

웨이블릿 변환영역에서 칼라 히스토그램과 에너지 벡터를 이용한 컷 검출

김수정⁰ 정성환

창원대학교 전자계산학과
sue95@cdcs.changwon.ac.kr

Cut Detection Using Color Histogram and Energy Vector in Wavelet Transform Domain

Su-Jung Kim⁰ Sung-Hwan Jung

Dept. of Computer Science, Changwon National University

요 약

본 논문은 웨이블릿 변환영역 하에서 칼라 히스토그램과 에너지 벡터를 이용한 컷 검출 방법을 제안한다. 기존 컷 검출 방법들은 대부분 공간영역과 변환영역 각각에 대한 특징을 이용해 컷을 검출하였다. 그러나 본 논문에서는 웨이블릿 변환영역 하에서도 공간영역 특성을 유지하는 LL밴드 상의 칼라 히스토그램과 LH와 HL밴드의 에너지 값을 변환영역 특성으로 함께 고려하였다. 최근 영상 압축 표준에 웨이블릿을 이용한 압축기법이 사용되고 있으므로, 제안한 방법은 웨이블릿 압축 영상에서 압축을 해제할 필요 없이 컷을 검출하는데 사용되어질 수 있다.

제안한 방법의 성능평가를 위하여 광고, 뉴스, 스포츠, 영화 등 5개 분야의 다양한 TV 프로그램에서 약 10,000 개의 프레임으로 실험한 결과, *Recall*에서는 약 90%, *Precision*에서는 약 94%의 컷 검출 성능을 나타내었다.

1. 서 론

최근 대용량, 고속의 데이터 전송이 가능한 초고속 통신망 및 멀티미디어 관련 기술의 발달로 인하여 영상, 비디오, 음성 데이터와 같은 멀티미디어 정보의 효율적인 전송, 저장, 관리 및 검색이 중요한 핵심 기술로 대두되고 있다. 특히 동영상은 비정형 데이터이고 그 양이 방대하여 정보의 검색 및 브라우징이 매우 곤란하다는 문제점을 가지고 있다. 따라서 이러한 문제점을 해결하기 위해 동영상 데이터를 보다 관리하기 쉬운 단위로 분할하여 구조화할 필요가 있다.

동영상의 구조는 크게 4종류의 계층적 데이터 모임을 구성되는데, 최하위 계층은 개개의 정지영상을 가리키는 프레임(Frame)이고, 한 카메라의 연속적인 촬영에 의한 프레임의 모임을 샷(Shot)이라 한다. 그리고 내용상 관련이 있는 인접한 샷의 집합을 장면(Scene)이라 하고, 한 주제에 대한 연속적인 장면들의 집합을 동영상의 최상위 계층인 시퀀스(Sequence)라 한다.

이와 같이 동영상 데이터를 구조화함으로써 원하는 부분을 쉽게 찾아내거나 볼 수 있을 뿐만 아니라, 검색을 위해 주석을 달거나 특징을 추출하는 기본적인 단위로서 활용이 가능하다. 특히 동영상구조에서 샷은 앞서 언급했듯이, 하나의 카메라 동작이 끝나고 다른 카메라로 넘어가기까지의 프레임 집합이므로, 특징을 이용해 자동으로 분할할 수 있는 부분이다.

동영상 즉, 비디오를 분할하기 위한 연구는 크게 주석기반 방법과 내용기반 방법으로 나눌 수 있다.

주석기반 방법은 사용자가 직접 비디오를 보면서 내용을 텍스트 형식으로 입력하는 방법으로 비디오 내용을 정확하게 모델링하고 검색할 수 있으나, 많은 노력과 시간이 필요로 한다는 단점이 있다. 반면, 내용기반 방법은 주석기반 방법보다 부정확하게 비디오를 분류하지만, 자동적으로 분류할 수 있다는 장점을 가진다.

본 논문은 내용기반 비디오 검색의 기초가 되는 샷 경계 검출 방법에 대한 연구이다. 여기서 샷 경계를 컷(Cut)이라 하며, 이러한 컷을 검출하기 위한 기존의 연구방법으로는 공간영역의

특정인 각 화소간의 차를 이용한 방법[1], 히스토그램의 차이를 비교하는 방법[2,3], 분산과 표준편차 등의 통계적인 차를 이용한 방법[4], 움직임 벡터 혹은 형태 기반 검출을 사용하는 방법[1,5], 변환영역의 특징인 DCT 변환 후 DC계수 차를 이용한 방법[5,6] 등이 있다.

이러한 기존의 방법들은 공간영역과 변환영역에서 각각 특징을 추출하여 개별적으로 사용하거나 혹은 이들을 결합하여 컷을 검출하였다. 그러나 본 논문에서는 웨이블릿 변환영역 상에서 공간영역 특징인 칼라 히스토그램과 변환영역 특징인 에너지 벡터를 함께 이용한 컷 검출 기법을 제안한다. 이것은 현재 MPEG-4 VTC를 비롯한 압축 표준에서 웨이블릿을 기본적으로 사용하고 있으므로[7] 제안된 방법은 압축 도메인에서 압축을 해제할 필요 없이 컷을 검출하는데 사용되어질 수 있다.

본 논문의 구성은 다음과 같다. 2장에서는 기존의 제안된 컷 검출 방법들에 대해 간단히 살펴보고, 3장에서는 제안한 웨이블릿 변환영역에서의 칼라 히스토그램과 에너지 벡터를 사용한 컷 검출 방법에 대해 설명한다. 그리고 4장에서는 기존의 컷 검출 방법과 제안한 방법을 비교 실험한 결과를 제시하며, 마지막으로 5장에서는 결론 및 향후 연구과제 등을 기술한다.

2. 기존의 컷 검출 기법

2.1 화소단위 비교방법

화소단위 비교방법은 동일한 샷 내에서는 화소값의 변화가 적다는 성질을 이용한 것으로, 연속하는 한 쌍의 프레임에서 대응하는 화소값을 비교하여 얼마나 많은 변화가 발생하였는가를 측정한다[1]. 이 방법은 구현이 간단한 반면, 카메라 연산과 움직임에 민감한 반응을 일으킨다.

$$DP_i(x, y) = \begin{cases} 1 & \text{if } |F_i(x, y) - F_{i+1}(x, y)| > t \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

$$\frac{\sum_{x, y} X, Y DP_i(x, y)}{X \times Y} \times 100 > T \quad (2)$$

식(1)에서 $F_i(x, y)$ 는 i 번째 프레임에서 좌표 (x, y) 의 화소값이고, 이 때 인접하는 프레임의 대응하는 화소값 차가 최소 밝기값 t 를 초과할 경우 1로 세트한다. 식(2)에서 전체 화소에 대해 1로 세트된 화소의 비율이 특정 임계값 T 를 넘으면, 컷으로 간주한다.

2.2 히스토그램 비교방법

히스토그램 비교방법은 동일한 샷 내의 프레임들은 서로 유사한 색상분포를 가진다는 특성을 이용한 가장 보편적인 컷 검출 방법이다. 이 방법은 식(3)과 같이 비디오 영상에서 인접한 프레임들의 히스토그램 차이를 계산하여 주어진 임계값 T 와 비교함으로써 컷을 검출한다[2].

$$\sum_{j=1}^N |H_i(j) - H_{i+1}(j)| > T \quad (3)$$

식(3)에서 N 은 히스토그램 레벨 수이고, $H_i(j)$ 는 i 번째 프레임인 j 레벨의 히스토그램 값을 나타낸다.

히스토그램 비교방법은 화소단위 비교방법에 비해 물체의 움직임이나 카메라 이동에 덜 민감하지만, 빛의 영향에 민감한 반응을 보이고 다른 샷이지만 색상분포가 같은 경우를 검출해 내지 못하는 단점을 지닌다.

2.3 통계적인 차를 이용한 방법

통계적인 차를 이용한 방법은 이웃하는 프레임 간의 평균, 분산, 표준편차 등과 같은 통계적인 특징값을 기초로 하여 식(4)와 같이 컷을 검출한다[4].

$$\frac{[(\frac{\sigma_i^2 \times \sigma_{i+1}^2}{2}) + (\frac{\mu_i - \mu_{i+1}}{2})^2]}{\sigma_i^2 \times \sigma_{i+1}^2} > T \quad (4)$$

식(4)에서 μ_i, μ_{i+1} 는 연속하는 두 프레임에 대한 평균이고, $\sigma_i^2, \sigma_{i+1}^2$ 는 각 프레임에 대한 분산이다.

2.4 변환영역 특징을 이용한 방법

이 방법은 FFT, DCT와 같은 변환영역의 특징을 이용하여 컷을 검출하는 방법으로, MPEG과 같은 압축 도메인에서 디코딩을 하지 않거나 혹은 최소한의 디코딩만을 수행한 후, 식(5)와 같은 DCT변환의 DC계수를 이용해 컷을 검출할 수 있다[5].

$$C(u, v) = a(u) a(v) \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f(i, j) \cdot \cos(\frac{(2i+1)u\pi}{2N}) \cdot \cos(\frac{(2j+1)v\pi}{2N}) \quad (5)$$

단, $u, v, i, j = 0, 1, 2, \dots, N-1$

3. 웨이블릿 변환영역에서 칼라 히스토그램 및 에너지 벡터를 이용한 컷 검출

(그림 1)은 본 논문에서 제안한 웨이블릿 변환영역 하에서 칼라 히스토그램과 에너지 벡터를 이용한 컷 검출 방법의 전체 블록도이다.

입력받은 비디오 프레임을 RGB 채널별로 분리한 후, 각각 1단계 웨이블릿 변환을 수행한다. 그리고 (그림 1)에서와 같이 LL밴드와 LH 및 HL밴드로 분리 처리한 후, 통합 특징값 비교를 통하여 최종적으로 컷을 검출한다.

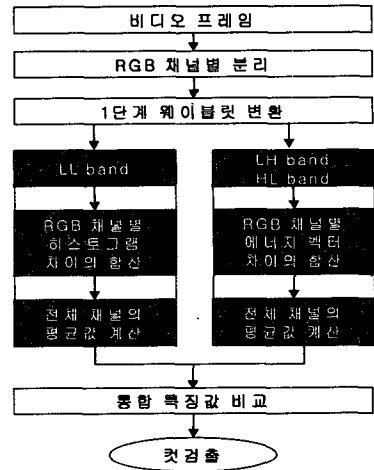
각 단계의 처리과정을 아래에서 기술한다.

3.1 웨이블릿 변환

웨이블릿 변환은 영상에 대하여 다해상도 분석이 가능하므로, 서로 다른 해상도로 영상을 분할하여 개별적으로 부영상을 분석할 수 있다는 특징을 가지고 있다.

예로서 1단계 웨이블릿 변환을 수행하면 원영상은 하나의 저대역 부밴드(LL)와 수평, 수직, 대각선의 방향 성분을 가지는 3개의 고대역 부밴드(LH, HL, HH)가 생성된다.

본 논문에서는 1단계 웨이블릿 변환을 수행한 후, 공간영역 특성을 가진 LL 저대역 부밴드에서 칼라 히스토그램을 추출하고, LH, HL 고대역 부밴드에서 에너지 벡터를 추출하여 컷 검출에 이용한다.



(Fig. 1) The block diagram of the proposed cut detection method

3.2 LL밴드에서 칼라 히스토그램 차의 평균 계산

칼라 히스토그램은 가장 일반적으로 사용되는 컷 검출의 방법으로, 동일한 샷 내에서는 특징값의 차이가 거의 없고 카메라 움직임의 변화에 덜 민감하다는 성질이 있다. 여기서는 웨이블릿 변환을 수행한 후, 식(6)에서와 같이 공간영역 정보를 가지는 LL밴드에서 칼라 히스토그램을 추출하여 연속한 프레임 간의 차이를 각 칼라 채널별로 계산하고 그 평균을 구하였다.

$$H_i = \frac{\sum_{j=1}^L (|R_i(j) - R_{i+1}(j)| + |G_i(j) - G_{i+1}(j)| + |B_i(j) - B_{i+1}(j)|)}{X \times Y} \quad (6)$$

식(6)에서 L 은 칼라 레벨의 수이고, $X \times Y$ 는 LL밴드의 크기를 나타낸다. 그리고 R_i, G_i, B_i 는 i 번째 프레임의 각 RGB 채널별 j 레벨의 히스토그램 값이다.

3.3 LH밴드와 HL밴드에서 에너지 벡터 계산

칼라 히스토그램은 조명의 변화에 민감하다는 단점이 있다. 이러한 단점에 의한 잘못된 컷 검출을 보완하기 위하여 본 논문에서는 웨이블릿 변환영역 상에서 에너지 벡터를 컷 검출의 특징값으로 추가하였다.

식(7)은 LH와 HL밴드의 칼라 채널별 에너지 벡터를 구하는 식을 나타낸다.

$$E_{k,c} = \frac{1}{m \times n} \sum_m \sum_n W_{k,c}^2 [m, n] \quad (7)$$

여기에서 $k \in K, c \in C$ 이고, $K = \{LH, HL\}, C = \{R, G, B\}$ 이다. $W_{k,c}$ 는 웨이블릿 변환 후 각 부밴드의 칼라 채널별 변환계수를 가리킨다. 그리고 m, n 은 변환계수의 위치, $E_{k,c}$ 는 각 부밴드의 칼라 채널별 에너지 벡터를 나타낸다.

본 논문에서는 위에서 구한 에너지 벡터를 이용하여 식(8)에서와 같이, 연속한 프레임 간의 에너지 벡터 차의 평균 (E_i)를 특징값으로 구하였다.

$$E_i = \frac{\sum_k \sum_c |E_{k,c}(i) - E_{k,c}(i+1)|}{6} \quad (8)$$

식(8)에서 i 는 프레임 번호를 가리킨다.

3.4 칼라 히스토그램과 에너지 벡터를 이용한 컷 검출

본 논문에서는 정확한 컷 검출을 위하여, 식(6)과 식(8)의 두 특징값을 이용한 다음의 두 조건을 적용하였다.

$$findCutH = H_i > T_h \quad (9)$$

$$findCut = findCutH \text{ and } E_i > T_e \quad (10)$$

여기서 T_h 와 T_e 는 각각 칼라 히스토그램에 대한 임계값과 에너지 벡터에 대한 임계값을 가리킨다.

식(9)는 칼라 히스토그램 차의 평균 (H_i)이 임계값 (T_h)보다 높을 경우 컷을 검출하는 조건이다. 식(10)은 식(9)에서 잘못 검출된 컷을 보완하기 위하여 에너지 벡터 차의 평균을 특징값으로 추가한 조건이다.

4. 실험 결과

본 논문에서는 160×120, 240×180 등 다양한 크기와 초당 15 또는 30 프레임을 가지는 AVI 비디오 테이터를 128×128 크기의 프레임으로 정규화하여 실험하였다. <표 1>은 각 실험 테이터의 실제 컷 수 및 프레임 수를 나타낸다.

<Table 1> Video data for the experiment

종류	AVI 파일 수	실제 컷 수	프레임 수
광고	9	112	2,722
뉴스	10	44	3,070
스포츠	7	39	1,803
영화	5	69	1,634
기타	3	23	654
합계	34	287	9,883

본 논문에서 제안한 두 특징값에 대한 컷 검출의 적합함을 알아보기 위하여 두 가지 실험을 하였다. 먼저 식(9)의 조건을 적용하여 변환영역에서의 칼라 히스토그램만을 이용한 컷 검출을 실험하였다. 다음으로 식(10)의 조건을 적용하여 변환영역에서의 칼라 히스토그램과 에너지 벡터를 특징값으로 이용한 컷 검출 실험을 하였다.

컷 검출의 성능 비교를 위해 *Recall*과 *Precision*의 두 가지 성능평가 척도를 사용하였다.

$$Recall = \frac{N_c}{N_c + N_m} \quad (11)$$

$$Precision = \frac{N_c}{N_c + N_f} \quad (12)$$

위 식에서 N_c 는 정확히 검출된 컷 수를 가리키며, N_m 는 검출되지 못한 컷 수, N_f 는 잘못 검출된 컷 수를 나타낸다.

식(11), (12)의 성능평가 척도를 기준으로 하여, 최근에 발표된 칼라 히스토그램과 DCT계수를 이용한 기존 방법[6], 본 논문에서 제안한 웨이블릿 변환영역에서 칼라 히스토그램만을 이용한 방법과 칼라 히스토그램과 에너지 벡터를 함께 이용한 방법의 성능을 각각 비교하여 <표 2>에 나타내었다.

<Table 2> Experimental result of cut detection methods

주제별 방법	칼라히스토그램과 DCT계수		칼라히스토그램		칼라히스토그램과 에너지 벡터	
	R	P	R	P	R	P
광고	0.97	0.90	0.78	0.91	0.86	0.96
뉴스	0.84	1.00	0.89	1.00	0.93	1.00
스포츠	0.79	0.86	0.95	0.95	0.95	0.93
영화	0.81	0.88	0.84	0.85	0.91	0.86
기타	0.83	0.86	0.96	0.81	0.91	1.00
검출률	88%	90%	85%	90%	90%	94%

*는 웨이블릿 변환영역에서의 방법, R은 *Recall*, P는 *Precision*을 가리킨다.

<표 2>에서, 기존 방법으로 검출한 결과가 웨이블릿 변환영역 상에서 칼라 히스토그램만을 이용해 컷을 검출한 결과보다 *Recall*이 약간 높고, *Precision*은 같았다. 그러나 변환영역 상에서 칼라 히스토그램에 에너지 특징값을 추가한 방법은 *Recall*과 *Precision*에서 각각 90%, 94%의 검출율을 보이면서, 기존 방법보다 2%, 4%로 향상된 결과를 나타냈다. 이는 변환영역 하에서 칼라 히스토그램과 에너지 벡터를 함께 사용함으로써 각각의 개별적인 영역의 특성을 사용하는 것보다 공간과 변환영역에서의 특성을 동시에 고려함으로써 개선된 성능을 얻을 수 있었다.

5. 결론 및 향후 연구과제

본 논문에서는 웨이블릿 변환영역에서의 칼라 히스토그램과 에너지 벡터를 이용한 컷 검출 방법을 제안하였다.

이 방법은 웨이블릿 변환을 수행한 후, 공간특성을 가지는 LL밴드에서 칼라 히스토그램을 추출하고, 동시에 LH와 HL밴드의 특징으로 에너지 벡터를 사용함으로써 기존의 방법에 비하여 효과적인 컷 검출이 가능하다.

본 논문에서 제안한 방법의 성능을 측정하기 위해 *Recall*과 *Precision* 성능평가 척도를 이용하여 다양한 5종류의 비디오 테이터에 대해 실험해 본 결과, *Recall*은 약 90%, *Precision*은 약 94%의 검출 성능을 보였다. 제안된 방법은 공간영역 혹은 변환영역에서의 특징값을 개별적으로 이용한 기존의 방법보다 좀 더 나은 결과를 나타내었다. 그리고 MPEG-4 VTC를 비롯한 최근 압축 표준들이 웨이블릿 변환을 이용하고 있으므로, 본 논문에서 제안한 방법은 압축된 영상에서 압축을 해제할 필요 없이 컷을 검출하는데 사용되어질 수 있다.

향후과제로는 검출 성능을 좀더 향상시키는 방법을 모색하면서, 비디오에서 유사한 주제를 가지고 있는 인접한 샷의 모임을 구성된 장면 경계 검출을 연구하는 것이다.

참고 문헌

- [1] John S. Boreczky, Lawrence A. Rowe, "Comparison of video shot boundary detection techniques," in Storage & Retrieval for Image and Video Database IV, Proc. of SPIE 2670, pp. 170-179, 1996.
- [2] Nilesh V. Patel and Ishwar K. Sethi, "Video shot detection and characterization for video databases," Pattern Recognition, Vol. 30, No. 4, pp. 583-592, April 1997.
- [3] 염성주, 김우생, "동적 영역 히스토그램을 사용한 동영상 데이터의 컷검출 방법," 정보과학회논문지(A) Volume 25, No.3 pp. 221-230 Mar. 1998.
- [4] I. K. Sethi and N. Patel, "A statistical approach to scene change detection," SPIE Storage and Retrieval for Image and Video Databases, Vol. 2420, 1995.
- [5] J. Meng, Y. Juan, and S-F Chang, "Scene change detection in a MPEG compressed video sequence," IS&T/SPIE Symposium Proceedings, Vol. 2670, San Jose, California, Feb. 1996.
- [6] 정해준, 이우선, 정성환, "칼라 히스토그램과 DC계수를 이용한 비디오 세그먼트이션," 정보처리학회, CD-paper #43, Oct, 1999.
- [7] Diego Santa-Cruz, Touradj Ebrahimi, "A study of JPEG 2000 still image coding versus other standards," To be published in the Proc. of the X European Signal Processing Conference, Tampere, Finland, September 2000. <http://Itswww.epfl.ch/~dsanta>
- [8] Arman, A.Hsu, and M.Y.Chiu, "Image processing on compressed data for large video databases," Proceedings of ACM Multimedia Conference, June 1993.
- [9] 고병철, 이해성, 변해란, "웨이블릿 변환을 이용한 비디오 파노라마 영상 구성," HCI '99 학술대회 발표 논문집, pp. 187-190, 한국정보과학회 HCI 연구회, 1999.