Video Authentication: Issues and Challenges

Saurabh Upadhyay[*], Sanjay Kumar Singh[†]

[*]Department of Computer Science & Engineering, SIT, Gujarat-India

[†]Department of Computer Engineering, IT-BHU, Varanasi-India

Abstract

Video authentication aims to ensure the trustworthiness of the video by verifying the integrity and source of video data. It has gained much attention in the recent years. In this paper we present the issues in the designing of a video authentication system. These issues include the classification of tampering attacks, levels of tampering attack and robustness. Further we present the categorization of existing video authentication techniques with their shortcomings. Moreover we have also given the challenging scenarios in which the video authentication would be a critical task.

Keywords: *Video Authentication, Fragile Watermarking, Digital Signature, Intelligent Techniques, Tampering Attacks*

1. Introduction

With the rapid innovation and development in digital technologies, video applications are infiltrating into our daily lives in breakneck speed from traditional television broadcasting to modern vulnerable communication media such as Internet/Intranet, wireless communication and consumer products such as VCD/DVD. In some applications the authenticity of video data is of paramount interest such as in video surveillance, forensic investigations, law enforcement and content ownership [33]. For example, in court of law, it is important to establish the trustworthiness of any video that is used as evidence. As in another scenario, for example, suppose a stationary video recorder for surveillance purpose, is positioned on the pillar of a railway platform to survey every activity on that platform along a side, it would be fairly simple to remove a certain activity, people or even an event by simply removing a handful of frames from this type of video sequences. On the other hand it would also be feasible to insert, into this video, certain objects and people, taken from different cameras and in different time. So video authentication is a process which ascertains that the content in a given video is authentic and exactly same as when captured.

1.1 Motivation behind Video Authentication

A video clip can be doctored in a specific way to defame an individual. In the recent years, several cases have been reported where the eminent personalities of the society were caught in illegal activities in the video recordings made by so called journalists. However in the absence of foolproof techniques to authenticate the video it is difficult to trust on such reports. On the other hand criminals get free from being punished because the video (used as evidence), showing their crime cannot be proved conclusively in the court of law. In the case of surveillance systems, it is difficult to assure that the digital video produced as evidence, is the same as it was actually shot by camera. In another scenario, a news maker cannot prove that the video played by a news channel is trustworthy; while a video viewer who receives the video through a communication channel cannot ensure that video being viewed is really the one that was transmitted [6]. In the scenario of sensitive cases where a video is produced as a witness in the court of law, even a small modification may not be acceptable. However there are some scenarios where editing also may be allowed while keeping intact the authenticity of the video. For example after shooting the video, a journalist may need to perform some editing before broadcasting it on a news channel. In such a case a video authentication system should be able to allow editing on the video up to a certain level ensuring the authenticity of the video [38]. These are the instances where malicious modifications cannot be tolerated. Therefore there is a compelling need for video authentication

Although traditional data authentication technology for message integrity was mature, video authentication is still in its early development stage and many fundamental questions remain open [28]. For example, for a number of different authentication algorithm developed over the past few years, it is difficult to affirm which approach seems most suitable for ensuring the integrity adapted to videos [28]. There is a need for synthesizing literature to understand the nature of the problem, identify the potential for research issues, standardize new research area and evaluate the relative performances of different approaches. The aim of this paper is to examine the status and issues of video authentication techniques and to assess their strengths and weakness in the reference of different tampering attacks. This paper is organized as follows. In Section 1 the notion of video authentication and framework are briefly introduced followed by a discussion of motivation behind video authentication. Section 2 explains the issue of robustness for video authentication. Security related issues are discussed in detail in section 3. Section 4 provides a concise review of existing techniques for video authentication. Some of the new challenging scenarios are

briefly introduced in section 5. Finally the summary and future research directions are discussed

2. Robustness

Any video applications may have at least three parties: Producer, receiver and the third party. The producer generates the video and the receiver receives the video from producer via third party. Here the third party is a general and wide concept. It could be either a storage device in consumer products (such as CD/DVD) or a busy and noisy channel in video transmission. Further a receiver can also be a third party if, after receiving the video, it forwards the video to any other party. The malicious attacker targets this third party category for altering the video content. Video authentication is the process which ascertains that the content in a given video is authentic and exactly same as when captured. Lin and Chang [8] classified the multimedia authentication techniques into two categories: Complete authentication and content authentication. The techniques which are proposed for complete authentication consider that the multimedia data, which have to be authenticated, have to be exactly the same as the original one. No change in the multimedia data is allowed. In content authentication, as long as the meaning of multimedia data remains unchanged, the received multimedia data is considered as authentic, regardless of the processing or transformation the multimedia data has undergone. Of course, video authentication should be content authentication because a receiver must not obtain an exact copy of the original video without any distortion, necessarily. For instance, due to its bigger size in storage, digital videos are usually compressed and most video compression, such as MPEG 1/2/4 are lossy compression. And definitely the de-compressed video is not identical to the original one. However, it should still be considered to be authentic. Another example is video transcoding in which the bit rate of a video stream is adjusted to adapt to variable transmission channel.

Thus a video authentication system theoretically should be robust enough to discriminate all normal video processing operations from malicious tampering attack. A robust video authentication system should tolerate the incidental distortion, which may be introduced by normal video processing such as compression, resolution conversion and geometric transformation, while being capable of detecting the intentional distortion, which may be introduced by malicious attack However, it is a difficult task to define all acceptable video processing operations due to the huge diversity of video applications. For example, the object based video processing operations such as rotation, scale and translation (RST) is very different from the traditional frame-based video processing operations. The video authentication system should also be sensitive to malicious manipulations.

3. Security issues of video authentication.

A continuous video sequence $V_c(x, y, t)$ is a scalar real valued function of two spatial dimensions x and y and time t, usually observed in a rectangular spatial window W over some time interval T. If $B(x, y, t)$ is modification vector then the tampered video$M_c(x, y, t)$ would also be a scalar real valued function of spatial dimensions x and y and time t as follows:

$$M_c(x, y, t) = B(x, y, t) + V_c(x, y, t)$$

When the content of information, being produced by a given video sequence is maliciously altered, then it is called tampering of video data. It can be done for several purposes, for instance to manipulate the integrity of an individual. Since a wide range of sophisticated and low cost video editing software are available in the market that makes it easy to manipulate the video content information maliciously, it projects serious challenges to researchers to be solved.

3.1. Video Tampering Attacks

There are several possible attacks that can be applied to alter the contents of a video data. Formally a wide range of authentication techniques have been proposed in the literature but most of them have been primarily focused on still images. In several applications, due to large availability of information in video sequences, it may be more significant if the authentication system can tell where the modifications happened (It indicates the locality property of authentication) and how the video is tampered [5]. On considering these where and how, the video tampering attacks can have different classifications. A lot of works have been done that briefly address the classification based on where [33], [5]. And some papers address the classification based on how [34]. In general, finding where the multimedia data is altered is more efficient than to find out how the multimedia data is tampered. When a video is being recorded by a video recording device, it captures the scene which is in front of the camera lens, frame by frame, with respect to time. Number of frames being captured by video recording device in a second depends on the hardware specification of the device. Thus a video sequence can be viewed as a collection of consecutive frames with temporal dependency, in a three dimensional plane. This is called the regional property of the video sequences. When a malicious alteration is performed on a video sequence, it either attacks on the contents of the video (i.e. visual information presented by the frames of the video), or attacks on the temporal dependency between the frames. Therefore based on the regional property of the video sequences, we can broadly classify the video tampering attacks into three categories: spatial tampering attacks, temporal tampering attacks and the combination of these two, spatio-temporal tampering attacks [5]. They can be further classified into their subcategories.

3.1.1. Spatial Tampering

In spatial tampering malicious alterations are performed on the content of the frames (X-Y axis). The operations that can be done as tampering attack in spatial tampering are cropping and replacement, morphing, content (object)

adding and removing etc [5]. These attacks can be efficiently performed with the help of video editing software as *Photoshop*, etc.

3.1.2. Temporal Tampering

In temporal tampering manipulation is performed on the sequence of frames. The focus is on the temporal dependency. Temporal tampering attacks are mainly affecting the time sequence of visual information, captured by video recording devices. The common attacks in temporal tampering are frame addition, frame removal and frame reordering or shuffling.

3.1.3. Spatio-Temporal Tampering

Spatio-temporal tampering attacks are the combination of the both kinds of tampering attacks. Frame sequences are altered as well as visual contents of the frames are modified in the same video. The authentication system should be able to identify both kinds of tampering.

All these tampering are further classified into their subcategories. Spatial tampering can be in effect either at block level or at pixel level. In both the cases the objects of the frames of the video are altered. Further the objects of the frames are classified into two categories: Foreground objects and Background objects. The foreground objects are those which are captured as individual elements, excluding the background, in a frame. And the background object is the background part of the frame excluding all of the foreground objects. The different pieces of visual information shown in the frames of the video are altered in spatial tampering. Basically the contents of the video frames are treated as objects. Based on these objects and their classification the spatial tampering can be further classified as following figure shows.

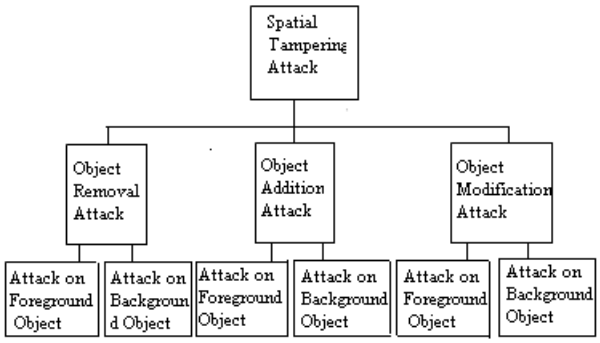


Fig.1. Spatial Tampering Classification

Fig. 1 shows an explicit classification of spatial tampering attacks in reference of objects of the frames.

3.1.1.1. Object Removal Attack

In object removal attack, the objects of the frames of the video are eliminated. This kind of attack is commonly performed where a particular person wants to hide his/her presence in a certain sequence of frames. With this kind of attack he /she may disappear in a specific time domain, recorded in the video. This attack can be performed with both kinds of object, foreground objects and background object, as shown in Fig. 2

3.1.1.2. Object Addition Attack

When an object is inserted in a frame or in a set of frames then there is a kind of spatial tampering attack: say Object addition attack. In any video sequence which can be treated as evidence, an additional object can be pasted in a frame or set of frames, with the help of sophisticated video editing software to mislead the investigation agencies as well as court of law. As shown in fig. 3, it can also be performed with both kinds of objects, foreground objects and background objects.

Original Frame Tampered Frame
Fig. 2 (a).

Original Frame Tampered Frame
Fig. 2 (b).

Fig.2. Example of object removal attack. Fig.2 (a) shows object removal attack with foreground object, where a small device is removed from the original frame in tampered frame. Whereas Fig.2 (b) shows the object removal attack with background object. Here a small object on the right side of the wall is eliminated from the original frame in tampered frame.

3.1.1.3. Object Modification Attack

In Object modification attack, an existing object of the frame(s) can be modified in such a way that the original identity of that object is lost, and a new object may be in appearance which is totally different from the original object. The object modification attacks can be existed in many prospects in the given video. For instance, the size and shape of the existing object may be changed, the colour of the object may be changed or it may be discoloured, and with the help of additional effect the nature of the object and it's relation with other objects also may be changed. In fact it is very hard to detect this kind of attack for authentication systems, since these attacks are performed at pixel level. The authentication systems should be robust enough to differentiate this kind of attack with the normal video processing operations. Fig.4 shows a typical example of object modification attack where the face of a person has been changed in such a way that a new person's face is introduced in the altered frame. These attacks can also be performed with both kinds of objects, foreground and background objects.

Besides spatial tampering, temporal tampering attacks have also sub classifications. Temporal tampering attacks can be performed at scene level, shot level and frame level, but the

primary focus is on attacking the temporal dependency of the frames of the video.


Original Frame


Tampered Frame

Fig.3. (a)


Original Frame


Tampered frame

Fig.3. (b)

Fig.3. Example of Object Addition attack. In original frame of Fig.3 (a) two persons are there as major foreground objects, while in tampered frame of Fig.3 (a) an additional person as a foreground object is added. In tampered frame of the Fig.3 (b), not only a foreground object is added but also an additional wall as a background object, in the middle of the frame, is added.


Original Frame


Tampered Frame

Fig.4. Example of Object modification attack. The face of the person in original frame is modified in tampered frame, in such a way that the new face of the person cannot be identified as the same as in original frame.

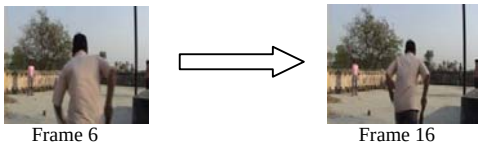
We call it 'Third dimensional (dimension with respect to time) attack' on the video sequences. Therefore based on this third dimensional attack we can classify the temporal tampering attacks into following categories.

3.1.2.1. Frame Addition Attack

In frame addition attack, additional frames from another video, which has the same statistical properties, are intentionally inserted at some random locations in a given video. This attack is intended to camouflage the actual content and provide incorrect information [33]. A typical example of the frame addition attack is shown in fig. 5.

3.1.2.2. Frame Removal Attack

In frame removal attack the frames of the given video are intentionally eliminated. In this kind of attack frames or set of frames can be removed from a specific location to a fixed location or can be removed from different locations. It depends upon the intention. Commonly this kind of tampering attack is performed on surveillance video where an intruder wants to remove his/her presence at all. Fig. 6 shows a typical example of frame removal attack


Frame 6 Frame 16


Frame 6 Added Frame Frame 17

Fig.5 Example of Frame addition attack. In first row the original frame sequence from frame 6 to frame 16 has been shown. After attack, the second row of the frames shows the altered frame sequence in which a new frame is inserted between frame 6 and frame 16. And frame 16 becomes frame 17.


Frame 14 Frame 22 Frame 30

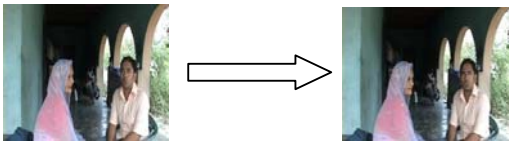

Frame14 Frame 29

Fig.6 Example of Frame removal attack. The first row of this figure shows the original frame sequence with frame 14, frame 22 and frame 30. In second row of the frame sequence, which shows the tampered frame sequence with frame removal attack, frame 22 is eliminated from the video and hence frame 30 becomes frame 29.

3.1.2.3. Frame Shuffling Attack

In frame shuffling attack, frames of a given video are shuffled or reordered in such a way that the correct frame sequence is intermingled and wrong information is produced by the video as compared to original recorded video. Fig. 7 shows a typical example of frame shuffling attack where two frames are shuffled.


Frame 13 Frame 20 Frame26


Frame13 Frame 20 Frame26

Fig.7. Example of Frame shuffling attack. The first row of this figure shows the original frame sequence with frame 13, frame 20 and frame 26. After the frame shuffling attack, the original frame sequence is tampered as shown in second row of the figure where the positions of frame 13 and frame 26 have been changed.

3.2. Levels of Tampering Attacks

In addition of these types of tampering attacks, tampering can be done at different levels in video sequences.

3.2.1. Scene Level

When the tampering attacks are performed at scene level then a whole scene of the video sequence is manipulated in such a way that, not even the scene itself is modified but also in the reference of the given video, the scene of that video is modified. It means spatial and temporal both kinds of tampering can be done at scene level.

3.2.2. Shot Level

In shot level tampering any particular shot of the given video is modified in reference to the given video. In shot level tampering a shot can be added or removed from the video. It can also be performed with all kinds of tampering attacks.

3.2.3. Frame Level

When frames of the given video are maliciously modified, then it is called tampering at frame level. Frame removal, frame insertion and frame shuffling are the common tampering attacks that can be performed at frame level. In other words, temporal tampering attacks are commonly performed at frame level.

3.2.4. Block level

In block level tampering, tampering attacks are performed on the blocks of the video frames. The content of the video frames are treated as blocks on which the tampering attacks are applied. Blocks (a specified area on the frame of the video) can be cropped and replaced, morphed or modified in any way in block level tampering. Spatial tampering attacks are commonly performed at block level.

3.2.5. Pixel level

In pixel level tampering contents of the video frames are modified at pixel level. This is the smallest level in video sequences at which tampering attacks can be performed. The video authentication system should be robust enough to differentiate the normal video processing operation and pixel level tampering, since many normal video processing operations are performed at pixel level. Spatial tampering attacks are commonly performed at pixel level. All these levels of tampering show the different aspects of tampering.

4. State of the art review

By definition, authenticity means sometimes “as being in accordance with fact, as being true in substance”, or “as being what it professes in origin or authorship, as being genuine” [30]. Another definition of authentication is to prove that something is “actually coming from the alleged source or origin” [31]. Video authentication, in general has received considerable attention by academia and practitioners over the last few years.

A typical video authentication system is shown in fig. 8. For a given video, authentication process starts with feature extraction. After that, with a specific video authentication algorithm, the authentication data H is generated using the features f of the video. This authentication data H is encrypted and packaged with the video as a signature or alternatively it can be embedded into the video content as a watermark. The video integrity is verified by computing new authentication data H` for the given video. The new authentication data H` is compared with decrypted original authentication data H. If both are matched, the video is treated as authentic else it is constructed to be tampered.

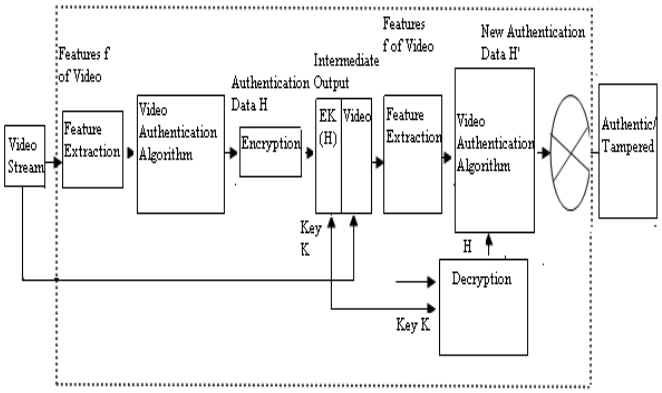


Fig. 8 A typical video authentication system.

4.1. Classification of Authentication Techniques

In past few years, watermarking and digital signatures have been widely used for the purpose of video authentication. Different techniques have their own advantages and shortcomings. Fig. 9 represents the tree structure of techniques which have been commonly proposed for the purpose of video/image authentication. In fact fragile watermarking and digital signatures are the two basic schemes for authentication [5]. Moreover there has also been worked on intelligent techniques for video authentication. Apart from these digital signature, fragile watermarking and intelligent techniques, some other authentication techniques are also introduced by researchers. We are giving here a brief classification of video authentication techniques.

4.1.1. Digital Signature

Integrity of multimedia data can be greatly verified by digital signature. For the authentication of multimedia data, it was first introduced by Diffie and Hellman in 1976[26]. For the purpose of authentication, digital signatures can be saved in two different ways. Either they can be saved in the header of the compressed source data, or it can be saved as an independent file. Further they can be produced for verification. In the prospective of robustness, since the digital signature remains unchanged when the pixel values of the images/videos are changed, they provide better

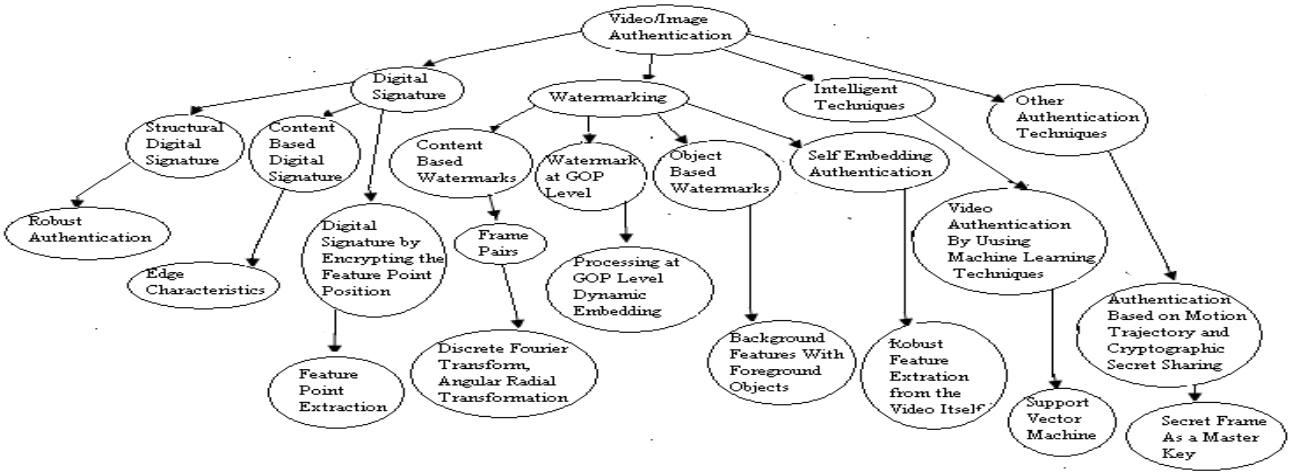


Fig. 9 Tree Structure of Authentication Technique

results. In the digital signature authentication, the digital signature of the signer to the data depends on the content of data on some secret information which is only known to signer [27]. Hence, the digital signature cannot be forged, and the end user can verify a received multimedia data by examining whether the contents of data match the information conveyed in the digital signature. Ching-Yung et al [8] proposed a scheme in which two types of robust digital signatures are used for video authentication in different kinds of situations. The first type of authentication signature is used in situation where the GOP (Group of Pictures) structure of the video is not modified, after transcoding or editing processes. The situation, where the GOP structure is modified and only the pixel values of picture will be preserved; a second type of digital signature is used. In another work, video authentication is done by generating digital signatures for image blocks and using them as watermarks [3]. In this approach localization packet, watermark insertion is done via LSB modification of pixel values. As compared to [2] where video tampering is identified through an analysis of watermark sequencing, here (explicit) block ID's are used for this purpose. The Johns Hopkins University Applied Physics Laboratory (APL) has developed a system for authentication of digital video [7]. The authentication system computes secure computer generated digital signatures for information recorded by a standard digital video camcorder. While recording, compressed digital video is simultaneously written to digital tape in the camcorder and broadcast from the camera into the Digital video authenticator. In this authentication system, video is separated into individual frames and three unique digital signatures are generated per frame-one each for video, audio and (camcorder) control data- at the camcorder frame rate. Here the key cryptography is used. One key, called a "private" key is used to generate the signatures and is destroyed when the recording is complete. The second a "public" key is used for verification. The signatures that are generated make it easy to recognize tampering. If a frame has been added it would not have a signature and will be instantly detected and if an original frame is tampered the signature would not match the new data and it will be detected in verification process. Ditmann [17] and Queluz [18] used the edge /corner of the image as the feature to generate the digital signature. They claimed this feature is robust against high quality compression and scaling but the problem is that the signature generated based on the edge is too long, and the consistency of the edge itself is also a problem. The, digital signature and watermarking, are able to detect regions that have been tampered, but often they are too fragile to resist incidental manipulations. For this type of incidental manipulations structural digital signature [23] can be used for image authentication. This approach makes use of an image's content to construct a structural digital signature (SDS) for image authentication. In this approach [23], many incidental manipulations which can be detected as malicious modifications in other digital signature verifications or fragile watermarking schemes, can be ignored. In the scenario of a station streaming video over network, it is significant for the audiences to have guarantees that the video stream they are watching is indeed from the station. Schemes that are used for this purpose can prevent the malicious parties from injecting commercials or offensive materials into the video streams. Actually this problem has been covered in information security called streaming signing [12] [13], which is an extension from message signing by digital signature schemes. A separate authentication code is written in [37] from the blocks of the video frames. Here the authors Po-Chyi Su et al use the approach of scalar/vector quantization on the reliable features. Once the authentication code is written, it is transmitted along with the video. Thus the authenticity of the given video content can be checked by matching the extracted feature with the transmitted authentication code. Navajit Saikia and Prabin K. Bora present a scheme for video authentication in [15] that generates the message authentication code (MAC) for a group of frames (GOF) using coefficients from the last but

one high pass band at full level of temporal wavelet decomposition. This digital signature based scheme uses temporal wavelet transform for the generation of message authentication code. After the extraction of $GOFs$ from the video, these $GOFs$ are recursively decomposed into high pass band up to a certain level using temporal wavelet transform. At this level the high pass band consists of two frames. In the signature generation process, these frames are divided into some blocks of fixed sizes. These blocks are randomly mapped on to a set of groups, using a mapping key in such a way that each group contains equal number of blocks. With the transform coefficients and these groups of blocks, a set of linear combination values is evaluated for each frame in the high pass band. And with these sets of linear combination values, message authentication code (MAC) is obtained for the GOF. In the signature verification process, the distances $d\ (\ MACi, 1, MAC`i, 1)$ and $d(MACi, 2, MAC`I, 2)$ are calculated where d is any distance measure and $\{\ MACi, 1, MACi, 2\ \}$ is the MAC of $i^{th}\ GOF$ of the original video and $\{MAC`i, 1, MAC`i, 2\}$ is the MAC of corresponding GOF calculated at receiver site. Here the GOF of the video would be authentic if these two distances are below some predefined threshold values, otherwise tampered. This authentication scheme would be advantageous for spatio-temporal manipulations, since it is effective for spatial tampering as well as for temporal tampering. Similar to Dittmann's [17] content based digital signature approach for image/ video authentication using edge characteristics, Bhattacharjee and Kutter [25] proposed a scheme to generate a digital signature by encrypting the feature points positions in an image. In this approach authentication is accomplished by comparing the positions of the feature point extracted from the targeted image with those decrypted from the previously encrypted digital signature.

4.1.2. Watermarking

Watermarking always remains a significant issue for solving authentication problems regarding digital multimedia data, in past few years. A wide variety of watermarking techniques have been proposed by various researchers in literature. Based on the application areas, watermarking can be classified in different categories [34]. Beside to ensure the integrity of the digital data and recognizing the malicious manipulations, watermarking can be used for the authentication of the author or producer of the content. Watermarks can be embedded with the multimedia data, without changing the meaning of the content of the data. The advantageous feature with the watermarks is that, they can be embedded without degrading the quality of multimedia data too much. Since the watermarks are embedded in the content of video data, once the data is manipulated, these watermarks will also be modified such that the authentication system can examine them to verify the integrity of data. In [4], authors describe the use of video authentication template, which uses a bubble random sampling approach applied for synchronization and content verification in the context of video watermarking. The authentication template is introduced in order to ensure temporal synchronization and to prevent content tampering in video sequences [4]. Basically in past few years, an increasing use of digital information in our society and availability of very sophisticated and low cost video editing software creates problems associated with copyright protection and authentication. The owners or producers of information resources are being worried of releasing proprietary information to an environment that appears to be lacking in security [9]. On the other hand with the help of powerful video editing software one can challenge the trustworthiness of digital information. In [9], M. P. Queluz presents the generic models with labelling and watermarking approaches for content authentication. In labelling based approach authentication data are written in separate file [9], while in watermarking based approach the authentication information is embedded in the frames. In this labelling-based authentication system, features C and C` are extracted from the original and modified pictures respectively as according:

$$C = f_c\ (\ I\ )\ ,\ \ C' = f_c(\ \hat{I}\ )$$ In order to assure the authenticity of the label content, it is signed in a trustworthy way, that is, the label is encrypted with a private key(K_{pr}). The label content is produced as:

$$L = EKpr\ (C, C_I)$$

Where C_I is optional information, say *Complementary Information*, about the frame and its author, assigned by an author society. In the authentication system the corresponding public key Kpu is used to decrypt the label, producing:

$$C, C_I = EK_{pu}(L)$$

Moreover in [9] M. P. Queluz presents two classical image features for image/video content authentication. The first image feature is concerned with second order image moments. The second feature relies on image edges and it takes the problem of image/video authentication from a semantic view [9]. In image moments feature, for a two dimensional continuous function *f(x, y)*, the moments of order $(p + q)$ is defined as

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^p\, y^q\, f(x,y)\, dx\ dy$$

$$For\ p, q = 0, 1, 2, \ldots\ldots\ldots\ldots$$

For a digital image the above equation would be as follows:

$$m_{pq} = \sum_i \sum_j i^p\, j^q\, f(i,j)$$

Where $f(i,j)$ represents image color values at pixel site (i, j). Moments are usually normalized dividing it by the image total mass, defined as$\sum_i \sum_j f(i,j)$. Chang-yin Liang, et al [16] proposed a video authentication system which is robust enough to separate the malicious attack from natural video processing operations with the cloud watermark. In the authentication system [16], first of all, the video sequence is split into shots and the feature vectors are extracted from each shot. Then the extracted feature is used to generate watermark cloud drops with a cloud generator

[16]. Here, for robustness, a content based and semi fragile watermark is used for authentication. In this authentication technique DCT coefficients are evaluated first by partially decoding the given video. After watermarking, the video is encoded again [16]. The extracted watermarks are compared with the features derived from the received video, to check the authenticity of the given video.

4.1.3. Intelligent Techniques

Intelligent techniques for video authentication use database of videos. The database comprises authentic video clips as well as tampered video clips. As in [33], the authors proposed an intelligent technique for video authentication which uses inherent video information for authentication, thus making it useful for real world applications. The proposed algorithm in [33] is validated using a database of 795 tampered and non tampered videos and the results of algorithm show a classification accuracy of 99.92%. The main advantage of intelligent techniques is that they do not require the computation and storage of secret key or embedding of watermark. The algorithm in [33] computes the local relative correlation information and classifies the video as tampered or non-tampered. Here the algorithm uses Support Vector Machine (SVM) for the classification of the tampered and authentic videos. SVM [10] is a powerful methodology for solving problems in non linear classification, function estimation and density estimation [11]. This algorithm [33] is performed in two stages: (1) SVM training and (2) Tamper detection and classification, usingSVM. In SVM training, the algorithm trains the SVM by using a manually labelled training video database, if the video in the training data is tampered, then it is assigned the label -1 otherwise the label is +1 (for the authentic video). From the training videos, relative correlation information between two adjacent frames of the video is computed, with the help of corner detection algorithm [14]. Then relative correlation information RC is computed for all adjacent frames of the video with the help of

$$RC = \frac{1}{m}\sum_{i=1}^{m} L_i$$

Where Li is local correlation between two frames $for\ i = 1,2,\dots m.$ and m is the number of corresponding corner points in the two frames. The local correlation information RC is computed for each video and the RC with the label information of all the training video data are provided as input to theSVM. With this information of all the video in video database, the SVM [10] is trained to classify the tampered and non tampered video data. Output of SVM training is a trained hyper plane with classified tampered and non tampered video data. In [24], authors integrate the learning based support vector machine classification (for tampered and non tampered video) with singular value decomposition watermarking. This algorithm is independent of the choice of watermark and does not require any key to store. This intelligent authentication technique embeds the inherent video information in frames using SVD watermarking and uses it for classification by projecting them into a non linear SVM hyper plane. This technique can detect multiple tampering attacks.

4.1.4. Other Authentication Techniques

Apart from digital signature, watermarking and intelligent techniques, various other techniques are proposed by researchers for authentication purpose of digital video in the literature. In [19], an authentication scheme for digital video is introduced which is based on motion trajectory and cryptographic secret sharing [19]. In this scheme, the given video is first segmented into shots then all the frames of the video shots are mapped to a trajectory in the feature space by which the key frames of the video shot are computed. Once the key frames are evaluated, a secret frame is computed from the key frames information of the video shot. These secret frames are used to construct a hierarchical structure and after that final master key is obtained. This master key is used to identify the authenticity of the video. Any modification in a shot or in the important content of a shot will be reflected as changes in the computed master key. Here trajectory is constructed, using the histogram energy of the frames of the video shot. Once the key frames are computed these are utilized to compute the secret frame by extrapolation. Now an interpolating polynomial $f(x)$ is computed by using key frames as follows.

$$\sum_{j=1}^{n+1}\prod_{i=1}^{n+1}\frac{x-x_i}{x_j-x_i}\ I_j$$

This is Lagrange interpolation formulation where the x_i position refers to each key frame and I_i is the pixel value of the key frames. By using this equation and extrapolation a frame at $x = 0$ is computed, which is regarded as the secret key. Considering the set of secret keys as another set of shares, the master key frame is computed for that particular video. With this scheme any video can be authenticated by comparing its computed master key with the original master key. This comparison can be performed by using the general cosine correlation measure given by:

$$sim = \frac{I_O\ .\ I_N}{|I_O\ .\ I_N|}$$

Where I_O and I_N are the original master key and the new master key cosidered as vectors. The similarity value would be in the range [0, 1] and if sim = 1, the two master keys would be the same, however if sim = 0, the two master keys would be different. In [20], the key frames are selected by deleting the most predictable frame. In the approach of [21], the key frames are extracted from a video shot based on the nearest feature line. The work in [22] authenticates a video by guaranteeing the edited video to be the subsequence of the original video using a special hash function. The MPEG video standard is one of the most popular video standards in today's digital era. In [35] Weihong Wang and Hany Farid have been worked on MPEG video standard(MPEG-1 and MPEG-2) in this paper they specifically show how a doubly compressed MPEG video sequence introduces specific static and temporal statistical perturbations whose presence can be used as evidence of tampering. In [36] Hany Farid describes three techniques to expose digital forgeries in

which the approach is to first understand how a specific form of tampering disturbs certain statistical properties of an image and then to develop a mathematical algorithm to detect this perturbation. These are Cloning, Lighting and Retouching. In Cloning, a digital image is first partitioned into small blocks of the regions. The blocks are then reordered so that they are placed a distance to each other that is proportional to the differences in their pixel colours [36]. Since it is statistically unlikely to find identical and spatially coherent regions in an image, therefore their presence can be used as evidence of tampering. In lighting approach the direction of an illuminating light source for each object or person in an image is automatically evaluated by some mathematical techniques. The retouching technique exploits the technology by which a digital camera sensor records an image, for detecting a specific form of tampering.

5. Challenging Scenarios for Video Authentication

In some of the surveillance systems storage and transmission costs are the important issues. In order to reduce the storage and transmission cost only those video clips which contain objects of interest are required to be sent and stored. Moreover in most of the surveillance applications, background object changes very slowly in comparison to foreground objects. A possible efficient solution in these scenarios is that only the objects of interest (mostly foreground objects) are sent out frame by frame in real time while the background object is sent once in a long time interval. In such surveillance applications, it becomes very critical to protect the authenticity of the video: the authenticity against malicious alterations and the authenticity for the identity of the transmission source (i.e. identifying the video source). In event based surveillance systems, the video sequences are captured when there is any kind of change in the scene (existence of an event) which would be captured by the camera. If there is uniformity in the scene in such a way that there is not any change in the scene then the surveillance camera does not capture any video sequence. This kind of surveillance system is used in military system for border security purpose. Authenticity for this kind of video sequences is a challenging issue because there is no proper time sequence in video sequences which are captured by surveillance camera. These are the scenarios which pose considerable challenges to the researchers for authentication.

6. Summary

Fig. 9 presents a tree structure of the methodologies that can be used for video authentication. The four children node of the root node covers almost all the methodologies. The leaf nodes of the tree structure show the key points of their grandparent node methodologies. This tree structure shows how all the methodologies use different approaches for video authentication. However many work has been done in watermarking and digital signature methodologies, other techniques (including intelligent technique) also produce better results for authentication purpose. There is no issue related with the size of authentication code in digital signature techniques, however, they provide better results regarding robustness, since the digital signature remains unchanged when there is a change in pixel values of the video frames. But if the location where digital signature is stored is compromised then it is easy to deceive the authentication system. On the other hand fragile watermarking algorithms perform better than algorithm based on conventional cryptography [32]. Fragile and semi fragile algorithm show good results for detecting and locating any malicious manipulations but often they are too fragile to resist incidental manipulations. Moreover embedding the watermark may change the content of video which is not permissible in court of law [33]. In addition of these techniques, intelligent techniques explore the new dimensions in video authentication. However learning based intelligent authentication algorithm does not require computation and storage of any key or embedding of secret information in the video data, it requires a large database of tampered and non tampered video to learn the algorithm so that it can classify whether the given video is authentic or not. These techniques are slower than some existing authentication techniques, since they use sufficient large database to learn the algorithm. In other techniques, most of the authentication techniques are established for specific attacks. For example motion trajectory based algorithm only detects the frame addition and deletion attacks (temporal attacks). Moreover compression and scaling operations also affect the performance of existing algorithms.

7. Conclusion

Video authentication is a very challenging problem and of high importance in several applications such as in forensic investigations of digital video for law enforcement agencies, video surveillance and presenting video evidence in court of law. However with growing development in video editing tools and wide availability of these powerful editing software video tampering attacks explores new dimensions in various fields. In future it is going to be a big menace for information security. By analysing the various video authentication techniques that were presented in this paper we can say that the authentication techniques are specific to the applications (surveillance, entertainment industry, medical, copyright...). As the time passes, we are getting more involved with video applications, in our daily lives. Our information systems are greatly dependent on video applications, now. This, with a wide range of tampering attacks, causes severe challenges on information security. In future robustness would be the key point for video authentication techniques, so that it can differentiate the acceptable video processing operations from malicious tampering attacks. However A perfect video authentication algorithm that detects all kinds of malicious manipulations and that can tolerate all content preserving manipulations is yet to be discovered. We can hope for the better in the future.

References:

[1] B.G Mobasseri, M.S.Sieffert, R.A.Simard, Content Authentication and tamper detection in digital video, Proc. IEEE International conference on Image Processing, Vancouver, September 10-13, 2000.

[2] B.G.Mobasseri, A.E.Evans, Content dependent video authentication by self water marking in color space, Proc. Security and watermarking of multimedia contents III, vol. 4314 pp.35-46, January 21-26, 2001.

[3] M.V. Celik et al, Video authentication with self recovery, Proc. Security and watermarking of multimedia contents IV vol. 4314, pp. 531-541, January 21-24, 2002.

[4] Fabrizio Guerrini, Reccardo Leonardi and Pierangelo Migliorati A new video authentication template based on bubble random sampling.

[5] Peng Yin, Hong heather Yu, Classification of Video Tampering Methods and Countermeasures using Digital Watermarking Proc. SPIE Vol. 4518, p. 239-246, Multimedia Systems and Applications IV

[6] Pradeep K. Atrey, Wei-Qi Yan, Ee-Chien Chang, Mohan S. Kankanhalli, A hierarchical signature scheme for robust video authentication using secret sharing.

[7] Johns Hopkins APL creates system to detect Digital Video Tampering. http://www.jhu.edu/

[8] Ching-Yung Lin, Shih-Fu Chang, "Issues and Solutions for authenticating MPEG Video" SPIE electronic Imaging 1999. San Jose.

[9] M. P. Queluz Authentication of digital images and video: Generic models and a new contribution.

[10] Vapnik VN (1995) The nature of statistical learning theory. Springer Verlag.

[11] Singh R., Vatsa M., Noore A (2006) Intelligent biometric information fusion using support vector machine. In soft computing in Image processing: Recent advances, Springer Verlag 327-350.

[12] R. Gennaro and P. Rohatgi, How to sign digital stream, Crypto' 97, pp. 180-197, 1997.

[13] J. M. Park, E. K. P. Chong and H. J. Siegel, Efficient multicast packet authentication using signature amortization, IEEE symposium on security and privacy, pp. 227-240, 2002.

[14] Kovesi PD (1999) Image features from phase congruency. Videre: Journal of Computer vision research, MIT Press 1(3).

[15] Navajit Saikia, Prabin K Bora, Video Authentication using temporal wavelet transform.

[16] Chang-yin Liang, Ang Li, Xia-mu Niu Video authentication and tamper detection based on cloud model.

[17] Ditmann, J.; Steinmetz, A; Steinmetz, R., Content based digital signature for motion pictures authentication and content fragile watermarking, Multimedia computing and systems, 1999. IEEE International Conference on, Volume: 2, 1999, Page(s): 209-213 vol. 2.

[18] Queluz, M. P., Toward robust, content based techniques for image authentication, Multimedia signal processing, 1998 IEEE Second workshop on, 1998 page(s): 297-302.

[19] Wei-Qi Yan an Mohan S Kankanhalli, Motion Trajectory Based Video Authentication ISCAS (3) 2003: 810-813

[20] Latechi L. Wildt D. and Hu J., Extraction of key frames from videos by optimal color composition matching and polygon simplification. Proceedings of MMSP' 2000, Cannes, France, October 2001

[21] Zhao L., Qi W., Li S., Yang S. and Zhang H., Key frame extraction and shot retrieval using Nearest Feature Line (NFL)., Proceedings of ACM Multimedia 2000.

[22] Quisquater J., Authentication of sequences with the SL2 Hash function application to video sequences, Journal of computer security, 5(3), pp: 213-223, 1997.

[23] Chun-Shien Lu and Hong Yuan Mark Liao, Structural digital signature for image authentication: An Incidental Distortion Resistant Scheme. *IEEE Trans. Multimedia*, vol. 5, no. 2, pp. 161-173, Jun. 2003.

[24] R. Singh, M. Vatsa, S.K. Singh, and S. Upadhyay, Integrating SVM Classification with SVD Watermarking for Intelligent Video Authentication, In Telecommunication Systems Journal - Special Issue on Computational Intelligence in Multimedia Computing, Springer, 2008.

[25] S. Bhattacharjee and M. Kutter, Compression tolerant image authentication, in IEEE International Conference on Image Processing, 1998, pp. 435-439.

[26] W. Diffie and M. E. Hellman, New Directions in cryptography, IEEE Trans. on Information Theory, Vol. 22, No. 6, pp.644-654, Nov 1976.

[27] P. Wohlmacher, Requirements and Mechanism of IT-Security Including Aspects of Multimedia Security, Multimedia and Security Workshop at ACM Multimedia 98, Bristol, U. K., Sep. 1998.

[28] Shui-Hua Han, Chao-Hsien Chu, Content based image authentication: current status, issues, and challenges. Int. J. Inf. Security (2010) 9:19-32, DOI 10.1007/s 10207-009-0093-2.

[29] S. Craver, N. Memon, B. Yeo and N. M. Yeung, Resolving Rightful Ownerships with Invisible watermarking Techniques: Limitations, Attacks and Implications, IEEE Journal on Selected Areas in Communications, Vol. 16, No. 4, pp. 573-586(1998).

[30] The Oxford English Dictionary, 2nd Edition, Oxford University, pp. 795-796, 1989.

[31] The Webster's New 20th Century Dictionary.

[32] Adil Hauzia, Rita Noumeir (2007) Methods for image authentication: a survey. In: Proceedings of the Multimedia Tools Appl (2008) 39:1-46, DOI 10.1007/s11042-007-0154-3

[33] S. Upadhyay, S.K. Singh, M. Vatsa, and R. Singh, Video authentication using relative correlation information and SVM, In Computational Intelligence in Multimedia Processing: Recent Advances (Springer Verlag) Edited by A.E. Hassanien, J. Kacprzyk, and A. Abraham, 2007

[34] Jana Dittman, Anirban Mukharjee and Martin Steinbach Media independent watermarking classification and the need for combining digital video and audio watermarking for media authentication. International conference on Information Technology: Coding and Computing, 2000.

[35] Weihong Wang, Hany Farid Exposing digital forgeries in video by detecting double MPEG compression.

[36] Hany Farid, Digital doctoring: How to tell the real from fake

[37] Po-chyi Su, Chun-chieh Chen and Hong Min Chang, Towards effective content authentication for digital videos by employing feature extraction and quantization

[38] Pradeep K. Atrey, Abdulmotaleb El Saddik, Mohan Kankanhalli, Digital Video Authentication, IGI Global, 2009.

Saurabh Upadhyay received the B. Tech. degree in computer science and engineering in 2001 and is currently working toward the Ph.D. degree in computer science at U.P. Technical University, India. He is an Associate Professor in the Department of Computer Science and Engineering, Saffrony Institute of Technology Gujarat, India. He is actively involved in the development of a robust video authentication system which can identify tampering to determine the authenticity of the video. His current areas of interest include pattern recognition, video and image processing, watermarking, and artificial intelligence

Sanjay K. Singh is Associate Professor in Department of Computer Engineering at Institute of Technology, BHU, India. He is a certified Novel Engineer and Novel administrator. His research has been funded by UGC and AICTE. He has over 50 publications in refereed journals, book chapters, and conferences. His research interests include computational intelligence, biometrics, video authentication and machine learning. Dr. Singh is a member of IEEE, ISTE and CSI.