

매크로블록과 프레임 정보를 이용한 컷 검출

전종대*, 정성일*, 류상률**, 김승호*
경북대학교 컴퓨터공학과*, 청운대학교 컴퓨터학과**

jdjeon@borami.knu.ac.kr, seongil@borami.knu.ac.kr
rsr@chungwoon.ac.kr, shkim@knu.ac.kr

Cut Detection Using Macroblock and Frame Informations

Joung-Dae Jeon*, Seong-Il Cheong*, Sang-Ryul Ryu**, Sung-Ho Kim*

Department of Computer Engineering, Kyungpook National University*
Department of Computer Science, Chungwoon University**

요 약

본 논문에서는 멀티미디어 데이터의 MPEG-1 비디오 데이터 정보 중 인트라 모드로 부호화된 매크로블록(macroblock) 타입 개수를 이용하여 후보 컷(cut)을 검출한다. 이렇게 검출된 후보 컷은 이전에 검출된 후보 컷과의 히스토그램 차와 화소 특징값의 분산값과 평균값을 이용하여 카메라 브레이크(camera break)와 같은 일반적인 장면전환(abrupt scene change)뿐만 아니라 커튼효과(curtaing effect)나 페이딩효과(fading effect)와 같은 점진적인 장면전환(gradual scene change)을 검출하는 방법을 제안하고 실험을 통하여 그 유용성을 보인다.

1. 서론

정보전달 매체 중에서 가장 효율적인 방식은 영상과 음향을 이용한 멀티미디어 데이터이며, 그 중에서 대표적인 미디어인 비디오 데이터는 텍스트나 정지영상 등과 같은 정적 데이터에 비해 쉽고 빠르게 인지할 수 있는 장점을 가지고 있다. 그러나 매체에 저장된 비디오 데이터의 크기가 너무 방대하므로 특정 데이터를 순차적으로 검색하거나 텍스트 등의 특정 키워드로 검색하는 방식은 적합하지 않다. 특히 비디오 데이터에서 원하는 장면(scene)을 보다 빠르고 정확하게 검색하려면 비디오 스트림(video stream)을 장면 단위로 분할하여 다른 내용의 장면으로 전환되는 위치들을 효율적으로 모두 검출해내는 장면 전환 검출기술이 필요하다.

장면 또는 샷(shot)이란 한 번의 카메라 촬영으로 생성될 수 있는 정지영상 또는 프레임(frame)의 집합을 말한다. 서로 다른 내용의 연속 장면을 가정할 때, 이전 장면의 마지막 프레임과 다음 장면의 첫 프레임은 서로 다른 내용과 특성을 가지게 되므로 전환된 장면의 첫 프레임을 컷(cut)이라 한다[1-3]. 장면전환 검출이란 비디오 데이터의 장면들의 특성을 비교하여 장면변화가 발생하는 프레임을 검출하는 것이다. 검출 방식에는 크게 프레임 기반의 장면전환 검출방식[3]과 압축 비디오 데이터 기반의 장면전환 검출방식[2,4,5]으로 나눌 수 있다.

프레임 기반의 장면전환 검출방식은 히스토그램(histogram)과 같은 프레임의 특징 값(feature)을 사용하므로 압축 비디오 데이터 기반의 검출방식보다 섬세하게 장면전환을 검출할 수 있다. 그러나 물체나 카메라의 빠른 움직임과 순간적인 밝기의

변화와 점진적인 장면전환에는 민감하지 못하며, 장면전환을 검출하는데 소요되는 연산 시간이 많이 든다는 단점이 있다. 프레임 기반의 검출방식의 대표적인 예로서 화소(pixel) 단위의 비교방식[1]과 히스토그램의 비교방식[3] 등을 들 수 있다. 압축 비디오 데이터 기반의 장면전환 검출방식은 프레임 기반의 방식보다 연산 시간이 훨씬 적게 소요되는 장점을 가지고 있으나, 실질적인 프레임의 특징 값이 아닌 압축된 비디오 데이터의 특징 값을 활용하므로 장면전환 검출의 정확성이 낮아질 수 있다. 압축 비디오 데이터 기반의 검출방식의 대표적인 예로서 DCT(Discrete Cosine Transform) 계수 정보를 이용하는 방식[2]과 움직임 벡터(motion vector) 정보를 이용하는 방식[5]을 들 수 있다.

본 논문에서는 MPEG-1 형식의 비디오 데이터를 이용하여 압축 비디오 데이터 기반의 장면전환 검출방식과 히스토그램 차를 이용하는 장면전환 검출방식을 복합한 효율적인 장면전환 검출방식을 제시한다. 장면전환이 발생한 프레임에 대한 움직임 보상(motion compensation)이 무효하므로 16×16 화소 크기의 매크로 블록들이 인트라 모드(intra-mode)로 부호화 된다. 이때 부호화된 매크로 블록의 개수를 비교하여 장면전환의 후보 컷들을 검출한 후, 후보 컷의 밝기 정보인 Y성분에 대한 히스토그램 차를 구하여 장면전환을 구분하는 컷을 검출한다.

2장에서는 본 논문에서 제안하는 매크로 블록의 정보를 이용한 후보 컷 검출방식을 제시하고, 후보 컷들의 히스토그램 차를 이용하여 컷을 검출하는 방식을 제시한다. 3장에서는 실험 결과를 제시하고 비교 분석한다. 마지막으로 4장에서는 결

론을 제시하고 차후 연구방향을 제시한다.

2. 매크로블록 정보와 히스토그램 차를 이용한 검출방법

2.1 매크로블록 정보를 이용한 후보 컷의 검출

MPEG-1에서 규정하고 있는 프레임 형태에는 I, P, B가 있다. I 프레임은 움직임 보상을 이용하지 않고 해당 프레임만을 DCT하여 부호화 한 것이고, P 프레임은 I 프레임이나 다른 P 프레임을 참조하여 움직임 보상을 적용하여 발생하는 차(difference)를 DCT한다. B 프레임에서는 P 프레임처럼 움직임 보상을 사용하지만 P 프레임과 달리 시간 축 상에서 앞과 뒤에 위치한 두 개의 참조 프레임(I 또는 P)으로부터 움직임 보상을 한다. 움직임 보상 방향은 그림 1과 같다.

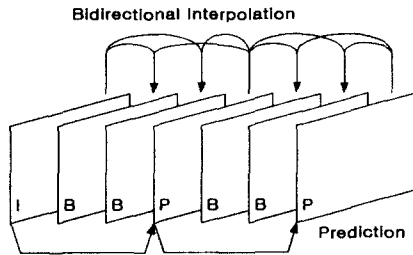


그림 1. 각 프레임의 움직임 보상 방향

MPEG-1 비디오 데이터는 매크로블록 단위로 부호화가 이루어지며, 각 프레임의 상황에 따라 매크로블록을 적절히 부호화 한다. 즉, 장면전환이 발생하여 움직임 보상이 무효한 경우나 해당 프레임 내에서 움직임 보상을 한 결과보다 인트라 모드로 부호화 하는 것이 더 좋은 경우는 프레임 종류에 관계없이 매크로블록을 인트라 모드로 부호화 한다.

본 논문에서는 인트라 모드로 부호화 된 매크로블록의 개수 θ 가 많은 프레임은 장면전환이 발생했을 가능성이 많으므로 해당 프레임을 후보 컷으로 검출한다. 검출된 후보 컷의 θ 는 히스토그램 차의 임계값 선택에 중요한 정보가 된다. 식 (1)은 후보 컷을 검출하기 위한 조건이다.

$$MB_{min} \leq \theta \tag{1}$$

임계값 MB_{min} 은 인트라 모드로 부호화 된 매크로블록의 최소 개수로 총 매크로블록 개수의 40%를 가정하였다. 후보 컷은 I 또는 P 프레임에서 많이 검출되었으며 B 프레임에서는 움직임 보상의 특성으로 인하여 후보 컷이 적게 검출되는 경향을 보였다.

2.2 히스토그램 차를 이용한 컷 검출

매크로블록의 정보를 MB_{min} 과 비교하여 검출된 후보 컷은 장면전환을 판단하는 절대적인 기준이 되지 않는다. 즉, 그림 2에서처럼 B 프레임에서 장면전환이 발생하면 B 프레임에 대한 움직임 보상의 특성상 P 또는 I 프레임과 같은 참조 프레임에서 약 40%이상의 매크로블록들이 인트라 모드로 부호화 되기 때문이다.

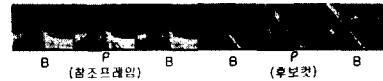


그림 2. 후보 컷과 장면전환 발생 프레임의 관계

그림 3의 컷 검출 프로세스는 후보 컷간의 밝기 정보인 Y 성분에 대한 히스토그램 차와 화소 특징 값의 분산값과 평균값을 이용하여 컷을 검출하는 과정을 제시한다.

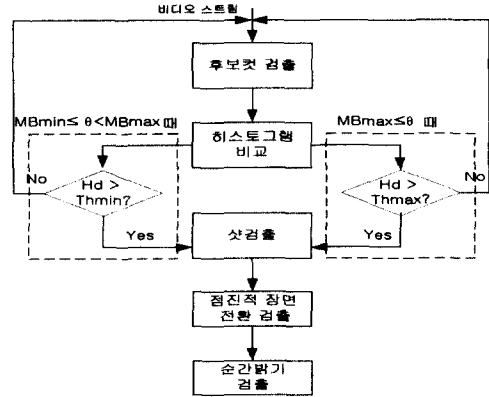


그림 3. 컷 검출 프로세스

매크로블록 정보로 검출된 후보 컷은 바로 이전에 검출된 후보 컷과의 히스토그램 차 H_d 를 구하여 장면전환의 발생 여부를 검사한다. 장면전환이 발생된 정확한 컷을 결정하기 위하여 후보 컷은 움직임 보상에 활용하는 참조 프레임(I, P)과 현재의 후보 컷 사이에 있는 프레임과의 히스토그램 차를 연속적으로 검사한다. 이때 사용되는 히스토그램 차의 임계값은 검사되는 후보 컷의 매크로블록 개수 θ 에 따라 달라지며 식(2)에서 임계값 적용 조건을 보여준다.

$$\begin{aligned} Th_{max} &: \text{if } MB_{max} \leq \theta \\ Th_{min} &: \text{if } MB_{min} \leq \theta < MB_{max} \end{aligned} \tag{2}$$

임계값 MB_{min} 은 후보 컷의 매크로블록의 40% 정도의 개수를 가지며 MB_{max} 는 후보 컷의 매크로블록의 70% 정도의 개수를 가진다. Th_{max} 는 후보 컷의 매크로블록 개수 θ 가 MB_{max} 이상일 때 적용되는 임계값이며, 이전 장면과 상당히 다른 장면을 검출할 수 있게 작은 임계값을 주었다. Th_{min} 는 θ 가 MB_{min} 와 MB_{max} 사이일 때 적용되는 임계값이며, 이전 장면과의 어느 정도 유사성을 가지는 장면전환도 검출할 수 있게 큰 임계값을 주었다.

히스토그램 차를 이용한 컷 검출은 일반적인 장면전환에는 좋은 결과가 나타나지만, 순간적인 밝기의 변화나 점진적인 장면전환에서는 같은 장면임에도 불구하고 컷으로 검출하는 오류를 범한다. 이러한 오류를 방지하기 위하여 후보 컷간의 화소 특징 값의 평균값과 분산값을 연속적으로 비교한다. 커튼 효과나 페이딩 효과가 발생하는 점진적인 장면전환에서는 샷의 프레임들이 점차적으로 단일 색 프레임으로 전환된 후, 새로운

샷의 프레임이 점진적으로 나타나는 특징을 가진다. 커튼효과와 순간적인 밝기변화 현상의 예는 그림 4에서 보인다

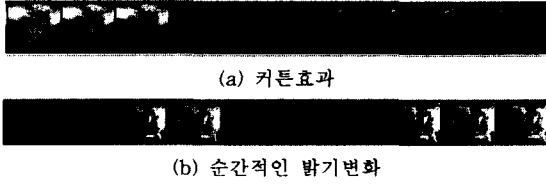


그림 4. 커튼 효과와 순간적인 밝기변화

점진적인 장면전환과 순간적인 밝기변화 현상의 검출은 후보 컷의 화소 특징값의 평균값과 분산값을 이용한다. 평균값은 후보 컷의 히스토그램을 이용하여 구할 수 있으며, 분산값 V 는 식(3)에서 구할 수 있다.

$$V = \frac{\sum_{j=0}^G H(j) \cdot (j - M)^2}{N} \quad (3)$$

$H(j)$ 는 후보 컷의 히스토그램 값을 나타내며, M 은 평균 화소값, N 은 프레임의 총화소수, G 는 화소값의 범위를 나타낸다. 점진적인 장면전환은 해당 후보 컷의 평균값과 분산값이 점진적으로 감소하여 그 값이 가장 작은 프레임이 점진적인 장면전환이 일어나는 부분이다. 따라서 후보 컷간의 분산값과 평균값을 감시하여 그 감소량이 임계값보다 작으면 점진적인 장면전환으로 보고, 임계값보다 크면 순간적인 밝기변화로 본다.

3. 실험 결과

본 논문에서 제안한 장면전환 검출기법의 성능을 평가하기 위하여 PentiumIII 600의 시스템 상에서 Visual C++ 6.0으로 구현하였으며 실험 동영상은 MPEG-1 형식의 비디오 데이터를 사용하였다.

표 1. 장면전환 검출방법 성능 비교

| 적용방식 동영상 (실제컷/프레임) | 히스토그램 차 | | 매크로블록타입 | | 제안방식 | |
|--------------------------|---------|--------|----------|--------|---------|--------|
| | T/ F | sec | T/ F | sec | T/ F | sec |
| 광고1(35/1358) | 33 / 2 | 154.02 | 39 / 4 | 47.36 | 35 / 0 | 66.64 |
| 광고2(34/1248) | 40 / 8 | 141.44 | 50 / 16 | 43.60 | 36 / 4 | 69.76 |
| 광고3(32/1341) | 38 / 7 | 151.43 | 51 / 19 | 47.23 | 35 / 3 | 71.83 |
| 순간불빛(38/2921) | 50 / 15 | 324.96 | 58 / 21 | 93.36 | 39 / 2 | 144.73 |
| 드라마(183/5083) | 186 / 4 | 575.99 | 208 / 26 | 171.43 | 184 / 5 | 283.67 |

대상 동영상은 일반적인 장면전환만을 가지는 영상(광고1)과 물체의 움직임이 큰 영상(광고 2), 순간적인 밝기변화를 가지는 영상(광고 3), 점진적 장면전환 중 커튼효과와 연속된 순간적인 밝기변화를 포함하는 영상(순간불빛)과 임의의 영상을 실험대상으로 사용하기 위해 드라마 영상을 사용하였다.

표 1은 적용방식에 따른 실험결과를 보여준다. T는 적용방식에 따라 검출된 컷을 의미하며, F는 검출된 컷 중 잘못 검출

된 컷을 의미한다. sec는 적용방식에 따른 컷 검출 수행 시간을 나타낸다. 히스토그램 차를 이용한 장면전환 검출결과는 매크로블록 검출방식에 비해 비교적 올바른 컷 검출을 보여주나 검출 소요 시간이 많이 걸리고, 매크로블록 타입 개수 방식은 히스토그램 차를 이용한 방식보다 검출 소요 시간은 현저히 줄어들었으나, 올바른 컷 검출에 문제점을 보이고 있다. 본 논문에서 제안한 방식은 이 두 방식의 단점을 보완하면서 기존의 방식보다 나은 결과를 보여주고 있다.

4. 결론 및 향후 연구 과제

본 논문에서는 MPEG-1 형식의 비디오 데이터를 이용하여 압축 비디오 데이터 기반의 장면전환 검출방식과 히스토그램 차를 이용하는 장면전환 검출방식을 복합한 효율적인 장면전환 검출방식을 제시한다. 제안된 방식은 장면들의 매크로블록의 정보를 이용하여 후보 컷을 검출하고, 후보 컷간의 밝기정보인 Y성분에 대한 히스토그램 차와 화소값의 평균값과 분산값을 이용하여 정확하게 컷을 검출하였다.

히스토그램 차를 이용하여 장면전환을 검출하는 방식은 많은 연산시간을 요구하는 단점이 있다. 그러므로 본 논문에서는 인트라 모드로 부호화 된 매크로블록 개수를 이용하여 장면 전환 가능성이 높은 후보 컷들을 검출하여 최종적으로 정확한 컷을 검출함으로써 연산 시간을 많이 감소시켰다. 그리고 압축 비디오 데이터에 매크로블록 타입 개수를 이용한 검출방식을 적용할 때 발생할 수 있는 오류를 보완하기 위하여 히스토그램 차를 이용하는 프레임 기반의 장면전환 검출기법을 적용하였다. 그 결과, 점진적인 장면전환이나 순간적인 밝기변화 장면 등에서도 좋은 결과를 보였다.

향후 연구방향은 적용했던 장면전환 검출방식들이 영상의 특성에 따른 임계값에 종속된 결과를 나타내므로 영상 특성에 따른 임계값 설정을 보완하는 연구를 제시한다.

참고 문헌

- [1] G. Ahanger and T. Little, "A Survey of Technologies for Parsing and Indexing Digital Video," *Journal of Visual Communication and Image Representation*, Vol. 7, No. 1, pp. 28-43, 1996.
- [2] N. Parel and I. Sethi, "Compressed Video Processing For Cut Detection," *IEE Proc, Video, Image and Signal Processing*, Vol. 143, pp. 315-323, 1996.
- [3] 이재연, 전병태, 배영래, "복수의 검출기를 사용하는 동영상의 샷 경계 자동검출 알고리즘," *정보과학회논문지(B)*, Vol. 25, No. 9, pp. 1352-1360, 1998.
- [4] H. Znag, C. Low and S. Smoliar, "Video Parsing and Browsing using Compressed Data," *Multimedia Tools and Applications*, Vol. 1, pp. 89-111, 1995.
- [5] 임영인, 남종호, "매크로블록 정보를 이용한 MPEG-1 비디오 스트림의 장면 전환 검출," *정보과학회논문지(A)*, Vol. 26, No. 4, pp. 527-538, 1999.