An Enhanced Query Model for Soccer Video Retrieval Using Temporal Relationships

Shu-Ching Chen[1], Mei-Ling Shyu[2], Na Zhao[1]
[1]*School of Computer Science, Florida International University, Miami, FL 33199, USA*
[2]*Department of Electrical & Computer Engineering, University of Miami, Coral Gables, FL 33124, USA*
[1]*{chens, nzhao002}@cs.fiu.edu,* [2]*shyu@miami.edu*

1. Introduction

Quite a few research efforts have been put on the sports video analysis. Specifically, there are systems developed for video annotation, summarization, or retrieval. Some of the existing works, such as IBM's MEPG-7 authoring system [6], deploy the event-based indexing, which is not suitable for automatic event detection target since it may require users' labor to manually annotate the video events and their boundaries. In some other approaches, such as SMOOTH [4] and GOALGLE [5], only the semantic queries and some simple temporal queries are supported, and furthermore, their visualized query interfaces are not powerful to perform all the possible queries. In addition, some of the proposed query languages including the one proposed in [3] are not practical because they assume the video events, objects and the related properties are all available, which is not true in the real world.

The focal goal of our research is to develop a general framework which can automatically analyze the sports video, detect the sports events, and finally offer an efficient and user-friendly system for sports video retrieval. In our earlier work [2], a novel multimedia data mining technique was proposed for automatic soccer event extraction by adopting multimodal feature analysis. Until now, this framework has been performed on the detection of goal and corner kick events and the results are quite impressive. Correspondingly, in this work, the detected video events are modeled and effectively stored in the database. A temporal query model is designed to satisfy the comprehensive temporal query requirements, and the corresponding graphical query language is developed. The advanced characteristics make our model particularly well suited for searching events in a large scale video database.

2. Temporal Query Model

Based on our observation, most of the significant events in sports video are captured within one shot, where a shot represents a video sequence that consists of continuous video frames for one camera action. Therefore, tedious manual work can be avoided by performing event detection based on the automatically segmented shots. In *SoccerQ*, video shots are stored, managed, and retrieved along with their corresponding features and meta-data.

The query language of *SoccerQ* provides three basic statements for retrieving the information:

```
select video from search_space [where condition];
select shot from search_space [where condition];
select variable from search_space [where condition];
```

The "`select`" clause specifies the search target, which could be videos, shots, or the values of some stored variables (e.g., the names of the participated soccer teams). The search space in the "`from`" clause can represent all the videos in the database, a list of selected videos, or one specific video file. The condition in the "`where`" clause can represent a single criterion, and a conjunction or disjunction of several rules.

In general, different soccer events may hold a temporal relationship between the cause and effect. Consequently, the users may want to retrieve the events based on some specific temporal relations. Most of the existing video query systems only model the relationships between two events based on Allen's 13 temporal relations [1], which is not a complete and convenient solution to handle event retrieval because sometimes the ranges of the events need to be considered. For example, given a query as follow:

Query 1: "Find all the event pairs (A, B) where event B occurs in the range of 10 minutes around event A."

In this case, there are several possibilities for the temporal relationship between A and B. As shown in Figure 1, both event B1 and event B2 satisfy the requirement but their relationships with event A are quite different – B1 *meets* A, while B2 *after* A.

Figure 1. Temporal relationship representation of Query 1

Therefore, in our approach, the relative temporal queries are extended by composing the following components – the absolute duration of video V; the relative duration and temporal position of event E; and the relative duration and temporal position of range R_E where E may occur. Assume that for a temporal interval T, T_0 and T_f represent the time positions of T's start point and end point, respectively. Accordingly, the temporal query criteria can be categorized into 5 different models. The parameters here are the variables that the users can specify for further quantifying purpose.

- Temporal relationship between event A and video V;
  A θ V, where θ ∈ {*starts*, *during*, *finishes*}
  Parameters: A_0, A_f
- Temporal relationship between the range of event A and video V;
  R_A θ V, where θ ∈ {*starts*, *during*, *finishes*}
  Parameters: R_{A0}, R_{Af}
- Temporal relationship between event A and event B;
  A θ B, where θ ∈ {*before*, *meets*, *equals*, *met by*, *after*}
  Parameters: $B_0 - A_0$, $B_f - A_f$
- Temporal relationship between event A and the temporal range of event B;
  A θ R_B, where θ ∈ {*before*, *meets*, *overlaps*, *finished by*, *contains*, *starts*, *equals*, *started by*, *during*, *finishes*, *overlapped by*, *met by*, *after*}
  Parameters: $R_{B0} - A_0$, $R_{Bf} - A_f$
- Temporal relationship between the temporal range of event A and the temporal range of event B.
  R_A θ R_B, where θ ∈ {*before*, *meets*, *overlaps*, *finished by*, *contains*, *starts*, *equals*, *started by*, *during*, *finishes*, *overlapped by*, *met by*, *after*}
  Parameters: $R_{B0} - R_{A0}$, $R_{Bf} - R_{Af}$

These relationships can be precisely described in terms of the temporal distances between their temporal starting points and ending points, where three units can be provided for the time interval computation – minutes, seconds, and shots.

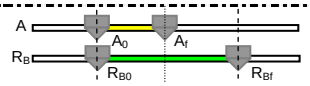
3. System Demonstration

A soccer video retrieval system named *SoccerQ* is developed to support both the basic queries and the relative temporal queries. The multi-threaded client/server architecture is deployed such that multiple requests from diverse users can be handled simultaneously. The client side application integrates both soccer event query and video browsing panels. The server side database engine performs the computation intensive functions including query optimization, processing, video supply, etc.

In particular, a graphical query language is designed for specifying the relative temporal queries. The basic idea is to compose the temporal query filters, sentential operators, and macro operators together to fulfill the requirements. Basically, the user is allowed to specify the search target and search space. After that, different events and the temporal relationship model types can be chosen. The double-thumbed sliders are utilized for the adjustments of event position, duration, and range specification. In addition, the sentential operators are provided such that different query rules can be combined. Given the following query, the corresponding visual query specifications are listed in Table 1.

Query 2: "Find all the *corner kick* shots from all the female soccer videos where the *corner kick* resulted in a *goal event* occurring in 2 minutes."

Table 1. Example mappings to the graphical query language

Events and Relationships:	Parameters (unit)	Graphical Query Filter
Query 2: A θ R_B, θ ="*starts*" A = "Corner kick" B = "Goal"	R_{B0}-A_0 = 0 R_{Bf}-A_f = 2 (minutes)	

As shown in Figure 2, the query criteria are specified in the upper-right graphical query panel while the search space is defined in the upper-left panel. The key frames of four result shots of Query 2 are displayed in the video browsing panel. The video shot can be displayed by double clicking the corresponding key frame.

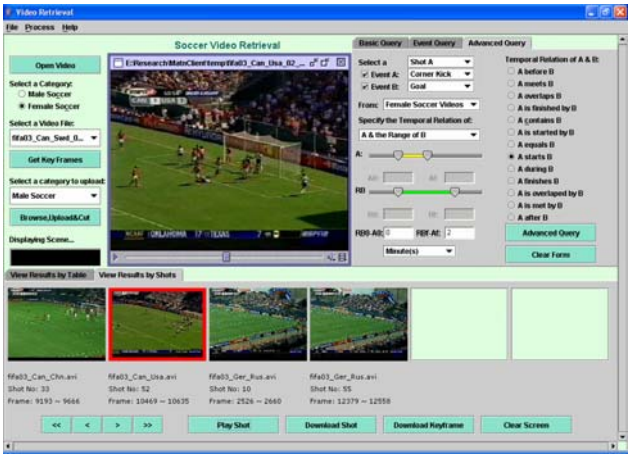


Figure 2. SoccerQ video retrieval interface for Query 2

4. Acknowledgement

For Shu-Ching Chen, this research was supported in part by NSF EIA-0220562 and NSF HRD-0317692. For Mei-Ling Shyu, this research was supported in part by NSF ITR (Medium) IIS-0325260.

5. References

[1] J. F. Allen, "Maintaining Knowledge about Temporal Intervals," *Communications of the ACM*, Vol. 26, Nov. 1983, pp. 832-843.

[2] S.-C. Chen, M.-L. Shyu, M. Chen, and C. Zhang, "A Decision Tree-based Multimodal Data Mining Framework for Soccer Goal Detection," *Proc. of IEEE International Conference on Multimedia and Expo*, June 2004, Taipei, Taiwan, R.O.C.

[3] M. E. Donderler, O. Ulusoy and U. Gudukbay, "Rule-Based Spatio-Temporal Query Processing for Video Databases," *VLDB Journal*, Vol. 13, No. 1, January 2004, pp. 86-103.

[4] H. Kosch, L. Böszörményi, A. Bachlechner, B. Dörflinger, C. Hanin, C. Hofbauer, M. Lang, C. Riedler, and R. Tusch, "SMOOTH - A Distributed Multimedia Database System," *VLDB'2001*, Rome, Italy, pp. 713-714.

[5] C.G.M. Snoek and M. Worring, "Multimedia Event based Video Indexing using Time Intervals," *IEEE Trans. on Multimedia*, 2005 (in press).

[6] IBM TRL's MPEG-7 Authoring System.
http://www.trl.ibm.com/projects/digest/authoring_e.htm