

A Novel Approach for Similarity Measure Schemes Based on Multiple Moving Objects in Video Databases

Choon-Bo Shim¹, Chang-Sun Shin¹, DongGook Park¹, and Won-Ho So²

¹ School of Information & Communication Engineering

² Dept. of Computer Education, Suncheon National University, Suncheon, Jeonnam
540-742, South Korea

{cbsim, csshin, dgpark, whso}@suncheon.ac.kr

Abstract. The general aim of this paper is to study the spatio-temporal modeling techniques which can efficiently represent multiple moving objects' in video databases. The traditional schemes only consider direction property, time interval property, and spatial relationship property for modeling moving objects' trajectories. But, our scheme also takes into account on distance property, conceptual location information, and related object information so that we may improve a retrieval accuracy to measure a similarity between two moving objects as well as them. As its application, we implement the Content- and Semantic-based Soccer Video Retrieval (CS²VR) system by using MS Visual C++ and DirectX for indexing and searching on soccer video data. The CS²VR helps users to easily extract the trajectory information of soccer ball from soccer video data semi-automatically as well as to conveniently retrieve the results acquired by sketching query trajectory with mouse button.

1 Introduction

The initial research issues on the content-based video retrieval have highly concentrated on data representation schemes which can efficiently model content itself extracted from video data [1-5]. However, for handling a large amount of multimedia data, it is required to provide schemes with good retrieval performance on a variety of user queries. Thus, we propose a new spatio-temporal modeling technique which can efficiently represent multiple moving objects in video databases. The traditional schemes only consider direction property, time interval property, and spatial relationship property for modeling moving objects' trajectories. However, our scheme also takes into account on distance property, conceptual location information, and related object information (e.g. player name having a soccer ball) so that we may improve retrieval accuracy to measure a similarity between two moving objects as well as them. Therefore, the proposed spatio-temporal scheme can support content-based retrieval using moving objects' trajectories as well as semantics-based retrieval using concepts which are acquired through the conceptual location information of moving objects. As its application, we design and implement the Content- and Semantic-based Soccer Video Retrieval (CS²VR) system. Finally, in our performance study, our scheme yields substantially better retrieval performance compared to existing related work in term of retrieval effectiveness.

This paper is organized as follows: Section 2 presents a new spatio-temporal modeling scheme for representing multiple moving objects; Based on our modeling scheme, the similarity measure functions for multiple moving objects is described in Section 3; For its applications, the Content-based Soccer Video Retrieval (CS²VR) system is presented in Section 4; and Section 5 covers conclusion with brief summary.

2 Spatio-temporal Modeling Scheme

Since video data have both spatial and temporal information, they should consider both spatial [6] and temporal [7] relationships to represent moving objects in an effective way. For this, we propose a new spatio-temporal representation scheme for modeling the multiple trajectories among several moving objects. At first, spatio-temporal representation scheme for single trajectory composed of only one moving object is introduced.

Definition 1. Motion property information for a moving object A over all the time intervals, MPS(A), is defined as follows:

$$MPS(A) = \{M_i(A) \mid i = 0, \dots, n-1\} = \{M_0(A), M_1(A), \dots, M_{n-1}(A)\}$$

Definition 2. A motion property for a moving object A over time interval I_i, M_i(A), is defined as follows:

$$M_i(A) = (R_i(A), D_i(A), I_i(A))$$

Here, R_i(A) is a moving direction over time interval I_i (= [t_i, t_{i+1}]) and is represented as a real angle with a range of 0 to 360 degree. D_i(A) is a moving distance over I_i and is described as an absolute Euclidean distance or a relative distance. I_i(A) means a time interval from the start time to the end time while the moving object A is moving.

Definition 3. Stationary property information for a moving object A over all the time instances, SPS(A), is defined as follows:

$$SPS(A) = \{S_i(A) \mid i = 0, \dots, n\} = \{S_0(A), S_1(A), \dots, S_n(A)\}$$

Definition 4. A stationary property for a moving object A at time t_i, S_i(A), is defined as follows:

$$S_i(A) = ([L_i(A)], [O_i(A)])$$

Here, L_i(A) is a location information of the moving object A. The location information describes a real location in coordinates or a semantic-based location according to a real application, e.g., penalty area or goal area in the soccer game. O_i(A) is an object information related with the moving object A, e.g., actor or owner having the moving object A. Here, [] means an optional operator.

For the single trajectory of a moving object A, it is possible to combine a motion property (Definition 1 and 2) with a stationary property (Definition 3 and 4). As a result, the information of a single trajectory is defined as follows.

Definition 5. For a given ordered list of time interval I_0, I_1, \dots, I_{n-1} , the single trajectory information of a moving object A, $ST(A)$, is defined as follows:

$$ST(A) = MPS(A) + SPS(A)$$

We define multiple trajectories as the trajectories of two or more moving objects. However, since the multiple trajectories can be represented by the combination of the trajectory between two moving objects, we first define a relationship trajectory between two objects.

Definition 6. Let at least one of object A and object B be a moving object. Motion property information for A and B over all the time interval, $MPM(A, B)$, is defined as follows:

$$MPM(A, B) = \{M_i(A, B) \mid i = 0, \dots, n-1\} = \{M_0(A, B), M_1(A, B), \dots, M_{n-1}(A, B)\}$$

Definition 7. Let at least one of object A and object B be a moving object. A motion property for A and B over time interval $I_i ([t_i, t_{i+1}])$, $M_i(A, B)$, is defined as follows:

$$M_i(A, B) = (D_i(A, B), I_i(A, B))$$

Here, $D_i(A, B)$ is a relative moving distance of A to B over I_i and is ranged from 0 to 100. That is, $D_i(A, B)$ is 50 in case the moving distance of A is the same as that of B. $D_i(A, B)$ is ranged from 51 to 100 in case the moving distance of A is greater than that of B while it is near to 0 as the moving distance of A is less than that of B. $I_i(A, B)$ is the same as single trajectory.

Definition 8. Let at least one of object A and object B be a moving object. Stationary property information for A and B over all the time instances, $SPM(A, B)$, is defined as follows:

$$SPM(A, B) = \{S_i(A, B) \mid i = 0, \dots, n\} = \{S_0(A, B), S_1(A, B), \dots, S_n(A, B)\}$$

Definition 9. Let at least one of object A and object B be a moving object. A stationary property for A and B at time t_i , $S_i(A, B)$, is defined as follows:

$$S_i(A, B) = ([L_i(A)], [O_i(A)], ([L_i(B)], [O_i(B)], T_i(A, B), R_i(A, B))$$

Here, $L_i(A)$ and $L_i(B)$ are the location information of moving object A and B, respectively. $O_i(A)$ and $O_i(B)$ are the actors having moving objects A and B, respectively. $T_i(A, B)$ is a spatial (topological) relations on XY-coordinates from A to B, being represented as one of seven topological relations operator : FA(FarAway), DJ(DisJoint), ME(MEet), OL(OverLap), CL(is-inCLuded-by), IN(INclude), and SA(SAme). Finally, $R_i(A, B)$ means a directional relations from A to B and is ranged from 0 to 360 degree.

For a relationship trajectory between A and B, it is possible to combine a motion property (Definition 6 and 7) with a stationary property (Definition 8 and 9). As a result, the relationship trajectory information is defined as follows.

Definition 10. Let at least one of object A and object B be a moving object. For a given ordered list of time interval I_0, I_1, \dots, I_{n-1} , the relationship trajectory information between A and B, $RT(A, B)$, is defined as follows:

$$RT(A, B) = MPM(A, B) + SPM(A, B)$$

Based on Definition 5 and 10, the multiple trajectory information of two or more moving objects, $MT(A_1, A_2, \dots, A_n)$, can be represented by a combination of the relationship trajectory information (RT) and the single trajectory information (ST).

Definition 11. Among objects A_1, A_2, \dots, A_n , let i be the number of moving objects and j be the number of stationary objects, i.e., $n=i+j$. The multiple trajectory information of A_1, A_2, \dots, A_n , $MT(A_1, A_2, \dots, A_n)$, is defined as follows:

$$MT(A_1, A_2, \dots, A_n) = \{ST(A_p) \mid p = 1, \dots, i\} + \{RT(A_q, A_{q+1}) \mid q = 1, \dots, k\} \quad , k = {}_n C_2 - j C_2$$

Here $ST(A_i)$ is the single trajectory information of an object A_i . $RT(A_k, A_{k+1})$ is the relationship trajectory information between object A_k and A_{k+1} where k is the number of relationship trajectories between two moving objects as well as between a moving object and a stationary object.

3 Similarity Measure

Based on our spatio-temporal modeling for moving objects’ trajectories, we define a similarity measure for a single trajectory and multiple trajectories, respectively. Since we measure a similarity between i -th motion in query trajectory Q and j -th motion in data trajectory S , we define a distance function between two motions.

Definition 12. A distance function, $dS_{df}(q[i], s[j])$, to measure the similarity between the arbitrary motion $s[i]$ of a data trajectory S and the arbitrary motion $q[j]$ of a query trajectory Q is defined as follows.

$$d_{dis}(s[i,2], q[j,2]) = |s[i, 2] - q[j, 2]|$$

$$\text{if } |s[i, 1] - q[j, 1]| > 180 \text{ then } d_{ang}(s[i, 1], q[j, 1]) = (360 - |s[i, 1] - q[j, 1]|)$$

$$\text{else } d_{ang}(s[i, 1], q[j, 1]) = |s[i, 1] - q[j, 1]|$$

$$dS_{df}(s[i], q[j]) = ((d_{ang} / 180) * \alpha) + ((d_{dis} / 100) * \beta)$$

Here, d_{ang} is a distance function for the direction (angle) property for all the motions of a trajectory and d_{dis} is a distance function for the distance property. $s[i, 1]$ and $s[i, 2]$ are the direction and the distance value of the i -th motion in a trajectory S , respectively. α and β mean the weight of the direction and the distance, respectively, when $\alpha + \beta = 1.0$.

Definition 13. Given a single trajectory $S = \{s[0], s[1], \dots, s[m]\}$ and a query trajectory $Q = \{q[0], q[1], \dots, q[n]\}$ ($1 \leq n \leq m$), the similarity between trajectory S and Q , $ST_{sim}(S, Q)$, by using definition 12 is defined as follows.

$$ST_{sim}(S, Q) = \text{MAX}_{i=1}^n \left\{ \frac{\sum_{j=1}^n dS_{df}(s[i + j], q[i])}{n} \right\} \quad (\forall_j, 0 \leq j \leq m - n)$$

Next, to measure similarity between multiple trajectories, we first define the similarity measure for a relationship trajectory between two trajectories. For this, we make use of topological relationships between multiple moving objects as well as moving direction and moving distance. Hence, we define a distance function using three-dimensional properties as follows.

Definition 14. A distance function, $dR_{df}(q[i], s[j])$, to measure the similarity between the arbitrary motion $s[i]$ of a data trajectory S and the arbitrary motion $q[j]$ of a query trajectory Q is defined as follows.

$$\begin{aligned}
 d_{top}(s[i,3], q[j,3]) &= (\text{top_dist}(s[i, 3], q[j, 3]))^2 \\
 d_{dis}(s[i,2], q[j,2]) &= |s[i, 2] - q[j, 2]| \\
 \text{if } |s[i, 1] - q[j, 1]| > 180 &\text{ then } d_{ang}(s[i, 1], q[j, 1]) = (360 - |s[i, 1] - q[j, 1]|) \\
 &\text{else } d_{ang}(s[i, 1], q[j, 1]) = |s[i, 1] - q[j, 1]|
 \end{aligned}$$

$$dR_{df}(s[i], q[j]) = ((d_{ang} / 180) * \alpha) + ((d_{dis}/100) * \beta) + ((d_{top}/25)*\gamma)$$

Here d_{ang} is a distance function for the direction (angle) property for all the motions of a trajectory, d_{dis} is a distance function for the distance property and d_{top} is a distance function for the topology property. $s[i, 1]$, $s[i, 2]$ and $s[i, 3]$ are the angle, the distance and the topology value of the i -th motion in a multiple trajectory S , respectively. α , β and γ mean the weight of the angle, the distance and the topology, respectively, when $\alpha+\beta+\gamma=1.0$. Also, $\text{top_dist}(a, b)$ means the similarity distance for topological relations between a and b .

Definition 15. Given a relationship trajectory $S=\{s[0], s[1], \dots, s[m]\}$ and a query trajectory $Q=\{q[0], q[1], \dots, q[n]\}$ ($1 \leq n \leq m$), the similarity between trajectory S and Q , $RT_{sim}(S, Q)$, by using definition 14 is defined as follows.

$$RT_{sim}(S, Q) = \text{MAX}_{j=1}^n \left\{ \frac{\sum_{i=1}^n dR_{df}(s[i + j], q[i])}{n} \right\} \quad (\forall j, 0 \leq j \leq m - n)$$

A list of multiple trajectories of at least two or more moving objects, $MT(A_1, A_2, \dots, A_n)$, can be represented as a combination of single trajectory(ST) and relationship trajectories(RT).

Definition 16. Let us suppose that i and j is the number of moving objects and stationary objects, respectively ($n=i+j$). The similarity for multiple trajectories of objects A_1, A_2, \dots, A_n , $MT(A_1, A_2, \dots, A_n)$, is calculated as follows:

$$MT_{sim}(A_1, A_2, \dots, A_n) = \frac{\sum_{p=1}^i ST_{sim}(A_p) + \sum_{q=1}^k RT_{sim}(A_q, A_{q+1})}{i + k} \quad , k = {}_n C_2 - {}_j C_2$$

Here $ST_{sim}(A_i)$ is the similarity value of single trajectory of object A_i . $RT_{sim}(A_k, A_{k+1})$ is the similarity value of relationship trajectories between objects A_k and A_{k+1} where k

is the number of relationship trajectories between two moving objects as well as between a moving object and a stationary object.

4 CS²VR System

We make use of soccer video data as its application since there are many trajectories data of salient objects such as soccer ball, player, and referee in the soccer ground field. Additionally, the trajectory information plays an important role in indexing soccer video data and detecting the scene with ‘goal in’ in soccer video database. Therefore, we implement the Content- and Semantic-based Soccer Video Retrieval (CS²VR) under Windows 2003 Server O.S with Pentium III-800 and 512 MB memory by using Microsoft Visual C++ compiler and DirectX so that we can support the indexing and searching on soccer video data. Figure 1 first shows a Graphic User Interface (GUI) for soccer video indexing which can help users to easily extract the trajectory information of soccer ball from soccer video data semi-automatically, e.g. moving direction, moving distance, # of frame, player name, and so on. We implement it so as to work well under Window platform. Our GUI for soccer video indexing is composed of two windows: main window and soccer ground window. The former is to browse raw soccer video data formatted as mpeg (*.mpeg) and to extract the trajectory information of soccer ball and main player. The latter is to transform the location of soccer ball on raw video through the main window data into an absolute location on the coordinate of soccer ground field. For this, we make use of ‘affine transformation algorithm’ which is mainly used in computer vision or image processing fields [8].

Figure 2 depicts a GUI for soccer video retrieval which can help users to retrieve the results acquired in soccer video databases. We implemented it by JAVA application in order to work well independently without regard to system platform. We can

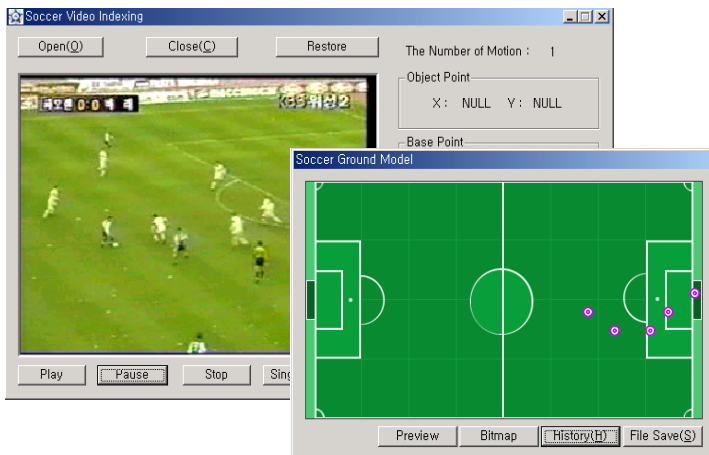


Fig. 1. GUI for soccer video indexing in CS²VR system

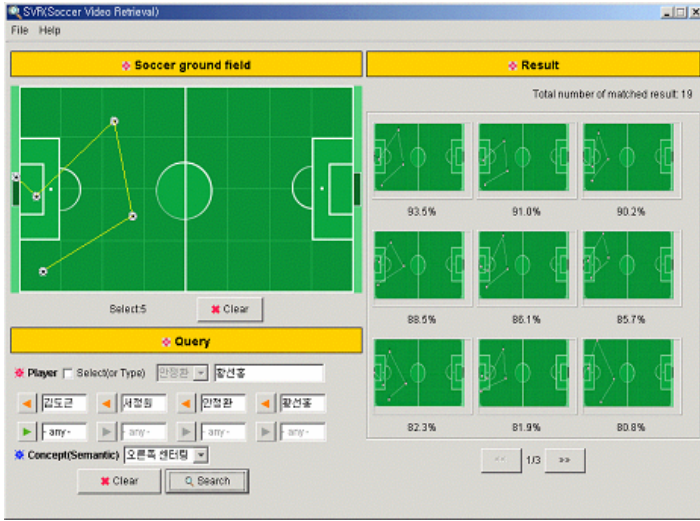


Fig. 2. GUI for soccer video retrieval in CS²VR system

provide three types of user query, that is, trajectory-based query, semantic-based query, and actor-based query as shown in the left part of Figure 2. The trajectory-based query is based on the trajectory of moving objects such as soccer ball and player as the following query: “Finds all video shots whose trajectory is similar to the trajectory sketched by a user on soccer video retrieval interface”. The semantic-based query is based on important semantics such as ‘penalty kick’, ‘corner kick’ and ‘goal in’ in soccer video databases: “Finds all video shots including a scene ‘goal in’”. Finally, the actor-based query is based on the interested player name such as ‘Ronaldo’, ‘Rivaldo’ and ‘Zidane’ in soccer video databases: “Finds all video shots including a scene ‘goal in’ by player name ‘Ronaldo’”. The retrieved results on a user query are provided in the form of trajectory images with the similar trajectory as shown in the right part of Figure 2. We can browse them in the order of the degree of relevance to a user query. The real soccer video shot corresponding to the trajectory image retrieved can be shown by clicking its trajectory image with the mouse button.

5 Conclusions

We proposed a new spatio-temporal modeling scheme which can efficiently represent multiple moving objects in video databases. In order to improve retrieval accuracy to measure a similarity between two moving objects, our scheme took into account on distance property, conceptual location information, and related object information while the traditional schemes only consider direction property, time interval property, and spatial relations property for modeling moving objects' trajectories. We designed and implemented the Content- and Semantic-based Soccer Video Retrieval (CS²VR) system as its application. As future study, it is necessary to research on indexing technique to support good retrieval efficiently when the amount of multiple moving

objects is very large and STQL(Spatio-Temporal Query Language) to make query formally for moving objects in spatio-temporal databases like SQL.

Acknowledgements

This work is financially supported by the Ministry of Education and Human Resources Development (MOE), the Ministry of Commerce, Industry and Energy (MOCIE) and the Ministry of Labor (MOLAB) through the fostering project of the Industrial-Academic Cooperation Centered University.

References

- [1] J. R. Smith, S. F. Chang, "VisualSEEK: a Fully Automated Content-Based Image Query System," in Proceedings of ACM Multimedia 96, pp. 87-98, 1996.
- [2] Virginia, E. Ogle and M. Stonebraker, "Chabot: Retrieval from a Relational Database of images," IEEE Computer, Vol. 28, No. 9, pp. 40-48, 1995.
- [3] G. Ahanger, D. Benson, and T. D. C Little, "Video query formulation," in Proceedings of SPIE Electronic Imaging Science and Technology, pp. 280-291, 1995.
- [4] A. Yoshitaka, M. Yoshimitsu, M. Hirakawa, and T. Ichikawa, "V-QBE: Video database retrieval by means of example motion of objects," in Proceedings of IEEE International Conference on Multimedia Computing and Systems, pp. 453-457, 1996.
- [5] Z. Aghbari, K. Kaneko, and A. Makinouchi, "Modeling and Querying Videos by Content Trajectories", In Proceedings of the International Conference and Multimedia Expo, pp. 463-466, 2000.
- [6] J. W. Chang, Y. J. Kim, and K. J. Chang, "A Spatial Match Representation Scheme Indexing and Querying in Icnoc Image Databases," ACM International Conference on Information and Knowledge Management, pp. 169-176, 1997.
- [7] J. F. Allen, "Maintaining Knowledge about Temporal Intervals," Communication of the ACM, Vol. 26, No. 11, pp. 832-843, 1983.
- [8] H.S. Yoon, J. Soh, B.W. Min, and Y.K. Yang, "Soccer image sequences mosaicing using reverse affine transform," In Proc. of International Technical Conference on Circuits/Systems, Computers and Communications, pp. 877-880, 2000.