

하이브리드 검출기법을 이용한 교양비디오의 장면 전환 검출

이지현*, 진송철*, 문종환*, 이양원*

*군산대학교 컴퓨터학과

e-mail : jhlee@kunsan.ac.kr

The Scene Change Detection of Cultural Videos Using Hybrid Detecting Techniques

Ji-Hyun Lee*, Song-Cheol Jin*,

Jong-Hwan Mun*, Yang-Won Rhee*

*Dept of Computer Science, Kun-San University

요 약

기존 장면 전환 검출 방법들은 대부분 특정 영역에 제한되어 사용할 수 있는 방법들이며, 많은 중요한 특징 정보들을 유실하여 장면 전환 검출에 효율적이지 못하였다. 또한 장면 전환 검출을 통하여 의미 정보를 추출하기가 어렵고, 카메라와 객체의 동작을 정확히 인식하지 못하기 때문에 하이브리드 장면전환 검출 기법을 적용하여 의미 있는 정보를 효율적으로 검출 하였다.

1. 서론

최근 멀티미디어 데이터의 자동화 또는 반자동화된 인식은 저장 공간과 처리할 수 있는 능력의 놀라운 발전으로 인하여 다양한 응용 분야에서 그 중요성이 날로 증대되고 있다. 전형적인 응용 분야로는 멀티미디어 데이터베이스 검색, 정보 필터링 (information filtering), 내용 기반 검색 (content-based retrieval), 대화식 미디어 서비스 (interactive media service) 등이 있다.

내용 기반 데이터 검색을 위해 적용되는 방법과 같은 전통적인 비디오의 이해는 비디오 색인화를 위하여 컬러, 모션, 그리고 텍스처와 같은 저수준 특징들의 사용에 초점을 두고 있다. 저수준 특징들을 위한 일반적인 유사성 측정 방법들의 직접적인 응용들이 유사한 시각적인 특징에 대하여 아주 좋은 결과를 제시할 수 있는 반면, 유사한 의미 클래스 (semantic class)의 식별을 위한 적용성은 매우 어렵고 애매모호한 실정이다. 이는 다양한 저수준 특징들을 효율적으로 결합하기 위한 매우 어려운 작업의 일부이며 다양한 비디오 응용을 위한 일반적인 모델들은 적용할 수 없다[1].

비디오를 의미 있는 단위로 분할하기 위하여 시각 정보 분석을 기반으로 하이브리드(hybrid) 형태의 급진적인 장면과 점진적인 장면에 대한 장면 전환 검출(scene change detection)을 실시하고, 실험을 통하여 비교평가 함으로써 우수성을 입증 하였다.

2. 관련연구

장면 전환 검출에 사용되는 가장 대표적인 기술들의 종류에는 픽셀의 차이, 통계적 차이, 히스토그램 비교, 에지 차이, 압축 차이, 그리고 모션 벡터 등이 있으며 대부분 이웃하는 프레임간의 근접 프레임 비교에 의해 검출하는 방법을 사용한다.

히스토그램은 샷 경계를 검출하기 위해 사용되는 가장 보편화된 방법이다. Ueda 등[2]은 샷 경계를 검출하기 위해서 컬러 히스토그램 변화 비율을 사용했고, Nagasaka 등[3]은 그레이 레벨과 컬러 히스토그램을 기반으로 한 몇 개의 간단한 통계학적 비교를 수행하였다. Swanberg 등[4]은 비디오 시퀀스에서 영역이 변화될 수 있는 정도에 대한 가중치를 부여함으로써 영역에 대한 그레이 레벨 히스토그램 차

이를 사용했다. Zhang 등[5]은 픽셀 차이, 통계 차이, 그리고 몇 개의 다른 히스토그램 방법을 비교했으며, 히스토그램 방법이 정확성과 속도 사이의 좋은 교환 요소임을 발견했다.

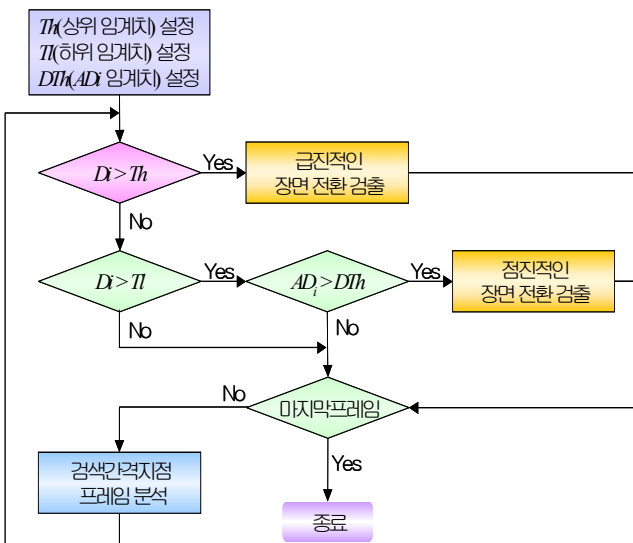
본 논문에서는 기존의 장면 전환 검출 방법들이 갖는 문제점들을 해결하기 위하여 컬러 히스토그램과 χ^2 히스토그램의 장점을 합성한 컬러- χ^2 명도 히스토그램을 이용하여 급진적인 장면 전환 검출 방법을 제시하고, 점진적인 장면들에도 중요한 키 프레임들이 존재함으로 이를 손실 없이 검출하고자 하이브리드 형태의 장면 전환 검출 방법을 제안하였다.

3. 제안된 장면 전환 검출 방법

교양 비디오는 사용자들에게 흥미와 감동을 주기 위한 점진적인 특수효과들이 등장한다. 따라서 교양 비디오는 급진적인 장면 전환과 점진적인 장면 변화에 의해 장면들이 진행된다.

교양 비디오를 분할하기 위한 장면 전환 검출은 일반적으로 임계치를 이용하여 임계치를 초과하는 프레임을 샷 경계로 설정한다. 즉, 주어진 임계치를 초과하는 첫 번째 프레임을 샷의 키 프레임으로 설정하는 것이다.

본 논문에서는 두 단계의 장면 전환 검출을 수행하는 방법을 다음과 같이 제시하였다. 첫 번째 단계는 급진적인 장면 전환을 검출하고, 두 번째 단계는 점진적인 장면 전환을 검출한다. 하이브리드 장면 전환 검출 방법에 대한 전반적인 흐름은 (그림 1)과 같다.



(그림 1) 하이브리드 장면 전환 검출 흐름도

첫 번째, 급진적인 장면 전환 검출을 위하여 본 논문에서는 식 (1)과 같은 컬러- χ^2 명도 히스토그램 차이 값을 이용한다.

$$d(I_t, I_{t-1}) = \frac{1}{3} \cdot \sum_{j=1}^N \left(\frac{(H_t^r(j) - H_{t-1}^r(j))^2}{H_t^r(j)} \times 0.333 + \frac{(H_t^g(j) - H_{t-1}^g(j))^2}{H_t^g(j)} \times 0.333 + \frac{(H_t^b(j) - H_{t-1}^b(j))^2}{H_t^b(j)} \times 0.333 \right) \quad (1)$$

식 (1)의 컬러- χ^2 명도 히스토그램 방법은 컬러 히스토그램을 R·G·B 각각에 대하여 산출함으로써 영상의 컬러를 구성하는 요소들을 신축성 있게 사용할 수 있으며, χ^2 히스토그램이 갖는 차이 값 강조 특징을 적용하여 보다 효율적으로 장면 전환을 검출할 수 있는 방법이다. 식 (1)에 곱한 세 개의 가중치는 영상을 RGB 컬러 모델에서 HSI 모델의 I(intensity) 요소로 바꾸기 위한 값이다. 즉, 컬러 모델을 명암도 등급으로 바꾸어 컬러를 구성하는 3원색을 컬러 분광 정도에 따라 명암도 단계의 데이터로 사용하기 위한 값이다.

본 논문에서는 임계값을 설정하기 위하여 제 262회 도전 지구탐험대에서 10분 분량의 비디오에 대하여 여러 개의 임계값을 적용하여 가장 적절한 급진적 장면 전환 검출 임계값을 선택하도록 하였으며, <표 1>은 임계값 적용에 따른 급진적 장면 전환 검출 결과를 비교한 것이다.

<표 1> 임계값별 급진적 장면 전환 검출 결과

| 구분 | 전체검출 | 미검출 | 오검출 |
|------------|------|-----|-----|
| 임계값1 (600) | 298 | 0 | 97 |
| 임계값2 (800) | 244 | 0 | 63 |
| 임계값3(1000) | 204 | 5 | 34 |
| 임계값4(1400) | 148 | 8 | 15 |
| 임계값5(1800) | 90 | 25 | 0 |

<표 1>에서 전체검출은 추출에 성공한 수를 의미하고, 미검출은 추출되어야 하지만 추출되지 않은 수를 의미하며, 오검출은 추출되지 않아야 하는데 추출된 수를 의미한다. 따라서 임계값을 너무 높게

설정하면 검출하지 못하는 장면 전환이 많고, 임계값을 너무 낮게 설정하면 정확한 장면 전환은 작게 검출하고 장면 전환이 아닌 오류를 많이 검출하게 된다. 결과에서 알 수 있는 것처럼 임계값4 (1400)의 경우가 추출에 성공하여 정확히 추출한 개수가 가장 많고, 미검출 개수와 오검출 개수가 적기 때문에 가장 만족스러운 임계값이다.

두 번째, 점진적인 장면 전환은 앞 단계에서 구한 값과 다음 식 (2)와 식 (3)를 적용하여 차를 구한 식 (4)의 방법을 이용하여 검출한다.

$$AI_i = \frac{\sum_{j=1}^{Bins} j * H_i(j)}{\sum_{j=1}^{Bins} H_i(j)} \quad (2)$$

$$AI_{i-1} = \frac{\sum_{j=1}^{Bins} j * H_{i-1}(j)}{\sum_{j=1}^{Bins} H_{i-1}(j)} \quad (3)$$

$$AD_i = AI_i - AI_{i-1} \quad (4)$$

식 (2), 식 (3), 식 (4)에서 AI_i 는 프레임 i 의 평균 컬러 값이고, AI_{i-1} 는 프레임 $i-1$ 의 평균 컬러 값이다. AD_i 는 프레임 i 와 $(i-1)$ 사이의 평균 컬러 값의 차이 값을 말한다.

점진적인 장면 전환은 컬러 히스토그램 차이 값 D_i 가 상위 임계치 Th 보다는 적고, 하위 임계치 Tl 보다는 크며 평균 컬러 값의 차이 값 AD_i 가 AD_i 임계치 DTh 보다 클 경우에 검출된다.

<표 2>는 임계값 적용에 따른 점진적 장면 전환 검출 결과를 비교한 것이다.

<표 2> 임계값별 점진적 장면 전환 검출 결과

| 구분 | 전체검출 | 미검출 | 오검출 |
|-----------|------|-----|-----|
| 임계값1 (60) | 14 | 0 | 6 |
| 임계값2 (70) | 11 | 0 | 3 |
| 임계값3 (80) | 9 | 0 | 1 |
| 임계값4 (90) | 4 | 5 | 0 |
| 임계값5(100) | 3 | 6 | 0 |

결과에서 알 수 있는 것처럼 임계값3 (80)의 경우가 추출에 성공하여 정확히 추출한 개수가 가장 많고, 미검출 개수와 오검출 개수가 적기 때문에 가장 만족스러운 임계값이다.

급진적인 장면 전환 검출의 미검출 수를 보완하기 위하여 점진적 장면 전환 검출을 혼합하여 적용한 결과 <표 3>과 같은 결과를 얻음으로 인하여 보다 효과적인 장면 전환 검출을 할 수 있었다.

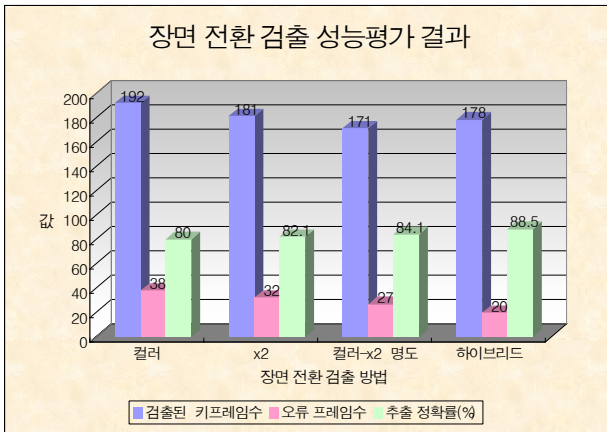
4. 성능평가

성능평가 비디오는 261회에서 263회까지 방영한 3회차의 도전 지구탐험대의 방송 10분을 대상으로 초당 10 프레임으로 캡처한 비디오 데이터에 대하여 컬러 히스토그램, χ^2 히스토그램, 컬러- χ^2 명도 히스토그램, 하이브리드를 이용하여 장면 전환 검출을 수행하였다. 그리고 이렇게 장면 전환 검출을 수행한 결과로써 검출된 키 프레임들의 수를 <표 3>과 같이 제시하였다. 검출된 키 프레임 수는 실험 비디오 데이터에서 추출한 총 키 프레임의 수를 말하고, 오류 프레임 수는 검출된 키 프레임에 포함된 오류 키 프레임의 수를 말한다. 그리고 추출 정확률은 검출된 키 프레임 수에서 오류 프레임 수를 뺀 프레임 수, 즉 정확히 추출한 프레임 수의 비율을 나타낸다.

<표 3> 장면 전환 검출 성능평가 결과

| 구분 | 검출방법 | 검출된 키프레임수 | 오류 프레임수 | 추출 정확률(%) |
|-------|------------------------|-----------|---------|-----------|
| 261 회 | 컬러 히스토그램 | 182 | 37 | 79.7 |
| | χ^2 히스토그램 | 167 | 31 | 81.4 |
| | 컬러 - χ^2 명도 히스토그램 | 160 | 28 | 82.5 |
| | 하이브리드 장면 전환 | 165 | 23 | 86.1 |
| 262 회 | 컬러 히스토그램 | 172 | 35 | 79.7 |
| | χ^2 히스토그램 | 161 | 27 | 83.2 |
| | 컬러 - χ^2 명도 히스토그램 | 148 | 23 | 84.5 |
| | 하이브리드 장면 전환 | 156 | 17 | 89.1 |
| 263 회 | 컬러 히스토그램 | 223 | 43 | 80.7 |
| | χ^2 히스토그램 | 214 | 39 | 81.8 |
| | 컬러 - χ^2 명도 히스토그램 | 204 | 30 | 85.3 |
| | 하이브리드 장면 전환 | 212 | 22 | 89.6 |
| 평균 | 컬러 히스토그램 | 192 | 38 | 80.0 |
| | χ^2 히스토그램 | 181 | 32 | 82.1 |
| | 컬러 - χ^2 명도 히스토그램 | 171 | 27 | 84.1 |
| | 하이브리드 장면 전환 | 178 | 21 | 88.3 |

본 논문에서는 장면의 변화를 보다 정확히 추출할 수 있도록 급진적과 점진적 장면 전환 검출을 하이브리드 형태로 이용함으로써 (그림 2)와 같은 결과를 얻었다.



(그림 2) 장면 전환 검출 성능평가 결과

(그림 2)는 도진 지구탐험대 3회차의 10분 비디오 데이터에 대한 추출 정확률(%)을 비교하였을 때 컬러 히스토그램 80%, χ^2 히스토그램 82.1%, 제안된 컬러- χ^2 명도 히스토그램 84.1%와 하이브리드 장면 전환 검출 88.5%로 검출률이 향상된 장면 전환 검출을 수행하였다.

5. 결론

검출된 샷들의 수를 비교하면 컬러 히스토그램을 이용할 경우 가장 많고, χ^2 히스토그램을 이용할 경우 컬러 히스토그램보다 적으며, 컬러- χ^2 명도 히스토그램과 하이브리드를 이용할 경우 가장 작았다. 그리고 잘못 추출된 샷들의 수도 검출된 샷들의 수와 마찬가지로 컬러 히스토그램을 이용할 경우 가장 많고, 컬러- χ^2 히스토그램과 하이브리드를 이용할 경우가 가장 작았다. 또한 검출된 키 프레임 수에 대한 추출 정확률도 컬러- χ^2 명도 히스토그램과 하이브리드를 이용할 경우 가장 높았다. 따라서, 제안된 컬러- χ^2 명도 히스토그램과 하이브리드를 이용한 장면 전환 검출 방법은 검출된 샷의 수가 가장 적고, 잘못 검출된 샷의 수도 가장 적으며, 추출 정확률도 가장 높기 때문에 컬러 히스토그램이나 χ^2 히스토그램을 각각 이용하여 검출하는 방법보다 컬러- χ^2 명도 히스토그램과 하이브리드를 이용한 장면 전환 검출 성능이 우수함을 입증하고 있다.

참고문헌

- [1] W. Zhou, S. Dao, and C.-C. J. Kuo, "Hierarchical Video Classification with Mixed Media Cues," *Proc. of SPIE, Storage and*

Retrieval for Media Databases 2001, Vol. 4315, pp. 280-291, 2001.

- [2] H. Ueda, T. Miyatake, and S. Yoshizawa, "IMPACT: An Interactive Natural-motion-picture Dedicated Multimedia Authoring System," in *proceedings of CHI, 1991 ACM*, pp. 343-350, New York, 1991.
- [3] Nagasaka, A. and Tanaka, Y., "Automatic Video Indexing and Full-Video Search for Object Appearances," in *Visual Database Systems II*, Knuth, E., Wegner, L., Editors, Elsevier Science Publishers, pp. 113-127, 1992.
- [4] Swanberg, D., Shu C. F., and Jain, R., "Knowledge Guided Parsing and Retrieval in Video Databases," in *Storage and Retrieval for Image and Video Databases, Proc. SPIE 1908*, pp. 173-187, February 1993.
- [5] X. D. Zhang, T. Y. Liu, K. T. Lo, J. Feng, "Dynamic selection and effective compression of key frames for video abstraction," *Pattern Recognition Letters* 24, pp. 1523-1532, 2003.