

MPEG 비디오 스트림의 칼라 히스토그램 정보와 매크로블록 정보를 이용한 새로운 샷 경계 검출 방법

정진국, 이화순, 낭중호, 김경수*, 하명환*, 정병희*
서강대학교 컴퓨터학과, KBS 기술 연구소*

(jjguk, hslee}@mlneptune.sogang.ac.kr, jhnang@ccs.sogang.ac.kr

A New Shot Change Detection Scheme Using Color Histogram and Macroblock Information of MPEG Video Stream

Jin-Guk Jeong Hwa-Sun Lee Jong-Ho Nang Gyung-Su Kim* Myung-hwan Ha* Byung-Hui Jeong*
Dept. of Computer Science, Sogang University, KBS Technical Research Institute*

요 약

최근 디지털 비디오 데이터의 사용이 급격히 증가하면서 보다 정확하게 샷을 검출하는 기법이 요구되고 있다. 비디오 정보를 이용하여 샷을 검출하는 연구는 크게 이산코사인 변환의 결과값을 이용하는 방법과 움직임 보상의 결과값을 이용하는 방법으로 그룹화할 수 있는데 전자의 방법은 점진적인 변화를 검출할 수 있는 반면에 전체적인 검출율이 떨어진다는 단점이 있고, 후자의 방법은 전체적인 검출율은 높지만 점진적인 변화를 검출할 수 없다는 단점이 있다. 본 논문에서는 실험을 통하여 이러한 두 가지 방법의 특징을 살펴본 후 이 방법들을 이용한 새로운 샷 경계 검출 방법을 제안한다. 전체적으로 검출율을 높이는 데 목적을 두었기 때문에 매크로블록 타입을 이용하는 방법을 기본으로 하면서 히스토그램을 이용하는 방법을 추가하여 precision을 높일 수 있도록 하였다. 히스토그램을 이용하는 방법에서는 단순히 프레임과의 비교를 하던 기존의 방법보다 프레임들간의 차이의 차이를 이용하여 성능을 높일 수 있도록 하였다. 본 논문에서 제안한 알고리즘을 이용하여 실험을 한 결과 평균 0.96의 recall과 0.96의 precision을 보이고 있음을 알 수 있었다.

1. 서론

최근 인터넷을 중심으로 디지털 비디오의 활용이 급속히 확산되고 대중화됨에 따라 VOD(Video On Demand)나 디지털 비디오 라이브러리와 같이 비디오 데이터를 이용하는 시스템의 개발이 활발히 이루어지고 있다. 이러한 시스템에서는 압축 형식의 비디오를 저장한 데이터베이스를 구축하여야 하는데 이를 위해서 비디오 데이터에 대한 인덱싱 과정이 필요하다. 특히 비디오 데이터의 가장 기본이 되는 단위가 샷이기 때문에 샷 경계 검출 과정을 자동화하는 작업은 비디오 데이터를 이용하는 시스템을 개발하는 데 있어 가장 기본이 되는 작업이다.

MPEG 형식 동영상에 대한 자동화된 샷 경계 검출 연구는 그 동안 많이 진행되어 왔다. 이러한 연구는 크게 두 가지 그룹으로 나눌 수 있는데, 첫째가 인코딩시의 이산코사인 변환(DCT)의 상수 정보를 이용하는 방법 [1,2,3,4,6,7,8,9,10]이고 둘째가 움직임 보상의 결과를

이용하는 방법 [1,5,11]이다. 첫째 방법은 점진적인 샷의 변화를 찾을 수 있는 장점이 있고, 둘째 방법은 비록 점진적인 샷의 변화는 찾을 수 없으나 첫째 방법에 비해 recall과 precision이 좋다는 장점이 있다.

본 논문에서는 위에서 설명한 두 가지 방법을 이용하여 새로운 알고리즘을 제안한다. 이 방법에서는 기본적으로 검출율을 높이는 데 목적이 있기 때문에 움직임 보상의 결과를 이용하는 방법을 기본으로 하면서 이산 코사인 변환의 결과를 이용하는 방법을 추가하여 precision을 높일 수 있도록 하였다. 또한 본 논문의 알고리즘에서는 기존의 방법들의 문제점으로 지적되어졌던 데이터의 장르에 종속적인 면을 보완하기 위해서 적응적으로 임계값을 설정할 수 있는 방법을 제안하고 있다. 본 논문에서 제안한 알고리즘을 적용하여 실험한 결과 평균 0.96의 recall과 0.96의 precision을 보였다.

2. 새로운 샷 경계 검출 알고리즘

본 논문에서는 위에서 언급한 두 가지 그룹의 알고리즘 - DCT 상수 정보를 이용하는 방법, 움직임 보상의 결과를 이용하는 방법 - 에서 좋은 성능을 보이는 알고리즘을 하나씩 골라 결합함으로써 더 나은 성능을 보이는 새로운 알고리즘을 제안한다. 이용한 두 알고리즘은 DC 이미지를 이용하는 방법[2,3,4]과 매크로블록 타입의 개수를 비교하는 방법[3,7]이다.

본 논문의 알고리즘에서는 성능을 높이기 위해서 DC 이미지를 이용하는 방법을 조금 변형시킨 후 이용하였다. 즉, 단순히 인접한 두 프레임의 칼라 히스토그램만 비교하는 것이 아니라 프레임간의 칼라 히스토그램 차이의 차이값을 임계값과 비교하도록 하였다. 카메라의 이동이 있는 경우를 예로 들어 보면 각 프레임간의 칼라 히스토그램 차이가 심하기 때문에 기존의 방법으로는 이 부분을 샷이라고 검출할 것이다. 하지만 변형된 방법을 쓰게 되면 비록 프레임간의 칼라 히스토그램의 차이는 심하지만 카메라는 어느 정도 일정한 속도로 움직일 것이기 때문에 칼라 히스토그램 차이의 차이값은 크지 않아서 샷으로 검출하지 않을 것이다. 이와 같이 변형된 알고리즘을 이용하게 되면 기존의 방법을 이용할 경우 발생했던 잘못된 검출의 수를 줄일 수 있게 된다.

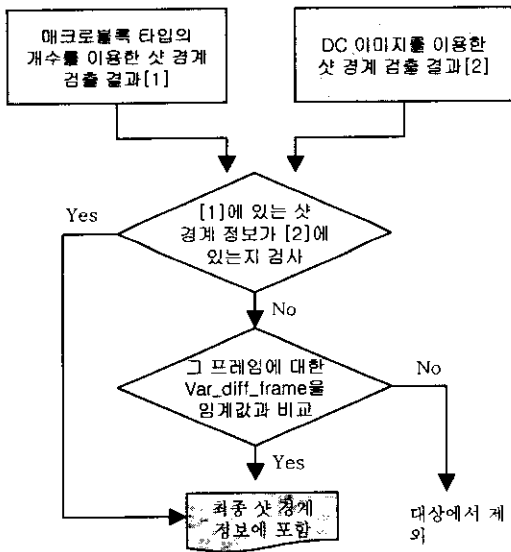


그림 2-1 제안하는 샷 경계 검출 알고리즘

<그림 2-1>은 본 논문에서 제안하는 샷 경계 검출 방법을 나타내고 있다. 기존 방법에 대한 실험 결과 매크로블록 타입의 개수에 의한 샷 경계 검출 방법이 대

부분의 장르에 대해서 높은 recall을 보였고 수행 속도도 빠름을 알 수 있었다. 반면 DC 이미지를 이용하는 방법은 점진적인 변환을 찾을 수 있다는 장점이 있기는 했지만 매크로블록 타입의 개수를 이용하는 방법보다는 recall과 precision이 떨어졌다. 비록 점진적인 변환을 찾을 수 없다고 하더라도 recall이 높은 것이 최종적으로 사용자에게 더 많은 샷 경계 정보를 줄 수 있기 때문에 본 논문의 알고리즘에서는 매크로블록 타입의 개수를 이용하는 방법을 우선시 하여 최종 샷 경계 정보를 생성하도록 하였다. 그리고 DC 이미지를 이용하는 방법은 매크로블록 타입의 개수를 이용하는 방법으로 찾을 수 없는 샷을 찾는 보조적인 수단으로 이용했다. 보다 자세히 알고리즘을 살펴보면 다음과 같다. 매크로블록 타입의 개수를 이용한 샷 경계 검출 방법에서 샷 경계로 찾은 프레임을 DC 이미지를 이용한 샷 경계 검출 방법에서도 샷으로 검출하였는지를 검사한다. 만약 검출이 되었다면 최종 샷으로 결정하고, 검출이 되지 않았다면 DC 이미지를 이용한 샷 경계 검출 방법 수행 중에 얻어지는 각 프레임들의 차이의 차이값 (var_diff_frame)을 임계값과 비교한다. 이 값이 임계값보다 크면 최종 샷으로 결정하고 그렇지 않으면 대상에서 제외하도록 하였다.

기존 연구들의 대부분은 임계값을 설정하는 부분에 대해 크게 고려하지 않았다. 그렇기 때문에 알고리즘을 적용하는 데 있어 샘플 데이터에 따라서 검출율의 변화가 심하게 나타났었는데 본 논문의 알고리즘에서는 데이터의 앞 300 프레임을 미리 살펴본 이후에 임계값을 자동으로 설정해주도록 함으로 해서 기존의 문제점을 많은 부분 보완할 수 있었다.

3. 성능 평가 및 분석

이 장에서는 3장에서 제안한 새로운 샷 경계 검출 방법을 여러 실험 데이터를 이용해 실험한 결과를 보이고 분석함으로 해서 그 유용성을 평가하도록 하겠다. 실험에서 사용한 비디오 데이터는 1.5Mbps의 비트율을 가지고 있고, 오디오 데이터는 44.1KHz의 샘플링 비율을 가지고 있다. 또한 비디오 데이터는 IBBPBBPBBPBBPBBPBBPBI의 GOP 형태를 가지고 있다.

표 3-1 샷 경계 검출 방법 속도 비교

방법	DC 이미지를 이용하는 방법	매크로블록 타입의 개수를 이용하는 방법	제안하는 알고리즘
속도 (sec)	183.56	156.00	339.56 + α

<표 3-1>과 <표 3-2>는 각각 샷 경계 검출 방법의

속도와 성능의 비교를 나타내고 있다. <표 3-1>에서 나타나는 바와 같이 수행 속도면에서는 매크로블록 타입만 이용하는 경우가 가장 좋았는데 이 방법에서는 복호화의 과정이 전혀 필요하지 않기 때문에 이러한 결과를 보이고 있다. 제안 알고리즘은 다른 두 알고리즘을 모두 수행한 이후에 수행이 되기 때문에 다른 두 알고리즘의 수행 시간과 비교 시간을 합한 시간 정도의 수행 시간을 보임으로 해서 가장 느렸다. 하지만 이 시간 또한 본 논문에서 이용한 실험 환경에서는 실제 재생 시간보다는 짧게 걸림을 알 수 있었다. Recall, precision을 통해 각 알고리즘이 샷을 얼마나 정확하게 찾는 지를 살펴보면 <표 3-2>에서 나타나는 바와 같이 본 논문에서 제안한 알고리즘이 거의 모든 경우에서 다른 알고리즘보다 좋은 성능을 보임을 알 수 있었다.

표 3-2 샷 경계 검출 방법 성능 비교

장 르 (총 30000 프레임)	DCT 계수 만 이용		매크로블록 타입만 이용		제안 알고리즘	
	r	p	r	p	r	p
시트콤	0.92	0.93	0.99	0.98	0.99	1
일반 영화	0.91	0.88	0.97	0.78	0.96	0.98
뉴스	0.91	0.89	0.99	0.71	0.98	0.95
액션 영화	0.85	0.77	0.99	0.78	0.98	0.94
연예 프로	0.89	0.81	0.93	0.69	0.93	0.93
스포츠	0.90	0.93	0.86	0.79	0.91	0.99
평균	0.89	0.86	0.95	0.79	0.96	0.96

r : recall, p : precision

본 논문에서 제안하는 알고리즘은 DC 이미지를 이용하는 방법에서는 찾을 수 없었던 샷을 매크로블록 타입에서는 찾아주기 때문에 recall이 높아질 수 있고, 또한 매크로블록 타입에서 잘못 찾았던 샷을 변형된 DC 이미지를 이용하는 방법에서 없애주기 때문에 precision이 높아질 수 있는 것이다.

4. 결론 및 앞으로의 연구 방향

최근 많은 멀티미디어의 사용으로 인해 정확한 샷 경계 검출 과정이 더욱 필요하게 되었다. 본 논문에서는 기존의 샷 경계 검출 알고리즘들을 구현, 분석한 후 그 중에서 좋은 성능을 내는 DC 이미지를 이용하는 방법과 매크로블록 타입의 개수를 이용하는 방법을 결합하여 새로운 알고리즘을 제안하였다.

실험 결과 본 논문에서 제안한 알고리즘은 비록 속도 면에서는 다른 알고리즘보다 뒤떨어지지만 recall과

precision 모두 0.96정도의 성능을 보임으로 해서 다른 알고리즘보다 좋은 성능을 보임을 알 수 있었다. 또한 적응적으로 임계값을 설정함으로써 해서 샘플에 종속적인 문제점을 줄일 수 있었다. 문제점으로 점진적인 샷의 변화가 있을 때 찾지 못하는 경우가 있었고, 플래쉬와 같이 같은 샷에서 갑작스러운 변화가 있을 때 잘못 찾는 경우가 있었는데 앞으로의 연구에서는 이러한 부분을 보완하여 보다 정확한 샷 경계 검출 방법을 제안하도록 할 것이다.

5. 참고 문헌

[1] H. J. Zhang, A. Kankanhalli and S. W. Smoliar, "Automatic Partitioning of Full Motion Video," *Multimedia Systems, Vol. 1*, pp. 10-28, July 1993.

[2] B. L. Yeo and B. Liu, "Unified Approach to Temporal Segmentation of Motion JPEG and MPEG Video," *Proceedings of International Conference on Multimedia Computing and Systems*, pp.81-88, May 1995.

[3] B. L. Yeo and B. Liu, "On the Extraction of DC Sequences from MPEG Compressed Video," *Proceedings of International Conference on Image Processing, vol. II*, pp.260-263, 1995.

[4] B. L. Yeo and B. Liu, "Rapid Scene Analysis on Compressed Video," *IEEE Transactions on Circuits and Systems for Video Technology, Vol. 5*, No. 6, pp.533-544, December 1995.

[5] H. J. Zhang, C. Y. Low and S. W. Smoliar, "Video Parsing and Browsing using Compressed Data," *Multimedia Tools and Applications, Vol. 1*, pp. 89-111, March 1995.

[6] J. Meng, Y. Juan and S.F.Chang, "Scene Change Detection in a MPEG Compressed Video Sequence," *Digital Video Compression: Algorithms and Technologies, Vol. SPIE 2419*, pp.14-25, February 1995.

[7] J.S. Boreczky and L.A. Rowe, "Comparison of automatic video segmentation algorithms," *In SPIE Photonics West, volume 2615*, pages 2-16, Philadelphia, 1995.

[8] C. Ngo, T. Pong and R.Chin, "A Survey of Video Parsing and Image Indexing Techniques in Compressed Domain," *Symposium on Image, Speech, Signal Processing, and Robotics vol. 1*, pp. 231-236 1998.

[9] H. Jiang, A. Helal, A. Elmagarmid and A. Joshi, "Scene change detection techniques for video database systems," *Springer-Verlay*, 1998

[10] J. Nang, S. Hong and Y. Ihm, "An Efficient Video Segmentation Scheme for MPEG Video Stream using Macroblock Information," *Proceedings of ACM Multimedia '99*, 1999

[11] H. C. Liu and G. L. Zick, "Scene Decomposition of MPEG Compressed Video," *Digital Video Compression : Algorithms and Technologies, Vol. SPIC 2419*, pp26-37, 1995