

# 비디오 시퀀스에서 장면 전환 검출과 시뮬레이터의 구성

## Scene change detection and simulation tool in video sequence

김 성 주, 강 응 관, 최 중 수

중앙대학교 공과대학 전자공학과 영상정보연구실

Sung Joo Kim, Eung Kwan Kang, Jong Soo Choi

Dept. of Electronic Engineering, Chung-Ang University

### 요 약

장면 전환 검출(scene change detection)은 영상 정보의 인덱싱 및 검색을 위한 전처리로서, 전체 검색 시스템의 성능을 좌우하는 중요한 기술로 현재 많은 연구가 진행되고 있다. 본 논문에서는 MPEG 표준으로 압축된 영상으로부터 얻은 DC 이미지를 이용한 장면 전환 검출 및 대표 프레임 검출에 대한 방법을 제안하고 이를 위한 시뮬레이터의 개발과 그에 대한 성능을 평가한다.

### 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][2]. 특히 영상 내 내용을 기반으로 한 시간 축 분할(temporal segmentation) 후 각 분할된 영역을 대표 프레임(representative frame) 등으로 표현함으로써, 전체 비디오 시퀀스를 효과적으로 표현하고 이를 바탕으로 인덱싱 및 검색에 이용하는 기술이 각광을 받고 있다.

비디오 시퀀스에서 샷(shot)은 필름이 끊기지 않고 시간적으로나 공간적으로 연속적으로 촬영된 일정 구간의 영상을 의미한다. 비디오 분할 기술은 장면 전환 효과에 의해 발생하는 샷 사이의 경계를 검출하는 방법으로, 샷은 비디오 분할의 기본 단위로 사용된다.

장면 전환의 종류로는 급격한 장면 전환(abrupt scene cut)과 점진적인 장면 전환(gradual scene cut)이 있으며, 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 많이 이용되고 있으며[3], 따라서 MPEG 압축 영역에서의 장면 전환 검출 또한 많이 연구되고 있다[2]. 본 논문에서는 MPEG 표준에 의한 비디오 시퀀스에서 일반적인 큰 움직임 및 조명 변화에서도 적용적으로 장면 전환을 검출하는 알고리즘과 함께 각 샷을 표현하는 키 프레임의 적응적인 추출에 대한 방법을 제안하고 이를 실제 시뮬레이션 프로그램으로 제작하여 성능을 평가해 보고자 한다.

### 2. 장면 전환 검출과 비디오 브라우저 시뮬레이터의 구현

#### 2.1. MPEG 비디오 시퀀스에서 DC 이미지의 추출.

MPEG 표준에 의한 비디오 시퀀스는 I, P, B type의 picture로 이루어져 있는데, 이로부터 간축된 움직임 보상에 의한 DC 이미지를 추출할 수 있는 알고리즘이 이미 제안된 바 있다[4][5].

본 논문에서는 이러한 방법 중에서 각 I 픽처로부터 DCT DC 계수에 의한 DC 이미지의 구성과 P, B 픽처에서의 움직임 보상과 예측 오차에 의한 DC 이미지의 구성에 의한 방법을 사용하였다.

#### 2.2 장면 전환 검출 알고리즘.

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

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

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

$$\chi_i^2 = \sum_{k=0}^{k=N} \frac{(H_{i-1}(k) - H_i(k))^2}{H_i(k)} \quad (3)$$

여기에서  $M, N$ 은 각각 DC 이미지의 가로, 세로 크기이며  $FD_i$ 는 DC 이미지의 frame difference,  $HD_i$ 는 DC 이미지의 histogram difference[6],  $\chi_i^2$ 는 DC 이미지의 히스토그램에 대한 chi-square이다. DC 이미지의 히스토그램

은 원영상의 히스토그램에 비하여 잡음 성질이 추가되어 있기 때문에 smoothing filter 를 거쳐 평활화한 후 사용하였다[7]. 그리고 이러한 프로세스를 수행함과 더불어 아래 파라미터를  $i-1$  번째 프레임까지 계산한다.

avr\_FD1 : shot boundary 가 아닌 프레임들의 이전 프레임과의 frame difference 의 평균.

avr\_FD2 : shot boundary 인 프레임들의 이전 프레임과의 frame difference 평균.

avr\_HD1: shot boundary 가 아닌 프레임들의 이전 프레임과의 histogram difference 평균.

avr\_HD2:shot boundary 인 프레임들의 이전 프레임과의 histogram difference 평균.

avr\_X1 : shot boundary 가 아닌 프레임들의 이전 프레임과의  $\chi^2$ 의 평균.

avr\_X2 : shot boundary 인 프레임들의 이전 프레임과의  $\chi^2$ 의 평균.

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

### 2.3 장면 전환 검출 알고리즘의 수행

위에서 계산된 파라미터들로부터 다음의 과정을 거쳐 장면 전환을 검출하게 된다.

STEP 1:  $FD_i > \alpha avr\_FD1$ ,  $HD_i > \alpha avr\_HD1$ ,  
 $\chi_i^2 > \alpha avr\_X1$ , shot duration > minimum shot duration

STEP 2:  $FD_i > avr\_FD2$  and  $\chi_i^2 > avr\_X2$

STEP 3:  $FD_i > \beta FD_{i-1}$  and  $FD_i > \beta FD_{i+1}$  and  
 $\chi_i^2 > \gamma avr\_X2$

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

STEP 5:  $HD_i > \beta HD_{i-1}$  and  $HD_i > \beta HD_{i+1}$

STEP 6:  $\chi_i^2 > \delta avr\_X2$

STEP 7:  $HD_i > \delta avr\_HD2$

위 과정 중 STEP 1 이 만족해야만 STEP 2-7 의 과정을 수행하며, 위 STEP 2 에서 STEP 7 까지의 과정 중 한 조건 이상을 만족하면  $i$  번째 프레임을 장면 전환으로 검출한다. 그리고, 위 판별 과정에 의해 장면 전환으로 검출된 프레임들

가운데 flash light 과 같은 급격한 조명 변화의 영향에 의한 shot boundary 인 경우를 배제하기 위해,  $FD_{i+1}$  이  $FD_i$  에 대하여 충분히 큰 값을 가지는 반면  $\chi_{i+1}^2$  는  $\chi_i^2$  에 비해 훨씬 작게 나타나는 성질을 이용하여 검사한다. 이로부터 검출된 flash light cut 은 장면 전환에서 제외한다. 한편 STEP 5 의 조건식만 고려할 경우, 즉 히스토그램이 일반적으로 움직임에 덜 민감한 반면 유사한 분위기나 배경을 갖는 첫 경계는 잘 찾아내지 못하기 때문에 본 논문에서와 같이 frame difference, chi-square 를 조합한 장면 전환 검출을 고안하였다.

또한 shot boundary 는 아니지만 STEP 1 을 만족하는 프레임에 대해 본 논문에서는 누적 histogram intersection measure 를 구하여 대표 프레임으로 선택할지를 결정하는 방법을 제안한다.

제안하는 누적 histogram intersection measure 의 정의식은 다음과 같다.

$$\overline{H_i(J)} = \min(\overline{H_{i-1}(J)}, H_i(J))$$

$$AHIM_i = \frac{1}{MN} \sum_{j=0}^{B-1} H_i(J) \quad (4)$$

여기에서  $H_i(J)$  는  $i$  번째 프레임에서 DC 이미지의 히스토그램의  $j$  번째 bin 값, B 는 bin 의 총 개수.

위 (4)에 의해 시간적인 히스토그램의 유사도의 변화를 측정함으로써 현재 프레임이 이전 프레임들에 대하여 얼마나 많은 변화를 가지는지를 효과적으로 판별할 수 있으며, 혹 검출되지 못한 중요한 영상 내용의 변화를 대표 프레임으로 반영하도록 한다.

더욱이 위 값을 원영상으로부터 구할 경우 장면 전환 검출 알고리즘에 사용 가능할 뿐만 아니라, 이의 1차 도함수를 이용하면 보다 좋은 결과를 얻을 수 있다. 그러나 DC 이미지 계산에 의한 노이즈 성분과 원영상이 아닌 축소된 형태의 영상 적용에 의한 노이즈 성분을 무시할 수 없으므로 본 논문에서는 DC 이미지의 히스토그램에 LPF 를 거쳐 (4)에 적용한다[7].

또한 불필요하게  $\overline{H_i(J)}$  의 레벨이 떨어지는 것을 막기 위하여  $HD_i$  가 일정 크기 이상인 경우에만 그 이전 계산된 값과의 누적 histogram intersection measure 를 구하

도록 하여 가능한 정확한 영상의 내용 변화를 추적할 수 있게 하였다. 그리고 셋 내에서의 영상 내용의 변화를 측정할 수 있도록 하기 위해 새로운 장면 전환이 나타나거나 대표 프레임이 검출되었을 경우 (4)식의  $\overline{H_i(t)}$  를 초기화한다. 이렇게 함으로써 셋 내의 유사도가 일정 수준 이하로 떨어짐을 검출할 수 있다.

만약 셋 내에서 위의 방법에 의한 대표 프레임이 출현하지 않을 경우 셋 내의 내용 변화가 별로 없는 것이므로 셋의 중간 프레임을 대표 프레임으로 선택한다. 또한 한 셋이 여러 개의 대표 프레임을 갖는 것을 허용하였다.

#### 2.4 장면 전환 알고리즘에 의한 Interactive Video Browser의 구현

그림 1은 장면 전환 검출과 대표 프레임 검출을 이용한 비디오 브라우저를 구현하여 실행한 모습이다. 본 프로그램은 MPEG 비디오 스트림을 재생할 수 있을 뿐 아니라 되감기, 빨리 감기, 역 재생, 일시 중지, 빠른 재생 등의 사용자 환경을 제공한다. 또한 제안된 알고리즘에 의해 장면 전환과 대표 프레임을 검출하여 그 결과를 사용자가 확인할 수 있으며, 비디오 시퀀스를 랜덤하게 혹은 셋 단위로 볼 수 있을 뿐만 아니라, 대표 프레임들만을 검색할 수도 있게함으로써 전체 내용을 쉽고 빠르게 찾을 수 있는 기능을 포함한다.

그림 1에서 네 개의 각 서브 윈도우에는 왼쪽 상단부터 시계방향으로 각각 디스플레이 윈도우, 셋 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, 알고리즘 시뮬레이션에 의해 계산되는 각종 파라미터들을 그래프로 볼 수 있는 윈도우, MPEG 파일 불러오기나 시작 프레임과 끝 프레임을 사용자가 지정 또는 확인할 수 있도록 하고 장면 전환 검출 알고리즘을 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 수행 할 지, 또는 디스플레이 없이 최대한 빠른 속도로 계산만 수행할 지를 선택할 수 있게 되어 있으며, 슬라이드 바를 이용하여 랜덤하게 영상을 액세스할 수 있도록 되어

있다.

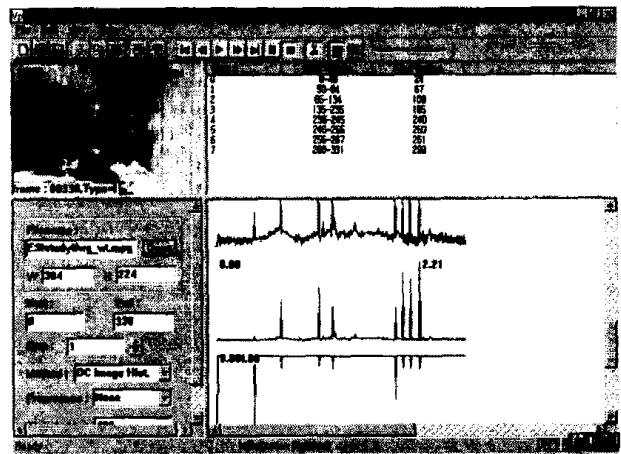


그림 1. 비디오 브라우저의 실행 모습

### 3. 실험 결과

테스트 영상에 대한 실험 결과를 표 1에 나타내었다. 실험에 사용된 테스트 영상들은 모두 인터넷 상에서 얻은 것들로 처음 두 영상은 각각 애니메이션과 영화의 일부분으로 상당히 짧은 반면 길이에 비해 많은 셋들을 포함한 영상들로 상당히 움직임이 많은 영상들이며 특히 두번째 영상은 급격한 조명 변화에 의한 flash cut 이 두 개 검출되었다.

세번째 영상은 영화 중 일부로 영상의 전반에 걸쳐 panning 이 일어나고 또한 움직임이 매우 많은 영상으로 표 1에서와 같이 비교적 떨어지는 결과를 보였다. 그러나 놓친 셋들 중 두 개의 셋은 대표 프레임으로서는 검출될 수 있었다.

#### 참고문헌

- [1] Borko Furht, Stephen W. Smoliar and HongJiang Zhang, "Video and Image Processing in Multimedia System", Kluwer Academic Publishers, 1995
- [2] Ahmed K. Elmagarmid et al., "Video Database Systems: Issues, Products, and Applications," Kluwer Academic Publishers, 1997
- [3] 대우전자 영상연구소, "MPEG 비디오", 연암출판사, 1995
- [4] K.R. Rao, J.J. Hwang, "Techniques & Standards for Image, Video & Audio Coding", Prentice-Hall PTR, 1996
- [5] Y. Nakajima, K. Ujihara, A. Yoneyama, "Universal scene

change detection on MPEG-coded data domain”, in *Proc. SPIE Visual Comm. and Image Proc.*, pp. 992-1003, 1997

and Characterization for Video Databases”, *Pattern Recognition Special Issue: Image Databases*, 1997

- [6] Omer N. Gerek, Yucel Altunbasak, “Key Frame Selection from MPEG Video Data”, in *Proc. SPIE visual Comm. and Image Proc.* pp.920-925, 1997
- [7] Nilesh V. Patel, Ishwar K. Sethi, “Video Shot Detection

표 1. 테스트 영상에 대한 실험 결과

테스트 영상	전체 컷 수	정확히 검출된 컷 수	잘못 검출된 컷 수	놓친 컷 수	키 프레임 수
“Walle and Gromit” 중에서 (330 frames)	8	8	0	0	10
“Poltergeist” 중에서 (335 frames)	9	8	0	1	13
“Blade Runner” 중에서 (570 frames)	15	11	2	4	18

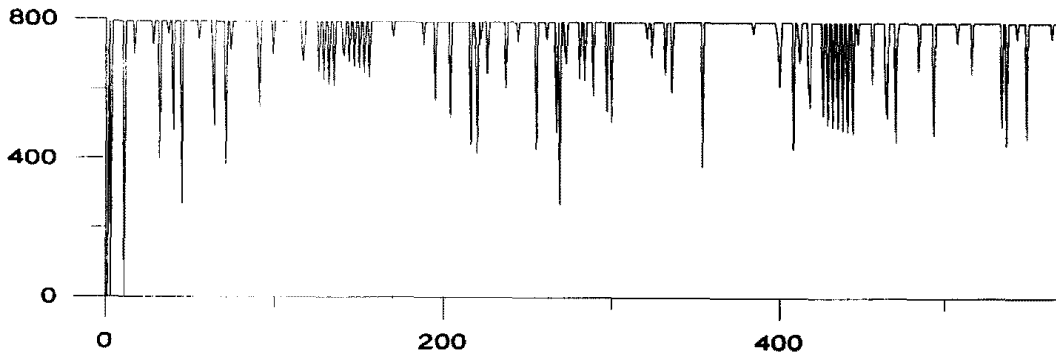


그림 2. “Blade Runner” 영상에 대한 누적 histogram intersection measure