

# A Survey on Event Recognition and Summarization in Football Videos

Dhanuja S Patil<sup>1</sup>, Waykar S.B<sup>2</sup>

<sup>1</sup>ME Student, Department of Computer Engineering, Sinhgad Institute of Technology, University of Pune, Maharashtra, India

<sup>2</sup>Department of Computer Engineering, Sinhgad Institute of technology, University of Pune, Maharashtra, India

**Abstract:** *There are large number of sports videos are available over the World Wide Web (www). But using these complete videos to get information is a hectic and time consuming job. So, the need of extraction on events from the videos aroused. Semantic analysis of videos and automatic event extraction plays a vital role in several applications; using content-based search engines, video indexing, and video summarization. As a powerful tool for learning complex patterns is the Bayesian network, this paper proposes a novel Bayesian network (BN) based method automatic event recognition and summarization in Football videos. This method includes efficient shot boundary detection algorithms, shot view classification, and the related Bayesian network construction. There are three main stages: The shot boundaries are detected in first stage. Using some model, the video is segmented into the play-break sequences, large but meaningful semantic units. Within the next stage, these play-break sequences are used to extract several key events. Lastly, in the final stage, the Bayesian network is used to get the high level semantic events. Constructing the Bayesian network is the basic part of the method. By applying family of Copula, the joint distributions of arbitrary variables of the network are modeled. Some of events that can be recognized by this method in Football videos are goal, card, corner, shots on goal, foul, offside, missed shots and non highlights. The users are more likely to be interested in these events and not in complete and large videos.*

**Keywords:** Bayesian network, shot boundary detection, play-break sequence, Copula distribution, Video indexing.

## 1. Introduction

The growing number of audio-visual data over internet has increased enormously over the last decade. Therefore the large capacity of storing these data has also increased. Thousands of GBs of audio-visual data is available across the internet, which is base for many applications those are used for different purposes. Thus, providing the appropriately indexed library is must. When the video data exceeds their limits, some technique is needed to summarize and extract the events from these videos and only necessary part should be given to the user as per their requirements. From the different genres of videos, Sport videos attract most of the users. Sports are basically composed of interesting events, largely called as highlights. Large number of users are interested in these highlighted events rather than the full length videos. There are some generic summarizations techniques are available and are sufficient, but some automatic technique will be more useful. Event detection and summarization makes it possible for the delivery of videos over narrow band networks, because the special events normally occupies a small part of whole content<sup>[1]</sup>. The main problem with the video event detection is "Event boundary detection". Methods like Dynamic Bayesian Networks, for detecting events, uses a Frame-based algorithm<sup>[2]</sup>.

The initial step for processing any video is Shot Boundary Detection, the process of detecting boundaries between shots automatically. It is a base step for any kind of video analysis, indexing, summarization, search or any other event based operation<sup>[3]</sup>. A given event is generally consist of multiple shots, like, long, medium, close-up and outfield. Likewise, a specific shot indicates a certain event at some probabilistic manner. At present, lot of work has been done on shot classification. Most of them are dedicated to play-

break sequence using shot classification<sup>[4]</sup>. Bayesian Networks, from the probabilistic Graphical Models family. These Graphical models are helpful to characterize the information about a tentative domain. In particular, random variables are represented by each node. On the other hand, the edge between nodes represents the dependencies between these nodes<sup>[5]</sup>. These Bayesian Networks will allow the system to trace the movements on the ground, which includes Players, ball, and referees. The players are needed to be traced for obvious reasons as they are the primary entities to generate or produce the events. The ball is needed to be traced to ensure the events like, Goal, shots on goal, missed shots, etc. The referees are needed to be followed for the events, such as Red or Yellow card shown to the players, Substitutions, and Offside's. There are three main stages. First stage is to detect the shot boundaries, and once the shot boundaries are detected, then segmenting the video into play-break sequence. These Play-break sequences are large semantic units, but are very meaningful according to the events that are been traced. The second stage uses these play-break sequences and extracts the key events from the video. Now, the key events in any Football match are likely to be, Goal, Red-Yellow cards, Fouls, Offside's, Shots on Goal, Injuries, Substitutions made, etc. The final stage is to get such high level semantic events from the given video as per the user requirements.

The important issue that needs to be solved in the near future is replays of events those are shown in the video. The replays are showing the already concluded events again and again. The Goal event, offside events, is such events are repeatedly shown in the video. The system must not consider these replays as occurrences of events. If this problem isn't solved, then output will be inappropriate. Another problem that can be faced is construction of Bayesian Networks. The faster algorithm is needed as the football is a sport, where the positions of all elements, like players, ball, and referees

will change in almost every frame. The remaining paper can be summed out as: Section II gives the related techniques for Shot boundary detection, which is the basic and important step for video summarization, and Shot view classification. Later in the section is the researches made on the important techniques necessary, the Bayesian networks and the Copula families. The play-break sequence is also described which is necessary for the segmentation. And extracting of key events is the last topic covered in the section. In Section III, given is the briefly reviewed conclusion of this proposed paper.

## 2. Literature Survey

The literature on the Event Recognition and Summarization of Football videos can be classified into: Shot boundary detection, Bayesian network construction, copula families, player and ball movement detection, extracting the key events.

### a. Shot Boundary Detection

P. P. Mohanta<sup>[6]</sup> proposed a Model based shot boundary detection technique. He used frame transition parameters for the technique. The previous and the next frame can be used to formulate frame estimation scheme. Frame transition parameters were based on this formulated frame estimation scheme. Z. M. Lu<sup>[7]</sup> proposed a Singular Value Decomposition and Pattern Matching based, Fast Video Shot Boundary Detection. Singular Value Decomposition (SVD) and Segment Selection was the key for this technique. Initially, a shot boundary position and gradual transition lengths were predicted. The prediction was based on adaptive thresholds. Some non-boundary frames were removed simultaneously.

### b. Bayesian Network Construction

M. Singh<sup>[8]</sup> proposed an algorithm to incorporate the Conditional Independence (CI) tests and ordering of nodes. The ordering of the nodes from database was generated by using the CI tests. These nodes are then used to construct the principal Bayesian network structure with the help of some non-CI based methods.

I. Ben-Gal<sup>[5]</sup> explained that, Bayesian Networks can be used to learn the informal relationships and to understand various problem domains. It can be also used to predict the potential data. All this features are also supported in the case of missing data.

F. Ya<sup>[9]</sup> established a multilayer dynamic BN model according to domain knowledge of sports video. This included primitive layer, characteristic layer and events layer. The direct mapping difficulties were avoided from high semantics and low characteristics by elementary composition and mapping back to characteristics.

### c. Copula Families

By R. B. Nelson<sup>[10]</sup> a copula is a function to couple a multivariate distribution function with respective 1-D marginal distribution function. The families of Copulas like, the Farlie-Gumbel-Morgenstern (FGM) family, Copulas cubic, normal Copulas are explained in brief. The Archimedean copulas are explained as well. Ian McHale<sup>[11]</sup> used Copulas to generate discrete distributions. And these

data is used for the videos from English Premier League. The main focus was on the Archimedean Copulas, because for this family of Copulas, a 1-Dimensional projection can be fully determined for dependency structure. Emil Stoica<sup>[12]</sup> explained that the Copula is a function, that describe dependency structure between marginal distributions. And that FGM family is an important parametric family of Copulas. Stoica explained the stability property of FGM copula family with help of some briefly explained operations.

### d. Player and Ball Movement Detection

Y. Huang [13] proposed that, using the Euclidean Distance Transform system can automatically detect and extract skeletons for almost every foreground spot. The shape analysis will then be performed on these spots to remove false shapes, i.e. non-player and non-ball shapes.

In Event detection, tactical analysis and in automatic summarization, detection and tracking of important object like, players and ball and referees is very important [14].

Xiao-Feng Tong<sup>[15]</sup> explained that, feature extraction from a frame is largely distributed in 4 operations: Field ratio, which deals with the understanding of the field. Texture, mostly used technique for separating players and game objects from the field. Head Area is important to determine as for the close up frames to identify the player. And the last is Object Ratio, using these differences between the long and medium range camera views can be understood.

### e. Extracting the Key Events

D. W. Tjondronegoro<sup>[16]</sup> proposed an automatic highlight classification model with the base as a universal scope of detection and set of audiovisual data. This model combined the statistics with rule based models. V. Naik<sup>[17]</sup> proposed a system that will using Fuzzy logic, try to implement human knowledge directly into the system. The event detection will take place in two phases. First is to identify the event frames and categorize them. In later phase, the significant events are extracted using the Fuzzy logic. H.G. Kim<sup>[18]</sup> introduced a highlight extracting system based on Audio contents. The three steps needed are, first, extraction of the audio from the video sequence. Secondly, using Hidden Markov Model and some feature extraction method, the event candidates are detected from the video. Finally, determine the events to be added to the highlights.

## 3. Conclusion

In this report, we presented a video summarization and event recognition techniques for Football videos. We have used some techniques for shot boundary detection to remove the unwanted portions from the video. Similarly, used Shot view classification. Bayesian Networks are constructed to trace the actual and complete movements of all players, the ball and the referees as well. Players are needed to be traced for obvious reasons, the movement of the ball is traced for the events like Goal and shots on goal. The referees are also followed by the system to mark out the events like Red-Yellow card shown, and Offside events. Generating the play-break sequence is the most important task, as it will provide us with the meaningful portions of the video. And

these sequences are further used to extract the several key events. From these key events, we will provide the required events to the user, as per their requirements. Still, this paper does not claim to have solved all the problems related to the event recognition and summarization in Football videos, we have indeed taken a step forward towards our goal.

using Fuzzy Systems”, Int. Journal of Emerging Technology and Advanced Engineering.  
[18] H.G. Kim, S. Roeber, A. Samour, T. Sikora, “Detection of Goal Events in Soccer Videos”, for Technical University of Berlin.

## References

- [1] Ekin, A. M. Tekalp, R. Mehrotra, “Automatic Soccer Video Analysis and Summarization”, IEEE Trans. Image Process.
- [2] L. Huang, H. C. Shih, C. Y. Chao, “Semantic Analysis of Soccer Video Using Dynamic Bayesian Network”, IEEE Trans. Multimedia.
- [3] F. Smeaton, P. Over, A. R. Doherty, “Video Shot Boundary Detection: Seven Years of TRECVID Activity”.
- [4] P. Xu, L. Xie, S. F. Chang, A. Divakaran, A. Vetro, H. Sun, “Algorithm and System for Segmentation and Structure Analysis in Soccer Video”, Proc. IEEE Int’l Conf. on Multimedia and Expo. 2001.
- [5] Irad Ben Gal, “Bayesian Networks”, Encyclopedia of Statistics in Quality and Reliability.
- [6] P. P. Mohanta, S. K. Saha, “A Model Based Shot Boundary Detection Technique using Frame Transition Parameters”, IEEE transactions on Multimedia.
- [7] Zhe Ming Lu, Yong Shi, “Fast Video Shot Boundary Detection Based on SVD and Pattern Matching”, Image Processing IEEE transactions.
- [8] M. Singh, M. Valtorta, “Construction of Bayesian Network Structures from Data: a Brief Survey and an Efficient Algorithm”, Int. Journal of Approximate Reasoning.
- [9] F. Ya, “The Modeling on Sports Video Based on Bayesian Network and Future Maps”, Huainan Normal University.
- [10] Roger B. Nelson, “Properties and Applications of Copulas: A Brief Survey”.
- [11] McHale, P. Scarf, “Modeling Soccer Matches using Bivariate Discrete Distributions with General Dependence Structure”, Centre for Optional Research and Applied Statistics.
- [12] Emil Stoica, “A Stability Property of Farli-Gumbel-Morgenstern Distribution”, Dragon European University of Lugo.
- [13] Y. Huang, J. Llach, S. Bhagavathy, “Players and Ball Detection in Soccer Video Based on Color Segmentation and Shape Analysis”, Thomson Corporate Research.
- [14] Wang J. R., Parameswaran N., “Survey of Sports Video Analysis: Research Issue and Applications”, Pan-Sydney Area Workshop on Visual Information Processing.
- [15] Xiao-Feng Tong, Qing-Shan Liu, Han-Qing Lu, Hong-Liang Jin, “Shot Classification in Sports Video”, National Laboratory of Pattern Recognition.
- [16] W. Tjondronegoro, Yi-Ping P. Chen, “Knowledge-Discounted Event Detection in Sports Video”, Senior Members, IEEE.
- [17] V. Naik, G. Rathod, “An Algorithm for Retrieval of Significant Events Near Goal Post from Soccer Videos