

A novel approach for semantic event extraction from sports webcast text

Chun-Min Chen · Ling-Hwei Chen

Published online: 18 December 2012

© Springer Science+Business Media New York 2012

Abstract Semantic event extraction is helpful for video annotation and retrieval. For sports video, most previous works detect events by video content itself. Some useful external knowledge has been researched recently. In this paper, we proposed an unsupervised approach to extract semantic events from sports webcast text. First, unrelated words in the descriptions of webcast text are filtered out, and then the filtered descriptions are clustered into significant event categories. Finally, the keywords for each event category are extracted. According to our experimental results, the proposed approach actually extracts significant text events, which can be used for further video indexing and summarization. Furthermore, we also provide a hierarchical searching scheme for text event retrieval.

Keywords Semantic event detection · Webcast text · Information retrieval · Video retrieval

1 Introduction

Due to the huge progress of science and technology in the last few decades, through the Internet, people can use computers or other handheld devices to acquire multimedia resources anytime and anywhere. Among these various resources, videos provide richer information, e.g., image, audio, text. Although there are a lot of videos on the Internet, we are only interested in highlight parts of videos. Accordingly, summarization and retrieval of significant events in different kinds of videos have been hot research topics.

In this paper, we focus on sports videos. A sports video usually lasts more than an hour. There are plenty of sports games held per day in different countries. Even if we narrow the scope down to basketball games in the United States, there still have five to ten games per day in National Basketball Association (NBA). For basketball fans, it is worthy to

C.-M. Chen · L.-H. Chen (✉)

Department of Computer Science, National Chiao Tung University, 1001 University Road, Hsinchu, Taiwan 300, Republic of China
e-mail: lhchen@cc.nctu.edu.tw

C.-M. Chen

e-mail: cmchen@debut.cis.nctu.edu.tw

summarize these games into highlights, so they can track daily results within just a few minutes. Many websites, such as NBA [4], ESPN [2], and Yahoo Sports [3], already make this kind of online service available. However, these online services, such as Game Recap [4], Daily Top 10 and Play of the Day [5], etc., are made by professional film editors and sports reporters by exhaustively watching the sports videos personally. People or fans can only see the unified version through online services. Fans, who want to practice certain basketball skills or imitate specific sports stars, have to spend a lot of time to download the whole game and personally search for certain moves made by certain players. Based on these facts, a personalized sports video summarization and retrieval system is desired by fans to make their own video highlights according to their different tastes and interests.

In video summarization and retrieval, a source video is first clipped into smaller videos representing significant events through a preprocessing, called Semantic Event Detection, which detects events occurred in a video and annotates events with appropriate tags. With finer results of the preprocessing, video summarization and retrieval can be completed efficiently and correctly. Most of existing event detection schemes [13, 14, 18] use video content as their resource knowledge. Chen and Deng [7] analyze low level features (e.g. color, motion, shot) to extract and index events in a basketball video. Some researches [8, 9, 12, 17] use audio-visual (AV) features to do video summarization. Hassan et al. [8] extract AV features and apply Conditional Random Fields (CRFs) based probabilistic graphical model for sports event detection. Kim and Lee [9] build an indexing and retrieving system for a golf video by analyzing its AV content. However, the schemes relying on video content encounter a challenge called semantic gap, which represents the distance between low level video features and high level semantic events. In sports video, two kinds of external knowledge can be used to bridge the gap.

One of the external knowledge is Closed-Caption (CC) [11]. CC is the transcript of speech and sound, and it is helpful for semantic analysis of sports videos. It is mainly used in aid of listening and language learning, but only available in certain videos and certain countries. Because CC completely records the sound in video, it contains a lot of redundant information and usually lacks of structure. The other external knowledge is webcast text. Comparing to CC, webcast text is the online commentary posted by professional announcers and focuses more on sports games. It contains more detail information (e.g., event name, time, player involved, etc.), which is difficult to extract from video content itself automatically. Xu and Chua [15] first use webcast text as external knowledge to assist event detection in soccer video. They proposed a framework that combines internal AV features with external knowledge to do event detection and event boundary identification. But the proposed model is inapplicable to other team sports. Xu et al. [16] apply probabilistic latent semantic analysis (pLSA), a linear algebra–probability combined model, to analyze the webcast text for text event clustering and detection. Based on their observation, the descriptions of the same event in the webcast text have a similar sentence structure and word usage. They use pLSA to first cluster the descriptions into several categories and then extract keywords from each category for event detection. Although they extend pLSA for both basketball and soccer, there are two problems in the approach.

- 1) The optimal numbers of event categories are nine for basketball and eight for soccer in the results, which is determined by minimizing the ratio of within-class similarity and between-class similarity. In fact, there are more event categories for a basketball or soccer game. For example, in a basketball game, many events, such as timeout, assist, turnover, ejected, are mis-clustered into wrong categories or discarded as noises. This may cause side effects degrading and limiting the results of video retrieval.

- 2) After keywords extraction, events can be detected by keywords matching. In Xu et al.'s method, they use the top ranked word in pLSA model as single-keyword of each event category. But in some event categories, the single-keyword match will lead to horrible results. For example, in their method for a basketball game, “jumper” event represents those jumpers that players make. Without detecting “makes” as a previous word of “jumper” in description sentences, the precision of “jumper” event detection is decreased from 89.3 % to 51.7 % in their testing dataset. However, the “jumper” event actually is an event that consists of “makes jumper” event and “misses jumper” event. The former can be used in highlights, and the latter can be used in sports behavior analysis and injury prevention. Accordingly, using single-keyword match is insufficient and some important events will be discarded.

To treat the above-mentioned problems, we propose a method to analyze sports webcast text and extract significant text events. An unsupervised scheme is used to detect events from the webcast text and extract multiple keywords from each event. A data structure is used to store these multiple keywords and to support a hierarchical search system with auto-complete feature for event retrieval. The word “hierarchical” means that a user can get more specific results by querying more keywords and the word “auto-complete” means that the system can give suggested keywords during the query step. According to our experiments, the proposed method keeps more natural event categories and works finely with objective classification results. This provides more flexibility and extensibility while summarizing or retrieving sports videos in different purposes. Our contributions are as follows: 1) detecting semantic events from webcast text in an unsupervised manner; 2) requiring no additional context information analysis; 3) preserving more significant events in sports games; 4) extracting multiple keywords from event categories to support hierarchical searching; 5) providing auto-complete feature for finer retrieval. The rest of the paper is organized as follows. Details of the proposed method are described in Section 2. Experimental results and conclusions are given in Section 3 and Section 4.

2 Proposed method

Webcast text comprises knowledge which is closely related to the game and is easily retrieved from websites. As can be seen in Table 1, it contains time tags, team names, scores, and event descriptions. The format is so organized that we can follow the time flow and understand how the game goes on. Among this well-organized text, it is apparent that event descriptions relate to semantic events the most. Our goal is to analyze event descriptions and automatically extract significant events from them. The block diagram of the proposed method is presented in Fig. 1. It can be seen that we first filter out unrelated words of webcast text and then cluster them into significant events. We store the extracted semantic information with a pair of index tables and build a hierarchical retrieval system by manipulating the two tables. The detail of each block will be described in the following subsections.

2.1 Unrelated words filtering

In webcast text, each description can be considered as an event. It contains many words and may include player name, team name, movement name, and whether the player or the team makes the movement or not. An example is given in Fig. 2, a player named “Peja Stojakovic” failed to make a movement called “10-ft two point shot.”

The number of descriptions in each basketball game is more than four hundred. The descriptions are readable and can be easily categorized into several events by human eyes.

Table 1 An example of basketball webcast text

1st Quarter Summary			
Time	NEW ORLEANS	SCORE	DENVER
12:00	Start of the 1st Quarter		
12:00	Jumpball: Tyson Chendler vs. Nene Hilario (Chris Paul gains possession)	0-0	
11:33	Peja Stojacovic misses 10-ft two point shot	0-0	
11:32	New Orleans offensive rebound	0-0	
11:32	shot clock violation	0-0	
11:19		0-2	Carmelo Anthony makes 10-ft jumper (Kenyon Martin assists)
11:03		0-2	Dahntay Jones personal foul (Chris Paul draws the foul)
10:53	Kenyon Martin blocks Peja Stojakovic's jumper	0-2	
10:52		0-2	Kenyon Martin defensive rebound
10:42		0-4	Kenyon Martin makes 20-ft jumper (Carmelo Anthony assists)
10:23	David West makes 12-ft two point shot	2-4	
10:08		2-6	Nene Hilario makes driving layup (Carmelo Anthony assists)
10:08	Peja Stojakovic shooting foul (Nene Hilario draws the foul)	2-6	
10:08		2-6	Nene Hilario misses free throw 1 of 1
10:07	David West defensive rebound	2-6	

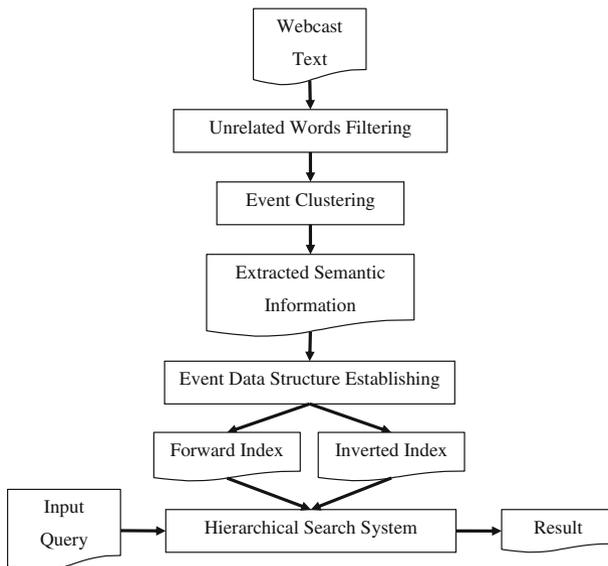


Fig. 1 Block diagram of the proposed method

	12:00	Jumpball: Tyson Chandler vs. Nene Hilario (Chris Paul gains possession)	0-0
Description	11:33	Peja Stojakovic misses 10-foot two point shot	0-0
	11:32	New Orleans offensive rebound	0-0
	11:32	shot clock violation	0-0
	11:19		0-2
Word	11:03		0-2
	10:55	Kenyon Martin blocks Peja Stojakovic's jumper	0-2
	10:52		0-2

Fig. 2 An example to illustrate description and word

But the task is not effortless for computer machines. According to our observations, words in each description consist of three mutually disjoint word sets: 1) stop words, 2) event keywords, and 3) names. Stop words are unrelated to event and should be discarded. Event keywords are closely related to event and should be kept for event detection. Names including team names and player names should be preserved for event annotation. Our objective is to extract event keywords and use these keywords to do event clustering. To achieve the objective, based on a reference stop word list and an online name information, an interactive system is first provided to establish a sports stop word list and an event keyword list. The system will be explained in Sections 2.1.1 and 2.1.2. According to these two lists, for each webcast text, an unrelated word filtering procedure described in Section 2.1.3 is next provided to filter out stop words and to preserve name words. The remaining keywords are then used for event clustering, which will be described in Section 2.2.

2.1.1 Stop words

In information retrieval, there are some words that occur very frequently (e.g. some articles, prepositions, pronouns, be-verbs) and are useless in document matching. These words are called stop words [10]. Due to the uselessness of stop words, filtering out them during both index step and query step can reduce the index size and query processing time. This technique has been used in search engines and can be implemented through predefining a stop word list. For the variety of applications, there is no standard stop word list. Many reference stop word lists [1, 6] have been proposed by using techniques about statistics and probability.

From Table 1, it can be seen that descriptions contain articles (e.g. “the”), prepositions (e.g. “of”), range of shot (e.g. “10-ft”), and points of shot (e.g. “two point”). Some words are details of events which decrease the connections between similar events. With the aid of reference stop lists, articles and prepositions can be easily filtered out from descriptions. However, the range of shot and points of shot are exceptions in reference stop lists. Moreover, in soccer webcast text, due to the relatively larger ground, there are more unrelated words to describe locations where an event happens. For example, right wing, left wing, inside the box, outside the box, left corner, right corner, etc. Accordingly, it is hard to automatically generate a sports stop word list for all kinds of sports. So we will provide an interactive system to establish a sports stop word list.

2.1.2 The proposed interactive system for establishing sports stop word list and event keyword list

As mentioned previously, an interactive system is proposed to establish the sports stop word list and the event keyword list for sports webcast text. First, webcast text descriptions of

several games are taken as training inputs, next some unrelated words are filtered out according to a reference stop word list [1] and a name word list (e.g., online box score in basketball and online player statistics in soccer). And then the system interacts with sports professionals, who will divide the remaining words into a black list and a white list. The black list contains stop words for sports, and the white list contains sports event keywords. Finally the black list is merged into the reference stop word list to get the sports stop word list. The block diagram of the interactive system is presented in Fig. 3.

Our training webcast text is conducted by 41 basketball games and 48 soccer games. After the reference stop words filtering and the name words filtering, the remaining words needed to interactively ask professionals are less than 100 in basketball and less than 200 in soccer. The responses from professionals may take just few minutes.

2.1.3 The proposed unrelated words filtering procedure

Figure 4 shows the block diagram of the proposed unrelated words filtering procedure. For a webcast text, the sports stop word list is first used to filter out unrelated words. Next the event keyword list is used to extract event keywords. Then the words with uppercase beginning in the remaining words are considered as reserved names for further indexing. According to our experiment results, the unrelated words filtering works well both in basketball and soccer.

2.2 Event clustering

After filtering, each description is reduced and almost exactly describes an event; for example, “misses shot” represents a missed shot. So a matching function is provided to cluster these filtered descriptions into event categories.

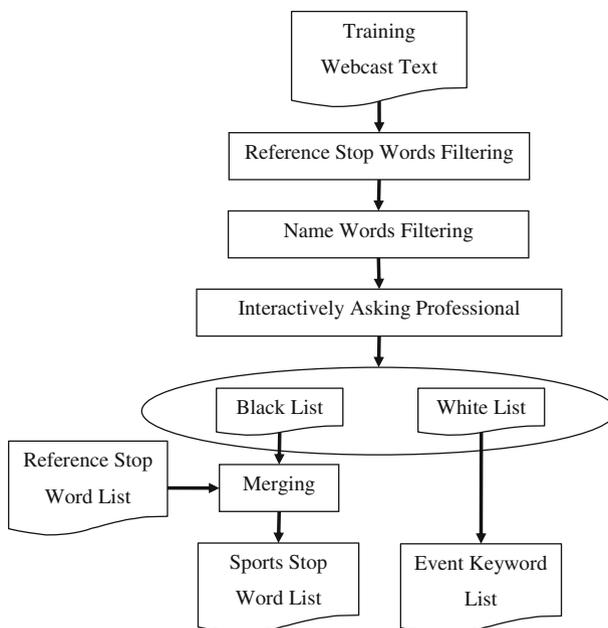


Fig. 3 The block diagram of the interactive pre-training system

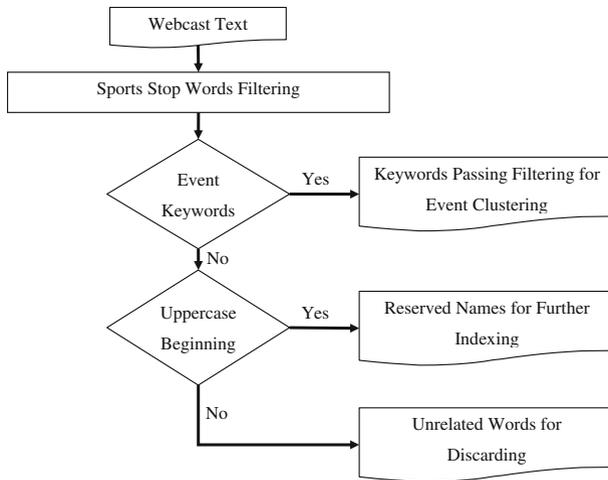


Fig. 4 Block diagram of unrelated words filtering procedure

Filtered descriptions can be represented as $FD = \{fd_1, fd_2, \dots, fd_N\}$, and event categories can be represented as $C = \{C_1, C_2, \dots, C_K\}$, where N denotes the number of descriptions in a game and K denotes the number of categories that the clustering step produces. Since a filtered description consists of some words, it can be considered as a set of words. Note that the number of keywords of an event category is not restricted to be single in our method. The matching function is defined as

$$Text_Match(x, y) = \begin{cases} 1, & \text{if } x = y \\ 0, & \text{otherwise,} \end{cases} \tag{1}$$

where x and y are two sets of words. Each filtered description, fd_i , can be clustered into one category based on the following function

$$\begin{aligned} Clustering(fd_i) &= \arg \max_m \{Text_Match(fd_i, Keywords(C_m)), m = 1, \dots, K\}, i \\ &= 1, \dots, N, \end{aligned} \tag{2}$$

where $Keywords(C_m)$ denotes the multiple-keywords set of category C_m . $Clustering(fd_i)=j$ means that description fd_i is clustered into category C_j . In order to avoid zero matching in (2), a flag function to examine whether the situation happens is defined as

$$Flag(fd_i) = \max_m \{Text_Match(fd_i, Keywords(C_m)), m = 1, \dots, K\}, i = 1, \dots, N. \tag{3}$$

The detail of the proposed clustering algorithm is given below.

2.3 Clustering algorithm

Step0: Initialization: Given $FD = \{fd_1, fd_2, \dots, fd_N\}$.
 Set $K=1, Clustering(fd_1)=1, Keywords(C_1)=fd_1, i=2$.

Step1: Cluster the description fd_i according to Functions (1), (2), and (3). The procedure includes the following pseudo code:

For $m = 1$ to K , use Function (1) to calculate $TMfd_{im}$

$$TMfd_{im} = \text{Text_Match}(fd_i, \text{Keywords}(C_m));$$

Let

$$\text{Flag}(fd_i) = \max_{m=1, \dots, K} \{TMfd_{im}\};$$

if ($\text{Flag}(fd_i) = 0$) then begin // fd_i cannot be clustered into any existing class

// create a new class for fd_i

$$K = K + 1;$$

$$\text{Keywords}(C_K) = fd_i;$$

$$\text{Clustering}(fd_i) = K;$$

else

// fd_i is clustered into one of the existing classes

Use Function (2) to calculate $\text{Clustering}(fd_i)$ as

$$\text{Clustering}(fd_i) = \arg \max_m \{TMfd_{im}\};$$

end

Step2: If any of the descriptions in FD is not clustered yet, set $i=i+1$ and go to Step1 for next iteration. Otherwise, end of iterations.

Once the clustering algorithm is completed, the filtered descriptions are clustered into event categories, and keyword extraction is done by using each keyword set as multiple keywords of the event. At the meantime, semantic event detection is accomplished. Then two data structures are built to recommend users for further queries and to support the hierarchical search.

2.4 Hierarchical search system

Figure 5 gives an example to show the concept of the proposed hierarchical search system. First, a user can query by one word to get rough results. Then he can continually query by more words to get into deeper levels for finer results. Here we implement the system by establishing a pair of index tables and manipulating them back and forth.

Here we build a forward index table and an inverted index table. The former records mappings from descriptions to event keywords, and the latter stores mappings from keywords to descriptions. Note that the forward index table is established automatically after applying the unrelated words filtering procedure. Based on the forward index table, the inverted index table can be established by sequentially scanning event keyword set of each description. An example is given in Fig. 6 to do clearer explanation. Suppose we have five descriptions as shown in Fig. 6a. After applying unrelated words filtering procedure to each description, we can obtain Fig. 6b. By scanning each row in Fig. 6b, for each row, we can obtain a description index (DI) and the corresponding event keyword set (EKS). Then DI is linked to each keyword in EKS. After scanning all rows sequentially in Fig. 6b, c is established. Both inverted index table and forward index table are referred to achieve the hierarchical search system. The inverted index table is used for returning query results by intersecting those description sets mapped by query keywords. The forward index is originally just an intermediate, but reused in our method for providing suggested query keywords, i.e. auto-complete feature.

In our system, a query is considered as a set of multiple words. The hierarchical feature means that a user can get more general results by querying fewer words or get more specific result by querying more words; for example, the results of querying “jumper” are those descriptions having the keyword “jumper”, and the results of querying “jumper makes” are those descriptions having both “jumper” and “makes.”

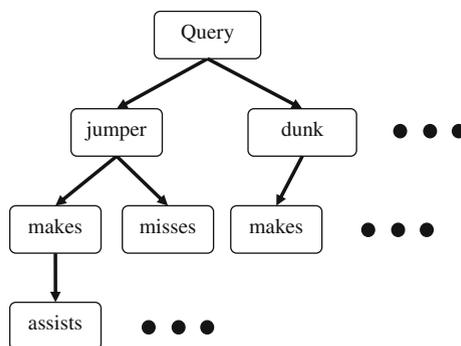


Fig. 5 An example to illustrate the concept of the proposed hierarchical search system

Webcast Text

Index of Description	Description
D1	Peja Stojakovic misses 10-foot two point shot
D2	David West misses jumper
D3	Peja Stojakovic makes 19-foot two point shot
D4	Trevor Ariza makes 19-foot jumper
D5	David West makes 17-foot jumper (Chris Paul assists)

(a) Descriptions and their indices.

Forward Index

Index of Description	Event Keyword Set
D1	misses, shot
D2	misses, jumper
D3	makes, shot
D4	makes, jumper
D5	assists, makes, jumper

(b) Mappings from description indices to event keywords.

Inverted Index

Keywords	Indices of Description Set
assists	D5
jumper	D2, D4, D5
makes	D3, D4, D5
misses	D1, D2
shot	D1, D3

(c) Mappings from keywords to description indices.

Fig. 6 An example to illustrate the data structure for hierarchical search

The query result is the intersection of description sets obtained through the keywords of query in the inverted index list. For providing suggested query keywords, the resulting intersection set is then used as another query for the forward index list. The keyword set of each description in the resulting intersection set are extracted. Finally, the union of all extracted keyword sets is considered as the suggested query keywords. The detail algorithm of the proposed search system is given below.

2.5 Hierarchical search algorithm

- Step1: A user types several query words.
- Step2: Look up the inverted index and get description sets mapped by the query words. Intersect these description sets to obtain a query result.
- Step3: Look up the forward index and get keyword sets mapped by the query result.
- Step4: Output the union set of these word sets. The user selects some keywords from output as query words. Perform Step2 and output the query result.

Here, we use Fig. 6 as an example to do explanation. Assume that a user types a query {jumper}, the system will look up the inverted index list and get a temporary result set {D2,

D4, D5}. Then, the system will look up the forward index list and recommend the user {assists, jumper, makes, misses}, i.e. the union set of {jumper, misses}, {jumper, makes}, and {assists, jumper, makes}. If the user changes his query to {jumper, makes}, the system will return {D4, D5}, i.e. the intersection set of {D3, D4, D5} and {D2, D4, D5}. Therefore, a powerful hierarchical search system with query recommendation function is built.

3 Experimental results

In most search systems, statistical analysis such as receiver operating characteristic (ROC) analysis or recall-precision is used to evaluate the performance. Through the analysis, the system degradation caused by misclassification can be estimated. However, as mentioned in Section 2.2, we cluster descriptions by an exactly matching function, so there is no misclassified event in our system. This means that both precision and recall rates of the proposed method are 100 %.

Researches aimed at detecting text events from webcast text are few. Xu and Chua [15] modeled webcast text as external knowledge in detecting events from football and soccer. The evaluation of the fusion video event detection was presented, but that of webcast text analysis alone was not. Xu et al. [16] proposed a framework to analyze webcast text and videos independently and align them through game time. According to the framework, the performance of video event detection mainly depends on webcast text analysis. Here we compare our method with Xu et al.'s work.

Our experiments are conducted by 25 NBA 2009–2010 games and 41 NBA 2008–2009 postseason games. The former are used as training database, and the latter are used as testing database to examine the reliability of the proposed method. We also collect 68 UEFA Champions League 2010–2011 soccer games, where 20 of them are used as training database and the other 48 are used as testing database. The webcast text from 134 games is acquired from ESPN website [2]. As can be seen in Table 2, hundreds of descriptions in a game are clustered into, in average, 44 semantic event categories for basketball and 20 semantic event categories for soccer.

From Xu et al.'s previous work, the pLSA, the optimal number of event categories is nine for basketball and eight for soccer. The top three keywords of each category are selected by a conditional probability. They use the top ranked keyword as single keyword during event detection. We map the top three results of pLSA to our multiple keywords categories in Tables 3 and 4. In Table 4, because “attempt” is chosen as a member of black list in the interactive system, we use “shot” as the single-keyword match for mappings from soccer events in pLSA to those in the proposed method. The words “missed” and “misses” refer to the same verb (e.g., miss) and have the same meaning in descriptions. We consider these two words as the same and use “missed(misses)” as their common representative. In order to achieve fine performance in detecting semantic events, Xu et al. not only use keywords

Table 2 Average number of sports event categories in 25 basketball training data and 20 soccer training data

	Mean	Variance	Standard deviation
Basketball	44.08	9.08	3.01
Soccer	19.85	5.40	2.32

Table 3 Mappings of basketball event categories from pLSA to the proposed method

Xu et al.'s Method (pLSA)	Proposed Method (Categories with Multiple Keywords)	
Category	Ranked Keywords	
Shot	shot pass bad	makes shot, misses shot
Jumper	jumper foot misses	jumper misses, jumper makes, assists jumper makes
Layup	layup driving blocks	layup makes, layup misses, driving layup makes, assists layup makes
Dunk	dunk makes misses	dunk makes, assists dunk makes, dunk makes slam, driving dunk makes, dunk misses
Block	blocks shot assists	blocks layup, blocks jumper, blocks driving layup, blocks hook shot, blocks shot, blocks dunk, blocks layup, blocks jumper, blocks driving layup, blocks hook shot, blocks shot, blocks dunk
Rebound	rebound defensive offensive	defensive rebound, offensive rebound
Foul	foul draw personal	draws foul shooting, draws foul personal, draws foul offensive, ball draws foul loose, foul technical, defense foul illegal person, draws flagrant foul type
Free throw	throw free makes	free makes throw, free misses throw
Substitution	enters game timeout	enters
N/A		bad pass, bad pass steals, bad lost steals, full timeout, official timeout, turnover, traveling, ejected, double dribble, defense illegal, clock shot violation

detection in description sentences, but also analyze context information in them. For example, in basketball, the top ranked keyword “jumper” is detected as “Jumper” event only if its previous word is “makes,” and other sentences containing word “jumper,” e.g., Kenyon Martin misses 22-ft jumper, are discarded. However, these discarded events are actually semantic events and can be valuable for further research, e.g., sports posture analysis, injury prevention, special highlight, etc. It can be seen from Tables 3 and 4 that every category of pLSA is mapped to several different semantic events of the proposed method. These several events are related but somehow different. For example, in basketball, “jumper misses” describes that a jumper is missed while “jumper makes” describes that a jumper is made successfully. In soccer, “blocked shot” describes that a shot attempt is blocked by an opponent while “missed(misses) shot” describes that a shot attempt is missed

Table 4 Mappings of soccer event categories from pLSA to the proposed method

Xu et al.'s Method (pLSA)	Proposed Method (Categories with Multiple Keywords)
Category	Ranked Keywords
Corner	corner, assisted corner saved shot, corner goal penalty shot, corner saved shot, conceded corner goal, assisted corner goal shot, assisted corner missed(misses), bottom corner goal shot, corner missed(misses) shot, assisted corner missed(misses) shot, corner free kick missed(misses) shot, assisted corner saved, corner free goal kick shot
Shot	blocked shot, assisted missed(misses) shot, assisted blocked shot, assisted goal saved shot, missed(misses) shot, assisted corner saved shot, assisted shot, right corner goal penalty shot, corner saved shot, assisted corner goal shot, corner footed goal shot, corner missed(misses) shot, goal saved shot, free kick shot, assisted goal shot, free kick missed(misses) shot, assisted corner missed(misses) shot, corner free kick missed(misses) shot, goal penalty saved shot, corner free goal kick shot, goal penalty shot
Foul	foul, card foul yellow, foul penalty, card foul
Card	card foul yellow, card yellow
Free kick	free kick, free kick shot, free kick missed(misses) shot, corner free kick missed (misses) shot, corner free goal kick shot
Offside	offside
Substitution	replaces substitution, injury replaces substitution
Goal	assisted goal saved shot, corner goal penalty shot, assisted corner goal, assisted corner goal shot, corner goal shot, goal saved shot, assisted goal shot, assisted goal saved, goal penalty saved shot, goal saved, goal, corner free goal kick shot, goal penalty shot, assisted goal
N/A	injury, assisted missed(misses), assisted blocked, penalty, assisted

by the kicker himself. Hence, misclassifying or discarding these events decreases the precision and recall rates. However, in our method, the precision and recall rates are both 100 %. With the support of hierarchical search system, we can query multiple keywords for more specific events, which is even better than pLSA with context information. Tables 3 and 4 also show those semantic event categories which are unavailable in Xu et al.'s method, but can be detected in our method, e.g., steal, timeout, turnover for basketball and injury, blocked, penalty for soccer. These semantic events are important for special highlights or injury prevention, and should not be ignored or misclassified. So, the proposed method is superior to pLSA.

Here we want to examine the reliability of the proposed method. For basketball, 25 NBA 2009–2010 games are taken as training data. After processing all the training data and

Table 5 Occurrences of exception basketball events from 41 testing games

Exception events	18679 basketball descriptions Number (Percentage)
10 s	3 (0.02 %)
backcourt	7 (0.04 %)
called full timeout	1 (0.01 %)
driving dunk misses	2 (0.01 %)
dunk misses slam	2 (0.01 %)
away ball draws foul	5 (0.03 %)
misses pointer	7 (0.04 %)
flagrant free misses throw	1 (0.01 %)
blocks driving dunk	1 (0.01 %)

gathering the extracted semantic events, we collect the union of these semantic events as a sample set with cardinality 82. Then we process the testing data, which are collected from 41 NBA 2008–2009 postseason games, and examine whether all the semantic events extracted from testing data are listed in the sample set or not. For soccer, we use 20 UEFA Champions League soccer games as training data and 48 UEFA Champions League soccer games as testing data. According to our examination, with sparse exceptions, almost all the semantic events extracted from testing data can be found in the sample set. Tables 5 and 6 show all exception events which are quite rare. These exceptions may be caused by different writing styles or some rarely happened events, and can still be collected in an interactive way if necessary. Therefore, the proposed method is very stable.

4 Conclusions and future work

In this paper, we have proposed an unsupervised approach for semantic event extraction from sports webcast text and made some contributions: 1) detecting semantic events from webcast

Table 6 Occurrences of exception soccer events from 48 testing games

Exception events	5727 soccer descriptions Number (Percentage)
card	6 (0.10 %)
corner penalty saved shot	2 (0.03 %)
missed(misses)	3 (0.05 %)
goal shot	1 (0.02 %)
assisted corner missed shot	1 (0.02 %)
missed shot	1 (0.02 %)
shot	4 (0.07 %)
corner missed(misses)	3 (0.05 %)
corner saved	2 (0.03 %)
assisted corner	1 (0.02 %)
blocked	1 (0.02 %)

text in an unsupervised manner; 2) requiring no additional context information analysis; 3) preserving more significant events in sports games; 4) extracting multiple keywords from event categories to support hierarchical searching; 5) providing auto-complete feature for finer retrieval. According to experimental results, the proposed method extracts significant semantic events from basketball and soccer games and preserves those events that are ignored or misclassified by previous work. Furthermore, the proposed method is reliable.

Because we have proposed a great filtering step, it is believed that we may extend our approach to other free-styled basketball webcast text in the near future. We will try to apply our method to other sports, e.g., football, baseball, etc. In the further future, we will combine text information with video to build a search-and-query system for sports video retrieval.

Acknowledgment This work is supported in part by National Science Council of Republic of China under grant NSC-100-2221-E-009-140-MY2.

References

1. [Online] Available: <http://armandbrahaj.blog.al/2009/04/14/list-of-english-stop-words/>
2. [Online] Available: <http://espn.go.com/nba/>
3. [Online] Available: <http://sports.yahoo.com/video/>
4. [Online] Available: <http://www.nba.com/video/highlights/>
5. [Online] Available: <http://www.nba.com/video/topplays/>
6. [Online] Available: <http://www.textfixer.com/resources/common-english-words.txt>
7. Chen YH, Deng LY (2011) Event mining and indexing in basketball video. Int. Conf. on Genetic and Evolutionary Computing (ICGEC)
8. Hassan E, Chaudhury S, Gopal M, Garg V (2011) A hybrid framework for event detection using multimodal features. Int. Conf. on Computer Vision Workshops (ICCV Workshops)
9. Kim HG, Lee JH (2011) Indexing of player events using multimodal cues in golf videos. IEEE Int. Conf. on Multimedia and Expo (ICME)
10. Manning C, Raghavan P, Schütze H (2008) Introduction to Information Retrieval. Cambridge University Press, pp. 27–28
11. Nitta N, Babaguchi N, Kitahashi T (2005) Generating semantic descriptions of broadcasted sports video based on structure of sports games and TV programs. *Multimed Tools Appl* 25:59–83
12. Schreer O, Feldmann I, Mediavilla IA, Concejero P, Sadka AH, Swash MR, Benini S, Leonardi R, Janjusevic T, Izquierdo E (2010) RUSHES - an annotation and retrieval engine for multimedia semantic units. *Multimed Tools Appl* 48:23–49
13. Shen J, Tao D, Li X (2008) Modality mixture projections for semantic video event detection. *IEEE Trans Circ Syst Video Technol* 18(11):1587–1596
14. Shyu M, Xie Z, Chen M, Chen S (2008) Video semantic event/concept detection using a subspace-based multimedia data mining framework. *IEEE Trans Multimed* 10(2):252–259
15. Xu H, Chua T (2004) Fusion of audio-visual features and external knowledge for event detection in team sports video. Proc. Workshop on Multimedia Information Retrieval (MIR'04)
16. Xu C, Zhang Y, Zhu G, Rui Y, Lu H, Huang Q (2008) Using webcast text for semantic event detection in broadcast sports video. *IEEE Trans Multimed* 10(7):1342–1355
17. Zhou H, Sadka AH, Swash MR, Azizi J, Sadiq UA (2010) Feature extraction and clustering for dynamic video summarisation. *Neurocomputing* 73:1718–1729
18. Zhu X, Wu X, Elmagarmid A, Feng Z, Wu L (2005) Video data mining: semantic indexing and event detection from the association perspective. *IEEE Trans Knowl Data Eng* 17(5):665–677



Chun-Min Chen was born in Taipei, Taiwan, Republic of China on September 19, 1983. He received the B.S. degree in Computer Science and Information Engineering from National Taiwan University of Science and Technology, Taipei, Taiwan in 2005. He is now a Ph.D. Candidate of College of Computer Science at National Chiao Tung University, Hsinchu, Taiwan. His major research interests include image processing, image/video retrieval, pattern recognition.



Ling-Hwei Chen was born in Changhua, Taiwan, in 1954. She received the B.S. degree in Mathematics and the M.S. degree in Applied Mathematics from National Tsing Hua University, Hsinchu, Taiwan in 1975 and 1977, respectively, and the Ph.D. degree in Computer Engineering from National Chiao Tung University, Hsinchu, Taiwan in 1987.