

MPEG 비디오 프레임에서 FCM 클러스터링 기법을 이용한 효과적인 장면 전환 검출

임 성 재, 이 배 호
전남대학교 컴퓨터공학과
jaeslim@ciscom.chonnam.ac.kr

Efficient Shot Change Detection Using Clustering Method on MPEG Video Frames

Seong-Jae Lim, Bae-Ho Lee
Dept. of Computer Engineering, Chonnam National University, RRC

Abstract

In this paper, we propose an efficient method to detect abrupt shot changes in compressed MPEG video data by using reference ratios among video frames. The reference ratios among video frames imply the degree of similarities among adjacent frames by prediction coded type of each frames. A shot change is detected if the similarity degrees of a frame and its adjacent frames are low. This paper proposes an efficient shot change detection algorithm by using Fuzzy c-means(FCM) clustering algorithm. The FCM clustering 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 검출은 히스토그램, 픽셀간 차이 등의 frame difference 기법을 이용한 비압축 영역에서의 작업과 DCT 계수, Motion vector 등을 이용한 압축 영역에서의 작업으로 나뉘어 연구하고 있다. 비압축 영역에서의 작업중 픽셀간 차이를 이용한 작업은 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 영역에서의 작업 중 픽셀값에 바탕을 둔 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 논문에서는 압축된 데이터에서 각각의 프레임들의 macroblock 부호화 타입을 이용하여 프레임들의 유사성을 측정함으로써 장면 전환을 검출하는 기법을 제안한다. 프레임들간의 유사성을 이용해 부호화된 macroblock의 타입을 보고 각각의 참조율을 구한다. 이렇게 구한 참조율은 mask matching 단계를 거쳐 장면 전환 확률값으로 정규화된다. 이렇게 구한 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. 2장에서는 프레임 참조율을 이용한 장면 전환 검출 알고리즘에 대해 설명하고, 3장에서는 실험 및 결론으로 구성한다.

2. 본론

2.1 프레임 참조 특성 분석

압축된 MPEG sequence는 macroblock의 참조 패턴에 따라 프레임 내 부호화 I, 프레임간 순방향 예측 부호화 P, 쌍방향 예측 부호화 B 프레임으로 구성되어 있다. I 프레임의 모든 macroblock들은 intra 부호화되기 때문에 다른 프레임의 참조 없이 독립적으로 부호화 된다. P 프레임은 이전의 I, P 프레임과 유사한 영상 패턴이 있을 때 macroblock들은 순방향 참조를 한다. 또한, B 프레임은 쌍방향 참조를 할 수 있어서 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 상관관

계가 있다. MPEG 부호화 비디오에서는 I, P, B 프레임의 시퀀스와 수가 미리 정해져 있다. 본 논문에서는 프레임 참조 특성 분석을 위해 IPB-ratio가 1:2:6의 형태를 갖는 부호화 시퀀스를 다루었다.

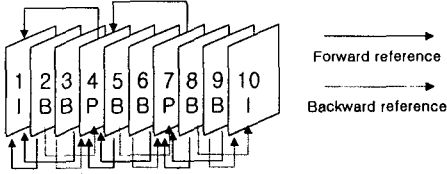


그림 1. 프레임 참조의 예

2.2 Mask Matching 기법에 의한 특징값 추출

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

1. 순방향 참조율(FRR) = R_f/N , R_f 는 프레임의 순방향 예측 부호화된 macroblock의 수, N 은 프레임의 전체 macroblock의 수
2. 역방향 참조율(BRR) = R_b/N , R_b 는 프레임의 역방향 예측 부호화된 macroblock의 수, N 은 프레임의 전체 macroblock의 수

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

장면 전환이 발생하였을 경우 그림 2와 같이 이전의 프레임들은 다음 장면에 낮은 BRR값을 갖고, 다음 프레임들은 이전의 프레임들에 대해 낮은 FRR값을 갖게 된다.

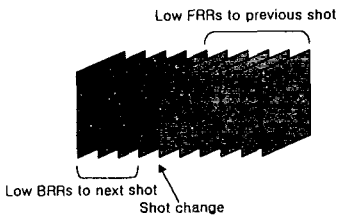


그림 2. 장면 전환이 발생한 경우의 참조율 예

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. 이 때 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 FRR의 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 그 값을 정규화 한다.

Mask 프레임 M_i 는 다음과 같이 표현된다.

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

IBP-ratio가 1:2:6인 비디오의 Mask는 다음과 같이 네 가지 경우를 갖게 된다. $M1$ 은 I 프레임, $M2$ 는 P 프레임, $M3, M4$ 는 B 프레임에서 장면 전환이 발생했을 경우의 Mask를 나타낸다.[3]

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

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 P 는 다음과 같다.

$$P = 1 - \frac{RR_1^2 + RR_2^2 + \dots + RR_n^2}{RR_1 + RR_2 + \dots + RR_n} \quad (1)$$

f_1, f_2, \dots, f_n 현재의 mask 프레임을 나타낸다. RR_{f_i} 는 mask 프레임 f_i 의 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 매크로블록이 존재한다. 만약 장면 전환이 일어나면 이전 프레임과 연속성이 없기 때문에 많은 매크로블록이 인트라 부호화된다.

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

R_{Fn} 은 순방향 예측 부호화된 macroblock 수이다.

B 프레임은 양방향 움직임 보상을 하기 때문에 순방향, 역방향 참조 타입, 인트라 타입이 존재한다. 1:2:6인 시퀀스에서 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 이용

하여 프레임 참조 macroblock의 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 macroblock 수, R_{Bn} 은 역방향 예측 부호화된 macroblock 수이다.

2.4 FCM 클러스터링 기법을 통한 장면 전환 검출

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

본 논문에서는 Fuzzy objective function 방법 중 Fuzzy c-means(FCM) 알고리즘을 적용하여 장면 전환 검출을 한다.

클러스터링은 주어진 데이터 집합의 패턴을 비슷한 성질을 가지는 그룹으로 나누기 위한 방법론이다. FCM 알고리즘은 클러스터 중심값과 원소 상호간의 가중 유사치 값을 기반으로 반복적 수행을 통해 objective function의 최적화를 추구하는데 objective function은 다음과 같이 정의된다.[6]

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

(5)식에서 μ_{ik} 는 i^{th} 클러스터 중심과 k^{th} 입력 데이터의 퍼지 소속값을 나타내며, 그 값은 fuzzy 분류법에서는 0과 1사이의 실수를 갖는다. d_{ik} 는 norm metric을 나타내며 Euclidean, Diagonal, Mathlanobis norm 등이 될 수 있다. m 값은 지수 가중치이고, 수렴된 클러스터의 fuzzy한 정도를 나타낸다. v 는 벡터값 형태의 클러스터 중심을 나타내는 행렬이고 U 는 fuzzy c-partition이다.

FCM 알고리즘은 다음과 같은 수행과정으로 구성된다.

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

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

5. $U^{(l)}$ 값을 갱신한다.

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

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

else, $\mu_{ik} = 1$ (8)

6. if $\|U^{(l)} - U^{(l+1)}\| \leq \epsilon$ then STOP

else $l = l + 1$ and return to step 4.

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. 본 논문에서는 장면 전환 존재 여부를 결정하기 위해 클러스터수(c)를 1로 잡는다. 먼저 프레임의 참조물에 기반한 mask matching 과정을 통해 얻은 특징값과 프레임 참조 차이 척도 특징값을 데이터 값으로 위 클러스터링 알고리즘을 적용하면 하나의 클러스터가 결정되고 이 클러스터를 중심으로 Euclidean 거리를 사용하여 장면 전환이 존재하는 프레임을 검출할 수 있다.

3. 실험 및 결론

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

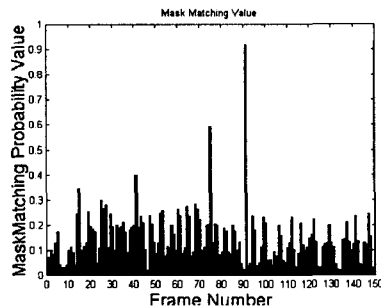


그림 3. Mask Matching Feature 값

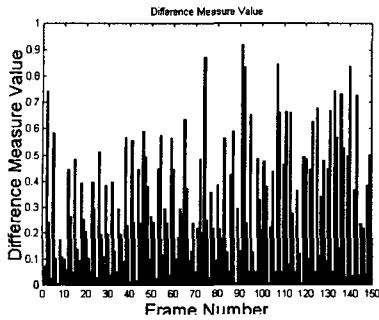


그림 4. Difference Measure Feature 값

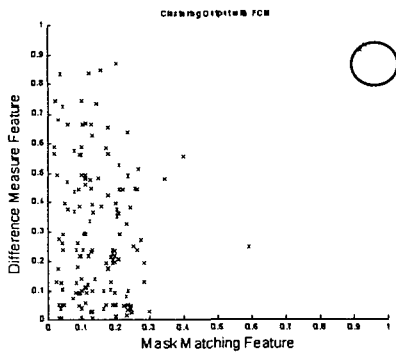


그림 5. 클러스터링 결과

그림5는 Mask Matching Feature와 Difference Measure Feature값을 이용해 FCM 클러스터링 한 결과이다. 그림에서의 검은점은 클러스터를 나타내고, 빨간 원안의 두 값이 장면 전환 검출된 프레임이다. 즉, FCM을 통해 구한 클러스터 중심값과 프레임간의 Euclidean 거리를 측정하여 일정한 임계값을 넘으면 장면 전환으로 검출한다.

<표 1>은 세 가지 테스트 시퀀스 stennis.mpg, laputa.mpg, jjackson.mpg를 사용하여 실험한 결과이다.

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

표 1. 장면 전환 검출의 실험 결과

시퀀스	프레임수	실제 장면 전환	바른 검출	오검출	미검출
stennis	150	2	2	0	0
laputa	56	3	3	1	0
jjackson	100	4	4	1	0
계	306	9	9	1	0

본 논문에서는 프레임간의 상관관계로부터 구한 두 특징값을 이용한 FCM 클러스터링 기법을 통해 장면 전환 검출을 하였다. 즉, 각각의 프레임간 참조율을 mask matching 하여 확률값으로 정규화 시킨 특징값과 참조 차이 척도를 이용한 특징값을 FCM 클러스터링 기법을 이용해 더 정확한 장면 전환 검출을 하였다. 이는 복원한 데이터에서 장면

전환을 검출한 기법보다 더 효과적인 장면 전환 검출을 할 수 있고, 하나의 특징값만을 이용하여 장면 전환 검출을 한 기법보다 더 정확한 결과를 얻을 수 있다.

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

감사의 글

본 논문은 정보통신분야 우수학교 지원사업 2000년도 전남대학교 연구과제사업 지원을 받아 수행된 것임.

참고 문헌

- [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