

객체 이동 궤적 기반 비디오의 검색

정희원 유웅식*, 이규원**, 김재곤*, 김진웅*, 권오석***

Video Retrieval based on Objects Motion Trajectory

Woong-Shik You*, Kyu Won Lee**, Jae-Gon Kim*, Jinwoong Kim**, Oh-Seok Kwon***

Regular Members

요약

객체의 이동 궤적에 대한 효율적 기술기(descriptor) 및 이를 이용한 내용기반 비디오 검색 기술을 제안하였다. 정지 배경의 비디오에서 이동 객체를 추출한 후, 객체의 중심 좌표의 시간에 따른 이동 궤적을 2차 다항식의 매개변수로써 표현하였다. 또한, 카메라의 움직임이 있는 유동 배경의 비디오에서 비디오의 전역 움직임을 어파인(affine) 모델링한 후, 어파인 매개변수로써 카메라의 움직임을 식별하고 색인화 하였다. 이동 궤적에 대한 질의 매개변수와 색인화 된 매개변수간 유사도 정합에 의하여 내용 기반의 비디오 검색을 실현하였다. 제안한 기술은 일반 비디오에 대한 신속한 검색뿐만 아니라 감시 비디오 시스템의 효율적 운영에 기여할 수 있을 것으로 기대된다.

ABSTRACT

This paper proposes an efficient descriptor for objects motion trajectory and a video retrieval algorithm based on objects motion trajectory. The algorithm describes parameters with coefficients of 2-order polynomial for objects motion trajectory after segmentation of the object from the scene. The algorithm also identifies types, intervals, and magnitude of global motion caused by camera motion and indexes them with 6-affine parameters. This paper implements content-based video retrieval using similarity-match between indexed parameters and queried ones for objects motion trajectory. The proposed algorithm will support not only faster retrieval for general videos but efficient operation for unmanned video surveillance system.

1. 서론

최근 인터넷의 급속한 확산, 디지털 방송의 시작, 대용량 저장매체의 대중화 등으로 인하여 많은 양의 멀티미디어 데이터가 생성되고 있으며, 이러한 데이터를 다양한 목적에 이용하려는 사용자의 요구가 점점 증가하고 있다. 사용자가 원거리에서 있는 멀티미디어 데이터를 이용하기 위해서는 먼저 원하는 자료를 검색하여야만 한다. 현재 텍스트를 기반으로 하는 검색 시스템은 상당한 수준의 검색 서비스를

제공하고 있으나, 멀티미디어 데이터의 경우 텍스트로 표현하는 데 있어서의 제약과 자료마다 일일이 주석을 기재해야 하는 부담으로 인하여 효과적인 검색에는 한계가 있다. 이러한 문제점을 해결하기 위하여 내용 기반의 검색 및 브라우징에 대한 관심이 고조되었고, MPEG(Moving Picture Expert Group)은 "멀티미디어 내용 기술(description) 인터페이스" (MPEG-7)라고 하는 새로운 표준화 작업을 시작하였다.

MPEG-7은 사용자가 원하는 다양한 형태의 데이터를 빠르고 효과적으로 검색하기 위한 내용 기술

* 한국전자통신연구원 무선방송기술연구소

** 대전대학교 정보통신공학과(kwlee@dragon.taejon.ac.kr),
논문번호 : 99473-1129, 접수일자 : 1999년 11월 29일

*** 충남대학교 컴퓨터공학과(oskwon@ce.cnu.ac.kr)

※ 본 연구는 정보통신부 출연금으로 수행되었습니다.

표준화를 목적으로 하고 있다^[1,2]. MPEG-7은 비디오, 오디오, 텍스트 정보 등을 포함하는 다양한 종류의 데이터를 대상으로 하고 있으며, 객체의 컬러, 색조, 모양, 이동 궤적등과 같은 시각적 정보에 대한 기술이 중요한 부분을 차지하고 있다. 그 중 객체의 이동 궤적은 객체가 갖는 중요한 특징 중 하나이며, 이를 이용하여 영상 자료의 내용을 기술하고 이를 검색에 이용하려는 연구들이 활발하게 이루어지고 있다^[3,4,5,6].

본 논문은 객체의 이동 궤적을 매개변수들로서 기술하고 이를 이용한 비디오 검색 기술을 제안한다. 우선, 객체 분리 알고리즘을 사용하여 객체의 중심 좌표를 추출하고 객체의 이동 궤적을 구한다. 객체의 전체 이동 궤적을 여러 개의 시구간으로 나누어 각 시구간 궤적을 2차 다항식으로 표현하며 그 계수들을 색인화하여 검색에 이용한다. 또한, 카메라에 의한 배경 움직임이 있는 영상의 검색을 위하여 카메라 움직임의 종류, 구간, 크기를 매개변수들로 기술하여 배경 움직임을 객체의 이동 궤적에 보상함으로써 이동 카메라 환경에서의 비디오 검색도 가능하도록 하였다. 이동 궤적에 기반한 다양한 질의에 대하여 정합 함수를 정의하고 질의로서 주어진 질의 매개변수와 색인화 된 매개변수 간 유사도 정합에 의하여 효율적, 효과적인 비디오 검색을 실현하고, 실험을 통하여 제안한 방법의 성능을 평가하였다.

논문의 구성은 다음과 같다. 2장에서는 관련연구에 대해서 간단하게 살펴보고, 3장에서는 객체 이동 궤적의 기술에 대해서, 4장에서 객체 이동 궤적을 이용한 질의 및 검색, 5장에서 실험 및 결과를 설명하고, 6장에서 결론을 맺는다.

II. 관련 연구

객체의 이동 궤적을 이용한 비디오 검색을 위한 연구는 내용 기반 비디오 검색 분야에서 비디오의 특성을 파악하는 다른 특징들과 함께 중요한 연구 이슈가 되고 있다.

Shan과 Lee는 객체 이동 궤적을 이용한 비디오 검색을 위해 전체 궤적의 각 세그먼트의 궤적 기울기를 구하고 이 기울기들의 리스트로 객체의 이동 궤적을 표현하였다^[5]. 이 방법은 객체 궤적의 공간 정보인 기울기를 기술하고 이를 검색에 이용하였으나, 객체 궤적에 대한 시간 정보를 기술하지 않음으로써 이를 이용한 질의를 지원하지 못하는 단점이

있다.

또한, Chang은 그의 연구에서 객체 이동 궤적을 표현하기 위해 각 프레임간의 객체 중심 좌표의 평균 변이를 구하고, 이의 벡터 리스트로서 전체 궤적을 기술하였다^[6]. 이 방법은 모든 프레임에서의 객체 위치 정보와 시간 정보인 프레임 정보를 같이 기술함으로써 이동 궤적에 대한 정보의 손실을 최소화하였으나, 객체 궤적이 길어질수록 기술되는 정보의 양이 매우 많고, 검색 시 점 대점 정합을 이용하므로 계산량이 많은 단점이 있다.

본 논문에서는 이러한 문제점들을 해결하기 위해 각 시구간에서의 객체 이동 궤적 정보를 소수의 시, 공간 매개변수를 이용하여 표현함으로써 기술되는 데이터의 양을 줄이고, 간단한 매개변수만의 정합을 통한 효율적인 검색을 위한 기술을 가능하게 하였다.

III. 객체 이동 궤적의 기술

1. 제안한 알고리즘 구성

그림 1은 객체 이동 궤적에 기반한 비디오 검색을 위한 전체적인 과정을 보여준다. 먼저 객체를 배경 화면으로부터 분리하여 중심 좌표를 추출한다. 객체의 분리를 위해서 배경의 움직임이 없는 영상은 통계적 움직임 모델에 기반한 영역 분할 알고리즘^[7]을 이용하였고, 배경 움직임이 있는 영상은 반자동 영역 분할 알고리즘을 이용하여 각각 추출하였다^[8]. 각각의 분할 방법을 이용하여 추출된 객체의 중심 좌표 열을 2차 다항식의 매개변수를 사용하여 이동 궤적을 기술한다.

또한, 카메라 움직임에 의한 배경 움직임을 검출하고 이를 정량화 하기 위해서 6개 매개변수를 포함하는 어파인 변환 모형을 사용한다. 어파인 변환 모형의 매개변수는 MPEG 스트림에 포함된 움직임 벡터로부터 추정한다. 배경 영역의 전역 움직임 검출 연산을 통하여 카메라 움직임의 종류와 그 구간 및 크기를 추출해 내고 이를 매개변수들로 기술한다. 배경 움직임이 있는 영상에 대해서는 각 프레임에서 객체의 중심 좌표에 카메라 움직임 매개변수들을 이용하여 배경 움직임을 보상하고 최종적인 객체의 이동 궤적을 기술하여 검색에 이용한다. 사용자 질의 인터페이스를 통하여 입력된 질의는 같은 방법으로 매개변수화 하여 저장되어 있는 객체 이동 궤적과 색인화 된 매개변수간 유사도 정합에 의하여 내용 기반의 비디오 검색을 실행하고 그 결과를 사용자 질의 인터페이스를 통하여 보여준다.

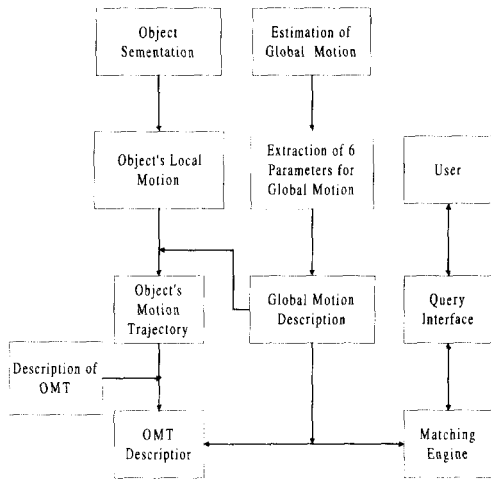


그림 1. 객체 기반 검색을 위한 프로시저

2. 고정 카메라 환경

고정 카메라 환경에서는 배경의 움직임이 없으므로 객체 이동 궤적의 기술을 위해 배경 움직임 보상을 위한 연산이 필요 없다. 우선 통계적 움직임 모델에 기반한 영역 분할 알고리즘을 이용한 객체 분리를 통하여 객체의 중심 좌표들을 추출한다. 이 방법은 유동적인 배경으로부터 안정된 움직임 추출을 위하여 배경과 이동 물체를 통계적 매개변수로 모델링하고, 배경만이 존재하는 초기 연속 영상 중 일부에 대하여 통계적 특징을 학습한다. 초기의 배경만이 존재하는 영상과 이후 영상의 통계적 특징으로부터 Bayes 결정에 의하여 이동 물체를 검출하고 물체의 영역 분할을 행한다^[7]. 이를 통해서 추출된 객체의 중심 좌표들을 연결하면 전체 객체 이동 궤적을 구할 수 있다. 객체의 전체 이동 궤적을 식 1의 2차 다항식으로 표현하여 다항식의 f, v, a 계수들을 객체 이동 궤적의 매개변수로 기술한다. 연속된 프레임에 대한 객체의 중심 좌표 열을 모두 저장하지 않고 이동 궤적에 근사화 된 매개변수만을 저장한다. 검색 시에는 저장된 매개변수들을 이용하여 객체의 원래 궤적을 복원하거나 매개변수 자체를 질의에 사용할 수 있다.

$$f(t) = f_i + v_i(t-t_i) + \frac{1}{2}a_i(t-t_i)^2 \quad (1)$$

- $f(t)$: 시간 t 에서 x, y 의 위치
- f_i : i 번째 시구간의 시작 시간에서의 x, y 의 위치
- v_i : i 번째 시구간의 시간 t 에서의 속도 성분
- a_i : i 번째 시구간의 시간 t 에서의 가속도 성분
- t_i : i 번째 시구간의 시작 시간

이러한 근사화를 통한 기술은 매우 적은 양의 데이터로써 객체 이동 궤적을 표현할 수 있는 장점이 있는 반면, 전체 궤적에 대해서 f, v, a 의 3개의 매개변수만으로 객체 이동 궤적을 표현한다면 실제 궤적과 많은 오차가 존재하게 된다. 이러한 오차는 객체의 이동 궤적이 길어질수록 궤적 근사화 과정에서 누적됨으로써 더욱 커지게 된다. 또한, 궤적 근사화 과정에서의 오차로 인하여 정확한 객체 궤적을 재구성하지 못하고, 정확한 검색도 기대할 수 없다.

본 논문에서는 객체 이동 궤적 재구성 과정에서 누적 오차를 줄이고 각각의 구간 정보만을 이용한 효과적인 질의를 위해서 전체 궤적 구간을 몇 개의 시구간으로 나누고, 각 시구간을 시작 프레임 번호와 시작 프레임의 객체 위치 f , 끝 프레임 번호와 끝 프레임의 객체 위치, 시구간의 v 와 a 들을 매개변수로 기술한다. 표 1은 비디오 객체의 이동 궤적을 기술하기 위한 데이터의 구조와 의미를 나타낸다.

그림 2는 연속된 프레임 축상에서 객체의 x 와 y 궤적의 변화를 그래프로 나타낸 것이다. 그림 2(a)는 객체의 실제 이동 궤적을, 그림 2(b)는 기술된 매개변수들을 이용하여 각 시구간별로 재구성된 궤적을 보여준다. 근사화된 매개변수들을 이용하여 재구성된 객체 궤적과 실제 객체 궤적이 거의 일치하고 있음을 확인할 수 있다.

표 1. 객체의 이동 궤적을 기술하기 위한 데이터 구조

요소	값	데이터형
Object_ID	객체 구별자	int
N_interval	시구간의 개수	int
Start_time[N_interval]	각 시구간 시작프레임 번호	int
End_time[N_interval]	각 시구간 끝 프레임 번호	int
Point_x[N_interval]	각 시구간 첫 프레임 객체 중심 x좌표	float
Point_y[N_interval]	각 시구간 첫 프레임 객체 중심 y좌표	float
Parameter_vx[N_interval]	각 시구간 x좌표 궤적의 속도	float
Parameter_vy[N_interval]	각 시구간 y좌표 궤적의 속도	float
Parameter_ax[N_interval]	각 시구간 x좌표 궤적의 가속도	float
Parameter_ay[N_interval]	각 시구간 y좌표 궤적의 가속도	float
Half_Width_of_Object	객체의 평균 너비	float
Half_Height_of_Object	객체의 평균 높이	float

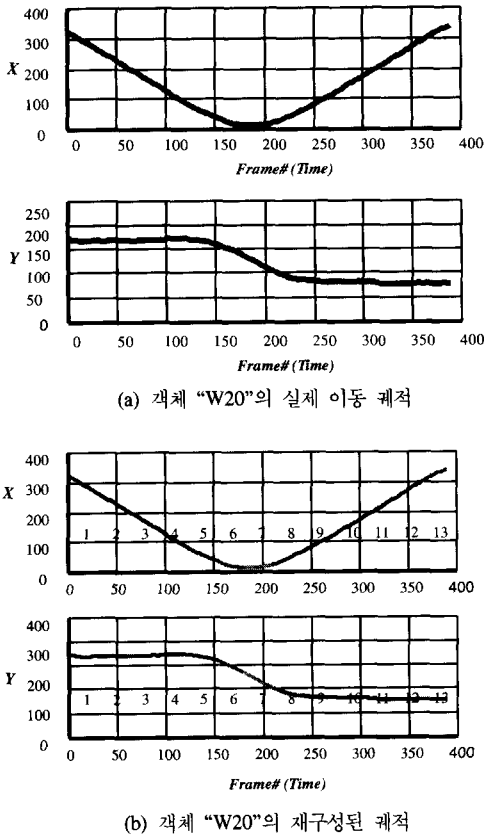


그림 2. 객체의 구간별 실제 궤적과 재구성된 궤적

3. 이동 카메라 환경

배경 움직임이 없는 고정 카메라 환경에서는 객체 분리를 통하여 객체의 중심 좌표들을 추출하면 이를 이용하여 객체의 이동 궤적을 기술할 수 있다. 그러나, 배경 움직임이 있는 영상은 각 프레임에 나타나는 객체의 중심 좌표들에 대한 기술만으로는 객체의 실제 이동 궤적을 표현할 수 없다. 배경 움직임이 있는 영상에서의 객체 이동 궤적은 카메라의 움직임에 따라서 실제 이동 궤적과 다르게 나타나기 때문이다.

이동 카메라 환경에서 객체의 실제 이동 궤적을 추출하기 위해서는 객체 분리를 통해 객체의 위치를 추출하여야 한다. 이동 카메라 환경에서의 객체 분리를 위해서 반 자동 영역 분할 알고리즘을 이용한다. 반 자동 영역 분할 알고리즘은 초기 영상에서 사용자가 분리하고자 하는 객체 영역의 경계(boundary)를 그리면 해당 영역의 객체가 배경 영상으로부터 분리되고, 분리된 객체 영역 정보를 이용하여 연속된 프레임상에서 자동으로 객체 분리를 행한다^[8].

또한, 배경 움직임이 있는 비디오의 검색을 위해서는 카메라 움직임의 종류와 그 구간 및 크기를 추출하여 이를 각 프레임상에 나타나는 객체의 움직임에 보상해주는 과정이 필요하다.

1) 카메라 움직임 기술

카메라 움직임 정보를 추출하기 위하여 MPEG 스트림에 포함된 움직임 벡터로 각 프레임의 움직임 벡터장을 구성하고 움직임 벡터장으로부터 어파인 변환의 6-매개변수를 계산한다^[9]. MPEG 스트림의 움직임 벡터가 실제 모션 플로우(motion flow)를 나타내는 것은 아니지만 대부분의 경우에서 카메라 매개변수를 계산하기에 충분하고, 객체 혹은 배경과 카메라 사이의 거리가 클 때 6-매개변수 어파인 변환을 이용하여 각 프레임의 배경 움직임을 표현할 수 있다^[10]. 각 프레임에서의 움직임 벡터는 식 2와 같이 나타난다.

$$\begin{bmatrix} U \\ V \end{bmatrix} = \begin{bmatrix} 1 & x & y & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & x & y \end{bmatrix} \cdot [a_1 a_2 a_3 a_4 a_5 a_6]^T \quad (2)$$

- (x, y) : 현재 프레임에서의 각 매크로 블록의 좌표
- [u v] : 각 매크로 블록의 움직임 벡터(U)
- $[a_1 a_2 a_3 a_4 a_5 a_6]^T$: 어파인 변환 벡터 (\vec{a})

MPEG 스트림의 움직임 벡터는 매크로 블록들의 블록 매칭에 의해서 구해진다. 블록 매칭에 의해 각 매크로 블록의 움직임 벡터를 구한 후 최소 자승 오차방식을 이용하여 식 2로 표현되는 움직임 벡터와 MPEG 스트림으로부터 추출된 움직임 벡터간의 오차를 최소로 하는 어파인 변환 벡터 \vec{a} 을 구할 수 있다^[10]. 어파인 변환 벡터 \vec{a} 의 각 매개변수가 갖는 의미는 다음과 같다^[11].

- a_1 : 카메라 움직임의 panning 성분. 부호는 pan left/right를 나타낸다.
- a_2 : 카메라 움직임의 tilt 성분. 부호는 tilt up/down을 나타낸다.
- $0.5 \times (a_3 + a_6)$: 카메라 움직임의 zooming 성분. 부호는 zoom in/out을 나타낸다.
- $0.5 \times (a_3 - a_6)$: 카메라 움직임의 rotation 성분.

식 2의 6-매개변수 어파인 변환 벡터를 이용하여 카메라 움직임을 구간별로 나누고, 각 카메라 움직임 구간의 시간 정보인 시작과 끝 프레임, 카메라 움직임의 종류, 각 구간에서의 6개 매개변수의 평균 값을 매개변수로 기술한다. 표 2는 카메라 움직임의

기술을 위한 데이터 구조를 보여준다.

표 2. 카메라 움직임의 기술

요 소	값	데이터형
N_camera_motion	카메라 움직임의 수	int
Start_time [N_camera_motion]	각 카메라 움직임 구간의 시작 프레임 번호	int
End_time [N_camera_motion]	각 카메라 움직임 구간의 끝 프레임 번호	int
Camera_motion_type [N_camera_motion]	각 구간 카메라 움직임의 종류	int
Affine [N_camera_motion]*6	어파인 매개변수 각각의 평균	float*6

표 2에 기술된 카메라 움직임은 카메라 움직임 보상을 통하여 객체의 실제 궤적을 구하는데 사용되며, 그 자체만으로도 카메라 움직임에 기반한 검색에 이용될 수 있다.

2) 카메라 움직임의 보상

(Global Motion Compensation)

카메라 움직임에 대한 기술이 이루어지면 이를 이용하여 식 3과 같이 전역 움직임을 보상함으로써 객체의 실제 이동 궤적을 구한다. 카메라 움직임에 의한 배경의 움직임을 객체의 움직임에 보상해주면 객체의 실제 움직임(local motion)과 이동 궤적을 구할 수 있다. 식 2에서 $[u \ v]$ 는 카메라의 움직임에 의해 영상의 각 매크로 블록이 어느 방향으로 얼마만큼 이동하였는지를 의미하였다. 그러나, 카메라 움직임 보상 과정에서의 매크로 블록은 현재 프레임에서 객체가 위치하고 있는 매크로 블록을 의미하므로, $[u \ v]$ 는 현재 프레임에서의 객체 영역이 속한 매크로 블록의 움직임 방향과 크기를 나타낸다. (x, y) 는 현재 프레임에서의 매크로 블록의 위치, 즉 객체 영역의 위치를 의미한다. 현재 프레임에서의 $[u \ v]$ 와 (x, y) 를 이용하여 연속된 프레임 상에서 그림 3과 같이 (x_0, y_0) 에서로의 객체의 실제 움직임 M 을 구할 수 있다.

$$M = X * \vec{a} - U \tag{3}$$

$M(m_x, m_y)$: 연속된 두 프레임에서 객체의 실제 움직임 벡터

U : 연속된 두 프레임에서의 움직임 벡터

\vec{a} : 연속된 두 프레임에서 배경의 움직임 벡터

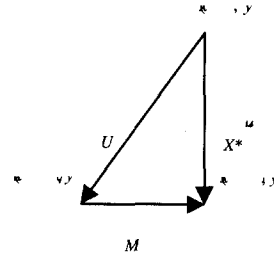


그림 3. 배경 움직임 $X * \vec{a}$, 객체의 실제 움직임 M , 움직임 벡터 U 의 관계

IV. 질의 및 검색

1. 질의 및 검색

질의는 사용자 질의 인터페이스를 통하여 이루어진다. 질의 인터페이스를 통하여 객체의 이동 궤적과 카메라 움직임을 이용한 다양한 형태의 질의가 가능하고, 질의에 대한 검색 결과를 확인하고 검색 구간을 직접 접근할 수 있다. 각각의 질의는 사전에 정의된 기술자로 변환되어 저장된 기술 데이터와의 정합이 이루어지고, 검색 결과는 유사도에 의해 순위화 되어 질의 인터페이스를 통하여 보여진다.

논문에서는 객체 이동 궤적을 위한 다양한 질의와 정합 함수를 정의하고 이를 구현하였다. 또한, 검색을 위하여 다양한 종류의 질의 특성에 따라 다른 정합 함수를 사용하고 있다. 각각의 질의를 위해서 사용한 용어의 정의와 정합 함수(matching function)는 다음과 같다.

1) 용어 정의

Q : 질의 매개변수

S : 저장되어 있는 매개변수

f : 객체의 위치

v : 객체의 이동 속도 성분

a : 객체의 이동 가속도 성분

α, β, γ : 궤적 모양, 속도, 가속도에 대한 가중치

$d(Q,S)$: 질의 매개변수와 저장된 매개변수 간의 Euclidean 거리

i : 각 시구간의 인덱스

2) 정합 함수

- $Q1$: 객체의 이동 방향에 기반한 질의
 객체의 이동 방향을 이용한 질의이다. 그림 4와 같이 객체의 질의 이동 방향을 8방향으로 나누고, 질의 이동 방향과 저장된 객체의 실제 이동 방향을 식 5의 정합 함수를 이용하여 비교한다. 질의 이동

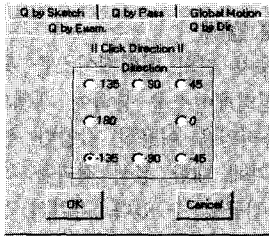
방향과 같거나 유사한 비디오 객체들을 검색하여 사용자 인터페이스를 통하여 결과를 보여준다. 객체 궤적의 시작과 마지막 위치를 기준으로 하는 객체의 전체 이동 궤적 방향 검색과 각 시구간에서의 이동 방향을 이용하는 검색 모두 가능하다. 전체 이동 궤적 방향을 이용한 정합 함수는 식 5와 같다.

$$d(Q, S) = d(\theta_q, \theta_s) \tag{5}$$

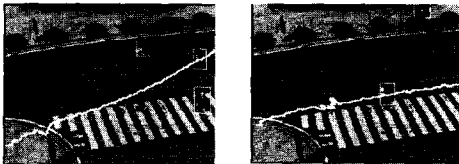
$$\theta = \arctan(dy/dx)$$

$$dx = Point_x[N_interval-1] - Point_x[0]$$

$$dy = Point_y[N_interval-1] - Point_y[0]$$

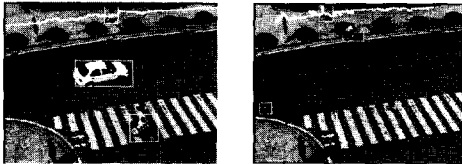


(a) 질의 (-135)



(b) 1순위(R15)

(c) 2순위(R12)



(d) 3순위(R2)

(e) 4순위(R7)

그림 4. 객체 이동 방향에 기반한 질의와 검색 결과

- Q2 : 이동 속도에 기반한 질의

객체의 이동 속도에 기반한 질의는 객체의 이동 속도를 프레임당 이동 픽셀의 수(pixels per frame)로 환산하여 질의하고 질의 이동 속도에 해당하는 구간을 검색하여 준다. 그림 5에 나타난 질의 예는 일정 속도 이하의 저속도 값의 질의를 가하였을 때, 객체가 정지한 구간이 검색된 결과를 보여 준다. 이동 속도에 기반한 질의를 위한 정합 함수는 v_q 가 질의 속도라 할 때 식 6과 같다.

$$d(Q, S) = |v_q| - \sqrt{v_x^2 + v_y^2} \tag{6}$$



(a) 객체 Y7(Fr. 937) (b) 객체 Y7(Fr. 950) (c) 객체 Y7(Fr. 966)

그림 5. 객체의 이동 속도에 기반한 질의와 검색 결과 ($v \leq 0.2$ ppf, Y7(Fr. 937-966))

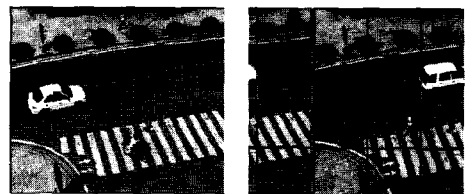
- Q3 : 객체가 지나는 임의의 점/영역에 기반한 질의
영상의 임의의 지점이나 영역이 질의로써 주어지면 질의 지점 혹은 영역을 지나는 객체를 검색하는 질의이다. 그림 6(a)와 같이 영상의 한 지점 혹은 그림 6(b)의 16개의 영역 중 한 영역을 질의로써 입력한다. 객체의 이동 궤적 중 식 7의 정합 함수 조건을 만족하여 질의 지점 C_q 을 통과하는 구간만을 객체 식별자와 함께 인터페이스를 통하여 보여주고 검색 구간을 직접 접근하여 확인할 수 있다. 그림 6(a)는 질의 지점을 지나는 객체의 검색 결과의 예를, 그림 6(b)는 질의 영역을 지나는 객체의 검색 결과의 예를 각각 보여주고 있다.

$$d_x \leq |Half_Width_of_Object| \ \& \ d_y \leq |Half_Height_of_Object| \tag{7}$$

$C_q = (C_{qx}, C_{qy})$: 질의 좌표

C_s : 재구성된 각 프레임의 객체 중심 좌표

$$d_x = |C_{qx} - C_{sx}|, \quad d_y = |C_{qy} - C_{sy}|$$



(a) 질의 좌표(73,111)를 지나는 객체(R5)

(b) 질의 영역을 지나는 객체(R13)

그림 6. 객체가 지나는 임의의 지점 또는 영역에 기반한 질의와 검색 결과의 예

- Q4 : 비디오 객체 예제에 기반한 질의

비디오 객체가 질의 예제로써 주어지면, 주어진 예제 객체와 유사한 궤적, 속도 및 가속도를 갖는 비디오 객체들을 검색하는 질의이다. 비디오 객체 예제에 기반한 질의는 표 3에 주어진 정합 가중치 α, β, γ 의 값을 조정함으로써 궤적의 모양, 객체의 이동 속도와 가속도를 이용하여 다양한 사용자 질의가 가능하다. 이를 위한 정합 함수는 식 8과 같

고 그림 7은 질의와 검색 결과의 예이다.

$$d(Q, S) = \sum_i (\alpha * d_i(f_q, f_s) + \beta * d_i(v_q, v_s) + \gamma * d_i(a_q, a_s)) \quad (8)$$

표 3. 정합 함수에 사용한 가중치

가중치 \ 값	0	1	2	3	4
α	0	0.5	1.0	1.5	2
β	0	25	50	75	100
γ	1	1	1	1	1



(a) 질의예제(Y2) (b) 1순위(Y2) (c) 2순위(Y4)
그림 7. 비디오 객체 예제에 기반한 질의와 검색 결과 ($\alpha = 2, \beta = 0, \gamma = 1$)

- Q5 : 궤적 스케치에 기반한 질의

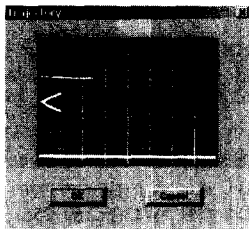
사용자가 직접 질의 궤적을 스케치하여 질의 궤적과 유사한 궤적을 갖는 비디오 객체들을 검색하는 질의이며, 질의 형태는 그림 8(a)와 같다. 스케치된 궤적상의 질의 점들의 좌표는 질의 궤적의 각 시구간 시작 프레임의 객체 좌표로 변환된다. 정합 시 스케치된 질의 궤적의 구간을 색인화 된 각 비디오 객체들의 구간에 맞게 정규화한 후 각 시구간의 매개변수간 유사도 정합을 통하여 유사한 궤적을 갖는 객체들을 검색한다. 이 질의를 위한 정합 함수는 식 9와 같고 그림 8은 질의의 예이다.

$$d(Q, S) = \sum_{i=0}^N d_i(f_q, f_s) \quad (9)$$

q : 질의 점들의 좌표

s : 저장된 객체의 각 시구간 시작 프레임들의 객체 좌표

$$\begin{aligned} & \text{if } (n(q) \geq n(s)) \\ & N = n(q) - n(s) \\ & \text{else } N = n(s) - n(q) \end{aligned}$$



(a) 스케치에 의한 질의 궤적

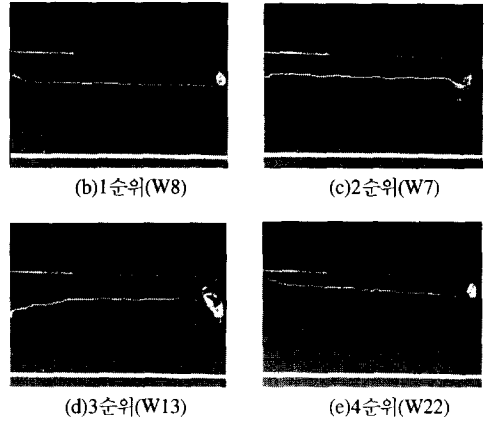
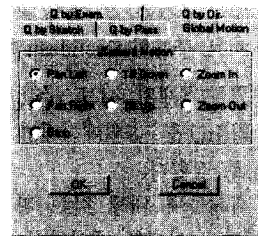


그림 8. 스케치에 기반한 질의와 검색 결과

- Q6 : 카메라의 움직임에 기반한 질의

카메라의 움직임을 그림 9(a)와 같이 7개의 종류로 나누고, 각 질의로써 주어진 카메라 움직임의 종류와 일치하는 카메라 움직임을 갖는 객체 구간을 검색하는 질의이다. 그림 9(b)~(e)는 질의의 결과를 전체 움직임 벡터와 함께 나타낸 검색 결과이다.



(a) 카메라 움직임에 기반한 질의(pan left)

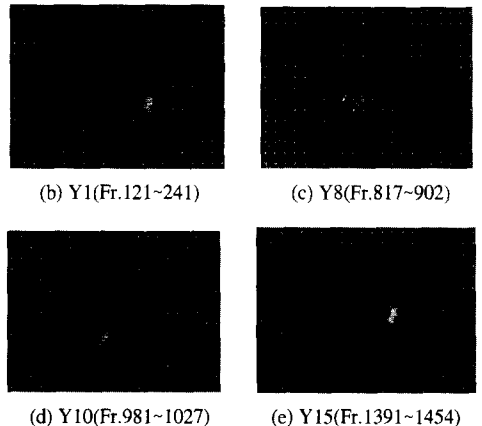


그림 9. 카메라 움직임에 기반한 질의와 결과

3) 정합구간의 정규화

검색 시 정합 되는 질의 구간과 색인화 된 객체의 궤적 구간의 수가 다를 때 효율적이고 정확한

검색을 위하여 정합 구간 정규화(normalization)가 필요하다. 정합 구간 정규화 과정은 다음과 같다.

첫째, 질의 궤적과 객체 궤적 중 짧은 구간을 기준 궤적 구간으로 정한다.

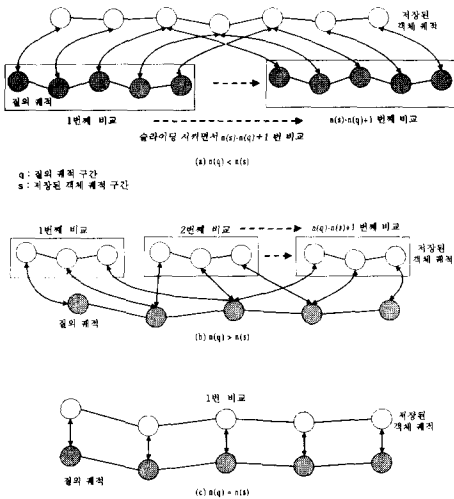


그림 10. 정합 구간 정규화

둘째, 기준 궤적 구간을 긴 궤적 구간에 슬라이딩 시켜가면서 각 서브 구간의 Euclidean 거리를 구한다.

셋째, $|n(q)-n(s)+1|$ 개의 서브 구간 Euclidean 거리 값들 중 최소값을 질의에 대한 해당 객체의 Euclidean 거리로 정한다.

넷째, 전체 객체들에 대해 같은 과정을 반복하여 질의 궤적에 대한 유사도를 계산한다.

그림 10은 질의 궤적 구간과 색인화 된 객체 궤적 구간간의 정합구간 정규화 과정을 보여준다.

V. 실험 및 결과

1. 실험데이터

본 논문에서는 고정 카메라 환경과 이동 카메라 환경에서 각각 비디오를 취득하여 MPEG-1 인코딩 한 실시간 동영상들을 실험 데이터로 이용하였다. 실험 데이터는 배경 화면의 움직임이 없는 동영상("Road", "Walkmen") 2개와 카메라 움직임에 의한 배경 화면의 움직임이 있는 동영상("Yard") 1개로 구성되어 있다. 그림 17, 18, 19는 각 동영상 시퀀스에서 나타나는 비디오 객체들과 객체들이 갖는 실제 궤적들을 보여준다.

2. 실험 결과 및 평가 방법

최근 객체 이동 궤적을 이용한 멀티미디어 데이

터 검색을 위한 연구가 활발히 진행되고 있으나, 연구를 위한 실험 데이터와 실험 결과를 평가하는 객관적인 방법이 아직 존재하지 않는다. 본 논문에서는 실험 결과를 평가하기 위한 수단으로서 다음과 같은 검색 정확도 평가 방법 1과 평가 방법 2을 이용하여 실험 결과를 평가하였다. ground truth를 정의하는 객관적 기준의 모호함으로 인해 제안한 평가 방법들이 갖는 제한에도 불구하고 실험 결과의 참고 자료로써 제시한다.

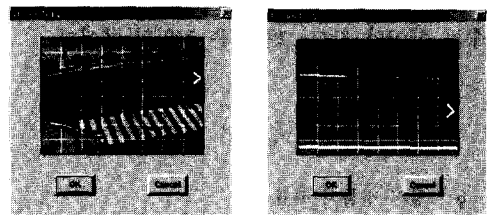
1) Retrieval Accuracy Measure-1

평가 방법 1은 다음과 같이 실험 결과를 평가한다.

첫째, 매개변수에 의해 재구성된 객체의 궤적과 질의 궤적간의 점 대점 정합을 통하여 두 궤적간의 Euclidean 거리를 구하고 이 결과로 나온 순위를 질의에 대한 ground truth로 삼는다.

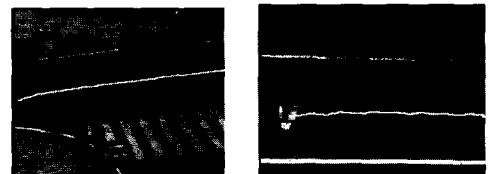
둘째, 실제 정합에 있어서는 논문에서 제안한 매개변수 정합만을 이용하여 검색 결과를 얻고, 이 결과를 ground truth와 비교하고 식 10을 이용하여 검색율을 구하는 방법이다.

그림 11과 12는 실험 평가에 사용된 질의들이다. 그림 13과 14는 각 질의들의 검색 결과를 평가 방법 1에 의해서 평가한 결과로서, 검색된 객체 수와 검색율을 각각 그래프로 나타낸 것이다. 그래프는 ground truth에 따라 반환되는 객체의 수를 하나씩 증가시킬 때, 반환되는 객체 집합 내의 ground truth와 동일한 객체 수의 비율을 보여준다.



(a) 스케치에 기반한 질의 궤적(Road)
(b) 스케치에 기반한 질의 궤적(Walkmen)

그림 11. 스케치에 기반한 질의



(a) 질의 객체 "R9"(Fr. 679~726)
(b) 질의 객체 "W1"(Fr.160~256)

그림 12. 객체 예제에 기반한 질의($\alpha=2, \beta=0, \gamma=1$)

표 4와 5는 실험에 사용한 질의들에 대한 ground truth, 검색 결과 및 유사 궤적을 갖는 객체 집합을 보여준다.

$$RR(q) = \frac{NR(q)}{NG(q)} \quad (10)$$

$NG(q)$: 질의 q 에 대한 ground truth의 객체 수
 $NR(q)$: 검색된 상위 k 안에 ground truth와 일치하는 객체 수
 k : 질의에 대한 결과로서 반환되는 객체의 수

표 4. 스케치에 기반한 질의의 결과

질의	Ground Truth	검색 객체 순위	유사궤적 객체집합
그림1 1(a)	R1-R9-R0-R13-R5-R20-R19-R8-R4-R18-R10-R11-R17-R16-R14-R7-R6-R12-R3-R15-R2	R20-R5-R19-R13-R1-R0-R9-R8-R10-R4-R18-R17-R11-R6-R7-R16-R14-R12-R15-R2-R3	R1-R9-R0-R13-R5-R20-R19
그림1 1(b)	W11-W1-W6-W17-W3-W15-W2-W12-W5-W0-W4-W16-W20-W14-W18-W22-W24-W9-W23-W7-W21-W10-W8-W19-W13	W1-W17-W11-W6-W3-W5-W15-W0-W12-W2-W4-W16-W14-W20-W18-W21-W9-W24-W22-W8-W23-W7-W10-W13-W19	W11-W1-W6-W17-W3-W15-W2-W12-W5-W0-W4

표 5. 객체 예제에 기반한 질의의 결과

질의	Ground Truth	검색 객체 순위	유사궤적 객체집합
그림12(a)	R9-R1-R0-R13-R19-R20-R5-R10-R8-R4-R11-R3-R17-R18-R6-R7-R16-R14-R12-R15-R2	R9-R1-R5-R0-R20-R19-R13-R8-R4-R18-R11-R10-R17-R7-R6-R16-R14-R3-R12-R15-R2	R9-R1-R0-R13-R19-R20-R5
그림12(b)	W1-W11-W6-W17-W2-W5-W15-W12-W3-W0-W4-W18-W19-W21-W14-W9-W24-W23-W10-W8-W13-W7-W16-W22-W20	W1-W3-W5-W11-W6-W17-W0-W12-W15-W2-W4-W16-W14-W18-W20-W21-W24-W9-W22-W23-W8-W7-W10-W13-W19	W1-W11-W6-W17-W2-W5-W15-W12-W3-W0-W4

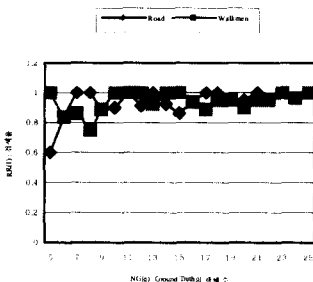


그림 13. 그림 11의 질의에 대한 검색을

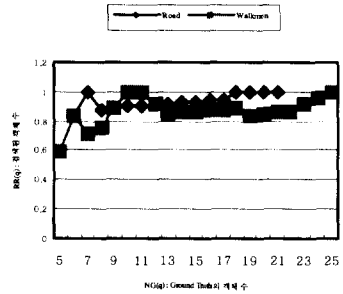


그림 14. 그림 12의 질의에 대한 검색을

2) Retrieval Accuracy Measure-2

평가 방법 1은 질의 궤적과 색인화 된 객체 궤적들 사이의 점 대점 정합에 의한 물리적 거리를 기준으로 한 정합 결과를 나타낸 것으로서, 질의와 유사한 궤적을 갖는 객체의 수와는 상관없이 검색 순위에 따른 검색율을 나타낸다. 평가 방법 2는 식 11과 같이 recall rate와 precision rate을 각각 정의하고, 이를 이용한 실험 결과를 제시한다. 실험 평가 방법 2는 다음 과정과 같이 이루어진다.

첫째, 재구성된 객체 궤적과 질의 궤적간의 점 대점 정합을 통하여 질의에 대한 ground truth를 정한다. 둘째, ground truth의 상위 k 개의 객체들을 유사 궤적을 갖는 객체 집합으로 정의한다.

셋째, 검색 시에는 질의와 저장된 객체 사이의 점 대점 정합이 아닌 매개변수간 정합을 행하고 유사 객체 집합과 비교하여 recall rate와 precision rate을 구한다.

평가 방법 2가 갖는 제약 조건은 질의 궤적과 유사 궤적을 갖는 객체 집합의 크기 k 를 정하는 데 있다. 궤적들간의 단순한 물리적 거리만을 고려해서는 어느 객체들의 궤적이 질의 궤적과 유사한지, 또는 상위 몇 개까지의 객체를 유사 궤적을 갖는 객체 집합으로 정할 것인지의 결정이 주관적이기 때문이다. 실험에서는 ground truth의 순위를 바탕으로 다수의 사람이 질의 궤적과 유사한 궤적을 갖는다고 판단한 순위까지의 객체들로 유사 궤적 객체 집합을 정의하고, 이를 검색된 결과와 비교하여 recall rate와 precision rate를 구하였다. 그림 15과 16은 평가 방법 2를 사용하여 그림 11과 12의 질의 결과를 recall/precision rate 그래프로 나타낸 것이다.

$$recall = \frac{n(R \cap T)}{n(R)}, \quad precision = \frac{n(R \cap T)}{n(T)} \quad (11)$$

R : 유사 객체 집합
 T : 검색된 객체 집합

실험 결과를 살펴보면 검색 순위가 ground truth의 순위와 전후로 뒤바뀐 경우는 존재하지만, 검색 결과 상위 개 내에 위치하는 모든 객체들이 사전에 정의한 유사 궤적 객체 집합에 포함되는 것을 확인할 수 있다. 검색 결과의 순위가 ground truth의 순위와 정확하게 일치하지 않는 것은 ground truth는 점 대점 정합을 통하여 정의되는 반면, 검색 시에는 효율적이고 빠른 정합을 위하여 간단한 매개변수간의 간단한 정합만을 이용하기 때문에 발생하는 결과로 여겨진다.

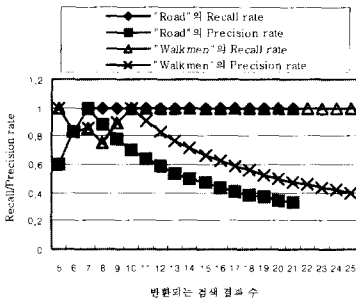


그림 15. 그림 11의 질의에 대한 recall/precision rate

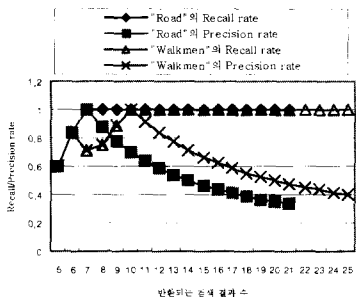


그림 16. 그림 12의 질의에 대한 recall/precision rate

IV. 결론

본 논문에서는 객체의 이동 궤적에 대한 효율적 기술기 및 이를 이용한 내용 기반 비디오 검색 기술을 제안하였다. 객체의 이동 궤적을 여러 개의 시구간으로 나누어 기술함으로써 객체 이동 궤적에 기반한 효율적인 비디오 검색을 가능하게 하였다. 또한, 카메라의 움직임에 의한 배경의 움직임도 기술하여 고정 카메라 환경과 이동 카메라 환경 모두에서 이동 객체 궤적에 기반한 검색이 가능하도록 하였다. 객체 궤적 기반의 다양한 질의 및 정합 함수를 정의하고 이를 검색에 이용함으로써 효율적인 내용 기반 비디오 검색이 가능함을 보였고, 실험을

통하여 제한한 검색 기술의 성능을 확인할 수 있었다. 제한한 기술은 일반 비디오에 대한 신속한 검색 뿐만 아니라 감시 비디오 시스템의 효율적 운영에 기여할 수 있을 것으로 기대된다. 향후, 객체의 컬러, 색조, 모양 등 다른 비디오 특징을 조합적으로 이용한 내용 기반 검색과 네트워크상의 검색이 고려되어야 할 것이다.

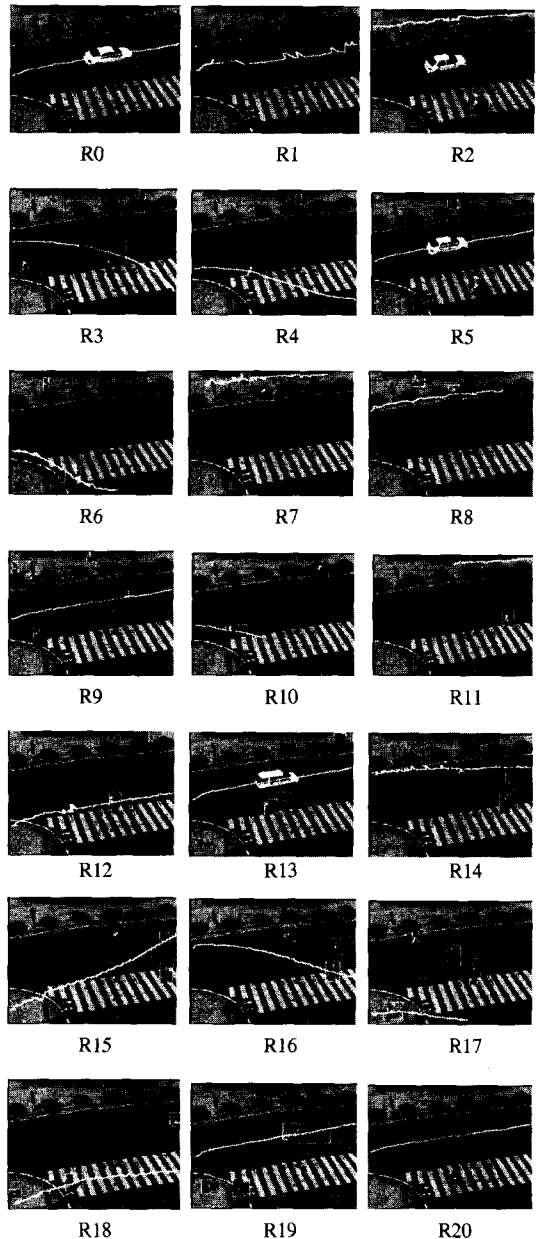


그림 17. 비디오 데이터 "Road"의 비디오 객체 (고정카메라환경)

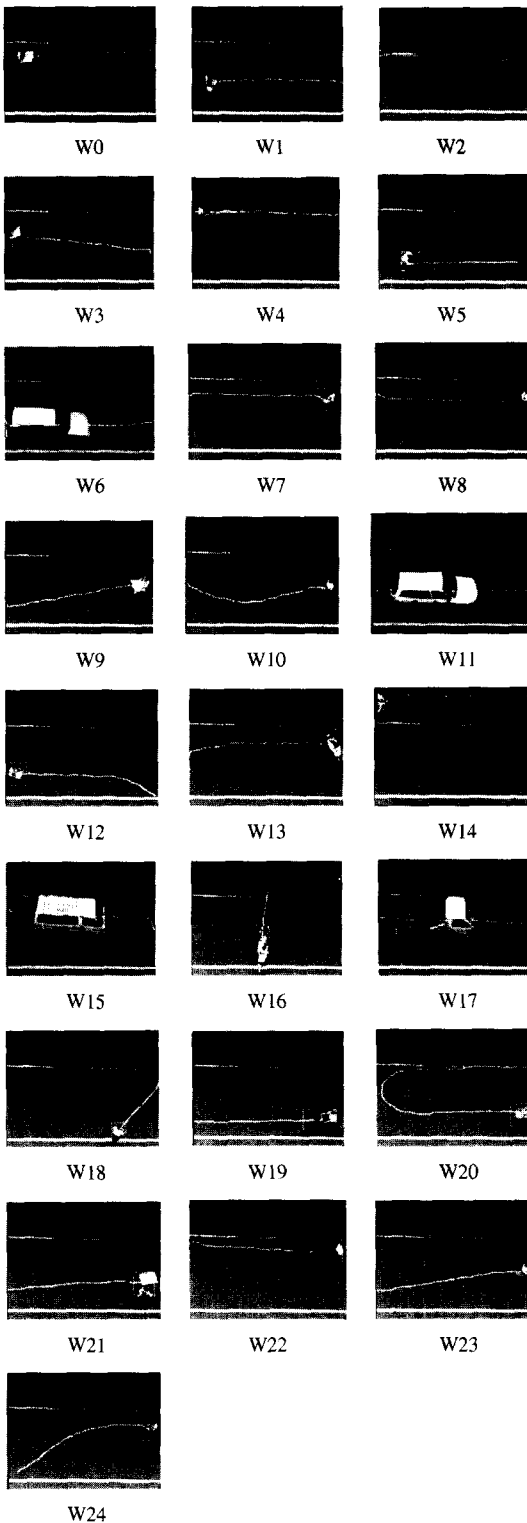


그림 18. 비디오 데이터 “Walkmen”의 비디오 객체 (고정카메라환경)

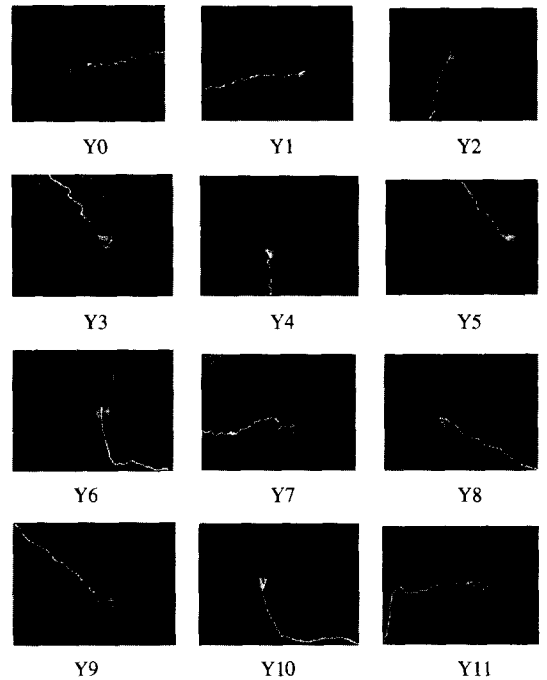


그림 19. 비디오 데이터 “Yard”의 비디오 객체 (이동카메라환경)

참고 문헌

- [1] “MPEG-7: Context and Objectives (version-10 Atlantic City),” *ISO/IECJTC1/SC29/WG11/ MPEG98/N2460*, Atlantic City, USA, Oct. 1998.
- [2] “MPEG-7 Requirements Document V.7,” *ISO/IECJTC1/SC29/WG11/MPEG98/N2461*, Atlantic City, USA, Oct. 1998.
- [3] Kyu Won Lee, Woong-Shik You and Jinwoong Kim, “Hierarchical Object Motion Trajectory Descriptor,” *ISO/IECJTC1/SC29/WG11/MPEG99/M46 81*, Vancouver, July 1999.
- [4] S. Jeannin, “MPEG-7 Visual part of eXperimentation Model Version 1.0,” *ISO/IECJTC1/SC29/WG11/MPEG99/N2695*, Seoul, March 1999.
- [5] Man-Kwan Shan and Suh-Yin Lee, “Content-based Video Retrieval via Motion Trajectories,” *SPIE*, vol. 3561, pp. 52-61, 1998.
- [6] Shih-Fu Chang, William Chen, Horace J. Meng, Hari Sundaram, and Di Zhong, “A Fully Automated Content-Based Video Search Engine Supporting Spatiotemporal Queries,”

- IEEE Trans. on CSVT*, vol. 8, No. 5, pp. 602-615, Sep. 1998.
- [7] K.W. Lee and J.W. Kim, "Moving Object Segmentation based on Statistical Motion Model," *IEE Electronics Letters*, vol. 35, No. 20, pp. 1719-1720, Sep. 1999.
- [8] Jinsuk Kwak, Jun Geun Jeon, Munchurl Kim, Myung Ho Lee, and Chieteuk Ahn, "A Semi-automatic video segmentation method for object oriented multimedia applications," *ITC-CSCC'99*, Niigata, Japan, pp. 114-117, July 13-15, 1999.
- [9] Jae-Gon Kim, Kyu-Won Lee, Jinwoong Kim and Hyung-Myung Kim, "Extraction of Moving Objects from MPEG-Compressed Video for Object-Based Indexing," in *Proc. 1st European Workshop on Content-Based Multimedia Indexing*, Toulouse, France, pp. 131-139, Oct. 1999.
- [10] Jianhao Meng and Shih-Fu Chang, "CVEPS-A Compressed Video Editing and Parsing System," *Proceedings, ACM Multimedia 96 Conference*, Boston, MA, Nov. 1996.
- [11] P. Bouthemy, M. Gelgon, and F. Gannansia, "A Unified Approach to Shot Change Detection and Camera Motion Characterization," *IEEE Trans. on CSVT*, vol. 9, No. 7, pp. 1030-1044, Oct. 1999.

유 웅 식(Woongshik You) 정회원



1997년 8월 : 충남대학교 컴퓨터 공학과 졸업
 2000년 2월 : 충남대학교 컴퓨터 공학과 석사
 2000년 4월~현재 : 한국전자통신연구원 연구원

<주관심 분야> 멀티미디어 검색, MPEG-7, 3DTV

이 규 원(Kyu Won Lee) 정회원

1986년 2월 : 연세대학교 전자공학과 졸업
 1988년 2월 : 연세대학교 전자공학과 석사
 1998년 2월 : 연세대학교 전자공학과 박사
 1988년 2월~1989년 8월 : LG산전(주) 연구원



1989년 9월~2000년 2월 : 한국전자통신연구원 선임연구원
 2000년 3월~현재 : 대전대학교 정보통신공학과 교수
 <주관심 분야> 멀티미디어검색, MPEG-7, 영상처리, Motion Analysis, 영상통신

김 재 곤(Jae-Gon Kim) 정회원



1990년 2월 : 경북대학교 전자공학과 졸업
 1992년 2월 : 한국과학기술원 전기 및 전자공학과 석사
 1992년 3월~현재 : 한국전자통신연구원 선임연구원

<주관심 분야> 비디오 신호처리, 비디오 인덱싱

김 진 웅(Jinwoong Kim) 정회원



1981년 2월 : 서울대학교 전자공학과 졸업
 1983년 2월 : 서울대학교 전자공학과 석사
 1993년 2월 : Texas A&M Univ. Dept. of Electrical Engineering 박사

1983년 3월~현재 : 한국전자통신연구원 책임연구원
 <주관심 분야> 멀티미디어 검색, MPEG-7, 신호처리, 영상통신

권 오 석(Oh-Seok Kwon) 정회원



1977년 2월 : 서울대학교 전자공학과 졸업
 1980년 2월 : 한국과학기술원 산업전자공학과 석사
 1992년 2월~현재 : 한국과학기술원 전산학과 박사과정

1980년~현재 : 충남대학교 컴퓨터공학과 교수
 <주관심 분야> 신경회로망, 패턴인식, 퍼지이론 및 응용