

Otsu 임계값 설정과 프레임 블록화를 이용한 샷 전환 탐지

김승현*, 황두성**

*단국대학교 전자계산학과

**단국대학교 컴퓨터학과

e-mail: pcpp13@naver.com

A Shot Change Detection Algorithm using Otsu Threshold and Frame Segmentation

Seung-Hyun Kim*, Doosung Hwang**

*Dept of Electronic Calculation Science, Dan-Kook University

**Dept of Computer Science, Dan-Kook University

요 약

본 논문에서는 프레임 블록화와 Otsu 임계값 설정 방법을 이용한 샷 전환 탐지 알고리즘을 제안한다. 제안 방법은 연속된 두 프레임을 일정 크기의 영역으로 분할하여 두 프레임 간 대응되는 영역의 히스토그램 차이를 이용해 샷 전환을 탐지한다. 또한 각 영상마다 Otsu 임계값 설정 방법을 이용하여 자동으로 임계값을 설정한다. 제안 방법의 실험은 영화, 드라마, 애니메이션 등 다양한 영상에 대해 테스트 되었으며, 기 연구된 샷 전환 탐지 알고리즘과 비교 시 우수한 탐지율을 보였다.

1. 서론

최근 영상 장비의 첨단화로 인해 영상 데이터의 생산량이 증가하고 있는 추세이다. 효율적인 영상 데이터의 관리를 위해서는 컴퓨팅 자원의 활용이 필요하다. 비디오는 영상과 음향으로 구성되어 있으며, 영상은 여러 개의 프레임으로 이루어져 있다. 하나의 카메라에서 촬영된 연속적인 프레임으로 이루어진 영상을 샷(shot)이라고 하며, 샷과 연결된 다음 샷 사이의 경계부분을 컷(cut) 또는 샷 경계(shot boundary)라고 한다. 컷 또는 샷 경계를 탐지하는 샷 전환 탐지(shot change detection)는 영상 데이터의 검색(search), 분할(segmentation), 색인(index)의 기반 기술이다.

샷 전환 탐지를 위해서는 각 프레임의 특징을 추출하여 비교해야 한다. 기 연구된 샷 전환 탐지 방법에서는 화소값의 차이[1,2] 또는 화소값의 분포를 나타내는 히스토그램의 차이[1,3]가 많이 사용되었으며, 로그정규분포[4], 프레임 내 수직·수평의 화소값을 이용한 방법[5]도 연구되었다.

본 논문에서 제안하는 샷 전환 탐지 방법은 Otsu 임계값 설정 방법[6,7]을 이용하여 임계값을 자동으로 설정하며, 프레임을 분할하여 분할 영역 내 히스토그램의 차이를 비교한다. 각 영상별로 알맞은 값을 추출하여 임계값을 자동으로 설정하게 되면, 다양한 종류의 영상에 대해 균일한 샷 전환 탐지 정확도를 얻을 수 있다는 장점이 있다.

2. 관련 연구

샷 전환 탐지 알고리즘은 N 개의 프레임으로 구성된 비

디오 $V = \{F_1, F_2, F_3, \dots, F_{N-1}, F_N\}$ 에 대해 샷 전환 탐지 결과를 나타내는 $D = \{d_1, d_2, d_3, \dots, d_{N-1}\}$ 을 출력한다. d_k 는 연속 프레임 F_k 와 F_{k+1} 사이의 샷 전환 탐지 결과를 나타내며 탐지 알고리즘을 통해 결정된다.

Zhang 등[1]과 Xiaoquan Yi 등[2]은 연속된 두 프레임 간 대응되는 좌표의 화소값 차이를 이용해 샷 전환을 탐지하는 방법을 제안하였다. $F_k(x, y)$ 와 $F_{k+1}(x, y)$ 의 화소값 차이를 임계값 ϵ_1 과 비교, ϵ_1 보다 큰 차이를 보이는 화소의 비율이 임계값 ϵ_2 이상일 경우 F_k 와 F_{k+1} 사이에서 샷 전환이 발생한 것으로 판단한다.

화소값 차이와 더불어 히스토그램의 차이 또한 샷 전환 탐지에 많이 사용되는 방법이다. 두 프레임 간의 GRAY 히스토그램 차이를 이용한 방법[1]과 RGB, HSV, YCrCb 등 다양한 색 공간의 히스토그램 차이를 이용한 방법[3]이 연구되었다. 프레임 F_k 와 F_{k+1} 의 히스토그램 정보를 추출한 후, 각 히스토그램 빈(bin)값의 차이를 계산한다. 히스토그램 빈값의 차이가 임계값 ϵ_1 보다 큰 비율이 임계값 ϵ_2 이상일 경우 F_k 와 F_{k+1} 사이에서 샷 전환이 발생한 것으로 판단한다.

프레임을 일정 크기로 분할한 후, 분할된 영역 내 특징을 비교하여 샷 전환을 탐지하는 방법도 연구되었다[8]. 프레임을 분할하여 샷 전환을 탐지할 경우 밝기의 급격한 변화나 영상 내 객체의 이동으로 인한 오 탐지를 줄일 수 있었다.

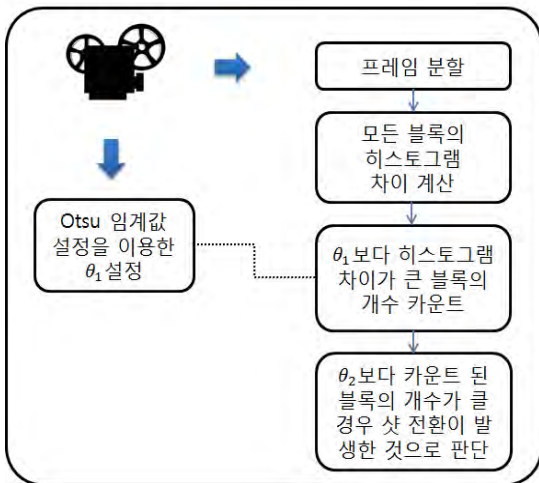
Lee 등[8]은 프레임을 분할하여 두 프레임 간 분할영역

을 비교할 때, 비교 대상인 $k+1$ 번째 프레임의 분할영역을 넓게 확장시켜 k 번째 프레임의 분할영역이 포함되도록 하였다. 객체가 급격히 이동을 하였다더라도 확장된 영역을 벗어나지 않았더라면 샷 전환이 발생한 것으로 판단하지 않기 때문에 오 탐지율을 줄일 수 있었다.

언급한 연구는 모두 임계 값을 사전실험을 통해 추출한 값으로 고정하여 사용한다. 따라서 샷 전환 탐지율은 임계 값에 많은 영향을 받는다. 본 논문에서 제안하는 방법은 Otsu 임계값 설정 방법을 이용하여 임계값 설정의 문제를 자동화 하였으며, 각 영상에 알맞은 임계값을 추출하여 설정한다.

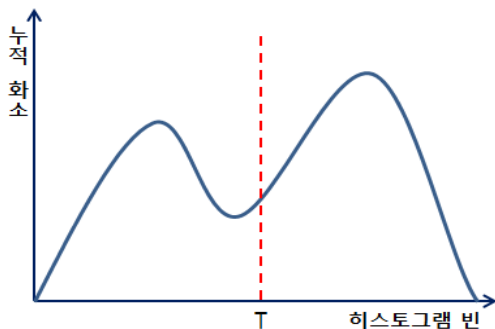
3. 샷 전환 탐지 알고리즘

본 논문에서는 프레임 블록화와 Otsu 임계값 설정방법을 이용한 샷 전환 탐지 방법을 제안하며, 각 영상에 알맞은 임계값을 추출하여 샷 전환 탐지의 정확도를 높이는 데 목적이 있다. 그림 1은 샷 전환 탐지 알고리즘의 개략적인 순서도이다.



(그림 1) 샷 전환 탐지 순서도

Otsu 임계값 설정 방법은 분산을 이용해 히스토그램을 두 개의 클래스로 나눌 수 있는 적절한 임계값을 추출하는 방법이다. 그림 2는 프레임 내 화소값을 기반으로 한



(그림 2) 쌍봉형 히스토그램의 Otsu 임계값 T

히스토그램에서 추출된 Otsu 임계 값을 나타낸다. 쌍봉형의 형태를 보이는 히스토그램에서 계곡점 주변이 Otsu 임계값으로 설정된다.

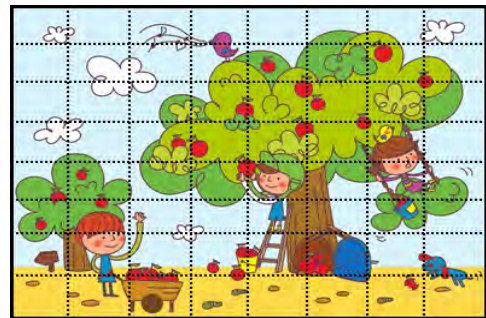
블록의 샷 전환을 판단하는 임계값 θ_1 은 Otsu 임계값 설정 방법[6,7]을 이용해 자동으로 설정한다. Otsu 임계값의 계산에 사용되는 클래스 내 분산 $\sigma_w^2(t)$ 은 식 (1)을 이용해 계산한다.

$$\sigma_w^2(t) = w_1(t)\sigma_1^2(t) + w_2(t)\sigma_2^2(t) \quad (1)$$

가중치 w_1 과 w_2 는 해당 화소가 프레임 내 존재할 확률이며, σ_1 과 σ_2 는 두 클래스의 분산 값이다. t 는 히스토그램의 각 빈 값이며, 모든 t 에 대해 $\sigma_w^2(t)$ 값을 계산한 뒤, $\sigma_w^2(t)$ 값이 가장 작을 때의 t 를 Otsu 임계값으로 설정한다.

임계값 θ_1 을 설정하기 위해 영상을 구성하는 모든 프레임의 Otsu 임계값을 계산하였다. 연속 프레임 F_k 와 F_{k+1} 의 Otsu 임계값의 차이가 3보다 클 경우, Otsu 임계값을 샷 전환 후보군으로 분류하였으며, 후보군의 평균을 계산하여 임계값 θ_1 으로 설정하였다.

샷 전환 탐지 과정에서 밝기의 급격한 변화나 객체의 빠른 이동은 오 탐지를 유발하는 가장 큰 요인이다. 이러한 오 탐지율을 줄이기 위해 본 논문에서는 프레임의 블록화를 이용한 샷 전환 탐지를 제안한다. 그림 3은 프레임을 일정크기의 블록으로 분할하는 예시를 나타낸다.



(그림 3) 프레임의 분할

본 논문에서는 사전실험을 통해 그림 3과 같이 프레임을 8×8 개의 블록으로 분할하였다. k 번째 프레임과 $k+1$ 번째 프레임을 비교하기 위해 각블록 내 화소값을 이용해 히스토그램을 계산한다. 대응되는 두 블록 간 히스토그램 차이가 Otsu 임계값 설정 방법을 이용해 설정한 θ_1 보다 클 경우 샷 전환이 발생한 블록으로 판단한다. 모든 블록에 대해 샷 전환 탐지가 완료된 후, 샷 전환이 발생한 것으로 판단된 블록이 θ_2 보다 많다면, 두 프레임 사이에서 샷 전환이 발생한 것으로 결정한다. θ_2 는 사전실험을 통해 전체 블록의 약 77%인 49로 설정하였다.

4. 실험

제안하는 알고리즘의 성능을 평가하기 위해 영화, 드라마, 애니메이션 영상을 실험영상으로 사용하였다. 실험 영상은 모두 5,000프레임으로 구성되어 있으며, 실험영상의 타입 및 실제 샷 전환의 수는 표 1과 같다.

<표 1> 실험영상 상세정보

타입	비디오	프레임	샷 전환 수
영화	V1	5,000	51
	V2	5,000	29
	V3	5,000	50
드라마	V4	5,000	34
	V5	5,000	55
	V6	5,000	67
애니	V7	5,000	46
	V8	5,000	47
	V9	5,000	40

제안하는 샷 전환 탐지 알고리즘과 기존 알고리즘의 탐지율을 비교하기 위해 프레임을 분할하지 않고 픽셀과 히스토그램 차이를 이용한 실험도 진행하였다. 실험영상에 대한 샷 전환 탐지 알고리즘의 결과는 표 2와 같다.

표 2의 Cut은 실험영상의 실제 컷 수, D 는 알고리즘이 탐지에 성공한 샷 전환의 수, M 은 알고리즘이 탐지하지 못한 샷 전환의 수, N 은 알고리즘이 오 탐지한 샷 전환의 수를 의미한다. 프레임을 분할하지 않았을 경우 화소 값의 차이를 이용한 알고리즘은 영상의 밝기 변화 또는 객체의 급격한 움직임에 민감하게 반응하여 상대적으로 오 탐지가 많이 발생하였다. 히스토그램의 차이를 이용한 알고리즘은 화소 값의 차이를 이용한 방법에 비해 오 탐지는 줄어들었지만 프레임 내 화소 값의 분포를 이용하기 때문에 비슷한 분포를 보이는 샷 전환은 탐지하지 못하였다. 제안하는 샷 전환 탐지 알고리즘은 프레임을 8x8개의 블록으

로 분할하여 블록별 히스토그램의 차이를 비교하기 때문에 화소 값 차이를 이용한 알고리즘에 비해 영상의 잡음에 덜 민감하며, 프레임 간 히스토그램이 비슷한 샷 전환도 탐지가 가능했다.

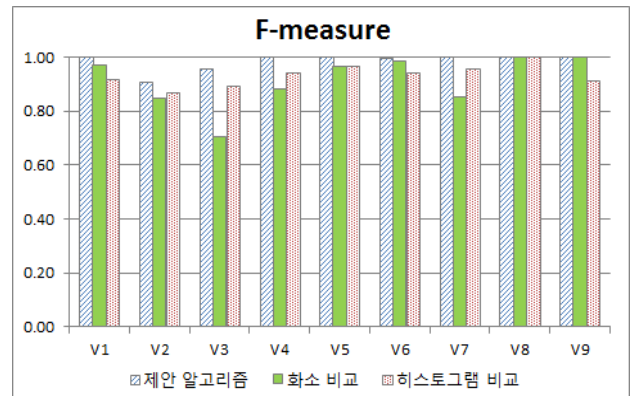
실험 결과의 일반화 성능 평가는 재현율(recall), 정밀도(precision), F -measure로 하였으며, 식 (2), (3), (4)에 제시한다.

$$rec = \frac{D}{D + M} \tag{2}$$

$$pre = \frac{D}{D + N} \tag{3}$$

$$F\text{-measure} = 2 \times \frac{pre \times rec}{pre + rec} \tag{4}$$

식 (2)에서 rec 값이 1이면 모든 샷 전환을 탐지한 것이며, 식 (3)에서 pre 값이 1이면 오 탐지된 샷 전환이 없다는 것을 의미한다. 식 (4)의 F -measure는 값이 1에 가까울수록 샷 전환 탐지 알고리즘의 정확도와 예측율이 높다는 것을 의미한다. 그림 4는 샷 전환 탐지 알고리즘의 F -measure 평가 결과이다.



(그림 4) 실험 영상의 F -measure 결과

<표 2> 실험영상의 샷 전환 탐지 결과

타입	비디오	Cut	제안 알고리즘			화소 비교			히스토그램 비교		
			D	M	N	D	M	N	D	M	N
영화	V1	51	51	0	0	51	0	3	43	8	0
	V2	29	28	1	5	28	1	9	22	7	0
	V3	50	50	0	5	49	1	40	49	1	11
드라마	V4	34	34	0	0	34	0	9	31	3	1
	V5	55	55	0	0	54	1	3	51	4	0
	V6	67	67	0	1	67	0	2	59	8	0
애니	V7	46	46	0	0	46	0	16	42	4	0
	V8	47	47	0	0	47	0	0	47	0	0
	V9	40	40	0	0	40	0	0	36	4	3

5. 결론

본 논문에서는 프레임 블록화와 Otsu 임계값 설정 방법을 이용한 샷 전환 탐지 알고리즘을 제안하였다. 기존의 샷 전환 탐지 알고리즘은 영상 내 밝기 변화, 객체의 급격한 이동 등 잡음으로 인한 오 탐지가 많았으며, 샷 전환을 판단하는 임계값이 고정되어 있어 다양한 타입의 영상에 대해 좋은 결과를 얻을 수 없다는 문제가 있었다.

제안하는 샷 전환 탐지 알고리즘은 프레임을 일정 크기의 블록으로 분할하여 잡음으로 인한 오 탐지율을 줄이는데 목적이 있다. 또한 영상의 Otsu 임계값을 이용해 영상에 적합한 임계값을 자동으로 추출한다. 추출된 임계값을 사용할 경우 다양한 타입의 영상에 대해 모두 비슷한 정확도를 보여 주었으며, 실험 결과를 통해 기존의 알고리즘보다 우수한 결과를 보였음을 증명하였다.

제안 알고리즘의 경우 기존 샷 전환 탐지 알고리즘에 비해 탐지율은 높아졌으나 여러 프레임에 걸쳐 진행되는 점진적인 샷 전환은 탐지하지 못하는 경우가 많았다. 점진적인 샷 전환의 탐지를 위해서는 2 프레임 이상의 프레임에 이용한 탐지 알고리즘이 필요할 것으로 보인다.

참고문헌

- [1] H. J. Zhang, A. Kankanhalli, S. W. Smoliar, "Automatic partitioning of full-motion video," *Multimedia Systems*, Vol. 1, No. 1, pp. 10-28, 1993.
- [2] Xiaquan Yi, Nam Ling, "Fast Pixel-Based Video Scene Change Detection," *IEEE Int. Symposium on circuits and Systems*, Vol. 4, pp. 3343-3346, May 2005.
- [3] S. Y. Tak, S. Yoo, B. R. Lee, W. J. Lee, K. S. Ryu, T. G. Kim, H. C. Kang, "Scene change detection of various color space using difference of histogram," *Processings of the Spring Conf. on The Korea Contents Association*, pp. 466-468, May 2010.
- [4] S. G. Hong, B. K. Cho, Y. S. Choe, "Adaptive Thresholding for Scene Change Detection," *IEEE Third Int. Conf. on Consumer Electronics-Berlin*, pp. 75-78, Sept. 2013.
- [5] J. K. Jin, J. H. Cho, J. H. Jeong, D. S. Jeong, "Fast Scene Change Detection Using Macro Block Information and Spatio-temporal Histogram," *Journal of The Institute of Electronics Engineers of Korea*, Vol. 48-SP, No. 1, pp. 141-148, Jan. 2011.
- [6] J. R. Parker, "Algorithms for Image Processing and computer vision" 2nd Ed. wiley publishing, pp. 149-151, 2011.
- [7] Puneet, Naresh Kumar Garg, "Binarization Techniques used for Grey Scale Images," *International Journal of Computer Applications*, Vol. 71, No. 1, pp. 8-11, June 2013.
- [8] J. G. Lee, K. S. Han, B. M. Yoo, D. S. Hwang,

"Shot Boundary Detection Algorithm by compensating Pixel Brightness and Object Movement," *Journal of The Korea Society of Computer and Information*, Vol. 18, No. 5, pp. 35-42, May 2013.