

압축된 비디오분할을 위한 평균값 비교 알고리즘

김영호^o, 이경화, 서석배, 성창우, 강대성
 동아대학교 전기전자컴퓨터공학부

The Mean Value Comparative Algorithm for MPEG Video Segmentation

Y. H. Kim^o, K. H. Lee, S. B. Seo, C. W. Seong, D. S. Kang
 School of Electrical, Electronic and Computer Eng., Dong-A Univ.
 E-mail: dskang@daunet.donga.ac.kr

요약

본 논문에서는 MPEG 비디오분할을 위한 평균값 비교 알고리즘을 제안한다. 장면 전환 검출은 MPEG 비디오 시퀀스에서 분할하는 가장 기본적인 작업이다. 일반적으로 많이 사용되는 분할 알고리즘은 이전 프레임과 현재 프레임을 비교하기 때문에 카메라의 움직임이나 화면 변화에 따라 오검출이 생기는 단점이 있었다. 본 논문에서는 기존 알고리즘에서 검출한 장면 전환 지점을 사용하여 이웃 프레임들과의 평균값 비교를 통해 재판별 하였다. 실험 결과 기존 알고리즘 보다 정확한 장면 전환 검출을 할 수 있었고, 오검출도 줄일 수 있었다.

I. 서론

최근 디지털 기술의 발전은 영상과 음성 데이터를 통신 및 컴퓨터와 결합하여 새로운 멀티미디어로 발전하고 있다. 이들 중에서 비디오와 관련된 연구로는 비디오 색인 및 검색에 관한 연구가 많이 진행되고 있다[1,2]. 비디오 데이터는 매우 복잡한 계층 구조를 가지는데 이를 효과적으로 표현하기 위해서는 여러 단계의 계층으로 나눌 필요가 있다. 비디오 데이터의 가장 기본이 되는 단위는 프레임(frame)이며 그 다음 상위 단

위는 필름이 끊기지 않고 시간적으로나 공간적으로 연속해서 촬영된 일정구간의 프레임들로 구성된 shot 이다. 비디오 데이터의 shot 경계 검출 즉, 장면 전환 검출은 연속되는 두 개의 shot 사이의 경계를 검출하는 것으로서 비디오 색인 및 검색의 첫 번째 단계로서 매우 중요한 역할을 한다. 최근 몇 년 동안 비디오 데이터의 장면 전환 검출에 관한 연구가 많이 수행되어 왔다[3-5]. 대부분의 연구는 화소 단위에 의한 검출방법, 부분 영역단위에 의한 검출방법 및 프레임단위에 의한 검출방법등 비압축 영역에서의 검출방법과 DCT 계수, 움직임 벡터 등을 이용한 압축 영역에서의 검출방법을 사용했다. 본 논문에서의 MPEG으로 압축된 비디오에 대해서 장면 전환 검출을 한다.

본 논문의 구성은 다음과 같다. I 장 서론에 이어 II 장은 MPEG 비디오 시퀀스에서 DC영상을 추출하고 III 장에서는 추출된 DC영상을 이용하여 새로운 장면 전환 검출 알고리즘을 제안하고 IV 장에서는 실험한 결과와 이에 대한 고찰을 하였고 V 장에서 결론을 맺었다

II. DC영상 추출

MPEG 비디오 시퀀스에서는 I(Intra), P(Predictive), B(Bidirectionally predictive) picture가 있다[6]. I picture는 inter frame 예측을 사용하지 않고 해당 화면 정보만으로 부호화하는 화면으로서 모든 MB(Macroblock) type은

본 연구는 정보통신부의 2001년도 대학기초연구지원사업의 지원으로 수행하였음.

intra frame이며, 채널 전환 시의 원영상 복구와 오류의 전파를 막기 위해 GOP(Group of Picture)내에 최저 한 장의 I picture를 필요로 한다. P picture는 I 혹은 P picture로 부터의 순방향 움직임 보상 예측 수행으로 생기는 화면으로서 MB type은 intra frame, forward inter frame, backward inter frame, interpolative inter frame 예측부호화가 수행된다. B picture가 삽입됨으로써 화면 처리 순서가 원 화면 순서와 달라져서 부호기에서는 B picture를 건너 뛰어 다음의 I, P picture를 우선 부호화하고, 그 후 사이에 있는 B picture를 부호화 한다. MPEG 비디오의 시퀀스층은 다수개의 GOP로 되어있다. 각 GOP는 다수개의 picture로, picture층은 다수개의 슬라이스로, 슬라이스층은 다수개의 MB로, MB는 휘도 신호 블록과 색차 신호 블록으로 구성되어 있다. 보통 한 GOP당 15개 이하의 picture들로 구성된다. 이때 인트라 부호화인 I picture에서 각 슬라이스를 추출하고 슬라이스의 MB에서 각 8×8 DCT영역의 DC성분을 추출한다. 이때 한 I picture당 1개의 DC영상이 생성된다.

본 논문에서는 I picture에 대해서 DCT DC 계수를 구하고 이를 정규화 시킨 후 DC영상을 구성하였다.

III. 장면 전환 검출 알고리즘

1. 기존의 장면 전환 검출 알고리즘[7]

MPEG 비디오 시퀀스의 I picture에서 DC영상을 추출하여 다음 4개의 feature들을 이용하여 장면 전환 검출을 하였다. 첫 번째 feature는 이전 영상과의 pixel간의 차분값으로서 서로 다른 영상의 유사도를 측정하는데 가장 기본이 되며, 전체적인 휘도 변화를 나타낸다. 이 feature를 구하는 수식은 아래와 같다.

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

여기서 M은 DC영상의 열의 개수이고, N은 행의 개수이다.

두 번째 feature는 일반적으로 사용되어 지고 있는 DC영상의 히스토그램에 대한 chi-square 값이다. 이전 영상에 대한 히스토그램의 변화를 나타낸다. 이 feature를 구하는 수식은 아래와 같다.

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

여기서, $H_i(k)$ 는 i 번째 DC영상 히스토그램의 k 번째 bin의 값을 나타낸다. n은 8bit 영상을 사용할 경우 256이다.

세 번째 feature는 이전 DC영상과 현재 DC영상과의 히스토그램 분산의 차분값이다. 이 feature는 히스토그램의 전체적인 분포에 대한 변화를 나타낸다. feature를 구하는 수식은 아래와 같다.

$$DiffD_i^2 = \frac{(D_{i-1} - D_i)^2}{D_i} \quad (3)$$

여기서, D_i 는 i 번째 DC영상의 히스토그램 분산 값이다.

네 번째 feature는 양자화한 영상의 히스토그램의 각 bin값들의 열과 행의 위치에 대한 분산의 chi-square 값이다. 이 feature는 히스토그램의 bin 값들을 이용함으로써 객체의 움직임에 강인하고, 각 bin 값들의 열과 행의 위치에 대한 분산의 chi-square 값을 구함으로써 칼라의 변화에 둔감하다. 이 feature를 구하는 수식은 아래와 같다.

$$PX_i^2 = \sum_{k=0}^{n-1} \frac{(\rho X_{i-1}(k) - \rho X_i(k))^2}{\rho X_i(k)} \quad (4)$$

$$PY_i^2 = \sum_{k=0}^{n-1} \frac{(\rho Y_{i-1}(k) - \rho Y_i(k))^2}{\rho Y_i(k)} \quad (5)$$

$$\rho X_i(n)^2 = \epsilon(|BX_i(n) - \overline{BX_i(n)}|^2) \quad (6)$$

$$\rho Y_i(n)^2 = \epsilon(|BY_i(n) - \overline{BY_i(n)}|^2) \quad (7)$$

식(4)는 열에 대한 분산의 chi-square 값이고, 식(5)는 행에 대한 분산의 chi-square 값이다. 식(6)은 i번째 DC영상의 n 번째 bin의 열에 대한 분산이고, 식(7)은 i번째 영상의 n 번째 bin의 행에 대한 분산이다. 이상과 같이 구해진 DC영상의 feature들로부터 아래와 같은 단계로 장면 전환 검출을 한다.

Step 1. 각 파라미터들의 전체 프레임에 대한 평균을 구한다.

$$\overline{DiffImg}, \overline{X}, \overline{DiffD}, \overline{PX}, \overline{PY}$$

Step 2. DC영상의 각 feature 값이 아래의 조건식에서 condition 1과 2를 동시에 만족하면서 condition 3 혹은 4가 만족할 경우 장면 전환으로 검출한다.

- condition 1. $DiffImg_i > \alpha \overline{DiffImg}$
- condition 2. $X_i > \beta \overline{X}$
- condition 3. $DiffD_i > \gamma \overline{DiffD}$
- condition 4. $PX_i > \delta \overline{PX}$ and $PY_i > \delta \overline{PY}$

여기서, 각 상수 $\alpha, \beta, \gamma, \delta$ 는 실험을 통해서 2.0, 2.0, 2.0, 2.0으로 설정되었다.

2. 제안한 알고리즘(평균값 비교 알고리즘)

I picture로부터 DC영상을 각각 추출한다. 기존의 4개의 feature를 순차적으로 비교하여 구한 두 DC영상이 장면 전환 지점인가를 판별해야 한다. 이는 장면 전환 지점의 이웃 프레임들의 특징치의 평균을 장면 전환 지점의 특징량과 비교함으로써 이루어진다. 카메라 정지 등에 의해 장면 전환이 일어나는 지점은 급격한 화면의 변화가 일어나므로 높은 평균값을 갖는다. 하지만 장면 전환 내의 프레임 중에서도 물체의 빠른 움직임, 큰 물체의 움직임, 카메라의 움직임, 카메라의 후레쉬의 섬광 등은 프레임 내 화소값에 크게 영향을 주므로 DC영상도 많이 변하게 된다. 이때 기존 알고리즘은 위와 같이 큰 물체의 움직임 또는 물체의 빠른 움직임 등의 프레임을 장면 전환으로 오검출 할 수 있다. 따라서 제안한 알고리즘은 이를 해결하기 위해 기존에 장면 전환 지점으로 판별된 장면 전환 지점을 실제의 장면 전환 지점인지 아닌지 이웃 프레임의 DC영상과 비교하여 재판별한다. 실제 장면 전환이 일어나는 지점은 이전 프레임들과 이후 프레임들은 크게 차이가 난다. 하지만 장면 전환 지점을 기준으로 이전의 프레임들과 이후의 프레임들은 각각 비슷하다. 또 실제로 장면 전환 지점이 아닌 지점은 각 프레임의 특징치가 조금 크다 하더라도 이웃의 프레임들과 특징 부분을 제외하고는 비슷하다. 그러므로 기존의 두 비교 프레임의 전과 후의 프레임을 선정하여 총 네 개의 프레임을 이용해 장면 전환 지점의 이웃 프레임들과의 특징치의 평균을 구하여 장면 전환 지점의 특징치와 비교하여 실제 장면 전환 지점인지 아닌지 판별한다. 두 프레임이 아닌 한 프레임 전이나 한 프레임 후의 프레임과 특징치를 각각 비교하여 평균값을 취하면 해당 프레임들의 평균 특징치가 된다. 이때 이 평균값보다 장면 전환 지점의 특징치가 크다면 실제의 장면 전환 지점으로 판별한다.

$$F_{i,i+1} \geq M_{i,i+1} : \text{장면 전환 지점}$$

$$F_{i,i+1} < M_{i,i+1} : \text{장면 전환 지점 아님}$$

$F_{i-1,i+1}$ 와 $F_{i,i+2}$ 의 비교 측정치를 이용하여 $M_{i,i+1}$ 을 구해서 $F_{i,i+1}$ 와 $M_{i,i+1}$ 을 비교한다. $F_{i,i+1}$ 이 평균치, $M_{i,i+1}$ 이상이라면 장면 전환 지점으로 판정한다.

IV. 실험결과 및 고찰

본 논문에서는 MPEG 비디오 시퀀스로부터 DC영상을 추출하여 기존 각 feature들에 의해 구하여진 장면 전환 검출 지점을 새로이 제안하는 알고리즘을 사용하여 이웃 프레임들간의 평균값의 비교를 통해 장면 전환 검출을 재판별 하였다. 실험으로 사용한 비디오는 352 × 240 크기를 갖는 7740 프레임(핑클 "영원한 사랑")과 8810 프레임(김민중 "비원")을 갖는 뮤직비디오 2편에 대해 256 그레이 레벨 해상도의 휘도 성분을 갖고 실험한 결과에 대해 기술한다. 그림 1은 MPEG 비디오 시퀀스로부터 구성된 DC영상을 나타낸다. (a)는 원 영상이고 (b)는 DC영상이다.



(a) (b)
 그림 1. MPEG 비디오 시퀀스로부터 구성된 DC영상
 (a) 원 영상, (b) DC영상





그림 2. 장면 전환 지점으로 검출된 프레임

그림 2는 제안한 알고리즘에 의해 검출된 장면 전환 프레임이다.

표 1. 장면 전환 검출 결과
뮤직비디오 1

	총 개수	검출개수	오검출	비율
기존 알고리즘	51	70	19	37.2%
제안 알고리즘	51	58	7	13.7%

뮤직비디오 2

	총 개수	검출개수	오검출	비율
기존 알고리즘	49	66	17	34.6%
제안 알고리즘	49	54	5	10.2%

표 1은 장면 전환 검출 결과를 나타낸 것이다. 뮤직비디오1에서는 실제 장면 전환이 발생하는 프레임 개수는 51개이고 뮤직비디오2에서는 49개였다. 기존 알고리즘에서 검출한 장면 전환 프레임은 뮤직비디오1에서는 70개, 뮤직비디오2에서는 66개이고, 새로이 제안한 알고리즘에서는 뮤직비디오1에서는 58개, 뮤직비디오2에서는 54개였다. 비율은 (오검출)/(실제 장면 전환 프레임 개수)*100으로 제안한 알고리즘이 뮤직비디오1에서는 37.2%에서 13.7%로, 뮤직비디오2에서는 34.6%에서 10.2%로 오검출을 줄일 수 있었다.

V. 결론

본 연구에서는 MPEG으로 압축된 비디오의 색인 및 검색을 위한 비디오의 최소 단위인 프레임들을 shot 단위로 분류하기 위한 새로운 알고리즘을 제안하였다. shot 경계 지점 즉, 장면 전환 지점을 검출하기 위해서 기존의 feature들을 이용하여 구해진 두DC영상을 바탕으로 이웃의 프레임들을 비교하여 좀 더 정확한 장면 전환 검출을 하였다. 향후 이러한 장면 전환 검출로 구해진 shot 단위 프레임들을 이용해서 MPEG 비디오의 색인 및 검색에 이용하는 연구가 필요하다.

VI. 참고문헌

- [1] "Efficient Scene Change Detection and Camera Motion Annotation for Video Classification," Wei Xiong and Chung-Mong Lee, Computer Vision and Image Understanding Vol.71, Nos 2/2, August, pp. 166-181, 1998.
- [2] "A hierarchical Multiresolution Video shot Transition Detection Scheme," Hong Heather Yu, Computer Vision and Image Understanding Vol.75, Nos 1/2, July/August, pp. 196-213, 1999
- [3] S. W. Smith and H. J. Zhang, "Content-Based Video Indexing and Retrieval", IEEE Multimedia, summer 1994, pp 62-72, 1994.
- [4] Boon-Lock Yeo and Bed Liu, "Rapid Scene Analysis on Compressed Video", IEEE Transaction on Circuits and Systems for Video Technology, Vol. 5, No. 6, pp 533-544, 1995.
- [5] Shih-Fu Chang, "Compressed-Domain Techniques for Image/Video Indexing and Manipulation", IEEE ICIP-95, Vol. 1, pp 314-317, 1995.
- [6] "그림으로 보는 최신 MPEG," 후지와라 히로시, ASCII Corporation.
- [7] 강대성, 김영호 한국신호처리·시스템학회 논문지 제 2권 1호 2001.1 .pp 38-43