

유사성 측정을 이용한 교양비디오의 개요 추출

이지현*, 문중환*, 진송철*, 이양원*

*군산대학교 컴퓨터학과

e-mail : jhlee@kunsan.ac.kr

The Abstraction of Cultural Videos Using Similarity Measures

Ji-Hyun Lee*, Jong-Hwan Mun*,
Song-Cheol Jin*, Yang-Won Rhee*

*Dept of Computer Science, Kun-San University

요 약

키 프레임들의 컬러 정보 차이 값과 표준편차 차이 값을 이용한 유사성측정 방법을 기반으로 클러스터링을 이용하여 의미 있는 단위로 분할되어 색인화 된 비디오에서 중요 장면들을 추출하고 추출된 키프레임들을 대상으로 중간(middle)샷 개요 생성 규칙을 적용하여 요약된 형태의 비디오 개요를 추출하여 사용자들에게 제공함으로써 비디오 선택의 폭을 넓힐 수 있도록 제안하였다.

1. 서론

의미적 내용 기반 모델(semantic content based model)[1]은 비디오 정보 시스템에서 가장 중요한 특징인 의미를 처리할 수 있는 방법이다. 현재 의미적 내용에 기반을 둔 모델은 많지 않으며, 의미적 묘사와 관련된 비디오 객체의 모델(video object model)을 다룬다. 이 방법은 비디오 데이터의 물리적 특징들 이면에 숨겨진 유연성과 역량 때문에 비디오 정보의 표현과 질의를 위한 의미적 기반을 마련하는데 효율적인 방법이다. 의미적 기반 모델은 크게 세그먼테이션 기반 모델(segmentation-based model)과 층화 기반 모델(stratification-based model)로 분류된다. 세그먼테이션 기반 모델은 먼저 비디오 스트림을 시간적인 순서에 따른 샷들의 집합으로 분할하고, 그들에 대하여 다양한 추상화(abstraction) 작업을 수행한다. 이 방법의 단점은, 유연성이 부족하며 중복된 세그먼트들에 존재하는 의미를 표현하는 것이 불가능하다는 것이다. 층화 기반 모델에서는 단순한 분할이 아닌 비디오의 구문 정보를 분할한다. 층(stratum)이라고 하는 비디오 단위들은 서로 중복되고 포함될 수 있으며 비디오 스

트림의 물리적 세그먼트와 일치하는 시간 간격과 관계된다. 현재는 대부분 기존의 층화 개념이 확장되어 다중 시간 간격과 관련된 이벤트를 묘사한다[2].

2. 관련연구

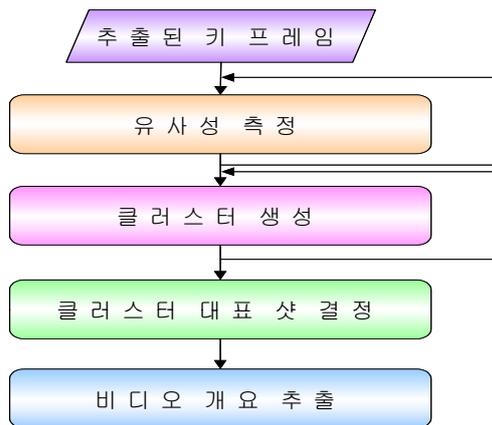
비디오 개요 추출에는 다큐멘터리나 뉴스 방송을 요약하기 위하여 비디오 스키밍(video skimming) 방법이 제시되었으며, 단어의 우선순위에 따라 비디오 클립들이 선택되어진다. Yeung 등은 장면 변화 그래프(scene transition graph)라는 스토리 흐름을 이용한 샷 기반 구조를 제시하였고, Hanjalic 등은 키 프레임을 추출하고 비디오 샷을 설정하며, 이들 키 프레임을 포함하는 비디오 요약 시퀀스를 생성하기 위하여 클러스터 유효성 분석(cluster validity analysis)을 사용하였다. Uchihashi 등은 코믹 만화 형태를 닮은 그림 형태의 비디오 요약 생성 방법을 제시하였으며, 이들은 세그먼트의 희소성과 지속성을 기초로 한 중요도 측정을 사용하였다. 또 다른 형태의 비디오 요약은 하이라이트 추출로써 신성운[3]등은 저수준 시각적/오디오 특징, 모션 정보, 그리고 컬러 정보를 탐색하여 무비 트레일러(movie

trailer)의 자동 생성 방법을 제시하였고, 이들을 디지털 비디오의 기본적인 물리적 파라미터들에 대한 경험적 특징을 사용하였다. Babaguchi[4]는 이벤트 기반 비디오 색인화 방법을 이용한 스포츠 비디오 요약 방법을 제시하였으며, 비디오 요약에 유용하긴 하지만 의미를 표현하는 중요한 특징들을 많이 유실한다.

따라서 비디오의 중요한 정보들을 유실하지 않고, 효율적으로 표현할 수 있는 새로운 비디오의 개요 추출 및 요약 방법의 개발이 요구됨에 따라 본 논문에서는 비디오의 중요한 정보들을 유실하지 않고, 개요 시퀀스를 효율적으로 생성할 수 있는 새로운 교양 비디오의 개요 추출 방법을 제시하고자 한다.

3. 제안된 비디오 개요 추출 방법

제안한 비디오 개요 추출 절차는 (그림 1)과 같이 나타낼 수 있다.



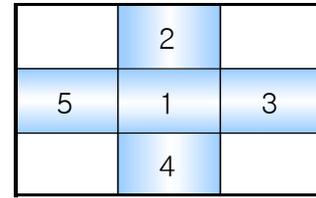
(그림 1) 제안된 비디오 개요 추출 과정

3.1 유사성 측정

클러스터링은 장면 전환 검출에 의하여 추출된 샷들에 대하여 유사성 측정에 의한 연속적인 병합을 통하여 유사한 샷들끼리 그룹화를 수행하는 것이다 [5]. 유사성 측정은 샷들을 구성하는 전체 프레임들을 비교하지 않고, 샷의 키 프레임들을 비교하여 수행한다.

컬러 히스토그램 공간에 대한 평균 중심점으로써 기준이 되는 영역에 대한 평균 컬러 값의 R·G·B 각각의 차이 합 값과 표준편차 차이 값을 비교하여 두 가지 조건을 모두 만족하는 조건으로 유사성을 측정하게 되는데 유사성 측정을 위한 영역은 프레임 전

체 영역을 대상으로 하지 않고, 일반적으로 프레임 내의 중요 정보가 중심영역을 기준으로 상·하·좌·우에 위치한다는 특징을 바탕으로 (그림 2)와 같이 비디오 프레임을 9개의 영역으로 분할하여 수행한다.



(그림 2) 기준 영역

중심 영역(1)에 일반적으로 카메라의 초점이 집중됨으로 60%의 가중치를 부여하고, 나머지 상(2)·하(4)·좌(5)·우(3) 영역에는 10%의 가중치를 부여하여 중심영역에 대한 중요도를 높임으로써 보다 정확한 유사성 측정을 할 수 있으며, 5개의 영역에 대한 R·G·B 컬러 평균값을 구하는 방법은 식 (1)과 같다. 식 (1)에서 C_a 와 C_b 는 이웃하는 두 샷의 키 프레임 a 와 b 에 대한 R·G·B 컬러 평균값을 나타낸다.

$$C_a = \frac{0.6C_{a_1} + \sum_{i=2}^5 0.1C_{a_i}}{5} \quad (1)$$

$$C_b = \frac{0.6C_{b_1} + \sum_{i=2}^5 0.1C_{b_i}}{5}$$

식 (2)는 두 키 프레임 a 와 b 에 대한 R·G·B 컬러 평균값의 차이를 계산하는 방법이다.

$$D(a,b) = |C_a - C_b| \quad (2)$$

5개의 영역에 대한 평균을 구하는 방법은 식 (3)과 같다. 식 (3)에서 평균은 표준편차를 구하기 위한 사전 계산으로 x_i 는 자료 값, f_i 는 도수, N 은 영역크기를 나타낸다.

$$\mu = \frac{\sum_{i=1}^n x_i f_i}{N} \quad (3)$$

5개의 영역에 대한 표준편차를 구하는 방법은 식

(4)와 같다. 식 (4)에서 σ_a 와 σ_b 는 이웃하는 두 샷의 키 프레임 a 와 b 에 대한 표준편차 값을 나타낸다.

$$\sigma_a = \sqrt{\frac{\sum_{i=1}^n x_i^2 f_i}{N} - \mu^2}$$

$$\sigma_b = \sqrt{\frac{\sum_{i=1}^n x_i^2 f_i}{N} - \mu^2}$$

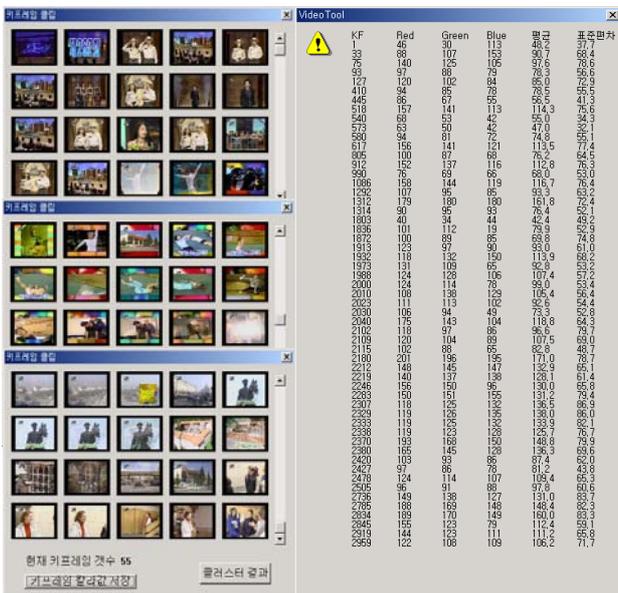
(4)

식 (5)는 두 키 프레임 a 와 b 에 대한 표준편차 값의 차이를 계산하는 방법이다.

$$\alpha(a,b) = |\sigma_a - \sigma_b|$$

(5)

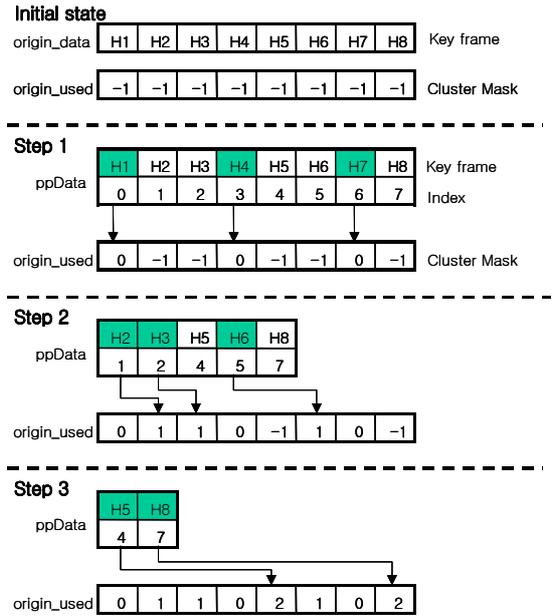
결국, R·G·B 컬러 차의 합 값을 이용한 유사성 거리 측정값은 ± 30 으로 설정하고, 키 프레임의 표준편차 값의 유사성 거리 측정값은 ± 5.0 으로 설정하여 적용한 결과는 (그림 3)과 같다.



(그림 3) 유사성 측정 값 산출 결과

3.2 클러스터링 생성

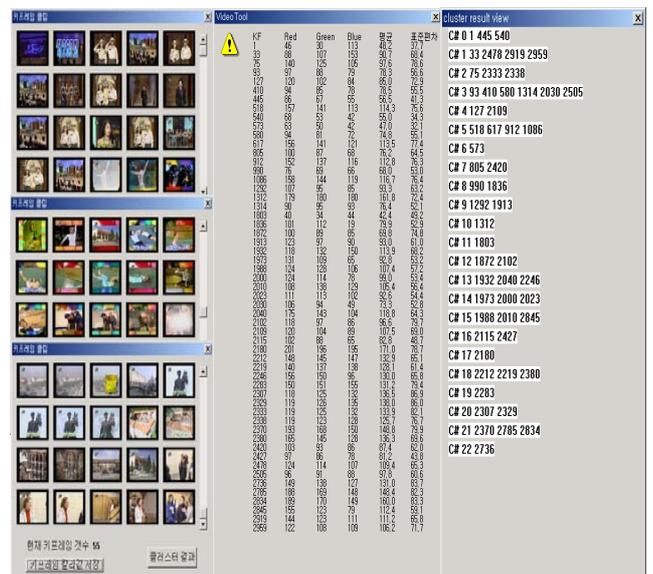
본 논문에서는 유사성 측정을 위한 비교 횟수와 기억 공간 할당을 단계별로 줄여가는 샷 병합 알고리즘을 이용한 클러스터링 방법을 이용하여 전체적인 샷 병합 과정은 (그림 4)와 같다.



(그림 4) 샷 병합 과정

(그림 4)에서는 8개의 샷들을 대상으로 유사성 측정을 위한 컬러 정보 값의 차이를 계산하여 샷들을 병합하는 과정을 통하여 최종적으로 클러스터 0, 1, 2의 세 개의 클러스터가 생성된 것이다. 또한 비교 횟수를 보면 Step 1에서는 7회, Step 2에서는 4회, Step 3에서는 1회로써, 단계별로 비교 횟수는 크게 감소함을 알 수 있고, 총 비교 횟수는 12회이며 단계별 평균 비교 횟수는 4회가 된다.

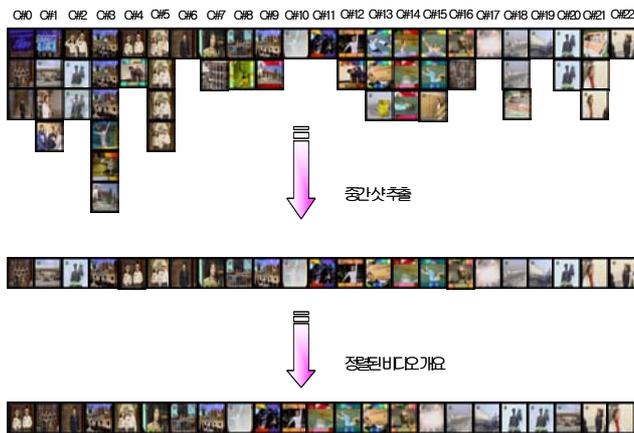
이와 같은 방법을 이용하여 샷들을 클러스터링 하면 (그림 5)와 같은 형태로 비디오가 클러스터 단위로 분할된다.



(그림 5) 클러스터링 결과

3.3 비디오 개요 생성

효율적인 비디오 요약을 위한 클러스터를 대표하는 샷들을 결정하기 위하여 첫 번째(first), 중간(middle), 마지막(end) 방법 중 전체적인 개요 파악이 용이 하며 중간 샷 계산(샷의 개수/2)과 정렬이 필요한 중간(middle) 방법[6]을 선택하여 적용한 결과는 (그림 6)과 같은 형태로 구성된다.



(그림 6) 중간 샷을 이용한 비디오 개요 추출

(그림 6)에서 나타내는 것처럼 각 클러스터를 구성하는 중간 샷들을 추출한 후 오름차순 정렬을 통하여 비디오 개요를 추출한다.

또한 “2001년 4월 1일 방영되었거나 방영 회차가 262회인 도전 지구탐험대의 개요를 검색하시오.”라는 질의를 통하여 (그림 7)과 같은 형태의 사용자 인터페이스를 제공한다.



(그림 7) 질의에 대한 검색 결과 브라우저

5. 결론

키 프레임의 열십자 영역 중 가운데 영역에 60%의 가중치를 부여하고, 나머지 4개의 영역에는 10%

의 가중치를 부여한 영역에 대하여 R·G·B 컬러 차이 합값(임계치 ± 30)과 표준편차 차이 합값(임계치 ± 5)을 모두 만족하는 조건의 유사성 측정을 통하여 유사한 키 프레임들의 클러스터링을 생성하고, 중간 샷 개요 생성 규칙을 적용하여 개요를 추출하였다. 추출된 비디오 개요는 장시간의 비디오를 시청할 필요가 없기 때문에 시청 시간을 줄일 수 있으며 전체 비디오를 시청하고자 원하는 사람들에게는 비디오 선택의 폭을 넓혀줄 수 있다.

참고문헌

- [1] Y. Day, S. Dagtag, M. Iino, A. Khoaklhar and A. Ghafoor, "A Multi-Level Abstraction and Modeling in Video Databases," *ACM Springer-Verlag Multimedia Systems*, 7(6), 1999.
- [2] C. Declair and M. S. Hacid, "A Database Approach for Modeling and Querying Video Data," *In IEEE Data Engineering*, Australia, 1999.
- [3] 신성윤, "시각정보와 자막정보를 이용한 축구 비디오 하이라이트 생성," *군산대학교 박사학위 청구논문*, 2003.
- [4] N. Babaguchi, "Towards Abstracting Sports Video by Highlights," *Proc. ICME'00*, Aug. 2000.
- [5] S. Antani, R. Kasturi, R. Jain, "A survey on the use of pattern recognition methods for abstraction, indexing and retrieval of images and video," *Pattern Recognition* 35, pp. 945-965, 2002.
- [6] J. Platt, "AutoAlbum : Clustering Digital Photographs using Probabilistic Model Merging," *IEEE Workshop on Content-Based Access to Image and Video Libraries*, 2000.