

내용 및 개념 기반 비디오 검색을 지원하는 시공간 표현 기법

심준보^o 장재우
전북대학교 컴퓨터공학과
(cbsim, jwchang)@dblab.chonbuk.ac.kr

A Spatio-temporal Representation Scheme for supporting Content- and Concept-based Video Retrieval

Choon-Bo Shim^o Jae-Woo Chang

Dept. of Computer Engineering, Chonbuk National University

요 약

비디오 데이터는 미디어의 특성상 시간의 흐름에 따라 객체의 위치가 변하는 움직임 객체(Moving Object)를 가지며, 움직임 객체를 기반으로 비디오 데이터를 색인 및 검색하기 위해서는 적합한 시공간 표현 기법이 요구된다. 따라서, 본 논문에서는 비디오 데이터가 지나는 움직임 객체의 움직임 경로(Trajectory)를 효율적으로 모델링하고 사용자에게 빠른 검색 결과를 제공할 수 있는 근사 탐색(Aproximation Search)이 가능한 새로운 시공간 표현 기법을 제안한다. 이는 사용자가 직접 스키프 한 움직임 경로를 토대로 검색을 수행하는 내용 기반 검색과 움직임 객체의 일련의 움직임 경로들과 움직임 경로가 일어나는 위치 정보를 통해 얻어진 개념(의미)을 이용한 개념 기반 검색을 지원한다.

1. 서론

텍스트나 이미지 데이터와는 달리 비디오 데이터가 지나는 가장 중요한 특징은 움직임 객체에 대한 움직임 정보이다. 이러한 움직임 정보는 각각의 프레임 내에서의 객체들간의 공간적인 정보와 일련의 프레임들간의 시간적인 정보가 결합된 시공간 관계성을 통해 표현될 수 있으며, 비디오 데이터에 대한 사용자의 내용 및 개념 기반 검색을 수행하는 데 있어 매우 중요한 역할을 한다. 비디오 데이터베이스에서 움직임 객체의 움직임 경로를 이용한 내용 기반 검색 질의로는 다음과 같은 질의가 가능하다. "사용자 인터페이스를 통해 스키프된 움직임 객체의 움직임 경로와 유사한 움직임 경로를 가진 모든 비디오 샷(Shot)을 찾아라." 또한 응용 분야에 따라 일련의 움직임 객체의 움직임 경로로부터 어떤 개념 혹은 의미(Semantic) 정보를 유추해 낼 수 있다. 가령, 축구 비디오 데이터베이스의 경우 일련의 축구공의 움직임 경로로부터 코너킥(Corner Kick), 페널티 킥(Penalty Kick), 센터링(Centering) 등과 같은 개념(의미) 정보를 대략적으로 얻을 수 있다. 따라서, 사용자는 비디오 데이터에서 움직임 객체의 움직임 경로를 이용한 내용 기반 검색과 더불어 다수의 움직임 경로로부터 얻어낸 개념을 통한 개념 기반 검색을 수행할 수 있다.

본 논문에서는 비디오 데이터가 지나는 움직임 객체의 움직임 경로를 효율적으로 모델링하기 위한 새로운 시공간 표현 기법을 제안한다. 아울러, 사용자가 원하는 객체의 움직임 경로를 기반으로 사용자 인터페이스를 이용한 검색을 수행할 수 있는 내용 기반 검색과 움직임 객체의 일련의 움직임 경로로부터 유추해 낸 개념을 이용한 개념 기반 검색을 지원한다.

본 논문의 구성은 다음과 같다. 2장에서 기존의 움직임 객체를 이용한 비디오 검색에 관한 관련 연구를 살펴보고, 3장에서는 내용 및 개념 기반 검색을 지원하는 시공간 표현 기법을 설명한다. 4장에서는 제안하는 시공간 표현 기법을 기반으로 축구 비디오 데이터베이스에서의 내용 및 개념 기반 검색에 대해서 기술한다. 마지막으로, 5장에서 결론 및 향후 연구 방향에 대해서 설명한다.

2. 관련 연구

첫째, John Z. Li[1]은 움직임 객체를 위해서 8개의 방향(NT, ET, ST, WT, NE, SE, SW, NW)을 고려하여 객체 움직임 경로를 표현하고 있다. 임의의 시간 간격 I_i 동안 객체 A의 움직임(Motion)은 (S_i, d_i, I_i) 로 표현하며, 여기서, S_i 는 객체 A의 변위(Displacement)이고 d_i 는 객체 A의 움직임 방향(Direction)을 의미한다. 시간 간격 I_k 동안 움직임 객체 A와 B사이의 시공간 관계성은 $A(\alpha, \beta, I_k)B$ 로 표현하며, 여기서, α 는 움직임 객체 A와 B사이의 위상 관계(Topological Relation) 즉, DJ, TC, EQ, IN, CB, CT, CV, OL 중의 하나를 의미하고, β 는 움직임 객체 A와 B의 방향 관계(Directional Relation)를 나타낸다. 둘째, Mohammad Nabil[2]은 비디오 데이터 내의 움직임 객체를 모델링하기 위해 기존의 이미지내의 객체들간의 공간 정보를 표현하기 위한 공간 매치 표현 기법인 2D-PIR(Projection Interval Relationships) 방법을 확장한 ST(Spatio Temporal)-PIR을 제안하고 그래프 기반 기법을 통해 이들을 표현하고 있다. 또한 움직임 객체를 이용한 검색을 위해 다수의 연산자 제공하고 있다. 셋째, M.K. Shan[3]은 움직임 객체의 움직임 경로를 일련의 세그먼트(Segment)의 집합으로 나타내고 각각의 세그먼트는 0~360까지의 각도로 표현하고 있다. 또한 움직임 경로에 대해서 근사 탐색(Aproximation Search)를 위한 OCM(Optimal Consecutive Mapption)과 OCMR(Optimal Consecutive Mapption with Replication)이라는 두 가지의 유사성 측정 알고리즘을 제안하였다. 마지막으로, Z.Aghbari[4]은 OL(Object Level), FL(Frame Level), SL(Shot Level) 3단계의 계층적인 표현 방법을 이용하여 비디오 데이터를 표현하고 있다. OL 단계에서는 각 객체가 가지는 특징인 색상, 모션, 위치 정보를 추출하여 feature trajectory를 구성하고, FL 단계에서는 각 객체들간의 방향 및 위상 관계를 추출하여 relationship trajectory를 구성한다. 그리고 SL 단계에서는 이전 단계에서 추출한 feature trajectory와 relationship trajectory를 결합한 content trajectory를 구성함으로써 비디오 데이터에 대한 내용 기반 검색을 지원하고 있다.

3. 내용 및 개념 기반 검색을 위한 시공간 표현 기법

비디오 데이터는 이미지와 달리 공간 정보와 시간 정보를 모두 포함하고 있는 특성을 가지고 있기 때문에, 움직임 객체를 보다 효율적으로 표현하기 위해서는 움직임 객체에 대한 공간 관계성(Spatial Relationship)과 시간 관계성(Temporal Relationship)을 모두 고려해야 한다. 따라서, 본 논문에서는 하나의 객체에 대해 움직임 경로를 나타내는 Single Trajectory(ST)와 둘 이상의 다수의 객체에 대해 움직임 경로를 나타내는 Multiple Trajectories(MT)를 위한 새로운 시공간 표현 기법을 제안한다. 아울러, 제안하는 시공간 표현 기법은 사용자에게 보다 빠른 검색 결과를 제공할 수 근사 탐색을 지원한다.

3.1 단일 움직임 경로(Single Trajectory)

[정의 1] 객체 A를 움직임 객체라고 할 때, 움직임 객체 A의 단일 움직임 경로(Single Trajectory), $ST^{(A)}$ 는 그림 1과 같이 IE와 ME로 구성된다.

$$ST^{(A)} = IE + ME$$

여기서, IE(Instantaneous Element)는 움직임 경로를 구성하는 전체 시점들 즉, t_0 에서 t_n 까지 각 시점에서의 위치 정보를 의미하며 $\{P_n^{(A)}\} = \{P_0^{(A)}, P_1^{(A)}, \dots, P_n^{(A)}\}$ 로 나타낸다. 여기서, 위치 정보는 일반적인 위치 정보 즉 좌표 정보가 아닌 본 논문에서 제안하는 개념 기반 검색을 위해서 정의한 새로운 개념 기반 위치 정보를 의미한다(자세한 설명은 다음 장을 참조). 한편, ME(Motion Element)는 움직임 경로를 구성하는 각 움직임(Motion)을 나타내며 $\{M_{n-1}\} = \{M_0, M_1, \dots, M_{n-1}\}$ 으로 표현한다. M_i 는 (a_i, D_i, I_i) 로 구성되며, I_i 는 움직임이 일어나는 시작 시점(t_i)과 끝 시점(t_{i+1})으로 구성된 시간 간격을 나타낸다. a_i 는 시간 간격 I_i 동안 움직인 방향으로서 $0^\circ \sim 360^\circ$ 까지의 실제 각도로서 나타내며, D_i 는 시간 간격 I_i 동안 움직인 거리(Distance)를 의미한다. 따라서, $ST^{(A)}$ 는 다음과 같다.

$$ST^{(A)} = \{P_n^{(A)}\} + \{M_{n-1}\}$$

주어진 시간 간격 리스트 $\{I_0, I_1, \dots, I_{n-1}\}$ 에 대해서, 움직임 객체 A의 움직임 경로 $ST^{(A)}$ 는 다음과 같이 표현된다.

$$ST^{(A)} = (P_0^{(A)}, P_1^{(A)}, \dots, P_n^{(A)}) +$$

$$\{(a_0, D_0, I_0), (a_1, D_1, I_1), \dots, (a_{n-1}, D_{n-1}, I_{n-1})\}$$

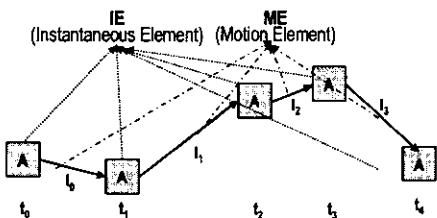


그림 1. 움직임 객체 A의 단일 움직임 경로

3.2 다수 움직임 경로(Multiple Trajectory)

[정의 2] 객체 A 또는 객체 B를 움직임 객체라고 할 때, 움직임 객체 A와 B 사이의 움직임 경로를 Relationship Trajectory, $RT^{(A,B)}$ 라고 하며 다음과 같다.

$$RT^{(A,B)} = IE + ME$$

여기서, IE(Instantaneous Element)는 단일 움직임 경로에서와 같이 움직임 경로를 구성하는 전체 시점들 즉, t_0 에서 t_n 까지 각 시점에서의 객체 A와 객체 B의 위치 정보, 위상 정보(R), 방향 정보(α)를 의미하며 $\{P_n^{(A,B)}\} = \{P_0^{(A,B)}, P_1^{(A,B)}, \dots, P_n^{(A,B)}\}$ 로 나타낸다. 여기서, $P_i^{(A,B)} = (P_i^{(A)}, P_i^{(B)}, R_i, \alpha_i)$ 이다. R_i 은 움직임 객체 A와 B사이의 위상 관계성 즉, FA(FarAway), DJ(DisJoint), ME(MEet), OL(OverLap), CL(Is inCLuded by), IN(INclude), SA(Same) 중에 하나의 값으로 표현한다. 그리고 α_i 는 움직임 객체 A(기준 객체)에 대한 객체 B(대상 객체)의 방향 관계성을 나타낸다. 한편, ME(Motion Element)는 움직임 경로를 구성하는 각 움직임(Motion)을 나타내며 $\{M_{n-1}\} = \{M_0, M_1, \dots, M_{n-1}\}$ 으로 표시한다. M_i 는 (D_i, I_i) 로 구성되며, I_i 는 단일 움직임에서와 같은 의미이고 D_i 는 시간 간격(I_i) 동안 객체 A에 대한 객체 B의 상대적인 이동 거리(Relative Moving Distance)로 0에서 100사이의 정규화된 값으로 나타낸다. 즉, 객체 A의 이동 거리와 객체 B의 이동 거리가 같을 때 D_i 를 50로 정하고, 이를 기준으로 객체 A의 이동 거리가 객체 B의 이동 거리보다 클 경우에는 50에서 100사이의 값을 갖으며, 그와 반대일 경우에는 0에서 50사이의 값을 갖는다. 따라서, $RT^{(A,B)}$ 는 다음과 같다.

$$RT^{(A,B)} = \{P_n^{(A,B)}\} + \{M_{n-1}\}$$

주어진 시간 간격 리스트 $\{I_0, I_1, \dots, I_{n-1}\}$ 에 대해서, 움직임 객체 A와 B사이의 Relationship Trajectory는 다음과 같이 표현된다.

$$RT^{(A,B)} = \{(P_0^{(A)}, P_0^{(B)}, R_0, \alpha_0), (P_1^{(A)}, P_1^{(B)}, R_1, \alpha_1), \dots, (P_n^{(A)}, P_n^{(B)}, R_n, \alpha_n)\} + \{(D_0, I_0), (D_1, I_1), \dots, (D_{n-1}, I_{n-1})\}$$

정의 1과 정의 2를 기반으로 둘 이상의 객체로 이루어진 다수의 움직임 경로인 Multiple Trajectory는 다음과 같이 정의한다.

[정의 3] 객체 A_1, A_2, \dots, A_n 중에 적어도 하나는 움직임 객체라고 할 때, 움직임 객체 A_1, A_2, \dots, A_n 으로 이루어진 다수 움직임 경로(Multiple Trajectory), $MT^{(A_1, A_2, \dots, A_n)}$ 는 다음과 같이 정의한다.

$$MT^{(A_1, A_2, \dots, A_n)} = \sum_{i=1}^n ST^{(A_i)} + \sum_{j=1}^{n-1} \sum_{k=1}^n RT^{(A_j, A_k)}$$

여기서, $ST^{(A_1)}, ST^{(A_2)}, \dots, ST^{(A_n)}$ 는 각 객체의 단일 움직임 경로를 나타내며, $RT^{(A_1, A_2)}, RT^{(A_1, A_3)}, \dots, RT^{(A_{n-1}, A_n)}$ 는 각각 객체들간의 Relationship Trajectory를 나타낸다.

4. 내용 및 개념 기반 검색 : 축구 비디오 데이터

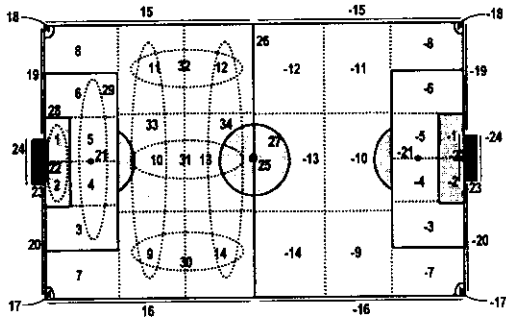
본 논문에서는 제안한 새로운 시공간 표현 기법의 유용성을 보이기 위해 축구 비디오 데이터베이스를 응용 분야로 선정하였다. 축구 비디오 데이터는 사용자의 주된 관심 객체인 축구공이 경기장을 배경으로 많은 움직임 정보를 지니고 있기 때문에 객체의 움직임 경로를 추출하는 데 있어 매우 용이할 뿐만 아니라, 축구공의 움직임 경로가 발생하여 위치에 따라 축구 경기에서 일어나는 개념(의미) 정보를 추출할 수 있다. 이를 위해, 내용 및 개념 기반 검색을 위한 축구 경기장의 개념 기반 위치 정보를 정의하며 그림 2와 같다.

만약 사용자 인터페이스(GUI)를 통해 그림 3과 같은 질의 움직임 경로 Q가 주어지면, 질의 움직임 경로 Q를 시공간 표현 기법을 이용해 다음과 같이 표현한다.

$$\begin{aligned} ST^{(Q)} &= \{P_n^{(Q)}\} + \{M_{n-1}\} \\ &= (P_0^{(Q)}, P_1^{(Q)}, P_2^{(Q)}, P_3^{(Q)}, P_4^{(Q)}, P_5^{(Q)}) + \\ &\quad (M_0, M_1, M_2, M_3, M_4) \\ &= (LRT, LCM, LCT, LLT, LLP, LGP) + \\ &\quad \{(240, 19, I_0), (105, 19, I_1), (150, 12, I_2), \\ &\quad (280, 32, I_3), (135, 16, I_4)\} \end{aligned}$$

1) 본 논문에서 사용하는 '+' 연산자는 산술 연산자가 아닌 연결(concatenation) 연산자를 의미한다.

위와 같이 표현된 질의 움직임 경로 $ST^{(Q)}$ 중에서 $\{Pn^{(Q)}\}$ 는 움직임 경로가 발생한 각 시점에서의 위치 정보를 나타내기 때문에 이를 이용하여 축구 비디오 데이터베이스에 대해 보다 빠른 검색을 가능하게 하는 근사 탐색(Aproximation Search)을 수행할 수 있으며, 또한 사용자가 보다 정확한 검색 결과를 원하면 $\{M_{n-1}\}$ 정보를 이용하여 근사 탐색을 통해 검색된 결과 중에서 보다 정확한 축구 비디오 샷들을 유사성을 계산하여 유사성이 높은 순으로 브라우징할 수 있다.



(a) 축구 비디오 데이터를 위한 축구 경기장 모델

1	Left Upper Goal	LUG(RUG)	18	Left Top Corner	LTC(RTC)
2	Left Lower Goal	LLG(RLG)	19	Left Top Goalline	LTG(RTG)
3	Left Bottom Penalty	LBP(RBP)	20	Left Bottom Goalline	LBC(RBG)
4	Left Lower Penalty	LPL(RLP)	21	Left Penalty Point	LPP(RPP)
5	Left Upper Penalty	LUP(RUP)	22	Left Goospost Line	LGL(RGL)
6	Left Top Penalty	LTP(RTP)	23	Left Goospost Inside	LGR(RGI)
7	Left Left Bottom	LLB(RRB)	24	Left Goospost Over	LGO(RGO)
8	Left Left Top	LLT(RRT)	25	Center Circle Point	CCP
9	Left Center Bottom	LCB(RCB)	26	Half Line	HFL
10	Left Center Middle	LCM(RCM)	27	Center Circle Area	CCA
11	Left Center Top	LCT(RCT)	28	Left Goal Area	LGA(RGA)
12	Left Right Top	LRT(RLT)	29	Left Penalty Area	LPA(RPA)
13	Left Right Middle	LRM(RLM)	30	Left Bottom Area	LBA(RBA)
14	Left Right Bottom	LRB(RLB)	31	Left Middle Area	LBA(RBA)
15	Left Top Sideline	LTS(RTS)	32	Left Top Area	LTA(RTA)
16	Left Bottom Sideline	LBS(RBS)	33	Left Left Area	LTA(RTA)
17	Left Bottom Corner	LBC(RBC)	34	Left Right Area	LRA(RLA)

(b) 내용 및 개념 기반 검색을 위한 위치 정보

그림 2. 내용 및 개념 기반 검색을 위한 축구 경기장의 위치 정보

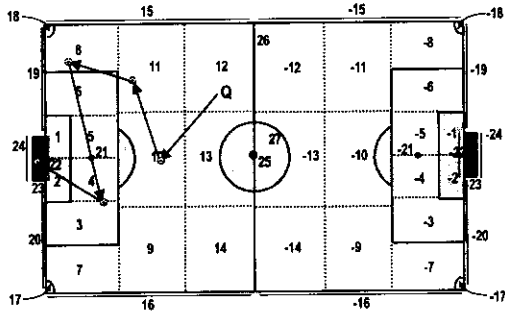


그림 3. 움직임 경로를 이용한 내용 기반 검색

또한 축구 비디오 데이터베이스에서는 일련의 축구공의 움직임 경로와 움직임 경로가 발생하는 대략적인 위치 정보를

토대로 코너킥(Corner Kick), 페널티 킥(Penalty Kick), 골킥(Goal Kick), 드로우 인(Throw In), 프리 킥(Free Kick), 센터링(Centering) 등과 같이 축구 비디오 데이터베이스에서 사용하는 개념(의미)을 유추해 낼 수 있다. 따라서, 본 논문에서는 움직임 경로를 이용한 내용 기반 검색 뿐만 아니라, 일반적으로 축구 비디오 데이터에서 중요시 여기는 개념들을 이용한 개념 기반 검색을 지원함으로써 검색의 정확성과 더불어 사용자의 편리성을 도모할 수 있다.

만약 사용자가 개념 기반 검색을 원하면, 이를 처리하는 시스템은 주어진 질의 개념에 해당하는 시공간 표현 기법으로 변환한 뒤에 데이터베이스를 검색하여 사용자에게 검색 결과를 제공한다. 다음 표1은 축구 비디오에서 주요 개념의 그에 해당하는 시공간 표현 기법을 나타낸 것이다. 단, 경기장 왼쪽 진영을 가정한 오른쪽 진영도 이와 유사하다.(여기서, '&', '|', '?', '['] 연산자는 각각 And, Or, Don't care, All을 의미한다.)

표 1. 축구 경기에서의 주요 개념과 그것의 시공간 표현

개념	시공간 표현
골 킥 (Goal Kick)	$(LGA LPA, [(LGA&LPA)]) + \{(0\sim 90 270\sim 360, ? , I_0)\}$
페널티 킥 (Penalty Kick)	a. $(LPP, LFG LBG LGO LGI) + \{(90\sim 270, ? , I_0)\}$ b. $(LPP, LPL, LGA LPA) + \{(90\sim 270, ? , I_0), (0\sim 90 270\sim 360, ? , I_1)\}$
프리 킥 (Free Kick)	$(LLB LLT LBA LMA LTA, LGA LPA) + \{(90\sim 270, ? , I_0)\}$
센터링 (Centering)	$(LLB LLT LCB LCT, LGA LPA) + \{(90\sim 270, ? , I_0)\}$
코너 킥 (Corner Kick)	a. $(LTC, LLT LGA LPA LTG LGL LGI) + \{(270\sim 360, ? , I_0)\}$ b. $(LBC, LLB LGA LPA LBG LGL LGI) + \{(0\sim 90, ? , I_0)\}$
드로우 인 (Throw In)	a. $(LBS, [LBS]) + \{(0\sim 180, ? , I_0)\}$ b. $(LTS, [LTS]) + \{(180\sim 360), ? , I_0)\}$
골 인 (Goal In)	$(?, \dots, LGI) + \{(?, ? , I_0), \dots, (?, ? , I_{n-1})\}$
골 아웃 (Goal Out)	$(?, \dots, LTG LBG) + \{(?, ? , I_0), \dots, (?, ? , I_{n-1})\}$

5. 결론

본 논문에서는 비디오 데이터가 지니는 움직임 객체의 움직임 경로를 효율적으로 모델링하고 사용자에게 빠른 검색 결과를 제공할 수 있는 근사 탐색이 가능한 새로운 시공간 표현 기법을 제안하였다. 이는 사용자가 직접 스케치한 움직임 경로를 토대로 검색을 수행할 수 있는 내용 기반 검색과 움직임 객체의 일련의 움직임 경로들과 움직임 경로가 일어나는 위치 정보를 통해 얻어진 개념(의미)을 이용한 개념 기반 검색을 지원한다.

참고문헌

[1] John Z. Li, M. Tamer Ozsu, Duane Szafron, "Modeling of Moving Objects in a Video Database", In Proceedings of the International Conference on Multimedia Computing and Systems, pp. 336-343 (1997).
 [2] M. Nabil, A.H. Ngu, J. Shepherd, "Modeling Moving Objects in Multimedia Databases", In Proceedings of 5th International Conference on Database Systems for Advanced Applications, pp. 1997
 [3] Man-Kwan Shan and Suh-Yin Lee, "Content-based Video Retrieval via Motion Trajectories", In Proceedings of the International Conference on SPIE, Vol. 3561, pp. 52-61, 1998
 [4] Z.Aghbari, K.Kaneko, A.Makinouchi, "Modeling and Querying Videos by Content Trajectories", In Proceedings of the International Conference and Multimedia Expo, pp. 463-466, 2000