

움직임 객체 모델링 및 검색

복경수^o 김미희 조기형 유재수

충북대학교 정보통신공학과 및 컴퓨터·정보통신연구소

{ksbok^o, mhkim}@netdb.chungbuk.ac.kr, {khjoe, yjs}@cbucc.chungbuk.ac.kr

Modeling and Retrieval of Moving Objects

Kyoung Soo Bok^o, Mi Hee Kim, Ki Hyung Cho, Jae Soo Yoo

Dept. of Computer & Communication, Chungbuk National University

요 약

본 논문에서는 시간의 변화에 따른 공간적 변화를 갖는 움직임 객체를 효과적으로 검색하기 위한 모델링 및 검색 기법을 제안한다. 멀티미디어 데이터에 포함된 움직임 객체는 객체의 움직임 방향뿐만 아니라 객체의 공간적인 크기 변화가 매우 중요하다. 제안하는 모델링에서는 객체의 공간적인 위치를 MBR로 표현하고 MBR의 변화에 따른 객체의 궤적 정보를 표현하여 이를 통한 다양한 검색을 수행한다. 또한 객체가 속한 프레임의 시각적인 특징과 객체의 시각적인 특징 변화를 표현하여 이를 통한 시각적인 유사도 검색을 지원한다.

1. 서 론

컴퓨터 기술의 발달과 더불어 멀티미디어 데이터에 대한 중요성이 증가되면서 이에 대한 기술들이 급속도로 발전하고 있다. 일반적으로 멀티미디어 데이터는 대용량이고 비정형화된 데이터 특성을 포함하고 있다. 특히 멀티미디어 데이터의 대표적인 비디오 데이터는 일반적인 멀티미디어 특성 뿐만 아니라 시각적인 특성(Visual-Audio feature)와 시공간적인 특성(Spatio-Temporal feature)을 포함하고 있다. 비디오의 시공간적 특성을 대표하는 움직임 객체(Moving Object)는 시간의 변화에 따라 공간적인 위치가 변화는 객체이다.

일반적으로 비디오에 나타난 모든 객체는 시각적인 특성과 공간적인 특성을 모두 포함하고 있다. 그러나 움직임 객체는 이러한 특성들이 시간의 변화에 따라 계속 변화되기 때문에 움직임 객체를 효과적으로 검색하기 위해서는 시간의 변화에 따른 변화를 추적할 수 있는 모델링 기법들이 필요하다. 그러나 기존의 움직임 객체 모델링은 움직임 객체를 비디오의 논리적인 구조 내에 표현하거나 객체의 움직임 방향에 집중되어 있었다. 따라서 객체의 크기 변화 또는 다른 객체와의 공간적인 거리 변화 등을 효과적으로 표현하지 못한다는 문제점이 있었다.

본 논문에서는 움직임 객체의 시각적인 특성 변화와 객체의 움직임 방향은 물론 객체의 크기에 따른 객체들 사이의 공간적인 거리 변화를 효과적으로 표현하고 검색할 수 있는 새로운 모델링 기법을 제안한다.

본 논문의 나머지 부분의 구성은 다음과 같다. 먼저 2장에서는 기존에 제안된 움직임 객체 모델링을 기술하고 3장과 4장에서 제안하는 움직임 객체 모델링 및 모델링을 통한 다양한 검색 방법에 대해 기술하고 마지막 5장에서 결론을 기술한다.

2. 관련연구

비디오에 나타나는 객체는 공간적인 위치뿐만 아니라 객체의 크기가 매우 중요하기 때문에 객체의 공간적인 위치를 MBR(Minimum Bounding Rectangle)로 표현한다.

※ 본 연구는 한국과학재단 목적기초연구(특정기초연구 과제번호 R01-1999-00244)지원으로 수행되었음

[1]에서는 비디오를 프레임의 연속으로 보고 프레임에 나타난 객체의 공간적인 위치를 MBR로 표현하였다. 그러나 [1]에서는 시간의 변화에 따른 객체의 위치 변화를 효과적으로 표현하지 못하기 때문에 이를 효과적으로 표현하고 검색하기 위해 [2, 3, 4, 5]와 같은 모델링 기법들이 제안되었다.

[2]에서는 12가지 방향 관계와 움직임 거리를 이용하여 객체의 궤적(Trajectory)를 표현하였고 [3]에서는 비디오의 논리적인 구조 내에 시각적 특징 변화에 의해 객체의 궤적을 표현하였다. [4, 5]에서는 객체의 움직임 방향을 각과 움직임 거리에 의해 표현하고 이를 통한 유사도 검색을 수행한다.

3. 움직임 객체 모델링

제안하는 모델링에서는 움직임 객체를 효과적으로 표현하기 위해 객체의 공간적인 움직임 뿐만 아니라 시간의 변화에 따른 객체의 크기 변화와 시각적인 특징 변화 등을 표현한다. 제안하는 움직임 객체 모델링은 원시 데이터 계층(Raw Data Level), 프레임 계층(Frame Level) 그리고 객체 계층(Object Level)의 세 개의 계층으로 구성된다.

3.1 원시 데이터 계층

원시 데이터 계층(Raw Data Level)은 모델링의 최하위에 존재하는 계층으로 실제 비디오에 대한 가공되지 않은 비디오 자체에 대한 정보를 표현하는 계층이다. 이러한 원시 데이터 계층은 다음과 같이 표현된다.

$\langle VideoID, Production_F, Physical_F \rangle$

VideoID는 비디오의 식별자, Production_F는 제작자, 비디오 이름, 감독과 같은 비디오 제작에 관련된 특성, Physical_F는 압축 방식, 저장방식, 초당 프레임 재생 수, 재생 시간과 같은 비디오의 물리적 특징을 표현한다.

3.2 프레임 계층

프레임 계층(Frame Level)은 비디오에서 특정 시간 간격 Δt 단위로 추출된 키프레임(Keyframe)에 대한 정보를 표현하는 계층이다. 프레임 계층은 프레임 자체에 대한 시각적인 특징은 물론 프레임 내에 포함된 객체들에 대한 정보를 표현한다. 프

레이 계층은 표현은 다음과 같다.

$\langle \text{KeyFrameId}, \text{Object}_L, \text{VisualFeature} \rangle$

*KeyFrameId*은 키 프레임 식별자, *Object_L*는 객체에 대한 특징 리스트, *VisualFeature*는 해당 키프레임 자체의 시각적인 특징을 나타낸다.

*Object_L*는 프레임 내에 존재하는 객체들에 시각적인 특징과 공간적인 특징을 나타낸다. 만약 해당 키프레임 내에 *m*개의 객체를 포함할 때 *Object_L*는 다음과 같이 표현한다.

$\langle (O_1, \text{MBR}_1, C_1, S_1, \text{OF}_1), (O_2, \text{MBR}_2, C_2, S_2, \text{OF}_2), \dots, (O_m, \text{MBR}_m, C_m, S_m, \text{OF}_m) \rangle$

*O_i*는 키프레임에 나타난 객체에 대한 이름, *MBR_i*는 해당 객체 *O_i*의 MBR로 표현된 객체의 공간적인 위치, *C_i*는 객체의 공간적 위치의 중심, *S_i*는 객체의 모양(shape), *OF_i*은 객체 *O_i*의 시각적인 특징을 RGB 값으로 나타낸다.

3.3 객체 계층

객체 계층(Object Level)에는 비디오에 나타난 움직임 객체에 대한 특징을 나타내는 계층으로 프레임 계층에 표현된 객체의 특징 정보를 이용하여 구성한다. 하나의 움직임 객체는 객체의 공간적인 위치를 추적할 수 있는 궤적 정보뿐만 아니라 객체의 시각적 특징에 대한 변화를 추적할 수 있도록 한다. 또한 시간의 변화에 따라 객체가 수행한 동작 정보와 객체에 관련된 사건에 대한 정보를 표현한다. 객체 계층의 표현은 다음과 같다.

$\langle \text{Oid}, \text{Name}, \text{Color}_L, \text{MBR}_L, \text{Camera}_L, \text{Action}_L, \text{Event}_L, \text{Trajectory}, I \rangle$

*Oid*는 객체 식별자, *Name*은 객체의 이름, *Color_L*은 시간의 변화에 따른 객체의 시각적 변화를 나타내는 칼라 리스트, *MBR_L*는 시간 변화에 따른 객체의 공간적인 위치 변화를 나타내는 MBR 리스트, *Camera_L*은 객체에 나타나는 시간 동안 객체에 사용된 카메라 기법에 대한 리스트, *Action_L*은 객체가 연속된 시간 동안 수행한 동작에 대한 리스트, *Event_L*는 객체와 연관된 사건의 리스트, *Trajectory*은 시간의 변화에 따른 객체의 위치 정보를 추적할 수 있는 궤적, $I=[t_s, t_e]$ 는 객체가 나타난 존속 구간(Duration)을 나타낸다. 이때, t_s 는 객체의 등장 시간, t_e 는 객체의 소멸 시간을 나타낸다.

제한하는 모델링에서는 객체의 시간적인 존속 구간을 명확히 판별하기 위해 동일한 이름을 갖는 객체에 대해 연속된 시간에 나타나지 않는 경우 서로 다른 식별자를 부여한다.

객체가 나타나는 존속 구간 *I* 동안 시간의 변화에 따른 객체의 특징 변화를 나타내는 *Color_L*과 *MBR_L*은 프레임 계층에 표현된 객체의 시각적 특징과 공간적인 위치를 이용하여 객체 단위로 구성한 것이다. *Color_L*은 시간의 변화에 따라 객체가 나타나는 시간 구간 *I* 동안 시각적인 특징 변화를 추적할 수 있는 시각적 특징 궤적이다. *Color_L*은 다음과 같이 표현한다.

$\langle (Col_1, I_1), (Col_2, I_2), \dots, (Col_m, I_m) \rangle$

*Col_i*는 칼라의 (*R_i*, *G_i*, *B_i*) 값, *I_i*는 *Col_i*를 갖는 시간 간격을 [*t_{si}*, *t_{ei}*]에 의해 나타낸다.

시간의 변화에 따른 객체의 공간적인 위치 변화를 나타내는 *MBR_L*은 시간의 변화에 따라 객체의 MBR 변화를 나타내며 다음과 같이 표현한다.

$\langle (\text{MBR}_1, C_1, S_1, I_1), (\text{MBR}_2, C_2, S_2, I_2), \dots, (\text{MBR}_m, C_m, S_m, I_m) \rangle$

*MBR_i*은 시간의 변화에 따른 객체의 MBR, *C_i*은 객체의 중심점, *S_i*은 MBR의 모양, *I_i*는 *MBR_i*가 나타나는 시간 간격으로 [*t_{si}*, *t_{ei}*]에 의해 나타낸다.

*Camera_L*는 객체가 나타나는 시간 구간 *I* 동안 객체에 사용된 카메라 기법에 대한 정보를 카메라 동작에 대한 변화를 시간 간격에 의해 다음과 같이 표현한다.

$\langle (a_1, I_1), (a_2, I_2), \dots, (a_m, I_m) \rangle$

*a_i*는 객체에 사용된 카메라 기법, *I_i*는 카메라 기법이 사용된 시간 간격을 나타낸다.

하나의 객체는 연속된 시간의 변화에 따라 다양한 동작을 수행하면서 비디오에 나타난다. 시간의 변화에 따라 객체가 수행하는 객체의 동작에 대한 변화를 표현하기 위해 *Action_L*를 사용한다. 이러한 *Action_L*는 다음과 같이 표현된다.

$\langle (\text{Action}_1, I_1), (\text{Action}_2, I_2), \dots, (\text{Action}_m, I_m) \rangle$

*Action_i*는 특정 시간 간격 동안 객체가 수행한 동작, *I_i*는 해당 동작을 수행한 객체의 시간 간격을 나타낸다.

또한 시간의 변화에 따라 객체에 관련된 사건에 대한 정보를 나타내는 *Event_L*은 *Action_L*과 유사하다.

3.5 움직임 객체 궤적

움직임 객체란 시간에 따라 공간상에서 위치나 모양이 변하는 객체이다. 따라서, 객체가 나타나는 특정 시간 구간 동안 객체에 대한 움직임을 추적할 수 있는 객체의 궤적이 필요하다. 시간 간격이 $\langle I_1, I_2, \dots, I_n \rangle$ 순서로 주어질 있을 때, 하나의 객체의 궤적 정보 θ 는 $\langle M_1, M_2, \dots, M_n \rangle$ 리스트로 다음과 같이 표현한다.

$\langle (R_1, V_1, D_1, I_1), (R_2, V_2, D_2, I_2), \dots, (R_n, V_n, D_n, I_n) \rangle$

*R_i*는 *I_i* 동안 객체가 움직임 방향, *V_i*는 *I_i* 동안 객체가 움직임 속도, *D_i*는 *I_i* 동안 하나의 객체가 움직임 거리, *I_i*는 객체가 움직인 시간 간격을 나타낸다.

객체가 $I_i=[t_{si}, t_{ei}]$ 동안 객체가 움직인 방향을 나타내는 *R_i*는 *t_{si}* 시점의 객체 중심점 $C_{si}=[C_x^{si}, C_y^{si}]$ 를 기준으로 *t_{ei}* 시점의 객체의 중심점 $C_{ei}=[C_x^{ei}, C_y^{ei}]$ 사이의 상대적인 각도로 -180~180도로 표현한다. 또한 객체의 움직임 거리와 속도를 나타내는 *D_i*와 *V_i*는 $I_i=[t_{si}, t_{ei}]$ 동안 하나의 객체가 *R_i* 방향으로 움직임 거리와 속도를 나타낸다.

4. 검색

4.1 시각적 특징 정보에 대한 유사도 검색

4.1.1 객체들 사이의 시각적 특징 유사도

두 객체 *O_v*, *O_j*의 존속 구간이 [*t_{sv}*, *t_{ev}*], [*t_{sj}*, *t_{ej}*]일 때, 특정 시점 *t*에서 객체의 시각적 특징을 (*R_{sv}*, *G_{sv}*, *B_{sv}*), (*R_{sj}*, *G_{sj}*, *B_{sj}*)라 할때, 특정 시점 *t*에서 객체의 시각적 특징 유사도 $SC(O_v, O_j, t)$ 는 (식 1)과 같다.

$$SC(O_v, O_j, t) = 1 - \frac{D_{color}(O_v, O_j, t)}{MaxColor} \quad \text{---- (식 1)}$$

이 때, *MaxColor*는 객체들 사이의 시각적 특징의 최대 차, $D_{color}(O_v, O_j, t)$ 는 (식 2)와 같이 두 객체의 RGB 값의 차이를 나타낸다. 단, *w_{ci}*는 가중치를 나타내며 $w_{c1} + w_{c2} + w_{c3} = 1$ 이다.

$$D_{color}(O_v, O_j, t) = \begin{cases} \sqrt{w_{c1}(R_v - R_j)^2 + w_{c2}(G_v - G_j)^2 + w_{c3}(B_v - B_j)^2}, & t \in [t_{sv}, t_{ev}], [t_{sj}, t_{ej}] \\ MaxColor, & \text{otherwise} \end{cases} \quad \text{---- (식 2)}$$

특정 시점 *t*에서 시각적 특징 유사도 $SC(O_v, O_j, t)$ 를 이용하여 주어진 시간 구간 $I = [t_s, t_e]$ 동안 두 객체 *O_v*, *O_j*의 시각적 특징 유사도 $SC(O_v, O_j, I)$ 는 (식 3)과 같다.

$$SC(O_v, O_j, I) = \frac{\sum_{t=t_s}^{t_e} D_{color}(O_v, O_j, t)}{t_e - t_s} \quad \text{---- (식 3)}$$

4.1.2 프레임 시각적 특징 유사도

프레임 k_i 와 k_j 의 시각적 특징이 $(C_i, T_i, S_i), (C_j, T_j, S_j)$ 일 때, 프레임의 시각적 특징 유사도 $SF(k_i, k_j)$ 는 (식 4)와 같다.

$$SF(k_i, k_j) = 1 - \frac{F_{color}(k_i, k_j)}{MaxF} \quad \text{---- (식 4)}$$

이때, $MaxF$ 는 프레임들 사이의 시각적 특징의 최대 차, $F_{color}(k_i, k_j)$ 는 (식 5)와 같다. 단, w_β 는 가중치를 나타내며 $w_\alpha + w_\beta + w_\gamma = 1$ 이다.

$$F_{color}(k_i, k_j) = \sqrt{w_\alpha (C_i - C_j)^2 + w_\beta (T_i - T_j)^2 + w_\gamma (S_i - S_j)^2} \quad \text{---- (식 5)}$$

4.2 객체의 거리 유사도 검색

객체 O_i, O_j 의 존속 구간이 $[t_{si}, t_{fi}], [t_{sj}, t_{fj}]$ 일 때, 특정 시점 t 에서 객체 O_i, O_j 의 MBR을 $MBR_{xi} = (X_{si}^t, Y_{si}^t), MBR_{xj} = (X_{sj}^t, Y_{sj}^t)$ 라 할 때, 객체들 사이의 거리 유사도 $SD(O_i, O_j, t)$ 는 (식 6)과 같다. 단, $X_{si} = (X_{si}^t, X_{si}^t), Y_{si} = (Y_{si}^t, Y_{si}^t), X_{sj} = (X_{sj}^t, X_{sj}^t), Y_{sj} = (Y_{sj}^t, Y_{sj}^t)$ 이다.

$$SD(O_i, O_j, t) = 1 - \frac{D_{mbr}(O_i, O_j, t)}{MaxDist} \quad \text{---- (식 6)}$$

이때, $MaxDist$ 는 객체들 사이의 최대 거리 차, $D_{mbr}(O_i, O_j, t)$ 는 객체들 사이의 거리를 나타내며 (식 7)과 같이 계산한다.

$$D_{mbr}(O_i, O_j, t) = \begin{cases} \sqrt{X^2 + Y^2}, & t \in [t_{si}, t_{fi}], [t_{sj}, t_{fj}] \\ MaxDist, & otherwise \end{cases} \quad \text{---- (식 7)}$$

$$X = \begin{cases} X_{sj}^t - X_{si}^t, & X_{sj}^t > X_{si}^t \\ X_{si}^t - X_{sj}^t, & X_{si}^t > X_{sj}^t \\ 0, & otherwise \end{cases}$$

$$Y = \begin{cases} Y_{sj}^t - Y_{si}^t, & Y_{sj}^t > Y_{si}^t \\ Y_{si}^t - Y_{sj}^t, & Y_{si}^t > Y_{sj}^t \\ 0, & otherwise \end{cases}$$

특정 시점 t 에서 거리 유사도 $SD(O_i, O_j, t)$ 를 이용하여 주어진 시간 구간 $I = [t_s, t_f]$ 동안 두 객체 O_i, O_j 의 거리 유사도 $SD(O_i, O_j, I)$ 는 (식 8)과 같다.

$$SD(O_i, O_j, I) = \frac{\sum_{t=t_s}^{t_f} SD(O_i, O_j, t)}{t_f - t_s} \quad \text{---- (식 8)}$$

4.3 객체의 궤적 유사도 검색

객체 O_i, O_j 의 존속 구간이 $[t_{si}, t_{fi}], [t_{sj}, t_{fj}]$ 일 때, 두 객체의 궤적 T_i, T_j 가 아래와 같다고 하자. 이때 R_{pq} 는 움직임 방향, V_{pq} 는 움직임 속도, D_{pq} 는 움직임 거리, I_{pq} 는 움직임 시간 구간이다. 또한 $m \leq n$ 이다.

$$T_i = \langle (R_{1i}, V_{1i}, D_{1i}, I_{1i}), (R_{2i}, V_{2i}, D_{2i}, I_{2i}), \dots, (R_{mi}, V_{mi}, D_{mi}, I_{mi}) \rangle$$

$$T_j = \langle (R_{1j}, V_{1j}, D_{1j}, I_{1j}), (R_{2j}, V_{2j}, D_{2j}, I_{2j}), \dots, (R_{mj}, V_{mj}, D_{mj}, I_{mj}) \rangle$$

두 움직임 객체 O_i, O_j 의 궤적 유사도 $ST(O_i, O_j)$ 는 (식 9)와 같다.

$$ST(O_i, O_j) = 1 - \frac{MIN \sum_{p=1}^m RDV_{p+1}(O_i, O_j)}{m} \quad \text{---- (식 9)}$$

$$\forall k, 0 \leq k \leq n - m + 1$$

이때, $RDV_{p,q}(O_i, O_j)$ 는 움직임 방향, 거리, 속도의 궤적 차이를 나타내며 (식 10)과 같다. 또한 w_{13} 는 방향, 거리, 속도 궤적에 부여된 가중치로 $w_{11} + w_{12} + w_{13} = 1$ 이다.

$$RDV_{p,q}(O_i, O_j) = w_{11}RR(R_i^p, R_j^q) + w_{12}RD(D_i^p, D_j^q) + w_{13}RV(V_i^p, V_j^q) \quad \text{---- (식 10)}$$

(식 10)에서 $RR(R_{pi}, R_{qj})$ 는 움직임 방향 차이, $RD(R_{pi}, R_{qj})$ 는 움직임 거리의 차이, $RV(R_{pi}, R_{qj})$ 는 움직임 속도의 차이로 (식 11), (식 12), (식 13)과 같다. 이때, $MaxD$ 와 $MaxV$ 는 객체들 사이의 최대 움직임 거리 차와 속도 차를 나타낸다.

$$RR(R_{pi}, R_{qj}) = \begin{cases} \frac{|R_{pi} - R_{qj}|}{180}, & |R_{pi} - R_{qj}| < 180 \\ 360 - \frac{|R_{pi} - R_{qj}|}{180}, & otherwise \end{cases} \quad \text{---- (식 11)}$$

$$RD(D_{pi}, D_{qj}) = \frac{|D_{pi} - D_{qj}|}{MaxD} \quad \text{---- (식 12)}$$

$$RV(V_{pi}, V_{qj}) = \frac{|V_{pi} - V_{qj}|}{MaxV} \quad \text{---- (식 13)}$$

4.4 가중치 혼합 유사성 검색

제안하는 모델링에서는 움직임 객체의 시각적 특징, 객체들 사이의 거리, 객체들 사이의 궤적을 이용하여 가중치 혼합 유사도 검색을 수행한다. 가중치 혼합 유사도 검색은 (식 14)와 같다. 이때, w_α 는 가중치에 해당하며 $w_\alpha + w_\beta + w_\gamma = 1$ 이다.

$$SS(O_i, O_j) = w_\alpha ST(O_i, O_j) + w_\beta SD(O_i, O_j, I) + w_\gamma SC(O_i, O_j, I) \quad \text{---- (식 14)}$$

5. 결론

본 논문에서는 비디오 데이터에 포함된 움직임 객체를 효과적으로 표현하고 검색하기 위한 새로운 모델링 및 검색 기법을 제안했다. 제안하는 모델링에서는 객체의 움직임 궤적은 물론 객체의 시각적 특징과 객체의 크기 변화에 따른 객체들 사이의 거리를 효과적으로 표현할 수 있다. 또한 이를 효과적으로 검색하기 위한 다양한 유사도 검색 기법을 제안했다.

참고문헌

- [1] Liang, R., Venkatesh, S. and Kieronska, D., "Video indexing with spatial representation", Third Australian and New Zealand Conference on Intelligent Information Systems, 99-104, 1995
- [2] J. Z. Li, M.T. Özsu and D. Szafron, "Modeling of Moving Objects in a Video Database", Proceedings of IEEE International Conference on Multimedia Computing and Systems, Ottawa, Canada, 336-343, 1997
- [3] Zaher Aghbari, Kunihiko Kaneko and Akifumi Makinouchi, "Modeling and Querying Videos by Content Trajectories". ICME2000, July 2000
- [4] Choon-Bo Shim and Jae-Woo Chang, "Spatio-temporal representation and retrieval using moving object's trajectories", ACM Multimedia Workshops 2000, 209-212, 2000
- [5] Mohammad Nabil, Anne H. H. Ngu and John Shepherd, "Modeling and Retrieval of Moving Objects", Multimedia Tools and Applications 13(1), 35-71, 2001