

화소와 히스토그램 정보를 이용한 샷 전환 탐지 알고리즘

이준구*, 한기선**, 유병문***, 황두성*

*단국대학교 컴퓨터과학과

**강동대학교 방송영상미디어과

***(주)엘앤와이비전

e-mail: leejg01679@hanmail.net

Shot Boundary Detection Algorithm By Using Pixel and Histogram Information

Joon-Goo Lee*, Ki-Sun Han**, Byoung-Moon You***, Doo-Sung Hwang*

*Dept. of Computer Science, Dan-Kook University

**Dept. of Visual Broadcasting Media, Gang-Dong College

***L&Y Vision Technologies, Inc.

요 약

비디오 데이터를 효율적으로 검색, 정렬, 탐색, 분류하기 위해서는 프레임 간의 샷 전환 탐지가 선행되어야 한다. 본 논문에서는 디지털 비디오 데이터의 샷 전환 탐지를 위해 비디오 스트림을 구성하고 있는 각 프레임들 간의 화소 밝기 차이와 히스토그램의 변화를 이용하였다. 플래쉬 등과 같은 인위적이고 급격한 화소 밝기변화에 의한 오류를 최소화하기 위해 샷 전환 탐지 이전에 각 프레임 간의 밝기 보상을 적용하였다. 밝기 보정된 프레임으로부터 프레임의 서브 블록 간의 지역적 화소 밝기 정보, 그리고 프레임의 화소 밝기 값 히스토그램을 비교하여 샷 전환을 탐지한다. 실험에서 제안된 알고리즘은 국가기록원 소장 비디오에 적용하여 효과가 있음을 보였다.

1. 서론

디지털 기술 발전은 비디오 데이터의 빠른 생성과 더불어 컴퓨팅 자원의 활용을 증가시켰으나, 비디오 데이터 관리를 위한 저장과 검색, 정렬과 탐색, 분류 등 관련 서비스가 효과적으로 제공되고 있지 않다. 비디오는 한 개 이상의 카메라를 사용하여 촬영한 장면(frame)들을 연속적으로 편집하여 구성이 되고, 비디오에서 하나의 카메라에 의해 촬영된 연속적인 장면을 샷(shot), 연속적으로 연결되어 있는 두 샷 사이의 전환되는 부분을 컷(cut)이라고 한다. 하나의 샷은 장면을 촬영한 카메라의 파라미터(동작, 확대, 축소, 초점)와 샷의 공간 및 시간적 연속성에 의해 비디오의 내용면에서 일관성을 가지는 최소의 단위가 된다. 샷 전환 탐지(shot boundary detection)는 대용량 비디오 데이터의 색인(indexing), 검색(browsing), 그리고 사건 탐지 등의 서비스를 효율적으로 제공하기 위한 기반이 된다[1].

샷 전환 알고리즘은 연속 프레임 간의 화소 비교 분석[2,3], 프레임 간 히스토그램 차이 분석[4,5], 객체의 에지 변화 비율(edge change ration) 비교[6], 칼라 모멘트(color

moment) 비교 방법[7] 등이 연구되었다. 샷 전환 탐지 기술은 비디오 데이터의 분할, 색인, 검색, 요약 등 비디오 분석 및 관리를 위해 반드시 갖추어야 할 요소기술이다. 연구된 알고리즘들은 제안하는 특징 벡터의 설정과 유사도 측정, 임계값의 설정, 정확률(precision)과 재현율(recall)을 이용한 테스트 등으로 요약된다. 지금까지 제안된 샷 전환 알고리즘은 임계값의 설정에 따른 탐지율에 의존되고 있으며, 높은 탐지율을 보장하는 적응적 탐지 방법에 대한 연구는 진행 중이다.

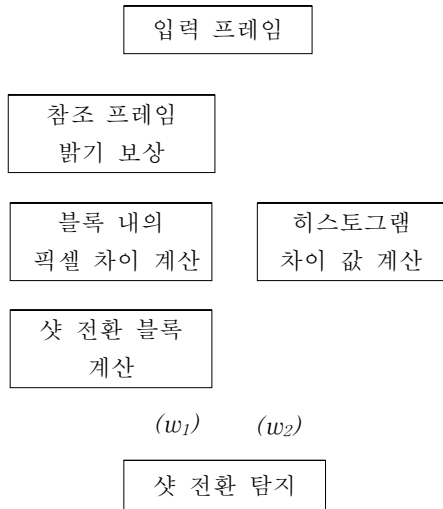
본 논문은 프레임의 지역적 화소 정보를 이용하여 프레임간의 밝기 변화를 보정한 후, 보정된 프레임의 지역적 화소 정보와 각 프레임의 히스토그램의 정보를 이용하는 샷 전환 탐지 알고리즘을 제안한다. 2절에서는 샷 전환 탐지를 위해 제안한 알고리즘을 설명하고, 3절에서는 제안한 알고리즘을 국가기록원 소장 디지털화 영상에 적용한 결과를 기술한다. 4절에서는 제안한 방법의 문제점과 개선 방향을 제시한다.

2. 샷 전환 탐지 알고리즘

본 논문의 샷 전환 탐지 알고리즘은 비디오 데이터 $V = (F_0, F_1, F_2, \dots, F_{N-1})$ 에 대해 샷 전환 벡터 $D = (d_0, d_1, d_2, \dots, d_{N-1})$ 를 출력한다. V 는 N 개의 프레임으로

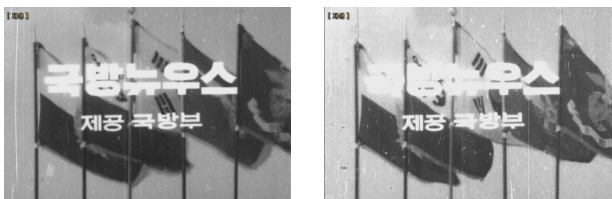
이 논문은 행정안전부 국가기록원 재원으로 2012년 기록보존기술 연구개발사업의 지원을 받아 수행된 연구임

구성된 비디오 데이터, F_k 는 크기가 $X \times Y$ 인 2차원 행렬, 그리고 $F_k(x,y)$ 는 k -번째 프레임의 좌표 (x,y) 의 화소 값이다. 제안한 샷 전환 탐지 방법에서는 주어진 시간 t 로부터 두 프레임 $F_t(x,y)$ 와 $F_{t+1}(x,y)$ 의 동일 좌표의 화소 값의 차이의 평균값을 이용해 프레임 간 밝기 변화를 보상 한다. 밝기가 보정된 프레임 F_k 는 일정 크기($X_b \times Y_b$)의 블록 $b_i^0, b_i^1, b_i^2, \dots, b_i^{B-1}$ 으로 분할시키고, 두 프레임 간 대응 블록 내의 화소의 차이와 프레임 간의 밝기 값 히스토그램의 차이를 구한다. 마지막으로 블록 내의 화소의 차이 값과 히스토그램의 차이 값에 가중치를 주어 샷 전환 여부를 결정한다. 이 순서도를 (그림 1)에서 보여주고 있다.



(그림 1) 제안하는 알고리즘의 순서도

프레임 간 화소 차를 이용하는 샷 전환 탐지 방법에서 지역 공간 위치에 따른 화소 변화의 차가 비슷한 경우 샷 전환 예측이 어렵다[8]. 이러한 단점을 보완하기 위해 전체 프레임을 서브 블록으로 분할하고 연속된 프레임의 동일한 위치의 블록의 화소 차를 이용한다. 또한 같은 장면임에도 불구하고 (그림 2)과 같이 연속한 프레임의 급격한 밝기 차이로 인해 오탐지가 발생 할 수 있다. 이러한 문제를 해결하기 위해서는 연속한 두 프레임의 밝기 차이를 구하여 보상해 주는 과정이 필요하다. 주어진 비디오 $V=(F_0, F_1, F_2, \dots, F_{N-1})$ 의 샷 전환 탐지 알고리즘은 다음과 같다.



(그림 2) 밝기 차이로 인한 검출 오류

프레임은 F , 시간에 따른 기준은 t , 프레임의 크기를 XY , 프레임의 한 픽셀의 좌표를 (x,y) 라 할 때, F_t 와

F_{t+1} 사이의 샷 전환 여부는 다음 절차에 따라 수행된다.

(단계 1) 프레임 간의 밝기 보상 값인 F_t 와 F_{t+1} 의 밝기 차이의 평균 $I(t,t+1)$ 계산

$$I(t,t+1) = \frac{1}{XY} \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} (F_t(x,y) - F_{t+1}(x,y))$$

(단계 2) 벡터 $M=(M_0, M_1, M_2, \dots, M_{B-1})$ 이 두 프레임 간 블록에서 샷 전환일 가능성이 있는 픽셀의 수라고 할 때, 각 M_i 를 계산

$$M_i = \sum_{x=0}^{X_b} \sum_{y=0}^{Y_b} 1(|b_i^i(x,y) - b_{i+1}^i(x,y) - I(t,t+1)| - \theta_1)$$

$, i=0, 1, 2, \dots, B-1$ (B 는 블록의 개수)

, 여기서 X_b 와 Y_b 는 블록의 크기, b^i 는 i 번째 블록, θ_1 은 샷 전환일 가능성이 있는 픽셀을 탐지하기 위한 임계값이고, $a > 0$ 이면 $1(a) = 1$, 그렇지 않으면 $1(a) = 0$ 이다.

(단계 3) 샷 전환일 가능성을 나타내는 블록 수 $n(t,t+1)$ 를 계산

$$n(t,t+1) = \sum_{i=0}^{B-1} 1(M_i - \theta_2)$$

, 여기서 θ_2 은 샷 전환일 가능성이 있는 블록 M_i 를 탐지하기 위한 임계값이다.

(단계 4) t 로부터 연속하는 L 개 프레임까지의 히스토그램 차이 $h(t,t+L)$ 계산

$$h(t,t+L) = \frac{1}{LC} \sum_{l=1}^L \sum_{c=0}^{C-1} 1(H_l(c) - H_{l+1}(c) - \theta_3)$$

, 여기서 L 은 계산에 필요한 프레임의 개수, C 는 히스토그램의 색상 수, θ_3 은 두 프레임 사이의 샷 전환 가능성을 탐지하기 위한 임계값이다.

(단계 5) 샷 전환을 나타내는 d_{t+1} 을 계산

$$d_{t+1} = \begin{cases} 1 & \text{if } z(t,t+1) > \theta_4 \\ 0 & \text{otherwise} \end{cases}$$

, 여기서 $z(t,t+1) = w_1 \times n(t,t+1) + w_2 \times h(t,t+1)$ 이고, $w_1 + w_2 = 1$ 이다. θ_4 는 샷 전환 프레임을 탐지하기 위한 임계값이다.

(단계 1)에서 밝기 보상 값의 계산은 카메라 플래쉬 같은 번쩍거림에 따른 급격한 변화를 극복할 수 있다. (단계 2)로부터 블록 단위로 F_t 와 F_{t+1} 의 픽셀의 차이를 구한 후 (단계 1)에서 구한 밝기 보상 값을 감산한다. 만약 감산한 각 픽셀의 차이 값이 임계값 θ_1 보다 크면 샷 전환을 나타낼 가능성이 있다. 블록 당 샷 전환을 나타내는 픽셀들의 개수를 이용해 샷 전환을 나타낼 가능성이 있는 블

록 M_t 를 계산한다. (단계 3)에서 각 M_t 가 임계값 θ_2 를 만족하는지를 검사하여 F_t 와 F_{t+1} 사이에 샷 전환을 나타낼 가능성이 있는 블록의 수 $n(t, t+1)$ 을 구한다. (단계 4)에서는 기준 프레임 t 의 밝기 정보 히스토그램과 $t+1$ 부터 L 까지의 프레임의 밝기 정보 히스토그램의 차이를 계산한다. 가능한 색상 수가 C 이면 $H_t(i)$ 은 프레임 F_t 의 색상 i 의 빈도수이다($0 \leq i < C$). 구해진 값이 히스토그램 차이의 임계값을 나타내는 θ_3 보다 큰 경우 두 프레임 사이에 샷 전환의 가능성이 있는 것으로 판단한다. $h(t, t+1)$ 은 프레임간 히스토그램의 불일치 정도를 나타낸다. 마지막으로 (단계 5)에서 $n(t, t+1)$ 과 $h(t, t+1)$ 에 가중치 w_1 과 w_2 를 고려하여 더한 값이 서로 다른 프레임 정도를 나타내는 임계값 θ_4 를 만족할 경우 새로운 샷의 시작을 의미하므로 $d_{t+1} = 1$ 로 설정하고, 그렇지 않은 경우 $d_{t+1} = 0$ 으로 한다.

각 단계에서 사용한 임계값과 가중치의 설정 방법은 다음과 같다.

- θ_1 대표 영상에서 샷 전환이 발생하는 프레임들의 밝기 차이의 평균
- θ_2 대표 영상에서 샷 전환이 발생하는 프레임들의 블록에서 차이가 있는 픽셀 수의 평균
- θ_3 대표 영상에서 샷 전환이 발생하는 프레임들의 히스토그램 차이의 평균
- θ_4 대표 영상에서 샷 전환이 발생하는 프레임들의 비율의 평균
- w_1, w_2 대표 영상에서 샷 전환이 발생하는 프레임들의 픽셀-블록 기반과 히스토그램 기반의 가중치

3. 실험 결과

본 논문에서 제안한 알고리즘의 성능을 평가하기 위하여 국가기록원이 소장하고 있는 대한뉴스를 사용하였다. 비디오 프레임의 수는 1,500이며, 6종의 비디오를 사용하였다. 영상의 특징으로 샷 전환 탐색의 오류에 크게 영향을 끼치는 흠집, 얼룩, 떨림, 찢김, 번쩍거림 등의 훼손이 나타났다. 사용된 비디오는 기록원 영상 처리 담당자와 대학원생이 프레임-프레임을 분석하여 샷 전환 여부를 표시

하였다. 최종 선정된 샷 전환 탐지 수는 74개이다. 테스트 평가는 정확률(precision)과 재현율(recall), F-measure를 이용한다.

테스트 비디오에 전역적 또는 지역적 화소 정보만 사용한 실험에서는 훼손 정도가 약하고 객체 또는 카메라 이동이 적은 경우 좋은 결과를 보였다. 그러나 훼손이 심하거나 객체 또는 카메라의 이동이 큰 경우 화소의 차가 크기 때문에 정확율과 재현율이 반비례하였다. 또한 히스토그램 정보만 적용했을 때, 객체 또는 카메라가 크게 이동하는 경우는 화소(블록)에 비해 잘 대처할 수 있었지만 밝기에 매우 민감하여 번색 또는 번쩍거림 등의 잡음에 크게 반응하였고, 샷 전환이 일어나도 밝기가 비슷한 경우 검출률이 떨어졌다<표 1>. 제안한 알고리즘에서는 화소 기반과 히스토그램 기반의 두 알고리즘 함께 사용하여 서로의 단점을 보완했다. 각각의 알고리즘을 따로 사용했을 때 보다 D가 증가하고 N과 M이 눈에 띄게 줄어들었다. 정확한 샷 검출의 예를 (그림 3)에 보였다. 정확율의 경우 화소(블록)만 사용한 것보다 12%, 히스토그램만 사용한 것보다 26% 증가했다. 재현율의 경우 화소(블록)만 사용한 것보다 16%, 히스토그램만 사용한 것보다 7% 증가했다. 실험에 사용한 임계값과 가중치는 $\theta_1 = 30, \theta_2 = 65, \theta_3 = 50, \theta_4 = 70, w_1 = .6$ 이다.



(그림 3) 제안 알고리즘을 적용하여 올바른 샷 전환을 검출한 예

<표1> 선택된 6종 비디오에 대한 실험 결과

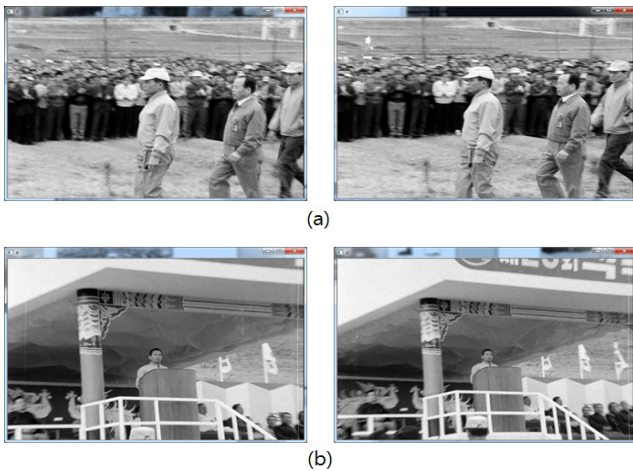
구분	T	화소						히스토그램						제안 알고리즘					
		D	N	M	pre	rec	f-m	D	N	M	pre	rec	f-m	D	N	M	pre	rec	f-m
V1	10	10	0	0	1.00	1.00	1.00	10	17	0	0.37	1.00	0.54	10	3	0	0.77	1.00	0.87
V2	12	5	0	7	1.00	0.42	0.59	11	1	1	0.92	0.92	0.92	12	1	0	0.92	1.00	0.96
V3	11	7	0	4	1.00	0.64	0.78	10	31	1	0.24	0.91	0.38	10	3	1	0.77	0.91	0.83
V4	12	11	58	1	0.16	0.92	0.27	11	0	1	1.00	0.92	0.96	12	2	0	0.86	1.00	0.92
V5	18	16	3	2	0.84	0.89	0.86	17	10	1	0.63	0.94	0.76	18	2	0	0.90	1.00	0.95
V6	11	11	10	0	0.52	1.00	0.69	8	8	3	0.50	0.73	0.60	10	0	1	0.91	0.91	0.95
평균					0.75	0.81	0.70				0.61	0.90	0.70				0.87	0.97	0.91

· T는 총 샷 전환의 수, D는 알고리즘을 통해 검출한 샷 전환의 수, N은 잘못 검출한 샷 전환의 수, M은 검출하지 못한 샷 전환의 수

4. 결론

본 논문에서는 화소와 히스토그램 정보를 이용하여 샷 전환을 탐지하는 알고리즘을 제안하였다. 제안한 방법은 프레임 간 전체적인 화소 차를 이용해 밝기를 보상하는 방법, 서브 블록을 이용한 지역적 화소 값의 차, 프레임의 밝기 값 히스토그램의 차이를 이용했다. 실험에서 두 알고리즘을 함께 사용한 경우가 한 가지 알고리즘만 사용한 경우보다 검출률과 정확도가 향상되었으나 개선의 여지가 필요하다. 알고리즘을 병합하여 사용함으로써 각 알고리즘에서 사용하는 임계값들을 모두 포함하게 되었다. 따라서 임계값의 설정에 민감하게 되었으며, 또 (그림 4)와 같이 객체 또는 카메라의 이동 폭이 매우 큰 경우는 샷 변화의 차이가 크기 때문에 잘못 검출하는 경우가 발생했다. 이를 해결하기 위해서는 적응적 임계값의 사용과 급격한 움직임에 대한 보상이 필요하다.

[7] Robert A. Joyce, "Temporal Segmentation of Video using Frame and Histogram-Space," IEEE Transactions on Multimedia, 8(1), 2006
 [8] In-Shik Jeong, Oh-Jin Kwon, "Video shot boundary detection using relative difference between frames", Optical Engineering, 42(3), pp.604-605, 2003



(그림 4) 객체 및 카메라의 급격한 이동으로 인한 오검출의 예 (a) 객체와 카메라의 급격한 이동 (b) 카메라의 급격한 줌-아웃

참고문헌

[1] Jinhui Yuan, Huiyi Wang, Lan Xian, Wujie Zheng, Jianmin Li, Fuzong Li, and Bo Zhang, "A Formal Study of Shot Boundary Detection," IEEE Transactions on circuits and systems for video technology, Vol. 17, No. 2, 2007
 [2] H. J. Jhang, A. Kankanhalli, S. W. Smoliar, "Automatic Partitioning of Full-Motion Video," Multimedia, Vol. 1, No. 1, pp. 10-28, 1993.
 [3] Xiaoquan Yi and Nam Ling, "Fast Pixel-Based Video Scene Change Detection," in IEEE Int. Symposium on Circuits and Systems, pp. 3443-3446, 2005.
 [4] B. H. Shekar, K. Holla K. Raghurama, Kumari M. Sharmila, "Video Cut Detection Using Chromaticity Histogram," International Journal of Machine Intelligence, Vol. 3, No. 4, pp. 371-375, 2011
 [5] Priyadarshinee Adhikari, Neeta Gargote, Jyothi Digge, and B.G. Hogade, "Abrupt Scene Shange Detection," in World Academy of Science, Engineering and Technology 42 2008.
 [6] R. Zabih, J. Miller, and K. Mai, "A feature-based algorithm for detecting and classification production effects," ACM Multimedia Syst, vol. 7, no. 1, pp. 119 - 128, 1999.