

계층적 클러스터링을 이용한 장면 전환점 검출

김중성, 홍승범, 백중환

한국항공대학교 정보통신공학과

전화 : 02-3158-5065 / 핸드폰 : 016-849-7822

Shot-change Detection using Hierarchical Clustering

Jong-Sung Kim, Seung-Pum Hong, Joong-Hwan Back

School of Electronics, Telecommunication and Computer Engineering Hankuk Aviation University

E-mail : kmjgsg@mail.hangkong.ac.kr

Abstract

We propose UPGMA(Unweighted Pair Group Method using Average distance) as hierarchical clustering to detect abrupt shot changes using multiple features such as pixel-by-pixel difference, global and local histogram difference. Conventional *k*-means algorithm which is a method of the partitional clustering, has to select an efficient initial cluster center adaptively. UPGMA that we propose, does not need initial cluster center because of agglomerative algorithm that it starts from each sample for clusters. And UPGMA results in stable performance. Experiment results show that the proposed algorithm works not only well but also stably.

압축 영역에서의 특징에는 DCT 계수를 이용한 방법과 움직임 벡터를 이용한 방법 등이 있다. 본 논문에서는 비압축 영역의 특징 중 PPD와 HD를 사용하였고 그 이유는 정규화에 대한 편이성 때문이다[2].

장면 전환점 검출을 위한 분류 방법에는 임계 값(threshold value)을 이용한 방법과 클러스터링을 이용한 방법이 있다. 임계 값을 이용한 방법에서는 임계 값을 정하기 위한 훈련 과정과 적용적으로 임계 값을 조절하여 검출 능력을 향상시킬 수 있지만 동영상의 종류에 따라 검출 능력의 변화가 심하다. 클러스터링은 각 샘플간의 거리를 측정하여 유사성에 의한 분류 방법이며 분류기를 훈련 할 필요가 없으며 전체 샘플들을 하위 그룹(subgroup)으로 분류 할 수 있다. 클러스터링에는 크게 계층적 클러스터링(hierarchical clustering) 과 분할식 클러스터링(partitional clustering)으로 나누어 볼 수 있다. 계층적 클러스터링에는 최소 거리측정 방법(single-linkage algorithm), 최대 거리측정 방법(complete-linkage algorithm), 평균 거리 측정 방법(average-linkage algorithm) or UPGMA:Unweighted Pair Group Method using Average distance), 최소분산 방법(Ward's method) 등이 있다. 또한 분할식 클러스터링 알고리즘에는 Forgy's 알고리즘, *k*-means 알고리즘, Iso-data 알고리즘 등이 있다[3].

I. 서론

장면 전환점 검출은 비디오 브라우징과 검색, 요약 등에 관한 많은 응용에 유용하다[1]. 샷(shot)은 연속적인 카메라 동작에서 만들어지는 프레임들의 시퀀스를 말하며 장면 전환에서는 프레임내의 내용이 급격하게 변화하는 특징이 있다. 대부분의 장면 전환점 검출 방법에 사용되는 특징량에는 크게 비압축 영역의 특징을 이용하는 방법과 압축 영역의 특징을 이용하는 방법이 있다. 비압축 영역의 특징량에는 프레임간 화소 값 차이(PPD:Pixel-by-Pixel Difference) 히스토그램 차이(HD:Histogram Difference)등이 있다.

본 논문에서는 계층적 클러스터링 알고리즘 중에서 평균거리측정 방법을 장면 전환점 검출 방법으로 제안하며 *k*-means 클러스터링과 검출 성능을 비교분석 하였다. 계층적 클러스터링은 수행시간이 비교적 오래 걸리는 단점이 있지만 *k*-means에서 초기 중심을 찾아야 한다는 단점에 비해 안정적인 검출 결과를 기대할

수 있다.

II. 특징량 추출

2.1 Pixel-by-Pixel Difference

화소 값 차(PPD)는 이전 프레임과 현재 프레임과의 각 픽셀에 대응하는 값의 차이를 특징량으로 한다. $N \times M$ 영상에 대하여 정규화된 PPD는 식 (1)로 구할 수 있으며 장면 전환점 검출에 있어서 좋은 성능을 보이지만 샷 내에서 객체 또는 카메라가 빠르게 움직일 때 장면 전환점이 아님에도 불구하고 장면 전환점으로 판단 할 수 있는 단점이 있다.

$$PPD = \frac{1}{(B \times N \times M)} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} |F_m(x, y) - F_{m-1}(x, y)| \quad (1)$$

$F_m(x, y)$: 프레임 m 의 (x, y) 에서의 화소 값

B : 최대 밝기 값

본 논문에서는 명암 성분(Luminance)에 대한 PPD만을 이용하였다.

2.2 Histogram Difference

히스토그램 차(HD)는 프레임간의 히스토그램의 차이를 의미하며 식 (2)로 정규화된 HD를 구할 수 있으며 PPD의 객체 또는 카메라의 빠른 움직임에 민감함을 보완하는 장점이 있다. 즉, 객체가 빨리 이동하여도 히스토그램은 큰 변화가 없다. 하지만 장면 전환이 있음에도 불구하고 명암 분포가 비슷할 때 동일한 샷으로 판단할 가능성이 있다. 본 논문에서는 명암(Luminance)과 색조(Hue)에 대한 히스토그램 차이를 특징량으로 이용하였다.

$$HD = \frac{1}{(2 \times M \times N)} \sum_{k=0}^B |H_m(k) - H_{m-1}(k)| \quad (2)$$

$H_m(k)$: 프레임 m 에서 k 의 히스토그램 분포

B : 최대 밝기 값

2.3 Local Histogram Difference

국부 히스토그램 차는 프레임을 일정한 크기의 블록으로 나누어 인접한 프레임간의 블록간 히스토그램 차를 구한 다음 하위 몇 개의 유사도를 갖는 블록들을 특징량으로 이용하는 방식이다. 이 특징량은 전역 히스토그램이 공간정보를 갖지 않음으로 인해서 비슷한 컬러 분포를 갖고 있으면 동일한 샷으로 판단하는 단점을 보완 한 것이다[4].

또한, 블록 단위로 유사도를 측정하고 유사도가 높은 상위 특징량을 제외시키기 때문에 부분적으로 발생하는 객체의 움직임에 덜 민감하게 반응한다. 본 논문에서는 320×240 의

프레임을 16개의 동일한 블록으로 상위 8개를 제외하고 하위 8개를 특징량으로 이용하였다.

$$LH_m = \frac{1}{8} \sum_{i=0}^7 \min[LHB_i]$$

$$LHB_i = \frac{1}{2|BS|} \sum_{k=0}^B |H_{m+1}(k, i) - H_m(k, i)|, 0 \leq i \leq 16 \quad (3)$$

$|BS|$: 블록내의 픽셀 수

III. 계층적 클러스터링(Hierarchical Clustering)

Unsupervised learning은 분류하고자 하는 샘플이 어떤 클래스에 해당되는지에 관한 사전지식이 없을 때 가능한 분류방법이다. 즉, 샘플에 대한 유사도를 측정하여 군집화하는 과정을 unsupervised learning이라 하며 클러스터링이 이에 해당한다. 클러스터링에는 크게 계층적 클러스터링과 분할식 클러스터링(Partitional Clustering)으로 나누어 볼 수 있으며 계층적 클러스터링은 하위에서 상위로 병합하는 병합식 또는 상위에서 하위로 분할하는 구분식으로 표현되어 질 수 있다[5]. 병합식 클러스터링의 알고리즘 중 최소/최대 거리를 유사도로 측정한다면 비정상 샘플에 대해 지나치게 민감하게 반응하는 경향이 있다. 따라서 이러한 문제점을 개선하기 위한 방법은 평균 거리를 유사도로 사용하는 UPGMA를 이용하는 것이다[6]. 분할식 클러스터링에는 클러스터의 중심이동(Moving Centers), 상호교환(Exchange), 연속배열(Seriation), Graph Partitioning 등의 방법이 있으며 가장 대표적인 MacQueen(1967)이 제안한 중심 이동 방식 중의 하나인 k -means 알고리즘을 본 논문에서는 비교하였다.

3.1 UPGMA(Unweighted Pair Group Method using Average distance)

병합식 클러스터링 과정은 그림 1의 트리 구조로 표현 할 수 있으며 알고리즘은 그림 2로 설명한다.

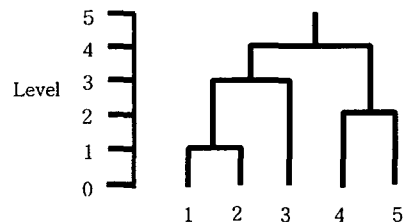


그림 1. 계층적 클러스터링의 구조

본 논문에서는 SC(Shot-change Cluster), NSC(No

SC) 두개의 클러스터이므로 아래 알고리즘에서 c 의 값은 2가 되며 n 은 샘플 수를 의미한다.

```

1. begin initialize  $c, c^* \leftarrow n, D_i \leftarrow \{X\}, i=1,2,\dots,n$ 
2. do  $c^* \leftarrow c^* - 1$ 
3.   가장 가까운 클러스터  $D_i, D_j$ 를 찾는다.
4.   클러스터  $i, j$ 를 병합한다.
5. until  $c = c^*$ 
6. return  $c$  클러스터
7. end
    
```

그림 2. 병합식 클러스터링의 알고리즘

UPGMA는 Single/Complete-linkage 알고리즘이 비정상 샘플에 민감함을 보완한 방법이며 클러스터간의 유사도를 식 (4)로 측정한다.

$$D_{avg}(C_i, C_j) = \frac{1}{n_i \times n_j} \sum d(x_i - x_j), \quad x_i \in C_i, x_j \in C_j$$

$d(x_i - x_j)$: 샘플 x_i, x_j 간의 거리 (4)

n_i : 클러스터 i 내에 있는 샘플 수

샘플 수 만큼의 클러스터로 시작하여 가장 작은 D_i, D_j 을 갖는 두 클러스터를 서로 병합하여 새로운 클러스터를 생성한다. 가장 작은 D_i, D_j 을 갖는다는 것은 두 샘플간의 유사도가 높다는 것을 의미하며 따라서 서로 병합하여 하나의 클러스터로 만들 수 있다. 이러한 과정을 하나의 클러스터, 또는 원하는 클러스터의 수만큼 생성 될 때까지 모든 샘플들을 서로 병합하여 계층적 표현이 가능하다.

3.2 k-means 알고리즘

k -means 알고리즘의 특징은 클러스터링을 수행하기 전에 초기 클러스터의 수 k 를 정해야 하는 것이다. k 는 보통 샘플 집단에서 처음 k 개의 샘플을 초기 클러스터로 선택되어진다. k 개의 클러스터로 분류한 다음 샘플들과 클러스터간의 유사도를 측정하여 가장 가까운 클러스터에 샘플을 분류하고 새롭게 생성된 클러스터의 중심을 갱신하여 다음 샘플에 대한 유사도를 측정하게 된다. 하지만 초기 클러스터의 수 k 의 설정이 결과에 많은 영향을 주며 초기 중심을 추정하기 위해 초기 클러스터링 과정을 수행해야 하는 단점이 있다[2]. k -means 알고리즘은 정해진 클러스터의 수에 대해서 식 (5)에 해당하는 평균 자승 오차 E 를 최소화하는 알고리즘이며 E 를 더 이상 감소시키지 못 할 때는 수행을 멈추게 된다. 평균 자승 오차가 Local minimum error에 빠져 클러스터링 결과가 나빠지는 경우가 있으며 이것은 초기 중심 k 를 설정하는 방법에 의존하여 해결 할 수 있다[3].

$$E = \sum_{i=0}^{N_j} \sum_{x \in S_i} (X - u_i)^2 = m\sigma^2 \tag{5}$$

$$u_i = \frac{1}{N_j} \sum_{x \in S_i} X$$

IV. 실험 결과 및 분석

UPGMA와 k -means 알고리즘에 대한 장면 전환점 검출을 비교하기 위해서 37개의 급격한 장면 전환점이 존재하는 다큐멘터리 비디오와 영화 매트릭스(48개의 장면 전환점), 미션 임파서블(30개의 장면 전환점)을 사용하였다.

4.1 k-means 알고리즘에 의한 결과

그림 3은 정규화된 특징 값 3개를 3차원 분포로 표현한 것이며 SC(Shot-change)에 해당하는 샘플들은 NSC(No shot-change)에 비해 큰 특징 값을 갖고 있다. 실제적으로 클러스터링을 수행한 것은 4개의 특징들을 이용하였으며 초기 클러스터링 값 k 는 2, 4로 각각 수행하였다.

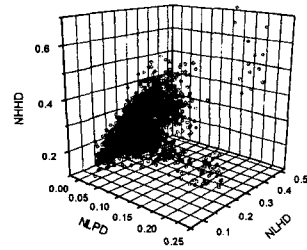


그림 3. 실험 영상 다큐멘터리의 정규화된 특징(LPD, LHD, HHD)값의 3차원 분포

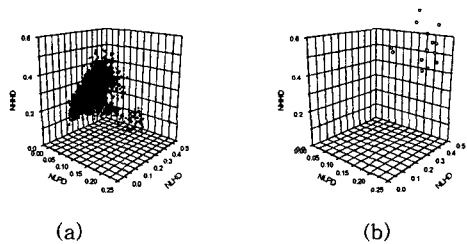


그림 4. k -means 알고리즘의 결과($k=4$) (a) NSC로 클러스터링된 샘플 (b) SC로 클러스터링된 샘플

그림 4에서 (a)는 장면 전환점이 없다고 판단된 클러스터에 속하는 샘플에 대한 분포이며 (b)는 장면 전환점이 있다고 판단된 클러스터에 속하는 샘플에 대한

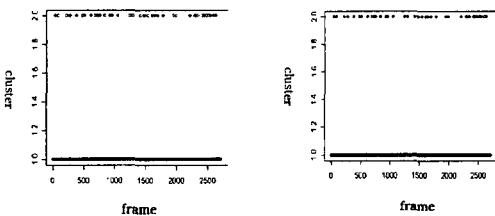
분포이다. 각각의 실험 영상에 관한 클러스터링 결과를 소환(Recall)/정확(Precision) 비율을 구하여 표 1에서 비교하였다.

표 1. 초기 클러스터 k에 따른 실험 영상별 성능 비교

실험영상 \ 성능	k=2		k=4	
	소환 비율	정확비율	소환비율	정확비율
다큐멘터리	1	0.032	0.97	0.94
매트릭스	0.83	0.701	0.479	0.522
미션 임파싸블	1	0.035	0.866	0.8387

표 1에서처럼 k에 따라 성능이 보장되어 있지 않았으며 비교적 k가 클수록 좋아지지만 영상에 따라서, 즉 특징 값의 분포에 따라 성능이 나빠지는 경우도 있다. 이러한 문제는 클러스터의 초기 중심을 적용적으로 변화하여 해결될 수 있지만 초기 클러스터링 과정을 수행해야 하는 단점이 있다.

4.2 UPGMA 알고리즘에 의한 결과



다큐멘터리 미션 임파싸블
 그림 5. UPGMA에 의한 샘플 분포

그림 5에는 SC, NSC에 관한 두개의 클러스터로 계층적 클러스터링을 수행한 샘플을 나타내고 있다. x축의 frame는 실험 영상의 프레임 번호를 나타내며, y축의 cluster=1은 NSC, cluster=2는 SC를 의미한다. 표 2는 클러스터링의 성능을 분석한 것이며 k-means 알고리즘에 비해 성능의 변화가 심하지 않았다. 매트릭스에서는 장면 전환점 검출 성능이 다소 떨어지는 이유는 동적인 영상이 많이 포함되어 있기 때문이며 다큐멘터리 영상과 같은 정적인 영상에서는 매우 좋은 성능을 보였다.

표 2. UPGMA에 의한 클러스터링의 성능

실험영상 \ 성능	소환 비율	정확 비율
다큐멘터리	1	0.9487
매트릭스	0.708	0.739
미션 임파싸블	1	0.857

IV. 결론

본 논문에서는 급격한 장면 전환점을 찾는 방법의 중점을 두었으며 계층적 클러스터링의 UPGMA 알고리즘을 이용하여 장면 전환점을 검출 하였다. k-means 알고리즘과 그 성능을 비교하여 볼 때 소환 비율은 비슷하나 그림 6에서처럼 정확 비율이 안정적이고 우수한 성능을 보였다. 비디오의 효과적인 요약(Summarization)을 위해서 점진적인 장면 전환점을 검출 할 수 있는 특징 정보 추출에 관한 연구가 필요하며 계층적인 트리구조에서 적절한 대표 프레임을 선정하는 연구가 추진되어야 할 것이다.

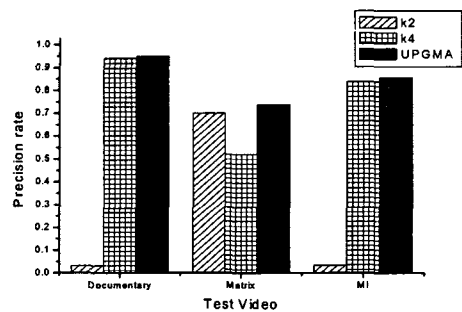


그림 6. k-means 알고리즘과 UPGMA의 정확비율

참고 문헌

- [1] Ullas Gargi 외 2명, "Performance characterization of video-shot-change detection methods", *IEEE Transaction on Circuits and Systems for video Technology*, vol. 10, pp. 1-13, 2000.
- [2] H.C.Lee 외 2명, "Abrupt Shot Change Detection Using an Unsupervised Clustering of Multiple Features", *IEEE* pp. 2015-2018, 2000.
- [3] Earl Gose and Richard Johnsonbaugh Steve Jost, "*Pattern Recognition and Image Analysis*", Prentice Hall, 1996.
- [4] G. Lupatini, C. Saraceno, and R. Leonardi, "Scene break detection: a comparison," *Proceedings of 8th International Workshop on Continuous-Media Databases and Application*, pp. 34 -41, 1998.
- [5] Boris Mirkin, "*Mathematical Classification and Clustering*", Kluwer Academic Publishers, 1996.
- [6] Richard O. Duda 외 2명, "*Pattern Classification*", A Wiley-Interscience Publication, 2001