Geetanjali Khanvilkar, Prof. B.B. Meshram / International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com
VOL. 1, ISSUE 4, PP. 1522-1526

Video Data Mining: Event Detection from the Association Perspective using FP-growth Tree

Geetanjali Khanvilkar
Computer Engineering Dept.,
Veermata Jijabai Technological Institute, Mumbai

Prof. B.B. Meshram
Computer Engineering Dept.,
Veermata Jijabai Technological Institute, Mumbai

Abstract-
Recent advances in computing, communication, and data storage have led to an increasing number of large digital libraries publicly available on the Internet. In addition to alphanumeric data, other modalities, including video play an important role in these libraries. Ordinary techniques will not retrieve required information from the enormous mass of data stored in digital video libraries. Instead of words, a video retrieval system deals with collections of video records. Therefore, the system is confronted with the problem of video understanding. The system gathers key information from a video in order to allow users to query semantics. The explicit definitions and evaluation measures for video associations are already introduced by integrating the distinct feature of video data. To explore associations among the audio and visual cues, a part of algorithm involves Apriori. In order to increase efficiency of algorithm, this paper is replacing Apriori with modified (to work on streamed data) Frequent Pattern-growth tree algorithm.
Index Terms- Video Mining, Multimedia Systems, Association Mining.

I. INTRODUCTION

Basically, content based video mining is achieved by first shot boundary detection in order to split video in different different shots. Then from each shot the key frame is extracted, which is assumed to be representative of that particular shot. Features from these extracted frames are obtained and used for video indexing purpose. When a piece of video is provided as input these indices are used for comparison and wherever match is found the belonging video get retrieved. But it limits accuracy, as it does not exhibit any temporal semantic. To this end, the semantic indexing and event detection from the association perspective is introduced to work with domain specific video. The association between different events occurring over time period related to specific domain video always provides the semantic insight to users.

In this paper, we are using fp-tree algorithm instead of Apriori in order to mine associations among video data. The use of fp-tree will improve the time efficiency. As well as to achieve parallelism, multicore capability of fp-tree is also employed.

The paper is structured in following manner: Knowledge based video indexing and system architecture are discussed in section 2. FP- growth to mine the associations between the different events is presented in section 3.

II. RELATED WORK

Video analysis and feature extraction form the core part of hierarchical semantic-sensitive video[7].

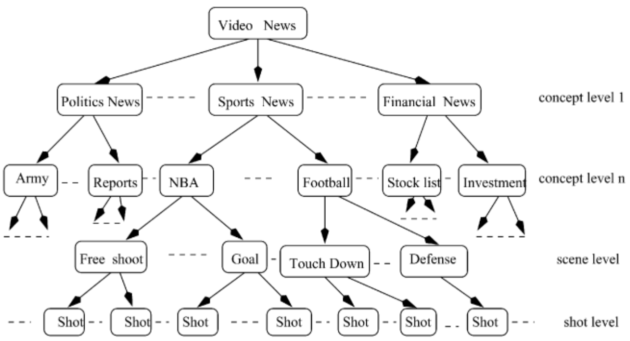

Fig. 1 Concept hierarchy for *video news* used in system

In general, videos are either contented, example news, documentaries or without contents example sport videos. In basketball videos, a series of actions, such as camera pan->camera still->camera zoom-in->applause -> scoreboard change, likely appear sequentially because they usually accompany a goal event. [6]

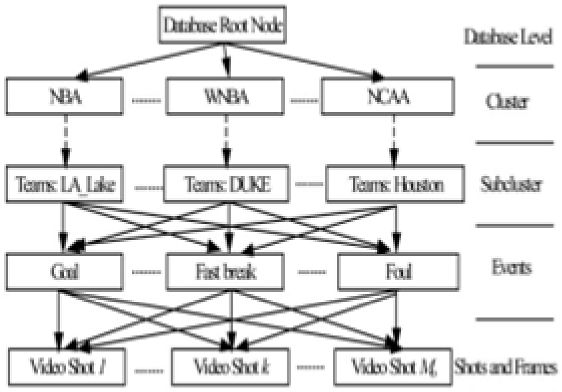

Fig. 2. Knowledge-based basketball video database management.

As shown in fig.2, the first level is the host association of the games, example NBA, NCAA. The second level consists of team names of each association such as Houston, Duke, where each video can be explicitly placed into one node. The interesting events from basketball can be used as nodes at third level of indexing structure. Index nodes with video shots are forming the lowest level of hierachy, where each shot may have more than one parent node as some shots may contain several events.

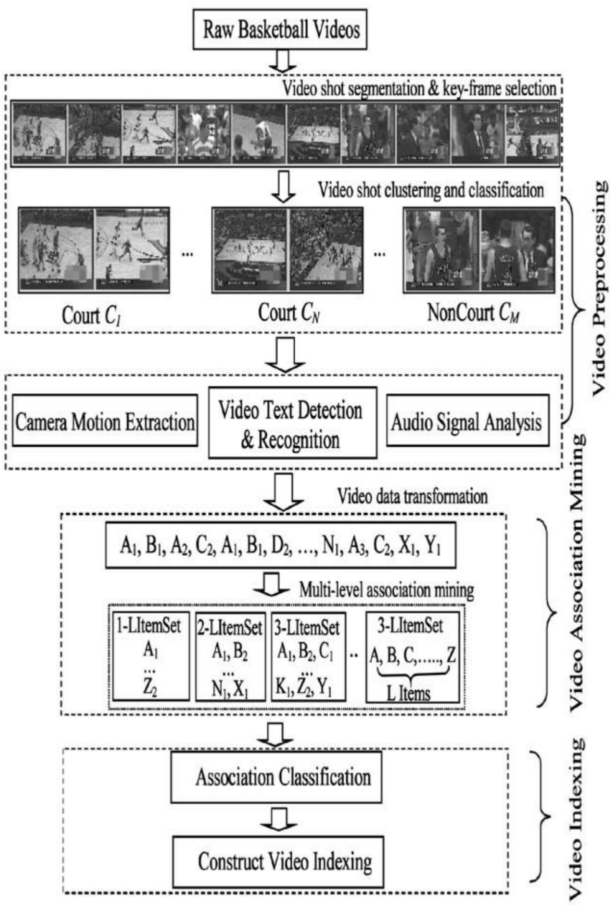

Fig. 3 The architecture of association-based video indexing

The architecture in fig.3 is implemented by following the steps:

a. Segment video into shots by using shot boundary detection techniques [10]. These shots are grouped into two: court and non-court with consideration of dominant color [11].

b. In order to construct second level index, team names are acquired using caption text detection. Caption text is detected by calculating edge difference between frames, the unchanged contents form the candidate text OCR (Optical Character Recognition) Engine, WOCR are used.

c. The camera motion in shots also gives some knowledge. To extract camera motions: Still, pan (left & right), zoom (in & out), a qualitative camera motion extraction method is used.[8]

d. Special audio events, e.g., audience applause & a referee's whistle help to get semantic cues. To detect audience cheering, the pitch of audio signal is used. To detect a referee's whistle, spectrum domain features are used.

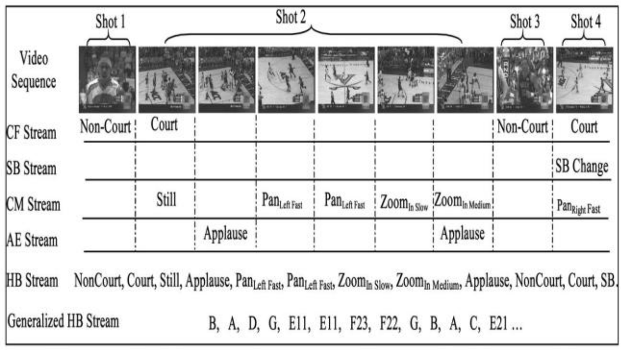

Fig. 4. Video data transformation and generalization.

e. The application of above mentioned techniques result in four separate streams. These streams combined in one hybrid stream.

Streams	CF Stream		SB Stream	CM Stream					AE Stream	
	Court	Non-court	SB Change	Still	$Pan_{Left Fast}$	$Pan_{Left Med.}$	...	$Zoom_{In Slow}$	Applause	Whistle
Generalization	A	B	C	D	E11	E12	...	F23	G	H

Fig. 5 A Mapping Table to Generalize Video Data

f. To mine associations, two measures are used. Temporal support and Temporal confidence, with respect of which semantic relationship can be achieved[6].

III. PROPOSED ASSOCIATION MINING WORK

To mine associations apriori is used. The drawback of Apriori is it works in iterative manner, where each iteration requires full database scanning. In order to lower the number of database scans instead of apriori, FP-growth algorithm is used. This will also lead to achieve efficient time and space complexity.

To have a compliance with temporal support, at some extent we will modify FP –tree algorithm. For the hybrid sequence fig.6, we will apply the first step of Apriori, to have the frequency of each event.Starting with null, whatever the temporal support accordingly we will partition the sequence.

Let TDT=2, then ABE FB BE (first two shots) is giving first window. By arranging these events in descending order of their frequency, we will have B:3, E:2,F:1, A:1 which gives rise to fig. 7a. Counter is used to keep track of occurrence of each event, every successive event is attached as child node to its parent node, more frequent one. While proceeding with next window, the overlapping shots are used to maintain temporal consistency among successive shots. The next window will be of second and third shots i.e. BE ADF. Arranging them in their order of frequency, we will get B:1, E:1, F:1, D:1, A:1 of which shot 2 with events BE is already processed, no counter increment for B, but with BEF as prefix, we need to add D as child node to F and then adding A as child node to D. Third and Fourth shots will form our third window to be processed. i.e. ADF EBDG, again by following same procedure as applied on previous windows, we will have B:1 E:1 D:2 A:1 F:1 G:1 of which D:1 F:1 A:1 events are already processed, we need to increment counters of B and E by one only and G should be added as child to D shown in fig 7c. The same procedure is repeated till the end of stream of video.

```
1. count [ ] = { countA, countB,……………….}
      countA: counts event A
      countB: counts event B and so on
2. scan the stream
      If (A) countA++
      If (B) countB++
      And so on
3. Arrange count[ ] in descending order.
      Sort the events and their frequencies in descending order.
4. Select TDT
5. ShotNo=1
6. root=null                                  //fp tree construction
7. scan shots(ShotNo+TDT)
8. arrange them in order found in step2
9. add first event as child node to root, second node as child node to first, and so on
10 ShotNo++
11.  while(end of stream)
12.     repeat steps 7 and 8
13.     if prefix is common follow the already available path else add it as new branch.
14.     ShotNo++
15. end of while
```

Fig. 6 FP-tree Algorithm

Starting with the leaf node, the conditional pattern base can be found. For let video data A, which is having two conditional pattern bases, <B:5 E:4 F:2 A:1> and <B:5E:4 F:2 D:1 A:1>.

As in each conditional pattern base the association is found for once and if support that mentions the least number of times the event should occur is 2, then no conditional fp-tree is obtained. This support threshold is different than the temporal support. In case of video data B, two conditional pattern bases are obtained, <B:5 E:4 F:2 D:1> and <B:5 E:4 D:2>. With second conditional pattern base D satisfies support 2, so can be used for determining conditional fp-tree. Similarly, for all nodes i.e. video data, conditional fp-tree are found. Then, applying temporal confidence the appropriate association rules are obtained.

shot1	shot2	shot3	shot4	shot5	shot6
ABEEFB	BE	ADF	EBDG	DC	CEG

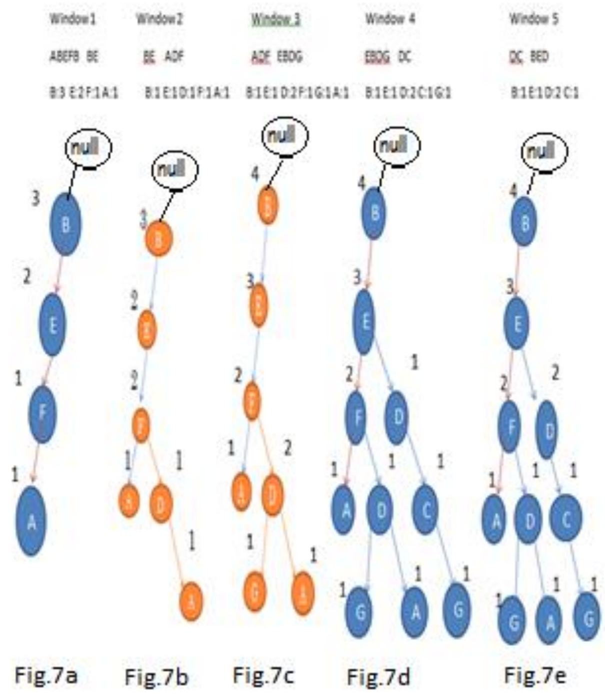

Fig. 7. FP-tree construction

Video Data	Conditional pattern base	Conditional fp-tree
A	< B:5 E:4 F:2 A:1> < B:5 E:4 F:2 D:1 A:1>	null
C	<B:5 E:4 D:2 C:1>	null
D	<B:5 E:4 F:2 D:1> <B:5 E:4 D:2>	<B, E, D:2>
E	<B:5 E:4>	<B, E:4>
F	<B:5 E:4 F:2>	<B, E, F:2>
G	<B:5 E:4 F:2 D:1 G:1> <B:5 E:4 D:2 C:1 G:1>	null

Fig. 8. Frequent Event Determination

By searching patterns from hybrid stream with constraints, appropriate association can be mined as well as manually these associations are evaluated to put in appropriate class of event.[6] Once event is detected for each of the sequence of associated video data, indexing is done.

To simplify the concept of video indexing, flow chart is specified, which shows the flow among different steps involved in it. As FP-tree requires only two passes, it is faster than Apriori. It takes time to build, but once it is bilt, frequent sequence of video data are read off easily.

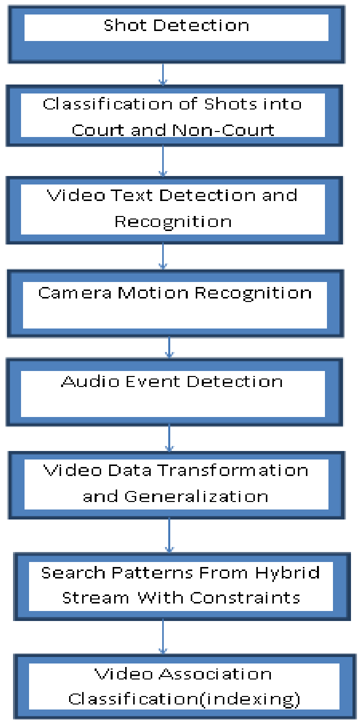

Fig. 9. Flow Chart for Video Indexing

IV. Future Works And Conclusion

In this paper, we have proposed the FP-tree on Streamed data in order to increase the efficiency of association mining algorithm. The strategies presented here are specific to basketball videos, but for any kind of videos from which associations among events are to be extracted FP-tree can be applied. As FP-tree construction does not involve candidate set generation as that of Apriori, less computations, less number of scanning and less amount of time it takes to mine associations among events. In future to achieve more efficiency, multithreading capability is to be employed.

References

[1] Jian-Kang Wu: "Content-Based Indexing of Multimedia Databases" November/December 1997

[2] Eung Kwan Kang, Sung Joo Kim, and Joon So0 Choi :"Video Retrieval Based On Scene Change Detection In Compressed Streams" August 1999

[3] Mohand-SaõÈd Hacid, Member and Jacques Kouloumdjian :"A Database Approach For Modeling And Querying Video Data"
Septeber/October 2000

[4] Jianping Fan, Ahmed K. Elmagarmid, Xingquan Zhu, Walid G. Aref and Lide Wu: "ClassView:Hierarchical

Video Shot Classification, Indexing, and Accessing" February 2004

[5] Domenico Cantone, Alfredo Ferro, Alfredo Pulvirenti, Diego Reforgiato Recupero, and Dennis Shasha : "Antipole Tree Indexing to

[6] Support RangeSearch and K-Nearest Neighbor Search in Metric Spaces" April 2005

[7] Xingquan Zhu, Ahmed K. Elmagarmid, Zhe Feng, and Lide Wu: "Video Data Mining: Semantic Indexing and Event Detection from the Association Perspective" May 2005

[8] M. M. Yeung and B.-L. Yeo : "Video visualization for compact presentation and fast browsing of pictorial content" Oct 1997

[9] X. Zhu, A.K. Elmagarmid, X. Xue, L. Wu, and A. Catlin: "InsightVideo: Towards Hierarchical Video Content Organization for Efficient Browsing, Summarization and Retrieval" 2004

[10] S. Nepal, U. Srinivasan, and G. Reynolds:"Automatic Detection of 'Goal' Segments in Basketball Videos"

[11] J. Fan, W.G. Aref, A.K. Elmagarmid, M. Hacid, M. Marzouk, and X. Zhu, "MultiView: Multi-Level Video Content Representation and Retrieval," J. Electronic Imaging, vol. 10, no. 4, pp. 895-908, 2001.

[11] L. Xie, S. Chang, A. Divakaran, and H. Sun, "Structure Analysis of Soccer Video with Hidden Markov Models," Proc. IEEE Int'l Conf. Acoustics, Speech, and Signal Processing (ICASSP), 2002.

Geetanjali Khanvilkar is a student of MTech, VJTI Mumbai. Her research interest include Data Mining, Business Intelligence.

Dr. B.B. Meshram is Head of Computer Engineering Department, VJTI Mumbai. His research interests include Web Security, Business Intelligence and Data Mining. He has a PhD in Computer Engineering from Mumbai University in Mumbai. He's a senior member of the ACM and a member of the IEEE Computer Society.