

컬러 히스토그램과 엔트로피를 이용한 동영상 컷 검출

°송현석, 안강식, 안명석*, 조석제
한국해양대학교 제어계측공학과, 컴퓨터공학과*

Cut Detection of Video Data Using Color Histogram and Entropy

Hyun-Seok Song, Kang-Sik An, Myung-Seok An, Seok-Je Cho
Dept. of Control & Instrumentation, Computer Engineering
Korea Maritime University
E-mail : wetjean@kmce.kmaritime.ac.kr

요약

내용에 기반한 동영상 검색에서, 대표 프레임을 자주 이용하는데, 이를 위해 우선적으로 동영상의 장면변화를 검출하는 기술이 필요하다. 일반적으로 컬러 히스토그램 비교방법이 많이 쓰이나, 급격한 조명변화에 민감하고 컬러 히스토그램 분포가 비슷한 부분의 장면전환을 놓칠 수 있다는 단점이 있다. 본 논문에서는 컬러 히스토그램 비교방법과 엔트로피를 복합적으로 이용하여 조명변화에 의해 장면전환이 잘못 검출되는 것을 막을 수 있다. 실험을 통해 제안한 방법은 컬러 히스토그램 비교방법보다 조명변화에 보다 강력함을 확인할 수 있었다.

Abstract

In content-based video data retrieval, the representative-frame is usually used. To do that, the skill of detection for scene change is needed. Generally the color histogram comparison is used, but sensitive to light variation and tends to miss the scene change of similar color histogram. This paper shows how to use both color histogram comparison and entropy to prevent the false-positive of scene change occurred by light variation. At the experiments, it is more powerful to light variation to use both color histogram comparison and entropy than to use only color histogram comparison.

1. 서론

VOD(Video On Demand)와 같은 서비스가 증가함에 따라 비디오 데이터가 급속히 커지고 있다. 늘어나는 비디오 데이터를 빠르고 효과적으로 저장, 색인, 검색하기 위한 기술이 필요하다. 이러한 비디오 정보 시스템의 인터페이스는 아직까지 키워드 기반에서 크게 벗어나지 못하고 있으며, 최근 내용에 기반한 동영상 검색 기법에 대한 연구가 활발히 진행되고 있다[1,2].

동영상 검색을 위해 우선적으로 동영상을 효과적으로

분할, 색인할 필요가 있다. 일반적으로 동영상 전체의 각 장면에서 나오는 대표프레임을 사용한다. 이 대표프레임 집합만으로 긴 동영상을 간략히 표현할 수 있기 때문이다. 연속되는 여러 프레임을 대표할 수 있는 하나의 프레임을 대표프레임이라 하며, 보통 하나의 대표프레임은 연속되는 샷(shot) 중의 하나로 정한다. 동영상에서 급격한 장면전환이 이루어지는 부분을 컷(cut)이라 하며, 컷과 컷 사이의 프레임 집합을 샷이라고 한다.

대표프레임을 얻기 전 꼭 필요한 단계가 컷 검출인데, 컷이 정확하게 검출되어야 대표프레임도 정확하게 검출된다는 사실은 명백하다. 컷 검출을 위해 일반적으로 쓰이는 방법으로 컬러 히스토그램을 이용한다[1-5]. 이 방법은 빠른 속도로도 월등한 능력을 보여준다. 하지만 공간정보를 포함할 수 없고, 갑작스런 조명변화에 잘못된 컷을 검출할 수 있는 단점이 있다. 뮤직비디오, 뮤지컬, 공연 같은 동영상은 수많은 장면전환 뿐 아니라 조명효과나 섬광, 카메라 플래쉬와 같은 밝기변화도 상당히 많기 때문에 컬러 히스토그램 비교방법만으로는 이러한 조명효과가 포함되는 영상에 대해 잘못 검출되는 컷이 너무 많다. 본 논문에서는 컬러 히스토그램을 이용한 방법의 문제점을 해결하기 위해 엔트로피(entropy)를 이용한 방법을 제안한다. 엔트로피는 정보의 복잡도에 관한 법칙으로[6], 조명이 갑자기 변화하였을 때에도 영상의 복잡도는 크게 변하지 않는다. 실험 결과를 통해, 제안한 방법이 조명변화에 더욱 민감하지 않아 잘못된 검출을 상당히 줄일 수 있음을 확인하였다.

II. 컬러 히스토그램을 이용한 컷 검출

컬러 히스토그램은 한 영상의 컬러분포를 나타낸 것이다. 연속되는 두 프레임 간의 컬러분포는 장면전환이 일어나지 않는 한 유사하며, 반대의 경우 두 프레임의 컬러 히스토그램 차이가 커진다. 이런 관점에서 컬러 히스토그램이 많이 사용된다. 식 (1)은 히스토그램 차이를 나타낸 식이다.

$$D_h = \frac{\sum_{j=1}^K |H_m(j) - H_{m+1}(j)|}{K} \quad (1)$$

여기서 $H(j)$ 는 히스토그램, K 는 화소값의 범위, 그리고 m 은 프레임을 나타낸다.

히스토그램 비교방법은 물체의 움직임과 카메라 동작 및 약간의 가려짐에도 강력한 효과가 있으나, 영상의 공간정보를 포함하지 못하며, 밝기변화에 민감하다는 단점이 있다. 즉, 갑작스런 조명변화나 섬광, 비슷한 배경이나 분위기에서 장면전환이 일어날 경우 잘못된 컷을 검출하거나 올바른 컷을 놓칠 수 있다. 그림 1의 #1997 프레임과 #1998 프레임이 컷은 아니지만 갑작스런 조명변화로 인해 컷으로 잘못 검출된 하나의 예이며, 그림 2는 컷이지만 컬러 히스토그램이 유사한 이유로 검출되지 못한 예를 보여주고 있다. 이런 문제점을

위해 본 논문에서는 갑작스런 조명변화에 강인한 엔트로피를 추가하여 컷을 검출하는 방법을 제안한다.

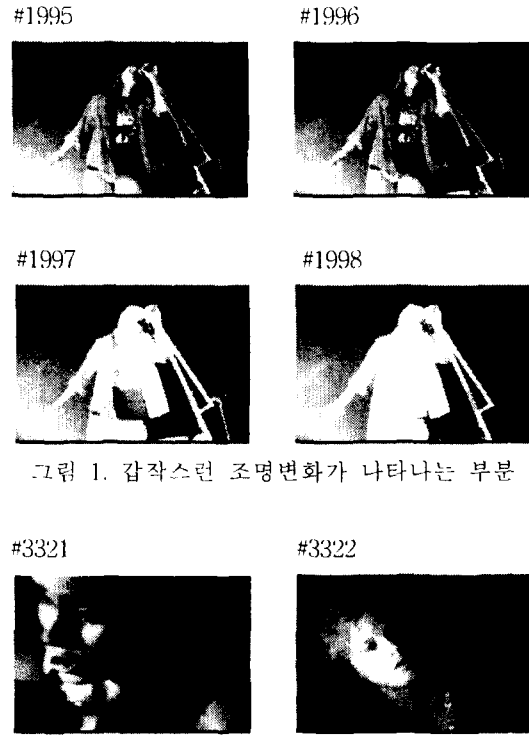


그림 1. 갑작스런 조명변화가 나타나는 부분

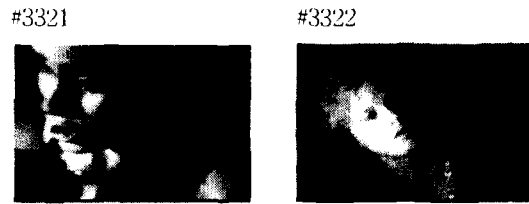


그림 2. 컬러 히스토그램이 유사한 컷

III. 컬러 히스토그램과 엔트로피를 이용한 검출

3.1 엔트로피

영상 정보의 복잡도를 확률에 기반한 엔트로피로 나타내면 식 (2)와 같다. 이 엔트로피를 이용하면 영상의 복잡도를 측정할 수 있다.

$$E = - \sum_{j=1}^K P(a_j) \log P(a_j) \quad (2)$$

여기서 a_j 는 j 번째 화소값의 개수이며, $P(a_j)$ 는 a_j 의 확률이다. 그림 1에서 연속되는 네 프레임의 엔트로피를 식(2)를 이용해서 구해보면 2.166, 2.187, 2.229, 2.064 이다.

동영상에서는 연속되는 두 프레임간의 엔트로피 차이를 측정하여 프레임 차이를 판별한다. 갑작스런 조명변화가 생겼을 때 컬러 히스토그램의 분포는 크게 차이가 난다. 그러나 물체나 배경, 혹은 영상 전체의 밝기변화

가 생기더라도 그 영상의 복잡도는 크게 변하지 않기 때문에 엔트로피 값이 크게 변하지 않는다. 따라서 이 방법을 사용하면 조명변화로 인한 잘못된 검출을 막을 수 있다.

3.2 제안한 방법

그림 2와 같이 장면전환이 일어났음에도 불구하고 컬러 히스토그램 분포가 비슷하다면 컷을 놓칠 것이다. 하지만 이런 극단적인 경우는 빈번하게 일어나지 않으므로 고려하지 않고, 급격한 밝기변화만 다루기로 한다. 뮤지컬 공연이나 실황방송, 뮤직비디오 같은 경우 장면 전환뿐 아니라, 수많은 조명변화와 카메라 플래쉬, 스포트라이트가 포함되어 있다. 컬러 히스토그램 비교방법은 이러한 영상에 적용하기에는 부적합하다. 제안한 방법으로 얻은 엔트로피를 이용하여 연속되는 두 프레임간의 엔트로피 차이를 얻을 수 있다. 본 논문에서는 컬러 히스토그램과 엔트로피 비교방법을 함께 사용한다. 두 프레임간의 컬러 히스토그램과 엔트로피의 차이를 구하여, D_{total} 이 임계값을 넘을 때에 컷으로 간주한다.

식 (3)은, 연속되는 두 프레임의 엔트로피 차이를 나타내며, 식 (4)는 식 (1)의 컬러 히스토그램 차이 D_h 와 식 (3)의 엔트로피 차이 D_e 를 결합한 전체 차이를 나타낸다. 즉 컬러 히스토그램의 차이와 엔트로피 차이의 가중치 합으로 나타낼 수 있다.

$$D_e = E_m - E_{m+1} \quad (3)$$

$$D_{total} = D_h + wD_e \quad (4)$$

여기서 w 는 가중치이다.

두 프레임의 차이가 식 (5)와 같이 어느 문턱치를 넘으면 컷으로 간주한다.

$$D_{total} > \tau \quad (5)$$

IV. 실험 및 결과

실험에 사용된 동영상은 현란한 조명효과와 섬광, 격렬한 몸동작이 주를 이룬 '전지현'의 비디오, 조명효과가 별로 없는 '광고', 상당한 개수의 컷은 물론 전지현의 비디오처럼 조명, 밝기변화가 상당히 많은 '뮤직'비디오, 그리고 시간도 짧고 움직임이 적은 '뉴스'영상으로 이루어져 있다. 사용된 각 동영상의 특성은 표 1과 같다. 이들 데이터 중 뉴스를 제외한 나머지 데이터는

짧은 시간이지만 수많은 장면전환과 조명변화, 빠른 동작을 포함하여 충분히 정확한 결과를 산출할 수 있었다. 엔트로피는 256단계 명암도(gray-level)를 사용하였고, 컬러 양자화는 64단계로 하였다.

표 1. 실험 데이터

	시간(분:초)	크기	컷의 개수
전지현	02:05	320×240	101
광고	00:46	352×240	39
뮤직	05:04	352×240	225
뉴스	00:38	320×240	9

문턱치는 여러 차례의 실험을 통해 적절한 결과가 나오는, 가장 큰 차이값의 15%로 정하였고, 가중치는 1로 정하였다. 실험결과는 표 2와 같이, 컬러 히스토그램 비교방법만을 사용하는 것보다 엔트로피 비교방법을 함께 사용하는 것이 잘못 검출된 컷의 개수가 놓친 검출의 개수에 비해 현저히 줄었다는 것을 알 수 있다. 전지현의 비디오는 많은 밝기변화와 조명, 거울에 반사된 조명, 몸동작이 포함되어 있다. 따라서 컬러 히스토그램 비교방법만 사용할 경우 컷으로 잘못 검출될 가능성이 매우 높다. 실험결과, 컬러 히스토그램 비교방법에서는 조명변화에 민감하기 때문에 잘못된 검출을 많이 일으켜 제안한 방법과 차이를 보여준다. 뮤직비디오의 경우도 마찬가지로 컬러 히스토그램 비교방법 만으로는 잘못 검출된 컷이 상당히 많았다. 그림 5와 6은 각각, 뮤직비디오에서 현란한 조명이 자주 나오는 부분의 3000 프레임을 컬러 히스토그램 차이와 컬러 히스토그램+엔트로피 차이의 그래프로 나타낸 것이다. 그래프를 보면, 컬러 히스토그램에 엔트로피를 이용함으로써 전체적으로 값이 줄어드는 경향이 있으나 조명효과로 인한 잘못된 컷을 상당수 없애주는 효과가 있다. 광고는 조명변화가 많지 않아서 컬러 히스토그램 비교방법과 차이가 별로 없으며, 뉴스의 경우도 컷의 구분이 확연하고 정적인 영상이 주를 이루므로 기존의 방법과 차이를 보이지 않았다.

표 2. 실험 결과

(a) 컬러 히스토그램 비교방법

	검출	잘못검출	놓침	문턱치
전지현	81	94	20	0.15
광고	35	27	4	0.15
뮤직	206	371	19	0.15
뉴스	9	0	0	0.15

(b) 컬러 히스토그램 + 엔트로피 비교방법

	검출	잘못검출	놓침	문턱치
전지현	75	51	26	0.15
광고	39	30	0	0.15
뮤직	181	181	44	0.15
뉴스	9	0	0	0.15

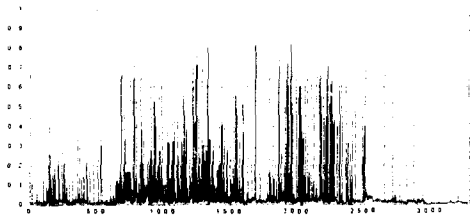


그림 3. 컬러 히스토그램 차이 (뮤직비디오)

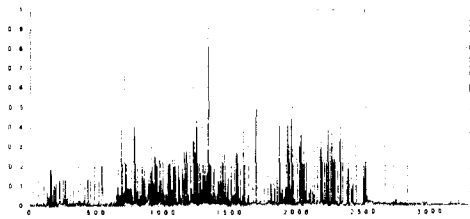


그림 4. 컬러 히스토그램 + 엔트로피 차이 (뮤직비디오)

V. 결론 및 향후과제

동영상 검색에 있어서 선행되어야 할 과제가 내용에 적합한 장면분할이다. 영상에서 엔트로피 값은 밝기변화에 민감하지 않다. 본 논문에서는 갑작스런 조명변화에서도 원활한 장면전환 검출을 위해 컬러 히스토그램 비교방법에 엔트로피를 이용하는 방법을 제안하였다. 실험을 통하여, 컬러 히스토그램 비교방법만을 사용하

는 것보다 제안한 방법이 더 우수함을 보여주었다. 차후 가중치와 문턱치의 최적값을 산출하는 법도 연구되어야 하며, 제안한 방법을 확장하여 급격한 장면전환만이 아닌, 점진적인 장면전환도 검출하는 방법이 연구되어야 하겠다.

참고문헌

- [1] G. Ahanger and T.D.C. Little, "A survey of technologies for parsing and indexing digital video," *Journal of Visual Communication and Image Representation* 7(1), pp. 28-43, 1996.
- [2] Irena Koprinska and Sergio Carrato, "Temporal Video Segmentation : A Survey," *Signal Processing : Image Communication*, v.16, pp. 477-500, 2001.
- [3] H. J. Zhang, A. Kankanhalli, and S.W. Smoliar, "Automatic partitioning of full-motion video," *Multimedia Systems* 7, pp. 119-128, 1993.
- [4] J. Y. Lee, S. Y. Jeong, B. T. Chun, and Y. J. Bae, "A Method for Shot Boundary Detection and R-Frame Selection of Digital Video," *Proc. of Visual Database Systems 4(VDB 4)*, pp. 361-375, 1998.
- [5] Ramin Zabih, Justin Miller, and Kevin Mai, "A feature-based algorithm for detecting and classifying production effects," *Multimedia Systems* 7, pp. 119-128, 1999.
- [6] Rafael Gonzalez and Richard Woods, *Digital Image Processing*, Addison-Wesley, 1992.
- [7] Michael J. Swain and Dana H. Ballard, "Color Indexing," *International Journal of Computer Vision*, 7:1, pp. 11-32, 1991.
- [8] John S. Boreczky and Lawrence A. Rowe, "Comparison of video boundary detection techniques," *IS&T/SPIE*, vol. 2670, pp. 170-179, 1996.