

MPEG 비디오 시퀀스에서 비디오 요약을 위한 장면 전환 검출 및 대표 프레임 추출 알고리즘

강 응 관[†]

요 약

장면 전환 검출 (scene change detection)은 영상 정보의 색인 (indexing) 및 검색을 위한 전처리로서 비디오 검색 시스템의 전체 성능을 좌우하는 중요한 기술로 현재 많은 연구가 진행되고 있다. 본 논문에서는 급격한 움직임 및 조명 변화, 빠르게 연속적으로 등장하는 장면 전환에 강건한 급격한 장면 전환 검출 기법과 디졸브, 페이드 등의 장면 전환을 보다 효과적으로 검출할 수 있는 점진적인 장면 전환 검출 기법 및 추출된 섷 (shot)의 내용을 DCT DC 이미지 버퍼링과 누적 히스토그램 교차 함수 (accumulative histogram intersection measure)에 의해 대표 프레임으로 추출하는 새로운 알고리즘을 제안하고, 모의 실험 결과 기존 방식에 비해 보다 우수한 성능을 보임을 입증하였다.

Scene Change Detection and Representative Frame Extraction Algorithm for Video Abstract on MPEG Video Sequence

Eung Kwan Kang[†]

ABSTRACT

Scene change detection algorithm, which is very important preprocessing technique for video indexing and retrieval and determines the performance of video database system, is being studied widely. In this paper, we propose a more effective abrupt scene change detection, which is robust to large motion, sudden change of light and successive abrupt shot transitions rapidly. And we also propose a new gradual scene change detection algorithm, which can detect dissolve, and fade in/out precisely. Furthermore, we also propose a representative frame extraction algorithm which performs content-based video summary by novel DCT DC image buffering technique and accumulative histogram intersection measure (AHIM).

Key words: 비디오 요약, 장면 전환 검출, 비디오 색인 및 검색, 비디오 브라우저

1. 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[1-3]. 특히 영상 내 내용을 기반으로 시간축 분할 (temporal segmentation)을 수행한 후, 각 분할된 영역을 대표 프레임

(representative frame) 혹은 키 프레임 (key frame) 등으로 표현함으로써 전체 비디오 시퀀스를 효과적으로 요약하고, 이를 바탕으로 색인 및 검색에 이용하는 기술이 활발히 대두되고 있다.

비디오 분할 기술은 장면 전환 효과에 의해 발생하는 섷 사이의 경계를 검출하는 방법으로, 섷은 비디오 분할의 기본 단위로 사용된다. 장면 전환의 종류로는 급격한 장면 전환 (abrupt scene change)과 점진적인 장면 전환 (gradual scene change)이 있으며, 비압축 영상에서의 장면 전환 검출과 MPEG 및 Motion-JPEG과 같은 압축 비디오 시퀀스에서의 장

† 접수일: 2002년 10월 5일, 완료일: 2003년 3월 3일

정희원, 전주대학교 정보기술컴퓨터공학부 멀티미디어 공학과 조교수

면 전환 검출이 있다[3-5].

한편, 뉴스, 영화, 뮤직 비디오, CF 등과 같은 영상물에서 비교적 빈번하게 나타나는 급격한 조명 변화(flashlight) 및 이와 유사한 것으로 10프레임 이내의 짧은 구간내에서 발생하는 연속적인 장면 전환에 대해 기존 알고리즘[6,7]은 효과적으로 대처하지 못하며, 이와 같은 영상에 대해 높은 오검출율(false detection ratio)을 보인다. 또한 기존의 MPEG 압축 스트림에서의 점진적인 장면 전환 검출 기법은 영상내 움직임에 민감한 단점을 가진다. 물론 [6]에서는 DCT 계수와 영상 밝기와의 관계를 밝히고 MPEG으로 압축된 데이터로부터 DCT 계수와 움직임벡터를 이용함으로써 빠른 시간 내에 셋 경계를 검출할 수 있다는 장점을 갖기도 한다. 특히 [8]에서는 복호화 과정이 단순해서 빨리 처리할 수 있다는 장점을 갖지만, 매크로블록의 타입 정보와 이웃하고 있는 P 혹은 B 프레임들 사이의 특정 타입을 갖는 매크로블록 수를 이용함으로써, 실제로는 다른 셋에 포함되지만 특정 타입의 매크로블록 수가 서로 유사할 경우 이들 프레임 사이의 셋 경계를 미검출할 가능성이 크다는 문제점을 갖는다. 본 논문에서는 부분 디코딩(partial decoding)에 의해 처리 데이터량을 현저히 줄이고 프레임 차분(frame difference)과 히스토그램 차분(histogram difference), [6]의 D_i^* , 픽셀간 차분(inter-pixel difference) 및 움직임 벡터 정보를 이용한 장면 전환 검출 기법, 누적 히스토그램 교차함수(accumulative histogram intersection measure) 및 DC 이미지 버퍼링 알고리즘에 의한 효과적인 대표 프레임 추출 알고리즘을 제시하고 전체 시스템의 성능을 평가한다.

본 논문의 구성은 다음과 같다. 2장에서는 제안하는 장면 전환 검출 및 대표프레임 추출 알고리즘에 대해 기술하고, 3장에서는 제안한 알고리즘을 이용한 성능평가 및 분석 내용을 검토하고, 4장에서 결론 및 향후 연구 방향에 대해 언급한다.

2. 장면 전환 검출 알고리즘

MPEG 표준에 의한 비디오 시퀀스는 I, P, B-type의 picture로 이루어져 있는데, 각 프레임에 대한 DCT DC 계수는 I-type picture의 경우 바로 압축계수들로부터 부분 디코딩을 거쳐 구하고 P, B type

picture의 경우 각 매크로블록에 대한 순방향, 역방향의 움직임벡터로부터 DC 계수값을 추정하는 방법이 이미 제안된 바 있으며 본 논문에서도 이 방법을 이용하였다.

이 장에서는 보다 효과적인 장면 전환 검출 기법과 대표 프레임 추출 기법을 기술한다.

2.1 급격한 장면 전환 검출 알고리즘을 위한 파라미터

먼저 이웃 프레임 간의 DC 이미지에 대하여 다음과 같은 파라미터들을 계산한다.

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

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

여기서 M, N 은 각각 DC 이미지의 가로, 세로 크기이며 FD_i 는 DC 이미지의 프레임 차분, HD_i 는 DC 이미지의 히스토그램 차분이다. 또한 B 는 히스토그램 bin의 개수이다. DC 이미지의 히스토그램은 원영상의 히스토그램에 비하여 잡음 성질이 추가되어 있기 때문에[9] smoothing filter를 거쳐 평활화한 후 사용하였다. 그리고 이러한 프로세스를 수행함과 더불어 아래의 파라미터를 $i-1$ 번째 프레임까지 계산한다.

avr_FD1 : 셋 경계가 아닌 프레임들의 FD의 평균
avr_FD2 : 셋 경계인 프레임들의 FD의 평균
avr_HD1 : 셋 경계가 아닌 프레임들의 HD 평균
avr_HD2 : 셋 경계인 프레임들의 HD 평균

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

2.2 급격한 장면 전환 검출 알고리즘

장면 전환은 다음의 과정을 거쳐 검출하게 된다.

STEP 1: $FD_i > \alpha \text{avr_FD1}$ and $HD_i > \alpha \text{avr_HD1}$ and shot duration > minimum shot duration

STEP 2: $FD_i > \beta \text{avr_FD2}$ and $HD_i > \beta \text{avr_HD2}$

STEP 3: $FD_i > \gamma FD_{i-1}$ and $FD_i > \gamma FD_{i+1}$ and

$HD_i > \delta \text{avr_HD1}$

STEP 4: $HD_i > \gamma HD_{i-1}$ and $HD_i > \gamma HD_{i+1}$ and $FD_i > \delta \text{avr_FD1}$

STEP 5: $HD_i > \epsilon \text{avr_HD1}$

STEP 6: if shot duration < minimum shot duration, then calculate flashlight detection

여기서 α 는 1에서 2, β 는 0.8정도, γ 는 3, δ 는 2, ϵ 는 3에서 4 정도의 값을 갖는다. 또한 최소 셋 구간은 10 프레임 정도가 적당하다. 특히 α , β 등과 같은 파라미터 값은 절대적인 수치(absolute value)가 아니라 영상에 따른 통계적 특성을 이용하여 선정된 상대적인 율(relative ratio)로, 비디오 전체 내용에서 장면전환이 급격히 자주 계속해서 발생하는 경우는 사실상 거의 드물고 어느 정도 연속된 프레임이 지속되어 나가면서 급격한 혹은 점진적 장면 전환이 이뤄지는 것이 일반적인 경우에 해당하므로 두 가지 경우, 즉 셋의 경계가 아닌 프레임들과 셋경계인 프레임들의 평균값에 대해 각각 200% 그리고 80% 이상 차이가 날 때를 장면 전환으로 판단되는 기준 시점으로 삼았다. 또한, α 를 너무 크게 하면 아주 현격한 차이가 나는 장면 전환만을, 그리고 β 를 작게 하면 너무 잦은 장면 전환을 검출할 가능성이 있으므로 앞서 기술한 일반적인 영상의 통계적 특성을 고려해서 설정하였다. 또한 이들 파라미터 값은 영상 내부적인 특징 변화에 그다지 민감하게 작용하지 않기(거의 무관하기) 때문에 다른 일반적인 영상 데이터에 대해서도 동일하게 적용 가능하다는 장점을 갖는다.

STEP 1은 나머지 Step을 거치기 위해 이웃 영상간의 변화가 일정이상으로 큰 지를 검사하고 프레임이 이전 프레임에 비해 변화가 적은 경우 불필요하게 다음의 과정을 수행하지 않게 한다. 그리고 급격한 장면 전환이 발생하면 이웃 프레임에 비해 영상 내용이 갑자기 크게 변하게 되므로 FD_i , HD_i 는 큰 값을 갖게되어 장면 전환으로 검출한다.

위 과정 중 STEP 1이 만족해야만 STEP 2에서 STEP 5까지의 과정을 수행하며, 위 STEP 2에서 STEP 5까지의 과정 중 한 조건 이상을 만족하면 i 번째 프레임을 장면 전환으로 검출한다. 또한 flashlight이나 연속된 장면 전환이 짧은 시간에 걸쳐 발생하면 장면 전환으로 검출되는 길이가 짧아지므로

경우, flashlight 혹은 짧은 구간의 연속된 장면 전환인가를 검사하는 프로세싱을 수행할지 결정한다.

STEP 2에서 STEP 5까지 여러 단계를 거치는 것은 히스토그램과 프레임 차분의 특성, 즉 히스토그램이 일반적으로 움직임에 덜 민감한 반면 유사한 분위기나 배경을 갖는 셋 경계는 잘 찾아내지 못하는 특성과 움직임과 조명 변화에 특히 민감한 프레임간 차분의 특성을 모두 고려한 것이다.

2.3 점진적인 장면 전환 검출

점진적인 장면 전환의 검출을 위해 다음과 같은 파라미터들을 사용한다.

$$D_i^k = d(I_i, I_{i-k}) \quad (3)$$

$$\overrightarrow{emu}_{iv} = \begin{cases} \overrightarrow{mv}_{iv}, & \text{if } |\overrightarrow{mv}_{iv}| > 5 \\ 0, & \text{elsewhere} \end{cases} \quad (4)$$

$$N_i^{emu} = \sum_{j=i-r}^i n_j \quad (5)$$

$$S_i^{emu} = \sum_{j=i-r}^i \frac{1}{N^{MB}} \sum_{v=1}^{N^{MB}} |\overrightarrow{emu}_{jv}| \quad (6)$$

(d : 프레임간 차분, D_i^k : i 번째 DC 이미지와 $i-k$ 번째 DC 이미지간의 프레임 차분, $\overrightarrow{emu}_{iv}$: 길이 5이상인 i 프레임 내 v 번째 유효 움직임 벡터, n_j : j 프레임 내 유효 움직임 벡터 수, r : 점진적인 장면 전환 길이, N^{MB} : 매크로 블록의 총 개수)

위 식을 적용함에 있어 움직임 벡터는 불필요한 잡음 성분의 제거를 위해 움직임 벡터에 3×3 median filter를 적용하여 사용한다[10]. 또한, N_i^{emu} , S_i^{emu} 는 P/B-type picture에 대해 전후 프레임과 현재 프레임간의 평균을 재계산하여 사용하며, I-type picture에서는 0을 갖는다. 아울러 $k=20$, $r=10$ 을 사용하였다.

식 (3)은 점진적인 장면 전환의 경우 급격한 장면 전환에서와 같이 이웃 프레임 간 변화량이 크게 나타나지 않으므로[6]에서 제안한 바와 같이 일정 길이를 건너뛰어 프레임 간 변화를 측정한다. 식 (3)은 급격한 장면 전환이나 flashlight이 구간 내에 있을 경우도 큰 값을 갖으나 이는 급격한 장면 전환 검출을 통해 제외시킬 수 있다. 또한 식 (4-6)을 통해 움직임 벡터를 이용하여 현재의 대상 구간이 움직임에 의해 저지저지하며 전환으로 검출되는 길이를 비가역적

기 어려운 단점이 있다.

위 파라미터를 사용하여 다음 조건을 만족하는 구간에 대해 점진적인 장면 전환이 일어난 것으로 검출한다.

$$D_j^k \geq l \times avr_FD2 \text{ and } D_j^k \geq m \times D_{static}^k \quad (7)$$

$$N_i^{mv} < n \text{ and } S_i^{mv} < s$$

(단 $j=i-r, i-r+1, \dots, i$)

위 식에서 $l=0.8, m=3$ 의 값을 가지며 n, s 는 실험에 의해 미리 정의된 값을 갖는다. D_{static}^k 는 장면 전환이 일어나지 않은 구간에서의 D_j^k 의 평균이다.

식 (7)은 D_j^k 가 점진적인 장면 전환에 의해 의미있는 큰 값을 가지고 구간 내 움직임이 적은 경우 점진적인 장면 전환으로 검출함으로써 [6]에서 제안한 알고리즘보다 움직임에 강건한 성능을 가진다.

또한 다음의 픽셀간 차분 (inter-pixel difference) 식을 사용하여 페이드 인 (fade in)/페이드 아웃 (fade out)의 검출에 사용한다.

$$PD_i = \sum_{x=0, y=0}^{x \leq N, y \leq N} |I_i(x, y) - I_i(x-1, y)| \quad (8)$$

점진적인 장면 전환 구간 ($i-r, i$)에서 $PD_i - PD_{i-r}$, 이 양의 임계값 이상의 값을 가지고 PD_{i-r} 이 충분히 작은 값일 때 페이드 인이며, $PD_i - PD_{i-r}$ 이 음의 임계값 보다 작은 값을 가지고 PD_i 가 충분히 작은 값을 가질 때 페이드 아웃, 그 이외에는 디졸브로 검출한다. 페이드 인인 경우 영상 내 복잡도가 거의 없는 단순한 섷으로부터 새로운 영상으로 천이되므로, 장면 전환 초기에는 픽셀간 차분값이 비교적 작은 값을 가지며 새로운 섷으로 진입함에 따라 선형적으로 큰 값을 갖게 되는 성질을 나타내며 페이드 아웃은 그 반대의 특성을 보인다.

페이드의 경우 영상 내 복잡도가 거의 없는 단순한 흰 배경이나 검은 배경의 섷과 의미있는 복잡도를 가지는 일반 섷 간의 디졸브이므로 식 (8)을 통해 영상 내 복잡도의 변화로부터 페이드를 판별할 수 있다. 반면 디졸브에서는 천이되는 두 섷의 PD_i 가 일반적으로 모두 크게 나타나므로 본 논문에서는 위에서 언급한 바와 같이 점진적인 장면 전환이 페이드인지 디졸브인지를 구분한다.

그러나 픽셀간 차분값이 같은 경우 같은 단순한

이 어렵다는 단점이 있다.

2.4 대표 프레임 추출 알고리즘

셋 경계가 아니며 프레임간 차분이 일정 크기 이상인 프레임에 대해 본 논문에서는 누적 히스토그램 교차 함수 (accumulative histogram intersection measure)를 구하여 대표 프레임으로 선택할지를 결정하는 방법을 제안한다.

제안하는 누적 히스토그램 교차 함수의 정의식은 다음과 같다.

$$\overline{H}_i(j) = \min(\overline{H}_{i-1}(j), H_i(j))$$

$$AHIM_i = \frac{1}{MN} \sum_{j=0}^{B-1} \overline{H}_i(j) \quad (9)$$

(단, $H_i(j)$ 는 i 번째 프레임에서 DC 이미지의 히스토그램의 j 번째 bin값, B 는 bin의 총 개수)

위 (9)에 의해 히스토그램의 유사도를 측정함으로써 현재 프레임이 이전 프레임들에 대하여 얼마나 많은 변화를 가지는지를 효과적으로 판별할 수 있으며, 혹 검출되지 못한 중요한 영상 내용의 변화에 대해서는 대표 프레임으로 반영하도록 한다.

또한 불필요하게 $\overline{H}_i(j)$ 의 레벨이 떨어지는 것을 막기 위하여 프레임간 차분이 일정 크기 이상인 경우에만 그 이전 계산된 값과의 누적 히스토그램 교차 함수를 구하도록 하여 가능한 한 정확한 영상의 내용 변화를 추적할 수 있게 하였다. 그리고 셋 내에서 영상 내용의 변화를 측정할 수 있도록 하기 위해 새로운 장면 전환이 나타나거나 대표 프레임이 검출되었을 경우 (9)식의 $\overline{H}_i(j)$ 를 초기화한다. 이렇게 함으로써 셋 내의 유사도가 일정 수준 이하로 떨어짐을 검출할 수 있다.

이와 같은 알고리즘을 적용함에 있어 장면 전환 검출에 의해 추출된 각 섷은 섷의 첫번째 프레임을 기본적인 대표 프레임으로 가지며 검출된 대표 프레임의 DC 이미지들은 동적으로 할당된 DC 이미지 버퍼에 보관된다. 따라서 새로 검출된 대표 프레임과 현재 소속된 섷 내의 이미 검출된 다른 대표 프레임들과의 히스토그램 교차 함수를 통한 유사도 비교 과정을 거쳐 새로 등장한 대표 프레임인지 아닌지를 판별한다. 즉, 셋 내에 이미 검출된 다른 대표 프레임이 존재한다면 이전 수준 이하이며 새로운 대표 프

DC 이미지는 DC 이미지 버퍼에 추가된다. 단, 점진적인 장면 전환 구간에서는 대표 프레임을 선택하지 않는다. DC 이미지 버퍼는 현재 대표 프레임 추출을 수행 중인 섷과 그 이전의 섷에 대한 대표 프레임의 DC 이미지 정보를 보관하는 동적인 버퍼이며 DC 이미지는 그레이레벨 원영상의 1/64의 정보만을 가지므로 이로 인한 메모리 사용량의 증가는 매우 적다.

이와 같은 DC 이미지 버퍼링 알고리즘을 통해 비교적 적은 메모리를 사용하면서 추출된 대표 프레임의 유효성 여부를 보다 정확하게 판별할 수 있다. 또한 DC 이미지 버퍼는 flashlight 검출단에서 단순한 flashlight인지 아니면 짧은 구간의 연속된 장면 전환 인지를 판별하는데 사용된다.

2.5 flashlight 및 짧게 연속적으로 등장하는 장면 전환 검출 알고리즘

뉴스와 같은 영상에서 많이 나타나는 카메라의 플래시나 영화에서의 급격한 조명 변화 등은 장면 전환으로 오검출되기 쉬우며 카메라의 플래시는 중요한 인물이나 사건과 관계될 수 있는 정보로도 활용 가능하다[6]. 또한 CF, 뮤직 비디오, 영화 등에서는 필름 편집 등을 통해 인위적으로 플래시와 비슷한 효과를 삽입하여 특수 효과로 많이 사용한다. 이와 같은 플래시 혹은 그와 유사한 종류의 빠르게 연속하여 일어나는 장면 전환을 보다 정확하게 검출하는 다음과 같은 알고리즘을 제안한다.

Step 1:

if the duration of current detected abrupt-shot is less than minimum shot duration threshold, then proceed Step 2, Step 3.

Step 2:

when $FD(k_{lastshot}, I_i) > \gamma \text{avr_FD1}$ and $HD(k_{lastshot}, I_i) > \gamma \text{avr_HD1}$ (A-B-C type),

if $PDI(I_{flashlight}) > P$ then successive abrupt scene change has been occurred,

else abrupt scene change with edited flashlight-like cut insertion has been occurred (A-inserted flashlight-B type).

Step 3:

if Step 2 is not satisfied, then natural flashlight has been occurred (A-flashlight-A' type).

(minimum shot duration threshold: 10 프레임, $k_{lastshot}$: DC 이미지 버퍼에 보관된 가장 최근에 검출된 섷의 제일 마지막 대표 프레임, I_i : i 번째 프레임의 DC 이미지, $I_{flashlight}$: flashlight로 예상되는 프레임의 DC 이미지, P : 실험에 의해 미리 정의된 inter-pixel difference의 임계값)

위 Step에서 A-B-C type은 완전히 서로 다른 영상 내용을 갖는 A, B, C의 섷이 연속되는 형태이다. 또한 A-flashlight-A' type은 flashlight이 발생하거나 또는 인위적인 필름 편집에 의해 flashlight과 같은 효과를 삽입하여 거의 같은 영상 내용 A, A'가 flashlight에 의해 구분되는 형태를 뜻한다. 이에 대한 예를 그림 1에 나타내었다.

본 논문에서는 위의 Step 1에서 10프레임 이하의 섷 구간을 가지는 경우 Step 2, Step 3을 통해 flashlight 혹은 짧은 구간의 섷을 검출한다. Step 2에서는 DC 이미지 버퍼에 보관된 가장 최근에 검출된 섷의 마지막 대표 프레임 $k_{lastshot}$ 와 현재 검출된 짧은 구간의 섷의 그 다음 구간에 속하는 첫 프레임 I_i 의 DC 이미지의 프레임간 차분 및 히스토그램 차분과 비교하여 서로 유사한 내용의 섷인지 아닌지를 판별한다. 또한 현재 검출된 짧은 구간의 섷에서 첫 번째 픽셀간 차분값 $PDI(I_{flashlight})$ 을 계산하여 작은 값을 가질 경우 인위적인 필름 편집에 의한 무의미한 섷으로 판정한다. 또한 Step 3에서는 자연적으로 영상 내용에서 발생한 flashlight을 검출한다. 그리하여 A-B-C type인 경우 모두 섷으로 검출하며 A-inserted flashlight-B type인 경우 A, B 섷만을 색인한다. 또한 A-flashlight-A' type인 경우 A' 섷의 첫 프레임을 대표 프레임으로 색인한다.



(a) A-B-C type



(b) A-flashlight-A' type

그림 1. Flashlight type의 예

그림 2는 본 논문에서 지금까지 제안한 전체 알고리즘의 흐름도를 나타낸 것으로 다음과 같은 단계를 통해 장면 전환 검출 및 대표 프레임 추출이 수행된다. 즉, 입력으로 들어오는 MPEG 시퀀스에 대해 부분 복호화를 통해 DC 이미지를 구성하고 이들간의 히스토그램 차분과 프레임 차분을 이용해서 급격한 장면 전환을 검출한다. 만일 검출된 장면 전환 구간이 너무 짧은 경우 flashlight 혹은 짧은 구간 내에서 연속적으로 바뀌는 장면 전환인가를 검사하게 된다. 아울러 픽셀간 차분 및 유효 움직임 벡터 정보를 이용한 점진적인 장면 전환 검출을 수행한다. 또한 AHIM을 이용하여 효과적으로 대표 프레임을 추출하게 된다.

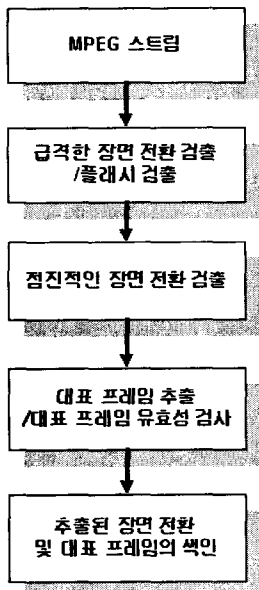


그림 2. 전체 알고리즘의 흐름도

3. 실험 및 결과

제안된 알고리즘은 Visual C++ 6.0에 의해 Windows NT 상에서 구현하였다. 시뮬레이션 틀은 그림 3과 같이 Interactive Video Browser로서의 기능을 갖춘 형태로 제작되어 모든 결과를 GUI환경에서 쉽게 볼 수 있으며 비디오 영상을 자유롭게 액세스(access)할 수 있고 검출된 장면 전환 및 대표 프레임을 통해 영상의 내용을 한 눈에 알아볼 수 있다.

알고리즘의 성능 평가에 사용된 영상은 320×240

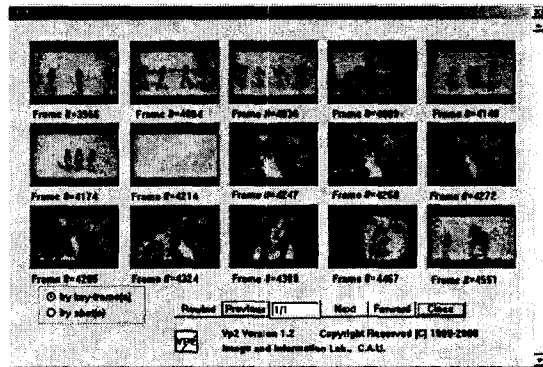
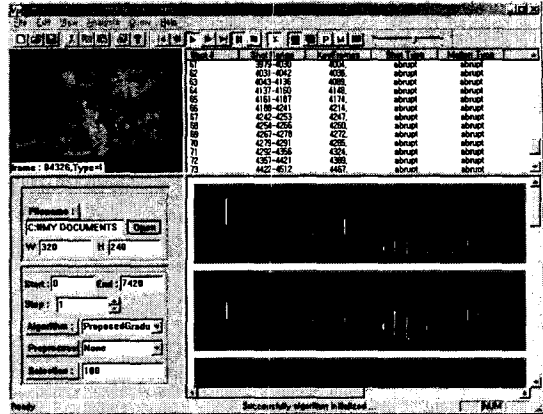


그림 3. 프로그램 실행화면 및 추출된 대표프레임 예

크기의 MPEG 부호화된 영화 및 뮤직 비디오 파일 등을 사용하였으며, 총 프레임 수는 15,245이고 급격한 장면 전환이 227개, 점진적인 장면 전환이 12개로 총 239개의 장면 전환이 포함되어 있다. 영화 영상의 경우 비교적 급격한 장면 전환이 많이 나타나는 전반부와 물체의 폭발 장면과 더불어 조명의 변화가 많으며 액션이 강한 후반부로 나뉜다. 뮤직 비디오 영상은 전반부에서는 계속해서 panning 및 tracking이 일어나며, close-up된 카메라 동작을 보여주는 부분이 있으며, 후반부에는 flashlight 썰에 이은 급격한 장면 전환이 반복되는 부분이 있다. 또한 전반적으로 많은 카메라 움직임을 포함하고 있으며 물체의 움직임은 보통 정도이다.

DC 이미지간의 히스토그램 차분을 이용한 고전적인 방법[3]에 의한 장면 전환 검출과 [6]의 알고리즘에서 Yeo & Liu가 제시한 최적의 파라미터를 그대로 적용한 장면 전환 검출 (파라미터를 달리 설정하여 실험하였으나 결과는 그들이 제시했던 최적—가장 우수한 성능을 나타내는 경우의 파라미터 조합—의

파라미터 설정값을 적용했을 경우의 결과가 표의 음영 부분에서와 같이 가장 우수한 성능을 보임을 확인할 수 있다. 예를 들면 표에서와 같이 파라미터 가운데 프레임 차분 계산을 위한 연속된 프레임 구간의 크기(m), flashlight 검출에 사용되는 상수값(ϵ 과 p)을 각각 달리 설정했을 때의 결과를 보여준다), 그리고 본 논문에서 제안한 알고리즘 (물론 제안한 방법에서도 파라미터 설정에 따라 성능에 있어 다소 차이가 있으나 잘못 검출되는 경우를 크게 줄일 수 있으며 이는 기존 논문에서도 유사한 형태로 제시되는 점을 감안할 때 본 논문에서 제시한 최적의 파라미터 값을 적용할 경우 표의 음영 부분과 같이 기존 방식에 비해 보다 우수한 성능을 보임을 알 수 있다)의 실험 결과를 표 1에 나타내었다.

[6]의 알고리즘은 일반적인 영상들에 대해서는 좋은 성능을 보이거나 움직임에 민감하여 움직임이 많을 경우 급격한 장면 전환이 검출되기 어려우며 flashlight에 대한 대응이 미흡하여 적은 수의 카메라 flashlight는 잘 검출하는 반면 자주 등장하는 flashlight들은 검출할 수 없고 아주 짧은 간격의 연속된 장면 전환을 flashlight로 잘못 검출함으로써 인해 장면 전환을 flashlight로 오검출하는 특성을 보였다. 아울러 점진적인 장면 전환에서는 [6]의 알고리즘은 매우 이상적인 형태의 디졸브만을 검출할 수 있었다.

점진적인 장면 전환의 경우[6]의 알고리즘은 총 12개의 페이드 및 디졸브 중 6개만이 검출 가능했으며 본 논문에서 제안한 알고리즘은 9개가 검출되었다. 특히 실험 영상에는 카메라 flashlight 12개와 10프레임 이하인 짧은 구간의 셋이 9개가 포함되어있는데 [6]의 알고리즘은 카메라 flashlight 8개만이 검출 가능하였으나 제안한 알고리즘은 flashlight 10개와 짧은

은 구간의 셋 9개를 검출할 수 있었다.

본 논문에서 제안한 알고리즘은 급격한 장면 전환 검출 성능은 물론 특히 flashlight이 많거나 짧은 간격의 연속적인 장면 전환에 대해 보다 정확한 검출이 이루어짐을 알 수 있으며 이로 인한 미검출율이 현저히 개선되었음을 알 수 있다. 또한 급격한 움직임에 민감한 기존 알고리즘에 비해 움직임에 더욱 강건한 점진적인 장면 전환 검출 성능을 보였다.

4. 결 론

본 논문에서는 방대한 MPEG 비디오로부터 부분적인 최소한의 디코딩만을 거친 후 단순 파라미터가 아닌 프레임 차분, 히스토그램 차분, 유효 움직임 벡터 등과 같은 파라미터들의 조합을 적응적으로 적용함으로써 급격한 장면 전환 및 점진적인 장면 전환을 행하는 보다 효과적인 알고리즘을 제안하였으며, 분할된 장면들로부터 전체 영상을 요약하기 위한 대표 프레임 추출 알고리즘도 제시하였다. 아울러 제안한 알고리즘을 기존 알고리즘과 비교 검토함으로써 그 우수성을 보였다.

이와 같은 연구 내용은 비디오 내용 기반 영상 검색 시스템, 디지털 비디오 편집기, 내용 기반 비디오 브라우저, 주문형 비디오 (Video on Demand) 등의 연구에 활용될 수 있으며, 본 논문에서 보인 비디오 브라우저가 그 좋은 예이다.

참 고 문 헌

[1] 이미숙, 황본우, 이성환, "내용 기반 영상 및 비디오 검색 기술의 연구 현황", *정보과학회지*, 제 15권 제9호, pp. 10-19, 1997.
 [2] Haitao Jiang, A. Helal, Ahmed K. Elmagarmid and Anupam Joshi, "Scene change detection techniques for video database systems", *Multimedia Systems*, pp. 186-195, 1998.
 [3] Borko Furht, Stephen W. Smoliar and HongJiang Zhang, *Video and image processing in multimedia system*, Kluwer Academic Publishers, 1995.
 [4] Wei Xiong and John Chung-Mong Lee, "Efficient scene change detection and camera

표 1. 기존 알고리즘과 제안된 알고리즘의 성능 비교

				correct	missed	false	hit ratio (%)
히스토그램 차분법 [3]				198	41	103	82.8
Yeo & Liu [6]	m=5	$\epsilon=0.1$	p=5	218	21	36	91.2
	m=7	$\epsilon=0.01$	p=4	212	27	39	88.7
	m=9	$\epsilon=0.001$	p=3	207	32	44	86.6
proposed	$\alpha=2$	$\beta=0.8$	$\gamma=3$	228	11	15	95.4
	$\alpha=1$	$\beta=1$	$\gamma=2$	216	23	20	90.4
	$\alpha=3$	$\beta=0.2$	$\gamma=1$	203	36	35	84.9

motion annotation for video classification”, *Computer Vision and Image Understanding*, vol. 71, no. 2, pp. 166-181, 1998.

[5] Chua Hock Chuan, Chen Yu and Yakooob M. Siyal, “A content-based MPEG video parser for video scene analysis”, *The Fifth International Conference on Control, Automation, Robotics and Vision*, Singapore, December, 1998.

[6] Boon-Lock Yeo and Bede Liu, “Rapid scene analysis on compressed video”, *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 5, no. 6, pp. 533-544, 1995.

[7] Y. Nakajima, K. Ujihara and A. Yoneyama, “Universal scene change detection on MPEG-coded data domain”, in *Proc. SPIE Visual Communication and Image Processing*, vol. 3024, pp. 992-1003, 1997.

[8] N. Gamaz, X. Huang and S. Panchanathan, “Scene change detection in MPEG domain”, in *Proc. IEEE Southwest Symposium on Image Analysis and Interpretation*, pp. 12-17, 1998.

[9] Nilesh V. Patel and Ishwar K. Sethi, “Video shot detection and characterization for video databases”, *Pattern Recognition*, vol. 30, no. 4, pp. 583-592, 1997.

[10] Ruggero Milanese, F. Deguillaume and A. Jacot-Descombes, “Video segmentation and camera motion characterization using compressed data”, *SPIE Multimedia Storage and Archiving Systems II*, vol. 3229, pp. 23-31, 1997.



강 응 관

1993년 2월 중앙대학교 전자공학과 학사
 1995년 2월 중앙대학교 대학원 전자공학과 석사
 1995년 1월~1997년 5월 현대전자 정보통신연구소 연구원

2001년 2월 중앙대학교 첨단영상대학원 영상공학과 박사
 2001년 9월~현재 전주대학교 정보기술컴퓨터공학부 멀티미디어공학과 조교수, 멀티미디어 기술사
 관심분야 : 동영상 부호화, 멀티미디어 신호처리, 멀티미디어 데이터베이스, 영상통신 등

교신저자

강 응 관 560-759 전북 진주시 완산구 효자동 3가 120C 전주대학교 정보기술컴퓨터공학부 멀티미디어공학과