha_x \alpha_y \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} I_{jk} \cos \frac{\pi(2j+1)x}{2J} \cos \frac{\pi(2k+1)y}{2k} \tag{3}$$

After embedding the bits of secret data in the DCT coefficients, inverse two dimensional DCT is applied on the frame T to generate the frame I using the equation,

$$I_{jk} = \sum_{x=0}^{J-1} \sum_{y=0}^{K-1} \alpha_x \alpha_y T_{xy} \cos \frac{\pi(2j+1)x}{2J} \cos \frac{\pi(2k+1)y}{2k} \tag{4}$$

where

$$\alpha_x = \begin{cases} \frac{1}{\sqrt{J}} & , x = 0 \\ \sqrt{\frac{2}{J}} & , 1 \leq x \leq J-1 \end{cases}$$

and

$$\alpha_y = \begin{cases} \frac{1}{\sqrt{K}} & , y = 0 \\ \sqrt{\frac{2}{K}} & , 1 \leq y \leq K-1 \end{cases}$$

Here, I_{jk} represents the pixel value in the cell $'jk'$ (column j and row K) of the frame I. Further, T_{xy} represents the transform coefficient corresponds to the cell $'xy'$ (column x and row y) of the 2D-DCT matrix.

Rajesh and Shajin [98] utilized DCT coefficients of the frames in raw video streams to embed the secret information. The secret data embedding procedure includes the following

steps; 1. extraction of the frame from the video stream, 2. Dividing the frames into image blocks of size 8×8, 3. Applying 2D-DCT on each image block and, 4. Embedding the secret data in less significant DCT coefficients. The less significant coefficients are detected using a predefined threshold value. To improve the security of the secret data embedded in the DCT domain, Mumthas and Lijiya [89] introduced RSA encryption, random DNA encryption, and Huffman encoding along with two-dimensional DCT-based videos steganography. The secret message is encrypted using the RSA algorithm. And the encrypted secret message is subjected to random DNA encryption followed by compression. The cover frame is divided into blocks of size 8×8 and 2D-DCT is applied on each block. The compressed and encrypted secret message is embedded in the LSBs of the transform coefficients and Inverse DCT is applied on each block to generate the stego-frame.

Mstafa et al. [85] proposed a method for hiding the data in the DCT domain. Firstly, the frames of the cover video are converted into YUV color space. Then, two-dimensional DCT is applied to each plane of the YUV color space. The secret data is encoded using two error-correcting codes, BCH codes and Hamming codes. The encoded data is embedded in the DCT coefficients, except for the coefficients with zero frequencies. Obtained evaluation results show the proposed scheme achieved high embedding capacity with minimal visual distortion in the video. Moreover, the presented method is robust against the salt & pepper attack, Gaussian white noise, and the median filtering attack.

Suresh et al. [113] proposed a data hiding approach based on shuffling the data on least significant DCT Coefficients. Initially, the scene change detection technique is implemented to select the cover frames. The scene changes are detected by the inter-frame difference value. After selecting the cover frame, each color channel in the cover frame is subdivided into 64 sub-images and the DCT coefficient is computed for each sub-image. Among obtained 64 DCT coefficients, 8 least DCT coefficients are selected for hiding 1 pixel of the secret image data. A random sequence generation-based shuffling is implemented to embed each bit of secret data randomly in the obtained 8 least DCT coefficients of the cover frame. The proposed shuffling approach improves the security of the steganography method. A critical analysis of different methods along with the evaluation metrics and remarks for transform domain video steganography is outlined in Table 3.

3 Video steganography in compressed domain

The majority of video steganography methods proposed in the literature for data hiding in the raw domain are simple and easy to implement. But, they are more prone to various attacks, especially compression attacks. Furthermore, currently, videos in compressed form are preferred for storing as well as transmission over the internet. The compressed video requires less storage space compared to uncompressed video. And transferring the videos in compressed form over the internet is quicker and requires less bandwidth. In this context, the data hiding techniques in the compressed video domain have gained popularity in the last two decades. On the other hand, compression causes the removal of redundant video data and reduces the space for hiding more data.

Among various available video compression coding standards, MPEG-X and H.26X are the widely utilized methods in recent years. Specifically, H.264/AVC a.k.a MPEG-4 Part-10 video coding standard is the popular and predominant video codec used in the literature by researchers for hiding data in the compressed domain. The H.264 video codec has multiple novel features compared to its predecessors and some of the novel features are

Table 3 Critical analysis of video steganography methods in transform domain

Method	Secret	Preprocessing	Psnr	Imperceptibility	Robustness	Capacity	Remarks	Ref.
Lazy lifting wavelet transform 3-LSB	Text	Encryption	31.23	↑	→	12.5%	(+) Simple method with two layered security	[95]
DWT, DCT & LSB	Text	Encryption	73.21	↑	↑	1.8 KB	(-) Security not tested	[2]
DWT & LSB	Video	Encryption	35.23	↑	→	N/A	(+) Evaluated in multiple video formats	[49]
DWT & LSB	Video	Nil	69.38	↑	→	N/A	(+) Able to hide multiple videos inside single video with minimal visual distortion	[114]
DWT	Image	Nil	45.60	↑	↑	N/A	(+) Biometric image sets are hidden inside video	[74]
Adaptive method (DWT & DCT)	Text	Encryption & ECC	48.67 (DWT) 49.01 (DCT)	↑	↑	3.4% (DCT) 3.46% (DWT)	(+) Secure against state of art steganalysis attacks.	[87]
Adaptive method (DWT)	Image	Nil	64.13	↑	↑	4.64 KB	(+) Robust against MPEG-4 compression	[57]
Adaptive method (DWT)	Image	Nil	53.5	↑	↑	2.78 KB	(+)Robust against MPEG-4 compression (-)Computationally expensive	[104]

Table 3 (continued)

Method	Secret	Preprocessing	Psnr	Imperceptibility	Robustness	Capacity	Remarks	Ref.
Adaptive method (DWT)	Text	ECC	41.5	↑	↑	4.4%	(+) Robust against artificial noises and attacks	[82]
DWT & Histogram shifting	Text	Nil	48.84	↑	↑	168.96 KB	(+) Simple , efficient & reversible method	[121]
Lifting wavelet transform	Image	Encryption	72.90	↑	↑	68.7%	(-) Steganalytic security not tested	[112]
DWT in YUV color space	Image	Encryption	58.67	↑	↑	4.8%	(+)Evaluated in SD, HD and full HD videos (+)Robust against H.264 compression	[15]
DCT	Text	Nil	36.9	↑	↑	N/A	(-) Steganalytic security not tested	[98]
DCT	Text	Encryption & Huffman encoding	37.29	↑	↑	N/A	(+) Multilayer security using RSA and DNA encryption	[89]
DCT	Text	ECC	40.73	↑	↑	27.53%	(-) Steganalytic security not tested	[85]

Secret: type of secret data used for hiding. Preprocessing : Preprocessing applied on secret data before embedding for improving its security, ↑ : High , ↓ : Low

"multiple frames reference capability", "flexible macroblock ordering", Intra prediction in intraframe, etc. Generally, H.264 video codec consists of multiple groups of pictures (GOP). And each GOP contains the intra-coded frames (I-frame), predicted frames (P-frame), and bidirectional predicted frames (B-frame). The I-frame a.k.a keyframe is the one that is independently coded and the first frame of each GOP. The P-frame contains only the difference between the current frame and the preceding frame. The B-frame holds only the changes in the current frame from both the previous and following frames.

The video encoding procedure in H.264 codec is shown in Fig. 5. During encoding, the initial frame (which contains all the important data and is considered the I-frame) is divided into macroblocks where each macroblock consists of 16×16 pixels. The data compression process comprises various steps such as prediction, domain transformation, and encoding. The prediction utilizes the temporal and spatial redundancy in the video data. Prediction allows encoding the difference between the previously coded data and the predicted data. There are two types of prediction: Intra prediction and inter-prediction. Intra prediction generates the prediction of macroblocks based on previously coded data in the current frame while inter prediction generates the prediction based on the data in the previously coded frames. Motion estimation and motion compensation techniques are utilized to predict the frame. The difference between the prediction and the current macroblock is known as residual. The block of residuals is subjected to domain transformation using integer transform. DCT is the most commonly used integer transform. The block of transformed coefficients is quantized to minimize the precision of the coefficients.

The final step of encoding converts the various values (quantized DCT coefficients, data required by the decoder to reconstruct the prediction, other data about the video sequence, etc.) obtained in the previous steps and syntax elements to binary codes.

In the compressed domain, the data hiding is implemented in two ways; data hiding along with the video encoding procedure and data hiding in the encoded bit stream. The data hiding techniques along with the video encoding procedure utilize various syntax elements related to the video coding task for embedding the secret data. A general overview of data hiding in syntax elements of the compressed domain is presented in the Fig. 6. The data
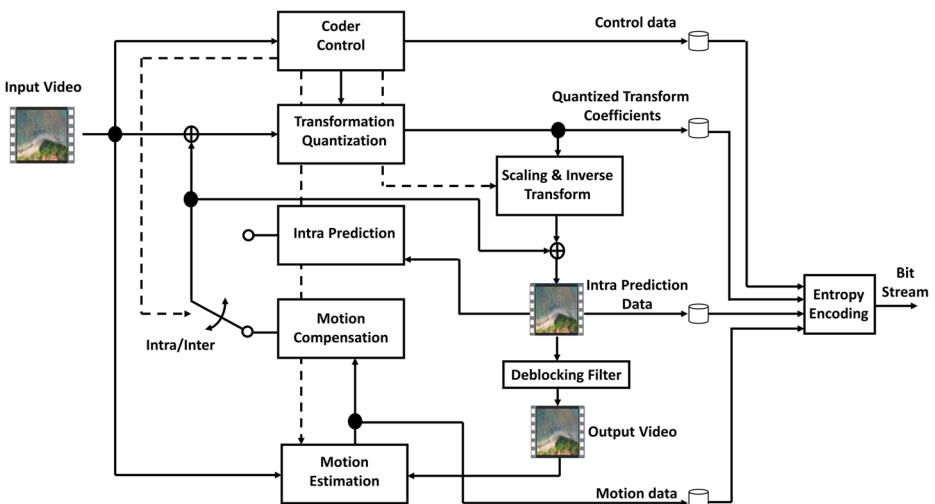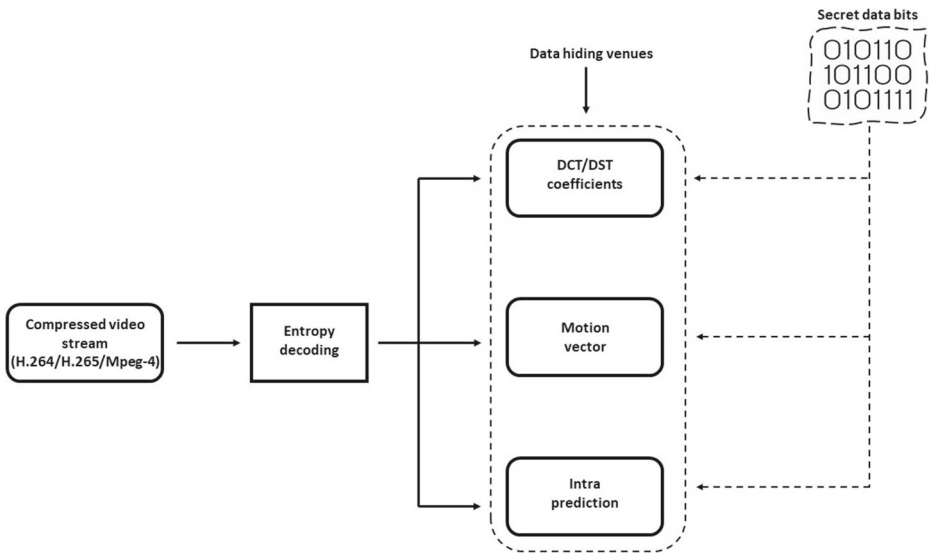

Fig. 5 H.264 video encoding

Fig. 6 An overview of data hiding in compressed domain

hiding in the encoded bit stream methods exploit the entropy coding modules to carry the secret data.

3.1 DCT/DST coefficients

In the literature, the quantized DCT coefficients obtained during the video encoding procedure have been utilized predominantly for hiding the secret data in H.264 videos. Generally, the secret data is embedded into the DCT coefficients of 4×4 luminance block of the cover frame (especially the I-frame). Intra-frame distortion drift is one of the main challenges faced while embedding secret data in the compressed video domain. In intra-frame prediction, the current prediction block will be the sum of residual values and predicted values. The predicted value is computed from the block samples of its neighboring encoded block. Suppose the neighboring encoded block is one of the venues for hiding the secret data, there should be distortion due to embedding the data. Since predicted values are computed from its neighboring blocks, this embedding-induced distortion in the previous block will propagate to the predicted block by intra-frame prediction. Most of the earlier works in literature have not considered the intra-frame distortion drift and experienced high visual distortion with less embedding capacity.

Ma et al. [76] proposed a novel method for hiding the secret data in quantized DCT coefficients with limited intra-frame distortion drift. The coefficients of 4× luma block in I-frames are utilized for embedding the secret data. After entropy decoding the H.264 bitstream, a pair of quantized DCT coefficients are selected from each 4× luma block to hide the secret data by controlling the intra-frame distortion drift. One of the DCT coefficients in the selected pair is used for embedding while another DCT coefficient is intended for compensating the intra-frame distortion drift. The correlation between DCT coefficients and distortion induced in the pixels (which are utilized in intra-frame prediction) is examined to select the appropriate pair of DCT coefficients. Based on [76], a method for data hiding

in quantized DCT coefficients without intra-frame distortion drift is proposed in [77]. Similar to [76], the proposed method utilized pair of DCT coefficients to embed the secret data as well as to accumulate the distortion drift. Moreover, the directions of intra-frame predictions are exploited to avert the distortion drift. Even though the approach of combining pair of DCT coefficients and the direction of intra-frame predictions can prevent the intra-frame distortion drift, the hiding capacity of the method was less. Only about 50 % capacity of the luminance blocks is utilized for data embedding.

To utilize the full capacity of luminance blocks for data hiding without any intra-frame distortion drift, a DCT-based perturbation method is proposed in [62]. In Ma et al. [77] approach, each 4*4 block is embedded with 3 bits of the secret data. Unlike [77], the proposed approach utilized quantized DCT coefficients of each $4\times$ block to embed 4 bits of secret data. Usually, embedding more data increase the visual distortion. To cope with the distortion induced in the video by increasing the amount of secret data hidden in each block, a DCT-based perturbation scheme is introduced. In the perturbation scheme, a new filtered 4×4 luma block is selected to hide the secret data by perturbing the related quantized DCT coefficients. The quantitative evaluation of the proposed DCT-based perturbation scheme shows that the embedding capacity is improved without compromising the visual quality. Reference [91] focused on further improving the hiding capacity of the quantized DCT coefficients-based method without harming the video quality. To embed the secret data and prevent the intra-frame distortion drift, the quantized DCT coefficients are classified into two different clusters. The first cluster is reserved to embed the secret data while the second cluster is utilized to prevent the intra-frame distortion drift. An embedding modification direction table is introduced to embed the secret data in DCT coefficients with minimal distortion. Initially, an embedding direction modification table for secret bits is created. Each element in the embedding direction modification table corresponds to each bit of the secret message. Then, the embedding modification direction value of the cluster reserved for embedding (embedding cluster) is calculated by a specific equation. After, the difference between the embedding direction modification value of the embedding cluster and decimal values of the n secret bits are computed to regulate the embedding and distortion.

In another work, Liu et al. [67] focused on improving the robustness of the data embedding approach [77] which uses quantized DCT paired coefficients and directions of intraframe prediction for preventing distortion drift. In that context, an error-correcting code 'BCH code' is used to encode the secret message before embedding it in quantized DCT coefficients. The robustness of the proposed approach is examined by exposing the secret data to re-encoding and re-quantization attacks. The obtained result of the quantitative examination shows that the robustness of the proposed method is improved by 25 % and 100 % of the secret message is recovered when the secret data is exposed to reconding attack. However, the recovery of the secret message bit is impossible if the frame drop happens. In [66], Liu et al. further improved the robustness of [77] to handle the frame loss problem. Shamir's (t,n)-threshold secret sharing is implemented to handle the secret data before embedding. The secret data is divided into the n-sub secrets with the help of Shamir's (t,n)-threshold secret sharing. The sub-secret are embedded in the quantized DCT coefficients of 4×4 luma blocks by following embedding conditions defined in [77] to prevent intra-frame distortion drift with improved robustness. Compared to [77], the implementation of Shamir's secret sharing for processing the secret message improved the survival rate by about 60% when the stego-video experienced frame loss. Reference [66] is further

extended in [65] to develop a robust reversible data hiding method without intra-frame distortion drift. The major contribution of the work is focused on recovering the original cover video completely after extracting the secret data.

In the last decade, most of the steganography techniques focused on embedding the data in quantized DCT coefficients have addressed the intra-frame distortion drift and proposed solutions to overcome this issue. To further improve the quality of the cover video, reduce the impact induced by embedding the secret data and increase the security of the data hiding approach, Syndrom trellis code (STC) is utilized by a few works in recent years [11, 122, 130]. Cao et al. [11] proposed a content-adaptive data hiding approach based on STC for H.264 videos. A new method called cover block decoupling is introduced to reduce the impact induced by the embedding. Two cover block decoupling strategies, passive strategy, and active strategy are presented in this work. The passive strategy is to select the non-referenced block (the blocks which are not referenced for Intra prediction) as the cover block for data embedding. Since H.264 coding contains only very few non-referenced blocks, choosing them alone for data hiding will severely affect the hiding capacity. To increase the capacity an active strategy that is focused on embedding the data in the first block of each macroblock and the rest of the blocks are utilized as the buffering zone to suppress the impact caused by the embedding. Most of the above-discussed methods in this section addressed mitigation of the inter-block distortion but not addressed the inner-block distortion. An STC-based method addressed both inter-block distortion and inner block distortion to further improve the video quality, as well as security [130]. The embedding of the secret data in DCT coefficients is based on three predefined strategies, 1. If the current block is not referenced for prediction, then all coefficients are used to embed the secret data independently, 2. If the pixels in the rightmost column or bottommost column of the current block are referenced for adjacent block prediction then paired coefficients are employed for embedding as well as compensating the distortion, 3. If the rightmost subblock and bottommost subblocks of the current block are referenced for the prediction of adjacent blocks, then four coefficients (one coefficient for embedding and the rest for compensating the distortion induced by the embedding) are selected for embedding and distortion compensation.

H.265/HEVC (high-efficiency video coding) is the most advanced and next-generation video coding standard which can provide more compression than H.264/AVC without affecting the video quality. H.265/HEVC encoding includes similar steps to H.264/AVC encoding. Analogs to macroblocks in H.264, a coding tree unit (CTU) of size up to 64×64 pixels is the basic coding unit in H.265/HEVC. The CTU can be further divided into multiple coding units of size 32×32 pixels, 16×16 pixels, 8×8 pixels, 4×4 pixels by following a quadtree structure. Furthermore, HEVC uses two different integer transforms which are based on DCT and DST to code the residual blocks. In the literature, few works utilized the quantized transform coefficients (DCT coefficients and DST coefficients) of intra-predicted residuals of H.265/HEVC bitstream to embed the secret data. Chang et al. [14] is the first method in the literature that utilized DST/DCT coefficient for information hiding in H.265/HEVC videos. The proposed method addresses the solution for intra-frame distortion drift as well as inter-frame distortion drift caused by embedding the secret message in DST/DCT coefficients. LSBs of the quantized transform coefficients are utilized for embedding the watermark data in H.265 video [116]. Only non-zero quantized transform coefficients are considered for embedding the watermark bits. The DST coefficients of the 4×4 luminance block are used for embedding the secret data in cite [70]. To handle the intra-frame distortion drift, the proposed approach introduced three conditions based on the

directions of intraframe prediction and used the multi-coefficient approach for data embedding. In the multi-coefficients approach, a triad of coefficients is used to handle embedding and distortion prevention. Among three coefficients, one coefficient is used for data embedding and the rest are utilized to reduce the distortion caused by the embedding of secret data. Liu et al.' s method [68] used 8×8 luminance quantized DCT coefficients for embedding the secret data. Two conditions based on the directions of the intraframe prediction and multi-coefficient-based embedding are employed for preventing the intra-frame distortion drift. Two types of multi-coefficient-based embedding approaches are used in the proposed work. The first type uses the combination of four coefficients where one coefficient is for embedding the secret data and the rest three are used to prevent distortion. The second type uses a pair of coefficients where one coefficient is reserved for embedding and the other one for compensating the distortion. Liu and Xu [71] proposed a robust steganography method for H.265 by utilizing the multi-coefficients of the selected 4×4 luminance DST blocks for data embedding. To enhance the robustness and security of the secret message, Shamir's (t, n)-threshold secret sharing is introduced to encode the secret message before embedding. Similar to their earlier work in [70], three conditions based on the directions of intraframe prediction and multi-coefficient approach-based data embedding are implemented to avert the intra-frame distortion drift.

3.2 Motion vectors

The motion vector is an essential syntax element in the process of video encoding and it is utilized for motion estimation as well as motion compensation to reduce the temporal redundancy. The motion compensation technique allows the prediction of the current block from previous or future blocks by accounting for the motion of objects in the frame. The motion vector is utilized widely for embedding the secret data in compressed videos. The earlier works in the literature considered embedding the secret data directly into the motion vectors. Most of the existing motion vector-based video steganography algorithms can be classified according to the modification approach employed in motion vectors for data embedding. Data embedding based on the modification of the magnitude of the motion vector and modification of the phase angle of the motion vector are the two approaches generally implemented in the literature. The data embedding approach based on the modification of the magnitude alters the magnitude value of the suitable motion vector by adding or subtracting 1 based on the secret data bits. The phase angle modification method converts the cartesian coordinate system to an imaginary system and considers each section of imaginary sections as 0 or 1. Based on the secret bit to be hidden, the selected motion vectors are rotated for data embedding.

Zhang et al. [137] and Xu et al. [126] considered some of the suitable motion vectors among all available motion vectors to embed the secret data bits. The suitable motion vectors are called candidate motion vectors and their magnitude is greater than a predefined threshold value. The phase angle of the motion vectors [23, 30] is utilized to hide the secret data bits. Initially, the candidate motion vectors are selected based on comparing the magnitude of the available motion vector with a predefined threshold value. Then the phase angle difference between successive motion vectors (candidate motion vectors) is utilized to embed the secret data bits.

A video steganography scheme based on the motion vector and matrix encoding is proposed in [59]. The proposed method is focused on hiding the data in macroblocks that are moving at high speed. The macroblocks moving at high speed were estimated based on the size of the motion vector. The motion vector of a macroblock with a higher amplitude than a

predefined threshold value is selected for embedding the data. The matrix encoding scheme is introduced to decrease the modification rate of the motion vector. Aly [5] et al. utilized the motion vectors used for encoding and decoding the P-Frame and B-Frame in the compressed video for hiding the secret data. Unlike [126, 137] which selected candidate motion vectors based on motion vector attributes such as angle or phase angle, the proposed method considered motion vectors that are associated with macroblocks of high prediction error as the candidate motion vectors. The data is hidden in both vertical and horizontal components of the candidate motion vector in P and B frames.

Cao et al. [135] proposed an adaptive approach for hiding the secret messages in MPEG-4 videos by utilizing the motion estimation process. A novel technique called perturbed motion estimation (PME) is implemented to estimate the motion as well as to embed the secret data. The novel technique, PME is inspired by Fridrich et al's perturbed quantization steganography based on the wet paper code. According to Fridrich et al's, the security of the video steganography method can be improved by embedding the secret data in adaptively selected components of the cover data without sharing the adaptive selection rule with the recipient. To achieve adaptive embedding without sharing the adaptive region selection rule, the wet paper code ("a simple variable-rate random linear code") is introduced. The proposed novel technique PME perturbs the motion estimation process associated with the encoding of some candidate macroblocks to hide the secret message. A selection rule based on the MSE value is introduced to select the candidate macroblocks. In traditional motion vector-based steganography methods, embedding the secret data in the motion vector induces shifting of the local optimal motion vector to non-optimal and thereby leaves the clues of data embedding. This issue will make the embedding methods vulnerable to motion vector-based steganalysis attacks. To improve the security of the motion vector-based steganography and robustness against the steganalysis systems, Cao et al. [12] introduced a data hiding scheme based on syndrome trellis code and uncertainty of motion estimation in H.264 videos. The syndrome trellis code is introduced to reduce the overall embedding impact while uncertainty in motion estimation is utilized to solve the distortion induced by the shifting of the motion vector from optimal to non-optimal. The proposed methods relate the embedding-induced distortion in a motion vector and its associated uncertainty to achieve higher security against steganalysis attacks. Yao et al. [133] proposed an effective method to enhance the security of the data hiding approach in the motion vectors associated with the process of H.264 encoding. A distortion function is defined to express the embedding impact on motion vectors. The distortion function is designed by considering the change in the prediction error due to modification of the motion vectors and change in the statistical distribution of the motion vectors. The proposed data embedding method consists of three stages. In the first stage, the embedding distortion for each motion vector in the cover frame is defined by using the designed distortion function. The second stage involves the data embedding procedure and modification of the motion vectors by limiting the changes. Two-layered syndrome trellis codes are introduced for data embedding to achieve minimal distortion steganography. And finally, the video is encoded using the modified motion vectors. The evaluation results show that the proposed approach has significantly improved the security of the motion vector-based data hiding approach. Zhang et al. [136] focused on preserving the local optimality of the modified motion vectors after embedding the secret data on them. By preserving the local optimality of the modified motion vectors, the traces or clues generated in the cover videos due to the data embedding can be minimized, and thereby the security of the steganography method can be improved. The proposed method considers a search area in the cover frame that consists of multiple

motion vectors. Before modifying a candidate motion vector, the local optimality of each candidate motion vector is evaluated and a candidate motion vector that contributes minimal degradation to the video encoding efficiency is selected for the modification.

A reversible data hiding algorithm using histogram shifting of motion vector values for H.264 videos is proposed in [127]. An error propagation control mechanism is employed to reduce the distortion induced by the modification of motion vectors. The motion vector values of the selected reference frame are not modified and the motion vectors associated with non-referenced frames are utilized to embed the secret data.

3.3 Intra prediction modes

Intra prediction in the video encoding procedure generates the prediction of macroblocks based on the previously coded data in the same frame. The video encoding process utilized multiple intra-prediction modes to encode the macroblocks. In H.264/AVC coding standard, there are 13 prediction modes, which include nine of 4×4 blocks and four of 16×16 blocks. The H.265/HEVC codec has 35 Intra prediction modes, where 33 are angular, and the rest are the planner and DC prediction modes. In the literature, the Intra prediction modes are utilized to embed the secret data in H.264 and H.265 videos by mapping the modes to secret data bits.

Hu et al. [33] proposed an intra-prediction mode-based data hiding scheme in H.264/AVC videos. The proposed method is configured to embed 1 bit of secret data in each 4×4 luma block by altering the 4×4 intra-prediction modes. A predefined rule is implemented to select the candidate luma blocks for embedding the secret data. The candidate luma blocks are modified to hide 1 bit of secret data based on a predefined mapping algorithm between the secret data bit and intra-prediction mode. The prediction difference of intra 4×4 blocks is utilized for data hiding in H.264 videos [140]. The proposed method initially groups the set of 4×4 prediction modes into two disjoint subsets (let it be A and B) of prediction modes based on the prediction difference of intra 4×4 blocks. The mapping rule between secret binary bits and prediction mode is given as,

$$\mod e \in \begin{cases} A & if\, s_i = 0 \\ B & if\, s_i = 1 \end{cases} \tag{5}$$

Where $\mod e$ is the coding mode of the current block and s_i is the hidden data bit. Further, logistic mapping rules are implemented to randomly choose the hidden location and improve the security of the proposed method.

Yang et al. [132] used the intra-prediction modes and matrix coding to hide the secret data in H.264/AVC video stream. Here, a 4×4 block is selected for embedding the secret data only if the respective block is a candidate block based on certain predefined rules and all of the 4×4 blocks within a 16×16 block are of different prediction modes. To implement secret data-Intra prediction mapping, two secret data bits are mapped to every three 4×4 blocks by matrix coding.

Zhang et al. [139] proposed an adaptive video steganography algorithm based on intra-prediction mode. The proposed scheme hides the secret data in selected 4×4 blocks altering the prediction mode based on the predefined mapping rules between secret data bit and intra-prediction mode. Unlike other intra-prediction mode-based methods which select hostable blocks/candidate blocks using the different scrambling algorithms, the proposed method introduced the Syndrome trellis code-based algorithm to choose the regions with more complex textures for data embedding.

3.4 Entropy coding modules

In the video encoding process, the entropy coding method is implemented to encode various parameters obtained in the previous stages of the encoding process including syntax elements, coded block patterns, motion vectors, residual data, reference frame index, etc. Context-based adaptive binary arithmetic coding (CABAC) and context-adaptive variable-length coding (CAVLC) are the two entropy coding schemes implemented in various video encoding standards including H.264/AVC and H.265/HEVC. The property of the entropy coding schemes is utilized in the literature for concealing secret data in compressed video domains.

Liao et al. [61] proposed an entropy coding scheme-based data hiding technique in h.264/AVC videos. The proposed method achieves data hiding by embedding the secret data bits in the trailing ones of 4×4 blocks during the CAVLC procedure. To choose block positions for embedding the secret data, chaotic map-based random numbers are generated initially. After blocks are chosen, the secret information bits are hidden in the trailing ones The proposed method achieved acceptable imperceptibility with low embedding capacity. Ke and Weidong [44] utilized the property of CAVLC for data hiding in H.264 videos. The data embedding is implemented by altering the sign flag of the trailing one and different levels' codeword parity flag in CAVLC. The proposed approach considers the trailing ones of non-zero coefficients with high frequency in luma components for data hiding. Based on the secret data bit the trailing ones are modified where even parity is given to codeword if the secret bit is '0' and odd parity if the secret bit is '1'.

Zhang et al. [138] proposed a data hiding scheme based on the trailing coefficients obtained in each DCT transform block during H.264 encoding process. The proposed approach applies DCT transform on frames of the video sequence to obtain DCT coefficients and scans the DCT blocks in order. Among the arranged blocks, odd number blocks are selected for embedding the secret data and even-numbered blocks are used as correcting blocks. In the odd blocks, the value of the trailing coefficient is negative when the secret bit is 0 and the value is positive when the secret bit is 1. The proposed method displayed acceptable robustness against various noise attacks.

Xu et al. [128] utilized codeword substitution to embed the secret data in encrypted streams of H.264/AVC standard video sequences. The sensitive parts of compressed video sequences including motion vector differences, intra-prediction modes, and residual coefficients are encrypted using the stream cipher. The code word substitution approach named bin string substitution technique is employed to hide the secret data bit in the encrypted domain. The proposed method managed to maintain the same bitrate even after encryption as well as data hiding. Reference [129] presented an improved version of data hiding in the encrypted stream [128]. The proposed method mainly focused on increasing the embedding capacity of [128] without affecting the visual quality. In [128] only the code-word of the level whose suffix Length is 2 or 3 is utilized for data embedding by single codeword substitution. But in the proposed method code word of level with suffix length 1 is also used. Paired codeword substitution is implemented for data embedding when the suffix length is 1. And for suffix lengths greater than 2, the multiple-based notational system is adopted instead of single code-word substitution to achieve data embedding. Critical analyses of different methods proposed in the compressed domain are given in Table 4.

Table 4 Critical analysis of video steganography methods in compressed domain

Method	Preprocessing	Psnr	Imperceptibility	Robustness	Security	Capacity	Remarks	Ref.
4× DCT	Nil	40.74	↑	↑	×	0.11 kb / I frame	(+) Acceptable embedding capacity with less visual distortion	[77]
4× DCT	Nil	42.21	↑	↑	×	54% of all luma blocks	(+)Improved the embedding capacity of [77] without error propagation	[62]
4× Integer DCT	ECC	41.71	↑	↑	×	1.17 KB/ 20 I-frames	(+) Introduction of BCH codes for encoding secret data improved the robustness as well as security	[67]
4× Integer DCT	Secret sharing	36.53	↑	↑	×	0.017 KB/ 20 I-frames	(+) Addressed the recovery of secret message when frame drop happens	[66]
Content adaptive embedding (QDCT coefficients)	Nil	39.98	↑	↑	✓	0.12 KB/I-frame	(+) First method in the literature for content-adaptive embedding in H.264 streams	[11]
4× DCT	Nil	36.92	↑	↑	✓	0.40 KB/I-frame	(+) Able to resist state of art steganalysis methods	[130]
4× luma DST & prediction modes	Secret sharing	35.7	↑	↑	×	1.05 KB	(+)Achieved high robustness by introducing Shamir's secret sharing method	[71]

Table 4 (continued)

Method	Preprocessing	Psnr	Imperceptibility	Robustness	Security	Capacity	Remarks	Ref.
Watermarking (motion vector components)	Nil	43.08	↑	↑	×	N/A	(-) Less embedding capacity	[137]
Motion vector's phase angle	Nil	53.29	↑	↑	×	8.2 %	(+) Blind approach with less computational time	[23]
Motion vector's phase angle difference	Nil	37.5	↑	↑	×	0.018 KB	(+) Embedding efficiency is improved by reducing the modification rate of MVs	[30]
Motion vector components	Nil	38.41	↑	↑	×	N/A	(+) Modification rate of motion vector is reduced and because of that imperceptibility increased	[59]
Motion vector and STC	Nil	35.47	↑	↑	✓	3 bits/MV	(+) Significant improvement in the security of the algorithm compared to related state of art literature	[133]
Modification of motion vector with preserved local optimality	Nil	38.78	↑	↑	✓	N/A	(+) High steganalytic security	[136]
Intraprediction mode & matrix coding	Encryption & Scrambling	N/A	↑	↑	×	N/A	(+) Minimal embedding distortion with slight increase in bitrate	[132]

Secret: type of secret data used for hiding. Preprocessing : Preprocessing applied on secret data before embedding for improving its security, ↑ : High , ↓ : Low , Security: Whether steganalytic security is verified or not

4 Features of video steganography

Evaluation of video steganography methods is important to determine the performance and efficiency of the method. The main features expected from good steganography methods are imperceptibility, hiding capacity, security, robustness, and resistance to other steganalysis attacks. In this section, each feature of the video steganography method is elaborated on in detail.

4.1 Imperceptibility

Imperceptibility is the capability of the method to hide secret information that is not visible to the human eyes. Humans should not be able to interpret the secret information hidden. This measure is more related to the visibility of the resulting videos. Higher imperceptibility means lower distortion and higher visual quality of the stego video. Many methods have displayed the constructed stego video results and the extracted secret image results to show the imperceptibility of the method.

Different evaluation metrics are used for measuring the imperceptibility of the method. The most commonly used measures are Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and Structural similarity index matrix (SSIM). MSE is calculated by taking the mean of the error between the input and output frames. PSNR is calculated with MSE values as the base. Equation (6) and (7) shows the formula used for calculating MSE and PSNR respectively. MSE and PSNR values are measured in decibels (dB) and are widely used because of their simplicity.

$$MSE = \frac{\sum_{R,C}[I_1(r,c) - I_2(r,c)]}{R * C} \tag{6}$$

$$PSNR = 10 * \log_{10} \frac{E_2}{MSE} \tag{7}$$

SSIM is another measure used in contemplating the cognitive degradation between images caused during compression or reconstruction. Loss between two instances of the images is calculated by keeping one of the images as a reference and compared against the processed image. In the case of video steganography, the input images are the reference and the reconstructed stego image and the extracted secret image are the processed image. The formula for calculating the SSIM value is given in (8). PSNR is the pixel-level difference between two images, however, SSIM calculates the visual difference between the two images. SSIM is considered a better metric for image degradation methods when compared with PSNR.

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1) + (2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{8}$$

4.2 Hiding capacity

Capacity is a metric to measure the amount of secret media that is hidden inside the cover image with minimum distortion. Hiding capacity is also called embedding capacity as it directly refers to the amount of secret information that can be embedded. Hiding capacity is calculated by dividing the amount of secret information by the total size of the cover video.

Equation (9) shows the formula to calculate the hiding capacity of the video steganography. Capacity is given in terms of bits per pixel (bpp) which indicates the number of secret bits that can be hidden inside each pixel of the cover frame.

$$C_H = \frac{C_s}{C_c} \tag{9}$$

C_H is the hiding capacity of the embedding algorithm, C_s is the size of the secret information, and C_c is the total size of the cover video.

Entropy is another measure used to calculate the embedding capacity of the method. It measures the total amount of information carried by the video, taking into account the information density. It also measures the randomness of the video and the values range between 0 and 1. Equation for calculating the entropy value is given in (10). M is the total intensity and N is the probability that a particular intensity will happen.

$$E = \sum_{m=1}^{M} N_m log_2(N_m) \tag{10}$$

4.3 Robustness

Robustness refers to the extent to which the secret media is embedded and retrieved without any loss of information. The secret information should be communicated across the users without any loss. The robustness of the steganography method can be measured using its resistance against different noise attacks. Many methods have subjected their steganography results to these noise attacks and measured the resistance of their algorithms against these attacks.

It is common to transfer the stego video through untrusted channels like the internet, wi-fi, and satellite. When being transferred, the image can be degraded because of the inclusion of noises through external disturbance. Many methods have tested the robustness of their method by adding noise to the stego videos and checking the security. Image noises with different levels of distortion densities are included and tested. In general, four different types of image noises are considered for testing, namely, Gaussian, salt and pepper, speckle, and periodic noises.

Salt and pepper noise is the common inclusion added to the image for testing the proposed method against steganalysis attacks. This noise is also called impulse valued noise, intensity spikes, and bipolar impulse noise. Salt and pepper noise happens when the original value of the pixels is replaced statistically with corrupted values. Salt and pepper noises are prone to occur during transmission of the video, malfunctioning of the camera sensors and memory. Even with the alterations of the original pixel values, the appearance of the images does not change.

Gaussian noises are the noises based on the Gaussian distribution function. Gaussian noise is also called statistical noise and is influenced by the probability density function and the normal distribution. Gaussian noises are caused naturally because of the image fluctuations. The main cause for Gaussian noise is during the image acquisition process like the faults in the sensors, changes in illumination, peak temperatures, and other electronic circuit noises. Spatial filters are used to smooth the Gaussian noises but they may affect the image quality due to the blurring of the edges. Gaussian noises can be modeled easily in images by replacing the original values with random values produced by a mathematical model.

Apart from the salt and pepper, Gaussian noises, speckle noises, and periodic noises are the other forms of noises available. Speckle noises are unwanted changes to the image signals caused by uneven changes in the scattering surface. Speckle noises are common in Synthetic Aperture Radar imaging, and laser and acoustic images. Speckle noises are multiplicative and exist in granular patterns in the form of artifacts, blurry edges and corners, and disturbing backgrounds. Speckle noise can be modeled by multiplying the original image pixel values with random values.

Periodic noises happen due to the electro-mechanical or electrical disturbance that occurs during the image acquisition process. Images affected with periodic noises resemble as if a layer of repeated image spikes are added to the original image. Passing the images with periodic noise through frequency domain filters can reduce the noise considerably. However, the level and type of frequency filter depend on the application. From the video steganography perspective, these noises are modeled mathematically and introduced on the stego image. The efficiency of the proposed steganalysis method is determined by its ability to detect and decode the secret message in the noise-infused stego image.

Another metric to measure the robustness of the embedding is the Bit-Error Rate (BER). BER is the metric used to measure the amount of distortions on an image during manipulation. Bit-Error rate and the Signal Noise Ratio (SNR) are inversely proportional. High values of SNR indicate higher similarity between the video transmitted and the video received. However, BER will be less when the SNR is high. The equation for calculating the BER is given in (11).

$$BER = \frac{\sum_{R,C}[I_1(r,c) \oplus I_2(r,c)]}{R * C} \qquad (11)$$

5 Steganalysis: an overview

While steganography is the process of hiding secret information, steganalysis is the process of breaking the steganography algorithm to detect and uncover the secret information embedded. Steganalysis is classified into active steganalysis and passive steganalysis. Passive steganalysis detects the presence of the secret information alone, but active steganalysis detects and decodes/modifies the hidden secret information.

Steganalysis is important for two reasons; one for reversible steganography where steganalysis is required at the receiving end to extract the secret information embedded. Another reason for using steganalysis is to prevent the transfer of illegal information. There are numerous steganography tools easily available online which makes the need for steganalysis crucial. Even a layman can easily access the steganography tools and use them for sending confidential and prohibited information without raising any suspicions to government officials.

The need for steganalysis is substantial, however, steganalysis is not easy. With the proper selection of cover media, even the best steganalysis tool may not be able to break the steganography. Since any details about the cover media is not available and without that knowledge, breaking the steganography is difficult. Especially with video as the cover media, steganalysis can be challenging as the correlation between the videos and the features is difficult. Steganalysis methods are developed under the notation that the characteristics and features of the cover media are modified when the secret information is ingrained. Many steganalysis methods work by comparing the features and other characteristics between the stego object and the cover object.

A thorough analysis of the existing steganography method and its advancement is required to formulate the steganalysis tool. A good steganalysis method should be able to break different steganography attacks and some of the existing steganalysis methods. Steganalysis is divided into two types, namely, specific steganalysis and universal steganalysis. Specific steganalysis methods are developed to deal with a particular type of steganography method. For example, reversible steganography is a specific steganalysis method, since this method can only break a particular steganography method and may not be able to work efficiently with other steganography methods. On the other hand, universal steganalysis methods aim at breaking all steganography methods. Specific steganalysis methods are possible, whereas, universal steganalysis methods are difficult to implement. The development of universal steganalysis methods should be in the direction of future works. Efforts for developing a single software that is capable of breaking any type of steganography should be invested.

5.1 Steganalysis techniques

Steganalysis methods are broadly classified into three types - signature steganalysis, statistical steganalysis, and feature-based steganalysis. Signature steganalysis, as the name suggests uses the signature left behind by the embedding method to detect the presence of the secret information. Statistical steganalysis uses statistical methods and mathematical formulas to detect and uncover secret information. Feature-based steganalysis methods extract features from the cover video and stego video for investigating the presence and thereby uncovering the secret information. A detailed branch diagram of the steganalysis techniques and their sub-branches is given in Fig. 7.

5.1.1 Signature steganalysis

Further, signature steganalysis methods are divided into visual attacks and structural attacks. The first and foremost feature expected from video steganography is imperceptibility. The distortion caused by the video steganography method has to be minimum, or else the traces of the hidden message may become visible to the Human Visual System (HVS). Visual attacks are the simple steganalysis technique that can break the steganography using the HVS. A stego frame and the cover frame are compared side-by-side with the naked eye to check for any visible changes [104, 125]. Though visual attacks are easy to implement, they
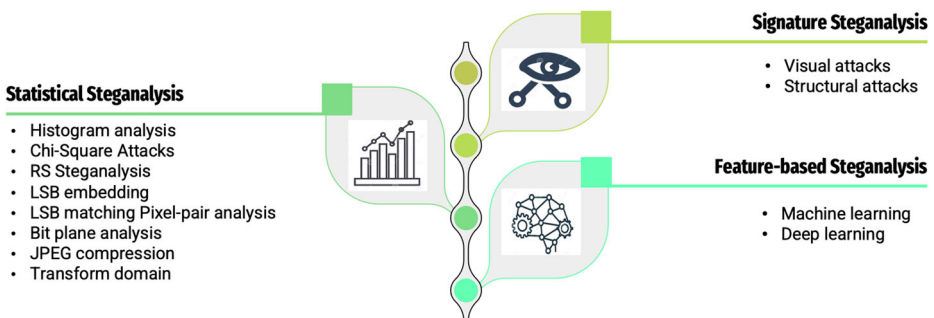

Fig. 7 Classification of different steganalysis techniques

are not reliable. Not only reliability but also the automation and the requirement of experts to perform the testing are other disadvantages of using the visual attack.

The characteristics and features of the cover video change after embedding the secret information. Some of the characteristics, features and other structural components of the stego video are taken into account to detect the presence of the secret information. This type of steganalysis attack is called the structural attack. One example is the file size comparison [52]. After embedding, the file size of the cover video is prone to changes. Similar to visual attacks, structural attacks are not reliable and experts in the domain are required. Not all tampered videos will undergo structural changes and can escape the structural attack methods.

5.1.2 Statistical steganalysis

Statistical steganalysis utilizes the values of the image pixels and analysis them for detecting the confidential content. Statistical steganalysis is pre-eminent compared to signature steganalysis. Statistical methods use the knowledge of the image pixel values and mathematical models to detect and recover the secret information. Statistical steganalysis can be grouped into Histogram analysis, Chi-Square Attacks, RS Steganalysis, LSB embedding, LSB matching Pixel-pair analysis, Bit plane analysis, JPEG compression and transform domain steganalysis.

Histogram steganalysis analysis the histogram of the cover video frame and the stego video frame to detect the presence of secret information. The histogram is a graphical representation of the pixel values of the image based on the distribution. When a cover video is manipulated to embed the secret information, the histogram of the stego video is affected, The embedding of the secret information may not be visible to the human eyes, but when the histogram is plotted and compared against the original cover video, even the slightest manipulation can be detected [32, 41, 52, 64], and [31].

The Chi-square test is a common steganalysis technique used to detect the presence of secret information. This test works by observing the similarity between the real-time event and the expected outcome. It uses the frequency distribution to determine the randomness in the videos. Lower values of the test indicate a higher degree of randomness, confirming the presence of the secret message. Higher values mean a lower degree of randomness and prove there is no tampering in the video [52, 52].

RS steganalysis is another powerful tool introduced by Jiri Fridrich et al. [25]. RS analysis is used to detect the secret information that was embedded using the LSB-based methods. It compares the pixel values of the image in the spatial domain. The selection of pixel pairs varies based on the method, sometimes the neighboring pixels are chosen and other times pixels from different blocks are chosen for comparison. These groups of pixels are called Singular groups (S) and Regular groups (R). The presence of the secret is determined by grouping the pixels based on the frequency distribution and analyzing the LSBs of the stego and the cover video. The LSBs are flipped and randomized to detect the secret message [75, 102]. RS analysis has better reliability compared to chi-square tests [106].

LSB embedding and LSB matching steganalysis methods are based on the working principle of the LSB steganography method. These steganalysis methods are popular since the usage of LSB steganography methods is wide compared to other steganography methods [92]. Transform domain steganalysis converts the image into the frequency domain with magnitude and phase. Magnitude represents the frequency count values of the image and phase represents the direction to restore the image to its original form. The commonly used approaches for transform domain are Wavelet, Fourier, and Cosine Transform [106].

5.1.3 Feature-based steganalysis

Features are an important part of an image. Feature-based steganalysis methods extract the features from the images and analyze the features to detect the presence of secret information. These features can be further used in training a classifier to automatically detect the secret information using machine learning algorithms [106]. Pixel value differencing (PVD) steganography methods hide more bits of the secret information in the smoother regions of the cover image than in the complex regions. Histogram analysis of the PVD steganography revealed a Laplace distribution. A feature-based steganalysis method is used to detect the presence of secret information. Since PVD has the Laplace distribution, the expected frequency distribution of the image is obtained using any randomness test. The expected values are compared against the observed values and the degree of similarity is calculated. If the similarity is below a certain threshold, then the image has not been tampered with, else it has some embedded information [63].

6 Discussion, challenges and future directions

The video steganography or data hiding in video sequences can be achieved in multiple ways. This work discussed various data hiding approaches proposed in the last two decades. The data hiding approaches are classified based on the data hiding venue used in each method. Among those discussed approaches, LSB substitution in the spatial domain is the simple, easy as well as predominant method employed by the researchers in the literature. Later complex methods are introduced to further enhance the performance of data hiding. However, there are many challenges to developing a precise steganography system. Further steganography is a fast-evolving field in the information security domain. Here we listed a few challenges as well as future directions for video steganography.

- Most of the LSB-based methods discussed in the work have displayed acceptable imperceptibility with high data hiding capacity. However, the LSB-based methods are not robust enough to withstand various attacks and noises. Moreover, the state of art steganalysis methods can easily detect the LSB substitution-based modification in video sequences. Because in LSB-based approaches the modification is made directly on the raw pixel values of the frames.
- Transforming the raw pixel values in the spatial domain to the frequency domain and hiding the secret data in the transformed frequency coefficients have displayed enhancement in security as well as robustness. DCT and DWT are two commonly used transform functions in steganography approaches for embedding in the transform domain. Different levels of DWT (first, second, third, etc..) were performed on the cover medium in a few methods to improve the robustness as well as security. However, applying multiple levels of DWT on the cover medium reduces the embedding capacity. In the future, the researchers can explore the effectiveness of other wavelet transforms other than DCT and DWT for data embedding. Reference [109] displayed the effectiveness of CvT over DCT and DWT. To the best of our knowledge, no other methods in the literature have used CvT function for embedding in transform coefficients.
- In raw domain steganography, instead of serially selecting the pixel values or transform coefficients for embedding the secret data bits, the researchers employed game theory, genetic algorithm, or random number generator for random selection of pixels

or transform coefficients. To an extent, the random selection of embedding venues has improved the security of the implemented method.

- Adaptive steganography methods have been implemented in the literature for improved robustness and security. The adaptive methods employed certain artificial intelligence algorithms for detecting moving pixels and the secret data is hidden in the raw moving pixels or transformed coefficients of moving pixels. Furthermore, the skin lesions available in the video frames are utilized as the venue for data embedding. Edges of the objects available in the cover frame are also a suitable venue for data embedding. The existing methods [74, 82, 87, 104] have utilized conventional methods for detecting moving objects or skin lesions. The latest deep convolutional neural network-based methods can be employed in the future for effectively detecting moving objects and skin lesions. Moreover, video summarization techniques [54–56] can be utilized to identify key frames, and later the detected keyframes can be used as the data hiding venue.

- The basic features of the steganography method are higher imperceptibility, higher security and robustness, and higher embedding capacity. But, it is not possible practically to achieve all the features. A threshold for the trade-off between imperceptibility, security, robustness, and capacity should be developed for practical use [28, 38] and [111]. To achieve better hiding capacity, more secret bits are embedded in the cover media, which may lead to the exposure of the presence of the secret media. Increasing the hiding capacity has compromised security, robustness, and imperceptibility. Based on the application scenario, which feature of the steganography algorithm can be compromised for the betterment of the other features can be decided.

- Artificial intelligence is a standard framework acclaimed in many computer vision and other multimedia applications. Steganography is an image/video reconstruction technique where the main goal is to reconstruct a stego frame which is a combination of the cover and secret media. Artificial intelligence methods like machine learning , neural networks are extensively used in image steganography and have proved to improve the imperceptibility, robustness, and computational cost. AI methods are optimal and provide adaptive solutions. However, AI methods are used in the field of image steganography method [110, 111]and only a handful of research [35, 79, 123] has focused on video steganography. Due to the generality of image steganography, the AI technology applied to image steganography has great reference value for video steganography.

- General Adversarial Networks (GAN) are powerful artificial intelligence network used in image reconstruction field [27]. Image steganography using GAN is popular and has achieved a greater performance with increased security, robustness, and capacity [119, 120], and [58]. Using GAN for video steganography is a field that is still not explored. More studies focusing on utilizing the GAN architectures for hiding secret information inside cover videos can be concentrated. Another method that can be explored is the coverless steganography where the cover object is generated or selected from the database based on the secret information [73] and [19]. Coverless steganography has added advantage as there is no need to transfer the original cover object for steganalysis.

- Robustness of the steganography method is measured by analyzing the resistance of the method against different noise attacks, compression attacks, and video/image manipulation attacks. The security of the method is measured by evaluating the resistance against steganalysis attacks. The proposed steganography method is subjected to these robustness attacks and against certain famous steganalysis techniques. The results are

reported and analysed to measure the robustness and security of the steganography method [2, 57, 67, 88] and [137]. However, there is no assurance the proposed method will be resistant to all possible robustness and steganalysis attacks. Moreover, many video steganography methods have not reported their resistance against robustness and security attacks. In that case, it makes it difficult to honestly judge the efficiency of the steganography method. There is no unified metric to compare the performance of the different methods. The evaluation metric used is based on the convenience of the authors, as there is no established evaluation metric for steganography.

- In the literature, several methods [67, 81, 85, 89] (including both row domain-based methods and compressed domain-based methods) leveraged the merits of cryptography and error-correcting codes for enhancing security as well as robustness. Encrypting the secret data before embedding is widely adopted for securing the secret data in video steganography methods. Integration of encryption schemes provides an additional security layer. However, the integration of encryption along with steganography will make the whole encoding and decoding process time-consuming. And these approaches are computationally expensive compared to the methods which just implemented the steganography method alone. Most of the ensemble methods proposed in the literature have not addressed or evaluated the time consumption issue. To achieve real-time encoding and decoding, future works can be focused on the parallelization of encryption and steganography methods. Or efficient secret sharing schemes [26, 51] can be integrated with steganography.

- The main consideration in video steganography is choosing a proper cover media. Signature steganalysis methods use the visible changes in the video to detect the presence of secret information. File size comparison is a technique used in signature steganalysis to compare the size of the cover video and the stego video. Bitrate is a measure to check the change in the bit rate after embedding. It is the difference between the bit rate before embedding and after embedding. It is not possible to maintain the same file size after embedding by all the steganography methods. The optimal prediction of the original cover video is destroyed after embedding. This increases the bitrate. The bitrate increase is more with fast motion videos and complex textural videos [93]. The changes in the bitrate can thus be controlled by carefully selecting the cover video.

- Reversible steganography is a steganography method where a steganalysis method to break the steganography method is developed alongside [20]. It has a sender-receiver kind of architecture with a steganography algorithm placed at the sender side and steganalysis at the receiving end. Reversible steganography is common in image steganography, however, video steganography methods do not focus on the steganalysis algorithm design. More focus is given to video steganography methods only, which makes the implementation of the sender-receiver architecture for videos. Video steganalysis methods are still not explored extensively. More attention to video steganalysis methods, and reversible video steganography can be given to developing better techniques.

- Specific steganalysis methods are easy to develop and universal methods are tough to develop. Universal steganalysis methods should be able to break the presence of secret information embedded using any kind of steganography method without the knowledge of the technique used. More universal steganalysis methods can be designed where one steganalysis method can work universally to detect the stego video irrespective of the steganography method with less computational work [92].

- Any digital video sequence can be used as the dataset for hiding the secret data and evaluating the proposed method. Most of the existing video steganography methods have used public video datasets available for different computer vision, multimedia, and machine learning tasks for their evaluation. Among those video trace library [107] , UCI [18], YFCC100M [118], and PETs 2009 [24] are the popular datasets used in video steganography tasks. The video steganography research domain still lacks a unified large dataset that is particularly developed for the evaluation of the data embedding problems.
- Encrypting the secret information is one way of adding an additional layer of security. Another way can be to perform bifold steganography where the steganography method is applied twice. Once to create the stego object from the original cover and secret information. The same steganography method can be applied once more on the stego object to provide additional security. The feasibility of using the bifold steganography can be studied. The computational cost and time, the imperceptibility, security and robustness, and the practical use of the method can be reported.
- The steganography methods can be classified into blind and non-blind methods based on the information required to decode the secret data. The blind methods do not require original cover media for the decoding task. On the other hand, the non-blind method requires original cover media for decoding the secret data. Thus non-blind methods have to transmit the original cover media along with the cover media holding secret data to the receiver side. Sending multiple copies of cover media makes the attacker suspicious and may expose the existence of secret data. In this context, blind steganography methods are more secure compared to non-blind methods.
- Steganography is a very powerful tool that can help in communicating secret confidential information between parties. With the advancements, these tools can be easily exploited by terrorists and other anti-government bodies. Government should provide more regulations and restrictions on these kinds of tools to prevent them from falling into the hands of people with wrongful intentions.

7 Conclusion

This work provides a comprehensive summary of recent improvements in video steganography systems. Based on the data hiding venues, we classified the existing methods into different categories. Along with the detailed explanation of various video steganography methods proposed in the last decades, this work also looked at its benefits and drawbacks. Further, this work discussed various features to be considered while designing an efficient and effective data hiding method. A brief introduction to steganalysis techniques is also provided in this work. The article is concluded by discussing various challenges and potential research directions for future research in video steganography.

Declarations

Conflict of Interests The authors declare no conflict of interest.

References

1. Abbas SA, El Arif TI, Ghaleb FF et al (2015) Optimized video steganography using cuckoo search algorithm. In: 2015 IEEE Seventh International Conference on Intelligent Computing and Information Systems (ICICIS). IEEE, pp 572–577
2. Ahmed EA, Soliman HH, Mostafa HE (2014) Information hiding in video files using frequency domain. Int J Sci Res (IJSR) 3(6):2431–2437
3. Al-Khater WA, Al-Maadeed S, Ahmed AA et al (2020) Comprehensive review of cybercrime detection techniques. IEEE Access 8:137,293–137,311
4. Alavianmehr MA, Rezaei M, Helfroush MS et al (2012) A lossless data hiding scheme on video raw data robust against h. 264/avc compression. In: 2012 2nd International econference on computer and knowledge engineering (ICCKE). IEEE, pp 194–198
5. Aly HA (2010) Data hiding in motion vectors of compressed video based on their associated prediction error. IEEE Trans Inf Forensics Security 6(1):14–18
6. Balu S, Babu CNK, Amudha K (2019) Secure and efficient data transmission by video steganography in medical imaging system. Clust Comput 22(2):4057–4063
7. Banik BG (2019) Exploring recent advances in digital video steganography and future scope. Intell Innov Multimed Data Eng Manag 88–115
8. Battisti F, Carli M, Neri A et al (2006) A generalized fibonacci lsb data hiding technique. In: 3rd International conference on computers and devices for communication (CODEC-06), institute of radio physics and electronics, University of Calcutta
9. Bhawna KS, Singh V et al (2021) Information hiding techniques for cryptography and steganography. In: Singh V, Asari VK, Kumar S (eds) Computational methods and data engineering. Springer Singapore, Singapore, pp 511–527
10. Böhme R (2010) Principles of modern steganography and steganalysis. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 11–77. https://doi.org/10.1007/978-3-642-14313-7_2
11. Cao Y, Wang Y, Zhao X et al (2018) Cover block decoupling for content-adaptive h. 264 steganography. In: Proceedings of the 6th ACM workshop on information hiding and multimedia security, pp 23–30
12. Cao Y, Zhang H, Zhao X et al (2014) Covert communication by compressed videos exploiting the uncertainty of motion estimation. IEEE Commun Lett 19(2):203–206
13. Cetin O, Akar F, Ozcerit A et al (2012) A blind steganography method based on histograms on video files. Imaging Sci J 60(2):75–82
14. Chang PC, Chung KL, Chen JJ et al (2014) A dct/dst-based error propagation-free data hiding algorithm for hevc intra-coded frames. J Vis Commun Image Represent 25(2):239–253
15. Dalal M, Juneja M (2019) A robust and imperceptible steganography technique for sd and hd videos. Multimed Tools Appl 78(5):5769–5789
16. Dalal M, Juneja M (2020) Evaluation of orthogonal and biorthogonal wavelets for video steganography. Inf Security J Global Perspect 29(1):40–50
17. Dasgupta K, Mandal J, Dutta P (2012) Hash based least significant bit technique for video steganography (hlsb). Int J Security Privacy Trust Manag (IJSPTM) 1(2):1–11
18. Dua D, Graff C (2017) UCI machine learning repository. http://archive.ics.uci.edu/ml. Accessed 30 June 2021
19. Duan X (2018) Coverless steganography for digital images based on a generative model. Comput Mater Continua 55(3):483–493
20. Duan X, Jia K, Li B et al (2019) Reversible image steganography scheme based on a u-net structure. IEEE Access 7:9314–9323

21. Elharrouss O, Almaadeed N, Al-maadeed S (2020) An image steganography approach based on k-least significant bits (k-lsb). In: 2020 IEEE international conference on informatics, iot, and enabling technologies (ICIoT). IEEE, pp 131–135

22. Eltahir ME, Kiah LM, Zaidan BB et al (2009) High rate video streaming steganography. In: 2009 International conference on information management and engineering. IEEE, pp 550–553

23. Fang DY, Chang LW (2006) Data hiding for digital video with phase of motion vector. In: 2006 IEEE international symposium on circuits and systems. IEEE, pp 4–pp

24. Ferryman J, Shahrokni A (2009) Pets2009: dataset and challenge. In: 2009 Twelfth IEEE international workshop on performance evaluation of tracking and surveillance. IEEE, pp 1–6

25. Fridrich J, Goljan M (2002) Practical steganalysis of digital images: state of the art. In: Security and watermarking of multimedia contents IV, international society for optics and photonics, pp 1–13

26. Gadicha AB, Gupta VBB, Gadicha VB et al (2021) Multimode approach of data encryption in images through quantum steganography. In: Multidisciplinary approach to modern digital steganography. IGI Global, pp 99–124

27. Goodfellow I, Pouget-Abadie J, Mirza M et al (2014) Generative adversarial nets. In: Advances in neural information processing systems, pp 2672–2680

28. Gupta S, Gujral G, Aggarwal N (2012) Enhanced least significant bit algorithm for image steganography. IJCEM Int J Computat Eng Manag 15(4):40–42

29. Hanafy AA, Salama GI, Mohasseb YZ (2008) A secure covert communication model based on video steganography. In: MILCOM 2008-2008. IEEE military communications conference, IEEE, pp 1-6

30. He X, Luo Z (2008) A novel steganographic algorithm based on the motion vector phase. In: 2008 International conference on computer science and software engineering. IEEE, pp 822–825

31. Htet TT (2012) Digital video steganalysis based on statistical features. Tenth Int Conf Comput Appl (ICCA 2012)

32. Htet TT, Mya KT (2013) Video steganalysis using histogram and texture features. Eleventh Int Conf Comput Appl (ICCA 2013)

33. Hu Y, Zhang C, Su Y (2007) Information hiding based on intra prediction modes for h. 264/avc. In: 2007 IEEE international conference on multimedia and expo. IEEE, pp 1231–1234

34. Hu SD et al (2011) A novel video steganography based on non-uniform rectangular partition. In: 2011 14th IEEE international conference on computational science and engineering. IEEE, pp 57–61

35. Jaiswal A, Kumar S, Nigam A (2020) En-vstegnet: video steganography using spatio-temporal feature enhancement with 3d-cnn and hourglass. In: 2020 International joint conference on neural networks (IJCNN). IEEE, pp 1–8

36. Jangid S, Sharma S (2017) High psnr based video steganography by mlc (multi-level clustering) algorithm. In: 2017 International conference on intelligent computing and control systems (ICICCS). IEEE, pp 589–594

37. Jha VK, Mukherjee S, Roy S et al (2017) Video steganography technique using factorization and spiral lsb methods. In: 2017 International conference on computer, communications and electronics (Comptelix). IEEE, pp 315–320

38. Johnson NF (1998) Exploring steganography: seeing the unseen. Computer 31(2):26–34. https://doi.org/10.1109/MC.1998.4655281

39. Kar N, Mandal K, Bhattacharya B (2018) Improved chaos-based video steganography using dna alphabets. ICT Express 4(1):6–13

40. Kaur M, Kaur A (2014) Improved security mechanism of text in video using steganographic technique. Int J 2(10)

41. Kaur R, Kaur S (2016) Xor-edge based video steganography and testing against chi-square steganalysis. Int J Image Graph Signal Process 8(9):31

42. Kaur K, Kaur B (2018) Dwt-lsb approach for video steganography using artificial neural network. Int Adv Res J Sci Eng Technol, IARJSET

43. Kaur R et al (2016) A hybrid approach for video steganography using edge detection and identical match techniques. In: 2016 International conference on wireless communications, signal processing and networking (WiSPNET). IEEE, pp 867–871

44. Ke N, Weidong Z (2013) A video steganography scheme based on h.264 bitstreams replaced. In: 2013 IEEE 4th International conference on software engineering and service science, pp 447–450. https://doi.org/10.1109/ICSESS.2013.6615345

45. Kelash HM, Wahab OFA, Elshakankiry OA et al (2014) Utilization of steganographic techniques in video sequences. Int J Comput Netw Technol 2(01)

46. Khan N, Gorde KS (2015) Video steganography by using statistical key frame extraction method and lsb technique. Int J Innov Reas Sci Eng Technol 4(10)

47. Khupse S, Patil NN (2014) An adaptive steganography technique for videos using steganoflage. In: 2014 International conference on issues and challenges in intelligent computing techniques (ICICT). IEEE, pp 811–815

48. Kishor SN, Ramaiah GNK, Jilani SAK (2016) A review on steganography through multimedia. In: 2016 International conference on research advances in integrated navigation systems (RAINS), pp 1–6. https://doi.org/10.1109/RAINS.2016.7764373

49. Kolakalur A, Kagalidis I, Vuksanovic B (2016) Wavelet based color video steganography. Int J Eng Technol 8(3):165

50. Koppanati RK, Kumar K (2020) P-mec: polynomial congruence-based multimedia encryption technique over cloud. IEEE Consumer Electron Mag 10(5):41–46

51. Koppanati RK, Kumar K, Qamar S (2019) E-moc: an efficient secret sharing model for multimedia on cloud. In: International conference on deep learning, artificial intelligence and robotics. Springer, pp 246–260

52. Kordov K, Valchev G (2019) Video steganography with steganalysis, vol 5, pp 15–22

53. Korgaonkar VV, Gaonkar MN (2017) A dwt-dct combined approach for video steganography. In: 2017 2nd IEEE international conference on recent trends in electronics, information & communication technology (RTEICT). IEEE, pp 421–424

54. Kumar K, Shrimankar DD (2017) F-des: fast and deep event summarization. IEEE Trans Multimed 20(2):323–334

55. Kumar K, Shrimankar DD (2018) Esumm: event summarization on scale-free networks. IETE Tech Rev

56. Kumar K, Shrimankar DD, Singh N (2017) Event bagging: a novel event summarization approach in multiview surveillance videos. In: 2017 International Conference on Innovations in Electronics, Signal Processing and Communication (IESC). IEEE, pp 106–111

57. Kumar P, Singh K (2018) An improved data-hiding approach using skin-tone detection for video steganography. Multimed Tools Appl 77(18):24,247–24,268

58. Kuppusamy P, Ramya K, Rani SS et al (2020) A novel approach based on modified cycle generative adversarial networks for image steganography. Scalable Comput Practice Exp 21(1):63–72

59. Li-Yi Z, Wei-Dong Z et al (2011) A novel steganography algorithm based on motion vector and matrix encoding. In: 2011 IEEE 3rd international conference on communication software and networks. IEEE, pp 406–409

60. Liao YC, Chen CH, Shih TK et al (2009) Data hiding in video using adaptive lsb. In: 2009 Joint Conferences on Pervasive Computing (JCPC). IEEE, pp 185–190

61. Liao K, Lian S, Guo Z et al (2012) Efficient information hiding in h. 264/avc video coding. Telecommun Syst 49(2):261–269

62. Lin TJ, Chung KL, Chang PC et al (2013) An improved dct-based perturbation scheme for high capacity data hiding in h. 264/avc intra frames. J Syst Softw 86(3):604–614

63. Lin WB, Lai TH, Chang KC (2021) Statistical feature–based steganalysis for pixel-value differencing steganography

64. Lin WB, Lai TH, Chou CL (2021) Chi-square-based steganalysis method against modified pixel-value differencing steganography. Arab J Sci Eng 1–9

65. Liu Y, Chen L, Hu M et al (2016b) A reversible data hiding method for h. 264 with shamir's (t, n)-threshold secret sharing. Neurocomputing 188:63–70

66. Liu Y, Ju L, Hu M et al (2016a) A new data hiding method for h. 264 based on secret sharing. Neurocomputing 188:113–119

67. Liu Y, Li Z, Ma X et al (2014) A robust without intra-frame distortion drift data hiding algorithm based on h. 264/avc. Multimed Tools Appl 72(1):613–636

68. Liu S, Liu Y, Feng C et al (2020) A hevc steganography method based on qdct coefficient. In: International conference on intelligent computing. Springer, pp 624–632

69. Liu Y, Liu S, Wang Y et al (2019) Video steganography: a review. Neurocomputing 335:238–250

70. Liu Y, Liu S, Zhao H et al (2019) A new data hiding method for h. 265/hevc video streams without intra-frame distortion drift. Multimed Tools Appl 78(6):6459–6486

71. Liu S, Xu D (2020) A robust steganography method for hevc based on secret sharing. Cogn Syst Res 59:207–220

72. Liu S, Yao H, Gao W et al (2007) Minimizing the distortion spatial data hiding based on equivalence class. In: International conference on intelligent computing. Springer, pp 667–678

73. Liu M, Zhang M, Liu J et al (2017) Coverless information hiding based on generative adversarial networks. arXiv:171206951

74. Lu Y, Lu C, Qi M (2010) An effective video steganography method for biometric identification. In: Advances in computer science and information technology. Springer, pp 469–479

75. Luo W, Huang F, Huang J (2010) Edge adaptive image steganography based on lsb matching revisited. IEEE Trans Inf Forensics Security 5(2):201–214

76. Ma X, Li Z, Lv J et al (2009) Data hiding in h. 264/avc streams with limited intra-frame distortion drift. In: 2009 International symposium on computer network and multimedia technology. IEEE, pp 1–5

77. Ma X, Li Z, Tu H et al (2010) A data hiding algorithm for h. 264/avc video streams without intra-frame distortion drift. IEEE Trans Circuits Syst Video Technol 20(10):1320–1330

78. Manisha S, Sharmila TS (2019) A two-level secure data hiding algorithm for video steganography. Multidim Syst Sign Process 30(2):529–542

79. Mishra A, Kumar S, Nigam A et al (2019) Vstegnet: video steganography network using spatio-temporal features and micro-bottleneck. In: BMVC, p 274

80. Moon SK, Raut RD (2013) Analysis of secured video steganography using computer forensics technique for enhance data security. In: 2013 IEEE second international conference on image information processing (ICIIP-2013). IEEE, pp 660–665

81. Mstafa RJ, Elleithy KM (2014) A highly secure video steganography using hamming code (7, 4). In: IEEE long island systems, applications and technology (LISAT) conference 2014. IEEE, pp 1–6

82. Mstafa RJ, Elleithy KM (2015) A novel video steganography algorithm in the wavelet domain based on the klt tracking algorithm and bch codes. In: 2015 Long Island Systems, Applications and Technology. IEEE, pp 1–7

83. Mstafa RJ, Elleithy KM (2015) A high payload video steganography algorithm in dwt domain based on bch codes (15, 11). In: 2015 wireless telecommunications symposium (WTS). IEEE, pp 1–8

84. Mstafa RJ, Elleithy KM (2015) An efficient video steganography algorithm based on bch codes. In: ASEE northeast section conference 2015, Asee

85. Mstafa RJ, Elleithy KM (2016) An ecc/dct-based robust video steganography algorithm for secure data communication. J Cyber Security Mobility 167–194

86. Mstafa RJ, Elleithy KM (2017) Compressed and raw video steganography techniques: a comprehensive survey and analysis. Multimed Tools Appl 76(20):21,749–21,786

87. Mstafa RJ, Elleithy KM, Abdelfattah E (2017) A robust and secure video steganography method in dwt-dct domains based on multiple object tracking and ecc. IEEE Access 5:5354–5365

88. Mstafa RJ, Younis YM, Hussein HI et al (2020) A new video steganography scheme based on shi-tomasi corner detector. IEEE Access 8:161,825–161,837

89. Mumthas S, Lijiya A (2017) Transform domain video steganography using rsa, random dna encryption and huffman encoding. Proc Comput Sci 115:660–666

90. Narayanan K, Prabakaran G, Bhavani R (2012) A high capacity video steganography based on integer wavelet transform. J Comput Appl 5:358–365

91. Nguyen DC, Nguyen TS, Hsu FR et al (2019) A novel steganography scheme for video h. 264/avc without distortion drift. Multimed Tools Appl 78(12):16,033–16,052

92. Nissar A, Mir AH (2010) Classification of steganalysis techniques: a study. Digital Signal Process 20(6):1758–1770

93. Niu K, Li J, Yang X et al (2019) Hybrid adaptive video steganography scheme under game model. IEEE Access 7:61,523–61,533. https://doi.org/10.1109/ACCESS.2019.2902464

94. Nyo HL, Oo AW (2019) Secure data transmission of video steganography using arnold scrambling and dwt. Int J Comput Netw Inf Security 11(6)

95. Patel K, Rora KK, Singh K et al (2013) Lazy wavelet transform based steganography in video. In: 2013 International conference on communication systems and network technologies. IEEE, pp 497–500

96. Raja Ratna S, Shajilin Loret J, Merlin Gethsy D et al (2019) A review on various approaches in video steganography. In: Intelligent communication technologies and virtual mobile networks. Springer, pp 626–632

97. Rajalakshmi K, Mahesh K (2017) Video steganography based on embedding the video using pcf technique. In: 2017 International conference on information communication and embedded systems (ICICES). IEEE, pp 1–4

98. Rajesh G, Nargunam A (2013) Steganography algorithm based on discrete cosine transform for data embedding into raw video streams. In: IET conference proceedings the institution of engineering & technology

99. Ramalingam M (2011) Stego machine–video steganography using modified lsb algorithm. Int J Inf Commun Eng 5(2):170–173

100. Ramalingam M, Isa NAM (2014) Video steganography based on integer haar wavelet transforms for secured data transfer. Indian J Sci Technol 7(7):897

101. Ramalingam M, Isa NAM (2015) A steganography approach over video images to improve security. Indian J Sci Technol 8(1):79

102. Roque JJ, Minguet JM (2009) Slsb: improving the steganographic algorithm lsb. In: WOSIS, pp 57–66

103. Sadek MM, Khalifa AS, Mostafa MG (2015) Video steganography: a comprehensive review. Multimed Tools Appl 74(17):7063–7094
104. Sadek MM, Khalifa AS, Mostafa MG (2017) Robust video steganography algorithm using adaptive skin-tone detection. Multimed Tools Appl 76(2):3065–3085
105. Saini A, Joshi K, Allawadhi S (2017) A review on video steganography techniques. Int J Adv Res Comput Sci 8(3)
106. Sairam T, Boopathybagan K (2019) Computational intelligence-based steganalysis comparison for rcm-dwt and pva-mod methods. Automatika: časopis za automatiku, mjerenje, elektroniku, računarstvo i komunikacije 60(3):285–293
107. Seeling P, Reisslein M (2011) Video transport evaluation with h. 264 video traces. IEEE Commun Surveys Tutorials 14(4):1142–1165
108. Sharma S, Kumar K (2018) Guess: genetic uses in video encryption with secret sharing. In: Proceedings of 2nd international conference on computer vision & image processing. Springer, pp 51–62
109. Shukur WA, Abdullah WN, Qurban LK (2018) Information hiding in digital video using dct, dwt and cvt. In: Journal of physics: conference series, IOP publishing, p 012035
110. Subramanian N, Elharrouss O, Al-Maadeed S et al (2020) Image steganography using auto encoder-decoder based deep learning method. In: International conference on interactive collaborative and blended learning. Springer, pp 520–530
111. Subramanian N, Elharrouss O, Al-Maadeed S et al (2021) Image steganography: a review of the recent advances. IEEE Access 9:23,409–23,423. https://doi.org/10.1109/ACCESS.2021.3053998
112. Suresh M, Sam IS (2020) Optimal wavelet transform using oppositional grey wolf optimization for video steganography. Multimed Tools Appl 79(37):27,023–27,037
113. Suresh M, Shatheesh Sam I (2018) High secure video steganography based on shuffling of data on least significant dct coefficients. In: 2018 Second international conference on intelligent computing and control systems (ICICCS), pp 877–882. https://doi.org/10.1109/ICCONS.2018.8662920
114. Sushmitha M, Suresh H, Manikandan J (2017) An approach towards novel video steganography for consumer electronics. In: 2017 IEEE international conference on consumer electronics-asia (ICCE-Asia). IEEE, pp 72–76
115. Swanson MD, Zhu B, Tewfik AH (1997) Video data hiding for video-in-video and other applications. In: Multimedia storage and archiving systems II, international society for optics and photonics, pp 32–43
116. Swati S, Hayat K, Shahid Z (2014) A watermarking scheme for high efficiency video coding (hevc). PloS One 9(8):e105–613
117. Thakur A, Singh H, Sharda S (2015) Secure video steganography based on discrete wavelet transform and arnold transform. Int J Comput Appl 123(11)
118. Thomee B, Shamma DA, Friedland G et al (2016) Yfcc100m: the new data in multimedia research. Commun ACM 59(2):64–73
119. Volkhonskiy D, Borisenko B, Burnaev E (2016) Generative adversarial networks for image steganography
120. Volkhonskiy D, Nazarov I, Burnaev E (2020) Steganographic generative adversarial networks. In: Twelfth international conference on machine vision (ICMV 2019), international society for optics and photonics, p 114333M
121. Wahab OFA, Badawy MB, Elshakankiry OA et al (2015) Utilizations of reversible lossless data hiding techniques in video sequences. Int J Comput Netw Technol 39(01)
122. Wang Y, Cao Y, Zhao X (2020) Minimizing embedding impact for h. 264 steganography by progressive trellis coding. IEEE Trans Inf Forensics Security 16:333–345
123. Weng X, Li Y, Chi L et al (2019) High-capacity convolutional video steganography with temporal residual modeling. In: Proceedings of the 2019 on international conference on multimedia retrieval, pp 87–95
124. Wu M, Liu B (2003) Data hiding in image and video. i. fundamental issues and solutions. IEEE Trans Image Process 12(6):685–695
125. Xu C, Ping X (2007) A steganographic algorithm in uncompressed video sequence based on difference between adjacent frames. In: Fourth international conference on image and graphics (ICIG 2007). IEEE, pp 297–302
126. Xu C, Ping X, Zhang T (2006) Steganography in compressed video stream. In: First international conference on innovative computing, information and control-volume I (ICICIC'06). IEEE, pp 269–272
127. Xu D, Wang R, Shi YQ (2013) Reversible data hiding in encrypted h. 264/avc video streams. In: International workshop on digital watermarking. Springer, pp 141–152
128. Xu D, Wang R, Shi YQ (2014) Data hiding in encrypted h. 264/avc video streams by codeword substitution. IEEE Trans Inf Forensics Security 9(4):596–606

129. Xu D, Wang R, Shi YQ (2016) An improved scheme for data hiding in encrypted h. 264/avc videos. J Vis Commun Image Represent 36:229–242

130. Xue Y, Zhou J, Zeng H et al (2019) An adaptive steganographic scheme for h. 264/avc video with distortion optimization. Signal Process Image Commun 76:22–30

131. Yadav P, Mishra N, Sharma S (2013) A secure video steganography with encryption based on lsb technique. In: 2013 IEEE international conference on computational intelligence and computing research. IEEE, pp 1–5

132. Yang G, Li J, He Y et al (2011) An information hiding algorithm based on intra-prediction modes and matrix coding for h. 264/avc video stream. AEU-Int J Electron Commun 65(4):331–337

133. Yao Y, Zhang W, Yu N et al (2015) Defining embedding distortion for motion vector-based video steganography. Multimed Tools Appl 74(24):11,163–11,186

134. Younus ZS, Younus GT (2019) Video steganography using knight tour algorithm and lsb method for encrypted data. J Intell Syst 29(1):1216–1225

135. Yun C, Xianfeng Z, Dengguo F et al (2011) Video steganography with perturbed motion estimation

136. Zhang H, Cao Y, Zhao X (2016) Motion vector-based video steganography with preserved local optimality. Multimed Tools Appl 75(21):13,503–13,519

137. Zhang J, Li J, Zhang L (2001) Video watermark technique in motion vector. In: Proceedings XIV Brazilian symposium on computer graphics and image processing. IEEE, pp 179–182

138. Zhang Y, Zhang M, Wang XA et al (2015) A novel video steganography algorithm based on trailing coecients for h. 264/avc. Informatica, vol 40(1)

139. Zhang L, Zhao X (2016) An adaptive video steganography based on intra-prediction mode and cost assignment. In: International workshop on digital watermarking. Springer, pp 518–532

140. Zhu H, Wang R, Xu D et al (2010) Information hiding algorithm for h. 264 based on the predition difference of intra_4× 4. In: 2010 3rd International congress on image and signal processing. IEEE, pp 487–490