

누적 히스토그램과 에지 정보를 이용한 장면 전환 검출

황 두 선, *이 종 설, *조 위 덕, 문 영 식
한양대학교 컴퓨터공학과, *전자부품연구원
전화 : 031-407-8991 / 핸드폰 : 019-255-7868

Scene Change Detection Using Cumulative Histogram and Edge Information

Doo Sun Hwang, Jong Sul Lee, We Duke Cho, Young Shik Moon
Dept. of Computer Science Engineering, Hanyang University
E-mail : dswang@cse.hanyang.ac.kr

Abstract

Automatic video partitioning is the first step for content-based indexing and retrieval of video data. In this paper, an efficient algorithm for scene change detection is proposed, where cumulative histogram and edge information are utilized. Experimental results have shown the effectiveness of the proposed algorithm.

I. 서론

멀티미디어 기술과 통신 기술의 발달은 컴퓨터의 이용폭을 훨씬 증가시켰고 이런 기술의 발전과 더불어 다양한 기술들을 하나로 통합시키려는 움직임이 활발히 전개되고 있다. 기존의 데이터베이스 분야와 멀티미디어 분야가 결합된 디지털 도서관(digital library)이나 VOD(Video on Demand)시스템, 영상 데이터베이스 시스템 등이 대표적인 통합기술의 예라 할 수 있다. 디지털 도서관은 영상과 소리 등 다양한 멀티미디어 데이터 정보를 통합시킨 데이터베이스 시스템으로 최근 멀티미디어 시대의 이슈로 떠오르고 있다. 정보화 사회에서, 낱말이 쏟아져 나오는 정보들을 다양한 형태로 잘 정돈해두어 원하는 정보를 빠른 시간 내에 찾아내도록 하기 위해서 디지털 도서관이나 영상 데이터베이스 시스템은 사용자에게 편리하고 쉬운 인터페이스를 제공해야 한다. 이를 위해 데이터베이스 시스템 내

에 구축될 색인(indexing)이 체계적이어야 한다.

특히 디지털 영상 도서관 시스템에 있어서 멀티미디어 정보인 영상이나 소리 데이터의 색인은 더욱 중요하다. 왜냐하면 멀티미디어 정보를 검색하기 위해 기존의 단순한 키워드 질의 방식으로 정보를 검색하기에는 멀티미디어 데이터가 가지고 있는 정보가 너무 크기 때문이다. 이런 문제를 고려하여 정보량이 많은 영상 데이터의 경우에는 내용 기반 색인(content based indexing)방법을 사용한다. 내용 기반 색인을 위해서 특히 대용량의 동영상 데이터는 전송 대역폭과 저장 용량의 한계 때문에 적절한 단위로 분할될 필요가 있다[1][2][3]. 결국 효율적인 동영상 분할과 분석 등을 통해 내용에 기반한 색인이 가능해지고 나아가 사용자 질의 방식에 유연성을 부여할 수 있다.

본 논문에서는 멀티미디어 데이터 중에서 동영상(video) 데이터를 대상으로 한 효율적이 장면 전환 검출(scene change detection) 방법에 대해 기술한다. 각 절에서 기술할 내용은 다음과 같다. 2절에서는 기존 연구와 문제점을 기술하고, 3절에서는 제안한 장면 전환 알고리즘을 설명한다. 4절에서는 제안한 동영상 분할 방법을 실험하고 기존의 장면 전환 검출 방법과 비교하여 그 효율성과 성능을 확인한다. 5절에서는 제안한 방법에 대한 결론과 향후 연구 방향을 기술한다.

II. 장면 전환 검출

동영상 데이터베이스 시스템에서 내용 기반 색인을 위해 동영상 데이터는 적절한 단위로 분할되어야 한다. 동영상은 정지영상과 달리 공간 정보와 시간 정보를 동시에 포함하고 있다. 따라서 공간 정보 분석을 위해서는 동영상 프레임(video frame)의 공간 밝기 분포 등을 분석하고 시간 정보 분석을 위해서는 동영상 프레임 상호 간의 시간적인 관계성을 분석해야 한다. 공간 정보를 포함하고 있는 한 장의 화면은 동영상 프레임 시퀀스 중 하나의 프레임에 해당하므로 각 프레임은 동영상 시퀀스를 구성하는 단위이다. 그러나 프레임 단위는 전체 동영상을 이해하는 단위는 되지 못한다. 따라서 동영상을 이해하기 위해서는 시간 정보를 분석해서 프레임간의 연속성을 파악하고 일정한 단위로 처리해야 할 필요가 있다[1][2][4].

한 예로 어떤 사용자가 비디오를 선택할 때 그 사용자는 하나의 프레임이 아닌 여러 장면들을 연상하게 될 것이기 때문에 영상 데이터베이스 시스템은 연속된 프레임 시퀀스를 브라우징하는 기능을 제공해야 한다. 이때 연속된 프레임 시퀀스는 한 샷(shot)이 되고 샷은 동영상 분할의 기본 단위가 된다. 여기에서 장면의 의미를 구체적으로 전달하기 위해 몇 개의 샷을 묶어 장면(scene)이나 에피소드(episode)를 구성하기도 한다. 동영상 분할에서 이런 계층적인 분할 및 분석 방법을 일컬어 구조적인 분석(structured analysis)이라 한다 [2]. 동영상의 최소 분할 단위를 샷이라고 하였는데 이는 카메라 조사가 시작된 시점으로부터 연속된 장면을 촬영 후 카메라 조사가 끝날 때까지 기록된 동영상 시퀀스를 말한다.

기존의 동영상 분할 연구 중 Zhang[2][5]은 픽셀 밝기 차이에 의한 방법을 제안하였다. 그리고 전처리 과정을 통해서 움직임 영향을 줄이려 했다. Sharary[5]는 픽셀 밝기 차이를 이용하였지만 영상을 동일한 영역으로 나누어 지역적인 변화도 함께 고려하였다. 픽셀 밝기 차이에 의한 방법은 대체적으로 움직임(motion)이 거의 없는 동영상 시퀀스에 대해서는 샷 경계를 잘 검출해 내지만 카메라나 물체의 움직임에 민감하게 반응한다는 문제점이 있다. 즉 샷 내부에서 카메라나 물체의 움직임 영향으로 발생한 픽셀 차이가 큰 경우 샷 경계로 오검출하는 경우가 많다. 픽셀 밝기 차이에 의한 방법은 차이 측정 속도가 간단하고 구현이 쉬우나 점진 컷을 검출할 만한 정교한 처리를 하지 못하는 문제가 있다. 그래서 히스토그램을 기초로 한 히스토그램 비교법에 관한 연구가 다양하게 진행되어 왔다. 히스토그램 비교법은 픽셀 밝기 차이에 의한 검출법에서 문제가 되었던 움직임에 민감하다는 단점을 어느 정도 극복하고 있다. 즉 히스토그램은 영상의 전체적인 분

위를 파악하는 것으로 영상을 이해하는데 있어서 영상 내에 물체나 카메라의 움직임이 존재하는 경우라도 밝기 분포나 색상 분포는 거의 일정하기 때문이다. Dugad[6]는 전후 프레임의 히스토그램의 누적값과 분산과 평균을 가지고서 두 개의 임계치 값(상한치, 하한치)을 정한다. 상한치보다 크면 장면 전환점이고, 상한치와 하한치 사이이면 블록으로 나누어서 장면 전환점을 찾는다. 히스토그램 비교법은 가장 많이 사용되고 있는 방법이나 공통적으로 몇 가지 문제점을 갖고 있다. 첫 째로 영상에서 차지하는 물체의 비율이 적고 셋 간에 배경의 변화가 없는 경우 또는 서로 다른 셋에 존재하는 프레임이라도 비슷한 밝기 분포를 갖게 되는 경우에는 히스토그램 비교법으로 샷 경계를 검출할 수가 없다. 그리고 히스토그램 비교법도 픽셀 밝기 차이에 의한 방법과 마찬가지로 점진컷을 검출하는 방법으로는 적합하지 않다. Arman[5]은 DCT 계수를 이용하여 프레임간의 상관도를 계산하였다. Yeol[7]은 DCT 계수와 영상 밝기와의 관계를 밝히고 MPEG으로 압축된 데이터로부터 DCT 계수와 움직임 벡터를 이용하여 빠른 시간 내에 샷 경계를 검출하는 방법을 제시하였다.

III. 제안한 방법

본 논문에서는 동영상 데이터의 종류에 상관없는 효과적인 장면 전환 알고리즘을 제안하고자 한다. 픽셀 밝기 차를 이용한 장면 전환 검출 방법의 단점과 히스토그램을 이용한 검출 방법의 단점을 보완하고 점진컷을 검출할 수 있는 방법을 제안한다. 이를 위해 본 논문에서는 Striker[8]가 제안한 누적히스토그램을 사용하였다. 누적히스토그램은 잠음이나 조명 변화에 강한 특징을 가지고 있다.

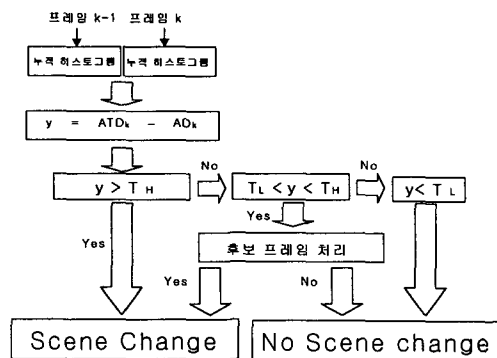


그림 3 장면 전환 검출 알고리즘의 전체 흐름도

컬러 좌표계는 YCbCr 좌표계를 사용한다. 먼저 YCbCr중에서 Y의 값만으로 누적 히스토그램을 구성한다. 정확한 섯을 검출하기 위하여 두 단계로 처리하며, 전체적인 알고리즘은 그림 1과 같다.

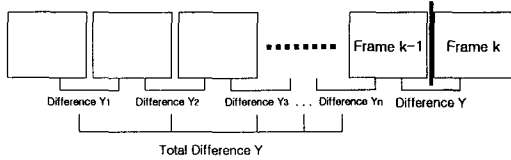


그림 4 누적히스토그램 차이

그림 1에서 장면 전환 검출의 기준이 되는 y 값은 아래와 같은 수식으로 계산된다.

$$CH_K(i) = \sum_{j=0}^i H_K(j)$$

$$D_K = \sum_{j=0}^G |CH_{K-1}(j) - CH_K(j)|$$

$$TD_K = \sum_{i=0}^K D_i$$

$$AD_K = \frac{D_K}{M \cdot N \cdot G}$$

$$ATD_K = \frac{TD_K}{M \cdot N \cdot G \cdot K}$$

$$y = ATD_K - AD_K$$

여기서 M, N은 영상의 크기, G는 최대 밝기 값, $H_K(j)$ 는 프레임 K에 대한 히스토그램, $CH_K(i)$ 는 프레임 K에 대한 누적 히스토그램을 의미한다.

위 수식으로 계산된 y값이 임계치 T_H 보다 크면 장면 변환점으로 판단한다. y값이 T_H 보다는 작지만 T_L 보다 클 경우는 일단 장면 전환의 후보점으로 간주하여 다음의 처리 과정을 거친다.

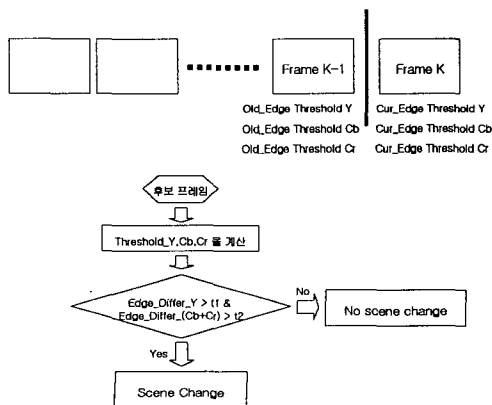


그림 3 후보 프레임 처리 알고리즘

두 번째 단계에서는 컬러 에지를 이용한다. 컬러 에지는 Fan[9]이 제안한 엔트로피를 이용한 컬러 에지를 찾기 위한 동적 임계값을 이용한다. 후보 프레임 처리를 위한 알고리즘은 그림 3과 같다.

점진식을 만들어내는 편집 효과에는 페이드인(fade in)과 페이드아웃(fade out), 디졸브(dissolve), 와이프(wipe)등이 있다. 이렇게 변하는 장면 전환점을 찾기 위하여 YCbCr 컬러 좌표계에서의 Y의 분산값과 에지를 이용한다.

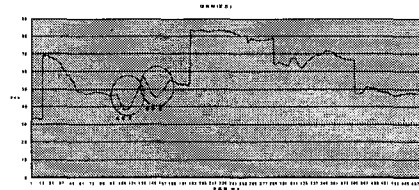


그림 4 프레임 별 Y의 분산값

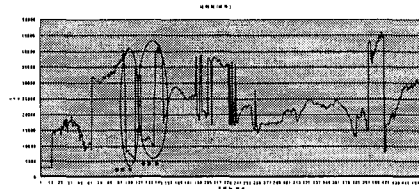


그림 5 프레임 별 에지 수

그림 4, 그림 5에서 보는 바와 같이 분산값이 순차적으로 작아 지다가 커지는 구간, 에지의 수가 급격히 낮아지다가 평탄구간이 발생하는 지점에서 점진적으로 장면 전환이 일어난다.

IV. 실험 결과

장면 전환 검출 실험을 위해 표 1에서 보는 바와 같이 세 가지 카테고리(뉴스, 상업적인 광고, 드라마)의 데이터를 사용하였다.

표 5 실험 사용된 비디오

종류	장면	크기	전체프레임	컷수
광고	남양여우야	352x240	462	11
	레모니아	352x240	461	13
	애니테이	352x240	913	23
	코렉스물	352x240	460	12
뉴스	SBS 2001/3/15	352x240	79841	465
드라마	그 여자네 집	352x240	2729	16

표2 에서 볼 수 있듯이 제안한 방법의 성능이 전반

적으로 더 우수하게 나오는 것을 알 수 있다. 광고(레모니아)에서는 카메라의 모션이 있어서 픽셀 방법의 결과에 오검출이 많이 발생했다. 광고(코렉스물)에서 점진 셋(페이드인, 페이드아웃)을 검출 해낼 수 있었다. 뉴스에서는 기자들의 취재 장면에서 많이 나오는 카메라 플래쉬에 의해서 오검출이 발생했다.

평균적으로, 제안한 방법의 precision은 81%로서 기존의 방법인 픽셀의 51%보다 30% 향상되었고, 히스토그램의 72%보다 9% 향상되었다. recall은 94%로서 기존의 방법과 거의 유사한 성능을 보이고 있다.

표 6 장면 전환 방법의 성능 평가

장면	방법	all	correct	false	miss	precision	recall
남양여우야	픽셀	11	11	0	0	1.0	1.0
	히스토그램	16	11	5	0	0.69	1.0
	제안한방법	12	11	1	0	0.92	1.0
레모니아	픽셀	45	13	32	0	0.29	1.0
	히스토그램	10	10	0	3	1.0	0.77
	제안한방법	13	12	1	1	0.92	0.92
애니메이션	픽셀	51	21	30	2	0.41	0.91
	히스토그램	30	18	12	5	0.6	0.78
	제안한방법	26	22	4	1	0.85	0.96
코렉스물	픽셀	42	11	31	1	0.28	0.92
	히스토그램	31	7	24	5	0.23	0.58
	제안한방법	26	9	17	3	0.35	0.75
뉴스	픽셀	1622	462	1160	3	0.26	0.99
	히스토그램	543	423	120	42	0.78	0.91
	제안한방법	519	460	59	5	0.89	0.99
드라마	픽셀	19	16	3	0	0.84	1.0
	히스토그램	15	15	0	1	1.0	0.94
	제안한방법	16	16	0	0	1.0	1.0

V. 결론 및 향후 과제

본 논문에서는 누적 히스토그램과 에지 정보를 이용한 장면 전환 검출법을 제안하였다. 실험 결과 제안한 방법의 성능이 우수함을 알 수 있었으며, 점진적인 장면의 검출과 조명에 대한 문제가 향후 연구할 과제로 남아 있다.

본 논문은 산업자원부의 중기거점사업으로 진행된 연구결과 이다.

참고문헌

[1] Yoshinobu Tonomura, Akihito Akutsu, Yukinobu Taniguchi, and Gen Suzuki, "Structured Video

Computing," IEEE Multimedia, pp. 34-43, Fall 1994.

[2] Stephen W. Smoliar and HongJiang Zhang, "Content-Based Video Indexing and Retrieval," IEEE Multimedia, pp. 62-72, Summer 1994.

[3] Howard D. Wactlar, Takeo Kanade, Michael A. Smith and Scott M. Stevens, "Intelligent Access to Digital Video : Informedia Project," IEEE Computer, pp. 46-52, May 1996.

[4] William I. Grosky, Ramesh Jain and Rajiv Mehrotra, The Handbook of Multimedia Information Management, Prentice Hall, 1997.

[5] John S. Boreczky and Lawrence A. Rowe, "Comparison of Video Shot Boundary Detection Techniques," Proc. of IS&T/SPIE 1996 Symp. on Elec. Imaging: Science and Technology, February 1996.

[6] Dugad R, Ratakonda K. and Ahuja N. "Robust Video Shot Change Detection", Proc. 1998 IEEE Second Workshop on Multimedia Signal Processing, 1998.

[7] Boon-Lock Yeo and Bede Liu, "Rapid Scene Analysis on Compressed Video", IEEE Trans. on Circuit and System for Video Technology, Vol. 5, No. 6, pp. 533-544, December 1995.

[8] M. A. Stricker and m. Orengo, "Similarity of Color Images", Proc. Storage and Retrieval for Image and Video Databases(SPIE), pp. 381-392, 1995.

[9] Jianping Fan, David. K. Y. Yau, Ahmed. K. Elmagarmid, Walid G. Aref, "Automatic Image Segmentation by Integrating Color-Edge Extraction and Seeded Region Growing", IEEE Trans. on Image Processing, Vol. 10, No. 10, pp. 1454-1466 October 2001.

[10] Dalong Li, Hanqing Lu, "Avoding False Alarms due to Illumination Variation in Shot Detection", Proc. 2000 IEEE Workshop on Signal Processing Systems, pp. 828-836 2000.