TAMPERED VIDEOS: DETECTION AND QUALITY ASSESSMENT

Synopsis of the Thesis to be submitted in fulfillment of the requirements for the Degree of

DOCTOR OF PHILOSOPHY

By

MANISH KUMAR THAKUR

Department of Computer Science Engineering and Information Technology

JAYPEE INSTITUTE OF INFORMATION TECHNOLOGY
(Declared Deemed to be University U/S 3 of UGC Act)
A – 10, Sector 62, Noida, India
October, 2013

TAMPERED VIDEOS: DETECTION AND QUALITY ASSESSMENT

1. Introduction

From the beginning of human civilization, visual information is the most often used medium to express knowledge, thoughts, evidences, *etc.* and represents one of the effective means for communication. Visual information has a capability to convey the broader spectrum of information because of its ease in acquisition, distribution, and storage. In the modern age, images and videos have become the main information carriers to disseminate knowledge and establish the bridge among several sources.

Developments in visual (video) technologies *viz.* compression, transmission, storage, retrieval, and video-conferencing have helped in many ways to the society. In the socio-economic knowledge and scientific development, the images and videos available at various video sharing and social networking websites (*viz.* YouTube, faceBook, *etc.*) are playing a significant role. Besides this, other applications like entertainment industry, video surveillance, legal evidence, political videos, video tutorials, advertisements, *etc.* signify their unprecedented role in today's context.

Apart from many good things, there are some darker sides of visual (video) information *viz*. misuse or the wrong projection of information through videos. One of them is video tampering, where a forger can intentionally manipulate real (actual or original) videos to create tampered or doctored or fake videos for malpractice [1-3] *viz.* video based forgery, doctored video evidences, criminal activities, *etc*.

Easy availability of many sophisticated video editing tools provides a platform for forger to manipulate real videos and create perceptually indistinguishable fake videos. Therefore, in many serious scenarios *viz.* court trials, law enforcement, defamation, politics, and defense planning, *etc.* authenticity of presented video needs to be examined.

Forensic tools and experts play a key role to examine the authenticity of videos by detecting traces (if any) of tampering. Here, success or failure of tools and experts depends on how intelligently tampering has been carried out by the forger. It is difficult for forensic

experts to detect tampering with videos if there are no (or little) traces left by forger while tampering.

Unfortunately, due to lack of established methodologies to examine the authenticity of videos [1], detection of tampering with videos have posed challenges before the scientific community, and its seriousness in many scenarios (*viz.* videos as evidence during court trials) seeks immediate attention.

Further, success or failure of forger depends on how intelligently fake videos have been made to deceive forensic tools and the expert eye. In literature, there are many counter-forensic (or anti-forensic) schemes available which facilitate forger to create fake videos to deceive forensic tools [2][4][5], whereas, to deceive perceptually, the least quality degradation in tampered videos is required.

On the basis of quality assessment between real (actual) and tampered videos, forger can decide the percentage of tampering needed to be introduced so that the perceptually indistinguishable fake (tampered or doctored) videos deceive expert (or casual) eye.

In addition, quality assessment is also helpful for forensic experts to analyze the authenticity of videos. Some schemes have been presented in [6-9] which traces the tampering based on quality assessment. Thus, the video quality assessment is used by forensic experts for tampering detection and by forger to ensure least quality degradation in tampered videos. But, lack of schemes to assess the video quality in tampered videos with respect to actual videos needs an urgent attention of the scientific community.

In this thesis, we developed schemes for video tampering detection and video quality assessment. The onward sections, describe the current problems in tracing the tampering with videos and assessing the video quality, challenges and currently known methods/technologies applied in this direction along with our proposed objectives.

2. Video: Tampering and Detection

However, video tampering is relatively new area, image doctoring is as old as the art of photography itself where we have numerous incidences of serious cases of fake photographs [1][10][11].

While tampering a video, objective of a forger is to create a tampered or doctored or fake video from real or actual or original video. These real videos are the source for creating

tampered videos. Tampering can be done either on a single video (*i.e.* single source) or on multiple videos (*i.e.* many sources) [2]. **In this thesis**, we considered the single source based video tampering and developed schemes for tampering detection in such tampered videos.

The seriousness of video tampering depends on how and where these tampered videos have to be used. Court trials are one of the most widely used application areas where these tampered videos are presented as evidence to mislead the court proceedings. Thus, whenever videos are presented as evidence during court trials, their authenticity are to be examined before considering them as evidence [1][2][12-15].

Here, forensic tools and forensic experts play the key role to examine the authenticity of video evidences. During examination, if it has been found as authentic (*i.e.* non-tampered or actual), experts generally embed watermark into authenticated videos such that whenever required its authenticity can be re-examined by retrieving the watermark [16]. In addition, if all parties (of court trials) agree, copies of watermark videos can be created. Such copies can be used during long court trials [17] in place of actual videos (with embedded watermark).

Thus, while examining videos, there may be following possibilities with forensic experts: **(a)** Forensic experts may need to blindly examine (*i.e.* trace the tampering if any) the videos *i.e.* no information is available about the original source video from which the tampered video was created, **(b)** Forensic experts may need to trace the tampering with a copy of watermark videos with reference to actual videos (with embedded watermark), and **(c)** Forensic experts may need to trace the tampering in actual videos (with embedded watermark).

Further, different domains in which tampering is possible are discussed in section 2.1, levels of tampering are described in section 2.2, and the most common tampering (temporal) has been presented in section 2.3. In section 2.4, various modes of tampering detection have been described.

2.1 Tampering Domain [18-23]

Depending upon the domain in which manipulation is done, there can be following types of video tampering

a. Tampering in spatial domain (*i.e.* Spatial Tampering)
b. Tampering in temporal domain (*i.e.* Temporal Tampering)
c. Tampering in spatio-temporal domain (*i.e.* Spatio-Temporal Tampering)

A forger can tamper source videos spatially (*i.e.* spatial tampering) by manipulating pixel bits within a video frame or in adjacent video frames. Fig. 1b presents a spatially tampered video created from the actual video of Fig. 1a.

Further, as presented in Fig. 1c, forger can tamper source videos with respect to time (*i.e.* temporal tampering) by disturbing the frame sequence through frames replacement, frames sequence reordering, frames addition, and by the removal of video frames.

Lastly, as presented in Fig. 1d, forger can tamper videos in combination of both spatial and temporal domain (*i.e.* spatio-temporal tampering) by manipulating pixel bits within a video frame or set of adjacent frames as well as disturb the frame sequence.

Over the past few years, significant contributions are available for detection of spatial tampering (or in other way, image tampering), whereas, relatively little contributions have been made for detection of temporal tampering.

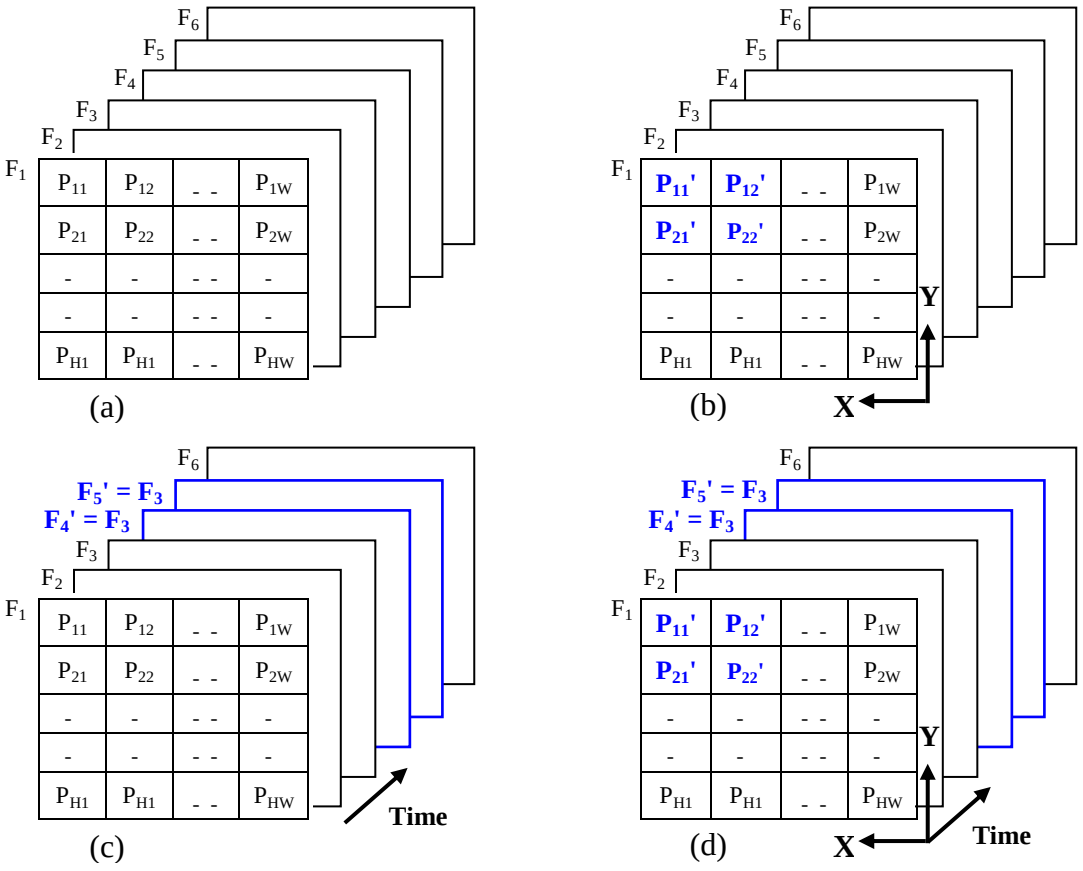

Fig. 1 An example of **(a)** Actual video **(b)** Spatially tampered video **(c)** Temporally tampered video and **(d)** Spatio-temporal tampered video. Where F_i represents the i^{th} frame and P_{ij} is the pixel intensity. H and W are frame height and width respectively. F_i' is the manipulated i^{th} frame and P_{ij}' is the manipulated pixel intensity.

Thus, **this thesis focuses to detect the temporal tampering** in spatio-temporal and temporally tampered videos. Next section describes the level at which temporal tampering is possible.

2.2 Levels of Temporal Tampering [18-20][24-26]

Videos can be manipulated (or tampered) by a forger at following levels

a. Frame level
b. Scene level or shot level
c. Video level

Videos can be manipulated at frame level *viz.* frames of a video (not necessarily of the same scene) can be deleted, or intermediate frames of a video scene can be removed. At the scene level, an entire scene of a video is manipulated *viz.* deletion of a video scene (*i.e.* a scene or shot cut), copying of a video scene to another place, *etc*. Finally, there can be manipulations at the video level *viz.* making copy of a video.

Although, videos can be temporally tampered at the video level, some temporal tampering is not feasible at the video level *viz.* deletion of frames at video level will result in the deletion of entire video. Thus, **this thesis focuses to detect temporal tampering** in such videos which were created by manipulating video frames **at frame and scene level**. Next section presents commonly used schemes in literature to perform temporal tampering with videos.

2.3 Most Common Temporal Tampering in Videos [2][13][18-22][24][25][27-46]

However, one can temporally manipulate videos in several ways, mostly addressed temporal tampering in literature are mentioned below.

a. Frame Drop or Frame Deletion or Frame Removal
b. Frame Swapping or Reordering of frame sequences
c. Frame Copy or Frame Addition or Copy-paste
d. Copy-move
e. Frame Averaging
f. Frame Replacement

While tampering a video (source), forger can drop (delete or remove) frames of his/her choice, resulting in tampered (or doctored) videos with reduced frame count. These frames can be the intermediate frames of a video scene or can be the set of frames spread into two scenes. Fig. 2a presents an example of frame drop at the frame level. Further, deletion can also be at the scene level, *i.e.* an entire scene is removed (or cut) by deleting all frames in that scene. Deletion at the scene level is often known as Scene or Shot Cut.

Unlike frame drop, frame count remains unchanged while swapping the video frames (or reordering frame sequences) to create a tampered (or doctored) video from actual (source) videos. At the frame level, these swapped frames can be some video frames of one/two scene(s), whereas, there will be swapping of entire scene at the scene level, *i.e.* all frames of one scene will be swapped with all frames of another scene. Fig. 2b presents an example depicting tampering of the frame swapping at the frame level.

Frame count will be increased if a source video is manipulated by copying the video frames and pasting to some other location in the source video. At the frame level, these copied frames may be the intermediate frames of a video scene, or at the scene level, an entire scene can be copied and pasted after another scene. Even, copying can also be at the video level, *i.e.* all frames of a video sequence can be copied and pasted into another video such that the copy of a source video is created. Fig. 2c presents an example depicting tampering of frame copying at the frame level.

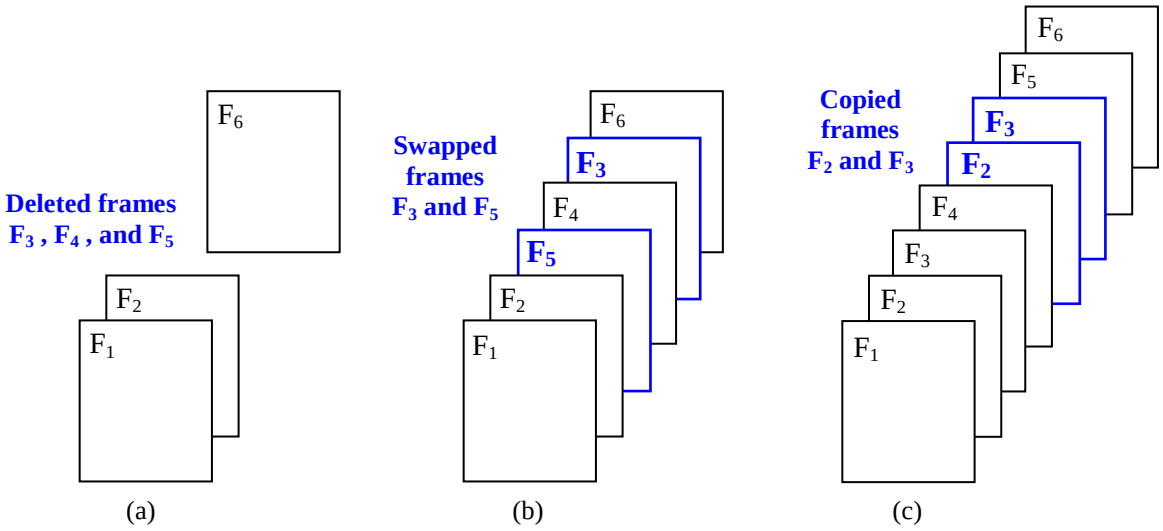

Fig. 2 An example of **(a)** Frame Drop, **(b)** Frame Swapping, and **(c)** Frame Copying where source is the video sequence presented in Fig. 1a

In copy-move tampering, frames of source videos are moved to another location, whereas in frame averaging, an average frame (created by averaging adjacent frames) is added between two video frames. In frame replacement, a video frame is replaced by another frame of the source video or by a foreign frame (a frame of another video).

This thesis focuses to detection of temporal tampering in tampered videos if created by **frame drop** (at frame level), **frame swapping** (at frame level), and **frame copying** (at scene level). Next section presents different modes to detect tampering with videos.

2.4 Modes in Detection of Tampering

[14][22][28][29][31][36][41][42][47][48]

Depending upon the availability of original (actual) videos as reference (or source), tampering with videos can be detected in following modes.

a. Full Reference (FR): In this mode, both, actual (source or reference) videos and tampered videos are available to forensic experts which need to examine the tampered video with reference to the actual video.

b. Reduced Reference (RR): Under NR mode, little information about actual (source or reference) videos are available to forensic expert for tracing (if any) the tampering in tampered videos.

   Tampering detection techniques (*viz.* detection of tampering using watermark, using properties of capturing device, *etc.*) under FR and RR mode are often called as active detection techniques. In literature for active tampering detection, generally watermark or digital signatures are initially embedded in original videos so that, later on, if required, its integrity and authenticity can be examined by retrieving the watermark [1][2][16].

c. No Reference (NR): Unlike FR and RR, in this mode, no other information available to forensic experts about the video whose authenticity is to be examined. Detection techniques under NR mode are often called as passive detection techniques.

This thesis focuses to detect the tampering in **full reference (FR) and no reference (NR)** modes. Next section presents the second issue (addressed in this thesis) related with tampered videos, *i.e.* video quality assessment.

3. Video Quality Assessment

Based on reviewed literature, we identified different types of quality assessment; modes of quality assessment; requirements of subjective experiments; and types of objective quality metrics.

Quality of a video can be measured in two ways *viz*. subjective assessment and objective assessment [49][50] where quality assessments are generally conducted in three modes *viz*. full reference (FR) quality assessment, reduced reference (RR) quality assessment, and no reference (NR) quality assessment [51][52].

Subjective quality assessment involves human subjects to measure the video quality (in all modes *i.e.* FR, RR, and NR), whereas, objective quality assessment involves various objective quality metrics *viz*. Mean Squared Error (MSE), Peak Signal to Noise Ratio (PSNR), Structural Similarity Index (SSIM), Video Quality Metrics (VQM), *etc.* to measure the video quality [53][54]. MSE, PSNR and SSIM are used in FR mode whereas VQM can be used in both FR and NR modes [52][55].

Subjective experiments are the basis of human visual system (HVS) and have been employed for designing many objective metrics. There are many requirements which need to be taken care of while conducting subjective experiments. These involve a set of reference (or source) and distorted videos; human subjects (trained, untrained or mixed); experiment procedures (*viz.* single stimulus continuous quality evaluation and double stimulus continuous quality evaluation); viewing conditions and setup; experiment sessions (around 30 to 40 minutes); scoring policy (*viz.* qualitative and quantitative); computation of the mean opinion score; and a subject rejection scheme [50][56][57].

Broadly, objective quality metrics are classified in two types *viz.* statistical metrics and perceptual metrics such as PSNR and SSIM respectively [52][53][54]. Statistical metrics employ the application of various mathematical functions like calculating pixel-by-pixel weighted differences (*viz.* MSE and PSNR) between reference and distorted video, whereas perceptual metrics use HVS characteristics [52][58]. Further, while accessing quality using objective quality metrics, it is always desired to meet the requirements of HVS.

Next section presents the critical review of some existing schemes related with video tampering detection and video quality assessment.

4. Critical Review of Some Existing Schemes

4.1 Schemes for Detection of Temporal Tampering

In recent years, many researchers contributed to develop FR and NR watermarking schemes to detect the embedded watermarks and classify videos as tampered or non-tampered accordingly. These schemes have been claimed to be robust against temporal tampering of frame drop, frame copying, frame averaging, and frame swapping [18][20][30-34][36][40]. The basic principle all these schemes is that tampering leads to change in the embedded watermark, so while retrieval, if there is change in embedded watermark or watermark is not retrieved back, then the video can be classified as a tampered video. Table 1 summarizes some of the reviewed FR schemes.

Although, these schemes can classify videos as tampered or non-tampered, they fail to identify the exact location of temporal tampering (if any) which is of equal importance because, one may be interested to know the locations of tampering if the source video is available. Considering a case where embedded watermark gets destroyed due to spatial tampering (by manipulating pixel bits) and temporal tampering (*viz.* by the removal of some video frames), watermark based schemes are likely to be inefficient to identify tampering locations.

Furthermore, in FR mode (where the source video is available) if identification of the tampering location is not required, embedding watermark (which degrades visual quality) to later on classify a video as tampered or non-tampered is not the only solution. In such a case (*i.e.* without identification of the tampering location), spatial tampering can easily be detected by comparing pixels of the source and tampered videos whereas, temporal tampering (due to frame drop or frame copying) can be detected by change in frame count.

Under NR mode where there is no information about actual (source or reference) video, tampering detection is based on abnormalities in the surrounding pixels (in spatial domain) or adjacent frames (in temporal domain). Unlike FR and RR (*i.e.* active tampering detection), NR tampering detection (often referred as passive tampering detection techniques) is relatively new research area and not much contributions are available. Due to diversity in type of tampering a single scheme may not be able to identify all types of tampering.

Table 1 Summary of surveyed papers to detect various temporal tampering

Temporal Tampering	Scheme; References; Claims
Frame Drop (or deletion or removal)	Watermarking based schemes have been presented in [18][20][30-34][36][40] and claimed to be robust against tampering of frame drop.
	Learning based scheme using Support Vector Machine (SVM) is presented in [22] and claimed to achieve around 99% accuracy to trace the tampering of frame drop.
	Threshold based schemes have been presented in [25][26] and claimed to trace the tampering of frame drop at scene level *i.e.* Scene (or Shot) Cut.
	Hashing based scheme is proposed in [24] and claimed to be robust against different tampering including frame drop.
	Compression properties have been explored using discrete cosine transformation in [29] to trace the tampering of frame drop.
	Threshold based schemes have been presented in [46] and claimed to detect the dropped frames at frame level.
Frame Swapping (or reordering)	Watermarking based schemes have been presented in [18][20] and claimed to be robust against tampering of frame swapping.
Frame Copying (or addition)	Learning based scheme using SVM classifier is presented in [22] and claimed to achieve around 99% accuracy to trace the tampering of frame addition (or copy).
	Watermarking based schemes have been presented in [18] and claimed to be robust against tampering of frame copy.
	Compression properties have been explored using histogram of oriented gradients in [21] to trace the tampering of frame copy (addition).
	Hashing based scheme is proposed in [24] and claimed to detect frame copy at the video level, *i.e.* Video Copy.
Others	Watermarking based schemes have been presented in [33-34][36] and claimed to be robust against tampering of frame averaging.
	Training based scheme using Gaussian Mixture Model is presented in [19] and claimed to trace tampering of Copy-move.
	Undecimated dyadic wavelet transform is used in [44] to trace the tampering of Copy-move in an image.
	MPEG double compression is detected in [46-48] to trace a tampering.

In recent years, researchers have proposed NR tampering detection schemes to detect and handle specific or set of different types of video tampering (*viz.* frame drop, frame copying, *etc.*). These include learning based schemes, threshold based schemes, and schemes based on codec (compression) properties [22][25][26][28][29][41-48].

Besides FR schemes, Table 1 also summarizes reviewed NR tampering detection schemes. In these schemes, the major challenge is to efficiently (or accurately) identify either inconsistencies (due to manipulations in pixel bit) in the surrounding pixels of a video frame (for spatial tampering detection) or inconsistencies (due to frame drop, frame copying, and frame swapping, *etc*.) in adjacent frames (for temporal tampering detection).

Further, tampering detection under NR mode has also been handled by detecting abnormalities in coding standards, especially MPEG. Schemes have been presented by researchers in recent years to detect MPEG double compression, quantization errors, break-points, and sub-sampling [29][43-45].

As to process a video, one is required to get it uncompressed, do the required processing and may get it recompressed (not necessarily with same codec) or keep it in raw format. Considering a video as tampered video, tampering detection based on abnormalities and inconsistencies in video formats (which may be ineffective if compressed with other codec or keep the video in raw format) is the matter of further discussion.

Based on these identified gaps, there is a strong need to propose efficient and reliable NR tampering detection schemes for various types of temporal tampering (*viz.* frame drop, frame copying, *etc*.) with videos.

4.2 Schemes for Video Quality Assessment

While objective quality assessment in the distorted (tampered) video with reference to the original (reference) video, it is always desired not to have much variation between measured quality and perceived quality.

This requires first to propose a subjective model (based on conducted subjective experiments) for specific or set of distortions and obtain the mean opinion score (MOS). Challenges lie into the proposal of an effective subjective model where obtained MOS should represent actual perceived quality. In addition, the obtained MOS is the basis to analyze the performance of many objective quality metrics (*viz.* PSNR, SSIM, *etc*.) to measure the quality degradation.

Two most commonly documented quality metrics *viz.* PSNR and SSIM have been employed for quality assessment of images. Currently, in the literature, these metrics have been effectively applied for video quality assessment where quality is measured frame by

frame (*i.e.* considering frame as an image) between reference and distorted videos and overall video quality is assessed by sample mean of the frame by frame assessed quality [59].

It is likely that just by taking average; these metrics might be inefficient to measure quality of temporally distorted / tampered videos [59-63] where sequence or count of frames may get changed compared with reference videos. Thus, the performance of PSNR and SSIM metrics to measure the quality of temporally tampered videos needed to be examined.

Through conducted experiments (for spatio-temporal distorted videos, where, frame drop is used as temporal distortion), we observed the huge variation between measured quality by often used objective metrics (PSNR and SSIM) and the proposed subjective model. This needs an attention to propose a video quality metrics which can efficiently measure the quality degradation in temporally distorted (tampered) video sequences.

As an outcome of literature review (tampering detection and quality assessment), next section presents identified research issues.

5. Identified Issues

Based on reviewed literature and experiments conducted for quality assessment, we identified following issues which require immediate attention.

Issue 1: Under FR tampering detection, there is lack of schemes which can efficiently identify the tampering locations in temporally tampered videos. Thus, there is a strong need to propose an efficient FR scheme to identify tampering locations in such videos.

Issue 2: Embedding of watermark for tampering detection leads to degrade the visual quality of source videos, thus, there is a need to propose an efficient FR scheme without compromising the quality of source videos.

Issue 3: Under NR tampering detection, a reliable and efficient scheme to trace the tampering is always desired. Thus, reliable and efficient NR schemes to trace and detect different types of temporal tampering needs to be proposed.

Issue 4: As after manipulations with malafied intensions, forger can store tampered videos in any format (may compress using different codec or keep in raw format), thus needs to propose format independent schemes which can detect tampering with videos.

Issue 5: Two most commonly documented quality metrics *viz.* PSNR and SSIM are inefficient to measure the quality degradation of spatio-temporal distorted videos, thus there is a need to devise a video quality metrics which can efficiently measure the degraded quality of spatio-temporal distorted videos.

6. Thesis Objective

Based on reviewed literature and issues identified, we considered to detect and trace the tampering of frame drop, frame swapping, and frame copying. We developed tampering detection schemes for frame drop under FR and NR modes, whereas frame swapping and frame copying were developed in NR mode.

Detection of frame drop has been considered to be developed for both spatio-temporal tampered and temporally tampered videos whereas frame swapping and frame copying have been considered to be developed only for temporally tampered videos. Frame drop and frame swapping have been considered to be addressed at frame level, whereas, frame copying has been considered to be addressed at scene level. To facilitate format independent tampering detection, all schemes have been developed for raw videos.

Further, we considered to address quality assessment for spatio-temporal distortion (tampering) where frame drop has been considered temporal distortion. In summary, following are the set of objectives:

Objective 1: Development of a full reference (FR) algorithm to identify the exact location of tampering in spatio-temporal tampered videos where temporal tampering has been caused due to frame drop.

Objective 2: Development of no reference (NR) algorithm (s) to classify a video as tampered video and identify the location of tampering in temporally tampered videos where temporal tampering has been caused either due to frame drop, frame swapping, or frame copying.

Objective 3: Development of a full reference (FR) video quality metric which is capable to measure quality degradation in spatio-temporal distorted (tampered) videos where temporal distortion has been caused due to frame drop.

7. Thesis Organization

Apart from introduction (presented in Chapter 1) and literature review (presented in Chapter 2), the remaining chapters are organized as follows.

In **Chapter 3**, we developed a full reference scheme using genetic algorithm to identify the location of dropped frames (considering that tampered videos were created by spatio-temporal tampering). Further, we introduced data-parallelism to make the scheme scalable.

Addressed problem statement in this chapter is as follows: Let us consider an original (actual) video V_0 with m video frames which is being tampered by dropping some video frames and manipulating the remaining frames spatially too, thus creates a spatio-temporal tampered video V_T with n video frames, where, $m \geq n$. For given V_0 and V_T (*i.e.* FR), objective is to identify frame indices (*i.e.* the exact location of tampering) in V_0 which were dropped to create tampered video V_T.

We simulated and tested the proposed scheme using various sets of tampered and actual videos. In temporal (frame drop) tampered videos, we achieved 100% accuracy to detect the dropped frame indices, whereas, depending upon the percentage of spatial tampering, achieved accuracy is in the range of 79% to 100% for spatio-temporal tampered videos. The data-parallel scheme had been simulated and tested for videos of duration up-to 540 seconds. Here, we achieved the average speedup of around 1.77, 2.81, and 3.35 for 2, 4, and 8 processors (or threads) respectively.

Lastly, we concluded the chapter with comparative observations between developed scheme and watermark based schemes.

Chapter 4 presents a learning based NR scheme which identifies the frame indices in tampered videos (created by deleting video frames from the original video) after which some frames may have been dropped, *i.e.* traces the tampering of frame drop.

Addressed problem statement in this chapter is as follows: Let us consider an original (actual) video V_0 with m video frames which is being tampered by dropping some video frames, thus creates a temporally tampered video V_T with n video frames, where, $m \geq n$. For given V_T (V_0 is not available to experts, *i.e.* NR), objectives are (a) to classify V_T as a tampered video and (b) to identify the location of tampering in V_T, *i.e.* frame index after which frames were dropped in V_T.

Here, the proposed scheme uses Support Vector Machine (SVM) as a classifier which is being trained using 1500 tampered videos and it is successfully able to identify the frame index after which there is a trace of tampering due to the frame drop.

Performance of the proposed scheme has been observed for four features *viz*. mean square error, entropy, average object area, and count of displaced blocks. Finally, we concluded the chapter with the comparative analysis between the developed scheme and the scheme presented in [22].

Chapter 5 discusses the limitation of the scheme presented in the chapter 4 and proposes a threshold based NR scheme which identifies the frame indices in tampered videos (created by deleting video frames from the original video) after which some frames may have been dropped, *i.e.* traces the tampering of frame drop.

Addressed problem statement in this chapter is as follows: Let us consider an original (actual) video V_0 with m video frames which is being tampered by dropping some video frames of a scene S_i, thus creates a tampered video V_T (having less than m video frames). For given V_T (V_0 is not available to experts *i.e.* NR), objectives are (a) to classify V_T as a tampered video and (b) to identify location of tampering in V_T, *i.e.* frame index after which frames were dropped in V_T.

As pre-processing steps, in this chapter, we first presented two schemes *viz*. identification of change of scene (abrupt) and count of displaced blocks in two frames, and then defined various thresholds. These pre-processing schemes and recommended values of thresholds have been used to identify the frame index after which there is a trace of tampering of frame drop.

Performance of the proposed scheme has been observed using 60 tampered videos created by introducing tampering of frame drop into 4 source (actual) videos where achieved accuracy is around 85% to detect the tampering of frame drop. Finally, a comparative analysis is presented between the proposed scheme and other schemes.

Chapter 6 presents an NR scheme to identify tampering of frame copy (at the scene level) in tampered videos created by temporal manipulations in original (actual) videos.

Addressed problem statement in this chapter is as follows: Let us consider an original (actual) video V_0 with m frames which is being tampered by copying some video frames/scene to another location in V_0 and thus creates a tampered video V_T with n video

frames, where, $n > m$. For given V_T (V_0 is not available to experts, *i.e.* NR), objectives are (a) to classify V_T as a tampered video and (b) to identify location (frame index) of tampering in V_T after which frame/scene were copied.

In our proposed solution, we first identified the change of scene in given videos using pre-processing schemes presented in the chapter 5, and then identified the tampering of frame copy at the scene level.

We simulated the proposed scheme using 40 tampered videos created by introducing the tampering of frame copy into 4 source (actual) videos and achieved an accuracy of around 89% to detect the tampering of frame/scene copying.

Chapter 7 proposes an NR scheme to identify tampering of frame swapping in tampered videos created by temporal manipulations in original (actual) videos.

Addressed problem statement in this chapter is as follows: Let us consider an original (actual) video V_0 with m frames which is being tampered by swapping some video frames of a scene S_i with frames of a scene S_j and thus creates a tampered video V_T with m frames (*i.e.* the frame count is unchanged). For given V_T (V_0 is not available to experts, *i.e.* NR), objectives are (a) to classify V_T as a tampered video and (b) to identify location (frame index) of tampering in V_T after which some frames of two scenes were swapped.

We presented a solution for stated problem and observed its performance with 60 tampered videos created by introducing the tampering of frame swapping into 4 source (actual) videos and achieved an accuracy of around 75% to detect the tampering of frame swapping.

Chapter 8 deals with conducted subjective experiments and presents the obtained mean opinion score (MOS) for spatio-temporal distorted (tampered) videos. Watermarks were embedded using the least significant bit scheme to distort videos spatially, whereas, some frames were dropped to introduce temporal distortion. Further, the MOS was used to analyze the performance of two objective quality metrics *viz*. PSNR and SSIM. Finally, we proposed a FR dropped frames video quality metrics and through various experiments observed least variation between assessed quality and obtained MOS.

Keywords: video tampering; frame drop; frame swap; frame copy; video forensics; passive detection techniques; full reference; no reference; video quality metrics

References

[1] Rocha, A., Scheirer, W., Boult, T., and Goldenstein, S., "Vision of the unseen: Current trends and challenges in digital image and video forensics," *ACM Computing Surveys*, Vol. 43, No. 4, Article 26, October 2011, pp. 1-42. DOI: 10.1145/1978802.1978805

[2] Redi, J. A., Taktak, W., and Dugelay, J. L., "Digital image forensics: a booklet for beginners," *Multimed Tools Appl*, Vol. 51, Issue 1, Jan 2011, pp. 133–162. DOI: 10.1007/s11042-010-0620-1

[3] Wang, W., "Digital video forensics," *Ph.D. dissertation*. Department of Computer Science, Dartmouth College, Hanover, New Hampshire, June 2009

[4] Stamm, M. C., and Liu, K. J. R., "Anti-Forensics for Frame Deletion/Addition in MPEG Video," in *Proc. IEEE International Conference on Acoustic Speech and Signal Processing* (*ICASSP 2011*), May 22-27, 2011, pp. 1876-1879. DOI: 10.1109/ICASSP.2011.5946872

[5] Harris R., "Arriving at an anti-forensics consensus: examining how to define and control the anti-forensics problem," *Digital Investigation* 3 (Supplement 1), September 2006, pp. 44–49

[6] Avcibas, I., Kharrazi, M., Memon, N., and Sankur, B., "Image steganalysis with binary similarity measures," *Journal of Applied Signal Processing*, 2005, pp. 2749-2757

[7] Bayram, S., Avcibas, I., Sankur, B., and Memon, N., "Image Manipulation Detection with Binary Similarity Measures," in *Proc. European Signal Processing Conference*, 2005, pp. 752-755

[8] Battisti, F., Carli, M., and Neri, A., "Image Forgery Detection By Means Of No-Reference Quality Metrics," in *Proc. SPIE 8303, Conference on Media Watermarking, Security, and Forensics 2012*, 83030K, January 2012. DOI: 10.1117/12.910778

[9] Zhou, Z. P. and Zhang, X. X., "Image Splicing Detection Based on Image Quality and Analysis of Variance," in *Proc. 2nd International Conference on Education Technology and Computer* (ICETC 2010), June 22-24, 2010, pp. V4-242-V4-246

[10] Top 10 Doctored Photos. http://www.youtube.com/watch?v=9UOHauPRKbw Accessed September 30, 2013

[11] Photo tampering throughout history. http://www.cs.dartmouth.edu/farid/research/digitaltampering/ Accessed September 30, 2013

[12] http://www.bbc.co.uk/news/science-environment-20629671 Accessed September 30, 2013

[13] Kobayashi, M., Okabe, T., and Sato, Y., "Detecting Video Forgeries Based on Noise Characteristics," in *Proc. Pacific-Rim Symposium on Image and Video Technology* (PSIVT 2009), LNCS 5414, January 13-16, 2009, pp. 306–317

[14] Farid, H., "Image Forgery Detection," *IEEE Signal Processing Magazine*, Vol. 26, Issue 2, March 2009, pp. 16-25

[15] Farid, H. and Lyu, S., "Higher-order Wavelet Statistics and their Application to Digital Forensics," in *Proc. Conference on Computer Vision and Pattern Recognition Workshop* (CVPRW 03), June 16-22, 2003, pp. 94-101

[16] http://www.videoforensicexpert.com/tag/digital-video-forensic-evidence/ Accessed September 30, 2013

[17] http://www.videoforensicexpert.com/services/forensic-video-authentication-video-analysis/ Accessed September 30, 2013

[18] Atrey, P. K., Yan, W. Q., and Kankanhalli, M.S., "A scalable signature scheme for video authentication," *Multimed Tools Appl* (2007) 34, pp. 107–135

[19] Upadhyay, S. and Singh, S. K., "Video Authentication: Issues and Challenges," *International Journal of Computer Science Issues*, Vol. 9, Issue 1, No. 3, January 2012, pp. 409-418

[20] Atrey, P. K., Yan, W. Q., Chang, E.C., and Kankanhalli, M. S., "A Hierarchical Signature Scheme for Robust Video Authentication using Secret Sharing," in *Proc. 10th International Multimedia Modeling Conference* (MMM'04), Jan 5-7, 2004, pp. 330-337. DOI: 10.1109/MULMM.2004.1265004

[21] Subramanyam, A. V. and Emmanuel, S., "Video forgery detection using HOG features and compression properties," in *Proc. IEEE 14th International Workshop on Multimedia Signal Processing* (MMSP 2012), Sept 17-19, 2012, pp. 89-94. DOI: 10.1109/MMSP.2012.6343421

[22] Upadhyay, S. and Singh, S. K., "Learning Based Video Authentication using Statistical Local Information," in *Proc. International Conference on Image Information Processing* (ICIIP 2011), Nov 3-5, 2011, pp. 1-6. DOI: 10.1109/ICIIP.2011.6108953

[23] Zhang, J., Su, Y., and Zhang, M., "Exposing Digital Video Forgery by Ghost Shadow Artifact," in *Proc. MiFor'09,* October 23, 2009, pp. 49-53

[24] Malekesmaeili, M., Fatourechi, M., and Ward, R. K., "Video Copy Detection Using Temporally Informative Representative Images," in *Proc. International Conference on Machine Learning and Applications* (ICMLA'09), Dec 13-15, 2009, pp. 69-74. DOI: 10.1109/ICMLA.2009.32

[25] Yu, J. and Srinath, M. D., "An efficient method for scene cut detection," *Pattern Recognition Letters*, 22 (2001), pp. 1379-1391

[26] Yusoff, Y., Christmas, W., and Kittler, J., "Video Shot Cut Detection Using Adaptive Thresholding," in *Proc. British Mission Vision Conference* (BMVC 2000), pp. 11-14

[27] Roy, S. D., Li, X., Shoshan, Y., Fish, A., and Yadid-Pecht, O., "Hardware Implementation of a Digital Watermarking System for Video Authentication," *IEEE Trans. Circuits and Systems for Video Technology*, Vol. 23, No. 2, February 2013, pp. 289-301

[28] Goodwin, J. and Chetty, G., "Blind Video Tamper Detection Based on Fusion of Source Features," in *Proc. International Conference on Digital Image Computing: Techniques and Applications* (DICTA 2011), Dec 6-8, 2011, pp. 608-613. DOI: 10.1109/DICTA.2011.108

[29] Su, Y., Nie, W., and Zhang, C., "A Frame Tampering Detection Algorithm for MPEG videos," in *Proc. 6th IEEE Joint International Information Technology and Artificial Intelligence Conference* (ITAIC 2011), Aug 20-22, 2011, pp. 461-464. DOI: 10.1109/ITAIC.2011.6030373

[30] Chen, S. and Leung, H., "Chaotic Watermarking for Video Authentication in Surveillance Applications," *IEEE Trans. Circuits and Systems for Video Technology*, Vol. 18, No. 5, May 2008, pp. 704-709

[31] Wang, Y. and Pearmain, A., "Blind MPEG-2 Video Watermarking Robust Against Geometric Attacks: A Set of Approaches in DCT Domain," *IEEE Trans. Image Processing*, Vol. 15, No. 6, June 2006, pp. 1536-1543

[32] Chan, P. W. and Lyu, M. R., "A DWT-based Digital Video Watermarking Scheme with Error Correcting Code," in *Proc. 5th International Conference on Information and Communications Security* (ICICS 2003), LNCS 2836, Oct 10-13, 2003, pp. 202–213. DOI: 10.1007/978-3-540-39927-8_19

[33] Li, Y., Gao, X., and Ji, H., "A 3D Wavelet Based Spatial-Temporal Approach for Video Watermarking," in *Proc. Fifth International Conference on Computational Intelligence and Multimedia Applications* (ICCIMA'03), Sept 27-30, 2003, pp. 260-265. DOI: 10.1109/ICCIMA.2003.1238135

[34] Xu, D., Wang, R., and Wang, J., "Video Watermarking Based on Spatio-temporal JND Profile," in *Proc. 7th International Workshop on Digital Watermarking* (IWDW 2008), Nov 10-12, 2008, pp. 327-341

[35] Little, T. D. C., "A Framework for Synchronous Delivery of Time-Dependent Multimedia Data," *Multimedia Systems*, Vol. 1, No. 2, 1993, pp. 87-94

[36] Raghavendra, K. and Chetan, K.R., "A Blind and Robust Watermarking Scheme with Scrambled Watermark for Video Authentication," in *Proc. IEEE International Conference on Internet Multimedia Services Architecture and Applications* (IMSAA 2009), Dec 9-11, 2009, pp. 1-6. DOI: 10.1109/IMSAA.2009.5439475

[37] Lu, W., Sun, W., and Lu, H., "Novel robust image watermarking based on subsampling and DWT," *Multimed Tools Appl*, Vol. 60, Issue 1, September 2012, pp. 31-46. DOI: 10.1007/s11042-011-0794-1

[38] Doerr, G. and Dugelay, J. L., "A guide tour of video watermarking," *Signal Processing: Image Communication* 18 (2003), pp. 263–282

[39] Kim, K. S., Lee, H. Y., Im, D. H., and Lee, H. K., "Practical, Real-Time, and Robust Watermarking on the Spatial Domain for High-Definition Video Contents," *IEICE Trans. INE & SYST*, Vol. E91-D, No. 5, May 2008, pp. 1359-1368

[40] Yang, L. and Guo, Z., "A Robust Video Watermarking Scheme via Temporal Segmentation and Middle Frequency Component Adaptive Modification," in *Proc. 5 International Workshop on Digital Watermarking* (IWDW 2006), LNCS Vol. 4283, Nov 8-10, 2006, pp. 150–161. DOI: 10.1007/11922841_13

[41] Muhammad, G., Hussain, M., and Bebis, G., "Passive copy move image forgery detection using undecimated dyadic wavelet transform," *Digital Investigation* 9 (2012), pp. 49–57

[42] Chetty, G., Biswas, M., and Singh, R., "Digital Video Tamper Detection Based on Multimodal fusion of Residue Features," in *Proc. 4th International Conference on Network and System Security* (NSS 2010), Sept 1-3, 2010, pp. 606-613. DOI: 10.1109/NSS.2010.8

[43] Sun, T., Wang, W., and Jiang, X., "Exposing Video Forgeries by Detecting MPEG Double Compression," in *Proc. IEEE International Conference on Acoustic Speech and Signal Processing* (ICASSP 2012), March 25-30, 2012, pp. 1389-1392. DOI: 10.1109/ICASSP.2012.6288150

[44] Wang, W. and Farid, H., "Exposing digital forgeries in video by detecting double MPEG compression," in *Proc. 8th ACM workshop on Multimedia and security* (MM&Sec '06), Sept 26-27, 2006, pp. 37-47

[45] Wang, W. and Farid, H., "Exposing digital forgeries in video by detecting double quantization," in *Proc. 11th ACM workshop on Multimedia and security* (MM&Sec '09), Sept 7-8, 2009, pp. 37-47. DOI: 10.1145/1597817.1597826

[46] Wolf, S., "A No Reference (NR) and Reduced Reference (RR) Metric for Detecting Dropped Video Frames," in *Fourth International Workshop on Video Processing and Quality Metrics for Consumer Electronics*, VPQM 2009, Jan 14-16, 2009

[47] Chetty, G., Singh, M., and White, M., "Blind Image Tamper Detection Based on Multimodal Fusion," in *Proc. 17th International Conference on Neural Information Processing: Models and Applications* (ICONIP 2010), Part II, LNCS 6444, Nov 2010, pp. 557-564

[48] Kancherla, K. and Mukkamala, S., "Novel Blind Video Forgery Detection Using Markov Models on Motion Residue," in *Proc. 4th Asian Conference on Intelligent Information and Database Systems* (ACIIDS'12), Vol. Part III, LNAI 7198, 2012, pp. 308-315

[49] Seshadrinathan, K., Soundararajan, R., Bovik, A. C., and Cormack, L. K., "Study of Subjective and Objective quality assessment of video," *IEEE Trans. Image Processing*, Vol. 19, No. 6, June 2010, pp. 1427-1441

[50] Moorthy, A. K., Seshadrinathan, K., Soundararajan, R., and Bovik, A. C., "Wireless Video Quality Assessment: A Study of Subjective Scores and Objective Algorithms," *IEEE Trans. Circuits and Systems for Video Technology*, Vol. 20, No. 4, April 2010, pp. 587-599

[51] Winkler, S. and Mohandas, P., "The Evolution of Video Quality Measurement: From PSNR to Hybrid Metrics," *IEEE Trans. Broadcasting*, Vol. 54, No. 3, Sept 2008, pp. 1-9

[52] Wang, Z., Sheikh, H. R., and Bovik, A. C., "Objective Video Quality Assessment," Chapter 41 in *The Handbook of Video Databases: Design and Applications*, B. Furht and O. Marqure, ed., CRC Press, September 2003, pp. 1041-1078

[53] Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P., "Image Quality Assesment: From Error Visibility to Structural Similarity," *IEEE Trans. Image Processing*, Vol. 13, No. 4, April 2004, pp 1-14

[54] Wang, Z., Lu, L., and Bovik, A. C., "Video Quality Assessment Based on Structural Distortion Measurement," *Signal Processing: Image Communication*, Vol. 19, No. 2, Feb. 2004, pp. 121-132

[55] Chikkerur, S., Sundaram, V., Reisslein, M., and Karam, L. J., "Objective Video Quality Assessment Methods: A Classification, Review, and Performance Comparison," *IEEE Trans. Broadcasting*, Vol. 57, No. 2, June 2011, pp. 165-182

[56] Hands, D. S., "A Basic Multimedia Quality Model," *IEEE Trans. Multimedia*, Vol. 6, No. 6, Dec. 2004, pp. 806-816

[57] Zhai, G., Cai, J., Lin, W., Yang, X., Zhang, W., and Etoh, M., "Cross-Dimensional Perceptual Quality Assessment for Low Bit-Rate Videos," *IEEE Trans. Multimedia*, Vol. 10, No. 7, Nov. 2008, pp 1316-1324

[58] Sun, H. M. and Huang, Y. K., "The difference between perceived video quality and objective video quality," J Vis (2010) 13, pp. 159–168. DOI: 10.1007/s12650-009-0013-6

[59] Yim, C. and Bovik, A. C., "Evaluation of temporal variation of video quality in packet loss networks," *Signal Processing: Image Communication* 26 (2011) 24–38

[60] Pinson, M. H. and Wolf, S., "A New Standardized Method for Objectively Measuring Video Quality," *IEEE Trans. Broadcasting*, Vol. 50, No. 3, September 2004, pp 312-322

[61] Gulliver, S. R. and Ghinea, G., "The Perceptual and Attentive Impact of Delay and Jitter in Multimedia Delivery," *IEEE Trans. Broadcasting*, Vol. 53, No. 2, June 2007, pp. 449-458

[62] Seshadrinathan, K. and Bovik, A. C., "Motion tuned spatio-temporal quality assessment of natural videos," *IEEE Trans. Image Processing*, Vol. 19, No. 2, Feb 2010, pp. 335-350

[63] Pinson, M. H., Wolf, S., and Cermak, G., "HDTV Subjective Quality of H.264 vs. MPEG-2, With and Without Packet Loss," *IEEE Trans. Broadcasting*, Vol. 56, No. 1, March 2010, pp. 86-91

Author's Publications

[1] Thakur, M.K., Saxena, V., Gupta, J. P., "Data-parallel full reference algorithm for dropped frame identification in uncompressed video using genetic algorithm," in *Proc. 6th International Conference on Contemporary Computing* (IC3 2013), August 8-10, 2013, pp. 467-471. [**Indexed in DBLP**]

[2] Thakur, M.K., Saxena, V., Gupta, J. P.," Performance Analysis of PSNR and SSIM Against Frame Drop and Its Subjective Score", *International Journal of Digital Content Technology and its Applications*, Vol. 7, No. 3, pp. 679-688, February 2013. [**Indexed in SCOPUS, DBLP**]

[3] Thakur, M.K., Saxena, V., Gupta, J. P., "A hybrid video quality metric for analyzing quality degradation due to frame drop," *International Journal of Computer Science Issues*, Vol. 9, Issue 6, No 1, pp.78-84, November 2012. ISSN (Online): 1694-0814. [**Indexed in DBLP**]

[4] Thakur, M.K., Saxena, V., Gupta, J. P.," A Full Reference Algorithm for Dropped Frames Identification in Uncompressed Video Using Genetic Algorithm", *International Journal of Digital Content Technology and its Applications*, Vol. 6, No. 20, November 2012, pp. 562-573. [**Indexed in SCOPUS, DBLP**]

[5] Thakur, M.K., Saxena, V., Gupta, J. P., "A performance analysis of objective video quality metrics for digital video watermarking," in *Proc. 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT 2010)*, Vol.4, July 9-11, 2010, pp.12-17. [**Indexed in SCOPUS**]

[6] Thakur, M.K., Saxena, V., Gupta, J. P., "Video authentication against set of temporal tampering," **under review**

[7] Thakur, M.K., Saxena, V., Gupta, J. P., "Learning based No Reference algorithm for dropped frame identification in uncompressed video," **under writing phase**

[8] Thakur, M.K., Saxena, V., Gupta, J. P., "Video Tampering: A Review," **under writing phase**

[9] Thakur, M.K., Saxena, V., Gupta, J. P., "A dynamic dual threshold scheme for scene change detection," **under writing phase**

Thesis Supervisors

Dr. Vikas Saxena
Associate Professor
Jaypee Institute of Information Technology, Noida, India

Prof. J. P. Gupta
Professor & Vice Chancellor
Sharda University, Greater Noida, India