DBpia

논문 99-24-9B-15 한국통신학회논문지 '99-9 Vol.24 No.9B

압축영역에서의 대표프레임 추출 및 영역분할기반 비디오 검색 기법

정회원 강 응 관*, 김 성 주*, 송 호 근**, 최 종 수*

Key Frame Extraction and Region Segmentation-based Video Retrieval in Compressed Domain

Eung-kwan Kang*, Sung-joo Kim*, Ho-keun Song**, Jong-soo Choi* ***Regular Members***

요 약

본 논문에서는 동영상 압축 부호화에 대한 표준안인 MPEG 기반의 압축 비디오 시퀀스로부터 DCT DC 계수를 추출하고, 이들로 구성된 DC 이미지로부터 제안하는 AHIM (Accumulative Histogram Intersection Measure)을 이용하여 장면 전환 검출을 수행한 후 대표 프레임을 추출하는 방법을 제시한다. 또한, 추출된 대표 프레임을 두 단계를 거쳐 데이터베이스의 색인 정보로 저장한 후, 입력되는 질의 영상에 대해 사용자가 원하는 검색 결과를 제시하는 방법에 대해 제안한다. 즉, 전처리 과정으로 추출된 대표 프레임에 대해 영역 분할을 한 후, 첫번째 단계에서 수평 투영된 결과를 히스토그램 분포 특성으로 변환시켜 데이터베이스의 색인 정보로 저장한다. 두번째 단계에서는 영상의 모멘트 특성을 거리함수 값으로 변환시킨다. 실험 결과 제안된 방법이 검색에 있어 우수한 성능을 갖추고 또한 상당한 양의 처리 시간과 메모리 공간을 줄일 수 있음을 확인하였다. 향후 제안한 방법은 색상과 같은 다른 색인 정보와 결합할 경우, 보다 나은 영상 색인과 검색 수단을 제공할 것이다.

ABSTRACT

This paper presents a new key frame extraction technique, for scene change detection, using the proposed AHIM (Accumulative Histogram Intersection Measure) from the DC image constructed by DCT DC coefficients in the compressed video sequence that is video compression standard such as MPEG. For fast content-based browsing and video retrieval in a video database, we also provide a novel coarse-to-fine video indexing scheme. In the extracted key frame, we perform the region segmentation as a preprocessing. First, the segmented image is projected with the horizontal direction, then we transform the result into a histogram, which is saved as a database index. In the second step, we calculate the moments and change them into a distance value. From the simulation results, the proposed method clearly shows the validity and superiority in respect of computation time and memory space, and that in conjunction with other techniques for indexing, such as color, can provide a powerful framework for image indexing and retrieval.

I. 서 론

최근의 압축 기술과 통신 기술의 급격한 발달에 따라 디지털 비디오 정보의 활용이 폭발적으로 증가하고 있으며 이를 효과적으로 관리하고 검색하기 위한 방법이 많이 연구되고 있다. 특히 영상 자체의 내용을 기반으로 시간축 분할 (temporal segmentation)한 후, 각 분할된 영역을 대표 프레임 (key frame)

* 중앙대학교 전자공학과 영상정보연구실(jschoi@candy.ee.cau.ac.kr), ** 한서대학교 컴퓨터과학과(hksong@gaya.hanseo.ac.kr)
논문번호 : 99066-0218, 접수일자 : 1999년 2월 18일
※ 본 연구는 정보통신연구진흥원 우수시범학교 지원사업에 의해 수행되었습니다.

www.dbpia.co.kr

등으로 표현함으로써, 전체 비디오 시퀀스를 효과적으로 표현하고 이를 바탕으로 색인 및 검색에 이용하는 기술이 활발히 연구되고 있다.

비디오 시퀀스에서 셧 (shot)은 필름이 끊기지 않고 시간적으로나 공간적으로 연속해서 촬영된 일정 구간의 영상을 의미한다. 비디오 분할 기술은 비디오 제작시 카메라의 기록과 멈춤에 기인한 장면 전환 효과에 의해 발생하는 셧 사이의 경계를 검출하는 방법으로, 셧은 비디오의 내용 해석을 위한 비디오 분할의 가장 기본적인 단위로 사용된다.

장면 전환의 종류로는 셧과 셧의 연결이 급격한 장면 전환 (abrupt scene cut)과 점진적인 장면 전환 (gradual scene cut)이 있으며, 데이터 유형에 따라 비 압축된 비디오 시퀀스에서의 장면 전환 검출과 압축된 영역에서의 장면 전환 검출로 나누어 볼 수 있는데, 특히 MPEG 비디오 시퀀스는 그 유연성과 범용성 및 높은 압축율로 인해 영상관련 응용 분야에서 폭 넓게 이용되고 있으며[1], 따라서 MPEG 압축 영역에서의 장면 전환 검출 또한 많이 연구되고 있다[2][3][4]. [4]에서는 FD(frame difference) 혹은 HD(histogram difference)와 함께 Plateau를 제안하여 사용하였는데, 이 방법은 급격한 장면 전환 검출시 영상 내 움직임에 특히 민감하여 오검출율 (false detection rate) 및 미검출율 (miss detection rate)이 높다. 또한 점진적인 장면 전환 검출 방법으로 사용한 Plateau는 미리 모델링된 이상적인 장면 전환 검출에 최적화되어 있어 영상 내 움직임이 어느 정도 있을 경우 검출이 어려운 단점이 있다. [3]의 알고리듬 역시 영상 내 급격한 움직임이나 조명 변화에 민감한 성능을 보이고 있다. 본 논문에서는 MPEG 표준에 의한 비디오 시퀀스에서 일반적인 큰 움직임 및 조명 변화에 대해 적응적으로 장면 전환을 검출하는 알고리듬과 함께 각 셧을 대표하는 대표 프레임 추출에 대한 방법을 제안하고, 이를 실제 시뮬레이션 프로그램으로 제작하여 그 성능을 평가한다. 또한 추출된 대표 프레임들을 영상 내용에 기반한 특징을 추출하여 데이터베이스의 색인 정보로 저장한 후, 입력되는 질의 (query) 영상에 대해 원하는 유사 영상을 제시하는 영상 검색 기법에 대해 제안한다.

논문의 구성은 다음과 같다. 먼저 2장에서 제안하는 장면 전환 검출 및 대표 프레임 추출 알고리듬에 대해 기술하고, 3장에서는 추출된 대표 프레임에 대한 단계별 내용기반 색인 및 검색 기법에 대해 설명한 다음, 단계별 수행에 대한 타당성 및 효용성에 대해 살펴본다. 4장에서는 제안한 방식을 적용한 실험 결과 및 분석 내용을 검토하고, 5장에서 결론 및 향후 연구 방향에 대해 언급한다.

Ⅱ. 장면 전환 검출 및 대표 프레임 추출 알고리듬

1. MPEG 비디오 시퀀스로부터 DC 이미지 추출

MPEG 표준에 의한 비디오 시퀀스는 I-, P-, B-type의 픽쳐 (picture)로 이루어져 있는데, 이들로부터 간축된 움직임 보상에 의해 DC 이미지를 추출하는 알고리듬이 제안된 바 있다[4][5][6].

본 논문에서는 DC 이미지 추출 알고리듬으로, I-type 픽쳐에 대해서는 DCT DC 계수에 의한 DC 이미지의 구성과 P-, B-type의 픽쳐에 대해서는 움직임 보상에 의해 DC 이미지를 구성하였다. 그림 1은 I-픽쳐에 대해 DC 이미지를 구성한 예를 나타낸 것이다.

(a) (b)

그림 1 MPEG 비디오 시퀀스로부터 구성된 DC 이미지 예 (a) 원영상 (I-픽쳐), (b) DC이미지

2. 장면 전환 검출을 위한 파라미터

먼저 이웃 프레임 간의 DC 이미지에 대해 다음의 파라미터들을 계산한다.

$$FD_i = \frac{1}{MN}\sum_{x=0,y=0}^{x<M,y<N} |I_i(x,y) - I_{i-1}(x,y)| \quad (1)$$

$$HD_i = \sum_{k=0}^{k<q} |H_i(k) - H_{i-1}(k)| \quad (2)$$

$$X_i^2 = \sum_{k=0}^{k<q} \frac{(H_{i-1}(k) - H_i(k))^2}{H_i(k)} \quad (3),$$

, 여기에서 M, N은 각각 DC 이미지의 가로, 세로

www.dbpia.co.kr

크기이며, FD_i는 DC 이미지의 프레임 차분 (frame difference), HD_i는 DC 이미지의 히스토그램 차분 (histogram difference)[7], X_i^2는 DC 이미지의 히스토그램에 대한 chi-square이다. 한편, DC 이미지의 히스토그램은 원영상의 히스토그램에 비하여 잡음 성질이 추가되어 있기 때문에 스무딩 필터 (smoothing filter)를 거쳐 평활화한 후 사용하였다[8]. 또한 이들 프로세스를 수행함과 더불어 아래의 파라미터를 i-1번째 프레임까지 계산한다.

avr_FD1 : 셧 경계 (shot boundary)가 아닌 프레임들의 이전 프레임과의 frame difference의 평균.

avr_FD2 : 셧 경계인 프레임들의 이전 프레임과의 frame difference의 평균.

avr_HD1 : 셧 경계가 아닌 프레임들의 이전 프레임과의 histogram difference의 평균.

avr_HD2 : 셧 경계인 프레임들의 이전 프레임과의 histogram difference의 평균.

avr_χ1 : 셧 경계가 아닌 프레임들의 이전 프레임과의 X^2의 평균.

avr_χ2 : 셧 경계인 프레임들의 이전 프레임과의 X^2의 평균.

위 파라미터들의 초기값은 해당 비디오 시퀀스에서 초기의 몇 프레임에 의해 미리 추정된 값을 갖는다.

3. 장면 전환 검출 및 대표 프레임 추출 알고리듬

위에서 계산된 파라미터들로부터 다음의 과정을 거쳐 장면 전환을 검출하게 된다.

STEP 1: $FD_i > \alpha\, avr_FD1$, *and*
$HD_i > \alpha\, avr_HD1$, *and*
$\chi_i^2 > \alpha\, avr_\chi 1$, *and*
shot duration > *minimum shot duration*

STEP 2: $FD_i > \beta\, avr_FD2$, *and*
$HD_i > \beta\, avr_HD2$

STEP 3: $FD_i > \gamma\, FD_{i-1}$, *and*
$FD_i > \gamma\, FD_{i+1}$, *and*
$HD_i > \theta\, avr_HD1$

STEP 4: $\chi_i^2 > 2\gamma\, \chi_{i-1}^2$ *and* $\chi_i^2 > 2\gamma\, \chi_{i+1}^2$

STEP 5: $HD_i > \gamma\, HD_{i-1}$ *and* $HD_i > \gamma\, HD_{i+1}$

STEP 6: $HD_i > \delta\, avr_HD1$

, 여기서 $\alpha, \beta, \gamma, \delta, \theta$는 실험에 의해 얻어진 최적화된 상수를 나타낸다.

한편, 위 과정에서 STEP 1이 만족해야만 STEP 2-6의 과정을 수행하며, STEP 2에서 STEP 6까지의 과정 중 한 조건 이상을 만족하면 i번째 프레임을 장면 전환으로 검출한다. 그리고, 위 판별 과정에 의해 장면 전환으로 검출된 프레임들 가운데 flashlight과 같은 급격한 조명 변화의 영향에 의한 셧 경계인 경우를 배제하기 위해, $FD_{i+m}(m = 1,2,3,4,5)$이 FD_i에 대하여 충분히 큰 값을 가지며 i-1번째 DC 이미지와 i+m+1번째 DC 이미지의 프레임간 차가 충분히 작은 경우 flashlight 혹은 아주 급격한 조명 변화에 의한 셧으로 판별하며, 이로부터 검출된 flashlight 셧은 장면 전환에서 제외한다. 한편 STEP 5의 조건식만 고려할 경우, 즉 히스토그램이 일반적으로 움직임에 덜 민감한 반면 유사한 분위기나 배경을 갖는 셧 경계는 잘 찾아내지 못하기 때문에 본 논문에서와 같이 frame difference와 chi-square의 조합을 이용한 장면 전환 검출을 고안하였다. 또한 셧 경계는 아니지만 STEP 1을 만족하는 프레임에 대해 본 논문에서는 *AHIM* (Accumulated Histogram Intersection Measure)을 구하여 대표 프레임으로 선택할지를 결정하는 방법을 제안한다. 제안하는 *AHIM*의 정의식은 다음과 같다.

$$\overline{H_i}(j) = \min(\overline{H_{i-1}}(j), H_i(j)),$$
$$AHIM_i = \frac{1}{MN}\sum_{j=0}^{B-1}\overline{H_i}(j) \tag{4}$$

, 여기에서 $\overline{H_i}(j)$는 i번째 프레임의 DC 이미지 히스토그램의 j번째 빈 (bin) 값, B는 빈의 총 개수를 나타낸다.

위 식 (4)에 의해 시간적인 히스토그램의 유사도 변화를 측정함으로써 현재 프레임이 이전 프레임들에 대해 얼마나 많은 변화를 가지는지를 효과적으로 판별할 수 있으며, 또한 혹 검출되지 못한 중요한 영상 내용의 변화를 대표 프레임으로 반영하도록 한다. 한편, 원영상이 아닌 축소된 형태의 DC이미지 적용에 의한 잡음 성분을 무시할 수 없으므로 본 논문에서는 DC 이미지의 히스토그램에 저역통과 필터 (Low Pass Filter)를 거쳐 식 (4)에 적용하였다[8]. 또한 불필요하게 $\overline{H_i}(j)$의 레벨이 떨어지는 것을 막기

www.dbpia.co.kr

위하여 HD_i가 일정 크기 이상인 경우에만 그 이전 계산된 값과의 $AHIM$을 구하도록 하여 가능한 한 정확한 영상의 내용 변화를 추적할 수 있게 하였다. 그리고 셧 내에서의 영상 내용의 변화를 측정할 수 있도록 하기 위해 새로운 장면 전환이 나타나거나 대표 프레임이 검출되었을 경우 (4)식의 $\overline{H_i(j)}$를 초기화한다. 이렇게 함으로써 셧 내의 유사도가 일정 수준 이하로 떨어짐을 검출할 수 있다. 만약 셧 내에서 위의 방법에 의한 대표 프레임이 출현하지 않을 경우 셧 내의 내용 변화가 별로 없는 것이므로 셧의 중간 프레임을 대표 프레임으로 선택한다. 또한 한 셧이 여러 개의 대표 프레임을 갖는 것을 허용하였다.

4. Interactive Video Browser의 구현

그림 2는 장면 전환 검출과 대표 프레임 추출을 이용한 비디오 브라우저를 구현하여 실행시킨 화면이다. 본 프로그램은 MPEG 비디오 스트림을 재생할 수 있을 뿐 아니라 되감기, 빨리 감기, 역 재생, 일시 정지, 빠른 재생 등의 사용자 환경을 제공한다. 또한 제안된 알고리듬에 의해 장면 전환과 대표 프레임을 검출하여 그 결과를 사용자가 확인할 수 있으며, 비디오 시퀀스를 랜덤하게 혹은 셧 단위로 볼 수 있을 뿐만 아니라, 대표 프레임들만 검색할 수도 있게 함으로써 전체 내용을 쉽고 빠르게 찾을 수 있는 기능을 포함한다.

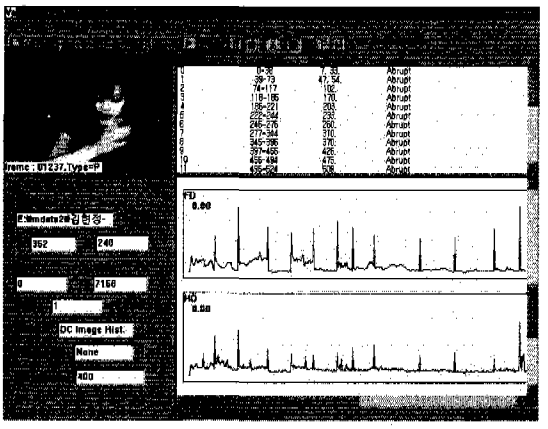


그림 2. 비디오 브라우저의 실행 화면

그림 2에서 네 개의 각 서브 윈도우에는 왼쪽 상단부터 시계방향으로 각각 디스플레이 윈도우, 셧 검색 윈도우, 시뮬레이션 그래프 윈도우, 컨트롤 패널 윈도우를 나타낸다. 이는 각각 MPEG 스트림의 디스플레이 화면을 볼 수 있는 윈도우, 장면 전환 검출 결과와 추출된 대표 프레임들을 확인할 수 있도록 한 윈도우, 알고리듬에 의해 계산되는 각종 파라미터들을 그래프로 볼 수 있는 윈도우, MPEG 파일 불러오기나 시작 프레임과 끝 프레임을 사용자가 지정 또는 확인할 수 있고 장면 전환 검출 알고리듬을 선택할 수 있도록 하는 등의 컨트롤 패널로 구성되어 있다. 또한 툴 박스에는 MPEG 시퀀스의 디스플레이와 관련된 버튼과 함께 장면 전환 검출을 디스플레이 하는 동안 함께 수행할 지, 또는 디스플레이 없이 최대한 빠른 속도로 계산만 수행할 지를 선택할 수 있게 되어 있으며, 슬라이드 바를 이용하여 랜덤하게 영상을 억세스할 수 있도록 되어 있다.

Ⅲ. 내용기반 대표 프레임 색인 및 검색 기법

현재까지 주로 연구되고 있는 내용기반 영상 색인 및 검색 기법으로 영상의 색상 (color), 형태 (shape), 질감 (texture) 등의 속성을 이용한 방식들이 제안되었다[9][10]. 그러나 색상 히스토그램을 사용할 경우 영상의 전체적인 색상 분포만으로 유사성을 계산하기 때문에 영상 내 객체들이 갖는 위치 정보는 전혀 고려하지 않는다는 단점을 갖고 있다. 또한 질감 정보를 사용한 검색 방법은 일반 자연 영상에 적용되기 보다는 영상 전체가 몇 개의 두드러진 질감 특징으로 구분되는 영상을 대상으로 적용된다는 제한성이 있다. 본 논문에서는 앞 장에서 기술한 장면 전환 검출 알고리듬에 의해 추출된 대표 프레임에 대한 영상 색인 및 검색 기법으로 영역 분할 (region segmentation) 기반 투영 히스토그램 및 모멘트 특성을 이용하는 방식에 대해 제안한다. 즉, 전처리 과정으로 추출된 대표 프레임에 대해 영역 분할을 한 후, 첫번째 단계로 수평 투영시킨 결과를 히스토그램 분포 특성으로 변환시켜 데이터베이스의 색인 정보로 저장한다. 다음 두번째 단계로 영상의 모멘트 특성을 거리함수 값으로 변환시킨다. 이에 대한 단계별 세부 사항을 기술하면 다음과 같다.

먼저 전처리 과정에 해당하는 영역 분할은 대상 이미지에서 한 물체에 속한 화소와 그 나머지의 화소를 구분하는 기준을 사용하여 유사한 특성을 가진 화소의 집단을 만드는 것으로, 대표적인 방법에 분할 및 합병 (split-and-merge) 방식이 있다. 이 때 중요한 고려 사항으로 동질성, 즉 공간적 인접성과 명암의 유사성 기준에 근거하여 동일 영역으로 합병 (merging)을 할 것인지 혹은 또 다른 영역으로 분할

www.dbpia.co.kr

(split)을 할 것인지를 결정한다. 본 논문에서 영역 분할을 수행하는 근본적인 이유는 화소 단위가 아닌 유사한 그룹인 영역 단위로 표현함으로써 영상의 대략적인 정보를 취득하기 위한 전처리 과정으로 사용하기 위함이다. 그림 3은 입력 영상에 대해 분할 및 합병 방식을 이용하여 영역 분할한 예를 나타낸 것이다.

(a) (b)

그림 3. 분할 및 합병 방식을 이용한 영역분할의 예 (a) 원영상 (b) 영역분할된 영상

다음으로 앞서 기술한 전처리 과정을 수행한 후 영상 색인 및 검색을 위한 첫번째 단계로 투영에 대한 이론적 배경 및 효용성을 고찰하면 다음과 같다. 원형 (original)의 고차원 데이터를 낮은 차원의 부공간으로 투영시키는 기법은 특히 영상 분석에 있어 유용한데, 즉 $n \times n$의 2차원 영상을 n개의 1차원 데이터 요소로 투영시킴으로써 상당한 양의 처리 시간과 메모리 공간을 줄일 수 있다. 특히, 2차원 좌표계에 있어서 그레이레벨 변화의 수평 투영 (X-projection)은 수평 방향으로 투영된 그레이레벨 변화의 개수로 정의되며, 영상 내부의 구조적인 특성을 반영시키는데 많이 이용된다[11]. 따라서 첫번째 단계인 수평 투영으로부터 취득한 히스토그램 분포 특성, 즉 영상의 대략적인 특징 정보를 데이터베이스의 색인 정보로 저장하고 또한 이를 활용함으로써 1차적인 영상 검색에 이용할 수 있다. 그림 4는 그림 3의 영상을 수평 투영시켜 히스토그램 분포로 나타낸 것이다.

한편, 히스토그램 분포에 대한 유사도를 비교하는 척도로 히스토그램 교차 함수 (histogram intersection measure) 방식[12]을 이용하였고, 이를 정규화한 수식은 다음과 같이 정의된다.

$$H(I,Q)=\frac{\sum_{i=1}^{n}\min(I_i,Q_i)}{\sum_{i=1}^{n}Q_i} \tag{5}$$

, 여기서, I_i, Q_i, 그리고 n은 데이터베이스 영상의 히스토그램, 질의 영상의 히스토그램, 그리고 히스토그램 빈 (bin)의 개수를 각각 나타낸다.

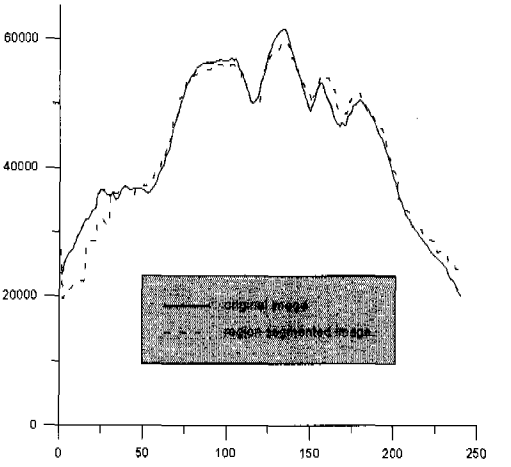


그림 4. 수평 투영 히스토그램

두번째 단계로 영상의 세분류를 위한 색인 및 검색에 사용된 특징 정보인 모멘트에 대한 특성을 기술하면 다음과 같다. 모멘트 특성은 데이터의 확률론적 기술 방법중의 하나로 군집 분포에 대한 평균 (mean), 표준편차 (standard deviation), 비대칭도 (skewness), 첨도 (kurtosis)와 같은 물리적인 특성을 각각 1차, 2차, 3차, 4차 모멘트로 정의하며, 각 모멘트에 대한 수식은 다음과 같다[13].

1차 모멘트: $\bar{x}=\frac{1}{N}\sum_{i=1}^{N}x_i$ (6)

2차 모멘트: $\sigma(x_1\cdots x_N)=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^2}$ (7)

3차 모멘트: $Skew(x_1\cdots x_N)=\frac{1}{N}\sum_{i=1}^{N}\left[\frac{(x_i-\bar{x})}{\sigma}\right]^3$ (8)

4차 모멘트: $Kurt(x_1\cdots x_N)=\left\{\frac{1}{N}\sum_{i=1}^{N}\left[\frac{(x_i-\bar{x})}{\sigma}\right]^4\right\}-3$ (9)

한편, 본 논문에서는 영상의 모멘트 특성을 비교하기 위한 유사도 평가 함수로 모멘트 값을 거리 값으로 변환시키기 위한 수식을 아래와 같이 정의하였다.

$$S_m^{(I,Q)}=\left|w_m^I\begin{bmatrix}1\text{차모멘트}\\2\text{차모멘트}\\3\text{차모멘트}\\4\text{차모멘트}\end{bmatrix}-w_m^Q\begin{bmatrix}1\text{차모멘트}\\2\text{차모멘트}\\3\text{차모멘트}\\4\text{차모멘트}\end{bmatrix}\right| \tag{10}$$

www.dbpia.co.kr

여기서 w_m^I, w_m^Q은 데이터베이스의 후보 영상 및 질의 영상에 대한 가중 매트릭스를 나타내며, 기본값으로 [1 1 1 1]을 갖는다. 또한 $S_m^{(I,Q)}$의 값은 후보 영상과 질의 영상 사이의 유사도를 나타내며, 값이 작을수록 두 영상간의 유사도는 높은 것으로 판정한다.

Ⅳ. 실험 및 고찰

본 장에서는 제안한 대표 프레임 추출 및 영상 검색 기법의 성능을 평가하기 위해, 테스트 영상으로 352×240 크기를 갖는 7157프레임의 뮤직 비디오 영상과 667프레임의 영화 영상에 대해 256 그레이 레벨 해상도의 휘도 성분을 갖고 실험한 결과에 대해 기술한다.

영화 영상의 경우 비교적 급격한 장면 전환이 많이 나타나는 전반부와 물체의 폭발 장면과 더불어 조명의 변화가 많으며 액션이 강한 후반부로 나뉜다. 다음으로 뮤직 비디오 영상은 전반부에서는 계속해서 panning 및 tracking이 일어나며, close-up된 카메라 동작을 보여주는 부분이 있으며, 후반부에는 flashlight 셧에 이은 급격한 장면 전환이 6번 반복되는 부분이 있다. 또한 전반적으로 많은 카메라 움직임을 포함하고 있으며, 물체의 움직임은 보통 정도이다. 2번의 fade-out/fade-in에 의한 장면 전환이 일어나며, 5회의 zoom-in이 있으며, panning이 영상 전반에 걸쳐 사용되었다.

본 논문에서 제안한 알고리듬에 의한 실험 결과, 영화 영상의 경우 급격한 조명 변화와 함께 일어난 장면 전환을 flashlight 셧으로 오판하여 놓친 것이 있으나, 그 외의 조명 변화나 많은 움직임에서의 장면 전환을 모두 검출해냈다. 뮤직 비디오의 경우, 급격한 장면 전환은 거의 모두 검출할 수 있었으며, flashlight 셧과 동시에 장면 전환이 6번 일어나는 부분에서도 2장에서 기술한 STEP 1의 조건과 flashlight 셧의 검출 조건에 의해 모두 장면 전환으로 검출할 수 있었다. 반면 두 개의 fade-out/in은 대표 프레임으로만 검출되었으며, 움직임이 많은 가운데 일어난 장면 전환도 일부 놓쳤으나 대표 프레임으로 대부분 흡수되었다. 그리고 셧으로 잘못 검출된 내용은 대부분 카메라의 움직임과 관련이 있다. 제안된 알고리듬에 사용된 *AHIM*, *FD*, *HD* 등은 계산량에 있어 기존 알고리듬과 비슷한 반면 비슷한 셧 간의 장면 전환 검출 및 조명 변화, 영상 내 움직임 등에 보다 강건함을 알 수 있다.

그림 5는 테스트 영상에 대해 제안하는 *AHIM*의 분포 특성을 나타낸 것이다. STEP 1을 만족하고 HD_i가 일정 수준 이상이면, 식(4)에 의한 *AHIM*을 계산하며 이 값이 최대값에 대하여 50% 이하로 떨어질 경우 대표 프레임으로 검출하였다.

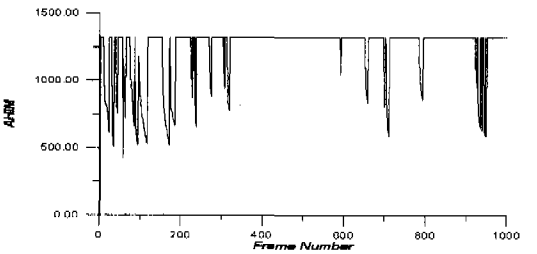


그림 5. 테스트 영상에 대한 AHIM

그림 6은 위의 *AHIM*에 의해 추출된 대표 프레임들의 예이다. 또한 표 1은 테스트 영상에 대해 장면 전환 검출 및 대표 프레임 추출을 수행한 통계 결과를 기술한 것이다.

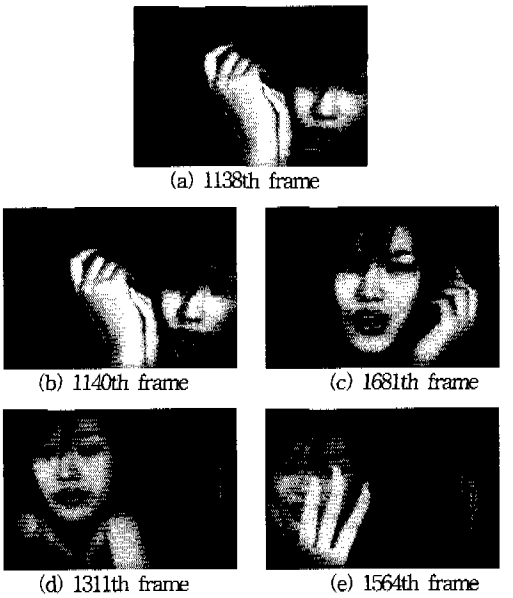
그림 6. 추출된 대표 프레임 예

한편, 그림 7은 제안한 영역분할기반 투영 히스토그램 및 모멘트 방식을 이용한 검색 결과를 가장 유사도가 높은 순서대로 나타낸 것이다. 그림 7에서 보듯이 임의의 질의 영상이 주어졌을 때 유사도가 높은 순으로 검색된 대표 프레임들을 사용자에게 제공함으로써 비디오 전체 내용을 검색하지 않더라도 원하는 셧을 찾아낼 수 있으므로 검색 효율 측면에서도 뛰어나다는 것을 알 수 있다. 그 밖의 실험에서도 검색 결과의 우수성을 확인할 수 있었고, 이상의 실험에 근거해 볼 때 제안한 방법은 효용성 및 검색에 있어 타당하고 우수한 성능을 갖는다는 것을 알 수 있다.

www.dbpia.co.kr

표 1. 장면 전환 검출 및 대표 프레임 추출 결과

Test sequence	Total shots	Correctly detected shots	Falsely detected shots	Missed shots	Extracted key frames	Shot detection ratio (%)
Movie (667 frames)	17	16	0	1	23	94
Music video (7157 frames)	113	107	4	6	128	95



(a) 1138th frame

(b) 1140th frame (c) 1681th frame

(d) 1311th frame (e) 1564th frame

그림 7. 제안한 방법에 의한 검색결과의 예 (a) 질의영상 (b)~(e) 검색된 대표프레임

V. 결 론

본 논문에서는 MPEG 비디오 시퀀스로부터 DCT DC계수를 추출하고, 이로 구성된 DC영상에 대해 *AHIM*함수를 이용하여 비디오 장면 전환 및 대표 프레임을 추출하는 알고리듬을 제안하고, 이를 수행하기 위한 브라우저를 구현하였다. 구현된 브라우저는 우수한 장면 전환 검출 성능을 갖고 있음을 모의 실험 결과로부터 확인할 수 있었다. 또한 영역분할에 기반한 투영 히스토그램 및 모멘트를 이용한 영상 검색 기법을 제시하였다. 모의 실험 결과 제안된 방법은 검색에 있어 우수한 성능을 보이며, 또한 상당한 양의 처리 시간과 메모리 공간을 줄일 수 있는 장점이 있다. 향후 제안한 방법은 색상과 같은 다른 색인 정보와 결합할 경우 보다 우수한 영상 색인과 검색 수단을 제공할 것이다.

참 고 문 헌

[1] 대우전자 영상연구소, *MPEG 비디오*, 연암출판사, 1995.

[2] Ahmed K. Elmagarmid et al.., *Video Database Systems: Issues, Products, and Applications,* Kluwer Academic Publishers, 1997.

[3] Borko Furht, Stephen W. Smoliar, and Hong Jiang Zhang, *Video and Image Processing in Multimedia System,* Kluwer Academic Publishers, 1995.

[4] B. -L. Yeo and B. Liu, "Rapid scene analysis on compressed video," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 5, no. 6, pp. 533-544, 1995.

[5] K.R. Rao and J.J. Hwang, *Techniques & Standards for Image, Video & Audio Coding,* Prentice-Hall PTR, 1996.

[6] Y. Nakajima, K. Ujihara, and A. Yoneyama, "Universal scene change detection on MPEG-coded data domain," in *Proc. SPIE Visual Comm. and Image Proc.*, pp. 992-1003, 1997.

[7] Omer N. Gerek and Yucel Altunbasak, "Key Frame Selection from MPEG Video Data," in *Proc. SPIE visual Comm. and Image Proc.* pp.920-925, 1997.

[8] Nilesh V. Patel and Ishwar K. Sethi, "Video Shot Detection and Characterization for Video Databases," *Pattern Recognition Special Issue :Image Databases*, vol. 30, no. 4, pp. 583-592, 1997.

[9] A. K. Jain and A. Vailaya, "Image retrieval using color and shape," *Pattern Recognition*, vol. 29, no. 8, pp. 1233-1244, 1996.

[10] G. L. Gimelfarb and A. K. Jain, "On retrieving textured images from an image database," *Pattern Recognition*, vol. 29, no. 9, pp. 1461-1483, 1996.

[11] J. K. Wu, C. P. Lam, B. M. Methere, Y. J. Gao, and A. Desai Narasimhalu, "Content-based retrieval for trademark registration," *Int.*

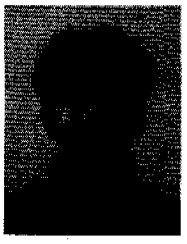
www.dbpia.co.kr

J. Multimedia Tools and Applications, vol.3, no.3, pp.245-267, 1996.

[12] M. J. Swain and D. H. Ballard, "Color indexing," *Int. J. Computer Vision*, vol. 7, no. 1, pp. 11-32, 1991.

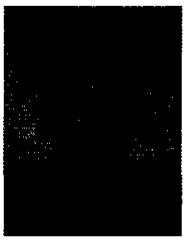
[13] William H. Press et al., *Numerical Recipes in C : the art of scientific computing*, 2nd Ed., Cambridge University Press, 1992.

강 응 관(Eung-Kwan Kang) 정회원



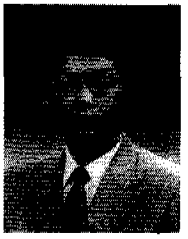
1993년 2월 : 중앙대학교 전자공학과 졸업(공학사)
1995년 2월 : 중앙대학교 대학원 전자공학과 졸업 (공학석사)
1995년 1월~1997년 5월 : 현대전자 정보통신연구소 연구원
1997년 9월~현재 : 중앙대학교 대학원 전자공학과 박사과정 재학 중
<주관심 분야> 동영상 부호화, 멀티미디어 데이터베이스, 영상통신, 컴퓨터비젼

김 성 주(Sung-Joo Kim) 준회원



1998년 2월 : 중앙대학교 전자공학과 졸업(공학사)
1998년 3월~현재 : 중앙대학교 대학원 전자공학과 석사과정 재학 중
<주관심 분야> 동영상 부호화, 멀티미디어 데이터베이스, 영상통신

송 호 근(Ho-Keun Song) 정회원



1991년 2월 : 중앙대학교 전자공학과 졸업(공학사)
1993년 2월 : 중앙대학교 대학원 전자공학과 졸업 (공학석사)
1997년 2월 : 중앙대학교 대학원 전자공학과 졸업(공학박사)
1996년 3월~현재 : 한서대학교 컴퓨터과학과 조교수
<주관심 분야> 패턴인식, 멀티미디어신호처리, 내용기반 영상검색 에지전트 시스템, 전자상거래를 위한 웹 호스팅 등임

최 종 수(Jong-Soo Choi) 정회원

1975년 2월 : 인하대학교 전기공학과 졸업(공학사)
1977년 2월 : 서울대학교 대학원 전자공학과 졸업(공학석사)
1981년 2월 : 일본 Keio대학 전기공학과 졸업(공학박사)
1981년 2월~8월 : Aloka(주) 연구소 연구원
1981년 9월~현재 : 중앙대학교 전자공학과 교수
<주관심 분야> 적외선 신호처리, 컴퓨터비젼, 영상압축 및 영상처리 등임

www.dbpia.co.kr