

내용 및 유사도 검색을 위한 움직임 객체 모델링

복경수[†], 김미희^{**}, 신재룡^{***}, 유재수^{****}, 조기형^{*****}

요 약

비디오 데이터에는 시간의 변화에 따라 공간적인 위치가 변화하는 움직임 객체를 포함하고 있다. 이 논문에서는 비디오 데이터의 움직임 객체에 대한 새로운 모델링 방법을 제안한다. 제안하는 모델링은 움직임 객체를 효과적으로 검색하기 위해 시간의 변화에 따라 공간적인 위치와 크기 변화를 표현한다. 또한 객체의 시간에 따른 시각적 특징 변화와 객체의 방향, 거리 그리고 속도를 고려한 궤적을 표현한다. 따라서 움직임 객체의 시각적인 특징 유사도 검색, 거리 유사도 검색, 궤적 유사도 검색을 수행할 수 있다. 또한 이들을 통합한 가중치 검색이 가능하도록 한다.

Moving Objects Modeling for Supporting Content and Similarity Searches

Kyoung Soo Bok[†], Mi Hee Kim^{**}, Jae Ryong Shin^{***},
Jae Soo Yoo^{****}, Ki Hyung Cho^{*****}

ABSTRACT

Video Data includes moving objects which change spatial positions as time goes by. In this paper, we propose a new modeling method for a moving object contained in the video data. In order to effectively retrieve moving objects, the proposed modeling method represents the spatial position and the size of a moving object. It also represents the visual features and the trajectory by considering direction, distance and speed of moving objects as time goes by. Therefore, It allows various types of retrieval such as visual feature based similarity retrieval, distance based similarity retrieval and trajectory based similarity retrieval and their mixed type of weighted retrieval.

Key words: Moving Object(움직임 객체), Content Search(내용 검색), Similarity Search(유사도 검색), Trajectory(궤적), Video Modeling(비디오 모델링)

1. 서 론

전세계적으로 인터넷과 초고속 정보 통신망의 구축 등으로 멀티미디어에 대한 관심이 집중되면서 멀

티미디어 컴퓨팅, 통신 그리고 이에 대한 응용 기술들이 급속도로 발전하고 있다. 멀티미디어 데이터들은 텍스트와 숫자 데이터에 비해 일정한 형식이 없는 대용량의 데이터 특성을 갖는다. 멀티미디어 데이터

※ 교신저자(Corresponding Author) : 유재수, 주소 : 충북 청주시 흥덕구 개신동 산 48번지(361-763) 전화 : 043)261-3230, FAX : 043)276-9961, E-mail : yjs@cbucc.chungbuk.ac.kr
접수일 : 2002년 11월 11일, 완료일 : 2003년 3월 17일

[†] 충북대학교 대학원 정보통신공학과 및 컴퓨터정보통신 연구소 박사과정

(E-mail : ksbok@netdb.chungbuk.ac.kr)

^{**} 충북대학교 정보통신공학과 공학석사

(E-mail : mhkim@netdb.chungbuk.ac.kr)

^{***} 광주보건대학 인터넷정보과 전임강사

(E-mail : sjr@www.kjhc.ac.kr)

^{****} 정회원, 충북대학교 정보통신공학과 및 컴퓨터정보통신 연구소 부교수

^{*****} 충북대학교 정보통신공학과 및 컴퓨터정보통신 연구소 교수

(E-mail : khjoe@cbucc.chungbuk.ac.kr)

※ 본 연구는 한국과학재단 목적기초연구(R01-2003-000-10627-0)지원으로 수행되었음.

의 대표적인 비디오 데이터는 대용량이고 비정형화인 데이터 특성뿐만 아니라 시청각 정보, 시공간 정보를 포함하고 있다. 또한 객체와 사건과 같은 의미 정보 등의 일반 데이터들이 가지지 않는 정보를 포함하고 있다. 이러한 비디오 데이터를 효율적으로 저장하고 검색하기 위해서는 응용 목적에 따라 사용자의 요구와 질의 형태를 분석하여 적절하게 모델링하는 작업이 필요하다. 또한 다양한 사용자의 요구를 처리하기 위한 새로운 검색 기법의 정비가 요구된다. 최근 몇 년간, 방대한 비디오 정보의 모델링과 검색 기법에 대해 많은 연구가 진행되고 있으며, OVID[1], VideoSTAR[2], JACOB[3], VideoQ[4], Informedial[5], AVIS[6]과 같은 많은 비디오 검색 시스템들이 개발되었다.

초기에는 비디오의 논리적인 구조 또는 비디오에 나타난 의미 정보와 시각적 특징들을 표현할 수 있는 모델링 기법들이 연구되었다. 이러한 모델링 기법들은 크게 구조 모델링(structure modeling)과 내용 모델링(content modeling)으로 분류될 수 있다. 구조 모델링은 비정형화된 비디오 데이터를 샷, 씬, 시퀀스와 같은 논리적인 단위들로 구조화하는 모델링이다[7,8]. 이에 반해 내용 모델링은 비디오에 포함된 시각적인 특징과 객체와 사건과 같은 의미적(semantic)인 정보를 중심으로 모델링한다[9,10]. 그러나 구조 모델링 기법은 비디오 데이터를 논리적인 구조 정보에 의해 모델링하므로 비디오의 내용 정보를 효과적으로 표현할 수 없다. 또한 대부분의 내용 모델링 기법은 비디오의 내용 정보를 논리적인 구조 계층 내에 표현하기 때문에 객체와 사건과 같은 의미적 내용이 구조 정보에 종속되어 표현된다는 문제점이 있다. 비디오의 구조 정보에 의해 의미 정보가 상실되거나 존속 구간을 판별하기가 어렵다. 따라서 비디오 내의 구조 정보와 관계없이 존속 구간을 가지고 나타나는 움직임 객체와 같은 의미 정보를 표현할 수 있는 새로운 모델링 기법들이 필요하게 되었다.

비디오에 존재하는 움직임 객체(moving object)는 시간의 변화에 따라 공간의 변화를 가지는 시공간적 특성을 가진다. 따라서 움직임 객체를 효과적으로 표현하기 위해서는 비디오의 구조 정보와 관계없이 객체의 존속 구간에 따른 시각적인 특징 변화를 나타낼 수 있어야 한다. 또한 움직임 객체의 움직임 모양을 파악할 수 있는 궤적(trajecory)을 표현하여 움직임 객체의 공간적인 위치 변화를 알 수 있어야 한다.

궤적의 명시를 통해 객체의 움직임 패턴을 분석하여 추후의 움직임 예측할 수 있다. 움직임 객체의 궤적을 이용한 응용들에는 도로 교통 시스템, 스포츠 경기, 동물 군의 움직임 관찰 시스템, 현미경 내의 세포 운동 관찰 등으로 의학, 동물학, 스포츠학 등의 많은 분야에 중요하게 적용되고 있다.

그러나 기존의 움직임 객체에 대한 연구들은 대부분이 객체의 시간에 따른 공간적인 위치만 표현하기 때문에 공간적 크기 변화 또는 다른 객체와의 거리 변화를 고려하지 않고 있으며 궤적을 표현하기 위해 일반적으로 방향만을 고려하고 있다는 문제점이 있다[11-13,18]. 또한 시각적인 특징 궤적만 지원하고 있다거나 공간적인 궤적만 지원하고 있다는 제한점을 가지고 있다.

이 논문에서는 새로운 움직임 객체 모델링 및 검색 기법을 제안한다. 제안하는 기법은 기존의 모델링 기법들의 문제점을 해결하고 제한점을 보완하기 위해 비디오의 구조적 정보와 관계없이 움직임 객체를 중심으로 표현한다. 객체는 공간적인 위치뿐만 아니라 공간적 크기 변화를 함께 표현한다. 또한 시간에 따른 시각적 특징 변화를 표현한다. 따라서 제안하는 모델링은 의미 기반 검색, 시공간적 특징을 기반으로 하는 시각적 특징 기반의 유사도 검색, 움직임 객체 간의 거리 유사도 검색, 그리고 움직임 객체의 방향, 거리, 속도를 고려한 궤적의 유사도 검색이 가능하도록 한다. 또한 움직임 객체간의 시각적 특징, 거리 그리고 궤적에 가중치를 부여한 가중치 혼합 유사성 검색을 지원한다.

이 논문의 구성은 다음과 같다. 2장에서는 움직임 객체를 표현하는 기존 모델링 기법에 대해 살펴본다. 3장에서는 제안하는 움직임 객체 모델링에 대해서 기술하고, 4장에서는 제안하는 모델링에서 지원하는 검색 기법에 대해 기술한다. 5장에서는 움직임 객체를 위한 검색 시스템을 구현한 내용과 실험 및 자체 성능 평가한 내용을 기술하고, 마지막으로 6장에서는 결론 및 향후 연구 방향에 대해 기술한다.

2. 관련연구

2.1 기존에 제안된 모델링

Zaher Aghbari는 객체의 시간의 변화에 따른 시각적 특징의 변화와 다른 객체와의 관계를 표현하기

위한 모델링을 제안하였다[11]. Zaher Aghbari은 비디오 데이터를 세 개의 구조적 계층으로 나누고 객체의 시각적 특징에 대한 궤적 정보를 표현하였다. 비디오 데이터는 객체에 대한 정보를 표현하는 OL (Object Level), 프레임 내의 정보를 표현하는 FL (Frame Level), 수십 개의 프레임들이 모여서 이론적인 단위 SL(Shot Level)로 구성된다.

OL은 모델링의 최하위 계층으로 각 객체들에 대한 시각적 특징을 표현한다. 시각적 특징의 요소에는 칼라, 모션 그리고 절대적 위치가 있다[14]. 칼라는 (식 1)과 같이 특정 시간 간격으로 추출한 키프레임 K_i 의 칼라 히스토그램 C 에 의해 표현하며 칼라 히스토그램은 객체를 표현하는 칼라의 최대 수인 l 만큼의 r_i, g_i, b_i 값으로 나타낸다. 모션은 (식 2)과 같이 객체의 움직임으로 키프레임에서 객체의 모션 히스토그램 M 에 의해 표현된다. 이 때, η 는 방향이고 v 는 속도이다. 절대적 위치는 키프레임에서 (식 3)와 같이 각 객체의 MBR 정보를 표현한다.

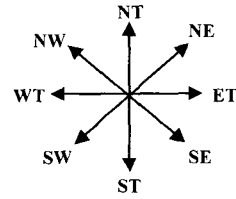
$$C : \{(r_1, g_1, b_1, p_1), (r_2, g_2, b_2, p_2), \dots, (r_l, g_l, b_l, p_l), K_i\} \quad (\text{식 1})$$

$$M : \{(\eta, v), K_i\} \quad (\text{식 2})$$

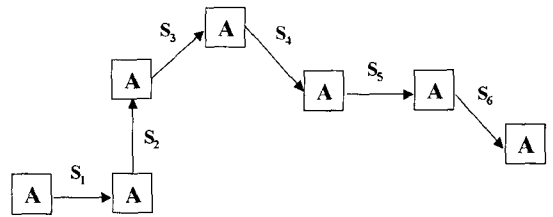
$$P : \{(x_1, y_1), (x_2, y_2), K_i\} \quad (\text{식 3})$$

FL은 각 프레임 내에 존재하는 객체들간의 시공간적 관계들을 표현하는 계층으로 VST 관계들(Visual-Spatio-Temporal relationships)[15]를 명시한다. 기본적으로 시간 관계와 공간 관계를 표현하기 위해서 Allen의 시간 간격 관계[16]과 Egenhofer의 공간 관계[17]를 사용한다. 또한 SL은 F(Feature trajectory), R(Relationship trajectory), θ (Content trajectory)를 명시함으로써 샷 내에서 객체의 시각적인 특징 변화와 객체들간의 관계 변화에 대한 궤적을 표현한다.

John. Z. Li는 Zaher Aghbari가 제안한 모델링과 달리 비디오의 구조적 정보와 관계없이 움직임 객체를 중심으로 표현하는 모델링을 제안하였다[12]. 움직임 객체를 비디오의 프레임 내에서 시간에 따라 위치가 변하는 물리적 객체로 "salient object"라 정의하고, 단일 움직임 객체의 궤적을 나타내었다. 방향은 그림 1의 (a)와 같이 8가지 방향 관계에 의해 나타내었다. 궤적은 시간 간격 $\langle I_1, I_2, \dots, I_m \rangle$ 이 주어졌을 때, (식 4)와 같이 시간에 따른 움직임 리스



(a) 방향 관계



$\langle (S_1, ET, l_1), (S_2, NT, l_2), (S_3, NE, l_3), (S_4, SE, l_4), (S_5, ET, l_5), (S_6, SE, l_6) \rangle$

그림 1. 움직임 객체 리스트의 예

트로 표현한다. 이 때, S_i 는 객체의 변위(거리), d_i 는 방향 관계, I_i 은 시간 간격을 의미한다. 예를 들어, 그림 1의 (b)와 같이 나타낼 수 있다.

$$\langle (S_1, d_1, I_1), (S_2, d_2, I_2), \dots, (S_m, d_m, I_m) \rangle \quad (\text{식 4})$$

John. Z. Li에 의해 제안된 모델링은 움직임 객체의 방향을 고려한 궤적의 유사도 검색을 지원한다. 두 객체 A와 B의 궤적을 $A = \{ T_1, T_2, \dots, T_m \}$ ($m \geq 1$), $B = \{ U_1, U_2, \dots, U_n \}$ ($m \leq n$)라 할 때, 객체 A, B의 궤적 유사도 $TrajSim(A, B)$ 는 (식 5)과 같이 나타낸다. 두 움직임 객체 A와 B의 궤적 유사도를 비교하기 위해 (식 6)와 같이 객체 A, B의 궤적 차를 구하고, 이를 이용하여 궤적 유사도를 계산한다. 유사도의 범위는 0에서 1로 나타나며 1에 가까울수록 유사하다.

$$TrajSim(A, B) = \frac{\max DiffTraj(A, B) - \min DiffTraj(A, B)}{\max DiffTraj(A, B)} \quad (\text{식 5})$$

$$\min DiffTraj(A, B) = \min_{i=1}^m \sum_{j=1}^m distance(T_i, U_{i+j}) \quad (\forall j, 0 \leq j \leq n - i) \quad (\text{식 6})$$

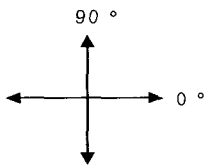
Choon-Bo Shim는 John. Z. Li와 같이 비디오의 구조적 정보와 관계없이 표현하고 단일 움직임 객체의 궤적을 움직임 거리와 움직임 방향으로 나타내는 모델링을 제안하였다[13]. 그러나 거리는 실제적 거리

를 1부터 100까지 범위에서의 상대적 거리로 나타내고, 움직임의 방향은 그림 2의 (a)와 같이 각을 사용하여 표현하였다. 궤적은 시간 간격 $\langle I_1, I_2, \dots, I_m \rangle$ 이 주어져 있을 때, (식 7)와 같이 시간에 따른 움직임 리스트로 나타낸다. 이 때, α_i 은 방향을 나타내는 각 ($1^\circ \sim 360^\circ$)이고, D_i 는 시간 간격 I 의 움직임 거리를 1에서 100까지의 상대적 거리로 나타낸 값이다. I_i 는 일정한 시간 간격을 나타낸다. 예를 들어, 그림 2의 (b)와 같이 나타낼 수 있다.

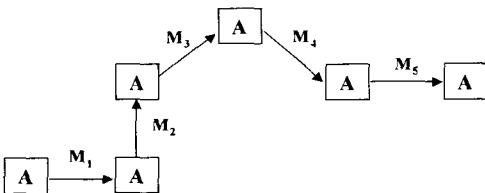
$$\langle (\alpha_1, D_1, I_1), (\alpha_2, D_2, I_2), \dots, (\alpha_m, D_m, I_m) \rangle \quad (\text{식 7})$$

Choon-Bo Shim에 의해 제안된 모델링은 움직임 객체의 방향과 거리를 고려한 새로운 궤적의 유사도 검색을 지원한다. 비디오 데이터베이스에서 단일 움직임 객체 V 의 궤적을 $V = \{ V_1, V_2, \dots, V_M \}$, 사용자 질의 Q 의 궤적을 $Q = \{ Q_1, Q_2, \dots, Q_N \}$ ($M \geq N$)라 할 때, 움직임 객체의 궤적 유사도 $SDST(V, Q)$ 는 (식 8)과 같이 방향(SR)과 이동한 거리(SD)의 차를 고려하여 계산한다. $SR_i(V_i, Q_i)$ 은 (식 9)과 같이 비디오와 사용자 질의간의 방향 차를 계산한다. $SD_i(V_i, Q_i)$ 는 (식 10)과 같이 비디오와 사용자 질의간의 이동한 거리 차를 계산한다.

$$SDST(V, Q) = \text{MAX} \sum_{j=1}^{M-N+1} \left(\frac{\sum_{i=j}^N SR_{i+j}^{(1-\omega)} * SD_{i+j}^{(1-\omega)}}{N} \right) \quad (\text{식 8})$$



(a) 방향 관계



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

(b) 움직임 궤적 리스트

그림 2. 움직임 객체 리스트의 예

$$SR_i(V_i, Q_i) = \frac{\cos(D_{ang}(V_i, Q_i)) + 1}{2} \quad (\text{식 9})$$

$$SD_i(V_i, Q_i) = 1 - \frac{|D_{R(i)}^V - D_{R(i)}^Q|}{\text{Max}(D_{R(i)}^V - D_{R(i)}^Q)} \quad (\text{식 10})$$

2.2 기존 연구의 문제점

Zaher AGHBARI는 비디오의 움직임 객체에 대한 시간의 변화에 따른 시각적 특징 정보의 궤적과 객체들간의 관계 궤적을 조합하여 나타내는 내용 궤적 모델링을 제안하였다. 이러한 정보를 나타내기 위해서 샷과 같은 논리적으로 나뉘어진 구조 계층 내에 움직임 객체의 변화를 표현하고 있다. 따라서 연속적인 시간의 흐름을 갖는 움직임 객체를 표현할 경우, 구조적 정보에 의해 객체가 종속되어 표현되고 객체의 정확한 궤적 표현과 검색이 어렵다. John. Z. Li와 Choon-Bo Shim는 움직임 객체에 대한 공간적 위치만을 표현하고 객체의 공간적 크기는 고려하지 않기 때문에 객체의 공간적 크기 변화와 객체와 객체간의 공간적 거리 변화를 표현하지 못한다. 또한 움직임 객체의 거리와 방향으로 궤적을 표현하기 때문에 시간의 변화에 따른 시각적 특징의 정보를 알 수가 없다.

John. Z. Li는 그림 1에서와 같이 8가지의 방향 관계만을 사용하고 있어서 8가지 이외의 방향은 정확하게 명시할 수 없다. 또한 궤적의 표현에서 객체의 움직임 방향만을 고려하고 있다. 이런 문제점들을 해결하고 제한점을 보완하기 위하여 제안하는 움직임 객체 모델링에서는 구조적인 정보와 관계없이 객체를 중심으로 모델링한다. 제안하는 모델링은 공간적인 위치뿐만 아니라 공간적인 크기 변화도 함께 표현한다. 이에 따른 검색은 의미 기반 검색, 시각적 특징 기반의 유사도 검색, 거리 유사도 검색, 궤적의 유사도 검색뿐만 아니라 이들을 혼합한 가중치 혼합 유사도 검색 등의 다양한 검색을 지원한다.

3. 제안하는 움직임 객체 모델링

3.1 설계시 고려 사항

비디오 데이터에 나타나는 객체의 움직임을 표현하기 위해서는 객체의 공간적인 위치의 변화뿐만 아니라 시각적인 특징의 변화를 효과적으로 표현할 수 있어야 한다. 또한 움직임 객체의 시간적인 존속 구간을 명확히 표현할 수 있어야 한다. 제안하는 움직

입 객체 모델링은 다음 과 같은 특징을 고려하여 객체를 표현하도록 한다.

첫째, 움직임 객체는 비디오의 구조적인 정보와 독립적으로 표현한다. 비디오 데이터를 검색하는 가장 기본적인 방법은 비디오의 구조 정보에 기반한 검색이다. 그러나 객체는 비디오의 구조 정보와 독립적으로 비디오 데이터에 나타난다. 따라서 연속된 시간 동안 나타나는 객체를 구조적인 정보에 기반하여 표현한다면 객체가 가지는 시간 정보는 구조적 특징에 의해 상실되게 된다. 따라서 제안하는 모델링에서의 객체는 구조 정보와 독립적으로 표현한다.

둘째, 움직임 객체의 시간적인 존속 구간을 명확히 표현한다. 비디오의 구조 정보와 독립적으로 표현한 객체가 불연속적으로 나타날 경우, 하나의 객체에 서로 다른 식별자를 부여한다. 만약 불연속적으로 나타나는 객체에 대해 동일한 식별자를 부여한다면 객체가 실제로 나타나지 않는 시간에도 객체가 존재하는 것으로 인식하기 때문에 실제적인 객체의 존속 구간을 명확히 판별할 수 없다. 또한 다른 객체와의 시간 관계 또한 불명확하다. 따라서 불연속적인 시간에 나타나는 객체에 대해 서로 다른 식별자를 부여하여 시간적인 존속 구간을 명확히 판별할 수 있도록 한다.

셋째, 시간의 변화에 따른 객체의 움직임에 대해 공간적인 위치뿐만 아니라 크기 변화도 함께 표현한다. 시공간 데이터베이스의 대표적인 응용인 GPS 시스템에서는 객체의 크기 변화보다는 위치 변화에 초점이 맞추어 있다. 그러나 비디오 데이터에서는 객체의 위치 변화뿐만 아니라 객체의 크기 변화도 매우 중요하다. 따라서 객체의 공간적 위치뿐만 아니라 MBR 정보도 함께 명시하여 객체의 크기 변화를 나타낼 수 있도록 한다.

넷째, 움직임 객체의 시간의 변화에 따라 수행하는 연속된 동작과 관련된 사건에 대한 의미 정보를 표현한다. 비디오에 나타난 객체는 시간의 변화에 따라 다양한 동작들을 수행한다. 또한 여러 객체들이 함께 특정 사건에 관계되어 나타나기 때문에 이에 대한 명시가 필요하다. 따라서 움직임 객체가 시간의 변화에 따라 수행하는 동작과 사건을 나타낼 수 있도록 한다.

다섯째, 움직임 객체의 궤적과 시각적인 특징의 변화를 표현한다. 시간의 변화에 따른 움직임 객체의 공간적인 위치를 표현하기 위해서는 특정 시간 간격 동안의 움직임 방향과 거리 그리고 속도의 변화를

나타내는 궤적 정보의 표현이 필요하다. 또한 시간의 변화에 따른 시각적인 특징 정보의 변화를 나타낼 수 있도록 한다.

3.2 제안하는 모델링

비디오는 수많은 프레임들로 구성된다. 일반적으로 비디오는 1초에 20~30정도의 프레임들로 구성된다. 또한 프레임 내에는 수많은 객체들이 포함되어 있다. 따라서 사용자의 질의에 대해 비디오의 모든 프레임을 검색하는 일은 어렵고, 짧은 시간 동안 프레임들에서 객체들의 변화는 미세하기 때문에 모든 프레임에 나타난 객체의 움직임을 표현하는 데는 문제점이 있다. 따라서 비디오를 구성하는 프레임들 중에서 일정한 시간 간격으로 키프레임(Keyframe)을 추출하고 비디오의 시간에 따른 객체들의 변화를 객체 단위로 나타낸다. 제안하는 모델링의 구조는 그림 3과 같이 원시 데이터 계층(Raw Data Level), 프레임 계층(Frame Level) 그리고 객체 계층(Object Level)의 세 개의 계층으로 구성한다. 비디오의 구조 정보와 독립적으로 움직임 객체를 중심으로 표현하기 위해서 비디오의 구조 레벨을 두지 않았다. 원시 데이터 계층은 비디오 자체가 가지고 있는 비디오 전체의 가공되지 않는 정보를 표현한다. 프레임 계층은 일정 시간 간격 Δt 단위로 비디오에서 추출한 키프레임에 대한 정보를 표현한다. 객체 계층은 실제적인 움직임 객체에 대한 정보를 표현한다.

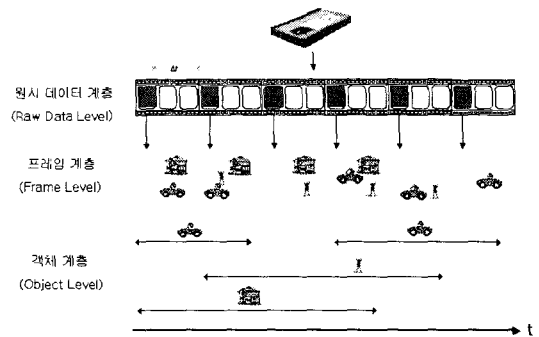


그림 3. 제안하는 모델링 구조

3.2.1 원시 데이터 계층

원시 데이터 계층은 제안하는 모델링의 최하위 계층으로 비디오에 대한 가공되지 않은 비디오 자체에 대한 정보를 표현하는 계층이다. 원시 데이터 계층은

(식 11)과 같이 비디오 식별자 $Video_{id}$, 비디오 제작에 관련된 특징 $Production_F$ 그리고 물리적 특징 $Physical_F$ 을 표현한다. 비디오 제작에 관련된 특징인 $Production_F$ 는 제작자, 비디오의 이름, 감독, 제목 등과 같은 정보들을 나타낸다. 이에 반해 비디오의 물리적 특징인 $Physical_F$ 는 저장 방식, 압축 방식, 상영 시간, 초당 재생 프레임 수와 같은 물리적인 특징을 나타낸다.

$$Raw\ data : \langle Video_{id}, Production_F, Physical_F \rangle \quad (식\ 11)$$

3.2.2 프레임 계층

프레임 계층은 비디오에서 특정 시간 간격 Δt 단위로 추출한 키프레임에 대한 정보를 표현하는 계층이다. 프레임 계층은 해당 키프레임의 시각적 특징과 키프레임 내에 나타난 객체의 시각적 또는 공간적인 특징 정보를 표현한다. 프레임 계층의 표현은 (식 12)과 같이 키프레임 식별자 $KeyFrame_{id}$, 프레임 내에 나타난 객체에 대한 정보 리스트 $Object_L$ 그리고 해당 키프레임의 시각적 특징 정보 $Visual_{Feature}$ 을 표현한다.

$$Frame : \langle KeyFrame_{id}, Object_L, Visual_{Feature} \rangle \quad (식\ 12)$$

$Object_L$ 는 객체 자체에 대한 정보뿐만 아니라 키프레임에서 객체가 가지는 공간적인 특징과 시각적인 특징에 대한 정보를 표현한다. 해당 키프레임에 m 개의 객체가 존재한다고 할 때, $Object_L$ 는 (식 13)과 같이 객체의 리스트 OL_i 들로 구성된다.

$$Object_L : \langle OL_1, OL_2, \dots, OL_m \rangle \quad (식\ 13)$$

OL_i 는 (식 14)과 같이 키프레임에 존재하는 객체의 식별자 O_i , O_i 의 공간적 위치 MBR_i 그리고 중심점 C_i 로 표현한다. 또한 O_i 의 모양 정보 S_i 와 시각적인 특징 정보 OF_i 로 표현한다. 객체의 공간적인 위치를 나타내는 MBR_i 은 (X_i, Y_i) 로써 $X_i = (X_{si}, X_{fi})$ 와 $Y_i = (Y_{si}, Y_{fi})$ 로 표현된다. 이 때, X_{si} 와 X_{fi} 는 객체 O_i 를 X축에 정사영시킨 시작 위치와 끝 위치를 나타낸다. 이와 동일하게 Y_{si} 와 Y_{fi} 도 Y축에 정사영시킨 시작 위치와 끝 위치를 나타낸다. C_i 는 중심점으로 $(\frac{X_{fi} - X_{si}}{2}, \frac{Y_{fi} - Y_{si}}{2})$ 이고, OF_i 은 O_i 의 RGB값인 (R_i, G_i, B_i) 로 표현한다.

$$OL_i : \langle O_i, MBR_i, C_i, S_i, OF_i \rangle \quad (1 \leq i \leq m) \quad (식\ 14)$$

$Visual_{Feature}$ 는 프레임 자체의 시각적인 특징으로 (식 15)와 같이 표현한다. 프레임의 칼라 C 와 프레임의 질감 T 그리고 프레임의 윤곽 S 에 대한 특징 벡터 값들로 나타낸다.

$$Visual_{Feature} : \langle C, T, S \rangle \quad (식\ 15)$$

3.2.3 객체 계층

객체 계층은 비디오에 나타난 객체에 대한 전체적인 정보를 나타내는 계층으로 프레임 계층에 표현된 객체의 특징 정보를 이용하여 객체 단위로 구성한다. 제안하는 모델링에서 객체 계층은 객체의 궤적 정보뿐만 아니라 시간의 변화에 따른 객체의 시각적 특징에 대한 변화를 표현한다. 또한 객체가 수행한 동작과 연관된 사건에 대한 정보를 표현한다. 객체 계층의 표현은 (식 16)과 같다. $Object_{id}$ 는 객체 식별자, $Name$ 은 객체의 이름, $flag$ 는 움직임 객체 또는 정지 객체인지를 구별하는 플래그이다. $Color_L$ 는 시간의 변화에 따른 객체의 시각적 변화를 나타내는 시각적 특징 리스트이다. MBR_L 는 시간 변화에 따른 객체의 공간적인 위치와 크기 변화를 나타내는 MBR 리스트이다. $Camera_L$ 는 객체가 나타나는 시간 동안에 객체에 사용된 카메라 동작 리스트이다. $Action_L$ 는 객체가 나타난 시간 동안에 수행한 동작 리스트, $Event_L$ 는 객체와 연관된 사건 리스트 그리고 $Trajectory_L$ 는 시간의 변화에 따른 객체의 공간적인 정보를 나타내는 궤적 리스트이다. I 는 객체가 나타난 존속 구간으로 $[t_s, t_f]$ 로 표현된다. 이 때, t_s 는 등장 시간이고 t_f 는 소멸 시간이다.

$$O : \langle Object_{id}, Name, flag, Color_L, MBR_L, Camera_L, Action_L, Event_L, Trajectory_L, I \rangle \quad (식\ 16)$$

$Object_{id}$ 는 객체의 고유한 식별자로 객체의 시간적인 존속 구간을 명확히 판별하기 위해서 객체가 불연속적인 시간 동안 나타나는 경우, 서로 다른 식별자를 부여한다. 예를 들어, 그림 4와 같이 비디오 내에 객체들 O_1, O_2, O_3 이 나타난다고 할 때, 객체 O_2 는 존속 구간 $[t_3, t_7]$ 의 연속된 시간 동안에 한 번 나타난다. 이에 반해 객체 O_1 와 O_3 는 이들이 가지는 존속 구간 $[t_1, t_5]$ 와 $[t_2, t_8]$ 에서 불연속적으로 두 번씩 나타나고 있다. 만약 이렇게 불연속적으로 나타난 객체를 동일한 객체로 인식하여 하나의 식별

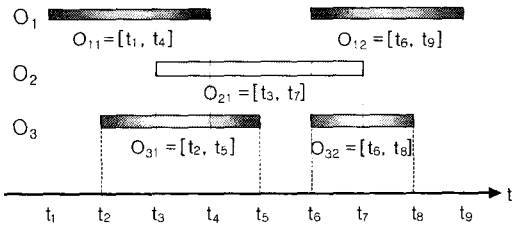


그림 4. 객체 식별자 부여

자로 부여한다면 실제로 객체가 존재하지 않은 시간에도 객체가 존재하는 것으로 표현된다. 따라서 객체 O_1 과 같이 t_1 에서 t_9 동안 불연속적으로 나타난 하나의 객체에 대해서 O_{11} , O_{12} 와 같이 서로 다른 식별자를 부여하여 객체의 존속 구간을 명확히 한다. O_3 도 이와 동일하다.

$Color_L$ 는 객체의 시각적 특징을 나타내는 칼라 리스트로 객체가 m 번의 변화를 가진다고 할 때, (식 17)와 같이 표현한다. Col_i 는 시간 간격 I_i 동안에 객체의 색상 정보를 나타내는 RGB 값 (R_i, G_i, B_i)로 표현하고, I_i 는 시간 간격으로 $[t_{si}, t_{fi}]$ 을 나타낸다.

$$Color_L : \langle (Col_1, I_1), (Col_2, I_2), \dots, (Col_m, I_m) \rangle \quad (식 17)$$

MBR_L 는 객체의 공간적인 위치를 나타내는 MBR 리스트로 객체가 m 번의 변화를 가진다고 할 때, (식 18)와 같이 표현한다. I_i 동안에 객체의 MBR 정보 MBR_i , 객체의 중심점 C_i 그리고 객체의 MBR의 모양 정보 S_i 로 표현한다.

$$MBR_L : \langle (MBR_1, C_1, S_1, I_1), (MBR_2, C_2, S_2, I_2), \dots, (MBR_m, C_m, S_m, I_m) \rangle \quad (식 18)$$

$Camera_L$ 는 객체에 사용된 카메라 동작 리스트로 객체에 m 번의 변화가 있을 때, (식 19)와 같이 표현한다. 이 때, $Camera_i$ 는 I_i 동안의 객체에 사용된 카메라 동작이다. 카메라 동작을 나타내는 $Camera_i$ 는 촬영의 기본인 고정(Fix), 팬닝(Panning), 틸팅(Tilting), 주밍(Zooming), 달리(Dolly), 트랙(Track), 엘리베이터(elevator)와 같은 7가지로 정의한다.

$$Camera_L : \langle (Camera_1, I_1), (Camera_2, I_2), \dots, (Camera_m, I_m) \rangle \quad (식 19)$$

객체는 연속된 시간의 변화에 따라 다양한 동작을 수행하면서 여러 사건에 연관되어 비디오에 나타난다. $Action_L$ 는 객체의 동작 리스트로 객체가 m 번의

변화를 가진다고 할 때, (식 20)와 같이 표현한다. 이 때, $Action$ 는 I_i 동안에 객체가 수행한 동작이다. $Event_L$ 는 객체가 연관된 사건의 리스트로 객체가 m 개의 사건에 연관된다고 할 때, (식 21)와 같이 표현한다. 이 때, $Event_i$ 는 I_i 동안에 객체가 연관된 사건을 나타낸다.

$$Action_L : \langle (Action_1, I_1), (Action_2, I_2), \dots, (Action_m, I_m) \rangle \quad (식 20)$$

$$Event_L : \langle (Event_1, I_1), (Event_2, I_2), \dots, (Event_m, I_m) \rangle \quad (식 21)$$

3.3 움직임 객체의 궤적

움직임 객체는 시간에 따라 공간상에서 위치나 모양이 변하는 객체이다. 움직임 객체의 시공간적 특성을 표현하기 위해서 시간에 따른 객체의 움직임에 대한 정보를 나타내는 궤적 정보의 표현이 필요하다. 움직임 객체가 시간 간격 $\langle I_1, I_2, \dots, I_m \rangle$ 이 주어지고 m 번 움직인다고 할 때, 객체의 궤적 정보 $Trajectory_L$ 는 (식 22)과 같이 표현한다.

$$Trajectory_L : \langle (R_1, D_1, V_1, I_1), (R_2, D_2, V_2, I_2), \dots, (R_m, D_m, V_m, I_m) \rangle \quad (식 22)$$

R_i, D_i, V_i 는 I_i 동안에 객체가 움직인 방향, 거리, 속도이다. I_i 는 시간 간격으로 $[t_{si}, t_{fi}]$ 을 나타낸다. R_i 는 객체가 움직인 방향으로 그림 5와 같이 객체의 움직임 변화를 각도로 표현한다. t_{si} 시점의 객체 중심점 $C_{si} = (C_{si}^x, C_{si}^y)$ 을 기준으로 t_{fi} 시점의 객체의 중심점인 $C_{fi} = (C_{fi}^x, C_{fi}^y)$ 사이의 상대적인 각인 $-180^\circ \sim 180^\circ$ 의 범위로 표현한다.

D_i 는 객체가 움직인 거리로 (식 23)와 같이 객체들의 중심간의 거리로 계산한다. V_i 는 객체가 움직인 속도로 (식 24)와 같이 계산한다.

$$D_i = \sqrt{(C_{fi}^x - C_{si}^x)^2 + (C_{fi}^y - C_{si}^y)^2} \quad (식 23)$$

$$V_i = \frac{D_i}{t_{fi} - t_{si}} \quad (식 24)$$

예를 들어 객체가 시간 간격 $\langle I_1, I_2, I_3, I_4 \rangle$ 동안 움직임 $\langle M_1, M_2, M_3, M_4 \rangle$ 을 가질 때, 궤적 $Trajectory_L$ 는 그림 6과 같이 표현할 수 있다. M_1 은 [1,4]의 시간 간격 동안에 (4, 7)에서 (6, 5)로 움직였다. 이 때, R_M, D_M, V_M 는 아래와 같이 계산하며 나머지 움직임

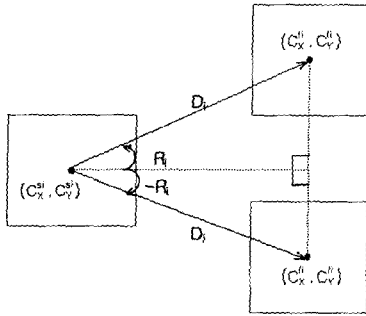
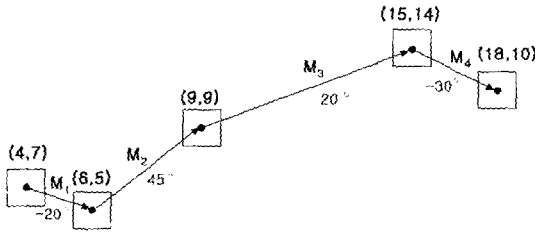


그림 5. 궤적의 방향과 거리 표현



$\langle (-20^\circ, 2.8, 0.94, [1, 4]), (45^\circ, 5, 1.67, [5, 8]), (20^\circ, 7.81, 1.12, [9, 16]), (-30^\circ, 5, 1, [17, 22]) \rangle$

그림 6. 움직임을 객체 궤적의 예

임 또한 동일한 방법으로 계산한다.

$$R_{M1} = -20^\circ$$

$$D_{M1} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \\ = \sqrt{(6-4)^2 + (5-7)^2} = 2.8$$

$$V_{M1} = \frac{D_{M1}}{I_{M1}} = \frac{2.8}{4-1} = 0.94$$

4. 움직임을 객체 검색

제안하는 움직임을 객체 모델링은 의미 기반 검색, 프레임 또는 객체의 시각적 특징 기반의 유사도 검색, 객체들간의 거리에 대한 유사도 검색, 움직임을 객체의 움직임을 방향과 거리 그리고 속도를 고려한 궤적 유사도 검색을 지원한다. 또한 객체들간의 시각적 특징, 거리, 궤적에 가중치를 부여한 가중치 혼합 검색을 지원한다. 이 장에서는 제안하는 모델링에 따라 지원하는 검색과 검색에 필요한 연산자를 정의한다.

4.1 의미 기반 검색

의미 기반 검색은 비디오 데이터에서 비디오에 포

함된 객체, 객체가 수행하는 동작 그리고 사건과 같은 의미 정보를 검색한다. 객체 계층에서는 키프레임 내에 포함된 객체들에 대한 객체의 동작, 사건, 존속 구간을 표현하였다. 이를 통해 의미 기반 검색을 수행한다. 질의 형태는 “주어진 동작을 수행하는 객체를 검색하라”, “주어진 사건과 연관된 객체를 검색하라”로 할 수 있다. 이러한 질의를 수행하기 위해서 객체 계층에 존재하는 객체의 동작 변화 리스트나 사건 리스트를 검색하여 주어진 질의를 만족하는 결과값을 반환한다.

4.2 시각적 특징 기반의 유사도 검색

시각적 특징 기반의 유사도 검색은 프레임이나 객체에서 추출되는 칼라, 모션, 질감, 모양 등의 시각적 특징과 유사한 프레임이나 객체를 검색한다. 프레임 내에 존재하는 객체의 시각적 특징 유사도 검색과 프레임 자체의 시각적 특징 유사도 검색이 있다.

4.2.1 객체의 시각적 특징 유사도

객체 계층은 키프레임 내에 포함된 객체들에 대한 시간의 변화에 따른 시각적 특징을 표현하였다. 이를 통해 객체의 시각적 특징 유사도 검색을 수행한다. 질의 형태는 “객체 O와 시각적 특징이 유사한 객체를 검색하라”로 할 수 있다. 이러한 질의를 수행하기 위해서 객체 계층에 존재하는 객체의 시각적 특징 변화 리스트를 검색하여 주어진 질의에 만족하는 결과값을 반환한다. [정의 1]은 시각적 특징 유사도를 정의한다.

[정의 1] 시각적 특징 유사도

두 객체 O_a, O_b 의 존속 구간이 $[t_s^a, t_f^a], [t_s^b, t_f^b]$ 일 때, 두 객체가 나타난 존속 구간에서의 시각적 특징 유사도 $SC(O_a, O_b)$ 는 (식 25)과 같이 정의한다.

$$SC(O_a, O_b) = \frac{\sum_{t=t_s}^{t_f} D_{color}(O_a, O_b, t)}{t_f - t_s} \\ t_s = \text{MIN}(t_s^a, t_s^b), t_f = \text{MAX}(t_f^a, t_f^b) \quad (\text{식 25})$$

이 때, $D_{color}(O_a, O_b, t)$ 는 특정 시점 t 에서 O_a 와 O_b 의 RGB 값의 차이이다. 만약 한 객체가 존속하는 구간에 다른 객체가 존속하지 않아 두 객체의 색상 유사도를 계산할 수 없거나 색상이 전혀 유사하지 않을 때에는 $MaxColor$ 가 된다. $SC(O_a, O_b)$ 는 O_a 와

O_b 의 존속 구간에서 시각적 특징 유사도로 두 객체의 움직임 전체에 대한 시각적 특징을 비교한다. 두 객체가 가지는 존속 구간 중에 최대인 객체의 소멸 시간 t_f 와 최소인 객체의 등장 시간 t_s 의 시간 간격으로 나누어 계산한다. $SC(O_a, O_b)$ 는 0에서 1의 범위를 가지며, 시각적 특징이 유사할수록 1에 가깝다.

4.2.2 프레임의 시각적 특징 유사도

프레임 계층은 키프레임 자체의 시각적 특징을 표현하였다. 이를 통해 프레임의 시각적 특징 유사도 검색을 수행한다. 질의 형태는 “프레임 F 와 유사한 프레임을 검색하라”, “객체 O 를 포함하는 프레임 중에서 프레임 F 와 유사한 프레임을 검색하라”로 할 수 있다. 이러한 질의를 수행하기 위해서 프레임 계층에 명시한 프레임의 시각적 특징을 검색하여 주어진 질의에 만족하는 결과값을 반환한다. [정의 2]는 프레임의 칼라, 질감, 모양의 시각적 특징에 대한 유사도를 정의한다.

[정의 2] 프레임의 시각적 특징 유사도

두 프레임 K_a 와 K_b 의 시각적 특징이 $\langle C_a, T_a, S_a \rangle$, $\langle C_b, T_b, S_b \rangle$ 일 때, 프레임의 시각적 특징 유사도 $SF(K_a, K_b)$ 는 (식 26)과 같이 정의된다.

$$SF(K_a, K_b) = 1 - \frac{D_{frame}(K_a, K_b)}{MaxF} \quad (\text{식 26})$$

이 때, $D_{frame}(K_a, K_b)$ 는 K_a 와 K_b 의 시각적 특징차이고 $MaxF$ 는 프레임의 시각적 특징의 최대 차이이다. $SF(K_a, K_b)$ 는 0에서 1의 범위를 가지며, 프레임의 시각적 특징이 유사할수록 1에 가깝다.

4.3 거리 유사도 검색

거리 유사도 검색은 특정 객체와 근접하게 움직이는 객체를 검색한다. 객체 계층에서 MBR 리스트를 표현하였다. 이를 통해 객체간의 거리 유사도 검색을 수행한다. 질의 형태는 “주어진 객체 O 와 가장 근접하게 존재하는 객체를 검색하라”로 할 수 있다. 이런 질의를 수행하기 위해서는 먼저 객체간의 떨어진 정도를 알기 위해 MBR 거리 계산을 해야 한다. [정의 3]은 두 객체의 거리 유사도를 정의한다. 그림 7은 객체의 MBR 간의 거리를 나타낸다. 객체 O_a 에 대한 객체들 $O_{b1}, O_{b2}, O_{b3}, O_{b4}$ 의 특정 시점 t 에서 MBR

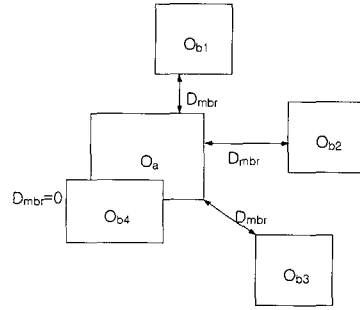


그림 7. 두 객체의 MBR간 거리

간의 거리는 두 객체의 MBR 의 영역 경계면 사이의 최소 거리로 정의한다. 만약 O_a 와 O_{b4} 와 같이 두 객체가 겹침이 발생한다면 $D_{mbr}(O_a, O_{b4}, t)$ 의 거리는 0이 된다.

[정의 3] 거리 유사도

두 객체 O_a, O_b 의 존속 구간이 $[t_s^a, t_f^a], [t_s^b, t_f^b]$ 라 할 때, 객체간의 거리 유사도 $SD(O_a, O_b)$ 는 (식 27)과 같이 정의한다.

$$SD(O_a, O_b) = \frac{\sum_{t=t_s}^{t_f} D_{mbr}(O_a, O_b, t)}{t_f - t_s},$$

$$t_s = \text{MIN}(t_s^a, t_s^b), t_f = \text{MAX}(t_f^a, t_f^b) \quad (\text{식 27})$$

이 때, $D_{mbr}(O_a, O_b, t)$ 은 특정 시점 t 에서 MBR 간의 거리이다. 만약 한 객체가 존속하는 구간에 다른 객체가 존속하지 않아 두 객체의 MBR 거리를 계산할 수 없을 때는 최대 거리값 $MaxDist$ 을 가진다. $SD(O_a, O_b)$ 는 두 객체의 움직임 전체에 대해 MBR 간의 거리를 비교하기 위해 두 객체가 가지는 존속 구간 중에 최대인 객체의 소멸 시간과 최소인 객체의 등장 시간의 시간 간격으로 나누어 계산한다. $SD(O_a, O_b)$ 는 0에서 1의 범위를 가지며 거리가 유사할수록 1에 가깝다.

그림 8은 두 객체 O_a, O_b 의 존속 구간이 $[t_s^a, t_f^a]$ 와 $[t_s^b, t_f^b]$ 일 때, 두 객체간의 거리를 나타낸다. 블록 부분은 두 객체가 동시에 존속하는 구간으로 $[t_s^a, t_f^a]$ 에 해당하며, 실제적 거리를 MBR 간의 거리로 구할 수 있다. 그러나 원으로 표시된 부분은 두 객체의 존속 구간이 서로 겹치지 않는 부분으로 $[t_s^b, t_f^a]$ 와 $[t_f^a, t_f^b]$ 에 해당하며 실제적인 거리를 알 수 없기 때문에 $MaxDist$ 값을 가진다. 거리 유사도 $SD(O_a, O_b)$ 는 두

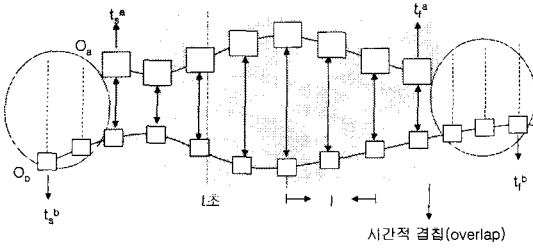


그림 8. 두 객체간의 거리 표현

객체를 포함하는 시간 영역인 $[t_s^a, t_f^a]$ 에서 t 의 변화에 따른 MBR 간의 거리를 구해 합한 값을 $[t_s^a, t_f^a]$ 의 시간 간격으로 나눈 평균값이 된다.

4.4 객체의 궤적 유사도 검색

객체의 궤적 유사도 검색은 특정 객체와 유사한 궤적을 가지는 객체를 검색한다. 객체 계층은 객체의 전체적인 움직임을 파악할 수 있는 궤적 정보를 표현하였다. 이를 통해 객체의 궤적 유사도 검색을 수행한다. 질의 형태는 “객체 O 와 가장 유사한 궤적으로 움직이는 객체를 검색하라”로 할 수 있다. 이러한 질의를 수행하기 위해서는 객체의 움직임 방향, 거리 그리고 속도를 고려한 객체간의 궤적 차를 구해야 한다. [정의 4]는 두 객체의 궤적 차를 정의한다. 이를 이용하여 [정의 5]에서 두 객체의 궤적 유사도를 정의한다. 두 객체 O_a 와 O_b 에 대해 m 과 n 번의 움직임을 가진다고 할 때, 궤적 T_a , T_b 는 아래와 같이 표현할 수 있다. 단, $(1 \leq m \leq n)$ 이다.

$$T_a = \langle (R_a^1, D_a^1, V_a^1, I_a^1), (R_a^2, D_a^2, V_a^2, I_a^2), \dots, (R_a^m, D_a^m, V_a^m, I_a^m) \rangle$$

$$T_b = \langle (R_b^1, D_b^1, V_b^1, I_b^1), (R_b^2, D_b^2, V_b^2, I_b^2), \dots, (R_b^n, D_b^n, V_b^n, I_b^n) \rangle$$

[정의 4] 움직임 궤적 차

두 객체 O_a 와 O_b 의 움직임 궤적차 $RDV_{i,j}(O_a, O_b)$ 는 (식 28)과 같이 정의한다.

$$RDV_{i,j}(O_a, O_b) = w_{\alpha}RR(R_a^i, R_b^j) + w_{\beta}RD(D_a^i, D_b^j) + w_{\gamma}RV(V_a^i, V_b^j) \quad (\text{식 28})$$

이 때, $RR(R_a^i, R_b^j)$ 은 움직임 방향차, $RD(D_a^i, D_b^j)$ 은 움직임 거리차 그리고 $RV(V_a^i, V_b^j)$ 는 움직임 속도차이다. w_{α} 는 가중치로 $w_{\alpha} + w_{\beta} + w_{\gamma} = 1$ 이 된다.

다. $RDV_{i,j}(O_a, O_b)$ 는 0에서 1의 범위를 가진다.

[정의 5] 객체의 궤적 유사도

두 객체 O_a 와 O_b 의 궤적 유사도 $ST(O_a, O_b)$ 는 (식 30)의 $Tra(O_a, O_b)$ 을 이용하여 (식 29)와 같이 정의한다.

$$ST(O_a, O_b) = 1 - Tra(O_a, O_b) \quad (\text{식 29})$$

$$Tra(O_a, O_b) = \text{MIN} \sum_{k=1}^m \frac{RDV_{i,i+k}(O_a, O_b)}{m} \quad (\forall k, 0 \leq k \leq n - m + 1) \quad (\text{식 30})$$

$ST(O_a, O_b)$ 는 객체의 궤적 유사도로써 0에서 1의 범위를 가지며, 유사할수록 1에 가깝다.

4.5 가중치 혼합 유사도 검색

제안하는 모델링에서는 움직임 객체간의 시각적 특징, 거리, 궤적에 사용자가 가중치를 부여함으로써 가중치 혼합 유사성 검색이 가능하다. 이러한 질의를 수행하기 위한 가중치 혼합 유사도는 [정의 6]에서 정의한다.

[정의 6] 가중치 혼합 유사도

두 객체 O_a 와 O_b 에 대해서 가중치 혼합 유사도 $SS(O_a, O_b)$ 는 (식 31)과 같이 정의한다.

$$SS(O_a, O_b) = w_{s1}SC(O_a, O_b) + w_{s2}SD(O_a, O_b) + w_{s3}ST(O_a, O_b) \quad (\text{식 31})$$

$SS(O_a, O_b)$ 는 앞에서 정의한 객체의 시각적 특징 유사도 $SC(O_a, O_b)$, 거리의 유사도 $SD(O_a, O_b)$ 그리고 궤적의 유사도 $ST(O_a, O_b)$ 에 사용자가 가중치를 부여하도록 한다. 이 때, w_{s1} 는 가중치에 해당하며 $w_{s1} + w_{s2} + w_{s3} = 1$ 이다.

5. 실험 및 성능 평가

이 장에서는 비디오 데이터에서 움직임 객체 검색을 수행하기 위해 제안하는 모델링의 특징에 따라 구현한 움직임 객체 검색 시스템과 이 시스템에서의 성능 평가한 결과를 기술한다. 그리고 마지막으로 다른 모델링 기법들과 비교 분석한다.

구현 환경은 128MB 메모리를 갖는 펜티엄-III 700 MHz에서 Microsoft SQL Server 2000 데이터베이스

이를 사용하였다. 구현 언어로는 Microsoft Visual C++ 6.0을 사용하였다. 성능 평가를 위한 실험 데이터는 avi 동영상 파일 포맷의 농구 비디오 데이터를 사용하였다.

5.1 움직임 객체 검색 시스템

그림 9는 움직임 객체 검색을 위해 구현한 검색 시스템의 사용자 인터페이스 화면이다. 움직임 객체의 검색을 수행하는 인터페이스 화면은 크게 검색 조건을 입력하는 부분과 검색 결과를 나타내는 두 부분으로 이루어져 있다. 검색 조건을 입력하는 부분은 의미 기반 검색, 거리 유사도 검색, 시각적 특징 기반 유사도 검색, 꺾적 유사도 검색 그리고 이들에 가중치를 부여한 가중치 혼합 검색의 5개 영역으로 구성된다. 검색 결과를 나타내는 부분은 검색 조건에 해당하는 검색한 결과를 사용자에게 프레임 계층의 대표 이미지와 텍스트 정보를 나타낸다. 검색된 이미지와 텍스트 정보는 이전 버튼과 다음 버튼을 통해 순차적으로 볼 수 있다.

의미 기반 검색 영역은 객체나 객체가 수행하는 동작 또는 사건을 검색하기 위한 조건을 입력하는 부분으로 이루어져 있다. 각 의미적 내용에 대한 검색 조건은 시간 관계에 기반한 검색을 수행하기 위한 기준을 입력하는 부분과 텍스트 기반의 검색을 수행하기 위한 텍스트 입력 부분을 포함하고 있다. 또한 객체 검색에 대해서는 공간 관계와 객체의 색상 정보를 이용한 유사도 검색을 수행하기 위한 색상 정보를 입력하기 위한 부분이 추가적으로 구성되어 있다. 객체의 색상 정보를 이용한 유사도 검색에 대해서는 가중치와 유사도를 입력할 수 있다. 그림 10은 의미 기반 검색과 시각적 특징 유사도 검색의 실행 예이다. 전체 검색창은 그림 9와 동일하며 여기서는 입력되는 부분과 결과 부분만을 캡쳐하였다. 입력 조건은 '마틴'이라는 객체가 '덩크슛'하는 동작을 하고, 색상이 70정도 유사한 객체를 검색하는 것이다. 결과창에는 검색 조건을 만족하는 객체의 대표 이미지와 텍스트 정보가 출력한다. 시각적 특징 기반 유사도 검색 영역은 예제 이미지를 선택하기 위한 부분과 선택한 이미지에 대해 가중치와 유사도를 입력할 수 있다. 거리 유사도 검색 영역은 객체와 근접한 거리에 있는 객체를 검색한다. 기준 객체와 유사도를 입력할 수 있다. 거리 유사도 검색에서는 사용자가 질의 조

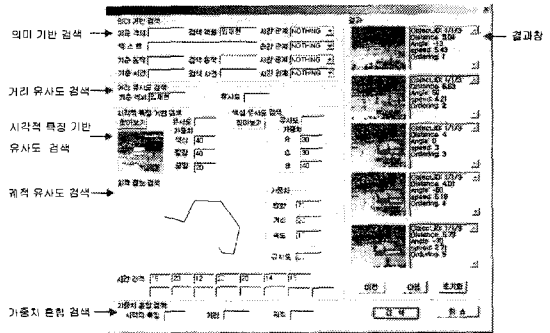
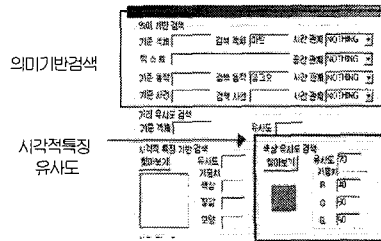
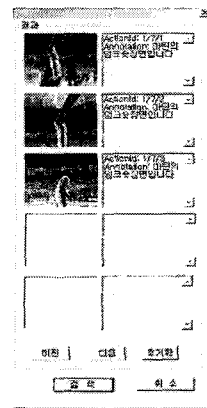


그림 9. 사용자 인터페이스



(a) 입력창



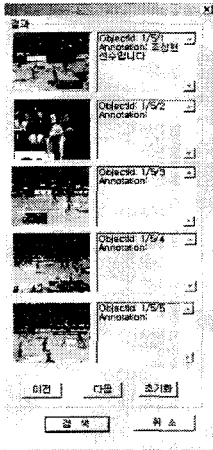
(b) 결과창

그림 10. 의미 기반 검색과 시각적 특징 유사도 검색 실행 예

건으로 객체를 입력하면 움직임 객체의 공간적인 위치 뿐만 아니라 공간적인 크기를 계산하여 사용자가 선택한 유사도를 만족하는 객체를 반환한다. 그림 11은 거리 유사도 검색의 실행 예이다. 입력 조건은 '김승현'과 유사도 80을 입력하면 질의 조건으로 입력한 '김승현'이라는 객체에 대한 MBR_L 을 검사하여 질의 조건 객체의 공간적인 위치와 크기를 계산하고 이를 통해 유사도 거리 내에 존재하는 객체를 반환한



(a) 입력창



(b) 결과창

그림 11. 거리 유사도 검색 실행 예

다. 질의 결과창에는 김승현과 유사도 거리에서 움직이고 있는 객체에 대한 대표 이미지와 텍스트 정보가 출력된다.

궤적 유사도 검색 영역은 사용자로부터 검색하고자 하는 궤적 모양을 스케치로 입력받아 수행하는 스케치 기반의 검색을 한다. 궤적 유사도 검색 영역은 사용자로부터 스케치를 입력받는 부분과 스케치된 각각의 움직임의 속도를 구하기 위한 시간 간격을 입력받는 부분이 포함되어 있다. 또한 제안하는 움직임 객체 모델링에서는 객체의 방향, 거리 그리고 속도를 고려하기 때문에 이들에 대한 가중치와 유사도

를 입력할 수 있다. 그림 12는 궤적 유사도 검색의 실행 예이다. 이 인터페이스는 궤적 계산 과정을 보여주기 위해 위의 인터페이스에 사용자가 스케치할 때마다 각 움직임에 대한 계산값들을 보여주는 영역을 포함시킨 것이다. 사용자에 의해 스케치될 때마다 좌표, 방향, 거리 그리고 속도가 계산된다. 즉, 계산된 값들이 다시 입력되어 사용자가 입력한 유사도에 가까운 궤적을 검색한다.

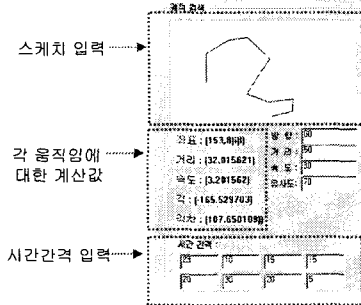
5.2 성능 평가

농구 비디오 데이터는 '경기장'이라는 한정된 공간을 배경으로 짧은 시간에 많은 움직임을 가지는 객체를 포함하고 있어서 움직임 객체의 궤적을 추출하는데 용이하다. 따라서 농구 비디오 데이터에서 표 1과 같은 특성을 갖는 실험 데이터를 사용한다. 본 논문은 객체를 인식하거나 궤적을 추출하는 것은 영상처리 기법에 의해 이미 수행되었다는 것을 전제로 하고 있다. 실험에서는 수작업을 통해 농구 비디오 데이터에서 객체의 내용 및 궤적을 표현하고 다음과 같은 조건에 의해 실험을 수행한다.

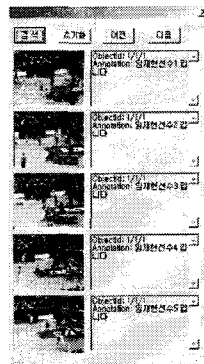
- 농구 비디오 데이터에서 선수 10명을 움직임 객체로 사용하였다.

표 1. 실험 데이터의 특성

포맷 형식	avi 동영상 파일
비디오 시간	54 분
파일 크기	544 MB
키프레임 포맷 형식	GIF 파일
거리 계산 단위	cm



(a) 입력창

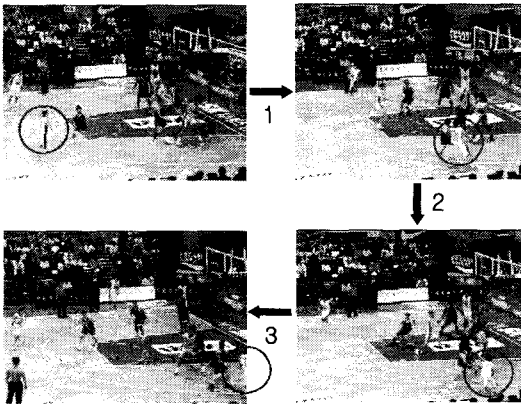


(b) 결과창

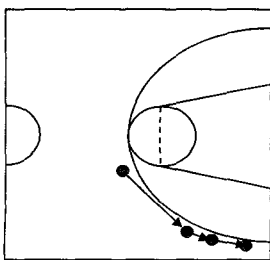
그림 12. 궤적 유사도 검색 실행 예

- 비디오 내의 각 객체는 5~15개의 움직임 정보들로 구성된다.
- 질의 형태는 스케치 기반 질의로 사용자로부터 입력받는다.
- 검색 효율의 측정을 위해 질의 데이터 표본 10개를 만들어 수행하였다.
- 질의 데이터 표본은 5~15개의 움직임 정보들로 구성된다.

비디오 데이터로부터 움직임 객체의 궤적 추출을 위해 비디오 데이터에서 1초 간격으로 키프레임을 추출하고 객체들의 움직임을 객체 단위로 표현하였다. 실제로 모든 객체는 3차원 공간에서 존재하지만 비디오는 2차원 평면으로 보여지기 때문에 객체는 2차원으로 정사영하여 표현하였다. 즉, 움직임 객체의 궤적을 표현할 때에는 위에서 아래로 내려다 봤을 때의 경기장을 축으로 하여 절대 좌표로 매핑시켜 객체의 위치 이동을 나타내었다. 움직임 객체의 공간적 크기 변화를 나타내는 MBR은 비디오 화면 자체를 축으로 하여 절대 좌표로 매핑시켜 위치 이동을 나타내었다. 그림 13은 농구 비디오 데이터로부터



(a) 프레임에서의 객체 움직임 추출



(b) 절대 좌표로 매핑

그림 13. 농구 비디오 데이터로부터 객체의 궤적 정보를 추출 예

객체의 궤적 정보를 추출한 예이다. 먼저 그림 13의 (a)와 같이 키프레임을 추출하여 키프레임에서 특정 객체의 움직임을 파악한다. 이를 통해 그림 13의 (b)와 같이 양쪽 골대를 기준으로 해서 (a)에 있는 특정 객체의 움직임을 표현하였다. 농구 경기장을 축으로 하여 절대 좌표로 매핑시켜 궤적에 필요한 거리, 방향 그리고 속도를 계산하였다.

움직임 검색 시스템에서 검색 효율의 성능을 평가하기 위해 정확율(Precision)과 재현율(Recall)을 사용하였다. 정확율과 재현율을 통해 사용자의 질의에 따른 시스템에서의 검색 결과를 평가하였다. 정확율과 재현율은 (식 32)과 (식 33)에 의해 계산한다. 이때, Similarity_retrieval_result는 질의에 대해 시스템이 데이터베이스에서 검색한 결과 중에서 유사한 검색 결과의 수이고, Retrieval_result는 질의에 대해 시스템이 데이터베이스에서 검색한 결과의 수이다. Real_retrieval_result는 주어진 질의에 적합하여 데이터베이스에서 반드시 검색되어야 하는 데이터의 수이다.

$$Precision = \frac{Similarity_retrieval_result}{Retrieval_result} \quad (\text{식 } 32)$$

$$Recall = \frac{Similarity_retrieval_result}{Real_retrieval_result} \quad (\text{식 } 33)$$

그림 14는 질의 표본 10개의 궤적을 가지고 움직임 객체 검색 시스템에서 실험한 결과에 대해 재현율과 정확율의 상관 관계를 나타낸 그래프이다. 질의에 대해 재현율을 0.1, 0.2, ..., 1의 10단계의 표준적인 값으로 미리 정해놓고 이에 이를 때까지 검색을 계속하여 정확율을 구하는 방법을 사용하여 이 시스템의 검색 효율을 나타내는 성능 곡선을 산출하였다. 재현율을 기준으로 정확율을 구하여 그래프를 나타낸

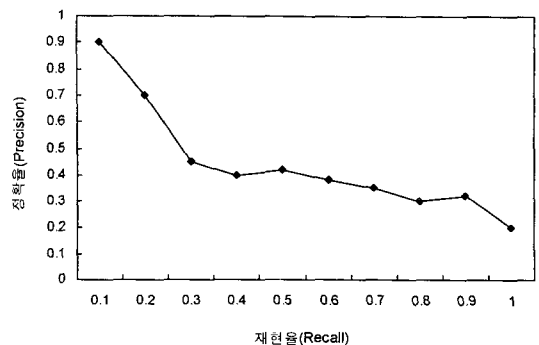


그림 14. 정확율-재현율

표 2. 모델링 기법들의 비교

비교대상 비교항목	Zaher AGHBARI	John. Z. Li	Choon-Bo Shim	제안하는 모델링
비디오 구조와 객체의 독립 여부	구조 내에 표현	독립적으로 표현	독립적으로 표현	독립적으로 표현
움직임 객체 표현	표현하지 않음	표현	표현	표현
궤적 표현 방법	표현하지 않음	거리, 방향(8가지 방향 관계)	거리, 방향(각)	방향(각), 거리, 속도
의미 기반 검색	제공하지 않음	제공하지 않음	제공하지 않음	제공
시각적 특징 유사도 검색	제공	제공하지 않음	제공하지 않음	제공
거리 유사도 검색	제공하지 않음	제공하지 않음	제공하지 않음	제공
궤적 유사도 검색시 고려 요소	시각적 특징	방향	방향, 거리	방향, 거리, 속도

결과, 전체의 궤적 중에서 질의에 대해 검색된 궤적의 수가 증가할수록 유사한 궤적이 많이 검색되므로 재현율이 증가하지만, 검색된 궤적의 수 중에는 유사하지 않은 궤적들 또한 포함되므로 정확율이 떨어지는 것을 볼 수 있었다.

5.3 모델링 기법들의 비교 분석

제안하는 움직임 객체 모델링을 기존의 모델링 기법들과 비교 분석을 수행하였다. 제안하는 모델링은 Zaher Aghbari[11], John. Z. Li[12] 그리고 Choon-Bo Shim[13] 모델링 기법과 비교한다. 표 2에서와 같이 제안하는 모델링은 기존의 모델링에 비해 움직임 객체를 효과적으로 표현할 수 있을 뿐만 아니라 움직임 객체에 대한 다양한 검색을 지원한다. 이를 위해 제안하는 모델링에서는 비디오의 구조적인 정보와 독립적으로 객체 중심으로 표현하며, 시간에 따른 공간적인 위치와 크기 변화를 함께 표현한다. 검색 유형은 의미 기반 검색과 시각적 특징 유사도 검색뿐만 아니라 객체의 공간적 크기까지 표현하기 때문에 기존의 모델링들에서는 제공하지 않았던 객체 간의 거리 유사도 검색을 지원한다. 또한 궤적의 유사도에서는 다른 모델들과는 달리 방향, 거리 그리고 속도를 고려하여 보다 정확한 궤적 검색을 수행한다. 그리고 가중치 혼합 유사도 검색을 지원함으로써 다양한 검색 유형을 제공한다.

6. 결론 및 향후 연구

이 논문에서는 움직임 객체를 효과적으로 표현하기 위한 새로운 모델링 기법과 다양한 검색 기법을 제안하였다. 또한 움직임 객체 검색을 위해 구현한 검색 시스템에서 입력받은 검색 조건에 대해 효과적으로 수행됨을 볼 수 있었다. 제안하는 모델링에서는 움직임 객체를 효과적으로 표현하기 위해서 움직임 객체의 공간적 위치와 공간적 크기를 함께 표현하였다. 또한 시간의 변화에 따른 객체의 시각적인 특징 변화와 움직임 객체의 궤적을 표현하였다. 제안하는 모델링에서는 움직임 객체의 특성에 따른 궤적에 초점을 맞춰 궤적의 유사도뿐만 아니라 시간의 변화에 따라 수행하는 객체의 동작과 연관된 사건에 기반한 의미 검색, 시각적 특징 유사도, 거리의 유사도 그리고 이들을 혼합한 가중치 혼합 유사도 검색의 다양한 검색 유형을 제공한다.

향후 연구 방향으로 특정 질의 유형에 대한 질의 처리 비용 및 정확도를 향상시키기 위한 연구와 함께 제안한 검색 기법을 실제적인 대용량의 비디오 데이터에 적용하여 성능 평가를 수행할 예정이다.

참고 문헌

[1] E. Oomoto and K. Tanaka, "OVID: Design and

- Implementation of a Video-Object Database System," IEEE Trans. on Knowledge and Data Engineering, Vol.5, No.4, pp.629-643, 1993.
- [2] R. Hjelsvold, "VideoSTAR - A Database for Video Information Sharing," Ph.D. Thesis, Norwegian Institute of Technology, Nov. 1995.
- [3] M. La Cascia and E. Ardizzone, "JACOB: Just a content-based query system for video databases," ICASSP 96, pp.7-10, 1996.
- [4] S. F. Chang, W. Chen, H. J. Meng, H. Sundaram and D. Zhong, "VideoQ: An Automatic Content-Based Video Search System Using Visual Cues," Pro. ACM Multimedia 97 Conference, Seattle, WA, November, pp.313-324, 1997.
- [5] A. Hauptmann and M. Witbrock, "Informedia: News-on-Demand Multimedia Information Acquisition and Retrieval," Intelligent Multimedia Information Retrieval, Mark T. Maybury, Ed., AAAI Press, pp.213-239, 1997.
- [6] S. Adali, K. S. Candan, K. Erol and V. S. Subrahmanian, "AVIS: An Advanced Video Information System," The Institute for Systems Research, University of Maryland, Technical Report, TR 97-44, pp.369-373, 1997.
- [7] Y. M. Park, Y. K. Kim, S. I. Jin and W. Choi, "Hierarchical Structure-based Metadata Model for Video Database Applications," ISCA COMPUTERS AND THEIR APPLICATIONS, pp. 242-245, March, 1998.
- [8] Yong Rui and Thomas Huang, "Chapter 9.2: A Unified Framework for Video Browsing and Retrieval," The Image and Video Processing Handbook edited by Alan Bovik, Academic Press, pp.705-715, 2000.
- [9] 윤미희, 윤용익, 김교정, "비디오의 의미검색과 유사성검색을 위한 통합비디오 정보 시스템", 정보처리논문지, 6권, 8호, pp.2031-2041, 1999.
- [10] M. Petkovic and W. Jonker, "A Framework for Video Modeling," Pro. Eighteenth IASTED International Conference Applied Informatics, Innsbruck, Austria, February 2000.
- [11] Zaher Aghbari, Kunihiko Kaneko and Akifumi Makinouchi, "Modeling and Querying Videos by Content Trajectories," IEEE International Conference on Multimedia and Expo (I) 2000, pp.463-466, 2000.
- [12] J. Z. Li, M.T. Ozsdu and D. Szafron, "Modeling of Moving Objects in a Video Database," Pro. IEEE International Conference on Multimedia Computing and Systems, pp.336-343, 1997.
- [13] Choon-Bo Shim and Jae-Woo Chang, "Spatio-temporal representation and retrieval using moving object's trajectories," ACM Multimedia Workshops 2000, pp.209-212, 2000.
- [14] Zaher Aghbari, Kunihiko Kaneko and Akifumi Makinouchi, "A Motion-Location Based Indexing Method for Retrieving MPEG Videos," DEXA Workshop 1998, pp.102-107, 1998.
- [15] Zaher Aghbari, Kunihiko Kaneko and Akifumi Makinouchi, "VST-Model: A Uniform Topological Modeling of the Visual-Spatio-Temporal Video Features," ICMCS, Vol.2, pp.163-168, 1999.
- [16] J. F. Allen, "Maintaining knowledge about temporal intervals," Communications of ACM, Vol.26, No.11, pp.832-843, Nov. 1983.
- [17] M. J. Egenhofer, "Point-Set Topological Spatial Relations," International Journal of Geographical Information Systems, Vol.5, No.2, pp.161-174, 1991.
- [18] M. Nabil, A. H. H. Ngu and J. Shepherd, "Modeling and Retrieval of Moving Objects," Multimedia Tools and Applications, Vol.13, No.1, pp.35-71, 2001.



복 경 수

1998년 북대학교 수학과 이학사
2000년 충북대학교 정보통신공학과 공학석사
2000년~현재 충북대학교 정보통신공학과 및 컴퓨터정보통신연구소 박사과정

관심분야: 자료 저장 시스템, 멀티미디어 데이터베이스, 이동객체 데이터베이스, 고차원 색인 구조, 시공간 색인구조 등



유 재 수

1989년 전북대학교 컴퓨터공학과 공학사
1991년 한국과학기술원 전산학과 공학석사
1995년 한국과학기술원 전산학과 공학박사
1995년~1996년 8월 목포대학교

전산통계학과 전임강사
1996년 8월~현재 충북대학교 정보통신공학과 및 컴퓨터정보통신연구소 부교수
관심분야: 데이터베이스 시스템, 정보검색, 멀티미디어 데이터베이스, 분산 객체 컴퓨팅 등



김 미 희

2001년 충북대학교 컴퓨터공학과 공학사
2003년 충북대학교 정보통신공학과 공학석사

관심분야: 데이터베이스 시스템, 멀티미디어 데이터베이스 등



조 기 형

1966년 인하대학교 전기공학과 공학사
1984년 청주대학교 산업공학과 공학석사
1992년 경희대학교 전자공학과 공학박사
1981년~1988년 충주공업전문대

학 조교수
1996년~1999년 전기전자공학부장
1988년~현재 충북대학교 정보통신공학과 및 컴퓨터정보통신연구소 교수
관심분야: 데이터베이스시스템, 화상처리 및 통신, 통신 프로토콜, 분산 객체 컴퓨팅 등



신 재 룡

1996년 충북대학교 정보통신공학 공학사
1998년 8월 충북대학교 정보통신공학 공학석사
2002년 8월 충북대학교 정보통신공학 공학박사
2003년~현재 광주보건대학 인터

넷정보과 전임강사
관심분야: 데이터베이스 시스템, 멀티미디어 데이터베이스, 실시간 시스템, 정보검색 등