

비디오 데이터에서 움직임 객체의 모델링을 위한 시공간 표현 기법

(A Spatio-temporal Representation Scheme for Modeling Moving
Objects in Video Data)

심 춘 보 [†] 장 재 우 ^{**}

(Choon-Bo Shim) (Jae-Woo Chang)

요약 비디오 데이터에서 움직임 객체에 대한 움직임 경로는 내용-기반 검색을 위해 비디오 데이터를 색인하는 데 있어 매우 중요한 역할을 한다. 따라서, 본 논문에서는 비디오 데이터에서 움직임 객체의 움직임 경로를 모델링하기 위한 새로운 시공간 표현 기법을 제안한다. 비디오 데이터를 위한 보다 효율적인 내용-기반 검색을 위해, 제안하는 기법은 시간, 공간 관계성과 더불어 일정 시간 간격 동안 움직인 객체의 이동 거리(moving distance)를 고려한다. 아울러, 제안하는 표현 기법에 기반하여 단일 움직임 객체의 움직임 경로와 다수 움직임 객체들의 움직임 경로를 위한 새로운 유사성 측정 알고리즘을 제시하며, 이들 알고리즘은 검색 결과에 대해서 유사성에 준하여 순위(Ranking)를 부여할 수 있다. 마지막으로, 성능 평가를 통하여 제안된 시공간 표현 기법은 기존의 Li 방법과 Shan의 방법에 비해 동등한 재현율을 유지하며, 정확율 측면에서 약 20%의 성능 향상을 보인다.

Abstract The trajectory of moving objects in video data plays an important role in video indexing for content-based retrieval. In this paper, we propose a new spatio-temporal representation scheme for modeling moving objects' trajectories in video data. In order to support content-based retrieval on video data very well, our representation scheme considers the moving distance of an object during a given time interval as well as its temporal and spatial relations. Based on our representation scheme, we present two similarity measure algorithms for both the trajectory of a single moving object and those of multiple moving objects, which provide ranking for the retrieved video results. Finally, we show from our experiment that our representation scheme achieves about 20% higher precision while it holds about the same recall, compared with its competitors, both Li's and Shan's schemes.

1. 서론

최근 들어, 컴퓨터 하드웨어, 컴퓨터 네트워크 그리고 저장 시스템 등이 급속한 발전을 거듭하면서 이미지, 오디오 및 비디오와 같은 대용량의 멀티미디어 데이터를 기반으로 하는 멀티미디어 데이터베이스에 대한 연구가 활발히 이루어지고 있다. 아울러, 멀티미디어 데이터의

수가 기하급수적으로 증가하면서 멀티미디어 데이터에 대한 내용-기반 검색 기술[1][2][3][4][5][6]에 대한 중요성도 빠른 속도로 증가하고 있는 추세이다.

한편, 텍스트나 이미지 데이터와는 달리, 비디오 데이터가 가지는 중요한 특징 중에 하나는 움직임 객체에 대한 움직임 경로이다. 이러한 움직임 경로[7][8][9]는 객체의 공간적인 속성과 시간적인 속성이 결합된 시공간 관계성을 통해 표현되며, 비디오 데이터에 대한 사용자의 내용-기반 검색을 수행하는 데 있어 매우 중요한 역할을 한다. 시공간 관계성에 기반한 사용자의 질의는 다음과 같다: "사용자에 의해 스케치된 움직임 객체의 움직임 경로와 유사한 움직임 경로를 가진 모든 객체들을 찾아라." 또는 "멀리 떨어져 있던 두 대의 자동차가 다가오면서 부딪치는 장면을 포함하는 비디오 샷(shot)

· 본 연구는 한국과학기술재단의 97목적기초과제(97-0100-0101-3)의 연구비에 의해 수행되었음.

[†] 비 회 원 : 전북대학교 컴퓨터공학과
cbsim@cblab.chonbuk.ac.kr

^{**} 종 신 회 원 : 전북대학교 컴퓨터공학과 교수
jwchang@cblab.chonbuk.ac.kr

논문접수 : 2000년 5월 22일

심사완료 : 2000년 10월 12일

을 찾아라.”

위와 같은 형태의 질의를 처리하기 위해서는 비디오 데이터에서 움직임 객체들 간의 시간 관계성(temporal relation)과 공간 관계성(spatial relation)에 대한 연구가 수반되어야 한다. 그 중에서 시간 관계성에 대한 연구는 대부분 Allen[10]의 13가지 시간 관계 모델에 기반하고 있으며, 공간 관계성에 대한 연구는[11][12] 공간 좌표를 이용하여 위상 정보와 방향 정보에 기반을 두고 있다. 또한 대부분의 기존 연구들은 이와 같이 움직임 객체들 간의 시공간 관계성만을 고려할 뿐, 일정 시간 간격 동안 움직임 객체가 움직인 이동 거리(moving distance)는 고려하고 있지 않다. 따라서, 비디오 데이터에서 움직임 객체에 대한 움직임 경로를 보다 효율적으로 모델링 하기 위해서는, 객체들간의 시공간 관계성 뿐만 아니라, 움직임 객체의 이동 거리 또한 고려해야 한다. 예를 들면, 축구 비디오 데이터의 경우에 사용자의 주된 관심 객체인 축구공의 움직인 이동 거리가 숏 패스(short pass)인지 롱 패스(long pass)인지를 판단하는 중요한 역할을 하기 때문이다.

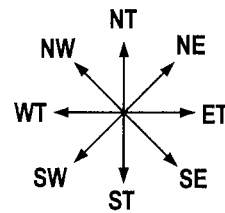
본 논문에서는 비디오 데이터에서 움직임 객체의 움직임 경로를 모델링하기 위한 새로운 시공간 표현 기법을 제안한다. 비디오 데이터를 위한 보다 효율적인 내용 기반 검색을 지원하기 위해, 제안하는 기법은 시간, 공간 관계성과 더불어 일정 시간 간격 동안 움직인 객체의 이동 거리를 고려한다. 아울러, 제안하는 표현 기법에 기반하여 단일 움직임 객체의 움직임 경로와 다수 움직임 객체들의 움직임 경로를 위한 새로운 유사성 측정 알고리즘, 즉 SDST(Similarity measure based on moving Distance for Single object's Trajectory)와 SDMT(Similarity measure based on moving Distance for Multiple objects' Trajectories)를 제안한다. 제안하는 유사성 측정 알고리즘은 사용자 질의에 대해서 유사성을 측정하여 검색 결과를 유사성에 기반하여 높은 순으로 브라우징 할 수 있다.

본 논문의 구성은 다음과 같다. 2장에서는 기존의 시공간 관계성을 이용한 내용-기반 비디오 검색에 대한 관련 연구를 살펴보고, 3장에서는 본 논문에서 제안하는 움직임 정보를 이용한 시공간 표현 기법을 설명한다. 4장에서는 제안하는 시공간 표현 기법에 근거하여 사용자 질의에 대한 유사성을 측정하는 유사성 측정 알고리즘을 기술한다. 5장에서는 제안하는 기법의 성능 평가를 위해, 기존의 기법들과의 성능 비교를 수행한다. 마지막으로, 6장에서는 결론 및 향후 연구 방향에 대해서 설명한다.

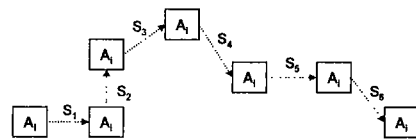
2. 관련 연구

이미지 내의 객체들 간의 공간 관계성을 이용한 내용-기반 이미지 검색 기술에 관한 연구는 현재 상당한 연구 결과를 보이고 있다. 그러나, 비디오나 애니메이션과 같은 멀티미디어 데이터를 기반으로 움직임 객체의 시공간 관계성을 이용한 내용-기반 비디오 검색에 관한 연구들은 그리 많지 않다. 따라서, 여기서는 본 논문과 밀접한 관련이 있는 두 가지 연구에 대해서만 기술한다.

첫째, John Z. Li(이하 Li)[7][8]은 어떤 일정시간 동안 객체의 위치가 변하는 객체를 움직임 객체(moving object)로 간주하고 이에 대해서 그림 1과 같이 8개의 방향 즉, North(NT), NorthWest(NW), West(WT), SouthWest(SW), South(ST), SouthEast(SE), East(ET), NorthEast(NE)을 고려하여 객체의 경로를 표현하고 있다.



(a) 방향(direction)



(b) 객체 A의 움직임 경로

그림 1 움직임 방향과 예

임의의 시간 간격 I_i 동안 객체 A의 움직임(motion)은 (S_i, d_i, I_i) 로 표현하며, 여기에서, S_i 는 객체 A의 변위(displacement)이고 d_i 는 객체 A의 움직임 방향(direction)을 의미한다. 따라서, 어떤 일련의 주어진 시간 간격 $\langle I_1, I_2, \dots, I_n \rangle$ 에 대해서 객체 A의 trajectory는 다음의 일련의 motion들로 표현한다.

$$\langle (S_1, d_1, I_1), (S_2, d_2, I_2), \dots, (S_n, d_n, I_n) \rangle$$

따라서, 그림 1(b)와 같이 객체 A의 trajectory를 표현하면 다음과 같다.

$\langle (S_1, ET, I_1), (S_2, NT, I_2), (S_3, NE, I_3), (S_4, SE, I_4), (S_5, ET, I_5), (S_6, SE, I_6) \rangle$

시간 간격 I_k 동안 이동 객체 A와 B사이의 시공간 관계성은 $A(\alpha, \beta, I_k)B$ 로 표현하며, 여기에서, α 는 객체 A와 B사이의 위상 관계(topological relation) 즉, DJ(DisJoint), TC(Touch), EQ(Equal), IN(Inside), CB(Covered_By), CT(ConTains), CV(Covers), OL(OverLap) 중의 하나를 의미하고, β 는 객체 A와 B의 방향 관계(directional relation)를 나타낸다. 따라서, 객체 A와 B사이의 시공간 관계성은 다음의 일련의 motion들로 표현한다.

$\langle (\alpha_1, \beta_1, I_1), (\alpha_2, \beta_2, I_2), \dots, (\alpha_n, \beta_n, I_n) \rangle$

I_i 에서는 표 1과 같이 위상 관계에 대한 거리(distance)와 방향 관계에 대한 거리를 이용하여 이동 객체 A의 trajectory와 객체 A와 객체 B사이의 시공간 관계성에 대한 유사성(Simialrity)을 계산한다. 객체 A의 trajectory와 객체 B의 trajectory 사이의 유사성 함수 $TrajSim(A, B)$ 는 다음과 같다.

$$\min Diff(A, B) = \min \sum_{i=1}^m distance(M_i, N_{i+j})$$

$$(\forall j, 0 \leq j \leq n - m)$$

$$TrajSim(A, B) = \frac{\max Diff(A, B) - \min Diff(A, B)}{\max Diff(A, B)}$$

둘째, Shan[9]은 내용 기반 비디오 검색을 위해 single motion trajectory와 multiple motion trajectory

를 이용한 유사성 검색(similarity retrieval) 알고리즘을 제시하였다. 먼저, single motion trajectory를 이용한 검색을 위해, 객체의 motion trajectory는 그림 2와 같이 일련의 세그먼트(segment)의 집합으로 나타내고 각각의 세그먼트는 $0^\circ \sim 360^\circ$ 까지의 각도로써 표현한다. 예를 들어, 그림 2에서 video motion trajectory V 는 다음과 같이 표현한다.

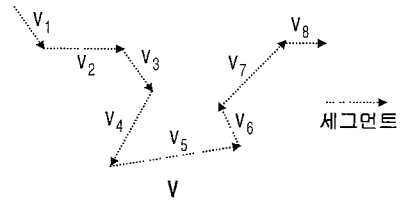


그림 2 움직임 경로의 예

$V = (310^\circ, 0^\circ, 310^\circ, 240^\circ, 5^\circ, 95^\circ, 45^\circ, 0^\circ)$ 사용자 질의 motion trajectory와 데이터베이스의 trajectory 사이의 유사성을 계산하기 위해 OCM(Optimal Consecutive Mapping)과 OCMR(Optimal Consecutive Mapping with Replication)이라는 두 가지의 유사성 측정 알고리즘을 제안하였으며, 이 알고리즘은 모두 single motion trajectory의 방향 정보만을 이용하여 유사성을 계산한다.

표 1 방향과 위상 관계에 대한 거리

(a) 방향 관계에 대한 거리

	NT	NW	NE	WT	SW	ET	SE	ST
NT	0	1	1	2	3	2	3	4
NW	1	0	2	1	2	3	4	3
NE	1	2	0	3	4	1	2	3
WT	2	1	3	0	1	4	3	2
SW	3	2	4	1	0	3	2	1
ET	2	3	1	4	3	0	1	2
SE	3	4	2	3	2	1	0	1
ST	4	3	3	2	1	2	1	0

(b) 위상 관계에 대한 거리

	DJ	TC	EQ	IN	CB	CT	CV	OL
DJ	0	1	6	4	5	4	5	4
TC	1	0	5	5	4	5	4	3
EQ	6	5	0	4	3	4	3	6
IN	4	5	4	0	1	6	7	4
CB	5	4	3	1	0	7	6	3
CT	4	5	4	6	7	0	1	4
CV	5	4	3	7	6	1	0	3
OL	4	3	6	4	3	4	3	0

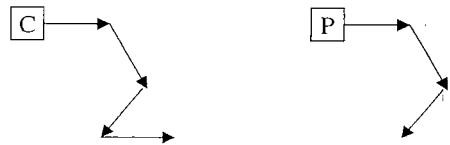


그림 3 다수(multiple) 움직임 경로의 예

그리고, multiple motion trajectory를 이용한 검색을 위해서는 단순히 기존의 2D 이미지 내의 객체 간의 공간 관계성을 위해 Chang[11]에 의해서 제안된 2D 스트링 방법을 이용하였다. 따라서, multiple motion trajectory는 일련의 심볼(symbol) 객체로 나타내며 각각의 심볼 객체는 2D 스트링으로 표현한다. 예를 들어, 그림 3은 “경찰이 자동차를 뒤쫓아 가고 있는 비디오 장면(샷)”으로 “경찰(P)”과 “자동차(C)” 두 객체의 trajectory로 구성된다. 그림 4는 그림 3과 같은 multiple motion trajectory를 2D 스트링으로 표현한 예이다. 그러나, 이 방법은 multiple motion trajectory를 위해 객체들 간의 공간 관계성과 시간 관계성을 전

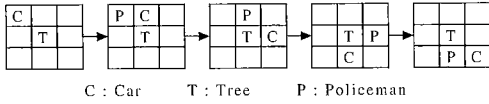


그림 4 2D 스트링 방법으로 표현한 예

혀 고려하고 있지 않는 단점이 있다.

3. 움직임 객체를 위한 새로운 시공간 표현 기법

Li와 Shan 기법은 움직임 객체들 간의 시공간 관계성만을 고려하고 있을 뿐, 일정 시간 간격 동안 움직임 객체가 움직인 이동 거리(moving distance)는 고려하고 있지 않다. 따라서, 비디오 데이터에서 움직임 객체에 대한 움직임 경로를 보다 효율적으로 모델링하기 위해서는 객체들간의 시공간 관계성 뿐만 아니라, 움직임 객체의 이동 거리 또한 고려해야 한다. 이를 위해, 본 논문에서는 비디오 데이터에서 움직임 객체의 움직임 경로를 모델링하기 위한 새로운 시공간 표현 기법을 제안한다. 제안하는 기법은 비디오 데이터를 위한 보다 효율적인 내용 기반 검색을 위해, 시간, 공간 관계성과 더불어 일정 시간 간격 동안 움직인 객체의 이동 거리를 고려한다. 또한 제안하는 시공간 표현 기법에 근거하여 움직임 경로를 통한 사용자 질의에 대한 유사성을 측정하는 새로운 유사성 측정 알고리즘을 제안한다.

먼저, 본 논문에서는 비디오 데이터내의 객체를 인식하기 위해 그림 5와 같은 최소 경계 사각형(MBR : Minimum Bounding Rectangle)을 이용한다. 움직임 객체(moving object)는 일정 시간 동안(time interval) 객체의 위치가 변하는 객체를 의미하며 이러한 움직임 객체 A에 대한 움직임 경로는 시간 $t_0, t_1, t_2, \dots, t_n$ 시점에서 객체 A의 중심 좌표를 나타내는 일련의 $[(x_0, y_0, t_0), (x_1, y_1, t_1), \dots, (x_n, y_n, t_n)]$ 로 표현된다.

만약에, 움직임 객체의 시작 시점 t_0 에서의 위치와 임의의 시간 후의 t_i 에서의 위치를 고려하면, 두 위치 사

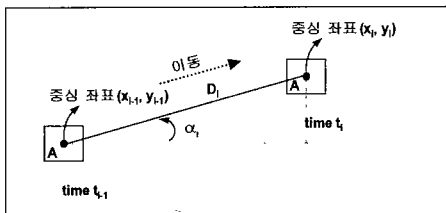


그림 5 MBR로 표현한 움직임 객체

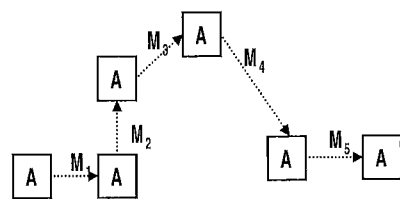
이의 방향(direction) 관계성을 구할 수 있으며, 여기에서 I_i 는 t_0 와 t_i 사이의 시간 간격(time interval)을 의미하며 또한 움직임 객체가 포함되어 있는 연속된 비디오 프레임들 중에서 마지막 프레임에서 시작 프레임의 차로 표현한다. 그리고 (x_0, y_0, t_0) 와 (x_i, y_i, t_i) 는 각각 t_0 에서의 움직임 객체의 위치와 t_i 에서의 움직임 객체의 위치를 의미한다. 이를 기반으로 본 논문에서는 단일 움직임 경로(single object's trajectory)를 정의 1과 같이 정의한다.

[정의 1] 객체 A가 움직임 객체일 때, 시간 간격 I_i 동안의 객체 A에 대한 움직임(M_i)은 (α_i, D_i, I_i) 로 표현된다. 여기에서, α_i 는 시간 간격 I_i 동안 움직인 방향으로서 $0^\circ \sim 360^\circ$ 까지의 실제 각도로서 나타낸다. 그리고 D_i 는 시간 간격 I_i 동안 움직임 객체가 움직인 이동 거리(moving distance)를 의미하며, 0에서 100까지의 상대적인 거리(relative distance)로서 표현한다. 따라서, 주어진 순서화된 시간 간격 리스트 $\langle I_1, I_2, \dots, I_n \rangle$ 에 대해서, 움직임 객체 A의 움직임 경로는 다음과 같이 일련의 움직임 리스트 $\langle M_1, M_2, \dots, M_n \rangle$ 로 표현된다.

$$\langle (\alpha_1, D_1, I_1), (\alpha_2, D_2, I_2), \dots, (\alpha_n, D_n, I_n) \rangle$$

[예 1] 그림 6에서 움직임 객체 A의 움직임 경로를 일련의 움직임(motion) 리스트로 표현하면 다음과 같다.

$$\langle (0^\circ, 15, I_1), (90^\circ, 15, I_2), (40^\circ, 23, I_3), (300^\circ, 32, I_4), (0^\circ, 15, I_5) \rangle$$



(a) 객체 A의 움직임 경로

$$\langle (0^\circ, 15, I_1), (90^\circ, 15, I_2), (40^\circ, 23, I_3), (300^\circ, 32, I_4), (0^\circ, 15, I_5) \rangle$$



(b) 시공간 관계성으로 표현한 움직임 리스트

그림 6 움직임 경로와 시공간 관계성으로 표현한 움직임 리스트

대부분의 비디오 장면에는 움직이는 객체(moving object)와 움직이지 않는 객체(stationary object)로 구성된 다수의 객체가 존재하며, 보다 효율적인 검색을 위해서는 이러한 객체들이 어떤 관계성을 가지면서 어떻게 움직이는가를 정확하게 표현할 수 있어야 한다. 따라서, 본 논문에서는 둘 이상의 객체로 이루어진 다수 움직임 경로(multiple objects' trajectories)를 모델링하기 위해 두 객체 사이의 공간 관계성과 시간 관계성을 모두 적용한다. 두 객체 사이의 공간 관계성을 표현하기 위해 FA(FarAway), DJ(DisJoint), ME(MEet), OL(OverLap), CL(Is inCLuded by), IN(INclude), SA(Same)의 7개의 위상 관계 연산자(topological relation operator)를 정의하며, 이러한 위상 관계 연산자들은 기존의 이미지 내의 두 객체간의 공간 관계를 표현하기 위해 Chang에 의해 제안된 SMR 방법[13]으로부터 생성된다. 그림 7은 1차원에서 나타낸 7개의 위상 관계 연산자를 그림으로 나타낸 것이다. 그러나, 실제로 비디오의 대표 프레임 장면에서는 각 객체가 2차원으로 나타내기 때문에, 두 객체간의 공간 관계를 표현하기 위해서는 2차원을 모두 고려해야 한다. 따라서, 2차원에서의 위상 관계 연산자는 표 2와 같다.

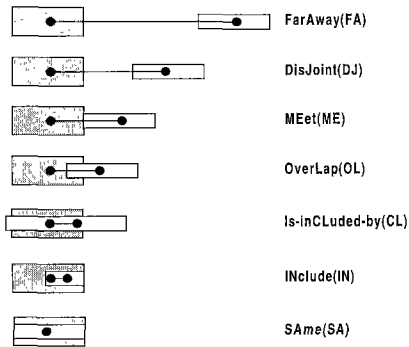


그림 7 7개의 위상 관계 연산자 (1차원)

표 2 2차원 상에서의 공간 관계

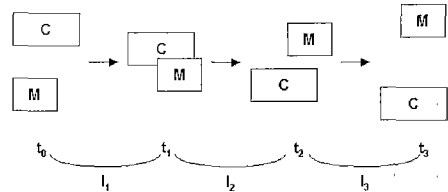
$R_{Y}^{AB} \setminus R_X^{AB}$	FA	DJ	ME	OL	CL	IN	SA
FA	FA	FA	FA	FA	FA	FA	FA
DJ	FA	DJ	DJ	DJ	DJ	DJ	DJ
ME	FA	DJ	ME	ME	ME	ME	ME
OL	FA	DJ	ME	OL	OL	OL	OL
CL	FA	DJ	ME	OL	CL	OL	CL
IN	FA	DJ	ME	OL	OL	IN	IN
SA	FA	DJ	ME	OL	CL	IN	SA

이를 기반으로 다수 움직임 경로는 정의2와 같다.

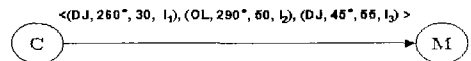
[정의 2] 객체 A와 객체 B가 움직임 객체일 때, 시간 간격 I_i 동안의 객체 A에 대한 객체 B의 시공간 관계(STR)는 $(R_i, \alpha_i, D_i, I_i)$ 로 표현된다. 여기에서, R_i 는 시간 간격 I_i 동안의 객체 A와 객체 B사이의 위상 관계를 의미하며, 위에서 언급한 7개의 위상 관계 연산자 중의 하나를 갖는다. α_i 는 시간 간격 I_i 의 시작 프레임(start frame)에서 객체 A에 대한 객체 B의 방향을 $0^\circ \sim 360^\circ$ 까지의 실제 각도로서 나타낸다. 그리고 D_i 는 시간 간격 I_i 동안 객체 A(기준 객체)에 대한 객체 B(상대 객체)의 상대적인 이동 거리(relative moving distance)로 0에서 100까지의 정규화된 값으로 나타낸다. 즉, 객체 A의 이동 거리와 객체 B의 이동 거리가 같을 때 D_i 를 50으로 정하고, 이를 기준으로 객체 A의 이동 거리가 객체 B의 이동 거리보다 클 경우에는 50에서 100 사이의 값을 갖으며, 그와 반대일 경우에는 0에서 50 사이의 값을 갖는다. 따라서, 주어진 순서화된 시간 간격 리스트 $\langle I_1, I_2, \dots, I_n \rangle$ 에 대해서, 객체 A로부터 객체 B에 대한 다수 움직임 경로는 다음과 같이 시공간 관계의 리스트 $\langle STR_1, STR_2, \dots, STR_n \rangle$ 로 표현된다.

$$\langle (R_1, \alpha_1, D_1, I_1), (R_2, \alpha_2, D_2, I_2), (R_3, \alpha_3, D_3, I_3), \dots, (R_n, \alpha_n, D_n, I_n) \rangle$$

[예 2] 그림 8은 “자동차(Car)와 오토바이(Motorcycle)가 경주를 벌이는 비디오 장면을 나타낸 것이다.”, 객체 C와 객체 M 간의 다수 움직임 경로를 표현하면



(a) 자동차(Car)와 오토바이(Motorcycle) 객체에 대한 움직임 경로



(b) 자동차(Car)로부터 오토바이(Motorcycle)에 대한 시공간 관계성

그림 8 다수 움직임 경로를 시공간 관계성으로 표현한 예

다음과 같다.

<(DJ, 260° , 30, I₁), (OL, 290° , 50, I₂), (DJ, 45° , 55, I₃)>

4. 유사성 측정 알고리즘

기존의 시공간 관계성을 위한 유사성 측정 알고리즘은 비디오 데이터에서 움직임 객체에 대한 사용자 질의와 데이터베이스간의 유사성을 측정하기 위해 대부분 공간 관계성만을 이용하고 있다. 그러나, 추후 비디오 데이터와 같은 응용 분야에서는 공간 관계성과 더불어 일정 시간 간격 동안 움직인 객체의 이동 거리가 사용자 질의와 데이터베이스간의 유사성을 측정하는 데 있어 매우 중요한 역할을 한다. 따라서, 본 논문에서는 움직임 객체의 움직임 경로에 대한 유사성 측정을 위해 공간 관계성과 움직임 객체의 이동 거리를 모두 고려하여 보다 정확하고 효율적인 유사성 검색이 가능한 새로운 유사성 측정 알고리즘을 제안한다. 아울러, 이는 사용자의 질의에 부합하는 결과를 유사성이 높은 순으로 사용자에게 제공할 수 있다.

본 논문에서 제안하는 단일 움직임 경로를 위한 유사성 측정 알고리즘은 SDST(Similarity measure based on moving Distance for Single object's Trajectory)라고 명명하며, 이는 다음과 같다.

[정의 3] 비디오 데이터베이스에서 주어진 단일 객체의 움직임 경로가 VS=<VS₁, VS₂, ..., VS_M>이고, 질의 객체의 움직임 경로가 QS=<QS₁, QS₂, ..., QS_N> (1 ≤ N ≤ M) 일 때, VS_i와 QS_j사이의 각도의 차 D_{ang}(VS_i, QS_j)는 다음과 같다.

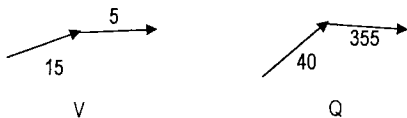
If |VS_i - QS_j| > 180° then (1)

$$D_{ang}(VS_i, QS_j) = (360^\circ - |VS_i - QS_j|)$$

Else

$$D_{ang}(VS_i, QS_j) = |VS_i - QS_j|$$

[예 3] 비디오 움직임 경로 VS=<15° , 5° >와 질의 움직임 경로 QS=<40° , 355° >가 아래와 같이 주어졌을 때, 각도의 차 D_{ang}를 구하면 다음과 같다.



$$D_{ang}(VS_1, QS_1) = |15^\circ - 40^\circ| = 25^\circ$$

$$D_{ang}(VS_2, QS_2) = (360^\circ - |5^\circ - 355^\circ|) = 10^\circ$$

[정의 4] 비디오 데이터베이스에서 주어진 단일 객체의 움직임 경로가 VS=<VS₁, VS₂, ..., VS_M>이고, 질의 객체의 움직임 경로가 QS=<QS₁, QS₂, ..., QS_N> (1 ≤ N ≤ M) 일 때, 비디오 객체의 움직임 경로 VS_i와 질의 객체의 움직임 경로 QS_j사이의 방향 관계에 대한 유사성 SR(VS_i, QS_j)는 다음과 같다.

$$SR(VS_i, QS_j) = \frac{\cos(D_{ang}(VS_i, QS_j)) + 1}{2} \quad (2)$$

여기에서, 각도의 차(0° ≤ D_{ang} ≤ 180°)에 코사인(cosine) 함수를 적용하면 -1에서 1까지의 값을 갖게 되며, 여기에 1을 더한 후에 2로 나눔으로써 SR(VS_i, QS_j)값은 항상 0에서 1사이의 값을 갖는다.

[예 4] 예 3 에서와 같이 Dang(VS₁, QS₁) = 25° 이고 D_{ang}(VS₂, QS₂) = 10° 라고 할 때, 비디오 객체의 움직임 경로 VS_i와 질의 객체의 움직임 경로 QS_j사이의 방향 관계에 대한 유사성 SR(VS_i, QS_j)를 계산하면 다음과 같다.

$$SR(VS_1, QS_1) = \frac{\cos(25^\circ) + 1}{2} = \frac{0.91 + 1}{2} = \frac{1.91}{2} = 0.96$$

$$SR(VS_2, QS_2) = \frac{\cos(10^\circ) + 1}{2} = \frac{0.98 + 1}{2} = \frac{1.98}{2} = 0.99$$

[정의 5] 비디오 데이터베이스에서 주어진 단일 객체의 움직임 경로가 VS=<VS₁, VS₂, ..., VS_M>이고, 질의 객체의 움직임 경로가 QS=<QS₁, QS₂, ..., QS_N> (1 ≤ N ≤ M) 일 때, 비디오 객체의 움직임 경로 VS_i와 질의 객체의 움직임 경로 QS_j사이의 이동 거리에 대한 유사성 SD(VS_i, QS_j)는 다음과 같다.

$$SD(VS_i, QS_j) = 1 - \frac{|VS_i - QS_j|}{Max(VS_i, QS_j)} \quad (3)$$

[예 5] 비디오에서의 단일 객체의 움직임 경로 VS=<(15° , 50, I₁), (5° , 50, I₂)>이고, 질의 객체의 움직임 경로 QS=<(40° , 45, I₁), (355° , 55, I₂)> 일 때, 비디오 객체의 움직임 경로 VS_i와 질의 객체의 움직임 경로 QS_j사이의 이동 거리에 대한 유사성 SD(VS_i, QS_j)를 계산하면 다음과 같다.

$$SD(VS_1, QS_1) = 1 - \frac{|50 - 45|}{50} = 1 - \frac{5}{50} = 1 - 0.1 = 0.9$$

$$SD(VS_2, QS_2) = 1 - \frac{|50 - 55|}{55} = 1 - \frac{5}{55} = 1 - 0.09 = 0.91$$

[정의 6] 비디오 데이터베이스에서 주어진 단일 객체의 움직임 경로가 VS=<VS₁, VS₂, ..., VS_M>이고, 질의 객체의 움직임 경로가 QS=<QS₁, QS₂, ..., QS_N> (1 ≤

$N \leq M$ 일 때, 식(2)와 식(3)을 이용하여 비디오 객체의 단일 움직임 경로 VS와 질의 객체의 움직임 경로 QS사이의 최종적인 유사성 SDST(VS, QS)는 다음과 같다. 여기에서, ω_1 과 ω_2 는 각각 방향 관계와 이동 거리에 대한 가중치 값을 의미한다.

$$SDST(VS, QS) = \text{MAX} \left\{ \frac{\sum_{j=1}^N SR_{i+j}^{(1-\omega_1)} * SD_{i+j}^{(1-\omega_2)}}{N} \right\} \quad (\forall j, 0 \leq j \leq M-N)$$

[정의 6]에 기반하여 SDST 알고리즘을 기술하면 다음과 같다.

SDST(VS, QS)

입력 : VS={VS₁, VS₂, ..., VS_M}, QS={QS₁, QS₂, ..., QS_N}

출력 : Similarity_Value

float Max, Direction, Distance ;
float Similarity_Value ;

Similarity_Value = 0.0 ;

if (N>M)

 return Similarity_Value ;

for (j=0 ; j<M-N+1 ; j++)

 Max = Direction = Distance = 0.0 ;

 for (i=1 ; i<N+1 ; i++)

 {
 Direction = $\frac{\cos(D_{ang}(VS_{i+j}, QS_i)) + 1}{2}$;

 Distance = $1 - \frac{|VS_{i+j} - QS_i|}{\text{Max}(VS_{i+j}, QS_i)}$;

 Max = Max + Direction^(1- ω_1) * Distance^(1- ω_2) ;

 }

 Max = Max / N ;

 if (Max > Similarity_Value)

 Similarity_Value = Max ;

 }

return Similarity_Value ;

아울러, 다수 움직임 경로를 위한 유사성 측정 알고리즘은 SDMT(Similarity measure based on moving Distance for Multiple objects' Trajectory)라고 명명하며, 이는 위상 관계, 방향 관계, 그리고 이동 거리를 모두 결합한 알고리즘으로 다음과 같다.

[정의 7] 비디오 데이터베이스에서 주어진 다수 객체의 움직임 경로가 VM=<VM₁, VM₂, ..., VM_M>이고, 질의 객체의 움직임 경로가 QM=<QM₁, QM₂, ..., QM_N> (1≤N≤M)일 때, VM_i와 QM_i사이의 위상 관계

에 대한 유사성 ST(VM_i, QM_i)는 다음과 같이 정의한다.

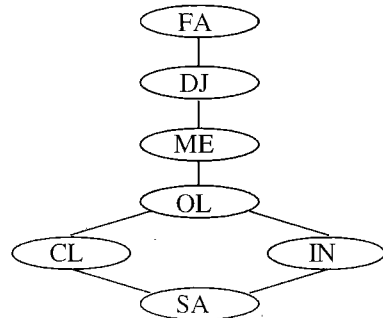
$$ST(VM_i, QM_i) = \frac{\lambda}{\lambda + \text{Sim_Dist}(VR_i, QR_i)^2} \quad (4)$$

여기에서, Sim_Dist(VR_i, QR_i)는 그림 9(a)에서와 같이 VR_i과 QR_i사이의 유사성 거리를 의미하며, 이는 그림 9(b)에서 각 노드와 인접한 노드 간의 거리를 1로 간주하고 노드와 노드 사이의 최단 거리를 구한 것이다. 예를 들어, FA와 DJ 간의 유사성 거리는 1, FA와 ME 간의 유사성 거리는 2, 그리고 FA와 SA 간의 유사성 거리는 5를 의미한다. 그리고 ST(VM_i, QM_i)의 값을 0에서 1 사이의 값으로 정규화하고 아울러, 위상 관계간의 유사성을 완만하게 나타내기 위해 λ 를 사용한다.

[예 6] 비디오에서의 다수 객체의 움직임 경로VM=<(DJ, 260° , 30, I₁), (OL, 290° , 15, I₂)>이고, 질의

D	FA	DJ	ME	OL	CL	SA	IN
FA	0	1	2	3	4	5	4
DJ	1	0	1	2	3	4	3
ME	2	1	0	1	2	3	2
OL	3	2	1	0	1	2	1
CL	4	3	2	1	0	1	2
SA	5	4	3	2	1	0	1
IN	4	3	2	1	2	1	0

(a) 위치 연산자간의 유사성 거리 (similarity distance:Sim_Dist)



(b) 위치 연산자 간의 유사성

그림 9 위상 관계 연산자간의 유사성 거리

객체의 움직임 경로 QM=<(ME, 230° , 40, I₁), (FA,

330°, 70, I₂)>일 때, VM_i와 QM_i사이의 위상 관계에 대한 유사성 ST(VM_i, QM_i)를 계산하면 다음과 같다.

$$ST(VM_1, QM_1) = \frac{10}{10 + (1)^2} = \frac{10}{11} = 0.91$$

$$ST(VM_2, QM_2) = \frac{10}{10 + (3)^2} = \frac{10}{19} = 0.53$$

[정의 8] 비디오 데이터베이스에서 주어진 다수 객체의 움직임 경로가 VM=<VM₁, VM₂, ..., VM_M>이고, 질의 객체의 움직임 경로가 QM={QM₁, QM₂, ..., QM_N} (1≤N≤M) 일 때, 비디오 객체의 움직임 경로 VM_i와 질의 객체의 움직임 경로 QM_i사이의 이동 거리에 대한 유사성 SD(VM_i, QM_i)는 다음과 같다.

$$SD(VM_i, QM_i) = 1 - \frac{|VM_i - QM_i|}{100} \quad (5)$$

[예 7] 비디오에서의 다수 객체의 움직임 경로 VM=<(DJ, 260°, 30, I₁), (OL, 290°, 15, I₂)>이고, 질의 객체의 움직임 경로 QM=<(ME, 230°, 40, I₁), (FA, 330°, 70, I₂)>일 때, VM_i와 QM_i사이의 위상 관계에 대한 유사성 SD(VM_i, QM_i)를 계산하면 다음과 같다.

$$SD(VM_1, QM_1) = 1 - \frac{|30 - 40|}{100} = 1 - \frac{10}{100} = 1 - 0.1 = 0.9$$

$$SD(VM_2, QM_2) = 1 - \frac{|15 - 70|}{100} = 1 - \frac{55}{100} = 1 - 0.55 = 0.45$$

[정의 9] 비디오 데이터베이스에서 주어진 다수 객체의 움직임 경로가 VM={VM₁, VM₂, ..., VM_M}이고, 질의 객체의 움직임 경로가 QM={QM₁, QM₂, ..., QM_N} (1≤N≤M) 일 때, 식(2), 식(4), 그리고 식(5)를 이용하여 비디오 객체의 움직임 경로 VM과 질의 객체의 움직임 경로 QM사이의 최종적인 유사성 SDMT(VM, QM)는 다음과 같다. 여기에서, ω₁, ω₂, ω₃는 각각 위상 관계, 방향 관계 그리고 이동 거리에 대한 가중치 값을 의미한다.

$$SDMT(VM, QM) = \text{MAX} \left\{ \frac{\sum_{i=1}^N ST_{i+j}^{(1-\omega_1)} * SR_{i+j}^{(1-\omega_2)} * SD_{i+j}^{(1-\omega_3)}}{N} \right\} \quad (\forall j, 0 \leq j \leq M-N)$$

[정의 9]에 근거하여 SDMT 알고리즘을 기술하면 다음과 같다.

SDMT(VM, QM)

입력 : VM={VM₁, VM₂, ..., VM_M}, QM={QM₁, QM₂, ..., QM_N}

출력 : Similarity_Value

float Max, Topology, Direction, Distance ;

float Similarity_Value ;

Similarity_Value = 0.0 ;

if (N>M)

return Similarity_Value ;

for (j=0 ; j<M-N+1 ; j++){

Max = Topology = Direction = Distance = 0.0 ;

for (i=1 ; i<N+1 ; i++)

{

$$\text{Topology} = \frac{10}{10 + \text{Sim_Dist}(VR_{i+j}, QR_i)} ;$$

$$\text{Direction} = \frac{\cos(D_{\text{ms}}(VM_{i+j}, QM_i)) + 1}{2} ;$$

$$\text{Distance} = 1 - \frac{|VM_{i+j} - QM_i|}{100} ;$$

$$\text{Max} = \text{Max} + \text{Topology}^{(1-\omega_1)} * \text{Direction}^{(1-\omega_2)} * \text{Distance}^{(1-\omega_3)} ;$$

}

Max = Max / N ;

if (Max > Similarity_Value)

Similarity_Value = Max ;

}

return Similarity_Value ;

5. 실험 및 성능평가

본 논문에서 제안하는 시공간 표현 기법의 유용성을 보이기 위해, 실험 데이터로서 축구 비디오 데이터를 선정하여 성능 평가를 수행하였다. 축구 비디오 데이터는 사용자의 주된 관심 객체인 축구공이 경기장을 배경으로 많은 움직임 경로를 가지는 특징을 지니고 있어 객체의 움직임 경로를 추출하는 데 매우 용이하다. 실험을 위하여 다음과 같은 데이터 집합을 사용한다.

- 실험 데이터로서 실제 축구 비디오 데이터 350개를 사용한다.
- 각각의 축구 비디오 데이터는 1개 ~ 15개의 움직임 정보들(motions)로 구성된다.
- 질의를 구성하기 위하여 축구 비디오 데이터로부터 추출된 40개의 질의 비디오 데이터를 만든다.
- 각각의 질의 비디오 데이터는 2 ~ 3개의 움직임 정보들(motions)로 구성된다.

실험에 사용되는 축구 비디오 데이터는 mpeg 동영상 파일 포맷으로 대부분이 “꼴을 넣는 장면”을 포함하고 있다. 또한 축구 비디오 데이터로부터 움직임 경로의 추출은 수작업을 통해 이루어지는 데, 정확한 움직임 경로의 추출을 위해 카메라가 Zoom Out된 상태 즉, 축구공,

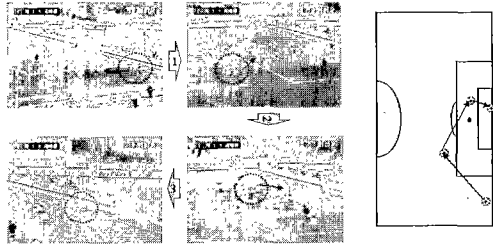


그림 10 축구 비디오 데이터에서 움직임 경로의 예

선수들, 골대 등의 객체가 전체적으로 화면상에 나타나는 샷들만을 선별하여 움직임 경로를 추출한다. 그림 10은 축구 비디오 데이터로부터 움직임 정보를 추출하는 예이다.

또한, 350개의 비디오 데이터로 구성된 축구 비디오 데이터베이스로부터 실제로 '골인'이 되는 대표적인 움직임 경로를 기준으로 축구 경기장 중앙선(하프라인)의 오른쪽 진영(right field)과 왼쪽 진영(left field)에서 각각 20개씩 총 40개의 질의를 추출하였다. 한편, 이들 질의는 일반적인 비디오 질의로는 부적당하며, 현재 실험을 위해 축구 비디오 데이터에 한정하여 사용한다. 추출된 질의는 질의 비디오 데이터에 포함된 움직임의 수(# of motion)와 움직임 경로가 일어나는 위치에 따라 크게 4가지로 분류된다.

- 질의 타입 1 : 2개의 움직임(motion)을 가지면서 오른쪽 진영에서 일어나는 10개 질의 (질의 1 ~ 질의 10)
- 질의 타입 2 : 2개의 움직임(motion)을 가지면서 왼쪽 진영에서 일어나는 10개 질의 (질의 11 ~ 질의 20)
- 질의 타입 3 : 3개의 움직임(motion)을 가지면서 쪽 진영에서 일어나는 10개 질의 (질의 21 ~ 질의 30)
- 질의 타입 4 : 3개의 움직임(motion)을 가지면서 왼쪽 진영에서 일어나는 10개 질의 (질의 31 ~ 질의 40)

그림 11과 그림 12는 각각 질의 타입 1과 질의 타입 3에 속하는 질의 비디오 데이터를 그림으로 나타낸 것이다.

성능 평가를 위하여 128MB 메인 메모리를 갖는 Pentium-PC 상에서 Microsoft Visual C++ 컴파일러를 사용하여 Li 방법과 Shan 방법 그리고 본 논문에서 제안하는 방법을 실제 프로그래밍 하여 실험을 수행하였으며, 검색 효율 측면에서 비교 수행하였다. 검색 효

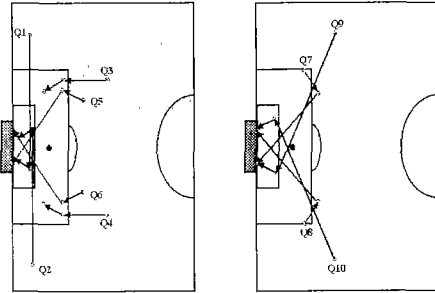


그림 11 타입 1에 속하는 질의 10개의 예

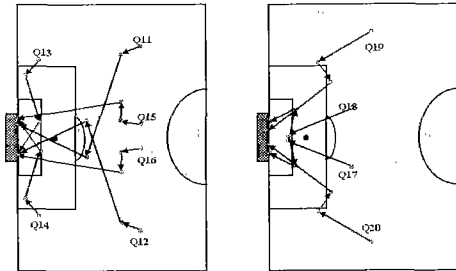


그림 12 타입 3에 속하는 질의 10개의 예

율을 평가하기 위하여 정확율(Precision)과 재현율(Recall)[13]을 사용하였다. 먼저, RVQ(Relevant Video data to Query)는 주어진 질의에 대한 시스템의 검색한 검색 결과의 수, RVD(Relevant Video data in Database)는 주어진 질의에 대해 유사성이 있어 데이터베이스로부터 반드시 검색되어야 할 검색 결과의 수, 마지막으로 RVR(Relevant Video data that are Retrieved)는 질의에 대해 시스템이 검색한 결과 중에서 유사성을 가지고 있는 검색 결과의 수 즉, RVQ 중에서 RVD에 속하는 결과의 수라고 정의한다. 아울러, RVD를 구하기 위해 10명의 컴퓨터 관련 대학원생을 통하여 수작업을 통해 주어진 질의에 대해 비디오 데이터베이스로부터 유사성이 높은 비디오 데이터를 선택하였다. 따라서, 정확율과 재현율은 다음 식 (6)으로 표현된다.

$$\begin{aligned} \text{정확율(Precision)} &= \frac{RVR}{RVQ} \\ \text{재현율(Recall)} &= \frac{RVR}{RVD} \end{aligned} \quad (6)$$

본 논문에서는 보다 정확한 정확율과 재현율을 구하기 위하여 가장 널리 사용되는 평균 11 point 방법을

사용하였다.[14] 표 3은 각각의 방법에 대한 평균 정확율과 평균 재현율을 나타낸다. 실험 결과를 통해 알 수 있듯이 제안하는 방법이 기존의 Li 방법보다 정확율과 재현율 모든 측면에서 우수함을 알 수 있다. 다시 말하면, 제안하는 방법이 정확율 측면에서는 약 20%, 재현율 측면에서는 약 10% 정도의 검색 효율이 좋아짐을 볼 수 있다. 또한, Shan 방법보다는 재현율 측면에서는 비슷하나 정확율 측면에서는 약 17% 정도의 더 나은 검색 성능을 보인다. 그림 13은 평균 11 point 방법을 이용하여 각각의 방법에 대해서 재현율을 기준으로 정확율을 구한 결과를 그래프로 나타낸 것이다.

표 3 검색 효율 비교

	검색 효율	
	평균 정확율	평균 재현율
Li 방법	0.23	0.42
Shan 방법	0.26	0.46
제안하는 방법	0.43	0.44

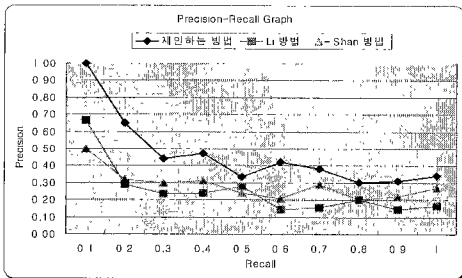


그림 13 재현율-정확율 그래프

6. 결론 및 향후 연구

본 논문에서는 비디오 데이터에서 움직임 객체의 움직임 경로를 모델링 하기 위한 새로운 시공간 표현 기법을 제안하였다. 비디오 데이터를 위한 보다 효율적인 내용 기반 검색을 지원하기 위해, 제안하는 기법은 시간, 공간 관계성과 더불어 일정 시간 간격 동안 움직인 객체의 이동 거리를 고려한다. 아울러, 제안하는 표현 기법을 기반으로 단일 움직임 객체의 움직임 경로와 다수 움직임 객체들의 움직임 경로를 위한 새로운 유사성 측정 알고리즘인 SDST와 SDMT를 제안하였다. 제안하는 유사성 측정 알고리즘은 사용자 질의에 대해서 유사성을 측정하고, 검색 결과에 대해 유사성이 높은 순으로 브라우징 할 수 있다. 또한 성능 실험을 통하여 제안한 기법이 기존의 Li 방법과 Shan의 방법에 비해 동등

한 재현율을 유지하며, 정확율 측면에서 약 20%의 성능 향상을 보임을 알 수 있었다.

본 논문의 향후 연구 방향으로는 본 논문에서 제안하는 새로운 시공간 표현 기법을 대용량의 비디오 데이터에 적용하여 추가적인 성능 평가를 수행하는 것이다.

참고 문헌

- [1] W. Niblack, et. al., "The QBIC project: Querying by Image Content Using Color, Texture, and Shape," in Proceedings of SPIE Storage and Retrieval for Image and Video Databases, pp.173-187, 1993.
- [2] 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.
- [3] T.D.C. Little, G. Ahanger, R.J. Folz, et. al., "A Digital On-Demand Video Service Supporting Content-Based Queries," in Proceedings of ACM Multimedia 93, pp. 427-436, 1993.
- [4] Virginia, E.Ogle and Michael Stonebraker, "Chabot: Retrieval from a Relational Database of images," IEEE Computer, Vol. 28, No. 9, pp. 40-48, 1995.
- [5] 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.
- [6] 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.
- [7] John Z. Li, M. Tamer Ozsu, Duane Szafron, "Modeling Video Temporal Relationships in an Object Database Management System," in Proceedings of Multimedia Computing and Networking(MMCN97), pp. 80-91, 1997.
- [8] John Z. Li, M. Tamer Ozsu, Duane Szafron, "Modeling of Video Spatial Relationships in an Objectbase Management System," in Proceedings of International Workshop on Multimedia DBMS, pp. 124-133, 1996.
- [9] Man-Kwan Shan and Suh-Yin Lee, "Content-based Video Retrieval via Motion Trajectories," in Proceedings of SPIE Electronic Imaging and Multimedia System II, Vol. 3561, pp. 52-61, 1998.
- [10] J. F. Allen, "Maintaining Knowledge about Temporal Intervals," Communication of the ACM, Vol. 26, No. 11, pp. 832-843, 1983.

- [11] S. K. Chang, Q. Y. Shi and C. W. Yan, "Iconic Indexing by 2D Strings," IEEE Trans. Pattern Analysis, Machine Intelligence, Vol. 9, No. 3, pp. 413-428, 1987.
- [12] J. W. Chang, Y. J. Kim and K. J. Chang, "A Spatial Match Representation Scheme Indexing and Querying in Iconic Image Databases," ACM International Conference on Information and Knowledge Management, pp. 169-176, 1997.
- [13] G. Salton, "A New Comparison between Conventional Indexing(MEDLARS) and Automatic Text Processing(SMART)," Journal of the American Society for Information Science, Vol. 23, No. 2, pp 75-84, 1972.
- [14] Salton, G., M. McGill, An introduction to Modern Information Retrieval, McGraw-Hill, 1993.



심 춘 보

1996년 전북대학교 컴퓨터공학과(공학사)
 1998년 전북대학교 컴퓨터공학과(공학석사). 2000년 전북대학교 컴퓨터공학과 박사과정 수료. 관심분야는 멀티미디어 데이터베이스, 멀티미디어 정보 검색, 데이터베이스 접근기법 등



장 재 우

1984년 서울대학교 전자계산기공학과(공학사). 1986년 한국과학기술원 전산학과(공학석사). 1991년 한국과학기술원 전산학과(공학박사). 1996년 ~ 1997년 Univ. of Minnesota, Visiting Scholar. 1991년 ~ 현재 전북대학교 컴퓨터공학과 부교수. 관심분야는 멀티미디어 데이터베이스, 멀티미디어 정보 검색, 하부저장구조.