퍼지 클러스터링 기법을 이용한 MPEG 비디오의 장면 전환 검출

임 성 재, 김 운, 이 배 호
전남대학교 컴퓨터공학과
전화 : 062-530-0719

Shot Change Detection Using Fuzzy Clustering Method on MPEG Video Frames

Seong-Jae Lim, Woon Kim, Bae-Ho Lee
Dept. of Computer Engineering, Chonnam National University, RRC
jaeslim@ciscom.chonnam.ac.kr

Abstract

In this paper, we propose an efficient method to detect shot changes in compressed MPEG video data by using reference features among video frames. The reference features among video frames imply the similarities among adjacent frames by prediction coded type of each frame. A shot change is detected if the similarity degrees of a frame and its adjacent frames are low. And the shot change detection algorithm is improved by using Fuzzy c-means (FCM) clustering algorithm. The FCM clustering algorithm uses the shot change probabilities evaluated in the mask matching of reference ratios and difference measure values based on frame reference ratios.

1. 서론

최근 장면 전환 검출에 관한 연구는 압축 영역 기법(히스토그램, 픽셀간 차이 등의 frame difference기법)과 비압축 영역 기법(DCT 계수, Motion vector 등)으로 크게 나누어져 활발히 연구되고 있다. 비압축 영역에서의 기법중 픽셀간 차이를 이용한 기법은 물체나 카메라의 움직임이 큰 경우에 성능이 저하되고, 히스토그램 차이를 이용한 작업은 장면이 비슷한 칼라 분포나 밝기 분포에서 성능이 저하되는 단점이 있다.[4] 또한 데이터를 복원하여 각 프레임마다 연산을 수행해야 하므로 많은 시간과 저장공간의 낭비를 가져온다. 압축 영역 기법 중 픽셀값 적용 방법은 조명의 변화에 민감하고, DC 성분에서 추출된 에지를 이용하는 방법은 뚜렷한 에지 검출이 어렵기 때문에 정확한 장면 전환 검출을 기대하기가 힘들다.[1] 이러한 단점들을 보완하기 위해서 본 논문은 압축된 데이터에서 각각의 프레임들의 macroblock(MB) 부호화 타입을 이용하여 프레임들의 유사성을 측정함으로 장면 전환을 검출하는 기법을 제안한다. 프레임들간의 유사성을 이용해 부호화된 MB의 타입을 기반으로 각각의 참조율을 구한다. 이렇게 구한 참조율은 mask matching 단계를 거쳐 장면 전환 확률값으로 정규화된다. 이렇게 정규화 된 장면 전환 확률값과 프레임들간의 유사성을 이용한 프레임 참조 타입의 차이 척도 값을 기반으로 FCM 클러스터링 기법을 이용하여 장면 전환을 검출하는 알고리즘을 제시한다.

본 논문의 구성은 다음과 같다. 2장에서 프레임 참조율을 이용한 장면 전환 검출 알고리즘에 대해 설명하고, 3장에서 실험 및 결론으로 구성한다.

2. 본론

2.1 프레임 참조 특성 분석

압축된 MPEG sequence는 MB의 참조 패턴에 따라 프레임 내 부호화 I, 프레임간 순방향 예측 부호화 P, 쌍방향 예측 부호화 B 프레임으로 구성되어 있다. I 프레임의 모든 MB들은 intra 부호화되기 때문에 다른 프레임의 참조 없이 독립적으로 부호화 된다. P 프레임은 이전의 I, P 프레임과 유사한 영상 패턴이 있을 때 MB들은 순방향 참조한다. 또한, B 프레임은 쌍방향 참조를 할 수 있어 이전에 위치한 I, P 프레임과 유사한 영상 패턴이면 순방향 참조를 하고 다음에 위치한 I, P 프레임과 유사한 영상 패턴이면 역방향 참조를 하게 된다. 이와같이 각각의 프레임은 전, 후의 프레임과 상관관계가 있다. MPEG 부호화 비디오에서는 I, P, B 프레임의 시퀀스와 수가 미리 정해져있다. 본 논문에서는

프레임 참조 특성 분석을 위해 IPB-ratio가 1:2:6의 형태를 갖는 부호화 시퀀스를 다루었다.

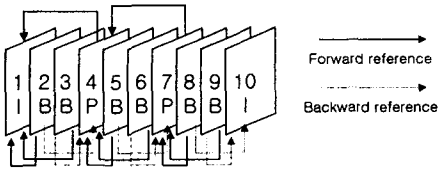

그림 1. 프레임 참조의 예

2.2 Mask Matching 기법에 의한 특징값 추출

인접한 프레임과의 유사성을 측정하기 위해 각 타입의 참조에 대한 MB의 수를 통해 참조율을 구할 수 있다. 참조율은 다음의 두 가지 타입으로 정의한다.

1. 순방향 참조율(FRR) = R_f/N, R_f는 프레임의 순방향 예측 부호화된 MB의 수, N은 프레임의 전체 MB의 수
2. 역방향 참조율(BRR) = R_b/N, R_b는 프레임의 역방향 예측 부호화된 MB의 수

P 프레임은 FRR값만을 갖고, B 프레임은 FRR, BRR 값을 갖는다. FRR값은 이전 프레임과의 유사 정도를 나타내며, BRR값은 다음 프레임과의 유사 정도를 나타낸다.

장면 전환이 발생하였을 경우 그림 2와 같이 이전 프레임들은 다음 장면에 낮은 BRR값을 갖고, 다음 프레임들은 이전 프레임들에 대해 낮은 FRR값을 갖게된다.

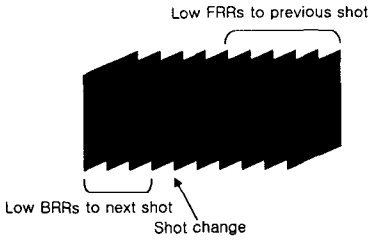

그림 2. 장면 전환 발생시 참조율 예

다음은 I, P, B 프레임에서 장면 전환이 발생하는 경우는 다음과 같다.

1. I 프레임에서 장면 전환한 경우 : I 프레임은 다른 프레임과 독립적으로 부호화되기 때문에 FRR, BRR 값을 갖지 않는다. 이때 I 프레임과 이전의 I, P 프레임 사이의 B 프레임들의 참조율을 계산 한다. 앞에 위치한 B 프레임들은 장면 전환이 발생한 현재의 I 프레임과는 유사한 패턴을 찾을 수가 없으므로 BRR의 값은 낮다.
2. P 프레임에서 장면 전환한 경우 : 이전의 I, P 프레임 사이의 B 프레임의 경우는 위와 같고, 이전의 I나 P 프레임에 대한 현재의 P 프레임의 FRR 값은 낮다.
3. B 프레임에서 장면 전환한 경우 : 현재의 B 프레임은 낮은 FRR 값을 갖게 되고, 이전의 I, P 프레임과 현재의 B 프레임 사이에 B 프레임이 존재했을 때, 그 프레임의 BRR 값은 낮게 된다. 다음에 위치한 I나 P 프레임과 현재의 B 프레임 사이에 B 프레임이 존재했을 때, 그 프레임의 FRR 값은 낮게 된다.

이렇게 구한 각각의 참조율 결과값들은 프레임의 장면 전환 가능성 여부를 결정하기 위해 mask matching 기법을 통해 정규화 한다.

Mask 프레임 M_i는 다음과 같이 표현된다.

$$M_i = FR, \quad where \quad F \in \{I, P, B\}, \quad R \in \{f, b\}$$

IBP-ratio가 1:2:6인 비디오의 Mask는 다음과 같이 네 가지 경우를 갖게 된다. M1은 I 프레임, M2는 P 프레임, M3, M4는 B 프레임에서 장면 전환이 발생했을 경우의 Mask를 나타낸다.[3]

$$M1 = \{I;\ (Bb, Bb, @I)\};$$
$$M2 = \{P;\ (Bb, Bb, @Pf)\};$$
$$M3 = \{B;\ (@Bf, Bf, Pf)\ or\ (@Bf, Bf, I)\};$$
$$M4 = \{B;\ (Bb, @Bf, Pf)\ or\ (Bb, @Bf, I)\};$$

위에서 @표시가 있는 Mask 프레임은 현재 프레임을 가리킨다.

장면 전환 가능성 여부를 결정하기 위한 mask matching의 결과값은 각 프레임의 장면 전환 확률을 나타내기 위한 값으로 정규화된다. 장면 전환 확률 함수 P는 다음과 같다.

$$P = 1 - \frac{\sum_{k=1}^{n} RR_{f_k}^{\ 2}}{\sum_{k=1}^{n} RR_{f_k}} \qquad (1)$$

f_k는 현재의 mask 프레임을 나타낸다. RR_{f_k}는 mask 프레임 f_k의 참조율을 나타낸다.

장면 전환 확률값은 0과 1사이에서 결정되며, 그 값이 커지면 그 프레임에서 장면 전환이 발생할 가능성이 높아진다.

2.3 차이 척도 측정에 의한 특징값 추출

P 프레임에는 순방향 참조 타입과 인트라 타입의 매크로 블록이 존재한다. 만약 장면 전환이 발생하면 이전 프레임과 연속성이 없기 때문에 많은 매크로 블록이 인트라 부호화된다.

$$D_P = \frac{Intra_n}{R_{Fn}} \qquad (2)$$

R_{Fn}은 순방향 예측 부호화된 MB수이다.

B 프레임은 양방향 움직임 보상을 하므로 순방향, 역방향 참조 타입, 인트라 타입이 존재한다. 1:2:6인 시퀀스

에서 이전 B 프레임에서 장면 전환이 발생했을 경우는 다음 B 프레임의 참조 성향이 이전 B 프레임과 비슷하게 된다. 또한, 뒤의 B 프레임에서 장면 전환이 발생하게 되면 이전 B 프레임과 반대의 성향을 갖게 된다. 이러한 특징을 이용하여 프레임 참조 MB의 수를 이용한 다음과 같은 식으로 차이 척도값을 구한다.[5]

$$D_{B1} = \left\{ D_{Bc} = \left(\frac{R_{Bn}}{R_{Fn}}\right) \right\} \cdot \left\{ D_{Bn} = \left(\frac{R_{Bn}}{R_{Fn}}\right) \right\} \quad (3)$$

$$D_{B2} = \left\{ D_{Bp} = \left(\frac{R_{Fn}}{R_{Bn}}\right) \right\} \cdot \left\{ D_{Bc} = \left(\frac{R_{Bn}}{R_{Fn}}\right) \right\} \quad (4)$$

D_{B1}은 앞 B 프레임의 측정값, D_{B2}는 다음 B 프레임의 측정값, D_{Bc}는 현재 B 프레임, D_{Bn}은 다음 B 프레임, D_{Bp}는 이전 프레임, R_{Fn}은 순방향 예측 부호화된 MB수, R_{Bn}은 역방향 예측 부호화된 MB수이다. 위의 식으로 구해진 특징값들은 클러스터링을 위해서 정규화된다.

2.4 FCM 클러스터링 기법을 통한 장면 전환 검출

본 논문에서는 위에서 구한 특징값들에 FCM 클러스터링 기법을 적용하여 장면 전환 검출을 한다.

FCM 알고리즘은 클러스터 중심값과 원소 상호간의 가중 유사치 값을 기반으로 반복적 수행을 통해 objective function의 최적화를 추구하는데 objective function은 다음과 같이 정의된다.[6]

$$J_m(U, v) = \sum_{k=1}^{n} \sum_{i=1}^{c} (\mu_{ik})^m (d_{ik})^2 \quad (5)$$

(5)식에서 μ_{ik}는 i^{th}클러스터 중심과 k^{th}입력 데이터의 퍼지 소속값을 나타낸다. d_{ik}는 norm metric을 나타내며 Euclidean, Diagonal, Mahalanobis norm 등이 될 수 있다. m값은 지수 가중치이고, 수렴된 클러스터의 fuzzy한 정도를 나타낸다. v는 벡터값 형태의 클러스터 중심을 나타내는 행렬이고 U는 fuzzy c-partition이다.

FCM알고리즘은 다음과 같은 수행과정으로 구성된다.

1. $c(1 \le c \le n)$값, $m(1 \le m \le \infty)$값을 결정한다.
2. fuzzy c-partition $U^{(l)}$행렬을 초기화한다.
   (1) $\mu_{ik} \in [0,1]; 1 \le i \le c,\ 1 \le k \le n$
   (2) $\sum_{i=1}^{c} \mu_{ik} = 1; 1 \le k \le n$
   (3) $0 < \sum_{k=1}^{n} \mu_{ik} < n; 1 \le i \le c$
3. 클러스터 중심 $U^{(l)}$값을 이용하여 $v_i^{(l)}$구한다. 각 클러스터 i에 대한 클러스터 중심은 다음과 같다.

$$i = \frac{\sum_{k=1}^{n} (\mu_{ik})^m x_k}{\sum_{k=1}^{n} (\mu_{ik})^m} \quad (6)$$

4. $U^{(l)}$ 값을 갱신한다.

$$\text{if } x_k \neq v_i^{(l)}$$

$$\mu_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{d_{ik}}{d_{jk}}\right)^{\frac{2}{(m-1)}}} \quad (7)$$

$$\text{else,} \quad \mu_{ik} = 1$$

5. if $\|U^{(l)} - U^{(l+1)}\| \le \varepsilon$ *then* *STOP*
   else $l = l+1$, and return to step 3.

여기서 c : 클러스터의 수
m : 지수 가중치
ε : 허용 오차 임계값이다.

본 논문에서는 참조 타입을 이용한 두 특징값에 상기 FCM을 적용하여 장면 전환 검출을 한다. 장면 전환 존재 여부를 결정하기 위해 클러스터수(c)를 변화시켜 가면서 모든 프레임들이 각 클러스터에 포함되게 한다. 그런 다음 각 클러스터에 속한 각 프레임들을 그 클러스터의 중심값과의 Euclidean 거리값이 일정한 임계값을 넘으면 장면 전환 후보로 결정한다. 얻어진 두 특징값을 기반해 후보 프레임을 장면 전환 프레임으로 결정한다.

3. 실험 및 결론

다음은 ten.mpg의 실험 결과이다. 먼저 두 특징 값들의 다이어그램이다.

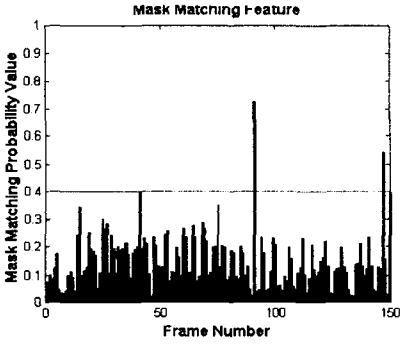

그림 3. Mask Matching Feature 값

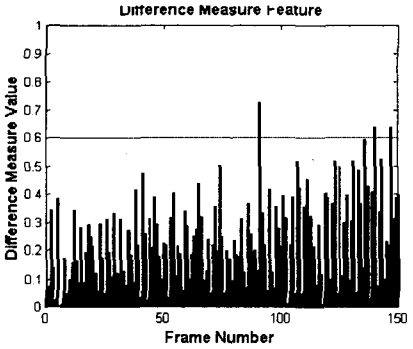

그림 4. Difference Measure Feature 값

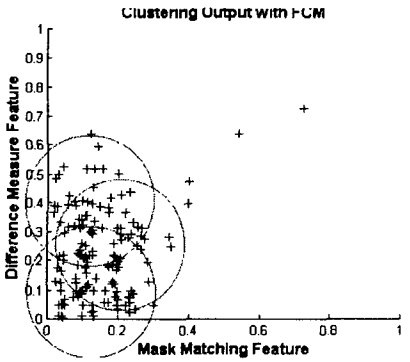

그림 5. 클러스터링 결과(c=3)

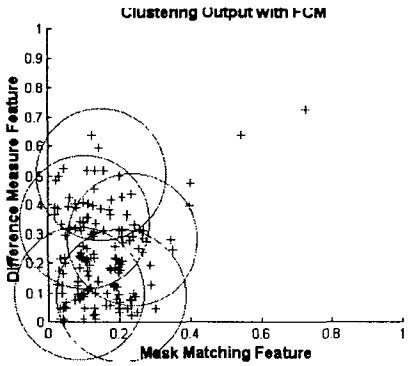

그림 6. 클러스터링 결과(c=5)

그림5와 그림6은 mask matching feature와 difference measure feature값을 이용해 FCM 클러스터링 한 결과이다. 그림5는 클러스터의 개수를 3으로 했을 때 각 클러스터에 포함되지 않은 4개의 프레임을 검출하였고, 그림6은 클러스터의 개수를 5로 했을 때 각 클러스터에 포함되지 않은 3개의 프레임을 검출하였다. mask matching probability value값에서 임계값 0.4일 때 3개의 프레임이 검출되었고, difference measure value에서 임계값 0.6일 때 3개의 프레임이 검출되었다. 이러한 결과를 기반으로 클러스터의 개수가 5개일 때 검출된 프레임이 장면 전환으로 검출하게 된다.

<표 1>은 네 가지 테스트 시퀀스 ten.mpg, laputa.mpg, jjackson.mpg, stennis.mpg 를 사용하여 실험한 결과이다.

표 1. 장면 전환 검출의 실험 결과

시퀀스	프레임수	실제 장면 전환	바른 검출	오검출	미검출
ten	150	3	3	0	0
laputa	56	3	3	1	0
jjackson	100	4	4	1	0
stennis	150	2	2	0	0
계	456	12	12	2	0

ten과 stennis는 평활한 영상에 어떠한 특수 효과가 첨가되지 않아 모두 바르게 검출하였고, laputa와 jjackson은 빠른 동작과 빛의 변화로 장면을 오 검출하는 결과가 발생하였다.

본 논문에서는 프레임간의 상관관계로부터 구한 두 특징값을 이용한 FCM 클러스터링 기법을 통해 장면 전환 검출을 하였다. 즉, 각각의 프레임간 참조율을 mask matching하여 확률값으로 정규화 시킨 특징값과 참조 차이 척도를 이용한 정규화된 특징값에 FCM 클러스터링 기법을 이용해 더 정확한 장면 전환 검출을 하였다. 일반적으로 하나의 특징값만을 이용하여 장면 전환 검출을 하는 경우엔 특별한 상황에서만 잘 적용되며, 임계값을 정하는 것이 문제가 된다. 제안한 기법은 이러한 문제점들을 보완하였으며, 복원한 데이터에서 장면 전환을 검출한 기법보다 더 효과적인 장면 전환 검출을 하였다.

앞으로 특수 효과가 사용된 다양한 데이터를 통한 실험으로 제안한 알고리즘의 성능 평가와 장면 전환 검출을 위한 다양한 특징값을 적용하는 연구를 수행할 것이다.

참 고 문 헌

[1] J. Meng, Y. Juan, and S. F. Chang, "Scene Change Detection in an MPEG Compressed Video Sequence", Digital Video Compression : Algorithm and Technologies, Vol. SPIE-2419, pp. 14-25, Feb. 1995

[2] R. Brunelli, O. Mich, and C. M. Modena, "A Survey of the Automatic Indexing of Video Data", Journal of Visual Communication and Image Representation, Vol. 10, pp.78-112, 1999

[3] Tony C. T. Kuo, Y. B. Lin, Arbee L. P. Chen, Shu-Chin Chen and C. Y. Ni, "Efficient Shot Change Detection on Compressed Video Data", Proc. of IEEE Workshop on Multimedia Database Management System, August 1996

[4] I. K. Sethi and N. Patel, "A Statistical Approach to Scene Change Detection", Storage and Retrieval for Image and Video Database III, vol. SPIE-2420, pp. 329-338, Feb. 1995.

[5] H. Zhang, C. Y. Low and S. V. Smoliar, "Video Parsing and Browsing Using Compressed Data", Multimedia Tools and Application, Vol. 1, pp. 89-111, 1995

[6] H. J. Zimmermann, *Fuzzy Set Theory and Its Applications*, Kluwer-Nijhoff Publishing, pp. 187-212, 1985