Asian Journal of Information Technology 13 (3): 112-118, 2014
ISSN: 1682-3915

Mining Semantic Patterns in Cricket Videos Using Fuzzy Based Classification

[1]D. Pushpa Ranjini and [2]D. Manimegalai
[1]Department of IT, PET Engineering College, Vallioor, India
[2]Department of IT, National Engineering College, Kovilpatti, India

Abstract: Video is an information-intensive media with much redundancy. Therefore, it is desirable to mine structure or semantics of video data for efficient browsing, indexing and highlight extraction. A naive user is interested in querying at the semantic level, rather than having to use features to describe his (her) concept. In most cases it is difficult to express concepts using feature matching and even a good match in terms of feature metrics may yield poor query results for the user. Researchers propose a novel video indexing using semantic pattern. First video preprocessing techniques are applied for processing the video to get audio and video cues. Then mining technique is applied to classify the events based on the semantic patterns. The task to discover useful and interesting patterns from a video is the main goal of video data mining. Data mining on multimedia is a challenging task, due to the complicated contents of multimedia data. Particularly the discovered patterns need to be supported by their semantic features. However, low-level features often have little meaning for naive users, who much prefer to identify content using high level semantics or concepts. This creates a gap between systems and their users that must be bridged for these systems to be used effectively. So, researchers propose to develop a knowledge-based video indexing using fuzzy based classification for mining semantic patterns.

Key words: Information-intensive, semantic, browsing, indexing, data mining

INTRODUCTION

Conventional relational databases are capable of storing and manipulating large volume of text data. However, current application requires facilities to store and retrieve text document, image and video data effectively.

In particular, digital video plays an important role in entertainment, education and other multimedia applications. Hence, it has become increasingly important to develop mechanisms that process, filter, search and organize the digital video information that is now accessible. Since, it is difficult to index and categorize video data automatically compared with similar operations on text, search engines for video data are still rare. Content-based video database modeling, representation, summarization, indexing, retrieving, navigating and browsing have emerged as challenging and important areas in computer vision and database management.

It has never been an easy operation to extract or explore knowledge from data (Zhang *et al*., 1993; Zhu and Wu, 2003a, b; Pan and Faloutsos, 2002; Xie *et al*., 2003). Recently, many data mining techniques (Fan *et al*., 2003, 2001; Snoek and Worring, 2005; Jiang and Elmagaramid, 1998; Zhu and Wu, 2003a, b; Han and Kamber, 2006) in exploring knowledge from large video data have been reported. Many approaches proposed for video mining are classified in to three categories:

- Special pattern detection (Fan *et al*., 2003) which detects special patterns as video events
- Video clustering and classification (Jiang and Elmagaramid, 1998; Zhu and Wu, 2003a, b) which clusters and classifies video units into groups
- Video association mining (Fan *et al*., 2001; Snoek and Worring, 2005) which explores video knowledge by association from video units

The solution for video data mining is applying the existing data mining techniques (Fan *et al*., 2000). The video and image databases are in different forms so, can not be able to figure out. Because, researchers may retrieve video frames from the video and acquire relationships from those video frames. Mining knowledge from the video database is harder compared to the general databases.

In this study, researchers use a knowledge based video indexing to access the video database. To explore video knowledge, researchers use the video preprocessing techniques to extract the information from

Corresponding Author: D. Pushpa Ranjini, Department of IT, PET Engineering College, Vallioor, India

the video data. Then applying video association mining that is the existing algorithms are integrated to mine video knowledge. Multiple information are extracted from the videos and form a data stream and the stream is processed for finding the sequential association. Here researchers use Cricket videos as the test bed.

NEED FOR VIDEO PREPROCESSING

A naive user is interested in querying at the semantic level rather than having to use features to describe his concept. In most cases, it is difficult to express concepts using feature matching and even a good match in terms of feature metrics may yield poor query results for the user. Management of digital video documents is becoming more problematic due to the ever-growing size of content produced. Video indexing has the advantage that it can profit from combined analysis of visual, auditory and textual information sources. Applications such as digital libraries, video-on-demand systems and interactive video applications introduce new challenges in managing large collections of audiovisual content. To help users find and retrieve relevant video more effectively and to facilitate new and better ways of entertainment, advanced technologies must be developed for indexing, filtering, searching and mining the vast amount of videos. Motivated by these demands, many video research efforts have been made on exploring more efficient content management systems. A simple framework is to partition continuous frames into discrete physical shots and extract low-level features from video shots to support activities like searching, indexing or retrieval. For the prior step of indexing, filtering, searching and mining the vast amount of video to explore video knowledge and to store it in the database.

In this research, first researchers perform the video preprocessing in order to implement data mining activities effectively.

PROPOSED VIDEO DATA MINING

Data mining is a technique to discover unknown and interesting patterns called semantic patterns from a vast quantity of data. Although, the data mining technique can be applied in various domains here researchers concentrate on video data. With the increasing amount of video data, video data mining plays an important role in efficient video data management. Along with the vast amount of data, a rapid method to mine semantic patterns is needed. Current content management system uses the retrieval for the low level feature. So, for naive users it is tedious to handle that. If retrieval is handled by semantic level it is very easy to handle for naive users also.

Since, the 1990s, data mining has been a very active area for research and applications. Many successful techniques have been implemented through academic research and industrial applications. However, these approaches deal with various databases (like transaction datasets) in which the relationship between data items is explicitly given. Video and image databases are different from these databases. The most distinct feature of video and image databases is that the relationship between any two of their items cannot be explicitly (or precisely) figured out. This inherent complexity of the multimedia data has suggested that mining knowledge from multimedia materials is even harder than from general databases.

In this study, researchers first used a knowledge-based video indexing framework to facilitate video database management and access. To explore video knowledge from this framework, researchers have implemented video association mining in which video processing and existing data mining algorithms are seamlessly integrated to mine video knowledge. Researchers have considered the definitions and evaluation measures (temporal distance, temporal support and confidence) for video associations by taking the distinct features of video data into consideration and then implemented them for mining sequential patterns from the video stream that usually consists of multiple information sources (e.g., image, audio and caption text). For indexing video researchers have performed the following operations:

- Video preprocessing
- Association mining
- Classification

General procedure:

Input: Video V
Output: Indexed video
  Video preprocessing
    a) Video shot segmentation and key-frame extraction
    b) Video shot clustering
    c) Camera motion extraction
    d) Text detection
    e) Audio event detection
Association mining from sequence D
Search interesting pattern
Classify the events using Fuzzy Based KNN

Figure 1 shows the overview of the system. The raw video sequences are processed through preprocessing techniques and association mining algorithms are applied and finally indexed video sequences are received. Figure 2 shows the overall procedure needed for video mining process. It covers two phases. The first phase deal with the preprocessing techniques.

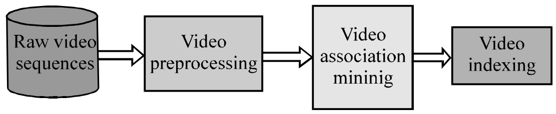

Fig. 1: Overall view

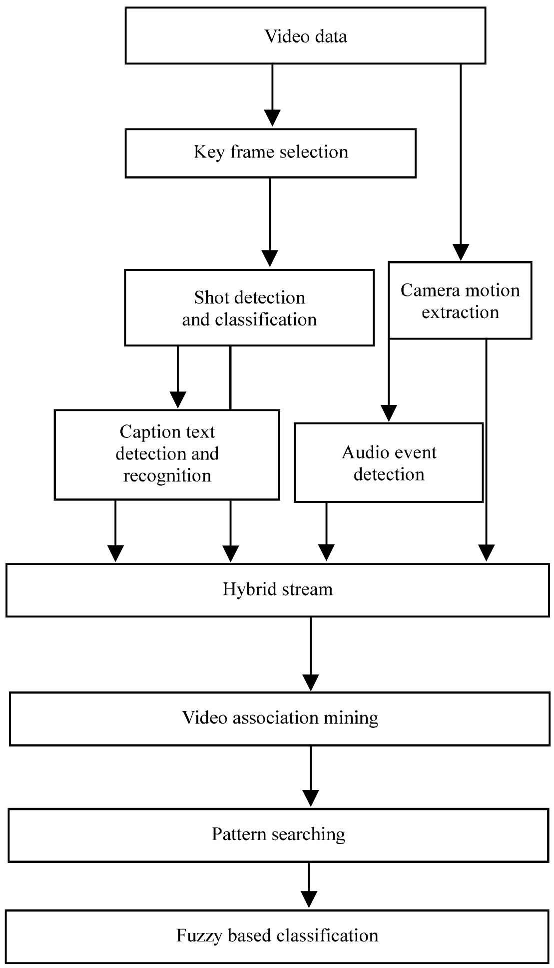

Fig. 2: Proposed system architecture

Video pre-processing: Video data sequences are stored in a data store. This is given as input for video preprocessing. If researchers are using a video file except MPEG format then the file should be converted in to MPEG. Many sub processes are used to perform video preprocessing. Initially the given video data file is converted into 'n' number of frames. Most schemes for video feature extraction begin by segmenting contiguous frames into separate shots and then selecting key-frames to represent shot content. With this scheme, a video database is treated somewhat like an image database, because the motion information in the video (or shot) is missed.

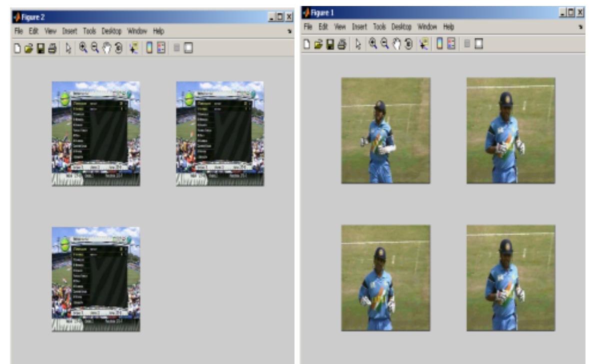

Fig. 3: Classified shots

Video shots which are directly related to video structures and contents are the basic units used for accessing video sources. An automatic shot detection technique has been proposed for adaptive video coding applications (Fan *et al.*, 2000; Kolekar and Sengupta, 2010). However, in this research researchers focus on video shot detection on compressed MPEG videos. Since, there are three frame types I, P and B in a MPEG bit stream (Kolekar and Sengupta, 2010) this technique is to used to detect the scene cuts occurring on I frames and the shot boundaries obtained on the I frames are then refined by detecting the scene cuts occurring on P and B frames. In the system, the motion information in the video is detected and extracted as a shot feature to help in identifying video content. Researchers first apply shot segmentation to the video. Researchers adopt a modified split-and-merge clustering algorithm by sequentially executing two major procedures: merging and splitting. Then researchers classify the shots in to two groups. In sports video any one of the color is dominant in the play field. So, seek the dominant color to identify the play field group and other colors are grouped as another group. Figure 3 shows the classification of shots of play field shots and score board shots.

Sports videos consist of two types of text. First is the text shown in the video scenes itself and is referred as scene text. Second is the text which is post processed and added into the video such as team names and their scores named as caption text. In sports video caption text is more important than scene text. Because scene text indicates some advertisement and etc. But caption text only denotes the team score and name of the team. So, efficient caption text algorithm is used to detect caption text. The current frame and succeeding frame edge pixels are compared to find the caption text pixel.

After detecting the candidate text regions, caption text is recognized using an existing OCR engine. It takes a binarized image as the input and yields an ASCII string.

Once it detects a score change, it adds symbolic tag at the corresponding place. Then the mpeg file is converted into non-relational data set and stored it in a text file for further processing. This data set is converted into relational data set. Motion characterization plays an important role in content-based video indexing. It is an essential step in creating compact video representation automatically. The image can represent a panning sequence and the frames before and after a zoom can represent the zoom sequence. So to find the zoom, pan, tilt and rotate representations of the frames. Also detect speaker excitation and crowd applause. Figure 4 shows the camera motion extraction of zoom, tilt, pan and rotation.

Video data transformation: From the video preprocessing it gives various individual streams. For performing association mining it is necessary to find out the most frequent item set. It is very tedious to seek the item set from multiple streams. So that researchers combined all stream into a single stream without affecting their original order. Hence, researchers are assigning the following tags to set the hybrid stream Fig. 5.

Fields:
a = Play field
b = Non play field

Camera motion:
c11 = Pan right slow
c12 = Pan right medium
c13 = Pan right fast
c21 = Pan left slow
c22 = Pan left medium
c23 = Pan left fast
e11 = Zoom in slow
e12 = Zoom in
e13 = Zoom in fast
e21 = Zoom out slow
e22 = Zoom out medium
e23 = Zoom out fast
f11 = Tilt up slow
f12 = Tilt up medium
f13 = Tilt up fast
f21 = Tilt down slow
f22 = Tilt down medium
f23 = Tilt down fast
o = Other type of camera motion

Audio cues:
h = Speaker excitation
g = Crowd applause

Text:
d = Scoreboard change

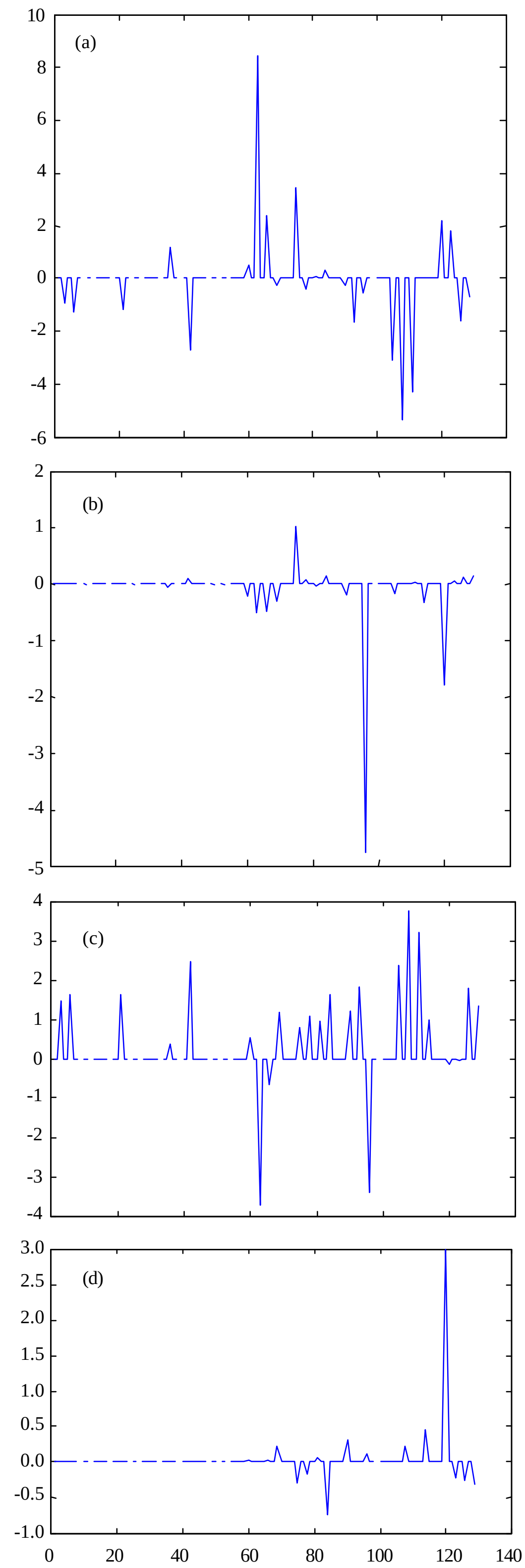

Fig. 4: Camera motion; a) pan, b) rotation c) zoom and d) tilt

Video association mining: Let J = $\{i_1, i_2, i_3, ..., i_m\}$ be a set of items. Let D, the task-relevant data, be a set of database transactions where each transaction T is a set of items such that T⊂J. Each transaction is associated with an identifier TID. Let A be set of Items. A Transaction T is said to contain A if and only if A⊂T. An association rule is an implication of the form A⊂B where A⊂J, B⊂J and A⊂B. The rule holds in the transaction set D with support s and confidence c. Rules that satisfy both a minimum support threshold and minimum confidence threshold are called strong. A set of item is referred as an itemset. An itemset that contains k item is a k-itemset. The occurrence frequency of an itemset is the number of transactions that contain the itemset. In this study, consider each item represents video item. From the hybrid stream researchers apply multilevel association mining and filter the dataset using Apriori algorithm to derive the frequent item sets. A Video database consists of enormous data. So, researchers have to find out the interesting patterns available in the hybrid stream.

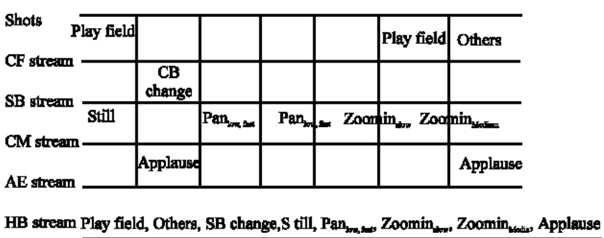

Fig. 5: Hybrid stream and tags

Hybrid stream
b b b a e21 b b b a c11 h b a d c23 c22 g
h b b b a d c11 b b b b a c11 a b a d c21

Filter data set
b b b a e2 b b b a c1 h b a d c2 c2 g h b
b b a d c1 b b b b a c1 a b a d c2

Fuzzy based classification: To apply video indexing, classify each association into a corresponding event and use detected events to construct video indices. Manually check the training videos to see what are the patterns indicate the respective event. Then apply fuzzy based KNN to classify those events Before that after checking the events, labeling the patterns with the respective event name. The test data set is having the respective labeled events. Based on the Euclidean distance the nearest neighbors are identified a group. Mine associations from training videos. For each association, manually go through the training data to evaluate what types of events associate with the appearance of this association. Count

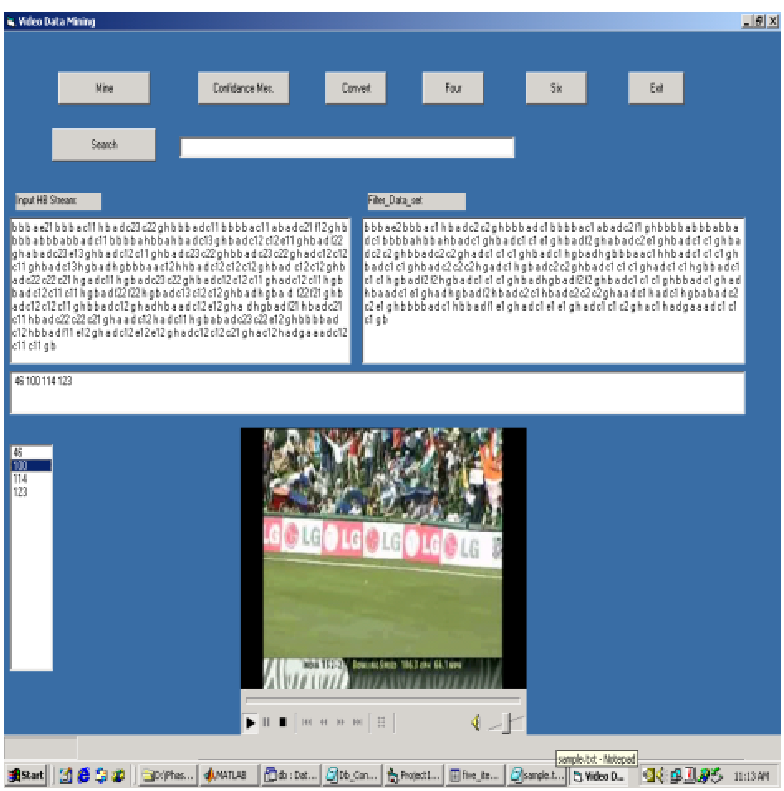

Fig. 6: Fuzzy based classification

the number and the types of events from all appearances and select the event with the largest number to label the association. Accordingly, each association will receive one class label. For each association in the test set, calculate its distance with associations in the training set and the class label of the association in the training set which has the smallest distance with {X} is used to lable {X}. In the case that multiple associations have the same smallest distance with {X}, all their class labels are used to label {X}. To calculate the distance between sequential associations, take the temporal order and the length of the associations into consideration and use the Longest Common Subsequence (LCS) between two associations to evaluate the association distances. Figure 6 shows the classified image.

PERFORMANCE ANALYSIS

Video preprocessing results: After the extraction of key frames from the video file, clustered them into groups. Then classify similar groups into one group and remaining in other group based on inter and intra cluster distance. Count the number of shots in those groups and named it as Group1 and Group2. Here researchers classified those groups as play field and others. To evaluate the performance of the clustering algorithm researchers are used, researchers manually classify the shots in to two groups. Count the number of shots in those groups and named it as Mgroups1 and Mgroup2. Calculate the accuracy by AccuracyA = G1/MG!; Accuracy B = G2/MG2. Table 1 and 2 shows the cluster accuracy and classification accuracy. Researchers tested accuracy with 2003 cricket world cup videos.

Table 1: Cluster accuracy

	Group 1			Group 2		
Videos	G1	MG1	Accuracy1	G2	MG2	Accuracy2
Ind vs. Pak	98	90	0.91	95	90	0.94
Pak vs. Ind	76	60	0.78	96	86	0.89

Table 2: Classification accuracy

	Group A			Group B		
Videos	GFA	GPNum A	AaccuracyA	GFB	GPNum B	AaccuracyB
Ind vs. Pak	14	10	0.71	13	10	0.76
Pak vs. Ind	10	8	0.80	12	8	0.66

Table 3: Event accurancy

	Events		
Videos	Mevent	Cevent	Accurancy
Ind vs. Pak	20	15	0.75
Pak vs. Ind	16	11	0.68

Video association mining and event detection: Event detection is a important procedure. Because researchers are classified events for indexing video shot. To evaluate the performance of event detection first manually checks what are the events available in the video. Count the number of events named as Mevent. Then classify the events with the algorithm researchers are used. Count the number of events named as Cevent. Table 3 shows the event detection accuracy.

CONCLUSION

The knowledge-based video indexing system has been designed and implemented to access video database management. This system can extract audio, visual cues from the video data file. Moreover, it combines the audio, visual cues in a single stream with their temporal order. So, that it is possible to view the events in correct sequence using association rule algorithm, researchers find the frequent item sets and search the interesting patterns. Finally, researchers classify the association with the respective events and carry out indexing.

Fuzzy K-Nearest Neighbor algorithm has advantages over the traditional K-Nearest Neighbor algorithm. While determining the class, the algorithm is capable of taking into consideration the ambiguous nature of the neighbors if any. But Fuzzy algorithms has been designed such that these ambiguous neighbors do not play a crucial role in the classification.

This research is only specifically used in cricket videos. But this can be extended to other domains also. Also it has only limited amount of events. Events can also be extended. The mining algorithms researchers used in the research is the modification of existing mining algorithm. So, this will be extended to find the new mining algorithm in video.

REFERENCES

Fan, F., D.K.Y. Yau, W.G. Aref and A. Rezgui, 2000. Adaptive motion-compensated video coding scheme towards content-based bit rate allocation. J. Electron. Imaging, 9: 521-533.

Fan, F., W.G. Aref, A.K. Elmagarmid, M.S. Hacid, M.S. Marzouk and X. Zhu, 2001. MultiView: Multilevel video content representation and retrieval. J. Electron. Imaging, 10: 895-908.

Fan, J., X. Zhu and X. Lin, 2003. Mining of Video Database. In: Multimedia Mining: A Highway to Intelligent Multimedia Documents, Djeraba, C. (Ed.). Kluwer Academic Publisher, Norwell, MA., USA., ISBN-13: 9781461354123, pp: 113-137.

Han, J. and M. Kamber, 2006. Data Mining: Concepts and Techniques. 2nd Edn., Morgan Kaufmann Publishers, San Francisco, CA., USA., ISBN-13: 978-1558609013, Pages: 800.

Jiang, H. and A.K. Elmagarmid, 1998. WVTDB: A semantic content-based video database system on the world wide web. IEEE Trans. Knowledge Data Eng., 10: 947-966.

Kolekar, M.H. and S. Sengupta, 2010. Semantic concept mining in cricket videos for automated highlight generation. Multimedia Tools Appl., 47: 545-579.

Pan, J.Y. and C. Faloutsos, 2002. GeoPlot: Spatial data mining on video libraries. Proceedings of the 11th International Conference on Information and Knowledge Management, November 4-9, 2002, McLean, VA., USA., pp: 405-412.

Snoek, C.G.M. and M. Worring, 2005. Multimodal video indexing: A review of the state-of-the-art. Multimedia Tools Appl., 25: 5-35.

Xie, L., S.F. Chang, A. Divakaran and H. Sun, 2003. Unsupervised Mining of Statistical Temporal Structures in Video. In: Video Mining, Rosenfeld, A., D. Doermann and D. DeMenthon (Eds.). Springer, New York, USA., ISBN-13: 9781441953834, pp: 279-307.

Zhang, H.J., A. Kankanhalli and S.w. Smoliar, 1993. Automatic partitioning of full-motion video. Multimedia Syst., 1: 10-28.

Zhu X. and X. Wu, 2003a. Sequential association mining for video summarization. Proceedings of the IEEE International Conference on Multimedia and Expo, Volume 3, July 6-9, 2003, Baltimore, MD., USA., pp: 333-336.

Zhu, X. and X. Wu, 2003b. Mining video association for efficient database management. Proceeding of the International Joint Conference Artificial Intelligence, August 12-15, 2003, Acapulco, Mexico, pp: 1422-1424.