19th International Computer Science and Engineering Conference (ICSEC)
Chiang Mai, Thailand, 23-26 November, 2015

# Automatic Soccer Archive Summarization using Time Constraint

Sutthikarn Bojukrapan
Department of Computer Sceince, Udon Thani Rajabhat University 64 Thaharn Road, Muang, Udon Thani 41000, THAILAND
sutthikarn@udru.ac.th

Punpiti Piamsa-nga
Department of Computer Engineering, Kasetsart University, Chatujak, Bangkok, 10900, THAILAND
pp@ku.ac.th

*Abstract*— **To manage a massive growth of sport videos, we require to summarize the content into more compact and dominant representation. The previous research projects proposed soccer summarization by using highlights or scene playing; however, they were unable to do it in a given time period. In this paper, we propose an automatic soccer video summarization that can be done within the time constraint. Our proposed approach consists of three parts: shot boundary detection, event detection, and summarization. Our proposed method evaluated by using 10 soccer matches from the FIFA World Cup 2014. Experimental results show the performance that the most users satisfy at "good" level.**

***Keyword:Automatic, soccer, summarization, time constraint.***

## I. INTRODUCTION

Sport archives analysis is used to extract significant information in order to add more values to news, advertisement, coaching and many other sport-based applications. There are many sport archives such as soccer match, tennis and racing. Each sport archive consist of interesting events such as goal, popular player and plays of offence and defense. If sport archives are segmented according to these semantically meaningful events, it can be used in numerous applications to enhance their values and enrich the user's viewing experiences [1]. Soccer, one of world most favorite sport, is popular team sports because the relative simplicity of its rules and using only a ball to play[2]. Each soccer match use a lot of time to search interesting event.

The soccer archive analysis has a good deal research and wide spectrum of possible applications have been considered. Traditionally, soccer were analyzed manually but it costs valuable time. Therefore, it is necessary to have automatic methods to analyze content of soccer matches. Each soccer match use time at least 90 minutes which contains only a few minutes of moment of interest such as goal, shoot, free kick, foul and red cards. Most people on news and the internet prefer a summarized version to the full-length match [14]. In recent soccer archive on broadcast and internet rapid increased. Therefore, automatic soccer video summarization is an essential method for ease and flexible accessibility to video content. It is clear that the ability to access semantic events among lengthy and voluminous videos by skimming non-interesting parts, it can help to reduce the time to review all of the matches[15]. A concise and informative video summary enable a user to quickly figure out the overview contents of a video and decide whether the whole video program is worth watching. To generate a perfect summarization of a given video requires good understanding of the semantic content. However, automatic understanding of the semantic content of general videos is still far beyond the intelligence of today's computing systems. Despite the significant advances in computer vision, image processing, pattern recognition, and machine learning algorithm, video summarizing aims to provide a summary of a long video for shortening the navigation and browsing the original video. Soccer archive matches always attract major sports audience. According to [4], there are two fundamental type of video abstracts viz. static video abstract and dynamic video skimming. A static abstract is a small collection of salient images extracted from the original video sequence. A dynamic skimming consists of both the image sequences and the corresponding audio abstract.

In this paper, we present a soccer archive summarizing technique. The output summary is an aggregate of clips containing a set of interest events from a user perspective, namely goals, corner kick, red card and penalty. The paper is organized as follow. In Section 2, we present an overview of the proposed system. Section 3 gives our experimental results. Finally, section 4 draws conclusions and discusses our future work.

## II. LITERATURE REVIEW

Several methods have been proposed for the event detection in sport archives using the audio track. For example, Coldefy and Bouthemy in[16] use the audio pitch and energy to detect the excitement parts of the sport archives. Then, they used color information to differentiate play events from advertisement. Cinematic features like replays are used as a highlight indicator. For example, Zhao et al.in [17] proposed a sport archives event detection and summarization method based solely on the detection of replays. The replays are detected depending on logo detection using rule-based approach. The detected parts are then ranked using audio energy and motion activity. Based on the common perception that sport archives have well-defined, repetitive structure, a number of methods have been developed to model and predicts sports events. For example, Ye et al. [18] use Support Vector

978-1-4673-7825-3/15/$31.00 ©2015

Classifiers to identify exciting events from soccer archives. They extract medium-level features to classify soccer shots.

A simple approach to extracting key-frames from a shot is based on frame content changes computed by features, such as color histogram or motion activity. Huiyo Zhou et al.[4] proposed an effective dynamic video summarization algorithm by using audio-visual features extracted from video. Audio, color and motion features are dynamically fused using an adaptively weighting mechanism. Dissimilarities of temporal video segments are formulated using the extracted features before these segments are clustered using fuzzy c-means algorithm with an optimally determined cluster number. Zhicheng Wei and Xue Yang [5] proposed a genetic-algorithm-based model for soccer archive summarization using time density function. There are three steps in our summary method. First, video time density function is employed to model each video sequence. Second, fitness function is used to measure important factors among alternative frames. Third, a genetic algorithm is developed to get the optimal group of key-frame. Hossam M. Zawbaa et al.[6] presented a system for soccer archive summarization using support vector machine (SVM). The system applies for emphasizing important segments with logo appearance with addition to detecting the caption region providing information about the score of the game. The system uses k-means algorithm and Hough line transform for detecting vertical goal posts and Gabor filter for detecting goal net and the system highlights the most important events during the match. Ekin[7] constructs comprehensive sport archives summarization and analysis using a structural semantic video model. C. Huang et al.[8] introduced semantic analysis based on Bayesian networks to indentify a group events (goals, penalties, card and corners).

Since sports events are basically composed of series of movement, a number of methods have been developed based on motion estimation and modeling; for example, Chen et al. [1] introduced a motion based method for event detection in tennis, baseball and soccer. They computed motion entropy as a function of time. Entropy curve is then series methods. A set of heuristics is then applied to the simplified entropy curve to identify sports events. Rule based methods have also been investigated in event detection and action recognition. For example, Tjondronegoro et al. [19] selected game highlights by segmenting the video into a number of plays and break shots. They classified the frames (global, zoom-in and close-up) and only used global views. Finally, based on statistics collected from training videos, they define rules for identifying highlight sequences of the video.

In this paper, we proposed a soccer archive summarization system, which relies on event detection that can be done within a given time constraint. The users satisfied sumarized results at "good" level.

## III. METHOD

The automatic video summarization using time constraint system propose in this paper is composed of four fundamental building phases; 1) *shot boundary* phase that segments the whole video stream into small video shots, 2) *event detection* phase resulted from the shot boundary phase, 3) *assignment weighting shot*, 4) generateing a clip using time constraint. The four phases are described in Figure 1.

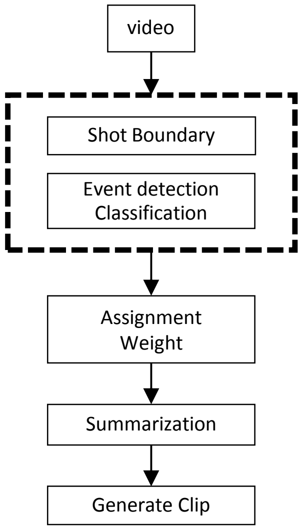



Figure 1. The overall structure of the proposed plan.

### *A. Shot boundary detection Phase*

The separated views that come from multiple cameras positioned at different locations. It can be realized that while changing from one camera to another, this indicates a start of a new shot and marks a boundary of a new shot. [8]-[10]. Step of shot boundary detection is each frame is divide into blocks with $m$ rows and $n$ columns. Then the difference of the corresponding blocks between two consecutive frames is computed. Finally, the final difference of two frames is obtained by adding up all the difference through different weights. Histogram plot is used to determine differences between blocks of two frame. Histogram is use to determine variation of pixel at different frequency of image. Let $F(k)$ be the $k^{th}$ frame in video, $k$=1, 2..$n$, the algorithm of shot boundary detection is described in [6] and as follows.

*Shot boundary detection algorithm*. For each frame convert the frame from RGB to HSI color space. Do frame skipping by step (k = 10 frames) to convert a gradual transition into a cut transition then calculate the mean change equal the difference between current hue frame and the next hue k-frame. Divide the original color frame into 32 x 32 blocks size. Compute the total percentage of changed blocks equal the total percentage of changing blocks between the current frame and the next k-frame. Compute the grass ratio equal the average difference of grass dominate color between the current frame and the next k frame and If (Mean change > T hr) and (total percentage of changing blocks > 0.25) and (grass ratio > 0.1) then mark a new shot.

### *B. Event detection Phase*

Most exciting events occur in the goal-mouth area such as goals, shooting, penalties, direct to free kicks, etc. Other non-exciting events such as dull passes in the mid-field. Defense and offense of some other shots to the audiences or coaches, are not considered as exciting as the former events[10]. Excitement event detection is based on seven features; namely 1) goal , 2) goal net, 3) replay, 4) attack

(a) (b) (c) (d)

Figure 2 Event interesting in soccer archive (a) goal net (b) penalty (c) goal (d) penalty

*1) Goal detection:*

A real time system for goal detection has been presented in [21] using four cameras with high frames rates (200 fps) in order to assure that high velocity shots would be captured. When the ball touches the under side of the bar and immediately bounces back into the field. The ball detection approach used by them is composed of two techniques: a circle detection algorithm (Circle Hough Transform) for ball candidate generation and a neural classifier to validate that candidate.

However, it is difficult to verify these conditions automatically and reliably by the state-of-the-art video processing algorithms. The occurrence of a goal event leads to a break in the game [23]. Figure 3 illustrates the sequence of cinematic features after scoring a goal.

Figure 3 goal image: examples.

*2) Goal net detection:*

Real time goalmouth detection has been performed in [22]. The authors estimate the dominant green color using the HSI space based on the assumption that field color may vary from stadium to stadium and the color is constrained to shadows and weather conditions. Coarse Spatial Representation (CSR) is also implemented for dominant green detection when coarse image resolution is found. The authors argue noise is reduced when applying the Hough transform within CSR green blocks. In order to find the dominant goal line orientation, their method starts with fixed parameters for the possible angle ranges, being refined during the first five minutes of the game. The authors also point out ways for making this calibration quicker in live game context. The two vertical bars are characterized by vertical strips of white followed by high contrast against the backdrop. For the horizontal cross-bar detection, it requires the goal line orientation previously detected in order to generate candidate lines connecting the vertical bars.

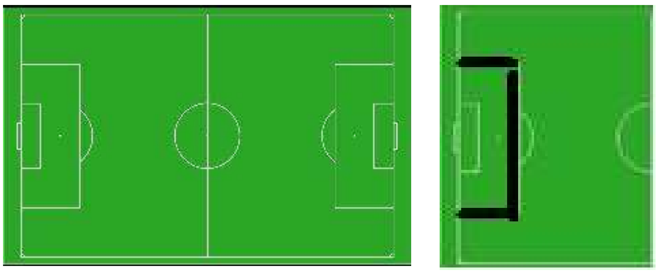
Figure 4 Goal net detection: an example

*3) Penalties box detection:*

In this section, we reduce the penalty box detection problem to the search for three parallel lines. In Figure 5(a), a view of the whole soccer field is shown, and three field lines, shown in figure 5(b), become visible when the action occurs around or within one of the penalty boxes. To detect lines, we use edge response, defined as the pixel response to the 3x3 Laplacian mask. The pixels with the highest edge response, the threshold of which is automatically determined from the histogram of the gradient magnitudes, are defined as line pixels.

$$h = \begin{vmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{vmatrix} \tag{1}$$

Then, three lines are detected by Hough transform that employs size, distance and parallelism constraints.



(a) (b)

Figure 5 Penalty area (a) Soccer field model and (b) three highlighted lines around goal area.

*4) replay detection:*

In most soccer matches, exciting events are often replayed. These exciting shots normally correspond to highlights in a game, e.g., actions near the goal posts in a soccer match. Replay is a video editing way that is often used to emphasize an important segment with a logo appearance for once or several times. The inputs to replay detection phase are the output shots from the shot boundary detection step of the pre-processing phase in order to extract the exciting events that are represented by the replay shots. In sport archives, there is often a highlighted logo that appears at the start and end of a replay segment, which indicates an exciting event within the soccer match. In the recent years, broadcasters use inserted logo sequence, as a digital video effect to replay the exciting and important events in soccer archives [20], as presented in figure 5 that shows an example of gradual logo appearance.

*5) Attack and other detection:*

Attack events may also match a lot of goal event features, although it is not as consistently as goals. The addition of attack events in the summaries may even be desirable since each of these events consists of interesting shots [7]. There are other interesting events such as fouls, cards, injure, or offside. The addition of these events in the summaries may even be desirable in order for each event to contain of interesting shots. Therefore, more users may enjoy watching interesting fouls and offside events.

## *C. Weight assignment*

The weight given to each shot acquisition of inquiries from users who satisfy watching football. The abbreviation in the first football match to see any of the shots required for a limited time. The weights shown in the table below.

$$w(e_i) = \sum_{j=1}^{n} w(x_j, e_i)/E(n) \tag{2}$$

When $w(e_i)$ is weight of each event $E_i$

## *D. Summarization using Knapsack problem*

The knapsack problem is actually rather difficult in the normal case where one must either put an item in the knapsack or not. However, in order to illustrate greedy algorithms, we consider a much simpler variation in which we can take part of an item and get a proportional part of the benefit. This is called the “fractional knapsack problem” since we can take a fraction of an item. (The more common knapsack problem is called the “0-1 knapsack problem” since we either take all (1) or none (0) of an item.

In symbols, for each even i we choose an amount $x_i$ $(0 \le x_i \le w_i)$ that we will place in the knapsack. We are subject to the constraint that the sum of the xi is no more than W since that is all the knapsack can hold. We again desire to maximize the total benefit. Since, for event i, we only put $x_i$ in the knapsack, we don't get the full benefit. Specifically we get benefit $\left(\frac{x_i}{w_i}\right) b_i$

Algorithm 1 FractionalKnapsack(S,W):

```
Input:  Set Event of items i with weight wi and
benefit bi all positive. Knapsack capacity
w > 0.
Output: Amount xi  of  i that maximizes the
total benefit without  exceeding the capacity.
for each i in Event do
   xi ← 0
   vi ← bi/wi
   w ← W
   while w > 0 do
     remove from Event item of maximal value
   xi ← min(wi,w)
   w ← w-xi
```

## IV. RESULTS

The proposed sport video summarization framework has been tested on several games of soccer. The test data set consisted of soccer archives 10 games from the FIFA World Cup 2014. Weight of each shot to calculated shown that table 1

Table 1 Weight of each shot

| Event NO. | Event | Weight |
|---|---|---|
| 01 | Goal | 1.00 |
| 02 | Goal net | 0.83 |
| 03 | Attack | 0.67 |
| 04 | Replay | 0.50 |
| 05 | Audience | 0.33 |
| 06 | Other | 0.17 |

Data of soccer 10 matches, each match has 6 different events, time and event as show that figure 6. The a number of matches in in x-axis and the event is on the Y-axis.

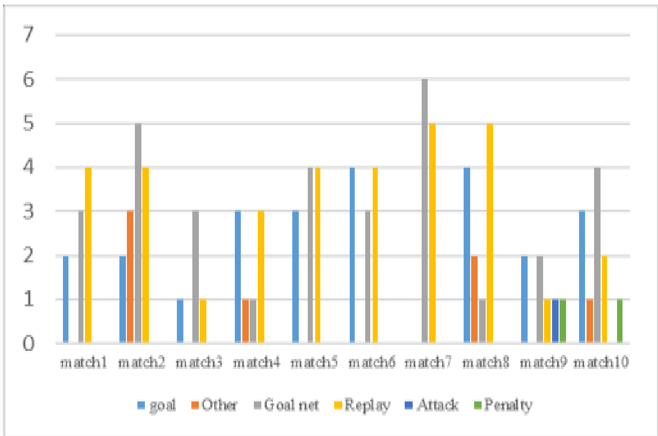



Figure 6 Number event detection

Result of soccer archive summarization from full-time 90-minute footage to a one-minute clip. The satisfaction survey of 10 members of the soccer audience found that the satisfaction level is good with the average of 4.07 as shows that figure 7.


Figure 7 Result of 10 soccer matches

## V. CONCLUSIONS AND DISCUSSION

The worldwide popularity of soccer between team sports is certainly due to its simplicity: there is no need of expensive kit and to enjoy watching the game takes very little time as grasping the basic rules is relatively straightforward compared to many other sports. For this reason, soccer archive summarization has attracted match research in the last decade. In this paper we present automatic video summarization using time constraint. The soccer content is very challenging for computer vision community since many practical issue must be considered. Depending on the range of application and interest, different methods seem to be more appropriate. In video summarization, the main problem is the extraction of significant events from broadcast images for automatic video clip generation. These systems do require of the images and event detection can be done by analyzing the common characteristics of soccer archives during and after significant events, such as goal, goal net , red-card and methods based on the extraction of low-level visual features and on mid-level processing to model the feature appearance such as camera motion. Since they are used for off-line processing. They do not consider the computational cost and often require manual intervention for the initial parameter estimation.

## REFERENCES

[1] Chen, C.-Y., Wang, J.3C, Wang, J.-F, Hu, Y.-H. (2008). Motion Entropy Feature and Its Applications to Event-Based Segmentation of Sports Video. EURASIP Journal on Advances in Signal Processing, Article ID 460913.

[2] D'Orazio, T., Leo, M. (2010). A review of vision-based systems for soccer video analysis. Pattern Recognition 43(8),2911-2926.

[3] Lotfi, E., Pourreze, H.R. (2007). Event Detection and Automatic Summarization in Soccer Video. 4th Iranian Conference on Machine Vision and Image Processing(MVIP 2007), Mashhad, Iran.

[4] Huiyo Zhou, Abdul H. Sadka, Mohammad R. Swash, Jawid Azizi, Umar A. Sadiq. (2010). Feature extraction and clustering for dynamic video summarization. Neurocomputing73 (2010) 1718-1729.

[5] Zhicheng Wei, Xue Yang. (2012). A novel Soccer Video Summarization Model Based on Video Time Density Function. International Journal of Digital Content Technology and its Applications(JDCTA). Volume6,Number 10,June 2012.

[6] Hossam M. Zawbaa, Nashwa El-Bendary, Abou Ella Hassanien. (2011). SVM-based Soccer Video Summarization System. 2011 Third World Congress on Nature and Biologically Inspired Computing. IEEE.

[7] Ekin, A. (2003). Sports Video Processing for Description, Summarization and Search. Phd Thesis. University of Rochester Rochester.

[8] Huang, C. –L., Shih, H.-C., and Chao, C.-Y. (2006). Semantic analysis of soccer video based on goalmouth detection. Asia-Pacific Workshop on Visual Information Precessing.

[9] Zawbaa, H.M., El-Bendary, N., Hassanien, A.E., Schaefer, G., and Yeo. (2011). Logo detection in Broadcast Soccer Videos Using Support Vector Machine. The 2011 Online Conference On Soft Computing in Industerial Applications WWW, (WSC16).

[10] Zhao, Z., Jiang, S., Huang, Q., and Ye, Q.(2006). Highlight summarization in soccer video based on goalmouth detection. Asia-Pacific workshop on Visual Information Processing.

[11] Sangwine, S., Horne, R. (1998). The Colour Image Processing Handbook. London: Chapman&Hall.

[12] Juan Rafael Nuñez, Jacques Facon, Alceu de Souza Brito Junior.(). Soccer Video Segmentation:referee and player detection.

[13] Soille Pierr.(1999). Morphological Image Analysis: Principles and Applications. Berlin: Springer-Verlag.

[14] Monireh-Sadat Hosseini, Amir-Masoud Eftekhari-Monghadam. (2013). Fuzzy rul-based reasoning approach for event detection and annotation of broadcast soccer video. Applied Soft computing 13 846-866.

[15] C.xu, et al., (2008). Using webcast text for semantic event detection in boardcast sport video, IEEE Transaction on Multimedia 10(7) 1342-1355.

[16] F. Coldefy and P.Bouthemy. (2004). Unsupervised soccer video abstracktion based on pitch, dominant color and camera motion analysis. In ACM MM'04, pp. 268-271.

[17] Zhao, Z., Jiang, S., Huang, Q., and Ye, Q.(2006).Highlight summarization in sports video based on replay detection. In multimedia and Expo,2006 IEEE International Conference on, pp 1613-1616.
[18] Q. Ye, Q.Huang. W.Gao, and S. Jiang. (2005). Exciting event detection on broadcast soccer video with mid-level description and incremental. In ACM MM'05, pp.455-458.
[19] D. Tjondronegoro, Y.P.P. Chen, and B.Pham. (2004). The power of play-break for automatic detection and browsing of self-consumable sport video highlights. In MIR'04.
[20] Ren R. and Jose J.M., Football Video Segmentation Based on Video Production Strategy. In: The 27th European Conference on IR Research (ECIR 2005), Santiago de Compostela, Spain, pp. 433-446, 2005.
[21] T. D'Orazio, M. Leo, P. Spagnolo, M. Nitti, N. Mosca, and A. Distante. A visual system for real time detection of goal events during soccer matches. Computer Vision and Image Understanding, 113(5):622–632, 2009.
[22] K. Wan, X. Yan, X. Yu, and C. X. Real-time goal-mouth detection in mpeg soccer video. In ACM international conference on Multimedia (MULTIMEDIA'03), pages 311–314, 2003
[23] Ekin,A.,Tekalp,A.M.,Mehrotra,R.(2003):AutomaticSoccerVideoAnalysis andSummarization. IEEE Transactions on Image processing 12(7)