DEMOCRATIC REPUBLIC OF ALGERIA
MINISTRY OF HIGH EDUCATION AND SCIENTIFIC RESEARCH

Jijel University
Faculty of Science Technology
Department of Computer Science

Memory
In order to obtain the Master's Degree in Computer Science:
Legal Informatics and Multimedia
Theme

Digital Video Tampering and Detection Techniques

Presented by

ABDOU Chaima

LOUNIS Selma

Supervised by

Dr LAHOULOU Atidel

College year :2016/2017

Remerciement

Nous venons en premier lieu remercions « ALLAH » qui nous ont donné la force pour terminer et d'avoir réussit dans nos études.

« Allhamdou li Allah »

Nous tenons à exprimer notre profonde gratitude à tous les enseignants du département pour leur soutien inestimable depuis nos premières années d'études.

Nous remercions nos famille plus particulièrement nos parents pour leur soutien permanent tout au long de ce mémoire et plus généralement au long de notre vie universitaire.

Nous aimerons aussi remercie notre encadreur , Dr ***Atidel lahoulou.***

Nos remerciements vont aux membres de jury qui nous a fait l'honneur de juger ce travail en tant que examinatrice

Nous tenons à remercie vivement les membres du département d'informatique

Nos remerciements vont également à monsieur Boubakir A et monsieur Mahrouk Z

Nous tenons aussi à dire un immense merci à tous nos collègues de promotion 2017 et tous nos amies

Enfin, nous adressons nous remerciements à toutes personnes ayant contribuée de près ou de loin à la réalisation de ce mémoire.

Chaima et selma

Dédicace

De mes profondes sensations de sagesse et sincérité je dédie ce travail

A ma mère

La plus belle créature que Dieu a créée sur terre, A cette source de tendresse, de patience et de générosité !

A mon père

Ce travail est le fruit de tes sacrifices que tu as consentis pour ma formation tout au long de ces années.

A mes chères sœurs et frères

Kenza, Roumaissa, Abd arrahmane, Mouhamed

A l'amour de ma vie

Nedjmo

A mes amis(e)

Aux personnes qui m'ont toujours aidé et encouragé, qui étaient toujours à mes côtés, et qui m'ont accompagnaient durant mon chemin d'études.

A toutes les personnes qui m'ont aidé par un mot, m'ont donné la force de continuer ....

CHAIMA

Dédicace

A ma chère mère

Pour son encouragement sa patience et surtout ses prières qui m'ont toujours guidé.

A mon père

Pour son soutien et sa patience

Mes sœurs *Lina, Mimi ,Nabiha,Samira et Rima*

Mes frères *Samir, Ahmed, Reda ,Omar et Karim*

Tous mes amies *Hadjer ,Radja,khadidja, Rawnak ,Widad*

Mon binômes *Chaima* pour sa patience durant toute cette années

A Ma tente *Naima*

Selma

Abstract

With the increasing accessibility of technology for everyday people, things are starting to get digitalized: digital camera, digital cable, digital sound, and digital video. It is no longer the case where a video production is only possible for specialized studios. The availability of various user-friendly, inexpensive tools is pushing motion pictures into individual computer owners. The Digital forgeries though not visibly identifiable to human perception it may alter or meddle with underlying natural statistics of digital content. Tampering involves fiddling with video content in order to cause damage or make unauthorized alteration/modification. Tampering impacts need to be studied and the applied technique/method is used to establish the factual information for legal course in judiciary. In this report we presenting different types of video forgery in the first chapter and detection techniques both in active and passive video in the second .
Keywords: video tempering, active forgery detection, passive forgery detection, v

Résume

Avec la croissante de la technologie électronique, tout commence à numériser dans notre vie quotidienne : appareil photo numérique, câble numérique, son numérique et vidéo numérique. La production de vidéo est donc possible pour tout le monde. La disponibilité de différents outils conviviaux et peu coûteux conduit à des falsifications numériques, non visiblement identifiables à la perception humaine. L'altération implique de jouer avec du contenu vidéo afin de causer des dommages de modification non- autorisé. Après des études sur les impacts de falsification la technique / méthode appliquée est utilisée pour établir les informations factuelles pour le cours légal dans le système judiciaire. Dans ce rapport, nous présentons dans le premier chapitre différent types de falsification vidéo et dans la seconde les techniques de détection dans la vidéo active et passive.
Mots-clés: video tempering, active forgery detection, passive forgery detection, video watermarking

Contents

Introduction

Video data has become more popular with the advancement of digital cameras and networking technologies with high speed bandwidths, as a result many systems make use of video data and rely on the accuracy of such data .on the other hand, an inevitable adverse effect of this critical nature of video data is video tampering.

with difficulty to detect a tampered video become more detection techniques and methods are implemented and developed despite that these methods does not include all types of tampering.

In the first chapter we present different illegal manipulation of videos within a brave definition of video, caracteristique of video forensic and video tampering .Then in chapter two video tampering detection technique Active and passive. For the state of art we chose chapter three to present different existing methods of copy-move detection and finally experimental results of similarity analysis

Chapter 1

Illegal manipulation of videos

1 Introduction

Advancements in video innovations compression , transmission, stockpiling, recovery combined with the availability of production /editing devices and fast growth of internet facilitate uploading ,downloading and sharing multimedia content (images ,music ,videos)especially with the advancement of social media websites like YouTube ,Facebook,Twitter.... Ect As a result, we have begun to see an apparition of some darker sides of video data and an increase in illegal uses and manipulation of videos like copy rights infringement, video tampering.

2 what is a video

A video is a succession of images. The fundamental principle of video is that the human eye has the possibility of retaining for a certain time (the order of a tenth of a second) any image printed on the retina.It is enough to scroll a sufficient number of frames per second .In the field of digital video, there are three essential spaces of color:

- RGB
- Gray-scale space
- YUV space

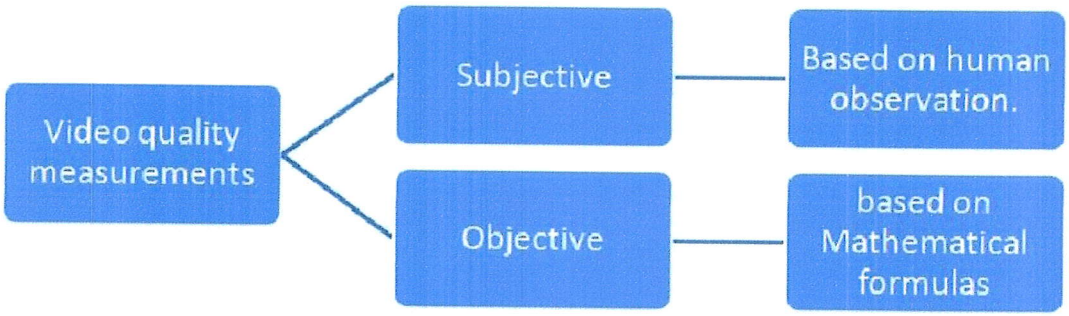


Figure 1.1: present video quality measurements

The digital video is too voluminous so compression is needed.

2.1 why we need compression

Digital television present a privileged position. The development of digital video has been possible mainly through the possibilities of compression. The compression of the video sequences is necessary in order to store the video sequences , and to be able to transmit digital video data through a limited bandwidth network or from a medium who have a transfer rate limit. The objective of video compression is to reduce the number of bits required to represent the information carried by a video sequence. The improvements brought by the compression are not simply due to the elimination of the redundant data but rather The discarding of information deemed irrelevant, for example information that are not perceptible to the naked eye The theory of information tries to extract the relevant signs of information and to abandon redundancy to reduce the flow or the storage capacity of a signal. It's a size problem Ex: 1h uncompressed video = 1GB, 1h uncompressed HDTV = 800GB

2.2 Hierarchy of video data

figure 1.2 shows a Hierarchy of video data

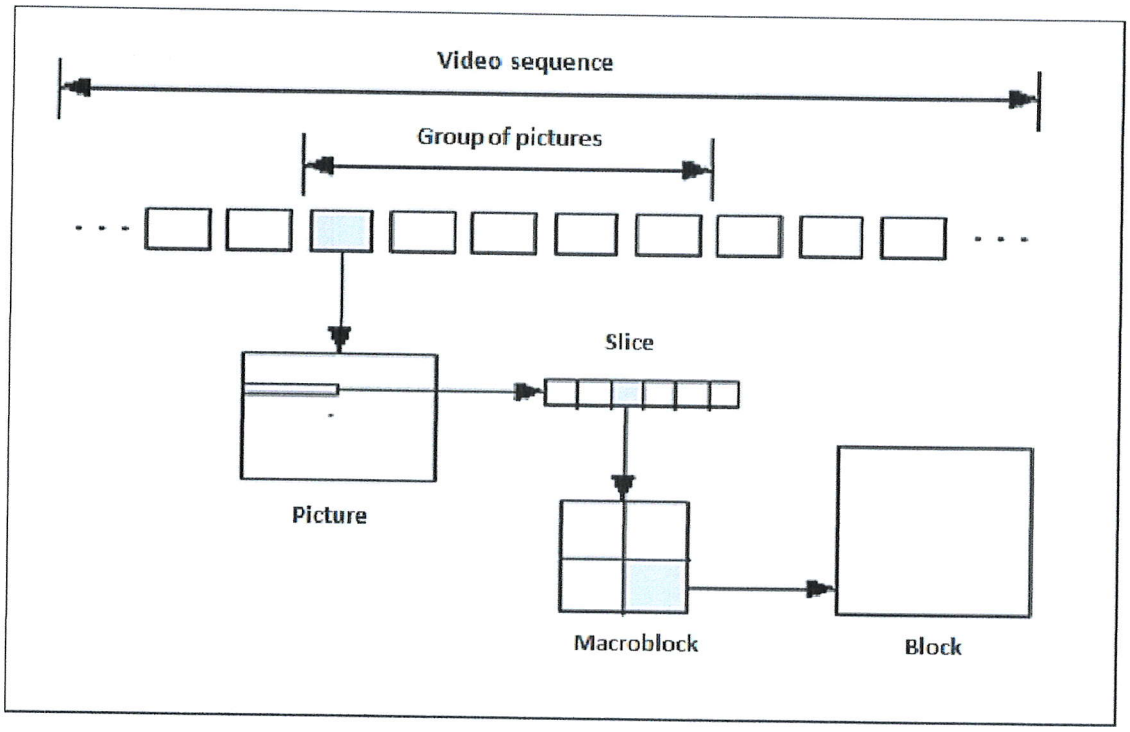


Figure 1.2: Hierarchy of video data

Sequence video : it starts with a header of sequence containing one or more groups of images and ends with a sequence end code.

Group of pictures : it groups together a header and series of one or more images allowing to access randomly.

Image : it is the elementary unit for the coding of the video sequence an image is a group of three rectangular matrices which represents the luminance (Y) and the chrominance (Cr and Cb), an element of the matrix representing a pixel. This representation YCrCb is equivalent to that of RGB. It is preferable because the eye is more sensitive to luminance than to chrominance, so it is not necessary to store as much information in the matrices Cb and Cr as In the matrix Y; Whereas in RGB, the three matrices are of the same size. The matrices Cb and Cr are twice smaller than the matrix Y.

Slice : The slice is one or more adjacent macroblocks ordered from left to right and then from top to bottom. These are important elements for error handling. If the data stream contains an error, the decoder can skip the slice and move to the beginning of the next one directly is processing the errors but it is a waste of space.

Macroblocks : it is a rectangular matrix of two dimension and made up of blocks. Which are pavers of 8 * 8 pixels. Therefore a macroblock covers 16 * 16 pixels in the luminance space and 8 * 8 pixels in the chrominance space.

Block : This is a set of luminance and chrominance values of 8 lines of 8 pixels

3 Video Forensic

Video forensic has become an important area of research in the last decade, by analyzing the video by extracting valuable information from it content its main objectives comprise source identification, hidden data detection and tampering detection.

3.1 Source identification

Multimedia content is produced using various devices, e.g: computer, camera, scanner, recorder, cell phone..etc. In general, each device has different characteristics that affect the generated multimedia content. This is based on the assumption that all multimedia content generated by a device will contain certain characteristics that are intrinsic to the device itself. The origin of the multimedia content or source device (mobile phones, camcorders, Cameras) can be identied by analyzing the characteristics of devices and the multimedia content they produce [1]. Video source identification is very important in validating video evidence, tracking down video piracy crimes and regulating individual video sources

3.2 Video hidden content detection

steganography was a way to secure data that was used in cryptography. Unfortunately it has been connected to the distribution of child pornography on the dark web and it's an advanced communications tools for terrorists and Drug-dealers ; In May 2011 a suspected AL-Qaeda member was arrested in Berlin According to the German news paper Die Ziet he was found in possession of a memory card with a password protected folder with what appeared to be a pornographic video called Kick Ass Within that video German forensic investigators discovered 141 separate text files ,containing what appear to be document detailing Al-Qaeda operations and plans for future operations among them three entitled "Future work ", "lessons learned" and "report on operations" [2]

In response to all those cyber crimes video hidden content detection or what we call Video Forensic steganalysis was needed [3]

3.3 Detection of illegal reproduction of videos

An important problem in copyright protection is the proliferation of bootleg videos[1] :many illegal copies of movies are available on the Internet even before their official release. which caused a huge financial losses these fake copies is produced by recording films with camcorders in cinemas. Video forensics help to reduce this problem by:

detection of re-acquisition(detection of re-projected video) Re-acquisition occurs when a video sequence that is reproduced on a display or projected on a

[1]The word "bootleg" originates from the practice of smuggling illicit items in the legs of tall boots, particularly the smuggling of alcohol during the American Prohibition era. The word, over time, has come to refer to any illegal or illicit product. This term has become an umbrella term for illicit, unofficial, or unlicensed recordings, and any other commercially sold media or materia

screen is recaptured. Several approaches based on watermark are developed for detecting re-projected video

detection of copying The most common approach in video copy detection is to extract salient features from visual content that do not depend on the device used to capture the video.

3.4 video tampering detection

video tampering detection techniques seeks to find evidence of tempering in a digital evidence. this category of video forensic is the main purpose of this thesis we will present video tampering detection techniques with details in the following subsections

4 video tampering

4.1 about video tampering

Tampering the digital video is modifying or changing the contents of video's frame or the order of the frames to alter or meddle with underlying natural statistics of digital content Tampering involves fiddling with video content in order to cause damage or make unauthorized alteration/modification This should be possible by different techniques which will be presented in the following subsection in order to create tampered (doctored,forgered or fake) video.

Level of tampering attack

1. **Scene level:** The Whole scene of the video sequence is manipulated in such a way that not the scene itself is altered but the scene of the video is altered. Copying of a video scene to another place or delete a scene. In spatial and temporal both kinds of tampering can be done at the scene level.

2. **Shot Level Tampering:** In shot level tampering, a particular shot of the given video is altered. In shot level tampering shot can be added or deleted from the video. It can be performed at spatial as well as temporal.

3. **Frame Level Tampering:** In Frame level tampering, the alteration is done on video's frames. The attacker may delete the frames, add the frames, reshuffle the sequence of frames, and replicate the frames from a given video to change the contents of the video. This can be done using temporal tempering.

4. **Block Level Tempering:** In Block level Tempering, the content of the video frames are treated as blocks. And on which the tampering attacks are applied. Blocks mean a specific part of the video's frame can be replaced, cropped, altered or modified in block level tampering. Block level tampering attacks are commonly performed at spatial tampering.

5. **Pixel Level Tempering:** In pixel level tampering, the content of the video frames are altered at the pixel level. The video authentication system should

be strong enough to differentiate the regular video processing operation and pixel level tampering since normal video processing operations are performed at the pixel level. Pixel level attacks are performed at spatial tampering.

4.2 Types of video tampering

There are several possible attacks that can be applied to alter the contents of a video data [4]., we can broadly classify the video tampering attacks into three categories: spatial tampering attacks wich happend on frame level it's done with manupilation of a region, temporal tampering attacks and the combination of these two, Spatio-temporal tampering attacks [1] In figure 1.3 present differents types of video tampering

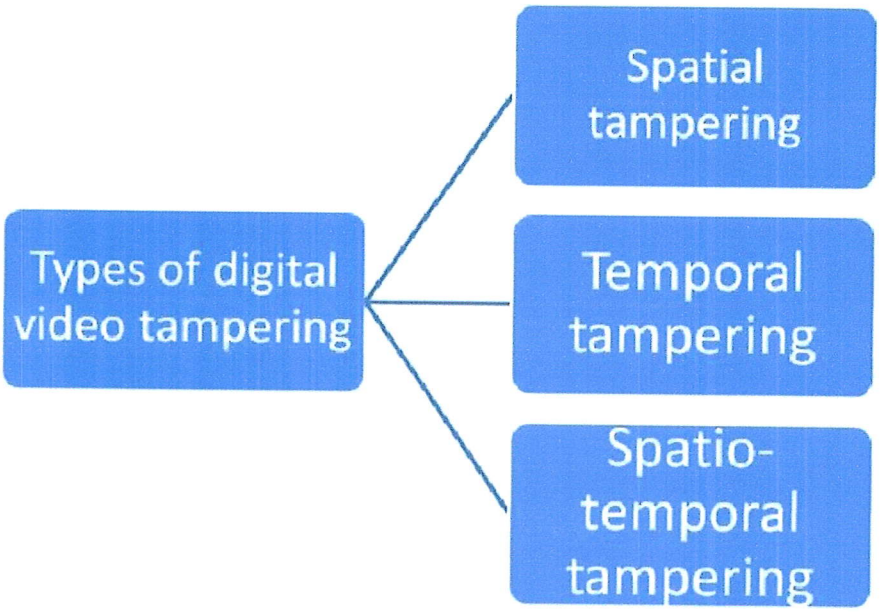


Figure 1.3: Types of Digital video tampering

4.2.1 Spatial tempering (intra frame tampering)

Alterations are performed on content of the frame (x- y axis) ,In spatial tampering with the help of digital video editing techniques and video editing software such as Photoshop and more,the attacker can easily affect many manipulations on the frame like adding , removing , deleting and modifying an object in a frame even a small manipulation of pixel bits causes a video tampering. as shown in figure 1.4 it's can be done on three levels:

Pixel level pixel is the smallest component of video so working with it makes the work more precise

Block level block is a specified area on the frame of the video, some attacks in spatial domain manipulate frame's block like removing, adding modification

Scene level the Whole scene of the video sequence is manipulated

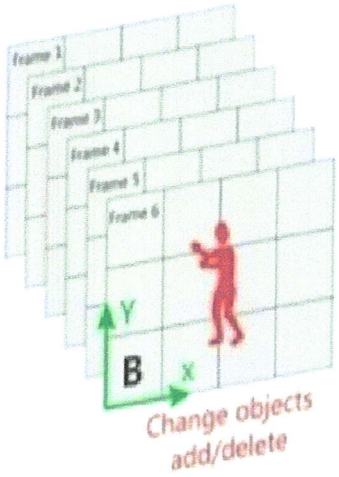


Figure 1.4: spacial video tampering

We can distingue several attacks in spatial domain like figure 1.5 shows

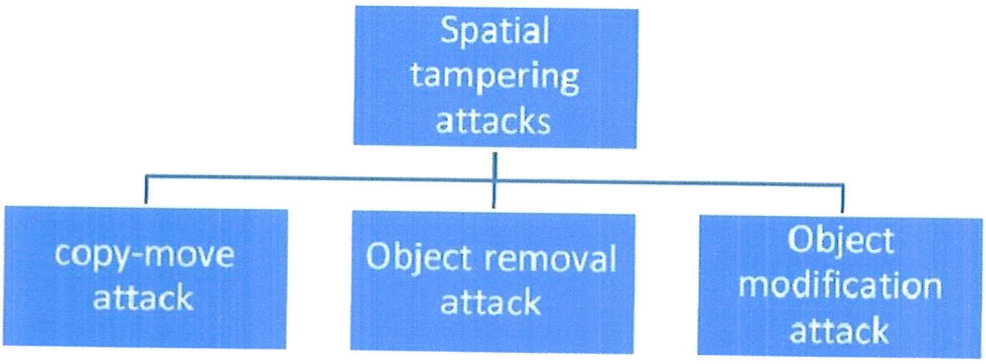


Figure 1.5: spacial video tampering attacks

4.2.1.1 copy-move attack In copy-move or copy-paste tampering technique, portions of the frame/image are cloned (Copy-paste) to expose a person or object in the scene.figure 1.6 shows a copy move

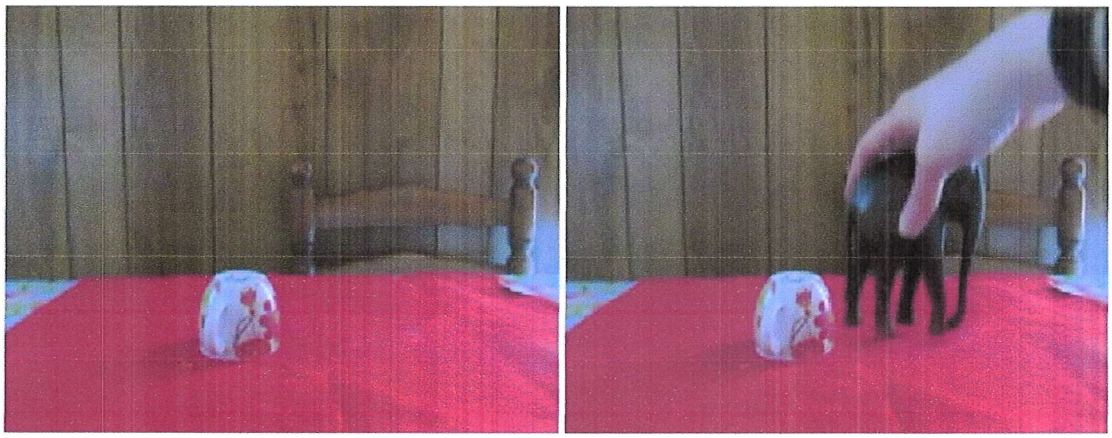


Figure 1.6: Object copy move

the copy-move attack can divided into two parts:

Object duplication attack Object duplication attack is Cloned regions (patches or blocks) can be of any shape and location in a frame possibly with some modifications. When copied region of some frames is pasted on different frames then it is known as video splicing. We can say that splicing is a special case of copy- move like figure 1.7 shows at the left original image and a tampered one the right

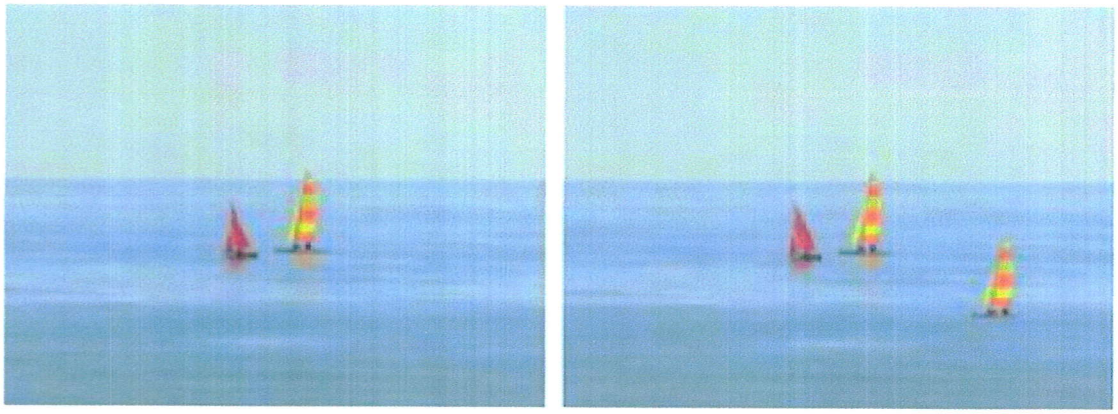


Figure 1.7: Object duplication attack

Object hiding attack on a frame an object can be kept out of sight by another object or even replace a person with another person

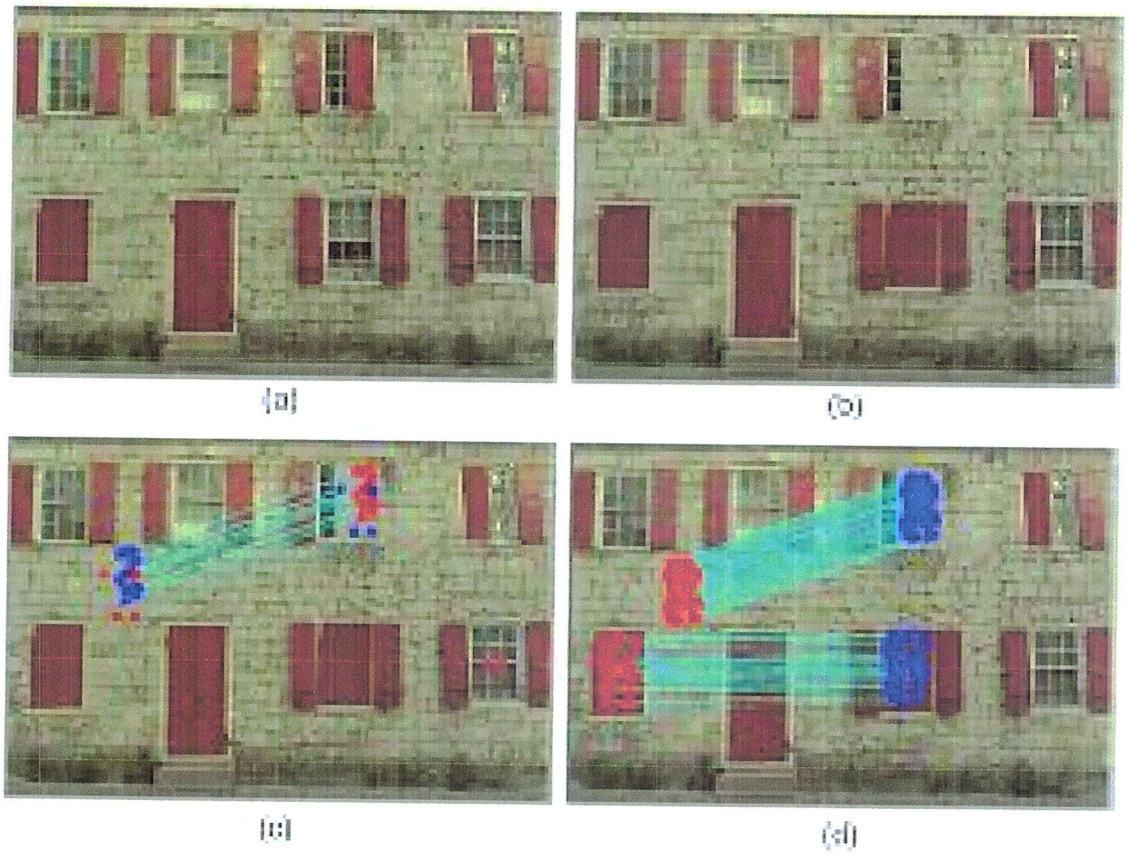


Figure 1.8: a)present an original frame b)the same frame after abject hiding attack c,d)shows the tampered places

4.2.1.2 Object removal attack In this attack an object of frame in video will be eliminated or deleted to hid it with malicious intention, the object who can be a car, a ball and even a person can be removed from one or more frames it depends on the number of appearance how much frames the objects appear on it. So the number of tampered frames with this attack equals number of appearance of that object . as shown in figure 1.9

Figure 1.9: Object removal attack

4.2.1.3 Object modification with this attack there are several modifications we can find in frames. Those frames are different of the original sequences cause they were manipulated with resampling which is resizing , rotation and other geographic transformation and splicing which is a Photomontage of two frames or images for example, the moving objects can be separated from stationary objects

Figure 1.10: Frame splicing attack

1.11 shows an example of resampling manipulation

Figure 1.11: Frame Resampling manipulation

4.2.2 Temporal tampering (inter frame tampering)

The tempering is done on the third dimension of the video(time) which make the difference between video forgery and image forgery, It is performed on sequence of frames. Focus on temporal dependency. Also called 'third dimensional attack' which done on frame level.this type of attack is shown in in figure 1.12

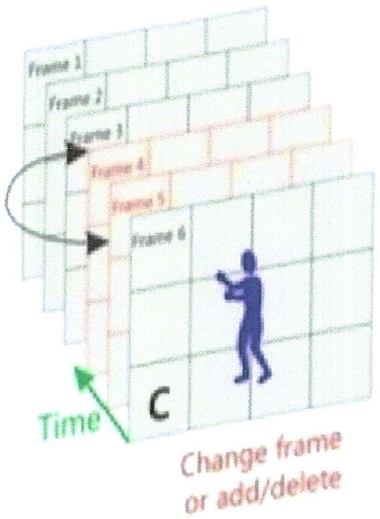


Figure 1.12: temporel video tampering

types of temporal tampering attacks like figure 1.13 shows

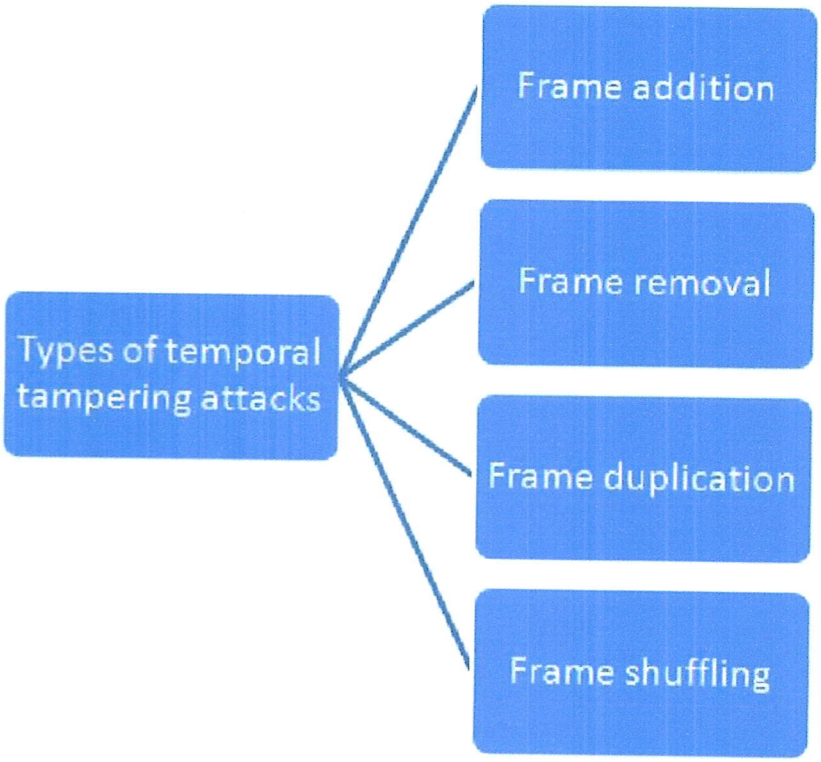


Figure 1.13: Types of temporel video tampering

4.2.2.1 Frame addition Additional frames from another video or in the same video, which has same statistical property, are intentionally inserted at some random locations in a given video. It intends to camouage the actual content and provides incorrect information [5]

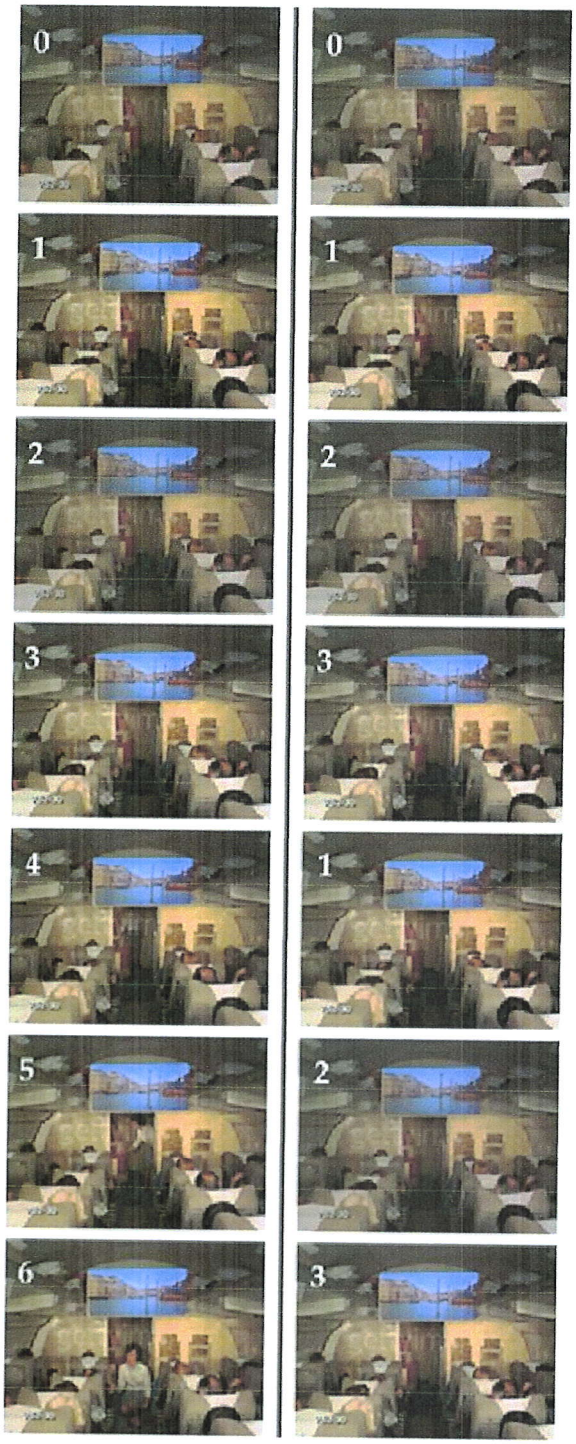


Figure 1.14: adding Frames in a video sequence

4.2.2.2 Frame removal Delete a frame from a specific place or different places when it appears many time, generally it's used on surveillance video when a person

want to eliminate his presence at a specific time from a specific place

4.2.2.3 Frame duplication In duplication of frame people hide the unwanted frame by duplicating another frame ,it's a temporal copy paste which is copy and paste a frame
In figure 1.15 the first frame has duplicated in order to hide the Passing of white van

Figure 1.15: Frame duplication

4.2.2.4 Frame shuffling Changing the order of video frames which can produce video's content different from the original.

This attack can be used when a suspect wants to prove that he was not present at a given moment in the crime scene but rather in another time. Like in figure 1.16 the order of frame has changing, in the original video the red car has passed first but in But in the tampered video the white car has passed first

Figure 1.16: Frame shuffling

4.2.3 Spacio temporel tampering

It is a combination of temporal and spatial tampering. Frame sequences and visual contents are modified in the same video [5],this attack is dementrated in figure1.17

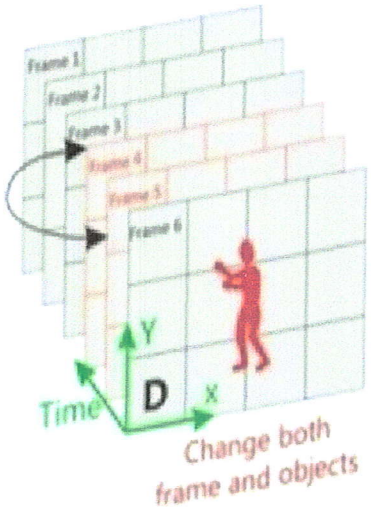


Figure 1.17: spacio temporel video tampering

in figure 1.18 the video has tampered both in temporal and spacial domain ,In temporal tampering the first frame has duplicated two times and in spatial domain the women has eliminated from the fourth frame

Figure 1.18: figure show a spacio temporel attack

5 Conclusion

Due to advantages of video content and Developments in visual video technologies such as compression, transmission, storage, retrieval. Videos are extensively used and shared therefore A serious problem is born that the integrity of digital video content s easily violated because digital video can be easily modified using editing tools. Systems and methods for verifying the integrity of video content, by detecting any changes in the content, are thus becoming required and increasingly important

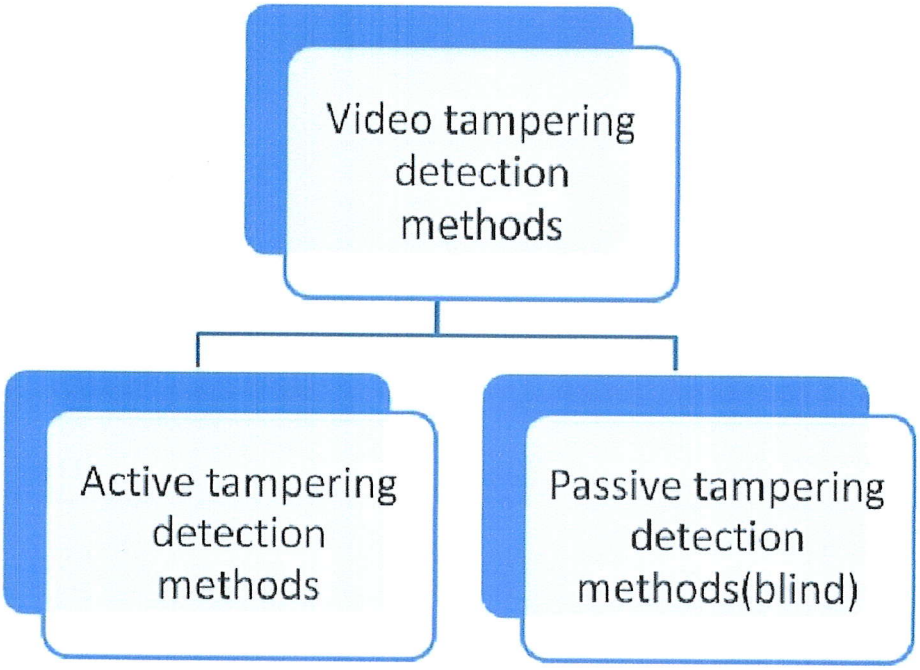


Figure 2.1: Types of Digital video tampering detection methods

Active methods consist to inserts information into the media prior to distribution, which can be later extracted to establish ownership. However, in passive methods the media (video, audio, image) contains enough unique information that can be used for detecting copies [6]. In this regard, considerable effort has been devoted to effective representation of video signatures and similarity matching.

3 Active tampering detection technics

In the case of an active video it is easy to detect tampering by using digital signature and digital watermark, but if there is no information about the source camera and video does not contain digital signature or digital watermark then it is very challenging to detect video tampering. Generally Internet streaming videos do not contain information regarding the source camera, digital signature and digital watermark.

figure 2.2 present Active detection methods

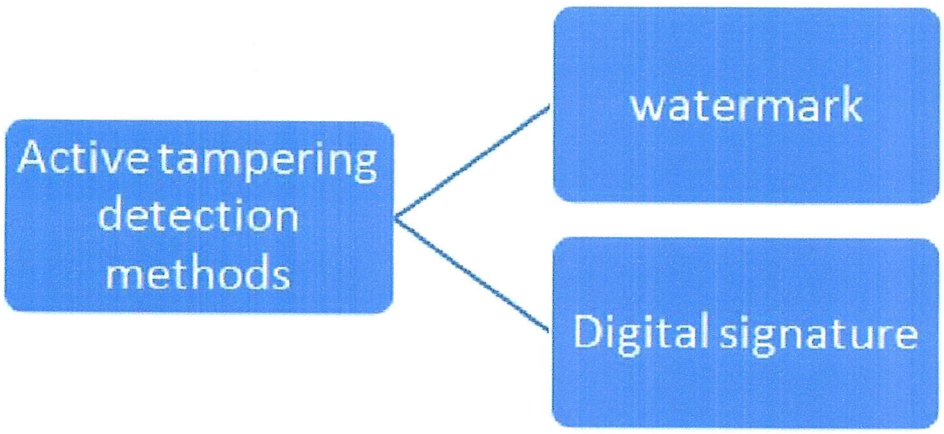


Figure 2.2: Active detection methods

3.1 Watermark

Digital watermark is a process of inserting secret information into digital multimedia, digital video watermarking has many aspects that must be considered during the design of any digital video watermark schemes such as imperceptibility, robustness, security, complexity, compressed domain , constant bit-rate ,interoperability [7] .

1. **Imperceptibility (Invisibility):** The digital watermark embedded into the video data should be invisible to the human observer.

2. **Security:** Unauthorized removal of the watermark must be impossible once it has been embedded, even if the basic scheme used for watermarking is known, as long as the exact parameters are unknown.

3. **Robustness:** It should be impossible to manipulate the watermark by intentional or unintentional operations on the uncompressed or compressed video without, at the same time, degrading the perceived quality of the video to the point of significantly reducing its commercial value. Such operations are, for example, addition of signals , filtering, cropping, encoding, or analog recording and playback.

4. **Complexity:** Watermarking and watermark retrieval should in principle have low complexity. Different applications do, however, pose different requirements on complexity. If watermarking is used for audit trail, each receiver has to retrieve the watermark, and watermark retrieval should be easy. If watermarking is used for embedding individual receiver identity labels, watermarking is performed on a large number of distributed video sequences, while watermark retrieval occurs only in cases where possible copyright violations have to be investigated. While the retrieval operation may be more complex in order to account for all possible kinds of attacks on the watermark, watermarking should be of low complexity in such cases.

5. **Compressed domain processing:** It can be assumed that the distributor or broadcaster of digital video will usually store the video in compressed format, for example on a video-on-demand server, or a World Wide Web server. Referring to the above complexity requirement, it should be possible to incorporate the watermark into the compressed video (the bit stream), because it is too complex and not feasible to decode and re-encode the video for watermarking the quality of decoded and re-encoded video can in general not be guaranteed.

6. **Constant bit-rate:** Watermarking in the bit stream domain should not increase the bit-rate, at least for constant bit-rate applications where transmission channel bandwidth has to be obeyed.

7. **Interoperability:** Even though many applications call for watermarking of compressed video, it would be a desirable property if uncompressed video could compatibly be watermarked without having to encode it first.

video watermarking techniques are divided into three main groups spatial domain,frequency domain and bit-stream domain as figure 2.3 shows

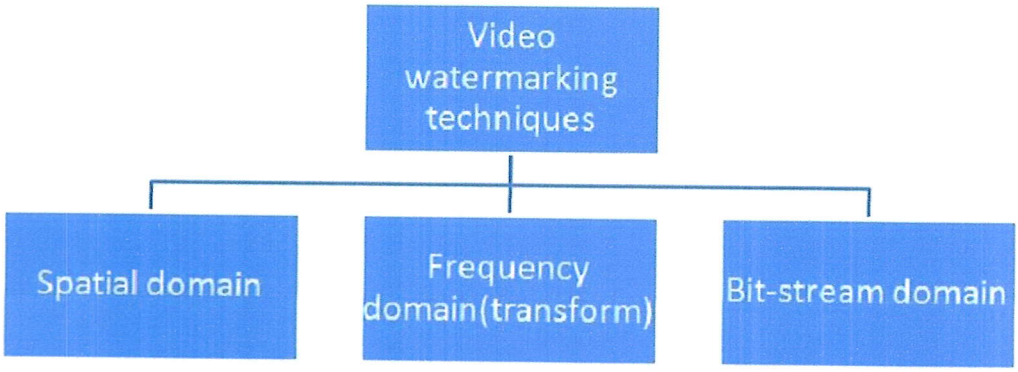


Figure 2.3: Video watermarking techniques

3.1.1 Spatial domain watermarking techniques

This techniques embed the watermark by modifying the pixel positions or pixel values of the video . The main advantages of using this technique are the small time complexities and simplicity of Implementation. However, these techniques have some disadvantages in providing robustness and meeting imperceptibility requirements [8]

The watermarking technique is implemented using the steps shown in figure 2.4

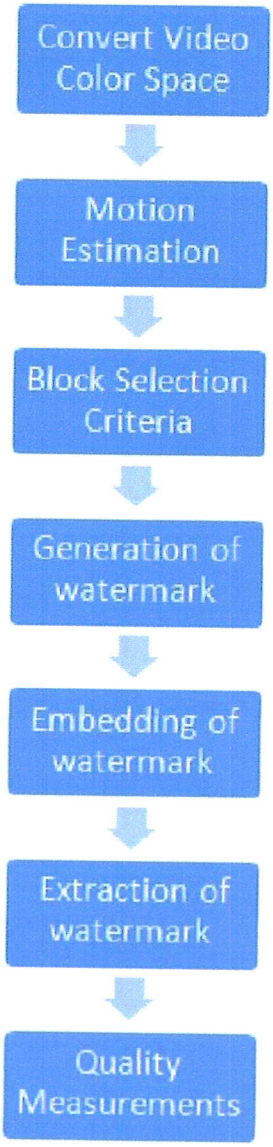


Figure 2.4: watermarking techniques steps in spatial domain

The authors in [9] have implemented the steps that are used to embed the watermark with the input video.as shows below :

- Extract loaded the color video into frames.
- Block matching motion applied in estimation techniques on the succeeding frames.
- Select only those frames that have enough number of motion blocks which is well-matched with the watermark size.

- From the selected frames to select the best blocks to use threshold during the matching process use a given threshold.
- Perform the wavelet transformation on the selected finest blocks.
- Random Gaussian distribution is embedded as a proposed watermark into the selected blocks (Apply only to the LH and HL wavelet bands).
- Extract the watermark which is embedded.
- Apply some attacks on the watermarked frames into the video.
- The conducted results are evaluated using PSNR for embedding and similarity for the extracting process before and after attacks.

The most important methods based on spatial domain watermarking are LSB and SSSC.

3.1.1.1 LSB(Least Significant Bit) The easiest watermarking method in spatial domain, this technique consist to replace every least significant bit(the most right bits) from the frame by bits from watermark (more details are shown in the example below)

example :

Supposedly we have the first pixel of frame is 25 and second =24

25 in binary =11001000

24 in binary =11000000

11001000 the watermark

By applying 1-LSB

1100100 became 11001001

11000000 became 11000001

By applying 2-LSB

1100100 became 11001011

11000000 became 11000000

3.1.1.2 SSSC(Spread Spectrum Signal Correlation) This method is based on adding a pseudo-random noise pattern to the luminance value frames in the spatial domain, and the correlation between the noise pattern

and possibly watermarked video for each frame is computed. If the correlation exceeds a certain threshold then, the watermark is detected.

3.1.2 frequency domain watermarking techniques(Transform)

It's also known as frequency domain watermarking in this techniques The host signal is transformed into a different domain according to a pre-determined transform (discrete cosine transform DCT, discrete wavelet transform DWT, discrete Fourier transform DFT) then the watermark is embedded in selective coefficients.

The majority of video watermarking techniques have been used in the frequency transform domain in order to overcome the main disadvantages of the spatial domain. Further, analysis of the bands in the frequency domain is a prerequisite to enhance watermark robustness and imperceptibility. [10]

3.1.2.1 Discrete cosine transforms DCT this technique is used to watermark compressed video streams ,Discrete cosine transformation transforms a signal from the spatial into the frequency domain by using the cosine function, it permits to device a frame into different frequency bands then the watermark is embedded into the middle Frequency bands of a frame. The middle frequency bands are chosen such that they have minimized to avoid the most visual important parts of the frame (low frequencies) [11]

In [12]authors proposed effective fragile video watermarking technique to embed and extract watermark in DCT domain with high capacity and transparency. Two watermarks are embedded into each frame. The first watermark is bits of the digital signature of hash value of the frame in frequency domain and second watermark is bits of micro-block numbers and frame numbers. The watermarks are embedded into video frames one by one in highest non-zero coefficient of quantized DCT coefficient. The first watermark is used to detect tampering and second watermark is used to localize the area being tampered. This technique causes significantly smaller video distortion as bits are embedded into the highest frequency coefficients. The embedded watermark is extracted and verified using public key. The block numbers and frame numbers are inserted in order to detect intra-frame and inter-frame tampering such as addition or removal of content within frames, frame reordering, dropping or addition of extra frames. If the video is being tampered we may extract one watermark correctly but other may get destroyed.

Note: A watermarking scheme that can detect malicious tampering and can be used in any modern video codec, such as H.264/AVC was introduced in [13] in this methods the watermark signals represent the macro block's and frame's indices, and are embedded into the nonzero quantized discrete cosine transform (QDCT)value of blocks, the nonzero are chosen such that they are able to detect spatial, temporal and spatiotemporal tampering.

3.1.2.2 Discrete wavelet transforms DWT It is a transform based on frequency domain, Due to its excellent spatio-frequency localization properties; the DWT is very suitable to identify areas in the video frame where a watermark can be embedded imperceptibly

3.1.2.3 Discrete Fourier transforms DFT A Discrete Fourier Transform (DFT) is performed on the original video data. The watermark is embedded in optimal selected frequency bands of the DFT. The DFT approach [14] has one advantage in comparison with the spatial domain methods. First, it is translation invariant and rotation resistant, which translates to strong robustness to geometric attacks. On the other hand, according to Raja et al. [15], fast Fourier transform (FFT) methods introduce round-off errors, which can lead to loss of quality and errors in watermark extraction. However, Cheddad et al. [16] states that this disadvantage is much more important for hidden communication than for watermarking.

An inverse Discrete Fourier Transform is performed into the watermarked frequency domain to reconstruct the watermarked in the original video .

3.1.3 Bit-stream domain watermarking techniques

In this technique The watermark is inserted into the compressed video bitstream. the major disadvantage of this technique is that strength of the embedded watermark is constricted by compressed bit-rate.Nandakishore et al. [17] proposes a hard video authentication and sender verification scheme for video sequences compressed using H.264/AVC Main Profile by using digital signatures and cryptographic hash. Features from the transform domain are used as the authentication data for a macroblock. The algorithm can also detect the cause of authentication failure and point out the location of the tampered frames in case of frame tampering.

The method proposed by Pradeep K.et al. [18] detects spatial cropping and temporal jittering in a video, yet is robust against frame dropping in the streaming video scenario. The authentication signature is compact and independent of the size of the video.

3.2 Digital signature

The digital signature invented by Diffie and Hellman in 1976. The digital signature shall depend on secret data which is known by signer [19]. So it cannot be forged and the judge can confirm that the content of video data matches the data contained in the digital signature. The sender first removes the key from the original video and then the data encrypted by a private key that give signature [20]. The receiver can use sender's public key to decrypt the signature to authenticate the received video. The signature is stored somewhere else than the media [19]. And it stored separately in user defined field.

Because the video is stored in a specific format, and the digital signature is being embedded in the video [20].

figure 2.5 shows how to generate a digital signature

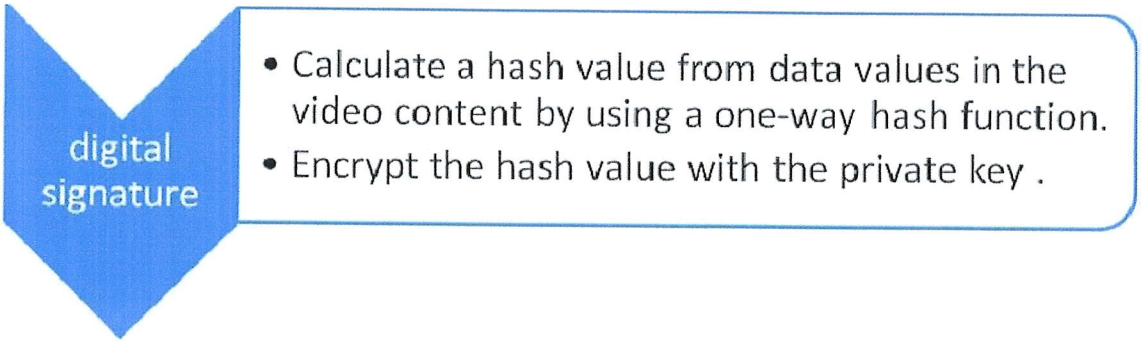


Figure 2.5: general process of generating digital signature

To verify a Digital signature those three steps :

(a) **Step 1:** Extract the encrypted hash value and decrypt the value with the corresponding public key.

(b) **Step 2:** Calculate a hash value of the content in the same manner in first Step of generation process.

(c) **Step 3:** Compare the value decrypted in Step 1(step 1 from verification process) with the one encrypted in Step 2. If the values match, the content integrity has been maintained. If they do not, it has been broken

Digital signature method In last two decades digital signature based technics have been widely used for the purpose of video authentication several sachems have been proposed ,in [21] Tzeng and Wen Hsiang Tsai exploits both color and geometric visual features and also made an attempt to reduce the signature size. The method for digital signature generation is as follows :

- A significant edge detection method was applied to an input image/video. The significant edge detection classifies each non overlapping block into two types, smooth blocks and edge blocks. If there is at least one connected component with size larger than 4 in the corresponding binary image, then block is classified as edge block or else it was called as smooth block.

- Smooth blocks were represented by their mean values. For edge block representation binary edge patterns are used as features. Use of binary edge pattern also prevents the size of The digital signature from explosion.

- At least the digital signature was constructed by cascading all the encoded features extracted from the blocks followed by an encryption key for security purposes which further encrypts the cascaded bit stream.

State of Art In [17] authors have proposed a video authentication and verification scheme for video Sequences compressed using H.264/AVC Main Profile by using digital signatures and cryptographic hash. Features from the transform domain are used as the authentication data for a macroblock. The algorithm can also detect the cause of authentication failure and point out the location of the tampered frames in case of frame tampering

Digital signature based on a hierarchical structure was proposed in [18] using three hierarchical levels of a video(key frame level, shot level, and video level) in a scalable manner. The proposed scheme suitably scales down the authentication process to shot level to achieve robustness against frame dropping in video streaming scenario. The algorithm can also detect speciied region tampering and has been successfully tested over face tampering in a video.

4 Passive tampering detection technics

Passive tampering detection techniques are methods used for detecting the authenticity of a video without depending on pre-embedded information, Works on the basic assumption that video contains naturally occurring properties . Passive video tampering detection methods are classified into the following three categories based on the type of tampering they address : into spatial tampering detection techniques, temporal tampering detection techniques and spatio-temporal detection techniques as figure 2.6 shows

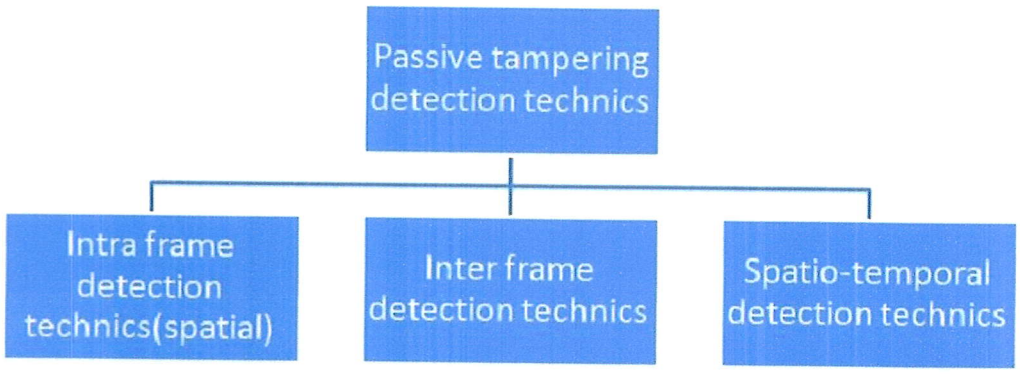


Figure 2.6: Passive tampering detection technics

4.1 intra frame tampering detection technics

In passive video intra frame tampering detection technics can be roughly categorized into five category by farid [22] figure 2.7 shows intra frame tampering detection technics

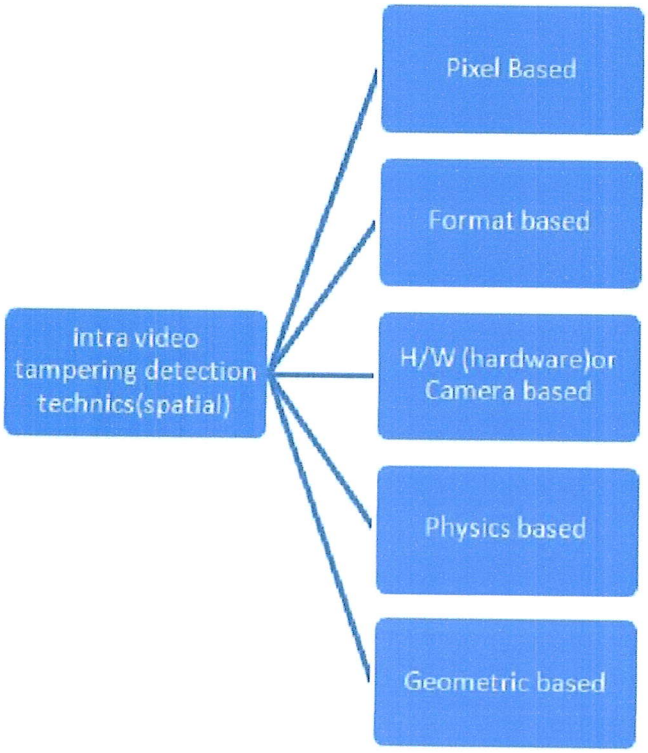


Figure 2.7: Intra frame tampering detection technics

(a) **Pixel Based :** This technique mains to detect statistical anomalies introduced at the pixel level, it is roughly categorized into three types copy move , splicing ,Sampling.

(b) **Format Based :** Format based technique include Double MPEG compression, Frame blocking.

- **Double MPEG compression :** double compression consists on decoding then alter the video and finally encoding the video more details on the double compression in the following sections.
- **Frame blocking :** The blocking artifacts uses pixel value difference within and across block boundaries. The pixel value difference within the blocks is smaller than the across the blocks. When frame is tampered a new set of blocking artifacts may be introduced that do not necessary align with previous block boundaries. These blocking artifacts provide clue for tampering detection [23]

(c) **H/W(hardware) or Camera based :** H/W or Camera based tampering detection uses :

- Sensor Noise
- Color filter array
- Camera response function
- Chromatic aberration
- White balancing
- gamma correction features of Camera used in shooting video.

i. **Sensor noise :** this technique focus on the principal that digital video moves from the camera sensor to the computer memory.

ii. **Camera response function CRF :** Because most digital camera sensors are very nearly linear, there should be a linear relationship between the amount of light measured by each sensor element, and the corresponding final pixel value. Most cameras, however, apply a point-wise non-linearity in order to enhance the final image. Differences in the camera response function across the frame are then used to detect tampering [23].

(d) **Physics Based :** When a splicing of two persons for example is done it is often difficult to exactly match the lighting effects under which each person was originally photographed. Differences in lighting across an image can then be used as evidence of tampering ; physic based methods are divided into two techniques Light direction and Light environment.

i. **Light direction :** The right side of the face in Figure 2.8 shows as in [24] (a) is more illuminated than the left, we can infer that a light source is positioned to the right. This observation can be formalized by making simplifying assumptions: the amount of light striking a surface is proportional to the surface normal and the direction to the light [21] estimation of light source direction in image is limited to 2D because it is usually difficult to determine 3D surface normal from a single image but in videos with knowledge of 3D surface normal, the direction to the light source can therefore be estimated.Ravi et al. [25] demonstrated the presence of ENF signals in video recordings and lightings using optical sensors. It is used as a natural timestamp for video recordings under fluorescent lighting indoors. Statistical correlations are high with respect to ENF signals from the main supply when video is untampered and discontinuities are found when it is tampered, which indicates a tampering.

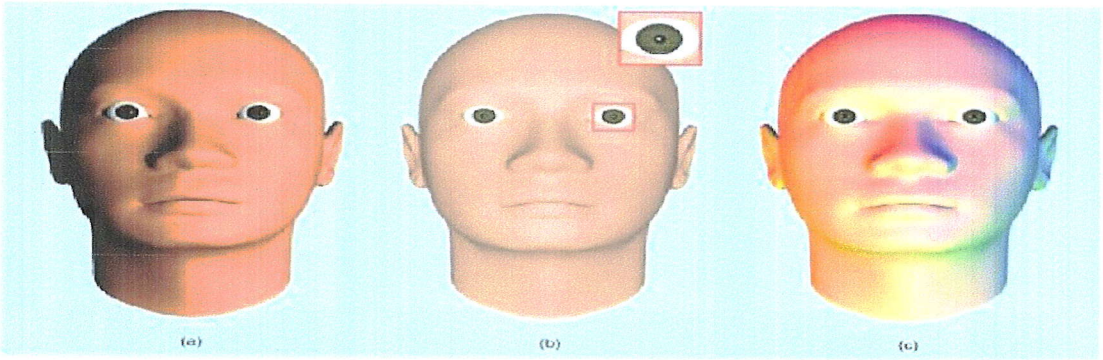


Figure 2.8: Shows light direction how (a)is more illuminated than the left

ii. **Light environment :** In Light direction a simplified lighting model consisting of a single dominant light source was assumed. In practice, however, the lighting of a scene can be complex: any number of lights can be placed in any number of positions, creating different lighting environments like. There are some woks based on this technique witch mainly search to estimate a low-parameter representation of such complex lighting environments.

In [26] Micah K. Johnson and Hany Farid proposed a method for estimating the light source based on assumption that's the light source in a scene gives rise to a specular highlight on the eyes .

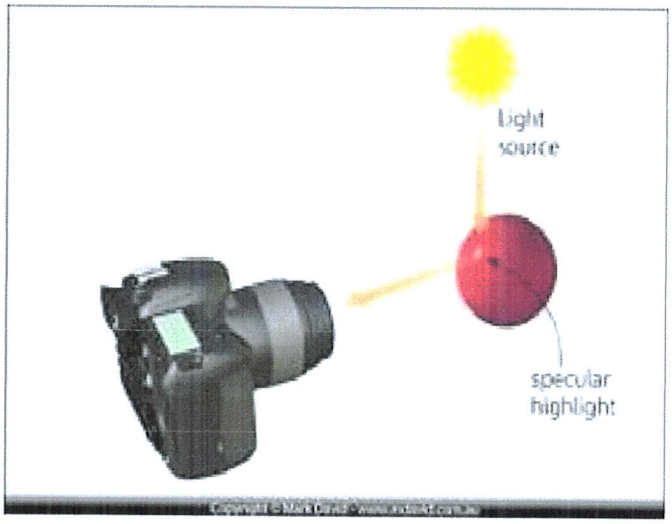


Figure 2.9: show how light source in a scene gives rise to a specular highlight

(e) **Geometric Based :** In geometric based tampering detection we mainly focus on principal point and Metric measurement. The temporal tampering detection in passive video then we can use the concept of motion compensated edge artefacts (MCEA) for I, P and B frames in video.

i. **principal point :** in authentic images, the principal point (the projection of the camera centre onto the frame plane) is near the

centre of the frame. When a person or object is translated in the video , the principal points moved proportionally. Differences in the estimated principal point across the frame can therefore be used as evidence of tampering [23]

ii. **Metric measurement :** These technique aims to rectifying planar surfaces under certain conditions, the ability to make real-world measurements from a planar surface.

4.2 Inter-frame tampering detection technics

As we said previously tampering detection techniques depends on the type of attack, Therefore we can classify temporal tampering detection methods in four categories as shown in figure 2.10

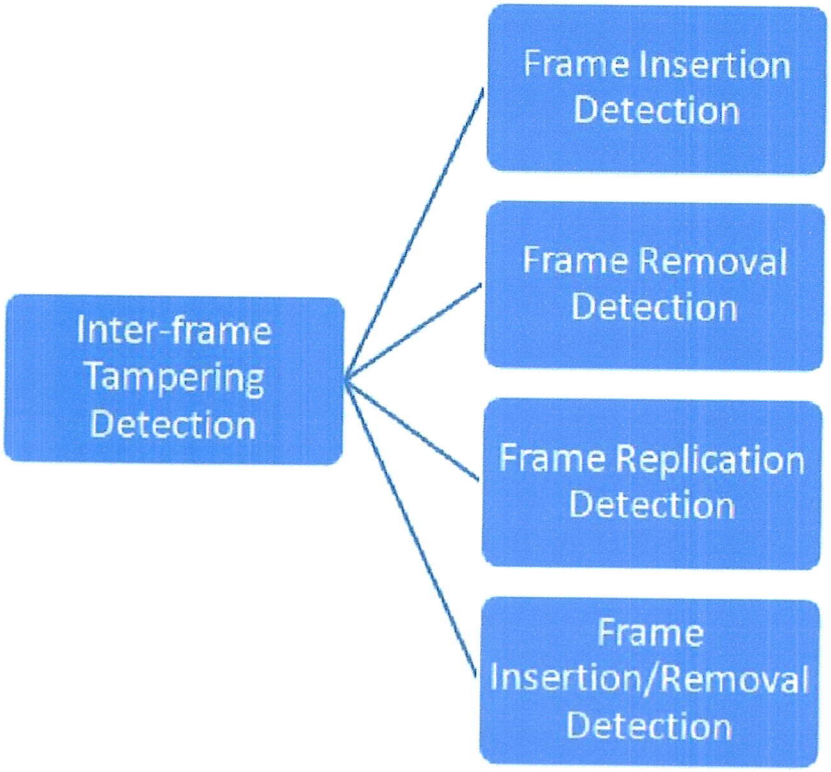
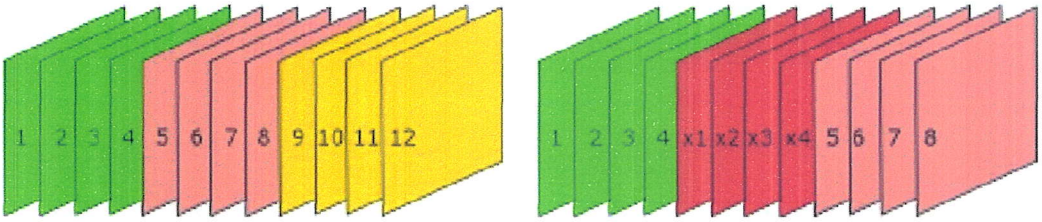


Figure 2.10: Inter-frame tampering detection technics model

4.2.1 Detection of frame Insertion

Frame insertion attack is to add new frame which was not present before to provide fake evidence or any other malicious activity . Detection using Block-wise Brightness Variance Descriptor (BBVD) this technique was proposed in [27] ,authors found that inter frame forgery disturbs the consistency ratio of Block-wise Brightness Variance Descriptor (BBVD) because of decrease in correlation

between adjacent frames. Most importantly, sub-sequence analysis instead of adjacent frame analysis increased the speed of forgery detection. Adaptive threshold was compared with BBVD ratio of each subsequence which generated two peak points at location of frame insertion.



(a) Original Video Sequence (b) Frame Insertion

Figure 2.11: shows (a)original video sequence (b)the same sequence with frame insertion attack

4.2.2 Detection of Frame Removal/Deletion

A technique was proposed for detecting frame deletion in MPEG-2 coded videos in7d based on [29] which utilized prediction error on VBR coded videos. The proposed technique used eight features which were based on prediction error energy, percentage of intra-coded macro-blocks, quantization scale values and estimated PSNR values. Results were analyzed using three types of classifiers (KNN, SVM and logistic regression) on 4 sets of MPEG-2 video sequences. The author in [30] determined the exact location of frame deletion by analyzing spikes in the fluctuation histogram of motion residual. Enhanced Fluctuation Feature (EFF) was measured for a range of frames selected in a window of variable sizes. The forged frames were classified using an adaptive threshold.General threshold of 1.2 and window size of 3 provided effective determination of deletion point. In the meantime, The author in [31] put forward a new feature named Sequence of Average Residual of P-frames (SARP) and analyzed it in time and frequency domain like in [32], [22] respectively. Periodicity of SARP of a video was analyzed in time domain and its spikes were detected in frequency domain using discrete time Fourier transform.

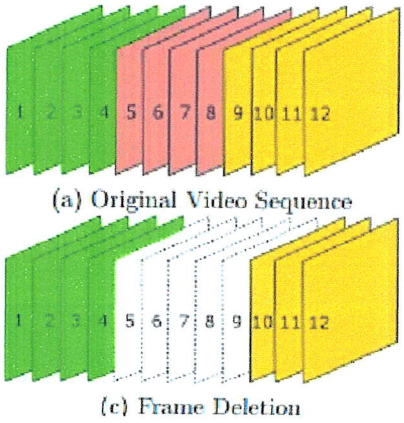


Figure 2.12: shows (a)original video sequence (c)the same sequence with frame Deletion attack

4.2.3 Detection of Frame Replication (duplication)

Frame duplication and region duplication were detected in [33] digital videos based on temporal correlation between all frame pairs in a subsequence and spatial correlation between all block pairs in a frame. Frame duplication was detected by comparing correlation coefficients of a video subsequence with an empirically selected threshold. Region duplication was detected by analyzing peaks in the inverse Fourier transform of power spectrum of two frames.[Base-inter] In [32] **Velocity field consistency** based technique was presented in [34] to detect frame deletion and frame duplication The generalized extreme studentised deviate ESD2 test is applied to identify the tampering types and locate the manipulated positions in the forged videos

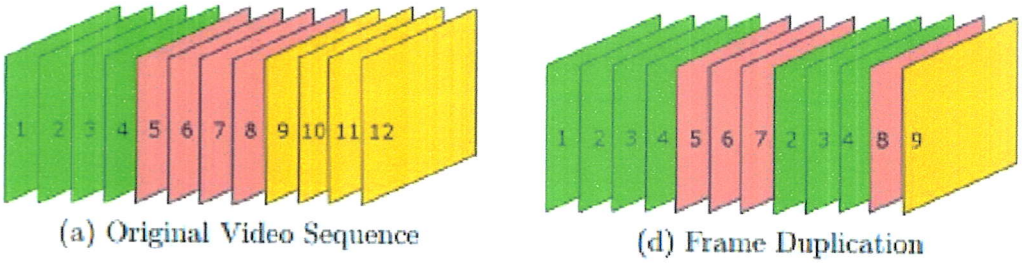
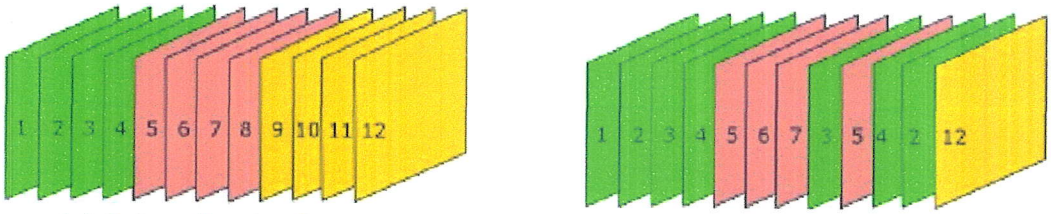


Figure 2.13: shows (a)original video sequence (d)the same sequence with frame duplication attack

4.2.4 Detection of Frame Shuffling

Frame shuffling is a change of the natural ordering of video frames it can be classify with the frame replication so the same detection methods can be applied and gives satisfying results.



(a) Original Video Sequence (e) Frame Shuffling

Figure 2.14: shows (a)original video sequence (e)the same sequence with frame shuffling attack

4.3 spatio-temporal detection technics

It is a combination of temporal and spatial tampering. Where frame sequences and visual contents of a video are modified that was done at scene level it's a intra and inter frame tampering combination many of tampering methods found in spatial and temporal tampering domains are proposed .

4.3.1 Double compression detection methods

Generally MPEG video contain three types frame: I frame , P frame, B frame. I frame is known as Intra frame and have least compression and high quality; P frame is known as predictive frame and have higher compression ratio and less quality in comparison to I frame; B frame is known as bidirectional frame and have highest compression ratio and least quality. I frame of any video is approximately equal to JPEG image. Generally to detect tampering in video the extraction of frames from video is the first step, then try to find some clues from that frames with application of several methods we will address it in the following.

nature of MPEG compression when a video is re-compressed after removing or adding a group of frames, a desynchronization will occur in the GOP pattern. Due to the predictive nature of MPEG compression, all the P frames in a GOP are correlated to the initial I frame. In the re-compressed sequence, some of the frames are likely to move from one GOP to another (last row of Figure 2.15), so their correlationwith the I frame of the newGOPwill be smaller, resulting in larger prediction errors. If a single set of frames is deleted, the shift of P frames will be the same throughout all the video sequence, and

the variability of prediction error in P frames along time will exhibit a periodic behavior. That is, smaller error values will result for frames that remained in the same GOP as the original video, and larger error for those that changed GOP. an example : the first six frames of the original MPEG compressed video (first row) are deleted, thus obtaining a newsequence (second row).When this sequence is re-compressed using MPEG, each GOP will contain frames that belong to different GOPs in the original video (frames highlighted in yellow in the third row).

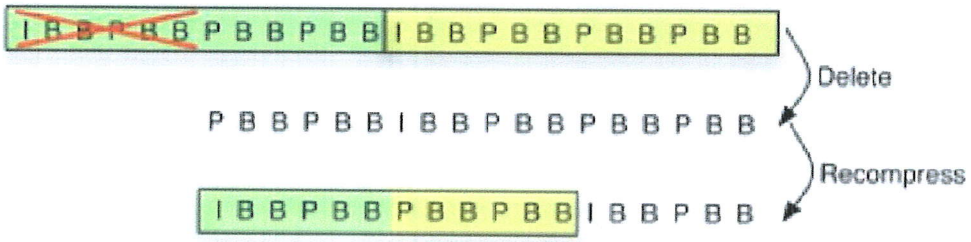


Figure 2.15: shows example of how the GOP become after a tampered alteration

4.3.1.1 Fixed GOP Based Approaches :

4.3.1.1.1 DCT Coefficient Analysis Frame based tampering leads to re-shuffling of frames amongst neighbouring GOPs which causes GOP desynchronization.As a result, DCT coefficient distribution is disturbed. The authors used different approaches [22], [35] to study this disruption in the distribution of quantized DCT coefficients; subsequently overcoming the limitation of the previous approaches. Wang and Farid first proposed a method [22] to detect double MPEG compression in video sequences by analyzing periodic pattern in the histogram of DCT coefficients of I-frames and motion error of P-frames. They did not provide any quantitative analysis of the performance of their approach but claimed that their method worked well if number of frames deleted or inserted was a multiple of 3. Furthermore, the technique's performance dropped if an entire GOP or multiples of GOP were deleted. Thorough analysis of this approach revealed that it also failed to detect forgery at macroblock level where different quantization scales were used during the first and second compression.

The same authors proposed another technique to detect double quantization by extending their former work [22] to macroblock level [36]. This technique detected doubly compressed videos which were either manually compressed or were the consequence of composition of two videos of different qualities, i.e. green screening. The authors here utilized the Gaussian distribution for doubly compressed DCT coefficients to detect double compression in every frame macro-block as small as 16×16 pixels. They used different quantization scales

(in the range 1-31) for the first and second compression; the technique's detection accuracy varied accordingly. The author in [35] analyzed convex pattern in the histogram of double quantized DCT coefficients of each macro-block rather than detecting peak and periodicity. A detection function was defined using empirically selected threshold of 0.1.After thorough analysis of the proposed technique, it was realized that selection of new thresholds for every new video dataset incurred a large computational overhead. The authors tested the technique on 100 video sequences using bit-rates in the range 4 Mbps-8 Mbps.

4.3.1.1.2 Usage of Benford's Law The author in [37] found that the disturbance in DCT coefficients due to double compression also violated the parametric logarithmic law for first digit distribution of quantized AC coefficients. For each GOP, a 36-D feature vector was computed using first digit probabilities of non-zero MPEG quantized AC coefficients and three goodness-to-fit statistics. SVM classifier was adopted to test the approach using different quantization scales in Variable Bit Rate (VBR) encoded videos and using different bit rates in Constant Bit Rate (CBR) encoded videos on three groups having 10 videos in each. Group 1 and Group 2 contained CBR encoded videos and Group 3 contained VBR encoded videos.

The methodology proposed in [38] utilized the same features as in22 and developed a new approach where instead of a 36-D feature, a 12-D feature vector was computed by considering the I-frames only. Experiments were performed on 12 video sequences considering bit-rates in the range 0.5 to 1.5 Mbps. The technique failed to generate accurate results when the target bit rate was smaller than the original bit rate of the given video. It extended their previous work of double JPEG compression detection [39] to videos. They generated a 63-D feature vector using first digit statistics. Videos violating Benford's law were further classified by applying a set of binary SVM classifier on the basis of k-means clustering. The technique detected up to three compressions in H.264/ AVC encoded video sequences.

4.3.1.1.3 Detection Approach using Markov Statistics A Markov statistics based double compression detection method for MPEG-4 videos was proposed in [34], where the authors assumed a fixed GOP pattern IPPPPP. The concept of JPEG recompression detection scheme used in [33] was applied to intra-coded frames of MPEG-4 video blocks. By extracting Markov features from an object based representation model of the given video and an empirically calculated threshold, final classification resulted. Comparative analysis of Markov statistics with first digit distribution demonstrated that Markov features performed better when quantization scales for second compression were an even multiple of the first quantization scale.

4.3.1.2 Variable GOP based Approaches All the techniques mentioned so far assumed a fixed GOP structure for the first and second compression and

as a result, did not work in scenarios where different GOP structures were used during the initial and any of the subsequent compressions. To overcome this limitation, the authors in [40] proposed techniques where the GOP structures were assumed to vary with every compression.

4.3.1.2.1 Detection using Block Artifact Strength (BAS) Compression introduces different block artifacts into video frames and recompression further disturbs the average of these artifacts; BAS score was used to quantify this variation. BAS depends on the number of deleted frames and type of GOP used in first and second compression. Along with the detection of videos re-encoded using different GOP structure, this approach also detected frame removal if number of deleted frames were not multiple of 3.

4.3.1.2.2 Detection using Variation of Prediction Footprint (VPF) The author in [41] proposed a technique utilizing the feature called VPF which was based on the variation in number of I and S macro-blocks in re-encoded P-frames which were I-frames in first encoding. This method also estimated the size of GOP used during first compression. Experiments were performed on videos encoded using three different encoders (MPEG-2, MPEG-4, H.264).

Each encoding was performed by specifying four different constant bit rates (100, 300, 500, 700 kbps). This method gave best results for short videos having uniform regions and for H.264 encoded videos re-encoded at high bit rates.

4.3.1.2.3 Detection using both BAS and VPF With MPEG-4 compression, discontinuity exists in 8×8 block boundaries as quantization and transform coding is different for each block. Instead of analysing variation in the number of macroblocks, VPF was measured by analyzing the variation in the block artifact strength.

After testing using bit rates of 100, 300, 500 and 700 kbps.it can be concluded that Chen and Shi [37] provided better accuracy for double compression detection in MPEG-1 and MPEG-2 videos recorded using both CBR and VBR modes. However, most of the techniques were not tested on MPEG-4 videos. On the other hand, The author in [40] detected double compression in MPEG-4 videos recorded in CBR mode with substantial accuracy, but this accuracy was dependent on the target bit rates of the final forged videos.

4.3.2 Region Tampering Detection

It methods that gives information about the location of tampering in the spatial as well as temporal domain such as : copy-paste/region duplication tampering and splicing in the video are discussed. By copy-paste or copy-move tampering, it means copying a small portion of the frame and pasting at an other location in the same frame, or copying particular regions from sequence

of frames and pasting it at another sequence of the same video as shown in figure 2.16 shows an example of copy-move tampering from the SULFA dataset provided by [42], the images in the left side of the figure correspond to succession of frames of an original video and those on the right correspond to the frames at the same position in a tampered video where the presence of a lady in the scene is concealed by copy-move tampering.

Figure 2.16: (a,c,e,g,i)frames of an original video, (b,d,f,h,j) corresponds to the same frames with the copy-move tamering attack

Splicing in video tampering indicates the insertion of foreign objects or copying of frame regions from a different video and pasting to the target video frames In this figure 2.17,(a)the red frame regions indicate tampering.and (b)green frame regions stand for the pasted portion which is copied from frames of different video.

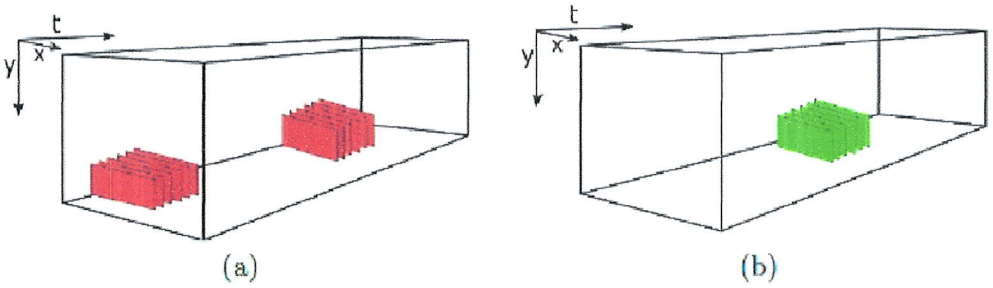


Figure 2.17: shows (a)the red frame regions indicate tampering.(b)green frame regions stand for the pasted portion which is copied from frames of different video.

5 conclusion

Because it's difficult to detect a tampered video everyday more detection technics and methods are implemented and developed despite that these methods does not include all types of tampering.However,copy-move attack is the most common attack and the more used there for we saved the next chapter about the state of art of copy-move video tampering detection technics and methods.

Chapter 3

Copy-Move Detection technics and methods

1 Introduction

Copy-move tampering is a common type of video tampering, containing two types: spatial tampering and temporal tampering. In spatial tampering, a region may be pasted to a different location on the same frame or other frames. In this way, the tampering aims to replace or hide the undesired object will be achieved. While in temporal tampering, multiframe is replaced by the copy of previous ones, having the scenes replaced without affecting the continuity of the video, or pasted to a different location having the scenes occurred ahead or delay. Now, different approaches are developed to detect copy-move tampering, and all of them are based on the same principle that a copy-move tampering brings a high correlation between the original frames and the duplicated ones.In allusion to the existing detecting approaches, high calculating complexity and high false alarm rate still exist. in this chapter, at beginning we are gone to discuss the copy-move detection methods that existe in earlier works then we will talk and explaine the method that we chose.

2 Existed methods of copy move detection

Different approaches are developed to detect video copy-move tampering, most of them are based on the same concept that a copy-move tampering brings a correlation between the original frames and the duplicated ones. In allusion to the existing detecting approaches, high calculating complexity and high false alarm rate still exist.

2.1 Detection of copy–move video tampering using histogram of orientated gradients(HOG)

It's a technique that detect copy-move tampering in spatial and temporal domain which is difficult and challenging to detect tampered video it's may drastically vary in terms of size, compression rate and compression type(I, B or P) or other changes such as scaling, filtering compression, translation. In this technique histograms of oriented gradients (HOG) is used , mainly The benefit of using HoG features is that they are robust against various signal processing manipulations

2.1.1 Principle of histograms of oriented gradients (HOG)

HOG was first developed by Dalal et al. [43] as a robust feature descriptor for object detection in computerized vision systems, they represented the frame by a set of local histograms. These histograms count the number of occurrences of gradient orientation in a local spatial region of the frame known as cell in HoG feature extraction. the RGB frame must convert to gray-scale then the frame will be convolved with horizontal mask [-1 0 1] and vertical mask [-1 0 1]^T.

$$Gx = [-1, 0, 1] \times I(x, y)]$$

Gx : horizontal gradient

$$Gy = [-1, 0, 1]^T \times I(x, y)$$

Gy : vertical gradient

The gradient magnitudes of the pixels in the cell are used to vote into the orientation of histogram. Several cells form a group, called a "block". The magnitude of the gradient at a given point (x, y) is determined as follows:

$$G(x, y) = \sqrt{Gx(x, y)^2 + Gy(x, y)^2}$$

The orientation of the edge at a given point (x, y) is derived as follows:

$$\theta(x, y) = \arctan \frac{Gy(x, y)}{Gx(x, y)}$$

The calculation of normalization is useful to contrast-normalize local responses to facilitate invariance to illumination and shadowing. This can be accomplished by

accumulating a measure of local histogram "energy" over "block". The result is used to normalize each cell in a given block, such that the cell appears several times in the final output vector with various normalizations. The normalized block descriptors are defined as the HOG.

And finally HOG descriptors are collected from all blocks within the detection window for assembly into a feature vector for use in detection.

2.1.2 Work process

Figure 3.1 shows work processing of HOG method

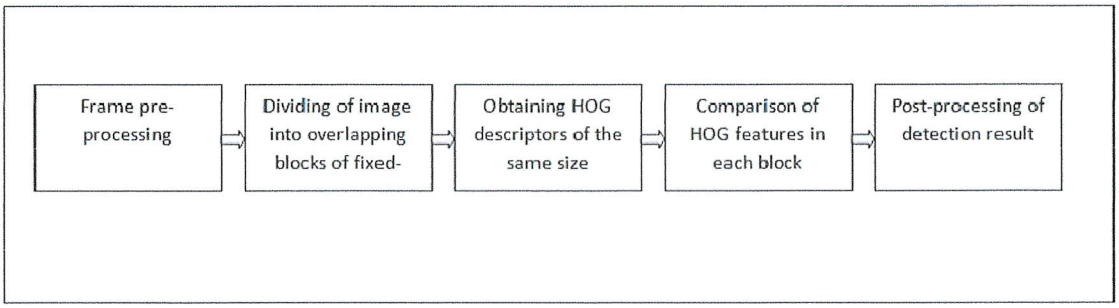


Figure 3.1: work process

frame pre-processing

Initially, RGB image is transformed into grayscale frame

$$Y = 0.59G + 0.29R + 0.12B$$

Y Represents the luminance. It contains the essential information of the image R, G, B are red, green, and blue

Dividing of image into overlapping blocks of fixed-size

To identify forged regions, the image is divided into overlapping square sub-blocks. The grayscale frame of for example M× N will be devided into sub-blocks of L× L for the calculation of HOG descriptors. The frame is then devided into (M-L+1)×(N-L+1) overlapping blocks.

Obtaining HOG descriptors of the same size

After applying HOG to each block, an HOG descriptor matrix the same size as the block is assembled to represent each corresponding block. in local histogram there is four bins. The channels of the histogram are evenly spread between 0 and 180 degree, such that each histogram bin corresponds to an orientation of 45 degree. The resulting cell histograms are then combined into a descriptor vector for each block, wherein four features can be used to represent each block. Thus, an LL block

is represented by a 14 feature vector vl = (x1, x2, x3, x4). For an frame of size MN, matrix A would include (M-L+1)(N-L + 1) rows and four columns, where four represents the number of features.

Comparison of HOG features in each block

Identifying duplicated block pairs with the same or similar features within a reasonable time frame can be exceedingly difficult.the lexicographical sorting of feature vectors, wherein similar features are located in different blocks. To reduce the time required for matching, similar feature vectors are stored in neighboring rows. Thus, detection can be achieved through lexicographical sorting of the rows in matrix A, such that the features of the duplicated block pairs appear successively. In this case, the lexicographically sorted matrix is denoted as b. With employment of block matching to match corresponding blocks and identify regions that are likely to have been forged. corresponding blocks are identified by estimating the Euclidean distances of the feature vectors. In order to accurately identify the forged region, the distance threshold Td and the threshold of similarity Ts should be predetermined. There are distinct similarities in the feature vectors of blocks with overlapping pixels; therefore, only blocks with a distance larger than the length of block L are compared. In this manner, distance threshold Td is defined according to the length of block L. The matching of the blocks begins in the first row of matrix β.

For a feature located in the ith row β_i , distances in the vicinity of ith rows are computed, and the smallest distance, denoted by D(i,Σ), between the ith row and nearby ith rows is obtained as follows:

$$D(i,\Sigma) = minD(i;i-j),\ldots,D(i;i-1),D(i;i+1),\ldots,D(i;i+j)$$

In this study, we set j = 5. If D(i,*Σ) is smaller than threshold Ts, then the corresponding blocks will be regarded as correctly matched, whereupon the locations of the two blocks are stored. The matching process is repeated for all rows of β. Finally, all of the matched block pairs are saved in set ω.

Post-processing of detection result

When all matched block pairs are saved in set x, the forged regions can be identified by marking the copied region and the modified region and removing the isolated blocks. Generally, all of the detected blocks, including the original blocks and forgery blocks, are marked to generate the final detection result this method is very performing but the high dimensional features of HOG lead to a higher complexity of the algorithm.and it demand an extensive background in computer vision and video encoding that's why we explore other options.

2.2 Copy Move Detection Technic with Automatic Threshold Determination

Beste et al.In [44] used DCT-phase terms to restrict the range of the feature vector elements' and Benford's generalized law to determine the compression history of the

image under test. The method uses element-by-element equality between the feature vectors instead of Euclidean distance or cross correlation and utilizes compression history to determine the threshold value for the current test image auto matically. Experimental results show that the method can detect the copied and pasted regions under different scenarios and gives higher accuracy ratios/lower false negative compared to similar works.

The method consists of four parts:figure 3.2 shows those parts

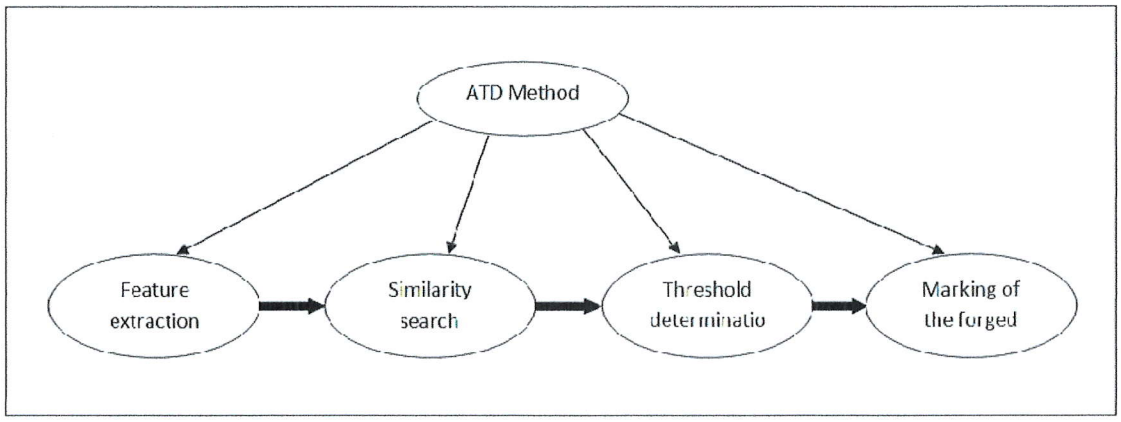


Figure 3.2: The four parts of the ATD method

Feature extraction: Input image is transformed into YCbCr space and divided into overlapping sub-blocks. Sign blocks are obtained from each overlapping block via DCT phase terms and zigzag scan extracts predetermined number of elements from the sign blocks to create the corresponding feature vectors. Feature vectors are lexicographically sorted and stored in a matrix.figure 3.3 gives the Steps of the feature extraction method

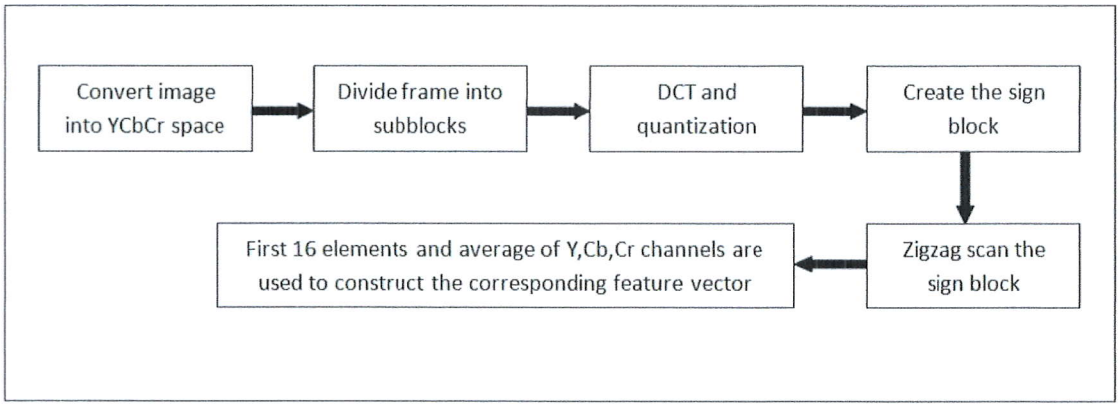


Figure 3.3: Steps of the feature extraction method

Similarity search: Each feature vector in the matrix is consulted to find the similar vectors. The method uses element-by-element comparison to test the similarity. Similarity search uses a threshold to judge whether

feature vectors represent the same blocks or not. The threshold will be determined automatically according to characteristic of the test image. If any two vectors are similar, corresponding shift vector will be calculated and recorded.

Threshold determination: The method determines whether the test image has been compressed before or not. If it is, the method finds the quality factor used for compression. Threshold value will be determined according to compression history.

Marking of the forged regions: If the number of shift vectors exceeds predefined threshold, corresponding regions are marked by the method.

2.3 Video Copy-Move Detection and Localization Based on Tamura Texture Features

Tamura et al. [45] proposed an image texture features descriptor. Tamura texture features consist of six components but only three of them are used in the Copy-Move forgery detection whitch is good expression of the visual image texture features so in this method the extraction of these three components of the frame as features for tamper detection. the eigenvector matrix of the video is made up by extracting the Tamura texture features of every video frames. Then the eigenvector matrix of video is sorted by dictionary ordering and computing the differences between the feature vector and the adjacent feature vectors. And when the differences between two feature vectors less than the threshold, compare their distance of the serial number with the threshold, record the pairs of the serial numbers which is greater than the distance threshold. Finally, locate the copy-move sequences. The method Is implemented in three phases like figure 3.4 Shows

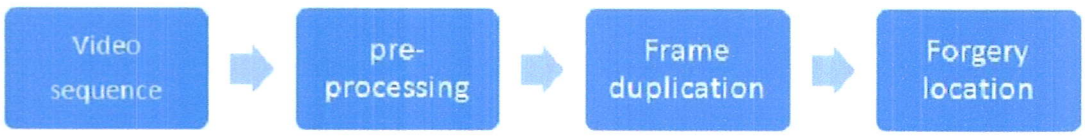


Figure 3.4: Steps of the feature extraction method

pre-processing Compose the video to a group of pictures size n, and then extracting the Tamura texture features, using the contrast, orientation, roughness three values constitute the corresponding three-dimensional feature vectors. Finally, dictionary sorting the eigenvector matrix composed of the corresponding feature vectors.

Frame-duplication Detection For detect whether two frames are duplicated It is necessary to calculate the distance Euclidean of their corresponding feature vectors for measure the similarity of those two frames

Forgery Location Calculate the similarity between the first and the last frame of the original sequence and the duplicated one with the adjacent frame respectively to locate the forged sequence

2.4 Detection of Duplication in Digital Video Tampering

Wang et al. [46] used the similarity in the temporal and spatial correlation matrices, embodying the correlations of short sub-sequences, as evidence of detect duplicated frames in a full-length video.there is two methods in this technics :frame duplication and region duplication

Figure 3.5 Shown in the left column are seven frames of an original video. Shown on the right are the results of duplicating three frames so as to remove the flight attendant from the scene.

Figure 3.6 : Shown in the left column are seven frames of an original video. Shown on the right are the results of region duplication so as to add another zebra into the scene.

2.4.1 frame duplication :

This type of manipulation is fairly easy to perform and can be difficult to detect visually particularly in a video taken from a stationary surveillance camera it would be computationally intractable to search for duplication by comparing all possible sub-sequences of arbitrary length and positions in time. An additional difficulty in searching for duplication is that compression artifacts (MPEG) introduce differences between initially identical duplicated frames. therefore, describe a computationally efficient algorithm for detecting duplicated video frames that is robust to compression artifacts the basic approach is to partition a full-length video sequence into short overlapping sub-sequences. A compact and efficient to compute representation that embodies both the temporal and spatial correlations in each sub-sequence is then extracted and compared throughout the entire video. Similarity in the temporal and spatial correlations are then used as evidence of duplication

2.4.2 region duplication :

A subset of pixel of unknown location is taken which are are duplicated and placed in another frame at a different spatial location, these pixels undergo no other geometric or significant intensity changes. Then estimate the shift by given a pair of frames and finally verify if that estimated shift corresponds to duplication.there are two ways to do this technics :the first is with stationary camera and the second with moving camera.

Figure 3.5: example of frame duplication

Figure 3.6: example of region duplication

2.5 Detection of video tampering using correlation of noise residue

A passive tampering detection in a digital video based on the statistical property of noise residue. We propose to analyze the temporal correlation of block-level noise residue to locate the tampered regions of a video. Our method does not need to pre-collect and pre-train the statistics of noise residue for specific video cameras as the noise residue information can be easily on-the-fly extracted from the video to be authenticated. Chih-Chung et al. [47]propose to model the distributions of temporal noise correlation values of video blocks in forged and normal regions using a GMM model. In this method, the GMM model parameters are estimated using the Expectation-Maximization (EM) algorithm so that optimum thresholds can be derived accordingly using a Bayesian classifier. Two video inpainting schemes are used to evaluate the performance of the proposed method.

work process

In the first step the noise residue of each video frame is extracted by subtracting the original frame from its noise-free version.In the second step, each video frame is first partitioned into non-overlapping blocks of size N $\times$ N. The correlation of the noise residue between the same spatially indexed blocks of two consecutive frames is then computed as illustrated The final step is to locate tampered blocks by analyzing the statistical properties of block-level noise correlations figure

3.7 shows the process of the proposed method

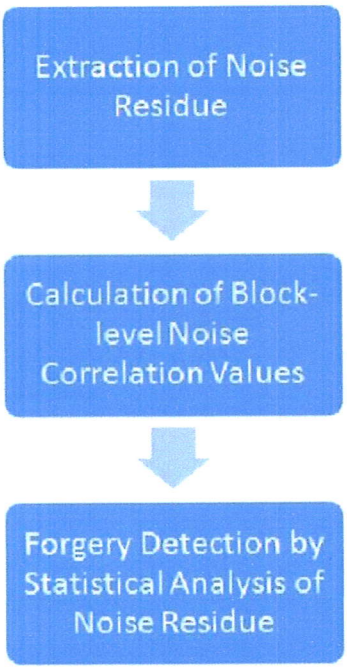


Figure 3.7: work process of the method

2.6 Video Copy-Move Detection and Localization Based on Structural Similarity

According to this method's name it's based on structural similarity. Structural similarity is extend to measure the similarity between two frames of a video. Since the value of similarity between duplicated frames is higher than that between the normal inter-frames, a temporal similarity measurement strategy between short subsequences.

2.6.1 MSSIM Mean Structral Similarity

The SSIM metric measures the similarity with three statistical components, which are luminance comparison, contrast comparison, and structural comparison. Let Y be the distorted image of X, for any two pixels x in X and y in Y, the SSIM metric is as follows :

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{3.1}$$

where $\mu_x$and μ_y are the means of the local windows with a size of 11×11 centered at x and y espectively.

σ_x and $\sigma_x$are the standard variance

$\sigma_{xy}$is the covariance of the two windows. C_1, C_2 are small constants

Then the mean SSIM (MSSIM) is used to evaluate the overall image quality :

$$MSSIM(x,y) = \frac{1}{M}\sum_{i=1}^{M} SSIM(x_i, y_i) \tag{3.2}$$

where M the number of local widows in the frame, the MSSIM can be used to evaluate the video sequence for finding whether copy-move operation is done or not.Figure 3.8 shows the procedure of similarity measure between two images.

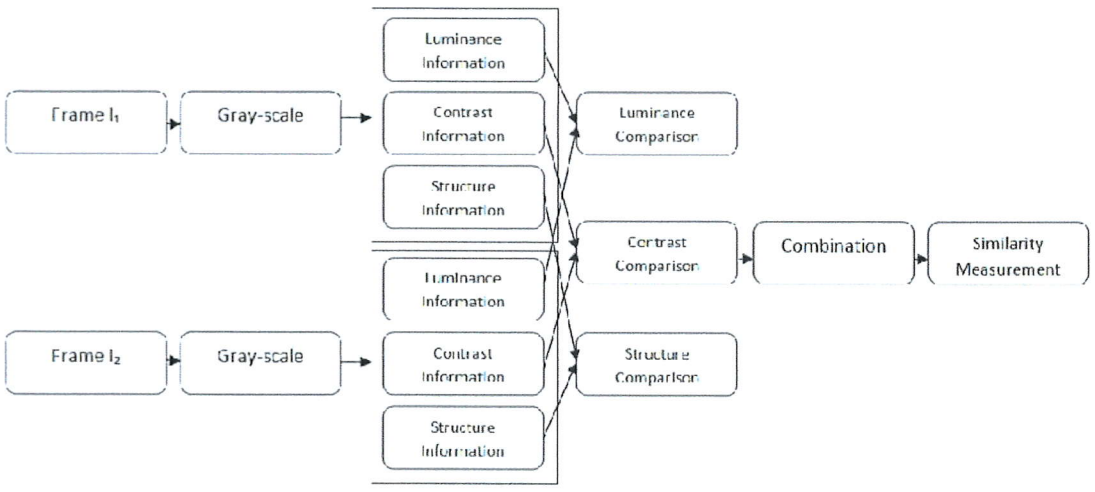


Figure 3.8: Similarity measure between two images

2.6.2 work principal

In temporal domain

The temporal copy-move tampering can be detected with this method.In [48] The input video is converted to frames and each frame is converted to gray scale images. Then these frames are grouped into overlapping blocks or sub-sequences. Each block is taken as a reference block and the similarity reference block and subsequent blocks are measured. The value of similarity measurement is compared against the threshold. If the MSSIM values between all the corresponding frames in reference block and subsequent block are greater than the threshold value, which is set by the user, the block is detected as copied.the process is shown in figure 3.9 diagram for temporal copy-move detection

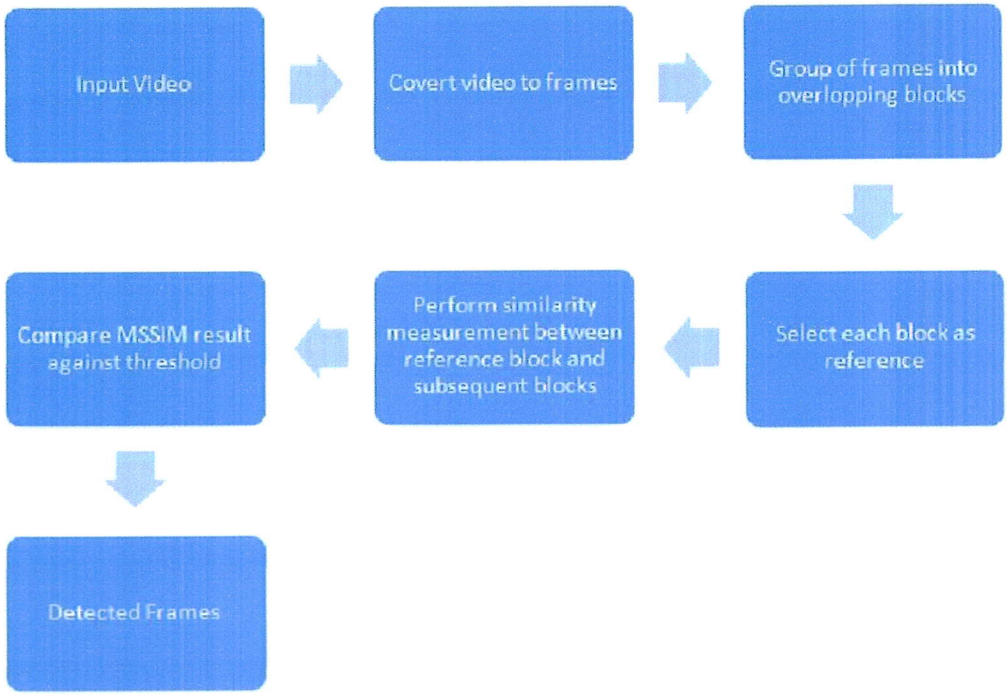


Figure 3.9: diagram for temporal copy-move detection

In spatial domain

For any of the two images I_1 and I_2 in the video sequence, after the convert of frames from color to gray-scale for reducing computation time, and the luminance information of gray-scale frames are extracted. Then we remove the luminance information of the images to calculate the contrast information. Finally, to calculate structural information divide by the contrast information is needed. The MSSIM will be obtained to measure the similarity between two images.Figure 3.10 shows diagram for spatial copy-move detection

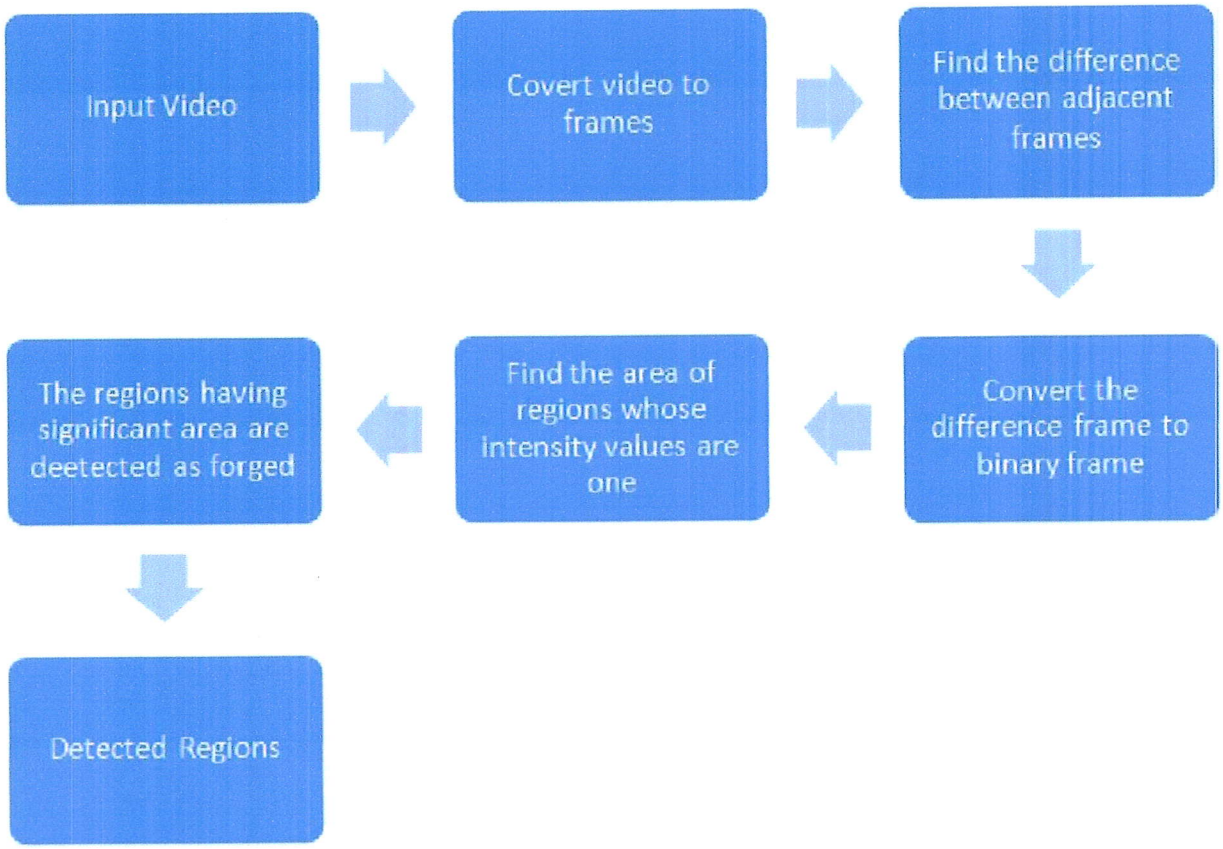


Figure 3.10: diagram for spatial copy-move detection

2.7 Similarity Analysis

This methods detect frame duplication attack it's based on Similarity analysis that is implemented in two stages. In the first stage, the features of each frame are obtained via SVD (Singular Value Decomposition)[1]. Next, the Euclidean distance is calculated between features of each frame and the reference frame. After dividing the video sequence into overlapping sub-sequences, the similarities between the sub-sequences are calculated, and then our algorithm identifies those video sequences with high similarity as candidate duplications. In the second stage, the candidate duplications are confirmed through random block matching.

General work process

The algorithm consists of two stages:

(a) selects the candidate duplicated sequences of the video with three steps :

[1]The singular value decomposition of a matrix A is the factorization of A into the product of three matrices A = UDV T where the columns of U and V are orthonormal and the matrix D is diagonal with positive real entries

i. obtain the features of each frame.

ii. calculate the Euclidean distance of the features.

iii. calculate the similarities.

(b) double-checking; is used to confirm the candidate duplicated sequences obtained from the first stage :

i. merge the sub-sequences.

ii. confirm the candidate duplications.

iii. perform localization.

Figure 4.4 present a flow chart of the process

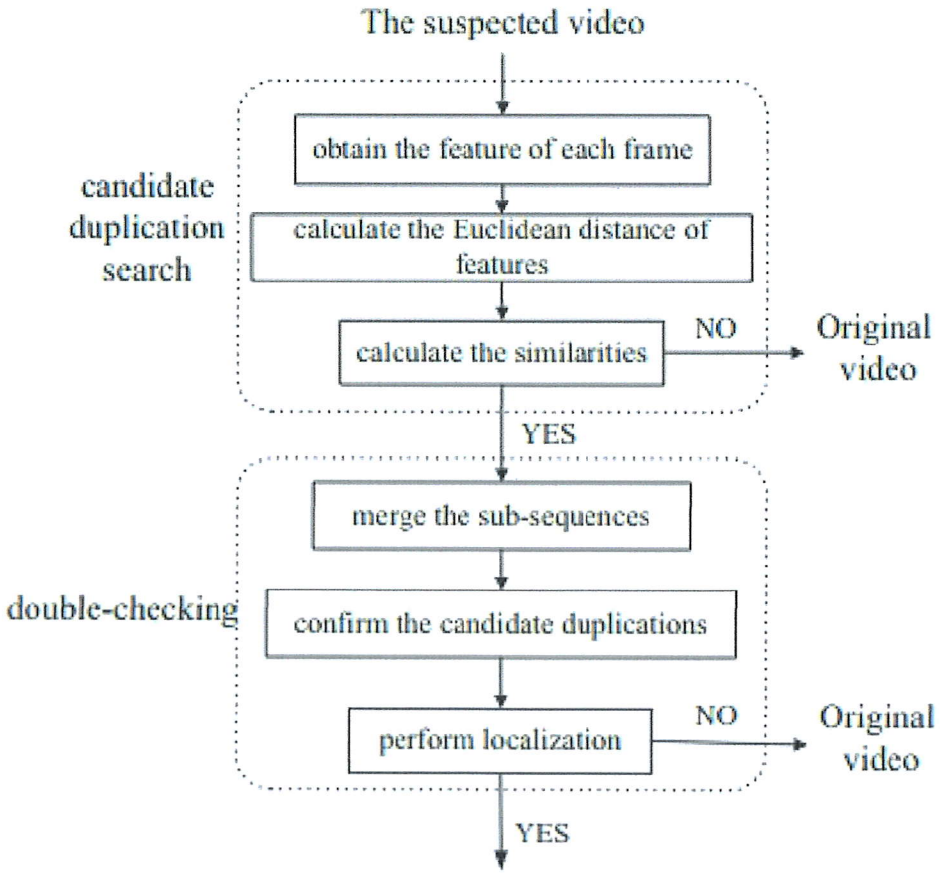


Figure 3.11: The flow chart of the similarity analysis algorithm

2.7.1 Candidate duplication search

The object of the candidate duplication search is to determine whether the input video has been alter with frame duplication by identifying the video sequences with high similarities to the candidate duplications.

(a) **Obtain the features of each frame :** Considering a video sequence of frames I , the i^th frame(random frame) is expressed as I_i. The reference frame is the first frame of the video, I_1 is expressed as I_refer after gray-scale of the frames SVD(singular value decomposition), givin by :

$$I_i = UX_iV^* \tag{3.3}$$

where U and V^* are the unitary matrices

X_i is the singular value of I_i which is a diagonal matrix.

From the diagonal elements of X_i a one-dimensional vector can be formed;the vector can be expressed as $X_i = x_{i1}, ..., x_{ir}$

X_i is considered a feature of I_i

(b) **Calculate the Euclidean distance of features :** After obtaining the features $X_i = x_{i1}, ..., x_{im} (i = 1, 2, ..., L)$ of I_i, the Euclidean distance between I_i and I_{refer} wich is the feature distance between them :

$$Fdis(i) = \left[\sum_k = 1^k = r(x_{ik} - x_{refer})^2\right]^{\frac{1}{2}} \tag{3.4}$$

$Fdis(i)$ is the new feature of I_i.

(c) **Calculate the similarities :** To calculate the similarities in the video and identify the sub-sequences with high similarities, the feature distances for all overlapping (by one frame) sub-sequences are first obtained.

Assuming that n is the lenght of sub-sequence and the lenght of video sequence is L that can be divided into m(m=L-n+1)sub-sequence, the feature distance of a sub-suquence in the video is:
g(i)=Fdis($\tau + i$)$i \in [0, n-1]$ Next, we calculate the similarities between these sub-sequences. Each sub-sequence has to be calculated the similarities with the rest of the sub-sequences. The correlation coefficient is used as a measure of similarity. The correlation coefficient between two vectors $\vec{u}$ and $\vec{v}$ (or matrices)is given by:

$$c(\vec{u}, \vec{v}) = \frac{\sum_i (u_i - mu_u)(v_i - \mu_v)}{\sqrt{\sum_i (u_i - \mu_u)^2}\sqrt{\sum_i (v_i - \mu_v)^2}} \tag{3.5}$$

where u_i and v_i are the i_{th} elements of $\vec{u}$ and $\vec{v}$

μ_u is the means of $\vec{u}$ and μ_v is the means of $\vec{v}$

Assume that g_i and g_j are the respective feature distances of the sub-sequence i^{th} and j_{th} sequence and calculate their similarity according to

formula 3.5. And define the feature distance correlation matrix Corr as an m×m symmetric matrix whose$(i,j)^{th}$ entry is the correlation coefficient between the i^{th} and j^{th} sub-sequence,and also it's the similarity between those two.The value of $Corr(i,j)$ is within [-1,1].

To solve the similarity of adjacent sub-sequences is high even though they are not duplicates, the weight deviation matrix Devi (m×m) eliminates the influence of adjacent sub-sequences

$$Devi = \begin{cases} 0.2 & \text{i=j+1;} \\ \frac{1}{(i-j)x} & \text{i > j > m} \end{cases}$$

(3.6)

$$CorrDevi = Corr - Devi$$

Any two sub-sequences with a correlation above a specified threshold (close to 1) are considered candidates for duplication.

2.7.2 Double-checking

Due to visual retention, a video is considered smooth only if the frame rate is greater than 16 fps. For most non-high-motion videos, we assume that the number of frames that will be altered in a video will cover at least more than 1 s or approximately 20 frames.

(a) **Merge candidate sub-sequences :** In this stage if two pairs of condidate sub-sequences obtained previously are adjacent the merge of those condidate sub-sequences gave a candidate sub-sequence, i.e. (sq_{t1},sq_{r1})and (sq_{t2},sq_{r2}) are two pairs of adjacent condidate sub-sequenceswith the merge (sq_t,sq_r).as figure 3.12shows

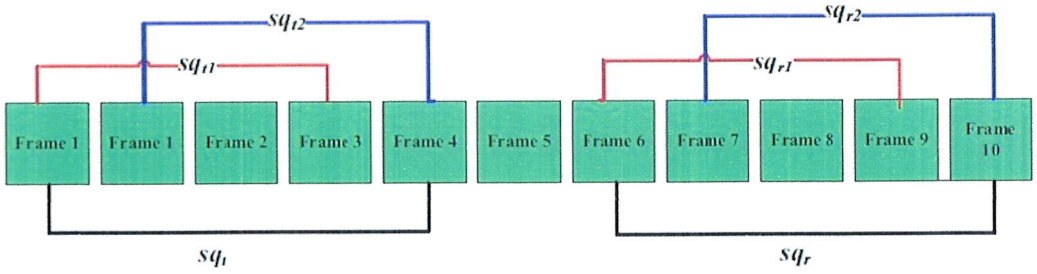


Figure 3.12: Merging of sub-sequences in a video sequence

(b) **Confirm duplicated sequences :** After merging the candidate sub-sequences each frame in each sub-sequence is divided into blocks of 8×8, then generate a random number R (number of block used)After computing the average gray value of each block selected.Finally,the calculate the

correlation between the two matrices to obtain the coefficient correlation $C(i)$(i is the duplicated sequences)if the matching frames are duplicated, then $C(i) = 1$; otherwise,it's not duplicated

(c) **Perform localization :** the similarities between adjacent video frames are stronger than those between non-adjacent frames.The SSIM [49](structural similarity index measurement) is used to measure the similarities $s[i]$ between two adjacent frames (I_i, I_{i+1}) as follows:

$$s[i] = SSIM(I_i, I_{i+1}) \tag{3.7}$$

where $i = 1...N-1$,I_i is the i^{th} frame in the video,and N is the total frame number in th video.

Assume that P_1 is the m^{th} frame of the video, which is denoted as I_m, and Q_1 is the n^{th}frame, which is denoted as I_n. Thus, the similarities $s[m-1]$ and $s[n-1]$ are calculated. If $s[m-1]$ is greater than $s[n-1]$, then P is considered the source sub-sequence and Q is the duplicate;otherwise, Q is a source sub-sequence and P is a duplicated sub-sequence.as figure 3.13 shows

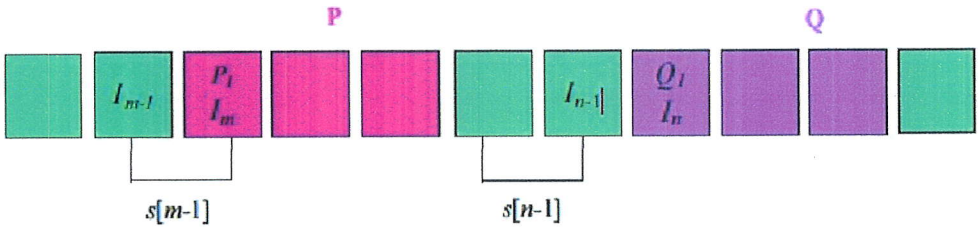


Figure 3.13: localization of duplications

3 Conclusion

this chapter present different methods and technics that are developed to detect the copy-move digital video tampering for both spatial and temporal tampering.Most of those methods still have high calculating complexity and high false alarm rate.The method based on Structural Similarity was our first choice but with the implementation of the algorithm calculating complexity was too high ,we finally chose the video Copy-Move Tampering Detection and Localization Based on Similarity Analysis method that is presented in section 2.7 to detect and localize the copied frames within video correctly.

Chapitre 4

Experimental results of Similarity Analysis

1 introduction

in the previous chapter we have presented a state of art about the copy-move detection technics and methods In this chapter we implemented the copy-move detection method based on the similarity analysis between frames of a video that is implemented in two stages : select the candidate duplicated sequences of the video, double-checking.

1.1 The computer configuration

processor : Intel(R) Core(TM)2 Duo CPU T6670 @ 2.20GHz 2.20GHz

RAM : 4Go

OS : Microsoft Windows 7 32 bits

Codding : Python 2.7

1.2 Daraset description

This dataset is composed by 20 video sequences used in [50] 10 original and 10 forged ones. Each sequence has a resolution of 320x240 pixels, and a frame-rate of 30 fps. Original sequences have been recorded using low-end devices, thus they have all been compressed at the origin (using h264 q30). Forged sequences have been saved as uncompressed file (RV24, 24 bit RGB).

videos were downloaded from SULFA (Surrey University Library for Forensic Analysis [51]

1.3 user interface presentation

After the click on Open buttom

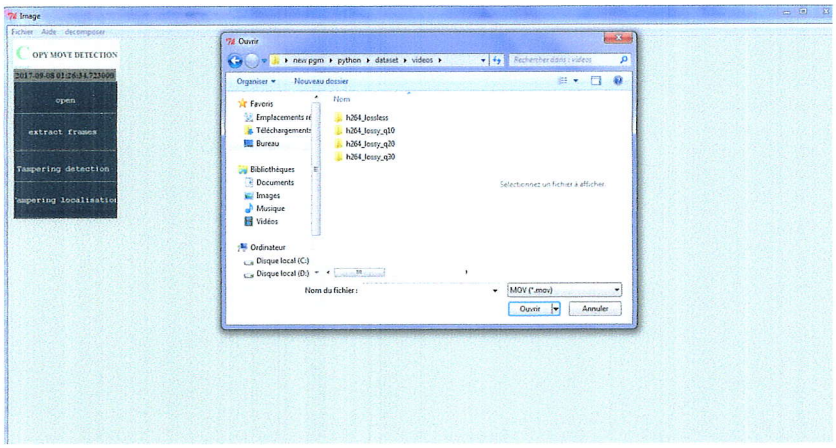


FIGURE 4.1 – user interface 1

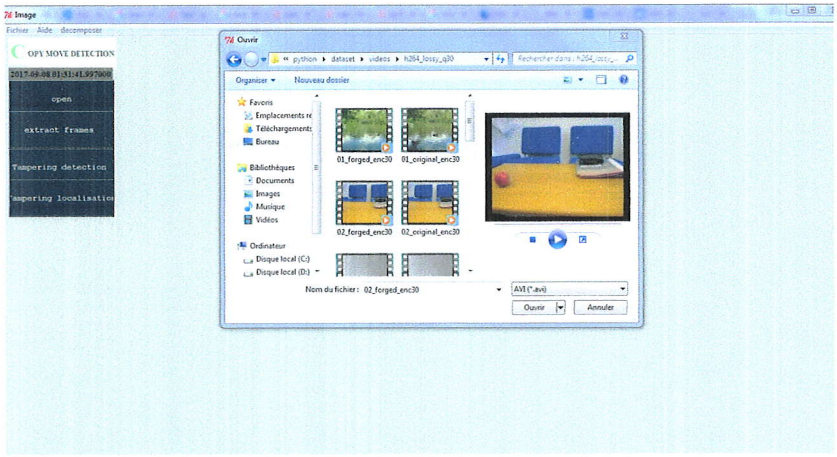


FIGURE 4.2 – user interface 2

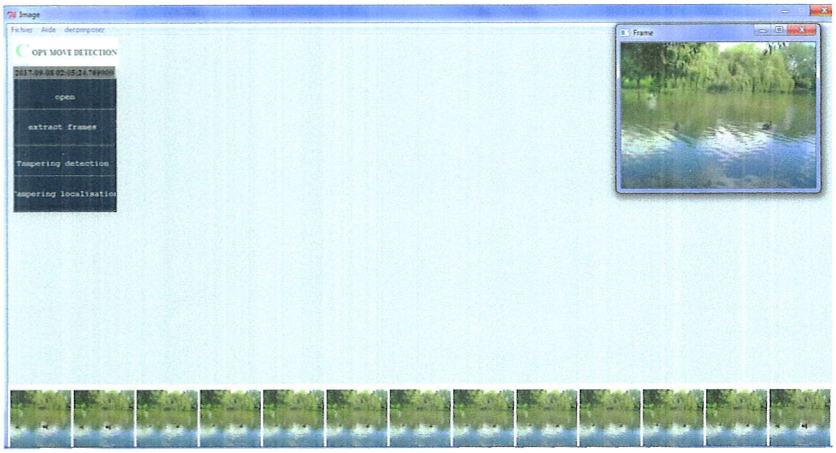


FIGURE 4.3 – user interface 3

2 implementation

FIGURE 4.4 – The flow chart of similarity analysis algorithm

2.1 selects the candidate duplicated sequences of the video

2.1.1 Obtain the features of each frame

After graying each frame we obtain the feature vectors via SVD, this step consists in decreasing the data to improve the response time

```
def svd_1d(A):
    ''' SVD
'''

    n, m = A.shape
    x = randomUnitVector(m)
    lastV = None
    V = x

    if n > m:
        B = np.dot(A.T, A)
    else:
        B = np.dot(A, A.T)

    iterations = 0
    while True:
        iterations += 1
        lastV = V
        V = np.dot(B, lastV)
        V = currentV / norm(currentV)
        if abs(np.dot(V, lastV)) > 1 - (1e-10):
            print("converged in {} iterations!".format(iterations))
            return V
```

FIGURE 4.5 – SVD function

2.1.2 Calculate the Euclidean distance of features

In this step we obtain one element within a new feature ; thus, the amount of computations is greatly reduced.

```
def euclidean0_1(vector1, vector2):
    '''calculate the euclidean distance, no numpy
    input: numpy.arrays or lists
    return: euclidean distance
    '''
    dist = [(a - b)**2 for a, b in zip(vector1, vector2)]
    dist =sqrt(sum(dist))
    return dist
```

FIGURE 4.6 – Euclidean distance between two vectors

```
def Euclidean_distance_of_features():
    liste_distance=[]#empty list in ordre to put distances
    L=list_frames()
    f=L[0]
    path2="C:\Users\HP\Downloads\Frames"
    lien_ref=path2+'\\'+f
    ref_image=Image.open(lien_ref)
    A=pixel_matrix(ref_image)
    B=SVD(A)
    l=get_diagonal_of_matrix(B)
    l=convert_array2float(l)
    liste_distance.append(euclidean0_1(l,l))
    for image in range(1,len(L)):
        lien=path2+'\\'+L[image]
        image=Image.open(lien)
        K=pixel_matrix(image)
        V=SVD(K)
        N=get_diagonal_of_matrix(V)
        N=convert_array2float(N)
        liste_distance.append(euclidean0_1(N,l))
    return liste_distance
```

FIGURE 4.7 – Euclidean distance between the SVD of two frames

2.1.3 Calculate the similarities

From above steps we know that only one element exists within each new feature, and we cannot calculate the similarity between two real numbers

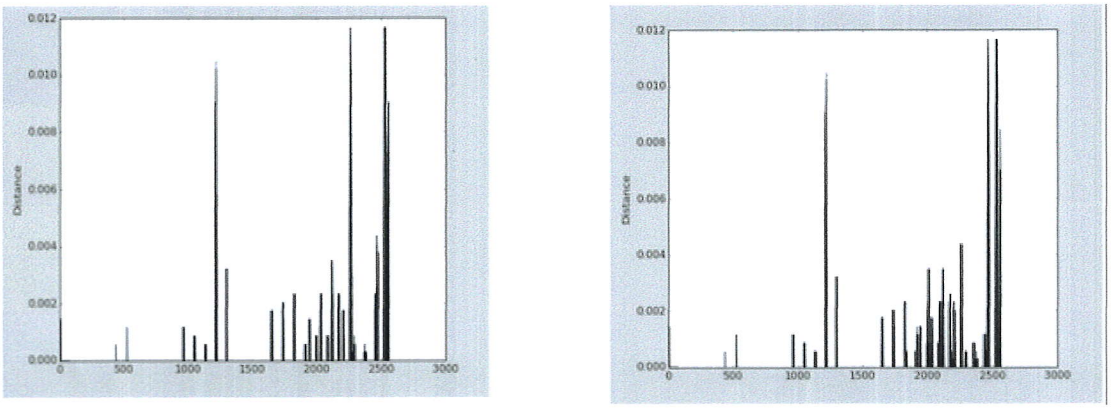


FIGURE 4.8 – Illustrations of the feature distances for both the original and forged videos : (a) feature distances for original video 1 ; (b) feature distances for the forged version of (a) by frame duplication ;

in Fig 4.8 we illustrate the feature distances for both the original and forged video ,We observe from this figure that the feature distances of the original video are different from the tampered one. the similarity in Fig 4.8 (a) is strong. Thus, the video is considered to have been tampered. Indeed, because of the length and the complexity of videos, it is difficult to find evidence of tampering in most cases of frame duplication based on feature distances.

To calculate the similarities in the video and identify the sub-sequences with high similar- ities, the feature distances for all overlapping (by one frame) sub-sequences are first obtained
It is assumed that the length of the sub-sequence is n and a video sequence of length L can be divided into m subsequences Assume that gi and gj are the respective feature distances of the ith sub-sequence and jth subsequence and calculate their similarity according to formula 3.5. We define the feature distance correlation matrix Corr as an $m * m$ symmetric matrix whose (i, j)th entry is the correlation coefficient between the ithand the jth sub-sequence, that is the similarity between the ith and the jth sub-sequence. The value of Corr (i, j) is within [1,1].

```
def Cor():
    dis= distances_into_sebsequence()
    vid=video_into_sebsequence()
    m=len(vid)
    Corr = np.zeros((m, m))

    for i in range(m):
        for j in range(i+1,m):
            Corr[i][j]=Corr[j][i]= CC(dis[i],dis[j])
    return Corr
```

FIGURE 4.9 – correlation coefficient

2.2 double-checking

As we said in in the previous chaptre the objective is to confirm and locate the duplicates using a method that can be subdivided into three stages : (a) merge candidate sub-sequence, (b) confirm duplicated sequences, and (c) perform localization.

2.2.1 Merge candidate sub-sequences

supposen that ($sqt1$, $sqr1$) and ($sqt2$, $sqr2$) are two pairs of candidate sub-sequences, and we need to merge the adjacent candidate sub-sequences to form a complete candidate sub-sequence, i.e., (sqt , sqr), as shown in Fig 3.13.

```
def new_seb():
    sub=mergeing()
    liste_original=[]
    liste_refrence=[]
    merged_original=[]
    merged_duplicated=[]
    index_original=[]
    index_duplicated=[]
    for item in sub:
        liste_original.append(item[0])
        liste_refrence.append(item[1])
    liste_original= sorted(liste_original)

    #liste_refrence= sorted(liste_refrence)
    if not liste_original and not liste_refrence:
        print "No duplication detected"
    else:
        L=video_into_sebsequence()
        merged_original=continuous_numbers(liste_original)
        merged_duplicated=continuous_numbers(liste_refrence)
        return merged_original
```

FIGURE 4.10 – merging adjacent sub-sequences

2.2.2 Confirm duplicated sequences

After merging the candidate sub-sequences, the goal is to confirm whether they are duplicated sequences. Therefore, a method of random block matching is designed for double checking (see chap 3 page 55).

```
def double_check():
    l,l1=after_merging()

    S=video_into_sebsequence()
    GP=[]
    GQ=[]
    C=[]
    D=[]
    R=[]
    C=[]
    folder="C:\Users\HP\Downloads\Frames"
    for s in range(1,len(l)):
        for i in range(M):
            lienp=folder+'\\'+str(l[s]+i)+'.jpg'
            lienq=folder+'\\'+str(l1[s]+i)+'.jpg'
            Gi=crop(lienp)
            G=crop(lienq)
            for k in range(500):
                GP.append(mu(Gi[k]))
                GQ.append(mu(G[k]))
            for j in range(len(GP)):
                if math.isnan(CC(GP[j],GQ[j]))==False:
                    if CC(GP[j],GQ[j])==1:
                        R.append(l[s]+i)
                        D.append(l1[s]+i)
```

FIGURE 4.11 – Double cheking duplicated frame

3 Experimental results

video	resolution	localisation of tapmpering
video1	320*240	between 6 to 25 and 99 to 119 between 105 to 124 and 252 to 274
video 2	320*240	Authentic
video 3	320*240	between 27 to 46 and 86 to 106 Between 52 to 72 and 146 to 167 Between 113 to 134 and 291 to 310 Between 148to 167 and 295 to 314
video 4	320*240	between 27 to 47 and 125 to 149 between 82 to 104 and 133 to 160 between 125 to 165 and 154 to 178 between 185 to 206 and 202 to 225 between 194 to 220 and 213 to 239 between 214 to 239 and 228 to 248 betweeen 281 to 300 and 277 to 297 between 353 to 382 and 286 to 308 between 372 to 397 and 345 to 379
video 5	320*240	between 176 to 195 and 204 to 224 between 201 to 221 and 225 to 244

we select from the dataset SULFA 5 videos , all of resolution of $320 * 240$ and compressed with h264 lossless compresion ,we found that the method is able to distinguish between original videos and tampered one and also localisate tapering but the precision rate is high which means that results gives false position or false alarms

4 conclusion

In this chapter we have implement an algorithm to detect video frame duplication the main advantage of this technique is that the response time is acceptable but precision rate (PR) is so high that sometimes the algorithm gives false result.

Conclusion

The availability of video editing makes videos very easily to tampered, in active video the digital signature and digital watermark is already exists so it is enough to match then to know which frame is genuine and which is tampered frame. But generally internet video do not contain digital signature or digital watermark therefore tampering detection in this video is very tough. There are various tampering detection techniques based on pixel, format, camera, physics and geometry to detect spatial tampering in passive and blind video.

Till now there is no universal method that can detect all types of tampering so video tampering detection still a challenge and suffers from various issues like:

New Tampering Attack Issue We have discussed in the first chapter different type of video tampering attack but it not means we have only such type of tampering attack, New tampering attack can arise at any time and tampering detection of such attack is very difficult

Hardware Issue Due to the continuous recording in surveillance video it requires huge memory and a lot of transmission power. If we try to make recording event wise then it will be tough to detect tampering in surveillance video because we do not have video frame sequence in continuous time space.

Technical Issue If we will not have digital signature and digital watermark in the video then proving authenticity of video or tampering detection will be very tough.

Tools Issue video and image editing tool are so advance that they visually left no clue for video . These tools are so advance that they create same size frame/ image .Tampering deletion in such frame / image required analysis till pixel level

Data Set Issue Video tampering is not as easy as image tampering, so create and maintain a data set of tampered video is very difficult.

Bibliography

[1] Lian, Shiguo, and Yan Zhang. "Multimedia forensics for detecting forgeries." *Handbook of Information and Communication Security* (2010): 809-828.

[2] Mehan, Julie. *CyberWar, CyberTerror, CyberCrime and CyberActivism: An i-depth guide to the role of standards in the cybersecurity environment.* IT Governance Publishing, 2014.

[3] Rocha, Anderson, et al. "Vision of the unseen: Current trends and challenges in digital image and video forensics." *ACM Computing Surveys (CSUR)* 43.4 (2011): 26.

[4] Kaur, Gurjinder, and Rishamjot Kaur. "A Survey study on video forgery detection using scale invariant feature detection Technique and DWT."

[5] Sowmya, K. N., and H. R. Chennamma. "A survey on video forgery detection." *International Journal of Computer Engineering and Applications* 9.2 (2015): 17-27.

[6] Chen, Li, and F. W. M. Stentiford. "Video sequence matching based on temporal ordinal measurement." *Pattern Recognition Letters*29.13 (2008): 1824-1831.

[7] Hartung, Frank, and Bernd Girod. "Watermarking of uncompressed and compressed video." *Signal processing*66.3 (1998): 283-301. des reference

[8] Abdullah, M. F. L., et al. "RECENT METHODS AND TECHNIQUES IN VIDEO WATERMARKING AND THEIR APPLICABILITY TO THE NEXT GENERATION VIDEO CODEC." *Journal of Theoretical Applied Information Technology* 74.1 (2015).

[9] Hussein, Jamal, and Aree Mohammed. "Robust video watermarking using multi-band wavelet transform." *arXiv preprint arXiv*:0912.1826 (2009)

[10] Abdullah, M. F. L., et al. "RECENT METHODS AND TECHNIQUES IN VIDEO WATERMARKING AND THEIR APPLICABILITY TO THE NEXT GENERATION VIDEO CODEC." *Journal of Theoretical Applied Information Technology* 74.1 (2015).

[11] Al-Taweel, Sadik AM, et al. "Digital video watermarking in the discrete cosine transform domain." *Journal of Computer Science* 5.8 (2009): 536.

[12] Liao, Sheng-Yang, and Tian-Qiang Huang. "Video copy-move forgery detection and localization based on Tamura texture features." *Image and Signal Processing (CISP), 2013 6th International Congress on.* Vol. 2. IEEE, 2013.

[13] ChandPandey, Ramesh, Sanjay Kumar Singh, and K. K. Shukla. "Digital Video Tampering Detection Techniques." *Encyclopedia of Information Science and Technology, Third Edition.* IGI Global, 2015. 1315-1325.

[14] Fard, Amin Milani, and F. Varasteh-A. "A new genetic algorithm approach for secure JPEG steganography." *Engineering of Intelligent Systems, 2006 IEEE International Conference on.* IEEE, 2006.

[15] Raja, K. B., et al. "A secure image steganography using LSB, DCT and compression techniques on raw images." *Intelligent Sensing and Information Processing, 2005. ICISIP 2005. Third International Conference on.* IEEE, 2005.

[16] Cheddad, Abbas, et al. "Digital image steganography: Survey and analysis of current methods." *Signal processing* 90.3 (2010): 727-752.

[17] Ramaswamy, Nandakishore, and K. R. Rao. "Video authentication for H. 264/AVC using digital signature standard and secure hash algorithm." *Proceedings of the 2006 international workshop on Network and operating systems support for digital audio and video.* ACM, 2006.

[18] Atrey, Pradeep K., Wei-Qi Yan, and Mohan S. Kankanhalli. "A scalable signature scheme for video authentication." *Multimedia Tools and Applications* 34.1 (2007): 107-135.

[19] Lin, Ching-Yung. *Watermarking and digital signature techniques for multimedia authentication and copyright protection.* Diss. Columbia University, 2001.

[20] Tang, Yuan-Liang, and Chih-Jung Hung. "Recoverable Authentication of Wavelet-Transformed Images." *ICGST International Journal on Graphics, Vision and Image Processing* 11 (2005): 61-66.

[21] Tzeng, Chih-Hsuan, and Wen-Hsiang Tsai. "A new technique for authentication of image/video for multimedia applications." *Proceedings of the 2001 workshop on Multimedia and security: new challenges.* ACM, 2001

[22] Wang, Weihong, and Hany Farid. "Exposing digital forgeries in video by detecting double MPEG compression." *Proceedings of the 8th workshop on Multimedia and security.* ACM, 2006.

[23] Abdullah, M. F. L., et al. "RECENT METHODS AND TECHNIQUES IN VIDEO WATERMARKING AND THEIR APPLICABILITY TO THE NEXT GENERATION VIDEO CODEC." *Journal of Theoretical Applied Information Technology*74.1 (2015).

[24] Sang, Jun, and Mohammad S. Alam. "Fragility and robustness of binary-phase-only-filter-based fragile/semifragile digital image watermarking." *IEEE Transactions on Instrumentation and Measurement*57.3 (2008): 595-606.

[25] Garg, Ravi, Avinash L. Varna, and Min Wu. "Seeing ENF: natural time stamp for digital video via optical sensing and signal processing." *In Proceedings of the 19th ACM international conference on Multimedia*, pp. 23-32. ACM, 2011.

[26] Johnson, Micah K., and Hany Farid. "Exposing digital forgeries in complex lighting environments." *IEEE Transactions on Information Forensics and Security*2.3 (2007): 450-461.

[27] Zheng, Lu, Tanfeng Sun, and Yun-Qing Shi. "Inter-frame video forgery detection based on block-wise brightness variance descriptor." *International Workshop on Digital Watermarking*. Springer International Publishing, 2014.

[28] Shanableh, Tamer. "Detection of frame deletion for digital video forensics." *Digital Investigation* 10.4 (2013): 350-360.

[29] Stamm, Matthew C., W. Sabrina Lin, and KJ Ray Liu. "Temporal forensics and anti-forensics for motion compensated video." *IEEE Transactions on Information Forensics and Security* 7.4 (2012): 1315-1329.

[30] Feng, Chunhui, et al. "Automatic location of frame deletion point for digital video forensics." *Proceedings of the 2nd ACM workshop on Information hiding and multimedia security*. ACM, 2014.

[31] Liu, Hongmei, Songtao Li, and Shan Bian. "Detecting frame deletion in H. 264 video." *International Conference on Information Security Practice and Experience*. Springer International Publishing, 2014.

[32] Luo, Weiqi, Min Wu, and Jiwu Huang. "MPEG recompression detection based on block artifacts." *Electronic Imaging 2008*. International Society for Optics and Photonics, 2008.

[33] Chen, Chunhua, Yun Q. Shi, and Wei Su. "A machine learning based scheme for double JPEG compression detection." *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*. IEEE, 2008.

[34] Jiang, Xinghao, et al. "Detection of double compression in MPEG-4 videos based on Markov statistics." *IEEE Signal processing letters* 20.5 (2013): 447-450.

[35] Su, Yuting, and Junyu Xu. "Detection of double-compression in MPEG-2 videos." *Intelligent Systems and Applications (ISA), 2010 2nd International Workshop on*. IEEE, 2010.

[36] Wang, Weihong, and Hany Farid. "Exposing digital forgeries in video by detecting double quantization." *Proceedings of the 11th ACM workshop on Multimedia and security*. ACM, 2009.

[38] Sun, Tanfeng, Wan Wang, and Xinghao Jiang. "Exposing video forgeries by detecting MPEG double compression." *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on.* IEEE, 2012.

[39] Milani, Simone, Marco Tagliasacchi, and Stefano Tubaro. "Discriminating multiple JPEG compressions using first digit features." *APSIPA Transactions on Signal and Information Processing* 3 (2014).

[40] He, Peisong, et al. "Double compression detection in MPEG-4 videos based on block artifact measurement with variation of prediction footprint." *International Conference on Intelligent Computing.* Springer International Publishing, 2015.

[41] Vazquez-Padin, David, et al. "Detection of video double encoding with GOP size estimation." *Information Forensics and Security (WIFS), 2012 IEEE International Workshop on.* IEEE, 2012.

[42] Qadir, Ghulam, Syamsul Yahaya, and Anthony TS Ho. "Surrey university library for forensic analysis (SULFA) of video content." (2012): 121-121.

[43] Dalal, Navneet, and Bill Triggs. "Histograms of oriented gradients for human detection." *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on.* Vol. 1. IEEE, 2005.

[44] Ustubioglu, Beste, et al. "A new copy move forgery detection technique with automatic threshold determination." *AEU-International Journal of Electronics and Communications* 70.8 (2016): 1076-1087.

[45] Tamura, Hideyuki, Shunji Mori, and Takashi Yamawaki. "Textural features corresponding to visual perception." *IEEE Transactions on Systems, Man, and Cybernetics* 8.6 (1978): 460-473.

[46] Wang, Weihong, and Hany Farid. "Exposing digital forgeries in video by detecting duplication." *Proceedings of the 9th workshop on Multimedia security.* ACM, 2007.

[47] Hsu, Chih-Chung, et al. "Video forgery detection using correlation of noise residue." *Multimedia Signal Processing, 2008 IEEE 10th Workshop on.* IEEE, 2008.

[48] Li, Fugui, and Tianqiang Huang. "Video copy-move forgery detection and localization based on structural similarity." *Proceedings of the 3rd International Conference on Multimedia Technology (ICMT 2013).* Springer Berlin Heidelberg, 2014.

[49] Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 13(4):600–612

[50] P. Bestagini, S. Milani, M. Tagliasacchi, S. Tubaro, "Local tampering detection in video sequences", 2013 IEEE 15th International Workshop on Multimedia Signal Processing (MMSP), Pula, Italy, 2013

[51] http://sulfa.cs.surrey.ac.uk/videos.php