

시공간 영상을 이용한 계층적인 장면 전환 검출

홍기진^{*)} 김영봉

부경대학교 전산학과

happyend@unicorn.pknu.ac.kr, ybkim@dolphin.pknu.ac.kr

Hierarchical Shot Boundary Detection Using Time-Space Image

Gi-Jin Hong^U Young-Bong Kim

Dept. of Computer Science, Pukyong National University

요 약

동영상 비디오 시퀀스에서 필요로 하는 장면을 빠르고 쉽게 찾을 수 있도록 해주는 내용 기반 검색에 대한 연구가 활발히 이루어져오고 있다. 특히, 내용 기반 검색 시스템의 기초 기술인 비디오 데이터의 샷(shot)에 따른 분할 연구는 다양한 방법으로 소개되었으나 정확도가 높은 분할 알고리즘이 아직 개발되지 않고 있는 실정이다. 본 논문에서는 비압축 비디오에서 컷(cut) 검출의 효율성을 향상시키기 위해 기존의 히스토그램 비교법과 시공간 영상을 활용하는 계층적인(hierarchical) 방법을 제안한다. 이를 위해 먼저 동영상의 각 프레임에서 한 행(row)씩 추출하여 동영상 전체를 대표하도록 시공간 영상을 생성하고, 생성된 시공간 영상에서 수평 에지(edge)를 이용한 프레임(frame) 특징값으로 장면 전환의 후보 영역을 선택하였다. 그리고 선택된 후보 영역을 히스토그램 비교법으로 분석하게 된다.

1. 서론

데이터 압축 기술과 통신 기술의 발달로 동영상 데이터(video data)를 이용한 다양한 서비스가 가능하게 되었다. 그러나 동영상 데이터의 시간적, 비정형적, 대용량적 특성으로 사용자가 필요로 하는 정보를 찾기 위한 효율적인 처리와 이용이 매우 어렵다. 이에 따라 동영상 데이터를 효율적으로 저장, 관리할 수 있는 내용기반 검색(content-based retrieval)에 대한 연구가 활발하게 이루어져오고 있다[1-3].

내용기반 검색은 동영상 데이터의 특징값을 활용하여 자동적으로 특정 영상을 찾는 방법으로 이를 위해서는 비디오 파싱(video parsing) 작업 중 샷(shot)을 기본 단위로 하는 비디오 분할(video segmentation)이 선행되어야 한다[4-6].

비디오 분할을 위해 많이 쓰이는 기법은 장면 전환 효과에 의해서 발생하는 샷 사이의 경계를 검출하는 것으로 히스토그램 비교법[1,7], 화소간 차이 비교법[1], 움직임 벡터 이용법[8], 데이터 압축율 이용법[9], 에지 정보 이용법[2] 등이 있다.

히스토그램 비교법은 다른 검출기법에 비해 비교적 정확하고 적은 시간을 소요하나 동일한 샷 내에서 두 프레임(frame) 사이의 심광 또는 선명한 조명이 있을 경우에 장면 전환으로 판단하는 오류를 발생시킬 수 있

고, 영상 내 객체의 위치는 전혀 고려하고 있지 않기에 거의 같은 색상 분포를 가지는 두 장면을 구분하지 못한다는 단점을 가지고 있다[10].

또한 카메라의 움직임을 분석하기 위해 Akutsu 등 [11]이 시공간 영상을 제안하였다. 몇 가지 종류의 카메라 동작을 분류하기 위해 수평과 수직 방향 그리고 기타 여러 방향으로 시공간 영상을 생성하여 물체의 이동이나 카메라의 동작을 Hough 변환으로 찾는 방법이다.

본 논문에서는 컷(cut)의 정확한 검출을 위해 특정한 패턴으로 객체의 움직임을 내포하는 2차원 시공간 영상을 생성하고 에지(edge) 정보를 사용함으로써, 객체의 움직임과 빛의 영향을 최소화하여 장면 전환의 후보 영역을 선택하는 방법을 제안한다. 그리고 선택된 후보 영역을 히스토그램 비교법으로 분석하여 샷의 경계를 검출하는 계층적인(hierarchical) 방법으로 처리 시간을 최소화한다.

2. 장면 전환의 후보 영역 선택

본 시스템은 두 단계로 컷 검출이 이루어진다. 첫번째 단계는 시공간 영상을 활용하여 장면 전환의 후보 영역을 선택하는 것이고, 두번째 단계에서 후보 컷들 중 히스토그램 비교법을 활용한 정확한 컷을 검출하게 된다. 본 장에서는 첫 단계인 장면 전환의 후보 영역을

선택하는 방법에 대해 기술한다.

2.1 시공간 영상의 생성

시공간 영상은 임의의 화소가 일정한 시간 동안 이동한 경로를 나타내는 영상으로, 각 프레임에서 동일한 위치의 화소들을 축적하여 생성할 수 있다. 즉 그림 1과 같이 동영상의 각 프레임에서 같은 위치에 있는 한 행(row)씩을 추출하여 시공간 영상을 생성하였다.

시공간 영상은 장면의 전환을 쉽게 보여줄 수 있도록 동영상의 각 프레임들로부터 구성되어야 한다. 이를 위해 본 연구에서는 물체의 존재 유무가 장면 전환에 결정적인 영향을 미친다는 생각하에 물체를 포함하는 한 행을 선택한다. 특히, 동영상에서 주인공은 일반적인 행에 위치하는 속성에 따라 각 프레임의 중간 행을 사용하였다.

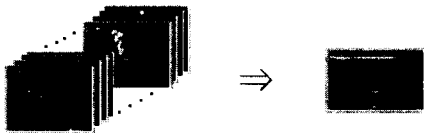


그림 1 시공간 영상의 구성

2.2 시공간 영상에서 특징값의 추출

생성된 시공간 영상은 각 행이 하나의 프레임을 대표하므로 각 행의 특징값을 해당 프레임에서의 특징값으로 하여 장면 전환의 후보 영역을 선택하게 된다.

우리는 특징값 추출을 위해 시공간 영상에서 sobel 필터를 적용하여 수평 성분만을 가지도록 에지 이미지를 생성하였다. 그리고 잡음(noise) 등에 의해 발생한 에지 성분의 제거를 위해 일정길이의 임계값을 사용하여 수평 에지를 임계값 이상인 에지들과 이하인 에지들로 분류했다. 본 연구에서는 의미있는 에지로 간주한 임계값 이상인 에지의 길이를 횡별로 누적시켜 각 프레임의 특징값으로 취하여, 그림 2와 같은 결과를 얻었다. 그림 2의 오른쪽 그래프가 각 프레임의 특징값이 된다.

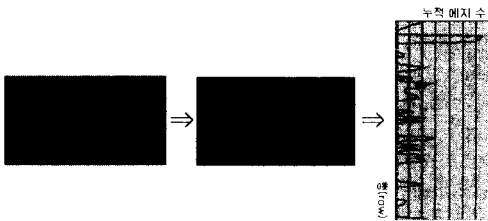


그림 2 시공간 영상에서 특징값의 추출

에지 영상을 이용함으로써 연산량은 증가하나 객체의 움직임을 고려할 수 있고 빛의 영향을 최소화하여

보다 정확한 검출을 할 수 있게 된다.

2.3 장면 전환의 후보 영역 선택

특징값 추출을 위해 3×3 sobel 필터를 적용하였으므로 컷과 인접한 영역에서는 값의 확산으로 인해 비슷한 특징값을 가지게 된다. 그리고 그 주변에서는 각각의 샷 내부에 속하므로 상대적으로 작은 특징값을 갖게 될 것이다. 따라서 인접한 프레임간 특징값의 차이를 구하면 장면 전환의 후보 영역은 V자를 이루게 된다.

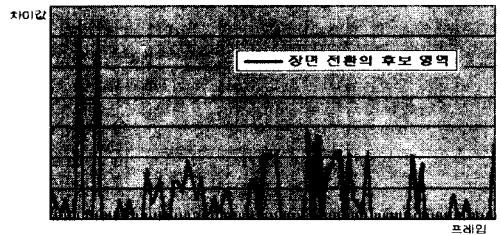


그림 3 프레임간 특징값 차 히스토그램

그림 3은 각 프레임간의 특징값 차이를 히스토그램으로 나타낸 것으로 프레임간 특징값 차이가 각각 임계값 이상이 되는, 그림 3에서 굵은 선으로 표시한, V자를 형성한 4개 지역이 장면 전환의 후보 영역으로 선택된다. 따라서 그림 3에서 보이는 것처럼 원래의 프레임 수 보다 훨씬 적은 양의 프레임을 장면 전환의 지점으로 조사하게 된다.

3. 장면 전환 지점의 검출

제안된 방법으로 선택된 장면 전환의 후보 프레임은 여러 장면 전환 검출 방법 중 프레임간 밝기 차를 이용한 일반적인 히스토그램 비교법으로 분석된다.

장면 전환의 후보 영역으로 선택되어진 프레임들은 먼저 식 (1)의 방법으로 RGB 컬러 공간에서 HSV(I) 공간 중 밝기(intensity)만을 가지도록 변환된다. 그리고 특징값으로 밝기를 가지는 프레임을 히스토그램 비교법으로 검사하여 정확한 장면 변화 지점을 검출한다.

$$I = \frac{1}{3}(R+G+B) \tag{1}$$

히스토그램 비교법을 적용시키는 데 사용되는 식은 다음과 같다

$$SD(i) = \sum_{j=0}^{255} |H_i(j) - H_{i+1}(j)| \tag{2}$$

여기서 $SD(i)$ 는 프레임간 유사성을 측정하는 함수이고 $H_i(j)$ 는 i 번째 프레임의 j 번째 히스토그램 요소 값이다. $SD(i)$ 가 주어진 임계값보다 크면 장면 전환이 발생한 것으로 정의한다.

4. 실험 및 평가

본 논문에서 제안하는 기법의 컷 검출 성능을 평가하

기 위해 Windows 98, Pentium 환경에서 Delphi 언어를 이용하였다. 사용된 동영상 데이터는 AVI형식의 비압축 파일로 160×128 해상도를 가지는 'debbieharry'와 160×120의 'simpsons', 352×240의 'table tennis', 352×240의 'flower garden' RGB 영상이다.

비교의 대상으로 채택한 검출기법은 가장 일반적인 히스토그램 비교법과 동영상 데이터를 균등 분할[12]하여 프레임을 추출하고, 추출된 프레임간 히스토그램 비교법을 통해 미리 정해진 n개의 컷을 검출하는 방법이다.

동영상	검출기법	검출	오검출	미검출	수행시간
debbie-harry(3)	히스토그램	6	4	1	139초500
	균등 분할	4	3	2	39초100
	제안 방법	2	-	1	23초
simpsons(8)	히스토그램	8	-	-	265초460
	균등 분할	8	-	-	84초530
	제안 방법	8	-	-	43초350
table tennis(1)	히스토그램	3	2	-	435초40
	균등 분할	1	-	-	140초400
	제안 방법	1	-	-	43초300
flower garden(0)	히스토그램	2	2	-	608초70
	균등 분할	-	-	-	156초
	제안 방법	-	-	-	16초510

표 1 동영상의 장면 전환 검출 비교

동영상	임계값	히스토그램	균등 분할	제안 방법
		오검출	오검출	오검출
table tennis	11000	2	0	0
	10000	3	0	0
	7000	30	0	0
flower garden	9000	2	0	0
	8000	3	1	0
	5000	57	7	0

표 2 임계값 변화에 의한 오검출

표 1은 동영상을 각 검출 기법으로 실험한 결과이다. 본 논문에서 제안한 기법이 수행시간에서는 확실히 더 좋은 결과를 보이면서, 정확도에서도 떨어지지 않는다는 것을 알 수 있다. 특히 표 2는 zooming과 panning의 카메라 기법이 들어간 영상에서 임계값의 변화로 인한 오검출을 보여주는 것으로 본 논문에서 제안한 방법이 다른 방법에 비해 안정되어 있다.

5. 결론

장면 전환 지점의 검출을 위해 본 논문에서는 시공간 영상과 히스토그램 비교법을 사용하는 계층적인 알고리즘을 제안하였다. 즉 동영상 데이터의 시공간 영상을 생성하고, 생성된 시공간 영상에서 에지 정보를 활용하여 장면 전환의 후보 영역을 선택한다. 그리고 선택된 후보 영역을 장면 전환 지점 검출을 위한 가장 보편적인 히스토그램 비교법을 이용하여 분석하였다.

실험 결과 제안된 방법은 비교적 정확하게 컷을 검출하면서 처리속도 또한 비교되는 다른 검출기법에 비해

약 2-38배 정도로 빠르다는 것을 알 수 있었다. 특히 오검출이나 미검출 양이 현저하게 줄어들었다.

6. 참고 문헌

- [1] John S. Boreczky and Lawrence A. Rowe, "Comparison of video shot boundary detection techniques", in *Storage & Retrieval for Image and Video Databases IV*, Proc. of SPIE 2670, pp.170-179, 1996.
- [2] Ramin Zabih, Justin Miller and Kevin Mai, "A feature-based algorithm for detecting and classifying scene breaks", in *Proc. of ACM Multimedia 95*, pp.189-200, 1995.
- [3] X. Sun, M. S. Kankanhalli, Y. Zhu and J. Wu, "Content-based representative frame extraction for digital video", in *Proc. of IEEE conf. Multimedia Computing and Systems*, pp.190-193, 1998.
- [4] A. Hampapur, R. Jain and T. Weymouth, "Digital video segmentation", in *Proc. of ACM Multimedia 94*, pp.357-364, 1994.
- [5] M. M. Yeung and B. Liu, "Efficient matching and clustering of video shots", in *Proc. of IEEE conf. Image Processing*, pp.338-341, 1995.
- [6] Y. Rui, T. S. Haung and S. Mehrotra, "Exploring video structure beyond the shots", in *Proc. of IEEE conf. Multimedia Computing and Systems*, pp.237-240, 1998.
- [7] H. J. Zhang, C. Y. Low, S. W. Smoliar, and J. H. Wu, "Video parsing, retrieval and browsing: an integrated and content-based solution", in *Proc. of ACM Multimedia 95*, pp.15-24, 1995.
- [8] 이재현, 장옥배, "움직임 벡터를 사용한 점진적 장면 전환 검출", *한국정보과학회 논문지(C)*, 제 3권 제 2호, pp.221-230, 1997.
- [9] Irena Koprinska and Sergio Carrato, "Video segmentation of MPEG compressed data", in *Proc. of IEEE conf. Multimedia Computing and Systems*, pp.243-246, 1998.
- [10] 이미숙, 황본우, 이성환, "내용 기반 영상 및 비디오 검색 기술의 연구 현황", *한국정보과학회지*, 제 15권 제 9호, pp.10-18, 1997.
- [11] A. Akutsu and Y. Tonomura, "Video tomography: an efficient method for camerawork extraction and motion analysis", in *Proc. of ACM Multimedia 94*, pp.349-356, 1994.
- [12] 김우생, "계승된 특징 차이를 이용한 효율적인 대표키 프레임과 관련 키 프레임 추출 기법", *한국정보과학회논문지(B)*, 제 26권 제 5호, pp.657-664, 1998.