

비디오 영상에서 지역적 움직임 특성을 표현할 수 있는 기술자

김형준, 김희율
한양대학교 전자공학과

A Descriptor for Characteristics of Local Motion in a Video

Hyoung-joon Kim and Whoi-yul Kim
Dept. of Electronics Engineering, Hanyang Univ.
E-mail : khjoon@vision.hanyang.ac.kr

요 약

본 논문에서는 비디오 영상에서 지역적 움직임 특성을 표현할 수 있는 지역적 움직임 활동(motion activity)에 관한 기술자(descriptor)를 제안한다. 제안된 방법은 화면 전체에 대해 지역적으로 높은 움직임 활동 정도를 갖는 영역에 대한 공간적 정보를 기술하고, 카메라 움직임에 무관하게 물체의 움직임 활동을 정확히 표현하기 위해 움직임 벡터의 통계적 특성과 화면 분할을 이용한다. 본 논문에서 제안하는 움직임 활동의 공간적 특성을 이용하면 동영상에서 화면의 일부에서 일어나는 움직임을 이용한 검색이 가능하고, 물체 추적, 감시 시스템에서도 활용이 가능하다. 실험으로 제안한 방법을 이용해서 움직임 활동이 높은 영역의 추출과정을 보이고, 이를 이용한 검색 결과를 보인다.

1. 서론

인터넷과 멀티미디어 기술이 발전함에 따라 멀티미디어 데이터의 수요가 기하 급수적으로 늘고 있다. 이에 따라 멀티미디어 데이터의 효율적인 관리가 요구되고 있으며, MPEG-7에서는 내용기반 멀티미디어 데이터의 검색에 관한 표준을 제정하기 위한 활동이 진행되고 있다[1].

MPEG-7에서는 멀티미디어 데이터를 내용 기반으로

기술하기 위한 여러 기술자들이 있는데 특히 움직임 정보를 이용한 동영상 기술방법 중에서 움직임 활동(motion activity)은 움직임의 격렬함을 기술하는 것이다[2,3,4]. 이는 사람이 동영상을 보면서 느리게 움직이는 장면 혹은 빠르게 움직이는 장면 등으로 구분하는 시각적 특성을 이용해서 스포츠처럼 움직임의 강도가 높은 장면이나 인터뷰 장면처럼 움직임의 강도가 낮은 장면 등으로 구분한다는 특성을 이용한 것이다.

움직임 활동 정도를 정의하기 위해 MPEG-7에서는 움직임 활동 기술자를 intensity, direction, spatial, temporal, 그리고 spatial localization의 5가지 특성으로 세분화한다. Intensity 특성은 움직임 활동 정도의 크기를 나타내고, direction 특성은 움직임 벡터의 대표 방향, spatial 특성은 움직임 활동의 공간적 상태, temporal 특성은 샷을 구성하는 프레임들의 intensity 분포를 나타내고, 마지막으로 spatial localization 특성은 시간에 대해 프레임에서 같은 위치에 놓인 단위 블록들의 움직임 크기의 분포를 나타낸다[3,5].

위와 같은 움직임 활동 정도의 특성을 나타내기 위한 방법들 중에서 공간적(spatial) 특성을 나타내는 방법으로 움직임이 없는 매크로블록의 수를 이용하는 방법이 제안되었다[6]. 이 방법은 움직임 활동 정도의 공간적 특성을 나타내기 위해 프레임의 평균 움직임 크기보다 작은 움직임 크기를 갖는 매크로블록들의 수평적으로 연속된 개수를 이용한다. 그러나 이 방법은 움직임 활동 정도의 위치 정보는 제공할 수 없고, 수평적으

로 연속된 특성만을 이용하기 때문에 실제 움직임이 없는 영역을 정확하게 표현할 수 없다. 또한 움직임 활동이 없는 매크로블록의 판단 기준으로 움직임 크기의 평균을 이용하기 때문에 카메라의 움직임이 있을 때 전체 움직임 크기가 커져서 상대적으로 움직임 크기가 작은 물체에 대한 움직임 활동 정도는 표현할 수 없다.

본 논문에서는 움직임 활동 정도의 지역적 특성을 기술하기 위해 움직임 활동이 높은 특정 블록들의 공간적 정보에 대해서 기술하는 방법을 제안한다. 프레임의 전체 움직임 크기에 대해 2차 모멘트를 구하고 이를 이용해서 움직임 활동이 적은 매크로블록들의 분포를 구한다. 움직임 크기에 대한 2차 모멘트를 이용하기 때문에 카메라 움직임이 있는 경우에도 상대적으로 움직임이 작은 물체에 대한 움직임 활동 정도를 표현할 수 있고, 공간적인 매크로블록의 분포를 기술함으로써 움직임 활동이 높은 영역을 정확히 표현할 수 있다.

본 논문의 구성은 다음과 같다. 2절에서 제안한 방법을 소개하고, 3절에서는 제안한 방법을 이용한 실험 결과를 보이고, 마지막으로 4절에서 결론을 맺는다.

2. Spatial motion activity

프레임의 움직임 활동 정도를 표현하기 위해 프레임 내의 모든 움직임 벡터들의 크기에 대한 2차원 모멘트를 사용한다[2,4]. 본 논문에서는 움직임 활동의 공간적 특성을 기술하기 위해 프레임에 대한 움직임 활동 정도를 기준으로, 이보다 큰 서브블록 내의 움직임 활동 정도를 갖는 서브블록들을 찾고 이 서브블록의 위치와 서브블록 내의 움직임 활동 정도를 기술한다.

(i, j) 위치의 매크로블록에서 움직임 벡터의 각각 수평성분과 수직성분을 $MV_{x,i,j}$, $MV_{y,i,j}$ 이라고 하고, 매크로블록의 수평 개수와 수직 개수를 M , N 이라 할 때, 각 매크로블록의 움직임 벡터의 크기는 다음과 같다.

$$R_{i,j} = \sqrt{MV_{x,i,j}^2 + MV_{y,i,j}^2} \quad (1)$$

프레임 내의 모든 매크로블록에 대해서 움직임 벡터 크기에 대한 1차, 2차 모멘트는

$$R_{avg} = \frac{1}{MN} \sum_{j=0}^{N-1} \sum_{i=0}^{M-1} R_{i,j} \quad (2)$$

$$\sigma_{Global} = \frac{1}{MN} \sum_{j=0}^{N-1} \sum_{i=0}^{M-1} (R_{i,j} - R_{avg})^2 \quad (3)$$

으로 표현되며, 이 때 식(3)에서 구한 움직임 벡터 크기에 대한 2차 모멘트를 전체 화면의 움직임 활동 정도에 대해 지역적으로 움직임이 높은 영역을 선택하는 기준으로 사용한다. 전체 화면의 움직임 활동 정도를 이용

해서 움직임 활동 정도의 지역적 특성을 구하기 위해서 화면을 4개의 서브블록으로 분할하고 분할된 서브블록 내에 존재하는 움직임 벡터에 대해서 다음과 같은 과정으로 각 서브블록의 σ_k ($k=1,2,3,4$) 구한다.

$$R_{avg,k} = \frac{1}{M_k N_k} \sum_{i,j \in Subblock_k} R_{i,j} \quad (4)$$

$$\sigma_k = \frac{1}{M_k N_k} \sum_{i,j \in Subblock_k} (R_{i,j} - R_{avg,k})^2 \quad (5)$$

이 때, M_k , N_k 는 서브블록에서 매크로블록의 수평, 수직 개수이다. 구해진 σ_k 에 대해서 σ_{Frame} 보다 큰 서브블록을 찾고, 그 서브블록에 대해서 서브블록에 포함되는 매크로블록의 개수가 15개 이상을 유지할 때까지 위 과정을 반복한다. 최종적으로 얻어진 서브블록에서의 σ 와 서브블록의 왼쪽 상단에 있는 매크로블록의 프레임 내에서 위치, 서브블록을 이루는 매크로블록의 수평, 수직 개수를 이용해서 지역적인 움직임 활동 정도를 표현한다[그림 1].

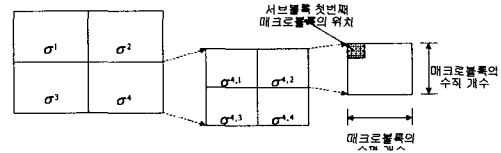


그림 1. 지역적 움직임 활동 정도를 표현하기 위한 화면 분할 (단, $\sigma_1, \sigma_2, \sigma_3 < \sigma_{Global}$, $\sigma_4 \geq \sigma_{Global}$)

$$\sigma_{4,1}, \sigma_{4,3}, \sigma_{4,4} < \sigma_{Global}, \sigma_{4,2} \geq \sigma_{Global}$$

위와 같은 방법으로 σ_{Global} 과 σ_{Global} 보다 높은 σ 를 갖는 서브블록들의 위치와 크기를 이용해서 움직임 활동의 공간적 특성을 정의한다. 이러한 일련의 과정들을 정리하면 그림 2와 같고 각 과정은 다음과 같다.

1. 전체 화면에서 움직임 벡터 크기에 대한 2차 모멘트 σ_{Global} 을 계산한다.
2. 화면을 4개의 서브블록으로 나누고 각 서브블록에서 움직임 벡터 크기에 대한 2차 모멘트 σ_{sub} 를 구하고 σ_{Global} 보다 큰 값을 갖는 서브블록들을 찾는다.
3. 만약 2번 조건에 해당하는 서브블록이 없으면 진행을 멈춘다.
4. 2번 조건을 만족하는 서브블록들에 대해서 2번과 같은 방법으로 다시 4개의 서브블록으로 나누어서 2차 모멘트 $\sigma_{sub,sub}$ 를 구하고 σ_{Global} 과 비교해서 큰 값을 갖는 서브블록들을 찾는다.

- 4번 조건을 만족하는 서브블록이 없으면 2번 과정에서 분할한 현재 서브블록에 대해서 σ_{sub} 와 왼쪽 상단의 매크로블록 위치, 매크로블록의 수평, 수직 개수를 기록하고 2번 조건에 맞는 다른 서브블록들에 대해서 반복한다.
- 4번 조건을 만족하는 서브블록들에 대해서 서브블록의 매크로블록 개수가 15개 이상을 유지할 때까지 다시 서브블록으로 나누고 과정 4, 5를 반복한다.

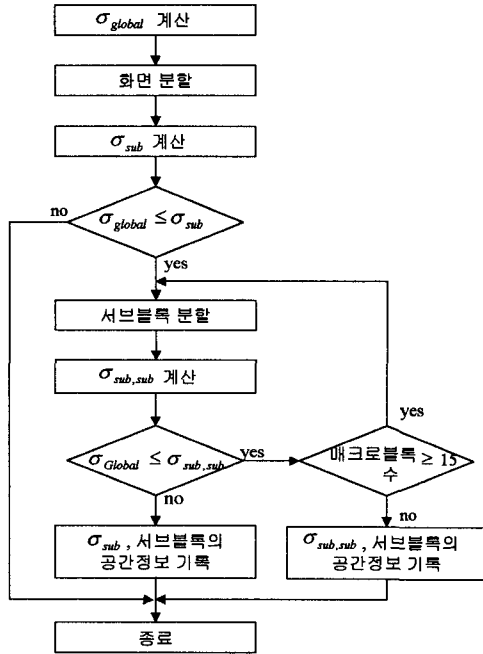


그림 2. 움직임 활동의 공간적 특성을 구하는 블록 다이어그램

3. 실험 및 결과

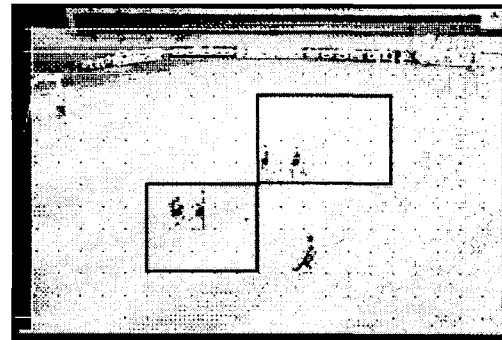
실험은 인간 시각적으로 지역적인 움직임 활동을 잘 표현 할 수 있는지 판단하기 위해 MPEG-7 테스트 셋에서 특정 부분에서 움직임이 있는 비디오를 선택해서 실험하였다.

그림 3은 제안한 방법을 이용해서 움직이는 물체에 의한 지역적 움직임 활동이 높은 영역을 표시한 것으로 (a)는 카메라 움직임이 없는 경우이고 (b)는 카메라 움직임이 있는 경우이다. 움직임이 없는 매크로블록의 수를 세는 기존의 방법보다 제안한 방법이 인간 시각적으로 지역적 움직임 활동 특성을 잘 표현할 수 있음을 알 수 있다.

그림 4는 Query-by-Example 방식으로 제안된 방법을 이용한 검색 결과이다. 여기서 좌측 상단의 영상이



(a) 카메라 움직임이 없는 경우



(b) 카메라 움직임이 있는 경우
그림 3. 지역적 움직임 활동이 높은 영역

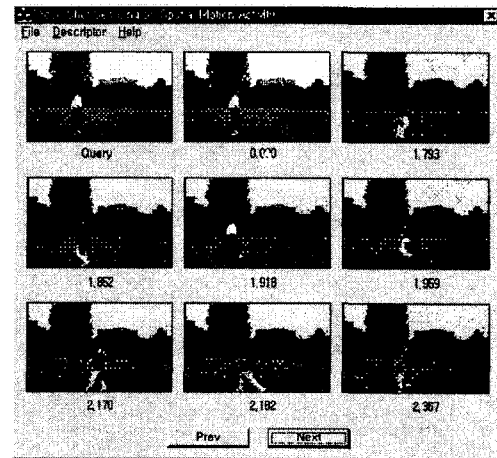


그림 4. 제안한 방법을 이용한 검색 결과

움직임 정보를 갖고 있는 query 영상이고, 영상 하단의 수치는 검색 결과 나온 영상들과 query 영상과의 움직임 활동의 공간적 특성을 이용해서 구한 거리이다. 움직임 활동의 공간적 특성을 이용하기 때문에 비슷한 위치에서 움직임 활동이 있는 장면 검색이 가능함을 알 수 있다.

4. 결론

본 논문에서는 비디오 영상에서 지역적 움직임 특성을 표현할 수 있는 지역적 움직임 활동 기술자를 제안하였다.

제안된 방법은 동영상으로부터 추출된 움직임 벡터의 통계적 특성을 이용하여 화면 전체에 대해 지역적으로 높은 움직임 활동 정도를 갖는 영역에 대한 공간적 특성을 기술함으로써 물체의 움직임으로 인한 지역적 움직임 활동 특성을 기술할 수 있다. 실험을 통해서 제안한 방법이 카메라가 움직이는 경우와 움직이지 않는 경우에 대해 물체에 의한 지역적 움직임 활동성을 잘 기술할 수 있음을 보였고, 이를 이용한 검색 결과를 보였다. 이러한 화면 전체에 대한 지역적 움직임 활동 특성을 이용하면 동영상에서 화면의 일부분을 이용한 검색, 물체 추적, 감시 시스템 등에 효과적으로 활용될 수 있다.

향후 연구 과제로는 효율적인 비디오 내용 기반 기술을 위해 프레임에서 정의된 움직임 활동의 공간적 특성을 샷으로 확장해서 기술하는 연구가 필요하다.

참고 문헌

- [1] M. I. Sezan and R. J. Qian, "MPEG-7 Standardization Activities," Proc. ICIP'1998, Vol. 3, pp. 517-520, Oct. 1998.
- [2] S. Jeannin, L. Cieplinski, J. R. Ohm, and M. C. Kim, "MPEG-7 Visual part of eXperimentation Model Version 6.0," ISO/IEC JTC1/SC29/WG11, N3398, Jun. 2000.
- [3] S. Jeannin, L. Cieplinski, J. R. Ohm, and M. C. Kim, "MPEG-7 Visual part of eXperimentation Model Version 7.0," ISO/IEC JTC1/SC29/WG11, N3521, Jul. 2000.
- [4] 심동규, 정재원, 오대일, 김해광, "비디오 검색을 위한 통계적 움직임 활동 기술자," 방송공학회 논문지, 제5권, 제1호, pp. 2-9, 2000년 6월.
- [5] A. Divakaran, H. Sun, H. Kim, C. S. Park, B. S. Manjunath, X. Sun, V. V. Vinod, G. Rattinasababady, D. Manoranjan, "A tentative merged Activity Descriptor based on the Motion Activity Core Experiment," ISO/IEC JTC1/SC29/WG11, M4811, Jul. 1999.
- [6] A. Divakaran, H. Ito, H. Sun, P. Akella, P. Bouklee, A. Vetro, and T. Poon, "A Bit Allocation Based Descriptor for MPEG-4/2/1 Compressed Video Sequences," ISO/IEC JTC1/SC29/WG11, P002, Jan. 1999.