

논문-00-5-1-01

# 비디오 검색을 위한 통계적 움직임 활동 기술자

심 동 규\*, 정 재 원\*, 오 대 일\*, 김 해 광\*\*

## Statistical Motion Activity Descriptor for Video Retrieval

Dong-Gyu Sim\*, Jae-Won Chung\*, Dae-Il Oh\*, and Hae-Kwang Kim\*\*

### 요 약

본 논문은 동영상으로부터 추출된 움직임 크기와 방향을 이용한 통계적 움직임 활동(Motion Activity) 특징 기술(Description) 방법과 이를 이용한 비디오 검색에 관한 것이다. 본 논문의 움직임 활동 기술자를 이용하여 동영상 전체, 대표 영상 사이, 시간상 특정 구간에 대한 시공간적 분포 및 변화 정도와 패턴 등에 대한 지각적 특징을 기술할 수 있어, 이러한 움직임 정도가 중요한 특징이 되는 동영상 검색(video retrieval), 원격감시(surveillance), 멀티미디어 데이터베이스, 방송 필터링(broadcasting filtering) 등의 디지털 비디오 서비스 응용들에 효과적으로 활용될 수 있다. 제안한 방법의 효율성을 보이기 위하여 MPEG-7 표준화 과정에서 사용된 다양한 비디오를 가지고 검색 결과를 보인다.

### Abstract

This paper presents a statistical motion activity description method and video retrievals by using the intensity and directions of the extracted motion vectors from video sequence. Since the proposed method can represent temporal and spatial cognitive characteristics of an entire video, several images between key frames, and images in a certain interval, it can be effectively applied to digital video services such as video retrieval, surveillance, multimedia database, and broadcasting filterings. In the paper, the effectiveness of the proposed algorithm is shown with a lot of shots of MPEG-7 video dataset.

## I. 서 론

최근들어 문자, 도형, 음성, 음향 그리고 영상 등을 모두 포함하는 표현 미디어와 통신 네트워크, 방송 네트워크와 같은 전달 미디어 그리고 이들 운용하는 시스템 성능이 지속적으로 발전하고 있다. 이러한 이유로해서 소용량 단일 미디어로 이루어진 모노미디어보다 복수의 모노미디어로 구성된 대용량 멀티미디어 데이터가 급속히 증가하고 있다. 멀티미디어 데이터 양의 증가에 비례하여, 데이

터를 사용 및 재사용하기 위하여 검색하는데 소요되는 시간과 비용 또한 증가하게 되었다. 따라서, 보다 빠르고 효율적인 검색을 위하여 현재 널리 이용되고 있는 문자기반 검색기술을 포함하여, 복합된 정보 속성을 갖는 멀티미디어 데이터의 효과적인 검색에 적합한 검색 기술에 대한 연구 및 개발이 활발하게 진행되고 있다.<sup>[1][2][3][4][5]</sup>

본 논문은 멀티미디어 중에서 가장 큰 부분을 차지하고 있는 비디오 검색에 관한 것이다. 기존의 비디오 검색 기법들은 비디오를 색상(Color), 질감(Texture), 모양정보(Shape) 그리고 움직임(Motion) 변화에 바탕을 두어 임의의 크기의 시간적 구간들로 구조화(Video Structuring)하여, 구간내의 영상들의 의미적/신호적 특성을 대표할 수 있는 몇 개의 대표영상(Key-Frame)을 선별하고, 선별된 대표 영

\* 현대전자산업 주식회사  
Hyundai Electronics Industries, Co. Ltd

\*\* 세종대학교 소프트웨어 공학과  
Dept. of Software Engineering, Sejong Univ.

상들의 정보 속성에 대하여 특징정보를 추출하여 색인하여 검색에 사용하였다<sup>[1][2]</sup>. 현재 멀티미디어 데이터 색인 및 검색에서 가장 많이 연구되고 있는 분야는 데이터 획득이 손쉬운 정지 영상 분야이다. 정지영상 색인 및 검색은 내용 기반 영상검색 방법으로서 이들의 주된 관심사항은 영상의 색(Color), 질감(Texture) 그리고 모양(Shape) 정보 등에 대해서 회전, 크기변환, 이동 등의 변화에 관계없이 일관된 특성을 보이는 특징정보 추출과 색인 그리고 이를 활용한 검색방법이다. 비디오 검색분야는 정지 영상에 비해 데이터의 획득이 쉽지 않고, 대용량의 데이터를 저장하고 처리해야 하는 관계로 그 응용이 제한적이었다. 비디오는 연속된 시간에서 영상 데이터를 획득하기 때문에 영상내의 공간적인 용장성(중복성)과 더불어 이웃하는 영상간의 용장성이 매우 커서 영상간 예측이 어느 정도 가능하다. 이러한 용장성에 기인한 비디오에서의 영상간 특성은 정지영상과는 크게 구별되는 것이다. 비디오 검색기법에서 상기한 영상간 유사성은 비디오 특징 정보추출에 있어 중요하게 활용된다.

동영상 기술방법에 대한 기존의 기술로는 문자, 대표정지영상(key frame) 등을 사용하여 왔으나, 이는 동영상의 고유한 특징인 움직임 활동 정도를 효과적으로 기술하지 못하는 단점이 있었다. 동영상에서의 움직임 활동 특징을 기술하는 종래의 기술로는 카메라 움직임이나 영상 객체의 궤적 등의 특징을 사용하는 방법이 있으나, 이는 동영상 화면 전체에

서 나타나는 전체적인 움직임 활동의 특징을 기술하지 못하는 단점이 있다. 또한 기존에 영상에서의 움직임의 크기 자체를 이용하여 움직임의 크기에 대한 성질을 나타낼 수 있었으나 영상의 움직임의 변화량의 개념인 움직임 활동성을 나타내는 특징이 제안되어 있지 않았다. 본 논문은 동영상의 움직임 활동(Motion Activity)을 기술하는 방법에 관한 것으로 움직임 활동 특징을 동영상으로부터 추출된 움직임 파라미터들의 크기와 방향의 통계적 특성으로 기술하는 것을 특징으로 동영상의 움직임 활동 특징 기술 방법에 관한 것이다.

본 논문의 구성은 II장에서 기존의 동영상 검색방법 기법들에 소개하고 III장에서 제안한 방법을 소개한다. 그리고 IV장에서 실험결과를 보이고, 마지막으로 V장에서 각각 실험결과와 결론을 기술한다.

## II. 기존의 움직임 활동자 기술방법

비디오 검색을 위하여 비디오의 특징중의 하나인 움직임 정보에 의한 비디오 검색은 인간의 시각적 인지도 관점에서 중요한 방법 중의 하나이다<sup>[6][7][8]</sup>. 본 논문에서는 움직임 자체를 특징으로하는 검색이라기보다 움직임 활동도에 의한 검색을 다룬다. 움직임 활동도는 움직임의 혼잡성의 정도에 따라 나타난다. 즉 움직임이 크다해도 움직임

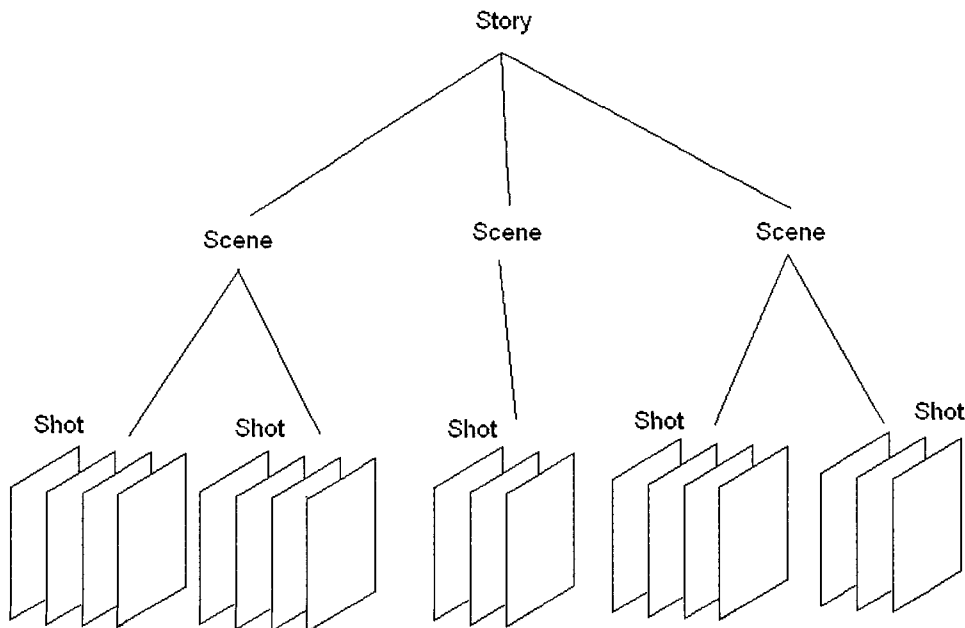


그림 1. 비디오 구조화  
Fig. 1. Video structure

의 변화가 적을 경우 적은 움직임 활동도를 가지며, 움직임의 변화가 심하여 복잡한 움직임을 가지는 비디오는 움직임 활동도가 크다고 할 수 있다. 이러한 움직임 활동도는 몇 개의 영상으로 구성된 비디오에서 정의되며, 이렇게 전체 비디오를 몇 개의 쇼트, 씬 등으로 분리될 수 있으며, 이것에 대한 특징을 기술해 줌으로써 비디오를 구조화하여 표현할 수 있다. 그림 1와 같이 비디오 구조화에 있어 일반적인 구조는 시간적으로 끊이지 않는 정지화상의 연속인 '쇼트(shot)'라는 단위를 기본 구조 단위로, '쇼트'의 시간적인 연속으로서 내용상 시공간적인 연속성을 갖는 단위인 '씬(Scene)' 및 '씬'들로 구성된 '기승전결' 수준의 이야기 전개 단위인 '스토리(story)' 등의 계층적 구조를 갖을 수 있다.<sup>[9][10][11]</sup>

본 논문에서 이러한 구조화된 비디오의 각 쇼트에 대한 움직임 활동 정도를 나타내는 특징의 추출을 목적으로 한다. 지금까지 이러한 목적을 달성하기 위하여 많은 알고리즘이 제안되어 왔다. 압축 비디오를 이용한 방법으로 할당된 비트의 수에 의한 움직임 없는 블록의 수에 의한 움직임 활동자 기술방법이 제안되었다<sup>[12]</sup>. 이 방법은 기존의 MPEG-1/2/4 그리고 H.263의 시간축 연속성에 기인한 블록의 움직임 벡터를 이용한다. 즉 각 매크로 블록의 움직임 벡터를 압축하기 위하여 필요한 비트 수가 임계치 보다 작을 경우 0으로 하는 방법으로 수평축으로 움직임이 없는 블록의 길이를 구하여 이 값을 비디오 검색에 이용하는 방법이다. 이 방법은 압축 평면에서 구현되어 계산량이 적다는 장점이 있는 반면 인간의 움직임 활동 특성과 무관한 압축 사용 비트수를 사용하여 인간 시각 특성과 다를 수 있으며, 움직임이 없는 블록은 수평축으로만 고려하여 실제 움직임이 없는 영역과는 다를 수 있는 단점이 있다. 또한 코딩모드에 바탕을 둔 움직임 양자화  $\alpha$ -히스토그램이 제안되어져 있는데<sup>[13]</sup>, 이것은 움직임이 없는 블록을 세는 방법으로 움직임 활동정도를 표현하는 방법이다. 비디오 쇼트에서 P프레임의 매크로 블록중에 움직임 보정을 하지 않은 블록의 수의 비율 이용하는 방법으로, 이 방법 또한 압축 평면을 이용한다는 점에서 계산량이 적은 장점을 가지고 있으나, 인간의 시각 특성과는 관계없는 압축여부를 가지고 판단하여 인간의 시각 특성과는 다른 특성을 나타낼 수 있는 단점을 가지고 있다. 또한 압축 비율에 따라 코딩 여부가 결정되어 실제 움직임과 다른 특성을 내는 경우도 많아 다른 비율로 압축된 비디오 간의 검색이 어려운 단점을 가지고 있다. 또한 움직임 벡터의 차의 평균크기를 움직임 활동정도로 기술하는 방법이 제안되었다<sup>[14]</sup>. 이 방법은 비디오에서 물체로 분할하거나

직사각형 형태의 블록에서 움직임 벡터에서 평균 움직임 벡터를 뺀 후 이 벡터의 크기의 평균값을 움직임 활동정도로 정의하였다. 이 연구에서 제안한 구현 방법으로 optical flow를 추정하고 이것으로 움직임 활동자를 구하는 방법으로 구성되어 있다. Optical flow를 사용함으로써 압축된 비디오를 사용할 경우 압축을 복원하여 다시 계산량이 많은 움직임 추정 기법을 적용해야 하는 단점이 있다.

### III. 제안한 방법

동영상은 일반적으로 기존의 아날로그 형태로 저장된 영상과 최근에 제안된 디지털에 바탕을 둔 영상으로 나눌 수 있다. 디지털 영상 압축 기술에 바탕을 둔 MPEG이나 H.263 등과 같은 미디어의 경우 움직임 파라미터를 비트열에 저장하고 있어 이를 이용하여 움직임 활동성을 쉽게 기술할 수 있다. 또한 아날로그 영상에서도 어떠한 움직임 추정 방법에 의하여 움직임 파라미터를 추정하고 이것을 이용하여 움직임 활동성을 기술하고 이 기술값에 따라 동영상을 검색할 수 있다. 특히 MPEG이나 H.263과 같은 동영상 압축부호화 방식에서는 영상을 여러 블록이나 객체로 나누어 부호화한다. 하나의 블록을 부호화 할 때, 영상정보의 시간적 중복성을 줄이기 위하여 시간적으로 이웃한 영상에서 부호화할 블록과 가장 유사한 참조 블록을 찾아내어 참조할 블록의 화면상의 상대적 위치를 움직임 파라미터로서 이를 부호화하고, 부호화할 블록과 참조 블록의 차이값을 부호화함으로써 압축효율을 높인다. k번째 화면의 j번째 블록의 움직임 파라미터  $MV_{j,k}=(MV_{x,j,k}, MV_{y,j,k})$ 으로 표현될 수 있다. 여기서  $MV_{x,j,k}$ 는 수평방향의 움직임 성분이고,  $MV_{y,j,k}$ 는 수직방향의 움직임 성분이다. 이 움직임 파라미터의 움직임의 크기와 방향은

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

$$\phi_{j,k} = \tan^{-1} \frac{MV_{y,j,k}}{MV_{x,j,k}} \quad (2)$$

으로 표현된다. 제안한 알고리즘은 움직임 활동 특징을 움직임 파라미터의 크기와 방향에 대한 통계적 특성으로 기술한다. 동영상의 각 화면이 M개의 블록으로 또는 객체로 구성될 때, 각 화면에서 움직임 파라미터의 크기와 방향에 대한 1차 통계값 및 2차 통계값을 추출한다. k번째 화면

에서의 움직임 파라미터 크기의 1차 통계값( $I_{av,k}$ )은

$$I_{av,k} = \frac{1}{M} \sum_{j=1}^M I_{j,k} \quad (3)$$

으로 표현되며, 그 화면에서 움직임 파라미터의 크기의 2차 통계값은

$$I_{dev,k} = \sqrt{\sum_{j=1}^M (I_{j,k} - I_{av,k})^2 / M} \quad (4)$$

으로 정의된다. 같은 방법으로 움직임 파라미터의 방향의 1차 통계값과 2차 통계값은 각각

$$\phi_{av,k} = \frac{1}{M} \sum_{j=1}^M \phi_{j,k} \quad (5)$$

$$\phi_{dev,k} = \sqrt{\sum_{j=1}^M (\phi_{j,k} - \phi_{av,k})^2 / M} \quad (6)$$

으로 정의된다. 위의 특징값은 한 화면내에서의 움직임 파라미터의 통계적 특성으로 동영상에서 움직임 활동자를 정의하기 위해서는 여러 연속연상에 위의 특징값을 추출한 후 그 값들의 통계적 특성을 다시 구하는 방법으로 쇼트에 대한 움직임 활동 특징값을 추출한다.

T개로 구성된 동영상에 움직임 파라미터의 특징값을 추출하기 위한 첫 번째 특징값으로 공간 움직임 크기의 1차 통계값을 T개의 영상에 대한 1차 통계값을

$$I_{av,av} = \frac{1}{M} \sum_{k=1}^T I_{av,k} \quad (7)$$

처럼 구한다. 이렇게 구한 특징값은 시간과 공간 축에서의 움직임 파라미터 크기의 평균값으로 동영상 전체의 평균 움직임을 나타내는 통계적값으로 동영상 검색 및 표현시 대표 특징으로 사용될 수 있다. 두 번째의 특징으로 T개로 구성된 동영상 각각에서 구한 움직임 파라미터 크기의 평균값들의 표준편차로 T개의 영상으로 구성된 동영상을 대표하는 통계적 값이

$$I_{dev,av} = \sqrt{\sum_{k=1}^T (I_{av,k} - I_{av,av})^2 / T} \quad (8)$$

으로 정의된다. 위 식에 의해 구한 특징값은 움직임 파라미터 크기의 시간적 활동성으로 시간적으로 변화하는 움

직임 활동성을 표현하는 통계적 특징값으로 동영상 검색에 유용하기 사용할 수 있다. 또한 이 값은 시간에 따른 움직임 파라미터의 크기의 표준편차로 (식) 7의 1차 통계값의 신뢰도를 표현할 수 있다. 예를 들어 (식) 8의 표준편차값이 큰 것은 (식) 7의 1차 통계값을 중심으로 움직임 파라미터의 크기의 분포가 크다는 것을 의미한다. 즉 이 값이 큰 것은 1차 통계값의 신뢰도가 적다는 것을 의미한다. 이러한 특성을 이용하여 이 값에 의하여 나누어 줌으로써 가중치를 줄 수 있다.

세 번째 특징값으로 T개로 구성된 동영상 각각에서 구한 움직임 파라미터 크기의 표준편차값들의 평균으로 T개의 영상으로 구성된 동영상을 대표하는 통계적 값의 하나가

$$I_{av,dev} = \frac{1}{T} \sum_{k=1}^T I_{dev,k} \quad (9)$$

으로 정의된다. 위식에 의해 구한 특징값은 공간적인 움직임의 활동성을 한 개 이상의 영상에서 평균한 값으로, 공간상의 움직임 활동 정도를 표현하는 통계적 특징값이다. 이러한 특징값으로 공간적인 움직임 활동 정도에 따라 동영상을 검색할 수 있다. 또한 네 번째 특징값으로 T개로 구성된 동영상 각각에서 구한 움직임 파라미터 크기의 표준편차값들의 표준편차로 T개의 영상으로 구성된 동영상을 대표하는 통계적 특징값으로, 식

$$I_{dev,dev} = \sqrt{\sum_{k=1}^T (I_{dev,k} - I_{av,dev})^2 / T} \quad (10)$$

이 사용된다. 이 특징값은 공간적인 움직임 활동 정도와 시간적 움직임 활동 정도를 함께 나타내는 특징값이다.

위의 네 개의 움직임 크기에 대한 특징값과 함께 움직임 벡터의 각도에 대한 특징을 정의한다. 움직임 활동 기술자의 다섯 번째로 T개로 구성된 동영상 각각에서 구한 움직임 파라미터 방향의 평균값들의 평균으로 T개의 영상으로 구성된 동영상을 대표하는 통계적 특징값이

$$\phi_{av,av} = \frac{1}{T} \sum_{k=1}^T \phi_{av,k} \quad (11)$$

으로 정의된다. 본 특징값은 움직임 파라미터의 방향의 공간 및 시간적 1차 통계값으로 한 개 이상의 동영상의 움직임 방향을 나타내는 특징값이다. 또한 여섯 번째 특징값

으로 T개로 구성된 동영상 각각에서 구한 움직임 파라미터 방향의 평균값들의 표준편차로 T개의 영상으로 구성된 동영상을 대표하는 통계적 기술자를

$$\phi_{dev, av} = \sqrt{\sum_{k=1}^T (\phi_{av, k} - \phi_{av, av})^2 / T} \quad (12)$$

으로 정의한다. 위 특징값은 움직임 파라미터값의 방향의 시간적 변화정도, 즉 활동정도를 나타내는 특징값으로 방향의 변화량에 따라 동영상을 검색할 수 있다.

일곱 번째 특징값으로 T개로 구성된 동영상 각각에서 구한 움직임 파라미터 방향의 표준편차값들의 평균으로 T개의 영상으로 구성된 동영상을 대표하는 통계적 기술자를

$$\phi_{av, dev} = \frac{1}{T} \sum_{k=1}^T \phi_{dev, k} \quad (13)$$

으로 정의한다. 정의된 특징값은 움직임 파라미터의 방향의 공간적 2차 통계값의 평균으로 움직임 파라미터의 공간적 활동도를 나타낸다. 이러한 특징값을 이용하여 공간적 방향 변화율에 따라 동영상을 검색할 수 있다. 마지막으로 여덟 번째 특징값으로 T개로 구성된 동영상 각각에서 구한 움직임 파라미터 방향의 표준편차값들의 표준편차로 T개의 영상으로 구성된 동영상을 대표하는 통계적 기술자로 사용한다. 이러한 특징을 위하여

$$\phi_{dev, dev} = \sqrt{\sum_{k=1}^T (\phi_{dev, k} - \phi_{av, dev})^2 / T} \quad (14)$$

을 사용하여 움직임 방향의 공간적, 시간적 활동 정도를 나타낸다. 위와 같은 방법으로 추출된 여덟 개의 움직임 활동 파라미터를 양자화하여 특징값으로 사용할 수 있다.

비디오 인덱싱 및 검색을 위하여 입력 비디오는 쇼트 단위로 나누고 각 쇼트에 대한 움직임 활동 파라미터를 추출한다. 이렇게 저장된 데이터 베이스를 인덱스 파일이라고 하며, 비디오 검색을 위하여 입력 영상의 활동 파라미터와 저장된 인덱스 파일의 특징값을 비교한다. 거리

$$Dis = W_{av, av} \|I_{av, av}^Q - I_{av, av}^I\| + W_{av, dev} \|I_{av, dev}^Q - I_{av, dev}^I\| + W_{dev, av} \|I_{dev, av}^Q - I_{dev, av}^I\| + W_{dev, dev} \|I_{dev, dev}^Q - I_{dev, dev}^I\| \quad (15)$$

를 이용하여 최소 거리를 갖는 쇼트부터 순차적으로 사용자에게 보여준다. 여기서 윗 첨자 Q와 I는 쿼리와 인덱스 파일을 나타낸다. 여기서 각각 특징의 가중치값은 각 특징값의 표준편차의 역수를 사용하여 모든 특징값을 정규화할 수 있다. 그러나 사용자가 각각의 특징값의 의미에 따른 검색을 위해서는 원하는 특징값의 가중치를 크게 설정하여 검색할 수 있을 것이다.

#### IV. 실험결과 및 토의

본 알고리즘의 유용성을 평가하기 위하여 MPEG-7 표준화 과정에서 사용한 MPEG-7 테스트 비디오와 MPEG-7 표준화 과정에서 사용된 평가 기준을 사용하였다. 본 실험에서는 MPEG-7 표준화 비디오에서 637개의 쇼트를 추출하여 데이터 베이스를 구축하였다. 이 비디오셋은 인간에 의하여 다섯 개의 움직임 레벨로 분류하였다. 본 실험에선 제안한 방법의 효율성을 보이기 위하여 기존의 방법으로 Run-length를 이용하는 방법<sup>[12]</sup>, 코딩모드를 고려하는 방법<sup>[13]</sup>, optical flow<sup>[14]</sup>를 이용하는 방법을 사용하였다. 또한 제안한 알고리즘으로 제안한 여러개의 파라미터 중에 Iavg.dev와  $\phi$  avg.dev를 사용하였다. 표 1에서 사용한 데이터 베이스를 보여주고 있다<sup>[16]</sup>. 본 데이터 베이스는 MPEG-7 표준화 과정에서 공동으로 사용한 50개의 비디오 CD 중의 몇 개를 선정하였다. 표에서 CD 번호는 50개의 CD중에 몇 번째 CD인지를 나타낸다. 표 1에서 추출한 비디오 쇼트를 여러 사람에게 움직임의 활동성을 다섯 개의 단계로 분류하라고 주지시킨 후 다섯 개의 움직임 활동레벨로 나누었다. 움직임 활동도란 움직임의 혼잡도로도 말할 수 있으며, 비디오 검색은 인간의 시각특성과 관련 있어 인간에 의한 분류가 주관적이지만 응용의 관점에선 가장 적당한 방법이다. 이렇게 나누어진 비디오 쇼트를 분류를 기준으로 하여 추정된 결과와 비교하는 것이다.

표 1. 실험에 사용한 비디오와 쇼트의 수

Table 1. Video and the number of shots used in experiments

CD Number	Test data stream	No of shots
14	Portugese/journaldanoite1.mpg	30
15	Portugese/journaldanoite2.mpg	30

CD Number	Test data stream	No of shots
17	News1.mpg	89
18	News2.mpg	192
19	Hallo/hallo.mpg	5
19	Singapore/waste.mpg	19
19	Rtve/don_qui.mpg	1
20	Misc1.mpg	11
21	Misc2.mpg	7
23	Riscos-sl	5
24	Docon	3
24	Playboy/lebalad1.mpg	4
25	Singapore/harmony	4
25	Singapore/lebalad2.mpg	4
26	kbs/golf.mpg	30
26	Kbs/basketball.mpg	40
26	Kbs/show.mpg	30
26	Nhk/nhkvideo.mpg	42
27	Ophthalmic/eyeexam	3
28	Lancaster/lanc.mpg	38
28	Samsung	5
29	Bigshowsic	5
30	Singapore/animals.mpg	6
30	Ucl/speedwa4.mpg	2
30	Ucl/speedwa5.mpg	4
30	Etri/ETRI_od_A.mpg	3
30	Etri/ETRI_od_B.mpg	7
30	Etri/ETRI_od_C.mpg	4

표 2에서는 각 활동도에 대하여 분류된 쇼트의 수를 보여주고 있다. 여기서 activity level 1은 움직임이 거의 없는 영상을 나타내며, activity level 5는 움직임의 변화가 급격함을 나타냅니다.

표 2. 각 활동도에 대한 분류된 shot의 수  
Table 2. Numbers of classified shots according to the activity

Activity level	Number of shots
1	114
2	198
3	180
4	103
5	28

표 3. 각 활동도에 대한 제안한 알고리즘과 기존의 알고리즘의 추정결과 비교

Table 3. Comparison of the estimated results of the conventional and proposed in terms of activity level

algorithm	Activity Level					Total Hits
	1	2	3	4	5	
Run-length	105/111 (94.6%)	56/194 (28.9)	23/179 (12.8%)	11/102 (10.8%)	4/28 (14.2%)	199 (32.4%)
Coding mode	0/111 (0%)	50/194 (25.8%)	90/179 (50.3%)	1/102 (9.8%)	0/28 (0%)	141 (23.0%)
Optical flow	78/111 (70.3%)	137/194 (70.6%)	68/179 (38.0%)	21/102 (20.6%)	3/28 (10.7%)	307 (50.0%)
Proposed	72/111 (64.9%)	94/194 (48.5%)	63/179 (35.2%)	47/102 (46.1%)	11/28 (39.3%)	287 (46.7%)

이렇게 분류된 비디오 쇼트를 이용하여 Run-length, 코딩 모드, optical flow 방법과 제안한 방법의 적중률을 표 3에 나타내었다. 표에서 보듯이 optical flow 방법이 전체 적중률에서 가장 좋은 결과를 보였고, 제안한 방법이 그에 근접하는 결과를 보였다. 그러나 활동도가 가장 큰 부분에서는 제안한 알고리즘이 크게 좋은 성능을 나타내 평균적으로 좋은 결과를 보였다. 그러나 optical flow를 이용하기 때문에 계산량이 과다한 단점을 가지고 있다.

표 4에서는 모든 비디오 클립에 대하여 추정된 활동도와 인간에 의하여 추정된 기준 활동도 사이의 차의 합을 평가 기준으로 한 비교결과를 보여준다. 표 4에서 보듯이 제안한 알고리즘이 기존의 방법에 비하여 기준 활동도와 가장 유사한 결과를 볼 수 있다. 즉 optical flow 방법은 적중률은 약간 높으나 적중하지 않을 경우 큰 오차를 나타냄을 알 수 있다. 이에 비하여 제안한 알고리즘은 모든 움직임 레벨에 대하여 안정적으로 좋은 추정결과를 볼 수 있다. 이러한 결과를 바탕으로 MPEG-7 표준화에서 움직임

표 4. 실제값과 추정된 활동도의 차의 합  
Table 4. Sum of the absolute difference of the estimated and ground truth

algorithm	Sum of the absolute difference between the ground truth and actual value
Run-length	574
Coding mode	838
Optical flow	394
Proposed	391

임 활동 레벨 기술자 파라미터 추출 방법으로 제안한 알고리즘이 추천되었다.

## V. 결론

본 논문은 압축 동영상에서 움직임 벡터의 통계적인 특성인 1차와 2차 모멘트를 비디오의 공간과 시간축에서 구함으로써 움직임 활동도를 추정하는 논문이다. 본 논문은 압축 평면에서 움직임 활동자를 추출함으로써 압축영상에 대하여 매우 적은 계산으로 움직임 활동자를 추출할 수 있고, 움직임 벡터의 통계적인 특성을 이용함으로써 인간의 시각에 보다 가까운 특징을 추출할 수 있었다. 본 실험을 통하여 기존의 압축코드와 같은 인간의 시각적 특성과 관계없는 특징에 비하여 좋은 성능을 나타냄을 볼 수 있었다. 앞으로 이러한 연구 결과를 바탕으로 움직임 활동자를 이용한 비디오 검색을 이용한 응용 예를 연구할 수 있을 것이다.

## 참고 문헌

- [1] J. M. Corridoni and A. Del Bimbo, "Structured Digital Video Indexing," *Proceedings of ICPR'96*, pp. 125-129, Aug. 1996.
- [2] M. Flicker, H. Sawhney, W. Niblack, J. Ashley, Qian Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, P. Yanker, "Query by Image and Video Content: The QBIC System," *Computer*, vol. 27, pp. 23-32, Sep. 1995.
- [3] C. Y. Low, Q. Tian, H. J. Zhang, "A Automatic News Video Parsing, Indexing, and Browing System," *ACM Multimedia '96, Boston*, Nov. 1996.
- [4] P. Nesi and A. DelBimbo, "A Vision System for Estimating People Flow," in *Image Technology*, pp. 179-201, Springer-Verlag, 1996.
- [5] M. J. Swain and D. H. Ballard, "Color indexing," *International Journal of Computer Vision*, vol. 7, pp. 11-32, Nov. 1991.
- [6] V. V. Vinod and H. Murase, "Video Shot Analysis Using Efficient Multiple Object Tracking," *Proceedings of IEEE Conf. on Multimedia Computing Systems*, pp. 501-508, Jun. 1997.
- [7] M. Yeung and B. L. Yeo, "Video Content Characterization and Compaction for Digital Library Applications," *Proc. of SPIE Storage and Retrieval of Image and Video Databases*, vol. 3022, pp. 45-58, Jan. 1997.
- [8] H. Z. Zhang, A. Kankanhalli, and S. W. Smoliar, "Automatic Partitioning of Full Motion Video," *ACM Multimedia systems*, pp. 10-23, 1993.
- [9] S. F. Chang, "Compressed-Domain Techniques for Image/Video Indexing and Manipulation," *IEEE Trans. Conf. on Image Processing*, vol. I, pp. 314-317, Washington, D.C., Oct. 1995.
- [10] A. Divakaran, H. Ito, H. Sun, and T. Poon, "Scene Change Detection and Feature Extraction for MPEG-4 Sequences," *Proceedings SPIE conference on Storage and Retrieval for Image and Video Databases*, San Jose, Jan. 1999.
- [11] D. Zhong and S. F. Chang, "Spatio-Temporal Video Search Using The Object Based Video Representation," *IEEE International Conf. on Image Processing*, vol. I, pp. 211-24, Santa Barbara, California, Oct. 1997.
- [12] 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, Lancaster, UK*, Jan. 1999.
- [13] X. Sun, B. S. Manjunath, P. Wu, Y. Deng, H. D. Shin, and Y. L. Choi, "Motion Quantized  $\alpha$ -Histogram as a Video Unit Descriptor," *ISO/IEC JTC1/SC29/WG11, P075, Lancaster, UK*, Jan. 1999.
- [14] D. Manoranjan and V. V. Vinod, "Video Segment Activity," *ISO/IEC JTC1/SC29/WG11, P627, Lancaster, UK*, Jan. 1999.
- [15] H. K. Kim and C. S. Park, "Temporally Aggregated Motion Histogram Descriptor," *ISO/IEC JTC1/SC29/WG11, P515, Lancaster, UK*, Jan. 1999.
- [16] A. Divakaran, H. Sun, H. K. Kim, C. S. Park, B. S. Manjunath, X. Sun, H. D. Shin, V. V. Vinod, G. Rattinassababady, and D. Manoranjan, "Report on the MPEG-7 Core Experiment on the Motion Activity Feature," *JTC1/SC29/WG11, m5030, Melbourne, AustralISO/IECia*, Oct. 1999.

저 자 소 개



심 동 규

1993년 : 서강대학교 전자공학과 졸업 (학사)  
 1995년 : 서강대학교 전자공학과 대학원 졸업 (석사)  
 1999년 : 서강대학교 전자공학과 대학원 졸업 (박사)  
 1999년~현재 : 현대전자산업 주식회사, 단말기 연구소, 선임연구원 근무 중  
 주관심분야: 영상처리, 컴퓨터비전, 영상코딩, MPEG



정 재 원

1989년 : 한양대학교 공과대학 전자통신공학과 졸업 (학사)  
 1991년 : 한국과학기술원 전기 및 전자공학과 졸업 (석사)  
 1998년 : 한국과학기술원 전기 및 전자공학과 졸업 (박사)  
 현재 : 현대전자산업주식회사 단말기 연구소, 책임연구원  
 주관심분야: 영상압축부호화 및 MPEG/JPEG 표준 방식, 멀티미디어 VLSI

오 대 일

1982년 : 서강대학교 전자공학과 졸업 (학사)  
 1991년 : 서강대학교 전자공학과 대학원 졸업 (석사)  
 1997년 : 서강대학교 전자공학과 대학원 졸업 (박사)  
 1984년~2000년 5월 : 현대전자산업(주) 정보통신연구소 수석연구원  
 2000년 5월 ~ : (주)오픈솔루션 연구소장  
 주관심분야: 광대역모뎀 VLSI, VLSI 구조 설계

김 해 광

1986년: 한양대학교 공과대학교 전자공학과 (학사)  
 1991년: 삼성전자 종합연구소 주임연구원  
 1997년: 프랑스 Paul-Sabatier대학 (전산학 D.E.A, 박사)  
 1997년~2000년: 현대전자 정보통신연구소 책임연구원  
 2000년~현재: 세종대학교 소프트웨어공학과 조교수  
 주관심분야: 멀티미디어 정보검색, 멀티미디어 콘텐츠 정보보호, 이동통신 프로토콜