

# Content Identification using Video Fingerprint based on Video Classification

Kulbhushan Abasaheb Choure  
M.E CSE

Department of Computer Science & Engineerin  
TPCT's College of Engineering,Osmanabad  
A/P: Osmanabad Tal and Dist:Osmanabad  
413501  
India.

Anilkumar N. Holambe, Ph.D.

Department of Computer Science & Engineering  
TPCT's College of Engineering,Osmanabad  
A/P: Osmanabad Tal and Dist: Osmanabad  
413501  
India.

## ABSTRACT

Internet usage has drastically increased in recent years. Based on the topic a tool for video content identification and video classification is presented. Tool uses MPEG-7 standard. Thus video streams are easily available in various forms. One must appreciate text-based search engines for immediate access to such a vast multimedia data on web. But currently, only text-based search engines are obtainable. So a tool that will search for a unique content in the video not in the name of the video is proposed. Tool will make use of video fingerprint. Video Fingerprint is a unique identity of a video most commonly called as video signature. The video signature has numerous applications like distribution management, corporate or personal database management, rights management and monetization, metadata association, and usage monitoring. We will be making use of this Video content identification tool for implementation of Topic linking concept, a method of organizing contents based on topics, presentation videos and provides such efficient way of browsing, such as news programs. This in turn helps to categorize the video according to its possessed content and store and manage the multimedia data in an easy way.

## Keywords

Topiclinking,VideoContentIdentification,MPEG-7,  
Videofingerprint, Videoclassification

## 1. INTRODUCTION

In recent years, the amount of video consumption has increased to a great level. On website like YouTube, every day hundreds of thousand videos are uploaded and approximately millions of videos are watched on online. Such a enormous amount of multimedia data is available on internet and for retrieval of some specific video from such huge data we must thankful to text-based search engines. An amount of videos from distributed sources has grown to large number. Video streams are easily accessible in large number of forms such as news video broadcasts, blogs, and podcasting because of the detonation of Internet bandwidth and broadcast channels. On websites like YouTube, many times it happens that you enter the name of the preferred video, then you get a list of related videos and you start watching it. Once you start the video you find some unrelated content. That means the video contained dissimilar data than what you insisted.Such videos had just same name as that of your search query. Survey found that currently there are only text based search engines are obtainable which competitions only the name of the search query to the name of the data available in database. So in this paper a tool is presented which will organize and manage the huge multimedia data according to its possessed content. To achieve this aim a tool is developed that will identify the specific content in the

video as per search query. This whole system will improve the performance of the search engine. The video content identification mainly has two main tasks, first one is development of unique and robust video fingerprint and second one is finding out specific content set in to an unrelated video. So this paper mainly aim at developing the complete retrieval system that will extract the video fingerprint, identify the specific content in the video and lastly will store and manage the all multimedia data according to its content.

## 2. MOTIVATION

Videos on commercial websites like YouTube, Metacafe are mainly textually tagged. Many times it is found that name of the video and actual content of the video has a vast difference. An efficient method of consuming such vast amounts of videos is essential. The proposed system implements the tool that identifies the content of the video and the concept of topic linking. Topic linking is a method of classifying contents based on topics, viewing videos and provides such efficient way of browsing, such as news programs. This leads to increased efficiency of system management, search engine etc. As the broadcasting channels and the Internet bandwidth has increased, videos are easily available in various instantly accessible forms like blogs, podcasting and news channel. It is very common that news channels explode with stories of the same topic when some critical issue is raised/happens from professional news or amateur or modified videos. Furthermore rapid progress in technology had made possible editing and duplication of the videos which may break the digital rights. Thus copyright protection has become critical issue because of vast amount of multimedia data. This made copyright detection as an active area of research. The video signature is a unique identity of a video. No two videos can have same video fingerprint. The video fingerprint is fundamental element for copyright protection. The video signature has a large number of potential applications like monetization, Distribution management and Rights management;video content based linking, Database management and reduplication, usage monitoring,.

## 3. RELATED WORK

Here, a key design choice in creating a video fingerprint and then study a number of notable studies in the field, both for localization and identification of a video entrenched in unrelatedvideo content and for near-duplicate video clip detection aredescribed.

In [3] this paper Hampaur et al. Studied detail and used

1. Block Based Ordinal Signature
2. Use of rank vector as the frame feature.

In analysis, it

- 1) Gave precise only for shorter queries
- 2) Not useful for detection of longer video content.

Detailed study of [4] this paper found that Law-To et al. used

1. Block Based Temporal Ordinal Signature
2. Signature is based on key-point temporal trajectory properties.

Evaluation showed that the Signature proved to be weak for authentication attacks.

Lastly detailed study of [5]. H. T. Shen, X. Zhou, Z. Huang, J. Shao, and X. Zhou, UQLIPS: A real-time near-duplicate video clip detection system, in Proc. 33rd Int. Conf. Very Large Data Bases, Sep. 2007, pp. 1374-1377. This paper found that it has designed near-duplicate video detection system. 2. Used two schemes BCS and ROC. Since the system is designed to match entire short clips, it does not address the issue of detecting and localizing a video segment inside a longer unrelated video.

In [1], near video-duplicate detection system is stated. In this system a dense spatiotemporal feature is used to represent a video. This compact spatiotemporal feature uses grey-level intensity distributions within a frame and temporal structure along the frame sequence. But this system failed to detect a video segment in a longer video segment.

Video fingerprint States in [2]. Video fingerprint is a unique identification mark for any video. Properties like Robustness to Fast matching, editing operations, Accurate localizing the embedded and duplicated content of video, Compactness for visual features, Fast extraction video into frames, No alteration of the content/data, Uniqueness in the result for detecting the duplicated video, Self-authority of the signatures, Independence on the content, codes, Partial matching must be satisfied by the video fingerprint.

For generating the fingerprints researchers had attempted a broad range of frame level features. Color histogram [9, 10] is widely used application. But, it lacks in terms of distinguishing as the spatial distribution of color is not held in the histogram.

Binary signature based on color histogram has been deployed in [6]. Luminance based descriptors [7], dominant color [12], gradient based features [13] and texture feature [11] are also attempted.

In [15], MPEG-7 video content identification tool has explained. It is the tool mainly useful for the broadcast channels, TV channels where the quality of videos is vary from sources to sources. MPEG-1, -2, -3 are the developed standards of MPEG.

- MPEG-1: It only specifies bitstream syntax and decoding process. It is used for the moving objects.
- MPEG-2: It's the genetic coding used for digital televisions in broad casting the channels. It also specifies the models and format that are compatible for DVD or other playing media.
- MPEG-4: Codes of MPEG-4 are used for audio and visual features of videos that are related for some format and technology.

The main function of MPEG-1, -2, and -4 standards is checks the availability of content in video where as MPEG-7 let one to find the content one is searching for video content with visual features.

MPEG-7: It gives the visual features for the users in terms of Images, Videos, Graphics and Video Activity. For content identification, [15] used three stage system

1. Signature extraction
2. Signature Compression
3. Matching.

### 3.1 Signature Extraction

A video signature is made up of two parts

1. fine signatures
2. Coarse signatures.

1. **Fine signature:** Fine signature is extracted from each individual frame. It is extracted from 32x32 pixel 8-bit luminance information. It is obtained by block averaging of luminance channel of a frame. It contains 3 sub parts
  - a) Set of local features. It is called as frame signature.
  - b) Words: It's a small representative subset of local features.
  - c) A global frame "Confidence".

Local feature are the calculations of local average intensities and differences at various levels of granularity in the frame. Each local feature is element of Frame Signature. Frame signature has 380 (19x20 matrix) elements out of which 32 are averages and 348 are differences. Averages and differences are subjected to the categorization. Categorization is important for ternarization. Ternarization gives improved result than binarization. A global frame confidence is 8 bit integer calculated from 348 differences categories D1 and D8. Fine signature requires 656 bit storage.

2. **Coarse Signature:** Coarse Signature is extracted from set of fine signatures. To obtain coarse signature fine signature is mandatory. The bag-of-words representation is extracted for temporal segments of 90 consecutive frames. In extraction of coarse signature five histograms are generated as each fine signature contains five words. Then each histogram is binarized.

This gives rise to five binary occurrence histograms. Each coarse signature needs 1215 bits for storage. In a video signature, all the coarse signatures first, followed by the fine signatures. This arrangement of a video signature gives the efficient bitstream access in sequence matching.

### 3.2 Video Signature Compression

Compression is basically applied to remove the redundancy. Temporal redundancy is present in fine signature to remove this redundancy lossless compression scheme is applied. Aligned with the coarse signature temporal segments, 45 consecutive frame signatures are subjected to compression. Each compressed temporal segment has key picture (KP), predicted picture (PP), Group of pictures (GOP).

### 3.3 Matching

Matching is performed in three stages. Output of this stage basically gives result that does the search query video matches any video in database. This stage is mainly identifies the video

content.

Stage 1: to recognize candidate matching signature, matching Coarse Signature uses coarse signature. It mostly measures the distance of the two functions of the word those are common.

Stage 2: to identify candidate parameters of frame rate ratio and temporal offset between the candidates matching by using fine signature.

Stage 3: In this stage frame by frame matching is done candidate matching intervals.

#### 4. APPLICATION

The video signature has a large number of applications like

Rights management and monetization: video fingerprint helpsto identify who is real owner.

Distribution management: Using video signature one can stopthe leakage/distribution of some sensitive data over a network or via email by coding video identification tool at firewall.

Usage monitoring: It helps to maintain statistics For example;an advertizing agent can check that does its material have distributed as expected using video fingerprint.

Video content-based linking: Video content in the web page can be linked to some other video.

Database management and deduplication: Mainly videocontent identification application is useful for large volume of data creators and owners like personal libraries, studios.

#### 5. PROPOSED WORK

Video classification using videocontent identification is the main purpose of this paper. Video content identification will be same as that of [15]. In this paper, for video content identification only fine signature will be used and with the spread of digital video recorders, storing large volumes of video has become recorders, storing large volumes of video has become simple. An economical method of overwhelming such vast amounts of videos is essential. Topic pursuit or topic linking is a method of organizing contents based on topics; provide such efficient approach of browsing and viewing videos, such as news programs. To observe visual duplicates between completely dissimilar contents that are together to form logical topics team, these techniques use visual options of videos. It utilizes the actual fact that the scenes depiction identical topic recurrently uses identical video segments from a typical supply even across totally different broadcast channels.

The key to topic linking is that the robust detection of visual duplicates from a similar supply, even with dissimilarities caused by modifications like caption overlay and compression, which are compulsory during the editing or storage of contents. For instance, totally dissimilar captions are overlaid at different areas on a similar video source for different programs or broadcast channels.

In this paper, we propose a topic linking for video archives using a robust video signature (video fingerprint) developed by the authors. The proposed video signature, which was accepted as part of a new ISO/IEC standard "ISO/IEC 15938-3/Amd 4MPEG-7 Video Signature Tools, is designed to be robust to a variety of alterations, and is suitable for duplicate detection of short segments in a large database. And also present our topic linking system which visually presents the produced topic groups to the users for proficient browsing and viewing of contents based on topics. Capitalize only the first word in a

paper name, apart from for proper nouns and element symbols. For papers published in translation journals, satisfy provide the English citation first, after that by the original foreign-language citation [10].

#### 5.1 Design

The implemented architecture of topic linking system is shown in fig. We implemented it in a matlab. This system records incoming broadcast videos and stores them in a hard disk. At the same time, the video signatures of the video are extracted simultaneously with the recording.

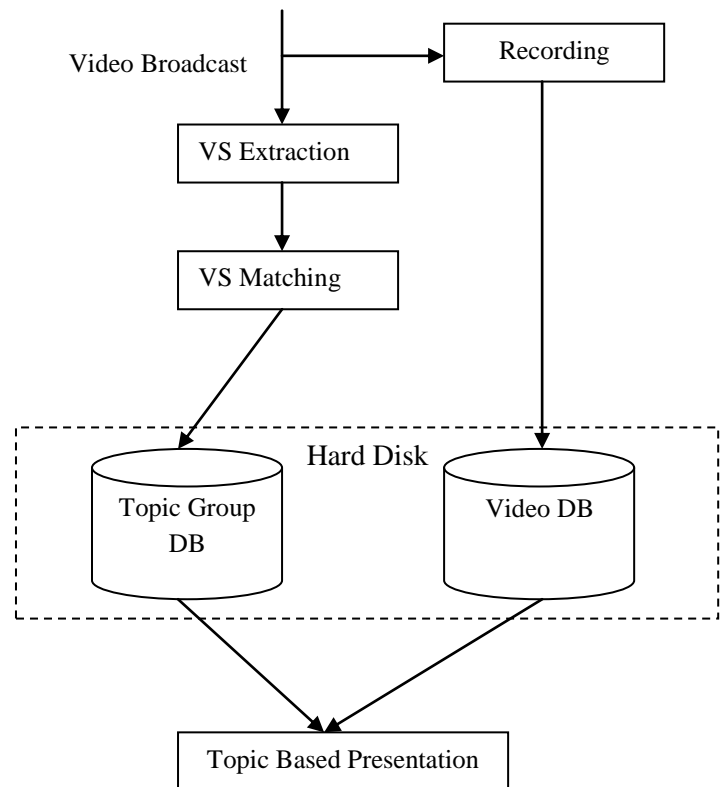


Figure: Topic linking Architecture

After extraction, video signature matching is carried out to detect duplicate segments between the newly recorded video and the videos already stored in the database. The duplicate segments are grouped together into topics, and the topic group information is stored in a hard disk.

#### 5.2 Mathematical Model

1. Frame features comprises of ternarized elements, i.e., quantized to three levels.

$$P = \begin{cases} 2, & \text{if } m > th \\ 1, & \text{if } m \leq th \\ 0, & \text{if } m < -th \end{cases}$$

th : threshold value for ternarization

m : Frame signature element.

p : a ternarized element.

2. Video Signature

$$sig = \sum_{k=0}^n average(x, y)$$

sig : stores the fine signature of a frame.

*average* : is a function which will calculate average the pixels around the current one.

$x, y$ : are co-ordinates of center pixel.

3. Distance between two frames

$$d = \sum_{x=20,y=19}^{x=0,y=0} \sqrt{(r1_{x,y} - r2_{x,y})^2 + (g1_{x,y} - g2_{x,y})^2 + (b1_{x,y} - b2_{x,y})^2}$$

Where,

$r1, r2$  are red pixels of 1st image and 2nd image respectively.

$g1, g2$  are green pixels of 1st and 2nd images respectively.

$b1, b2$  are blue pixels of 1st and 2nd images respectively.

## 6. RESULT

### 6.1 Dataset

We used the various types of footage of motion pictures as a dataset for experiments and result,. The footage we used is such as films, news, documentary, cartoons, sport, home video, and so on.

- Input*: We give a video as an input to the tool. We have a database that contains large number of videos. Database is used for matching purpose.
- Output*: The main output of the application will be putting query video under the section according to its content. Number of frames of input video with the specified regular interval, output of the matching process of query frame with the available in the database, name of the section that query frame belongs are the intermediate outputs.

### 6.2 Query Condition

The designed tools were classified into two query condition. The two tools are direct content matching and partial content matching. In the direct matching the content is matching directly, i.e. the whole query clip and a part of the original are matches with each other. In the partial content matching, as the name point out only the part of original and query clip are matched. For assessing performance of the application at dissimilar temporal granularities there are three query situations of each query which we are studied above.

### 6.3 Efficiency

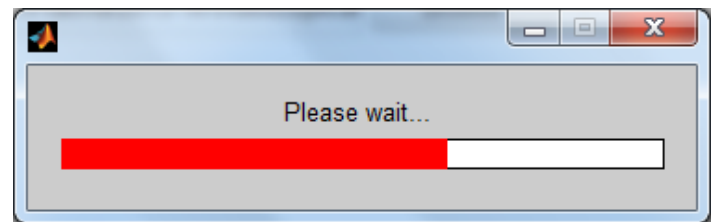
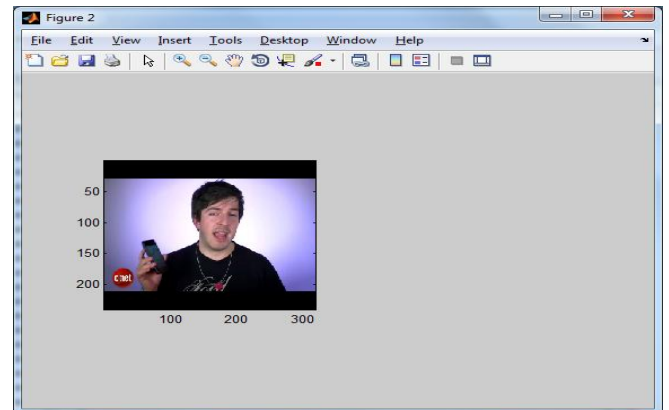
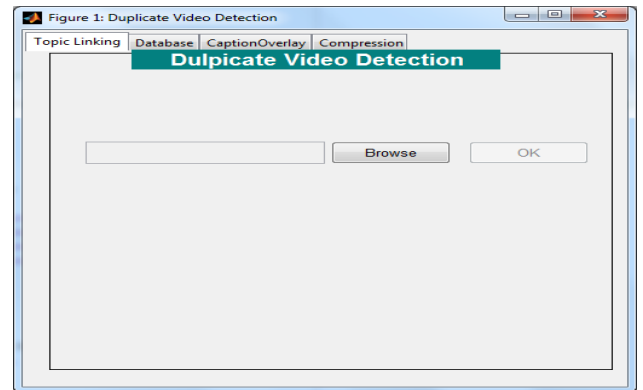
In Table I, we have considered a single video, after this we clutched the frames for dissimilar intervals as mentioned above. Number of frames is the column in which we have counted the frames extracted from input video. Number of frames matched is the column in which we counted the number of frames matched with the database frames. While in average success rate there is a ratio in % of the frames we have grabbed from the video and the frames matched.

**Table I**

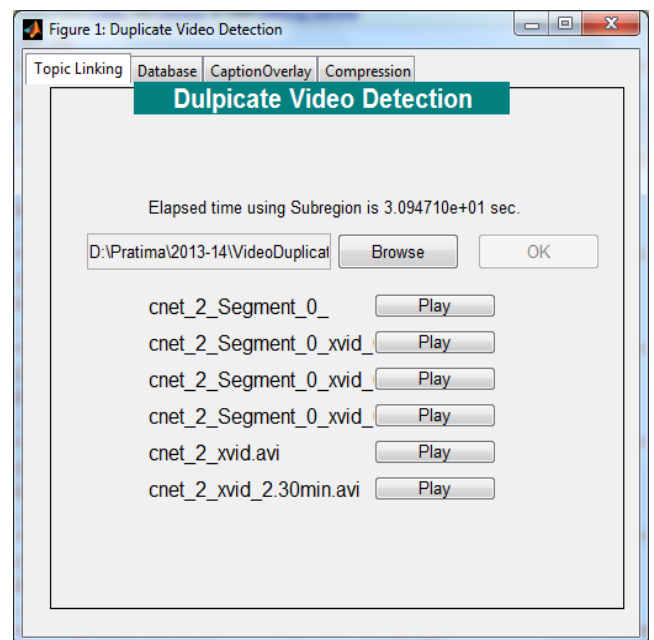
Matching Interval	No. of Frames	No. of frames matched	Average success Rate (%)
14s	7	5	78%
9s	15	12	84%
4s	24	20	92.64%
2s	36	34	98.47%

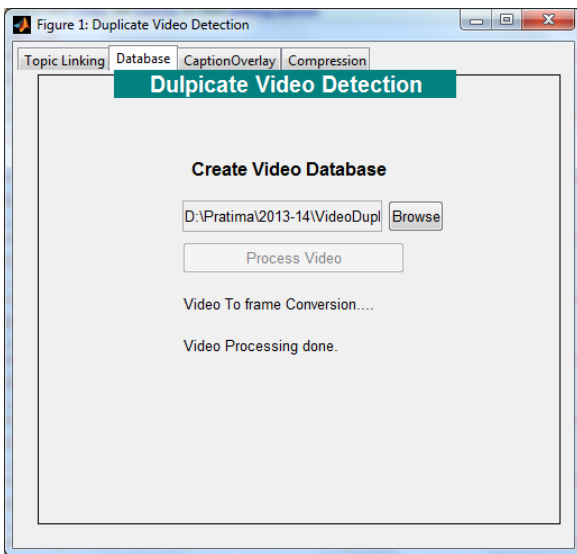
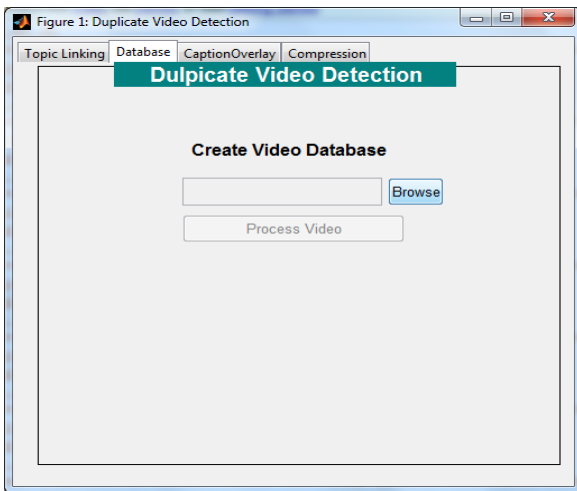
## 6.4 Result

### a. Topic Linking:

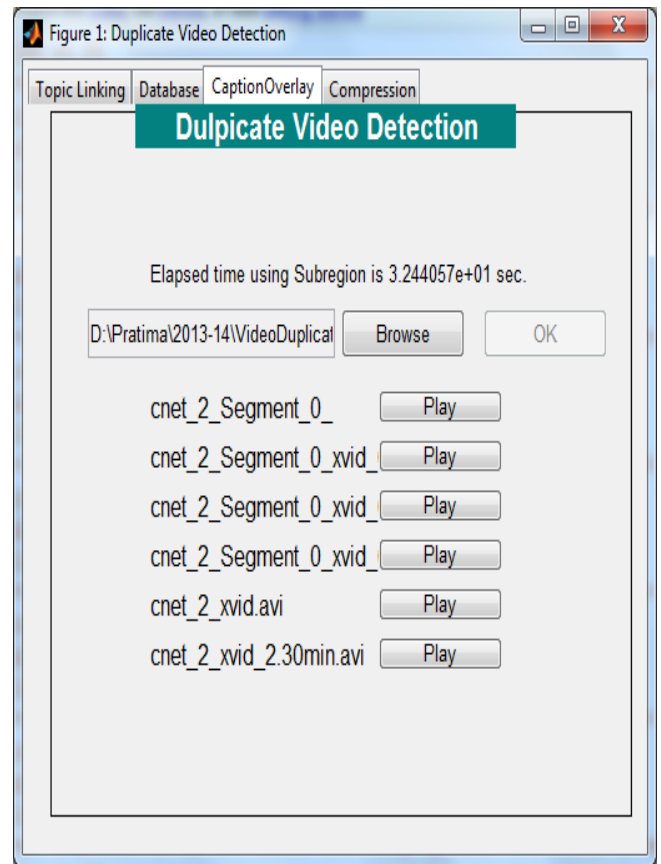
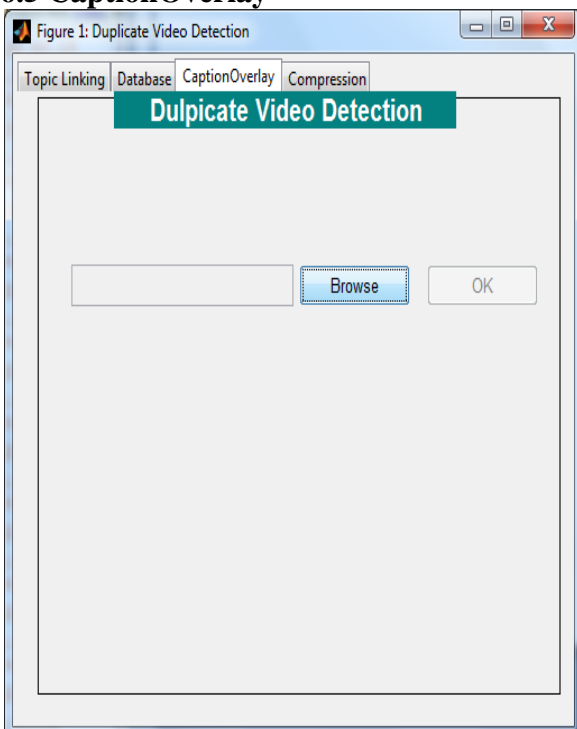


### b. Browse Database

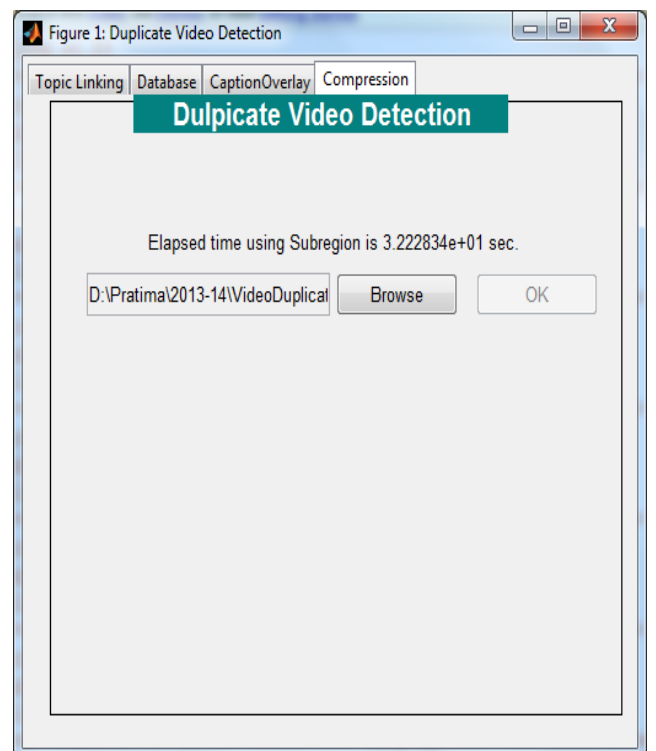




### 6.5 CaptionOverlay



### 6.6 Compression



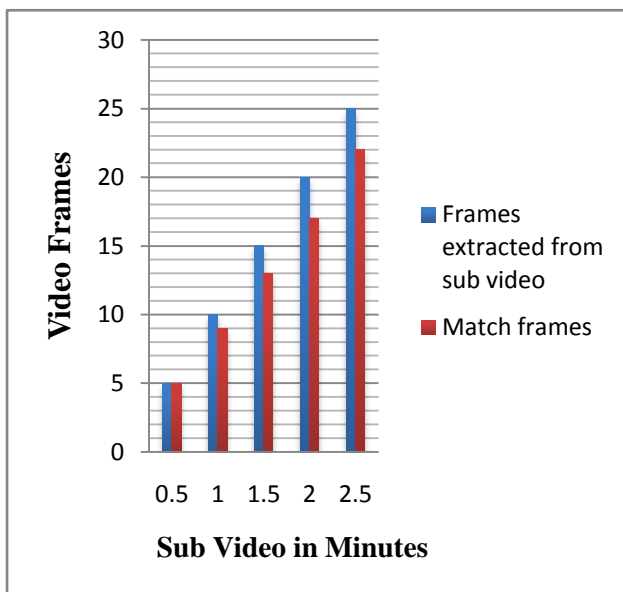
#### **Classification:**

Classification results will display the name of the topic that queries video contents. Then the video will be linked to accordingly to its contented topic's section.

## 6.7 Comparison

In our experiment, we analyzed our proposed method by giving input video to the system. First we get total number of frame from the input video (25 frames). We collect some sub videos of length ranging from 0.30sec to 2.30min. We have calculated the Total frames generation for this sub videos as shown in Table1. Initially, for first 30 sec sub video, 5 frames are generated and out of which 5 are matched with total frames generated for the input video. For next 1.00min sub video, 10 frames are generated and 9 frames are matched. At last interval for 2.30min we get 25 frames and according to which 22 are matched which give 88% accurate results. Therefore, the proposed method is efficient and a scalable for achieving topic linking of video archives.

Sub video in minutes	Total Frames extracted from sub video	Match frames with original video key frames
0.30	5	5
1.00	10	9
1.30	15	13
2.00	20	17
2.30	25	22



Graph1. Comparison between frames extracted from sub video and match frames in minutes.

## 7. CONCLUSION AND FUTURE WORK

The study gives brief idea about various works done by researchers in field of video processing and video protection. The used video signature will attain high level of robustness to common video editing operations and content identification will be accurately able to detect and locate a piece of video content embedded in a longer piece of unrelated video content are concluded by the researchers. Implementation of topic linking concept using video signature will prove to be effective to detect visual duplicates for grouping coherent

topics in a video archive. We implement this paper in matlab. A robust system for efficient browsing and viewing of videos using the detected topic links proved by Topic linking system. In the caption overlay and compression we are calculating the estimated time that required for time complexity calculation. By comparing the frames that are comes from video we set the threshold value as a benchmark for detection of duplicate video. Out of 380 frames 300 same frames shows the duplication after comparison. For every point of view from topic linking, caption overly and compression Key frame always plays important role in video duplication.

## 8. REFERENCES

- [1] L. Shang, L. Yang, F. Wang, K.-P. Chan, and X.-S. Hua, "Real-time largescale near-duplicate web video retrieval," in Proc. ACM Int. Conf. Multimedia, Oct. 2010, pp. 531–540.
- [2] Content Based Video Copy Detection: Issues and Practices. Sanjoy Kumar Saha, CSE Department, Jadavpur University, Kolkata, India.
- [3] Hampapur, K. Hyun, and R. Bolle, "Comparison of sequence matching techniques for video copy detection," in Proc. Conf. Storage Retrieval Media Databases, 2002, pp. 194–201.
- [4] J. Law-To, L. Chen, A. Joly, I. Laptev, O. Buisson, V. Gouet-Brunet, N. Boujemaa, and F. Stentiford, "Video copy detection: A comparative study," in Proc. 6th ACM Int. Conf. Image Video Retrieval, Jul. 2007,
- [5] H. T. Shen, X. Zhou, Z. Huang, J. Shao, and X. Zhou, "UQLIPS: A Real-time near-duplicate video clip detection system," in Proc. 33rd Int. Conf. Very Large Data Bases, Sep. 2007, pp. 1374–1377.
- [6] Y. Li, J. S. Jin, and X. Zhou. Video matching using binary signature. In Proc. Intl. Symp. on Intelligent Signal Processing and Comm. Systems, pages 317 – 320, 2005.
- [7] R. Radhakrishnan and C. Bauer. Robust video fingerprints based on subspace embedding. In Proc. ICASSP, pages 2245–2248, 2008.
- [8] S. Lee and C. D. Yoo. Video fingerprinting based on centroids of gradient orientations. In Proc. ICASSP, pages 401 – 404, 2006.
- [9] S.-C. S. Cheung and A. Zakhor. Efficient video similarity measurement with video signature. IEEE Trans. CSVT, 13 No. 1:59–74, 2003
- [10] M. Ferman, A. M. Tekalp, and R. Mehrotra. Robust color histogram descriptors for video segment retrieval and identification. IEEE Trans. on IP, 11(5):497 – 508, 2002.
- [11] L. Chen and T. S. Chua. A match and tiling approach to content-based video retrieval. In Proc. Intl. Conf. on Multimedia and Expo, 2001.
- [12] International Organization for Standardization Organization Internationale De Normalisation Iso/Iec Jtc1/Sc29/Wg11 Coding Of Moving Pictures And Audio.
- [13] S. Paschalakis, K. Iwamoto, P. Brasnett, N. Sprljan, R. Oami, T. Nomura, A. Yamada, and M. Bober, "The MPEG-7 Video Signature Tools for Content Identification".