

## 에지 이미지를 사용한 컷 검출의 구현

김설호<sup>○</sup>, 최형일<sup>\*\*</sup>, 김계영<sup>\*</sup>

<sup>○\*</sup>송실대학교 대학원 컴퓨터학과

<sup>\*\*</sup>송실대학교 대학원 미디어학과

e-mail: sulho0930@paran.com, {hic, gykim11}@ssu.ac.kr

## An Implementation of Cut Detection using Edge Image

Sul-Ho Kim<sup>○</sup>, Hyung-Il Choi<sup>\*\*</sup>, Gye-Young Kim<sup>\*</sup>

<sup>○\*</sup>Dept. of Computing, Soong-Sil University

<sup>\*\*</sup>Dept. of Media, Soong-Sil University

### ● 요약 ●

최근에는 텍스트정보 보다 동영상정보를 다루는 일이 많아졌고 그에 따라 동영상 데이터의 분할, 색인, 검색등을 위해 장면 전환 검출이 필요하게 되었다. 장면 전환 검출 기술은 비디오 데이터의 장면 변화가 발생하는 경계를 검출하는 기술이다. 본 논문에서는 에지 이미지를 이용한 장면전환 검출과 이를 위한 임계값 설정, 그리고 결과에서 중복된 이미지와 오검출된 이미지를 줄여줄 수 있는 구현에 대하여 실험결과를 바탕으로 설명한다.

**키워드:** 장면전환(Scene change), 컷 검출(cut detection), 샷 경계(shot boundary)

### I. 서론

장면 전환 검출 기술은 비디오 데이터의 장면 변화가 발생하는 경계를 검출하는 기술이다. 장면 전환 검출 방법은 비디오 데이터를 효과적으로 다루기 위하여 데이터의 분할, 색인, 검색, 요약, 브라우징등에서 많이 사용된다. 지금까지 장면전환 검출에 대한 많은 연구가 진행되어 왔고 진행 중이다[1]. 장면 전환은 크게 2가지로 나눌 수 있다. 첫 번째는 급진적인 장면전환이고 두 번째는 점진적인 장면전환이다[4]. 급진적인 장면전환이란 이전프레임과 현재 프레임 사이에 아무런 영상효과 없이 영상이 급진적으로 변하는 것을 말한다(그림1). 장면 전환의 2가지 대분류중 하나인 급진적인 장면전환에 대해서는 여러 가지 방법들 모두 뛰어난 검출 결과를 보여준다. 현대의 영상에는 디졸브, 오버랩, 페이드인-아웃, 와이프, 아이리스등 수많은 영상효과들이 적용되고 있다. 페이드인 페이드아웃은 영상이 점점 검게 변하거나 혹은 검은 영상에서 시작하여 점점 영상의 형태를 갖추어 나가는 효과를 말하며 디졸브는 서로 다른 두 영상이 겹쳐 이전 영상의 밝기는 점점 어두워지면서 이와 동시에 다음 영상의 밝기가 증가하면서 장면이 전환되는 효과(그림 1)를 말하며 겹치기(overlapping)와는 다르다[2]. 이러한 효과들이 적용된 점진적인 장면 전환 영상에 대한 장면 전환 검출 결과는 뛰어나지 않다고 알려져 있다. 실험에 사용된 영상은 디졸브 효과가 많이 적용되어 있는데 일반적으로 디졸브 효과에는 약 4~10프레임 정도가 이용되며 단, 1프레임으로 이루어진 경우, 10프레임 이상의 프레임을 사용하는 경우도 종종 있다.

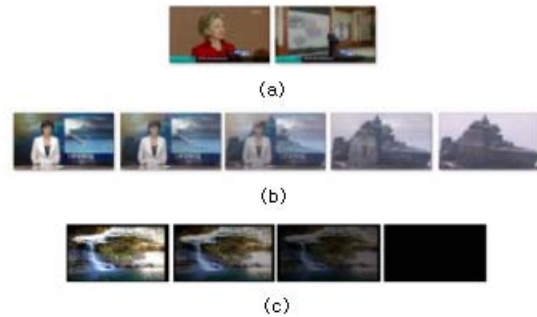


그림 1. 장면 전환의 종류  
(a) 급진적인 장면전환, (b) 디졸브, (c) 페이드아웃

### II. 본론

장면전환의 검출을 위해서는 첫 번째로 프레임의 특징을 추출할 방법이 있어야 한다. 두 번째로 추출된 특성에 따라서 장면의 변화 여부를 판단할 수 있는 임계값이 필요하다. 장면전환 검출 방법은 화소 값, 히스토그램, 움직임, DC계수기반의 방법등으로 나누어 볼 수 있다. 각각의 방법들은 프레임의 특징을 추출할 수 있으며 추출된 특성으로 장면전환 여부를 판단할 수 있는 임계값을 통해 장면전환을 검출하게 된다. Nagasaka와 Tanaka는 각 프레임의 전체 화소값의 합의 절대값 차이(simple interframe difference)로 프레임 특징을 추출하였고 아래의 식으로 표현된다[3].

$$\left( \left| \sum_{i=1}^X \sum_{j=1}^Y P(I_t, i, j) - \sum_{i=1}^X \sum_{j=1}^Y P(I_{t-1}, i, j) \right| \right) > T \quad (1)$$

X, Y는 프레임의 가로 세로사이즈이며, P(I, i, j)는 I번째 프레임에서 좌표(i, j)의 밝기 값을 나타내며 이 때,  $1 \leq i \leq X, 1 \leq j \leq Y$ 이다. 화소값을 이용한 방법은 객체의 움직임에 민감하다.

히스토그램 기반의 방법은 영상 픽셀정보의 분포도를 나타낸 히스토그램을 사용하여 영상의 차이를 비교하는 방법이다. 이 방법은 화소값을 이용한 방법의 단점을 어느정도 보완해 줄 수 있다. Tonomura와 Abe는 화소 밝기값 히스토그램에 기반한 방법을 제안하였고 아래의 식에 표현된다[5]. 이 방법외에도 HSV등 다른 공간에서의 히스토그램을 적용한 방법, 블록단위를 이용한 방법등도 있다.

$$\left( \sum_{v=0}^V |H(I_t, v) - H(I_{t-1}, v)| \right) > T \quad (2)$$

DC계수 기반의 방법은 대부분의 디지털 비디오가 8x8블록 단위의 DCT코딩되어 있음에 기반하는 방법으로 각 블록은 하나의 DC계수와 63개의 AC계수로 형성 되는데 DC성분은 블록의 평균 픽셀값의 8배와 같으므로 각각의 프레임에서 DC성분을 취하면 전체 데이터의 64배를 줄일 수 있기 때문에 빠른 구현에 유리하다고 한다[6][7].

장면과 장면사이에서는 색상, 히스토그램과 함께 에지의 특성도 급격하게 변화하게 된다. 이 변화량은 급진적인 장면전환에서 두드러지지만 점진적인 장면전환에서도 여전히 유효하다. 본 논문에서는 에지를 이용한 장면전환 검출 방법을 이용한다. 에지를 추출해내는 알고리즘은 많은 종류가 있지만 여기서는 John F. Canny가 제안한 Canny 에지 알고리즘[8] 사용하였다. 이 알고리즘을 사용하여 의미 없는 자료를 없애고 데이터의 크기를 줄일 수 있다. (그림 2.)에서 Canny 에지 알고리즘 적용후의 모습을 볼 수 있다.

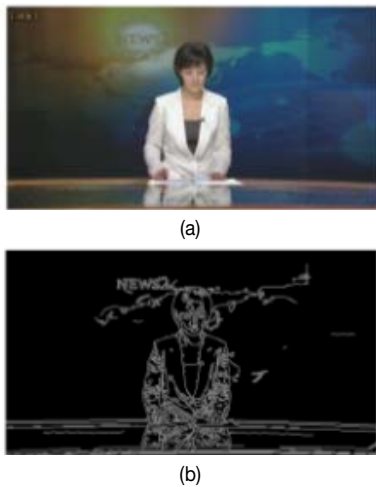


그림 2. Canny 에지 적용 전과 후  
(a)원 영상, (b)캐니에지영상

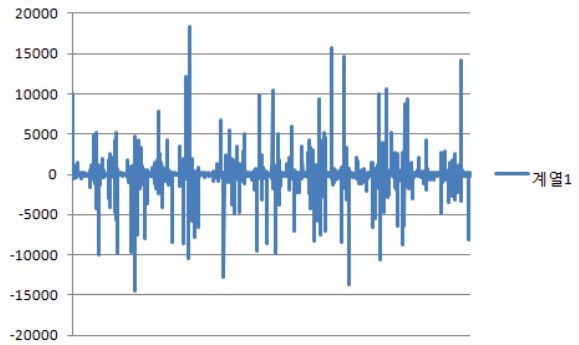


그림 3. 에지 총합의 차 그래프

입력영상에 Canny 에지 알고리즘을 사용하여 나온 출력영상은 이진화 되어있는데 이 이진화 된 영상에서 에지에 해당하는 픽셀의 총 합을 구할 수 있다. 현재 프레임의 에지 총합과 이전 프레임의 에지 총합의 차이를 이용하여 장면전환을 검출하는 방법을 사용하였다. 이러한 차이를 그래프로 나타내면 (그림 3)과 같다. X축은 순차적 차영상, Y축은 에지 총합의 차이 값이다.

### III. 실험 및 결론

실험에 쓰인 영상은 뉴스영상으로 영상의 구성은 아래 표 1과 같다.

표 1. 테스트 News 영상

항목	값
초당프레임	30fps
영상크기	640x360
재생시간	8min 24sec
용량	103,047KB
1프레임 디졸브 컷	92컷
10 프레임 디졸브 컷	11컷

구현은 C++와 OpenCV 2.0을 사용하였다. <표 1>의 1프레임 디졸브 컷은 이전 프레임과 현재프레임이 급진적으로 전환되는 듯이 보이지만 그 사이에 1프레임의 디졸브가 적용되어 있는 것을 말한다. 10프레임 디졸브 컷은 이전 프레임과 현재 프레임 사이에 총 8프레임의 디졸브가 적용된 것을 말한다. 이 두 가지 장면 전환의 에지 총합의 차 그래프는 아래와 같다. X, Y축은 (그림 3)과 동일하다.

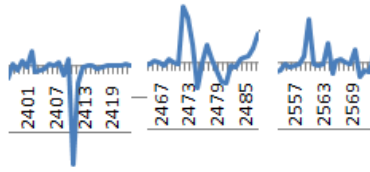


그림 4. 1프레임 디졸브 컷

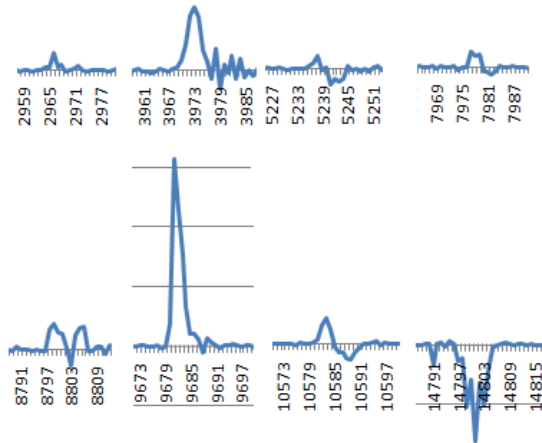


그림 5. 10프레임 디졸브 컷

Canny 에지를 이용한 알고리즘에서는 파라미터로써 2개의 값이 필요한데 이 비율은 1:2, 1:3이 좋다고 알려져 있다. 임계값은 3000이상 또는 -3000이하인 경우에 장면을 추출 하도록 임의 설정하였다. 그에 따른 실험결과는 아래 <표 2>와 같다.

표 2. News 영상을 이용한 실험결과

Canny 에지 설정	1프레임 디졸브	10프레임 디졸브	오검출	중복
100, 150	19/92	6/11	36	44
100, 50	10/92	4/11	44	53

Canny 에지를 100, 150으로 설정한 경우 1프레임 디졸브 92개중 19개를 찾지 못하였고, 10프레임 디졸브 11개중 6개를 찾지 못하였다. 오검출은 장면전환이 아닌 부분을 추출한 결과를 말한다. 중복은 장면전환과 1프레임 디졸브를 구별하지 못하고 두 가지 모두 장면으로 추출한 결과를 말한다. 1프레임 디졸브컷은 앞 프레임과 디졸브 프레임, 뒤 프레임을 합하여 총 3프레임으로 구성된다. 중복 추출된 결과에서는 이전 프레임을 제외한 차이값 변화가 큰 디졸브 효과가 적용된 프레임과 뒤 프레임이 동시에 검출된다. 이러한 문제는 추출된 결과 프레임의 간격이 1프레임 이하일 때 장면전환 효과가 적용되지 않은 뒤 프레임을 선택하도록 하여 해결할 수 있다. 또한 오검출에 해당하는 부분을 영상에서 살펴본 결과 과거의 전쟁 동영상이나 시민제보 동영상들이었다. 이 동

영상들에는 다수의 잡음이 있고 프레임이 부족한 경우가 많아 오검출을 하게되며 한 장면에서 계속적인 오검출이 발생하였다. 이는 추출된 장면과 장면사이의 프레임이 2프레임 이상 10프레임 이하일 경우 이전 프레임을 선택하게 하여 오검출을 줄일 수 있으며 <표 3>에서 이를 적용한 결과를 보여준다.

표 3. News 영상을 이용한 실험결과2

Canny 에지 설정	1프레임 디졸브	10프레임 디졸브	오검출	중복
100, 150	19/92	6/11	20	0
100, 50	10/92	4/11	21	0

위와 같은 설정으로 중복 검출된 영상을 제거하고 오검출을 줄일 수 있었다. 또한 Canny 에지 알고리즘의 설정만으로도 89.13% 확률로 1프레임 디졸브 컷과 54.55% 확률로 10프레임 디졸브 컷을 찾아낼 수 있었으며 오검출율은 19.09%이다.

### 참고문헌

- [1] Kim Won-Hee, A Study on Setting Adaptive Thresholds and Weighting Factors for Real-time Shot Change Detection, PUKYONG NATIONAL UNIVERSITY, 2009
- [2] Seok-joong Yoon. Gradual Scene Change Detection in a Compressed Video Sequence, KoSBE, 1998
- [3] A.Nagasaka and Y.Tanaka, "Automatic Video Indexing and Full-videoSearch for Object Appearances," inProc.IFIP Conf. VisualDatabaseSystems, Vol.A-7, pp.113~127, Oct.1991.
- [4] G.Camara-Chavez,F.Precioso,M.Cord,S.Phillip-Foliguet,and A.de A.Araujo,"ShotBoundary Detection by a Hierarchical Supervised Approach," Int.Works.Systems,Signalsand Image Processing andEURASIP Conf.Speechand ImageProcessing, Multimedia Communications and Services, pp. 197~200. Jun. 27~30,2007.
- [5] Y.Tonomura and S.Abe,"Content Oriented Visual nterface Using Video Icons for Visual atabase Systems," Journal of VisualLanguagesand Computing,Vol.1,No.2,pp.183~198, Oct.1990.
- [6] B.L Yeo, B. Miu, "Rapid Scene Analysis on Compressed Video", IEEE Transactions on Circuits and Systems for Video Technology, vol. 5. no. 6, pp.533~544, 1995
- [7] Sung-jun Park, Moon-Ho Song, Woon-Kyung M, Min-Gyo Chung Automatic Detection of Dissolving Scene Change in Video, IEEK, 1999
- [8] J.F. Canny, "A Computational Approach to Edge Detection," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 8, no. 6, pp. 679~698, Nov. 1986.