

## 화소 밝기와 객체 이동을 이용한 비디오 샷 경계 탐지 알고리즘

이준구\*, 한기선\*\*, 유병문\*\*\*, 황두성\*

# Shot Boundary Detection Algorithm by Compensating Pixel Brightness and Object Movement

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

### 요약

비디오 데이터를 효율적으로 검색, 정렬, 탐색, 분류하기 위해서는 프레임 간의 샷 전환 탐지가 선행되어야 한다. 프레임 간 화소 밝기와 객체 이동은 높은 탐지율을 보장하는 샷 탐지 알고리즘이 극복해야 할 문제이다. 본 논문에서는 프레임의 블록화 및 객체의 이동과 프레임의 밝기를 고려하는 샷 탐지 방법을 제안한다. 먼저 연속하는 두 프레임 사이에서 발생할 수 있는 객체의 이동을 고려하여 계산된 히스토그램과 밝기 차이를 반영하는 모폴로지 팽창 연산을 이용하는 알고리즘을 제안한다. 다음으로 화소 밝기 차를 보정한 프레임 블록의 화소 정보와 프레임의 전역적인 밝기 히스토그램의 변화를 함께 이용하는 샷 탐지 방법을 제안한다. 제안된 방법들은 국가기록원 소장 비디오 데이터에 대한 실험에서 화소 또는 히스토그램 기반 알고리즘에 비해 높은 샷 탐지율을 보였다.

▶ Keywords : 샷 경계 탐지, 밝기 보상, 움직임 보상, 블록 기반 방법, 비디오 인덱싱

### Abstract

Shot boundary detection is an essential step for efficient browsing, sorting, and classification of video data. Robust shot detection method should overcome the disturbances caused by pixel brightness and object movement between frames. In this paper, two shot boundary detection methods are presented to address these problem by using segmentation, object movement, and pixel brightness. The first method is based on the histogram that reflects object movements and the morphological dilation operation that considers pixel brightness. The second method uses the pixel brightness information of segmented and whole blocks between frames. Experiments on digitized video data of National Archive of Korea show that the proposed methods outperforms the

•제1저자 : 이준구 •교신저자 : 이준구

•투고일 : 2013. 3. 22, 심사일 : 2013. 4. 30, 게재확정일 : 2013. 5. 5.

\* 단국대학교 컴퓨터과학과(Dept. of Computer Science, Dankook University)

\*\* 강동대학교 방송영상미디어과(Dept. of Visual Broadcasting Media, GangDong College)

\*\*\* (주)엘엔와이비전(L&Y Vision Technologies, Inc)

existing pixel-based and histogram-based methods.

- ▶ Keywords : Shot boundary detection, brightness-adjusted matching, motion-compensated matching, segmentation

## I. 서 론

디지털 기술 발전은 비디오 데이터의 빠른 생성과 더불어 컴퓨팅 자원의 활용을 증가시켰으나, 비디오 데이터의 통합 관리를 위한 저장, 검색, 분류 등 연관 서비스가 효과적으로 제공 되지 않은 실정이다. 비디오는 시간상에서 나타나는 영상과 음성으로 구성되며, 일정한 형식에 따라 구조화된 멀티미디어 데이터이다. 비디오는 멀티미디어 데이터의 종류 중 크기가 가장 크며, 여러 대의 비디오카메라로부터 촬영된 연속된 샷(shot) 또는 프레임(frame)들을 문맥상에서 관련 있는 신(scene)으로 편집하여 생성된다. 그러므로 편집된 비디오는 순서가 있는 신의 집합이며, 하나의 신은 하나 또는 여러 개의 카메라로부터 촬영된 프레임들로 구성된다. 신과 음성은 비디오의 의미 정보를 제공하는 단위가 된다. 샷 경계(shot boundary)는 하나의 카메라에서 촬영된 비디오가 다른 카메라의 비디오와 편집 연결 시 발생되며, 신을 구성하는 샷을 추출하는 주요한 단위가 된다. 샷 경계 탐지는 비디오 데이터의 저장, 색인, 검색 등의 서비스를 효율적으로 제공하기 위한 기반기술 이다(1).

샷 경계 탐지의 판별 방법에는 두 프레임의 비유사도를 계산하기 위해 화소의 차이 측정(2,3), 히스토그램의 차이 측정(4-8), 에지 변화 비율(edge change ration)(9), 칼라 모멘트(color moment) 비교 방법(10), 모든 프레임의 정량 수치를 측정하는 방법 등이 제안되었고, 연속된 프레임 또는 현재 프레임과 현재로부터 k번째 떨어진 참조 프레임을 비교하여 제안된 비유사도 값의 차이가 미리 설정한 임계값(threshold) 이상이면 두 프레임 사이에 샷 전환이 발생된 것으로 탐지한다. 이들 방법은 두 연속 프레임(현재와 참조 프레임)의 전체 또는 일부분에서 급진적 샷 경계와 점진적 샷 경계로 나눌 수 있다. 급진적 샷 경계는 연속된 프레임간의 비유사도 값이 선정된 임계값 이상일 때 샷 전환을 비교적 쉽게 검출할 수 있다. 그러나 점진적 샷 경계는 비디오 편집자

가 신을 편집하는 방법에 따라 비유사도의 변화가 매우 낮아 샷 검출이 안 되거나 오 검출되는 경우가 많다.

본 논문에서는 프레임들 간의 화소 밝기, 히스토그램의 변화, 객체 이동 등 복합된 영상 정보를 이용하는 샷 전환 탐지 알고리즘을 제안하였다. 제안하는 방법은 연속된 전체 프레임의 영상 밝기 정보와 프레임 블록화 기반 지역 정보를 이용한다. 첫 번째 알고리즘은 히스토그램 기반에서 객체 이동을 이용한다. 객체 이동 영상 정보는 블록 영역 내 객체의 이동 거리를 반영하여 대응되는 참조 프레임 내 블록의 밝기 히스토그램 값을 비교한다. 두 번째 알고리즘은 화소와 히스토그램 기반 샷 경계 탐지를 수행한다. 연속 프레임의 지역 영역의 변화는 대응되는 블록의 화소 차로 탐지하며, 두 프레임의 전체적 변화는 히스토그램을 사용한다. 블록간의 급격한 화소 밝기변화에 의한 오류를 최소화하기 위해 참조 프레임의 블록의 화소의 변화를 보상시켰다.

이 논문의 2장에서는 샷 경계 탐지 관련 연구에 대해 살펴보고, 3장에서는 제안 알고리즘을 기술한다. 4장에서는 제안한 두 알고리즘을 국가 기록원 소장 디지털화 영상에 대하여 테스트한 결과를 기술하며, 화소 또는 히스토그램 차를 이용한 알고리즘과 성능을 비교한다. 마지막으로 제안하는 알고리즘의 문제점 및 개선 방향을 제시한다.

## II. 관련 연구

샷 경계 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 N개의 프레임으로 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 밝기 값이다. 선택된 비유사도 측정값  $D(k) = d_k$ 는 현재 프레임  $F_k$ 와 참조 프레임  $F_{k+1}$ 의 샷 경계 여부를 나타낸다.

Zhang 등(2)과 Xiaoquan Yi 등(3)은 샷 전환 탐지를 위해 두 연속 프레임사이에 서로 대응되는 화소 값의 차이를 이용한 프레임간의 평균 밝기 차이를 사용했다(1).

$$d_k = \frac{1}{XY} \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} 1(|F_k(x, y) - F_{k+\tau}(x, y)| < \epsilon) \quad (1)$$

여기서  $x > 0$ 이면  $1(x) = 1$  그 외는  $1(x) = 0$ 이다.  $\epsilon$ 은 화소 차에 대한 허용 임계값(tolerance threshold)이다.  $d_k$ 은  $k$  번째 현재 프레임과  $k + \tau$  번째 참조 프레임 간의 화소 변화의 평균이다.  $d_k$ 가 샷 경계 임계값  $\theta$  이상이면  $(k + \tau)$  번째 프레임이 새로운 샷의 시작이 된다. 만약  $\tau = 1$ 이면 참조 프레임은  $F_{k+1}$ 가 된다. 전체적인 화소의 변화를 고려하면 카메라와 객체 움직임에 의해 높은  $d_k$ 가 나타나는 단점을 보완하기 위해 평균 필터링 또는 움직임 보상 등이 이용되었다.

샷 전환 탐지의 비유사도 측정에 그레이, RGB, HCV, YCbCr 등 정보의 히스토그램이 이용되었다[2,4-8].  $C$ 개의 색과  $L$ 개의 빈(bin)을 고려 시  $H_k(c_{ij})$ 를  $k$ -번째 프레임에서 색  $c_{ij}$ 의 히스토그램 값일 때 식 (2)가 사용된다.[2.6,7].

$$d_k = \frac{1}{LC} \sum_{i=0}^{C-1} \sum_{j=0}^{L-1} 1(|H_k(c_{ij}) - H_{k+1}(c_{ij})| < \epsilon) \quad (2)$$

히스토그램 측정값은 화소 차의 비교 방법의 전체적인 데이터를 반영하는 단점을 극복할 수 있으나 선택하는 비유사도 방법에 따라 계산시간이 달라진다. 그러나 히스토그램 정보는 샷 경계 탐지에서 보편적으로 선택되었다.

프레임을 일정 크기의 중복이 없는 블록 영역으로 구분하고 대응 블록 영역의 화소 비교, 통계 값 비교, 히스토그램 비교 등을 이용한다.  $X_b \times Y_b$  크기의  $B$ 개로 프레임을 블록화한 경우  $F_k$ 의  $l$ -번째 블록 내 화소를  $b_{k,l}(x,y)$ 라 할 때, 블록 내 화소 차로부터 비유사도는 식 (3)과 같다.

$$d_k = \frac{1}{B} \sum_{i=0}^{B-1} \left[ \frac{1}{X_b Y_b} \sum_{x=0}^{X_b-1} \sum_{y=0}^{Y_b-1} 1(|b_{k,i}(x,y) - b_{k+1,i}(x,y)| < \epsilon) \right] \quad (3)$$

샷 경계 탐지 방법은 유사도가 가장 높은 매칭 블록을 이용하거나, 대응되는 블록의 유사도에 가중치를 부여하여 샷 경계를 탐지하는 연구 등이 진행되었다. 프레임 간 화소 차의 비교보다 블록 영역 내 차이를 반영하면 작은 객체 이동으로 인한 오탐지율을 감소시켰다고 보고되었다. Zabih 등(9)은 연속적인 두 프레임에서 에지들의 생성 또는 소멸 정도를 측정하는 에지 변화율(edge change ratio)을 사용하여 두 연속 프레임사이의 샷 전환을 탐지했다.

기 연구된 특징 벡터로부터 새로운 특징 벡터를 구성 후 기계학습을 적용하여 샷 경계를 탐지하는 연구가 진행되었다. Matsumoto 등[12]은 연속된 프레임 간의 화소 차이, 칼라, 에지 변화율, 영역 분할을 사용하는 여러 SVM 학습 기반 샷 탐지를 수행하였다. 학습 모델은 급진적 샷 탐지와 점진적 샷 탐지를 독립시켜 구성시키며, TRECVID 데이터의 실험에서 특징 벡터의 복합된 사용이 특징 벡터의 개별 이용보다 높은 탐지율을 보였다.

컬러 히스토그램과 통계학적 접근 방법을 합성하는 연구도 진행되었다.  $\chi^2$  히스토그램의 장점을 합성하는 방법[13]에서는 컬러 히스토그램의 차이를 계산한 후 NTSC 표준에 따른 명암도 등급변환 상수를 이용하여 프레임 간 차이를 강조한다. 이 방법은 급진적, 점진적 장면 전환 모두 검출 가능하다고 보고되었다.

연구된 알고리즘들은 제안하는 특징 벡터의 설정과 유사도 측정, 임계값의 설정, 정확율(precision)과 재현율(recall)을 이용한 테스트 등으로 요약된다. 두 연속 프레임사이의 동일한 화소 또는 선정된 블록을 비교하는 방법과 히스토그램을 이용한 방법은 계산이 간편하여 보편적으로 사용되나 객체의 이동이나 화소의 밝기 등의 변화에 취약하고, 에지변화율 특징을 이용한 방법은 밝기 변화나 객체 이동에는 강하나 탐지를 위한 계산 량이 많고 잡음에 민감한 문제점이 있다. 임계값의 설정에 따른 탐지율에 의존되고 있으며, 높은 탐지율을 보장하는 적응적 탐지 방법에 대한 연구는 진행 중이다.

### III. 샷 탐지 알고리즘

#### 1. 객체이동과 밝기 변화를 이용한 샷 전환 탐지 알고리즘

카메라에 의해 촬영된 샷은 일정한 시간차를 두고 획득된 여러 개의 비디오 프레임으로 구성이 되어, 현재 프레임과 참조 프레임사이에 존재하는 객체의 움직임은 샷 전환 탐지에서 오 탐지의 원인이 된다. 제안하는 샷 전환 탐지 방법은 프레임 블록화와 객체의 이동 정보를 반영하는 고정 영역 내 히스토그램을 이용한다. 현재 프레임의 블록 히스토그램은 참조 프레임의 대응 블록을 포함한 확장된 영역의 히스토그램과 비교한다. 참조 프레임의 확장된 영역은 그림 1과 같이 두 프레임 간 객체의 최대 이동을 고려하여 설정한다. 참조 프레임의 대응 블록의 히스토그램에 모폴로지 그레이-레벨 팽창(morphological gray-level dilation) 연산을 적용하여 프

레이프 간 밝기 변화를 보상한다.

객체 이동 거리  $\Delta$ 를 고려한 참조 프레임의 소구역의 크기는  $(X_b+2\Delta)*(Y_b+2\Delta)$ 이며, 참조 프레임의 소구역의 중심 영역은 현재 프레임의 소구역의 중심 영역과 일치하며 히스토그램 계산 시 대응 블록의 주변에  $\Delta$ 개의 중복 화소가 고려된다. 현재 프레임  $F_k$ 의  $i$ -번째 블록의 히스토그램  $H_{k,i}(\cdot)$ 와, 참조 프레임  $F_{k+1}$ 의 소구역으로부터 화소 밝기 히스토그램  $H_{k+1,i}(\cdot)$ 를 계산한다(그림 2). 히스토그램  $H_{k,i}(\cdot)$ 과  $H_{k+1,i}(\cdot)$ 에 대한 모폴로지 그레이-레벨 팽창 연산은 식 (4)와 같다.

$$H_{k,i}(j) = \text{Max}(H_{k+1,i}(j+x) \oplus S(x)) \quad (4)$$

여기서  $H_{k,i}(j)$ 는  $H_{k,i}(\cdot)$ 히스토그램의  $j$ 번째 빈,  $S$ 는 모폴로지 연산을 위한 1차원의 모폴로지 마스크, 그리고  $x$ 는 마스크의 크기이다. 만약 구성 요소  $S$ 가 7이고, 히스토그램의  $j$ -번째 빈에 있으면,  $H_{k,i}(j)$ 는  $H_{k,i}(j-3)$ ,  $H_{k,i}(j-2)$ ,  $H_{k,i}(j-1)$ ,  $H_{k,i}(j)$ ,  $H_{k,i}(j+1)$ ,  $H_{k,i}(j+2)$ ,  $H_{k,i}(j+3)$  중에서 최대치 값을 선택한다.



그림 1. 현재 프레임과 참조 프레임의 분할  
Fig. 1. Segmentations of current and reference frames

그림 2에서 참조 프레임의 대응 블록의 히스토그램(d)은 현재 프레임의 해당 블록의 히스토그램(c)보다 많은 정보를 포함하고 있다. 참조 프레임의 블록으로부터 얻은 히스토그램은 현재 프레임의 블록으로부터 얻은 히스토그램이 포함하고 있는 정보와 함께 객체의 움직임과 밝기 변화를 보상한 정보를 포함한다. 그러므로 만약 연속적인 두 프레임이 같은 샷에 속한다면, 현재 프레임의 블록으로부터 얻은 히스토그램은 참조

프레임의 블록으로부터 얻은 히스토그램에 포함되어야 한다.  $F_k$ 와  $F_{k+1}$ 에서 샷 전환의 여부는 식 (5)로 판별한다.

$$d_{k+1} = \frac{1}{BL} \sum_{i=0}^{B-1} \sum_{l=0}^{L-1} 1(H_{k+1,i}(l) - H_{k,i}(l) \geq \theta_1) \quad (5)$$

여기서  $L$ 은 히스토그램의 빈의 수,  $H_{k+1,i}(l) - H_{k,i}(l)$  값이 설정한 임계값  $\theta_1$ 보다 크면 참조 프레임의 블록이 현재 프레임의 블록이 갖고 있는 모든 정보를 포함하고 있다고 간주된다.  $d_{k+1}$ 가 샷 탐지 임계값  $\theta_2$ 보다 크면  $F_k$ 와  $F_{k+1}$ 은 동일한 샷으로 간주하고, 아니면 새로운 샷이  $F_{k+1}$ 에서 시작된다고 결정한다.

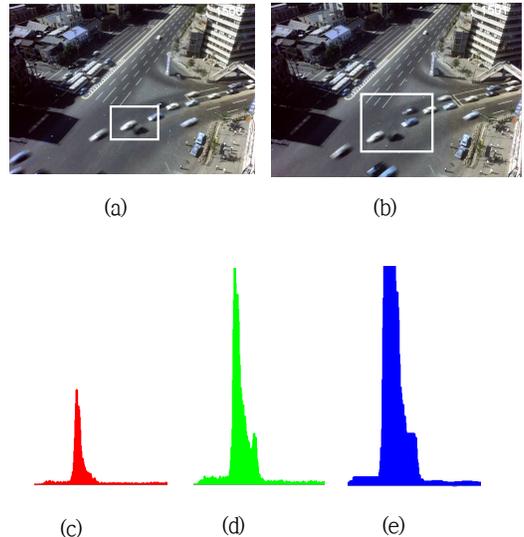


그림 2. 현재 프레임과 참조 프레임의 블록과 밝기 히스토그램의 예  
(a) 블록의 예를 보이는 현재 프레임, (b) (a)의 블록과 대응되는 블록을 보이는 참조 프레임, (c) 현재 프레임의 블록(a)의 밝기 히스토그램, (d) 참조 프레임의 블록(b)의 밝기 히스토그램, (e) (d)에 그레이-레벨 팽창 연산이 적용된 후의 밝기 히스토그램  
Fig. 2. Histogram examples of segmented blocks in current and reference frames (a) Segmented block of current frame, (b) Segmented block of reference frame, (c) Histogram of segmented block in current frame, (d) Histogram of segmented block in reference frame, (e) Morphological gray-level dilation of (d)

## 2. 화소와 히스토그램을 이용한 샷 전환 탐지 알고리즘

화소 기반 샷 탐지 알고리즘에서 전체 정보를 이용하면 객체의 이동 또는 급격한 밝기 변화에 민감하게 반응하여 오탐지율이 높게 나타난다[11]. 국가 기록된 소장 영상의 경우 웨

손으로 인해 밝기 차가 나타나며 이러한 경우를 오탐지의 주요 원인으로 분석한다. 프레임의 화소 밝기를 이용한 샷 탐지 방법은 그림 3과 같이 훼손으로 인한 밝기 차를 서로 다른 샷으로 오탐지 한다.

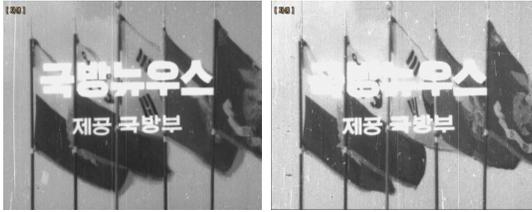


그림 3. 훼손에 의한 샷 탐지 오류  
Fig. 3. Shot boundary detection error due to damage

제안하는 샷 탐지 방법은 각 프레임 간 밝기 차의 평균, 프레임을 블록화한 후 지역적인 영역에서 나타나는 화소 밝기의 차, 그리고 연속한 여러 프레임의 글로벌적인 영역에서 나타나는 히스토그램 차를 사용한다. 그림 4에서 (a), (b)는 프레임의 블록화를 나타내고, (c)는 대응하는 블록을 기반으로 계산한 유사도 맵이며 (d)는 프레임 전체를 기반으로 하여 계산한 밝기 히스토그램이다.

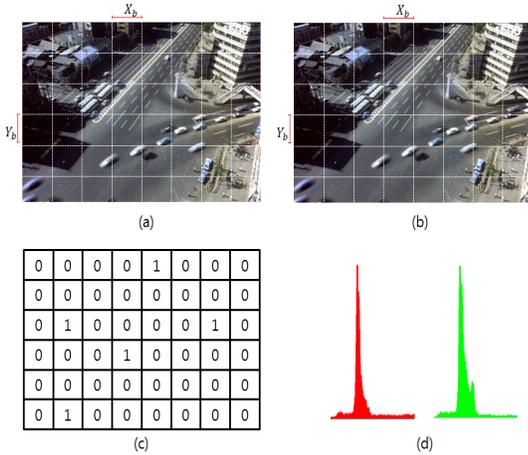


그림 4. 현재 프레임과 참조 프레임의 블록 차와 히스토그램 (a) 블록으로 나눈 현재 프레임, (b) 블록으로 나눈 참조 프레임, (c) (a)와 (b)의 화소 블록 정보를 이용해 계산한 결과 맵, (d) (a)와 (b)의 히스토그램

Fig. 4. Block difference and histogram examples (a) Segmentation of current frame, (b) Segmentation of reference frame, (c) Block difference of (a) and (b), Histograms of (a) and (b)

급격한 밝기 차이에 의한 오탐지를 보완하기 위해 프레임 전체를 기반으로 하여 계산한 화소 당 밝기 차의 평균을 블록 영역 내 화소의 차를 계산하는데 반영한다.  $F_k$ 와  $F_{k+1}$ 의 밝기 차이의 평균인  $I(k, k+1)$ 은  $F_k(x, y)$ 와  $F_{k+1}(x, y)$ 의 화소 값의 차이의 평균값 식 (6)을 이용한다.

$$I(k, k+1) = \frac{1}{XY} \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} (F_k(x, y) - F_{k+1}(x, y)) \quad (6)$$

연속된 두 프레임의 블록 간 차  $n(k, k+1)$ 는 프레임 간 평균 밝기  $I(k, k+1)$ 을 이용한다.

$$n(k, k+1) = \frac{1}{B} \sum_{i=0}^{B-1} \left[ \frac{1}{X_b Y_b} \sum_{x=0}^{X_b-1} \sum_{y=0}^{Y_b-1} 1 (|b_{k,i}(x, y) - b_{k+1,i}(x, y) - I(k, k+1)| - \theta_2) \right] \quad (7)$$

여기서  $\theta_2$ 은 샷 전환일 가능성이 있는 화소를 탐지하기 위한 임계값이다. 밝기 히스토그램은 객체의 이동 정보를 갖기 위해 객체를 추적할 수 있도록 연속된 프레임에 대해 계산한다. 추출된 히스토그램을 이용해  $k$ 로부터 연속하는  $L$ 개의 프레임까지의 히스토그램 차이  $h(k, k+L)$ 를 계산한다.

$$h(k, k+L) = \frac{1}{LC} \sum_{c=1}^L \sum_{i=c}^{C-1} 1 (H_k(c) - H_{k+i}(c) - \theta_3) \quad (8)$$

여기서  $L$ 은 계산에 필요한 프레임의 개수이며,  $\theta_3$ 은 두 프레임 사이의 히스토그램의 차를 탐지하기 위한 임계값이다. 마지막으로 샷 전환 여부를 나타내는  $d_{k+1}$ 을 식 (9)을 이용하여 계산한다.

$$d_{k+1} = \begin{cases} 1 & \text{if } z(k, k+1) > \theta_4 \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

여기서  $z(k, k+1) = \frac{n(k, k+1) + h(k, k+L)}{2}$  이고,  $\theta_4$ 는 샷 전환 프레임을 탐지하기 위한 임계값이다.

### IV. 실험결과

제안한 알고리즘의 성능을 평가하기 위하여 국가기록원의 디지털화 영화 필름을 사용하였다. 테스트에 사용된 영화는

1960년대부터 1970년대 사이에 촬영된 것으로 텔레시네 장치를 사용하여 아날로그 영상을 디지털로 변환된 영상이다. 영상의 특징으로 샷 전환 탐색의 오류에 크게 영향을 끼치는 흠집, 얼룩, 떨림, 찢김, 번쩍거림 등의 훼손이 존재한다. 사용된 디지털화 영화 필름의 프레임 수, 샷 전환수는 표1과 같다.

표 1. 실험 비디오  
Table 1. Test Videos

구분	프레임 수	샷 전환 수
V1	3,000	19
V2	3,000	13
V3	3,000	27
V4	3,000	22
V5	3,000	31
V6	3,000	29

제한된 알고리즘의 평가에서 사용된 디지털화 비디오는 흑백 비디오로 변환시켜 프레임간의 비교에는 그레이-레벨의 화소 밝기와 히스토그램이 사용되었다. 샷 탐지 알고리즘의 성능을 평가하는 기준으로는 재현율 rec과 정확도 pre와 F-measure를 사용한다.

$$rec = \frac{D}{D+M} \tag{10}$$

$$pre = \frac{D}{D+N} \tag{11}$$

$$F\text{-measure} = 2 \times \frac{pre \times rec}{pre + rec} \tag{12}$$

여기서 D는 탐지한 샷 전환 수, M는 탐지하지 못한 샷 전환 수, 그리고 N은 실제로 샷 전환이 발생하지 않았으나 오 탐지한 수를 나타낸다. 식 (10)에서 재현율이 1이면 모든 샷 경계를 탐지한 것을 의미한다. 식 (11)에서 정확도의 값이 1이면 모든 샷이 예측된 것이다. 따라서 식 (12)에서 F-measure의 값이 1에 가까울수록 모든 샷 경계를 탐지하고 오탐지의 경우가 없다는 것을 의미한다.

각 단계에서 사용한 임계값의 설정을 위해 컷의 발생 빈도

가 높고 종류가 다양한 표본영상 V를 사용하였다.  $\theta_1$ 은 V에서 샷 전환이 발생하는 프레임들의 소 영역 기반 히스토그램 차이의 평균,  $\theta_2$ 는 V에서 샷 전환이 발생하는 프레임들의 화소 밝기 차이의 평균,  $\theta_3$ 는 V에서 샷 전환이 발생하는 프레임들의 전역적 히스토그램 차이의 평균, 그리고  $\theta_4$ 는 V에서 샷 전환이 발생하는 프레임들의 비율의 평균으로 사용한다.

임계값의 설정 방법에 따라  $\theta_1 = 30, \theta_2 = 65, \theta_3 = 50, \theta_4 = 60$  으로 설정 했다. 그림 5와 같은 실험을 통해 제안 알고리즘 1의 프레임 간 객체 이동 거리  $L$ 를 20화소로 사용하였고, 밝기 변화를 보상하기 위해 길이가 10인 1차원 모폴러지 마스크를 사용했다. 제안 알고리즘 2의 객체 이동 거리 보상을 위한 L은 화소 기반의 알고리즘 보다 더 많은 객체의 움직임 정보를 갖기 위해 연속한 두 프레임보다 하나 더 많은 3프레임으로 사용하였다. 그림 6은 설정한 임계값에 대한 샷 탐지의 예이다.

본 실험에서는 블록 기반 히스토그램을 사용하는 제안 알고리즘1, 지역적 화소 정보와 전역적 히스토그램 정보를 함께 사용하는 제안 알고리즘2, 기존의 히스토그램 기반 알고리즘, 그리고 기존의 화소 기반 알고리즘을 적용한 결과를 표2에 나타냈다.

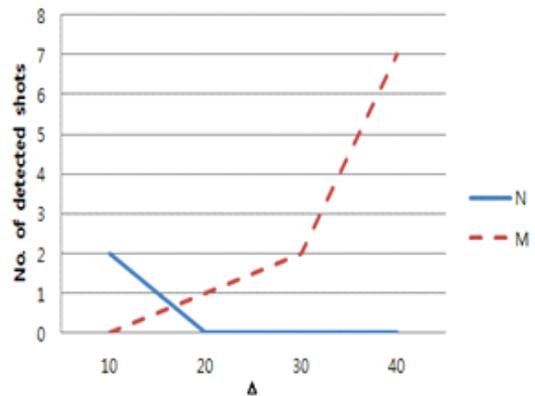


그림 5. 객체이동거리(Δ)에 따라 탐색한 N과 M의 수  
Fig. 5. Number of N and M for object-movement Δ

표 2. 사용된 비디오에 대한 실험 결과  
Table 2. result on experiments about used video

구분	제한 알고리즘1			제한 알고리즘2			히스토그램			화소		
	pre	rec	F-m	pre	rec	F-m	pre	rec	F-m	pre	rec	F-m
V1	0.95	0.82	0.88	0.61	0.92	0.73	0.67	0.63	0.65	0.56	1.00	0.72
V2	0.92	1.00	0.96	1.00	0.93	0.96	1.0	0.81	0.90	1.00	0.67	0.80
V3	1.00	0.87	0.93	0.96	0.87	0.91	0.96	0.67	0.79	0.97	0.55	0.70
V4	0.55	1.00	0.71	0.91	0.91	0.91	0.91	0.87	0.89	0.55	1.00	0.71
V5	0.97	0.88	0.92	0.94	0.85	0.89	0.90	1.00	0.95	0.94	0.49	0.64
V6	1.00	0.63	0.77	0.83	0.89	0.86	0.90	0.70	0.79	0.71	0.69	0.70
평균	0.90	0.87	0.88	0.88	0.90	0.89	0.89	0.78	0.83	0.79	0.73	0.76

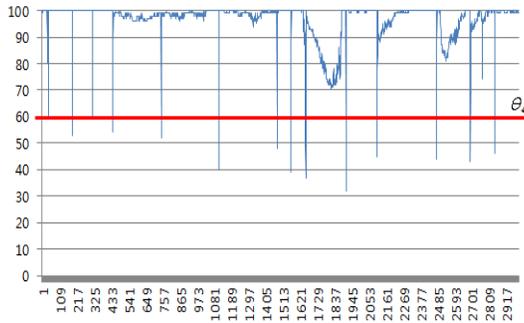


그림 6.  $\theta_4=60$ 으로 탐지된 컷의 예  
Fig. 6. Example of shot boundary when  $\theta_4=60$

훼손 정도가 약하고 객체 또는 카메라 이동이 적은 경우에 대해서는 전역적 히스토그램 또는 지역적 화소 정보만 사용해도 좋은 결과를 보였다. 그러나 훼손이 심하거나 객체 또는 카메라의 이동이 큰 경우 지역적 화소 정보만 사용하면 화소의 차가 크기 때문에 정확도와 재현율이 반비례하였다. 같은 경우에 대하여 전역적 히스토그램 정보만 사용하면 밝기에 매우 민감하여 변색 또는 변쩍거림 등의 잡음에 크게 반응하였고, 샷 전환이 발생해도 밝기가 비슷한 경우 탐지율이 낮았다. 히스토그램과 화소 기반 알고리즘 보다 제안 알고리즘 2에서 탐지한 샷 전환수가 증가하고 오탐지와 탐지하지 못한 샷 전환수가 줄었다. 평균적으로 두 제안 알고리즘 모두 기존의 알고리즘에 비해서 정확도와 재현율이 향상되었다. 특히 재현율은 약 10.0%, F-measure는 약 13.0% 향상되었다.

## V. 결론

기존의 히스토그램을 이용한 샷 전환 탐지 방법은 연속적인 두 프레임의 전체 또는 대응되는 동일한 크기의 블록의 히스토그램을 사용하며, 객체의 이동이나, 프레임의 밝기 변화에 대해서 취약한 문제점이 있었다. 본 논문에서는 객체 이동과 밝기 보상을 고려한 소 영역 히스토그램 기반의 알고리즘과, 밝기를 보상하는 지역적 화소 정보와 전역적 히스토그램 정보를 함께 사용하는 알고리즘을 통해 각각의 단점을 보완하였다. 실험에서 기존의 화소와 히스토그램 기반 알고리즘보다 제안된 방법은 임계값이 실험에 의해 정해진 범위 내에 있는 경우 샷 전환 탐지 결과가 우수했다.

## 참고문헌

- [1] Jinhui Yuan, Huiyi Wang, Lan Xian, Wujie Zheng, Jianmin Li, Fuzong Li, Bo Zhang, "A Formal Study of Shot Boundary Detection," IEEE Trans. on Circuits and Systems for Video Technology, Vol. 17, No. 2, pp. 168-186, Feb. 2007
- [2] H. J. Jhang, A. Kankanhalli, S. W. Smoliar, "Automatic Partitioning of Full-Motion Video," Multimedia Systems, Vol. 1, No. 1, pp. 10-28, June 1993.
- [3] Xiaoquan Yi, Nam Ling, "Fast Pixel-Based Video Scene Change Detection," IEEE Int. Symposium on Circuits and Systems, Vol. 4, pp. 3443-3446, May 2005.
- [4] A. Nagasaka, Y. Tanaka, "Automatic Video Indexing and Full-Video Search for Object Appearances," IFIP Proceeding of Visual Database Systems, pp. 113-127, 1992.
- [5] Young-Wook Han, Seong-Il Cheong, Seong-Jae Kim, Si-Young Lee, Sung-Ho Kim, "Improving Histogram Scene Change Detection Method Using Motion Vector," Proceedings of Fall Conf. on Korean Institute of Information Scientists and Engineers(KIISE-1999), Vol. 26, No. 2, pp. 410-412, 1999.
- [6] B. H. Shekar, Holla K. Raghurama, Kumari M. Sharmila, "Video Cut Detection Using Chromaticity Histogram," Int. Journal of Machine Intelligence, Vol. 3, No. 4, pp. 371-375, Dec. 2011.
- [7] Z. Cernekova, C. Kotropoulos, I. Pitas, "Video Shot Segmentation Using Singular Value Decomposition," Proc. 2003 IEEE Int. Conf. Multimedia and Expo, Baltimore, Maryland, Vol. 2, pp. 301-304, July 2003.
- [8] Priyadarshinee Adhikari, Neeta Gargote, Jyothi Digge, B. G. Hogade, "Abrupt Scene Change Detection," World Academy of Science, Engineering and Technology 42 2008.

[9] R. Zabih, J. Miller, K. Mai, "A Feature-based Algorithm for Detecting and Classification Production effects," ACM Multimedia System, Vol. 7, No. 1, pp. 119-128, March 1999.

[10] Robert A. Joyce, "Temporal Segmentation of Video using Frame and Histogram-Space," IEEE Trans. on Multimedia, Vol. 8, No. 1, pp. 130-140, Feb. 2006.

[11] In-Shik Jeong, Oh-Jin Kwon, "Video Shot Boundary Detection using Relative Difference between Frames," Optical Engineering, Vol. 42, No. 3, pp. 604-605, March 2003.

[12] Matsumoto, Kazunori, "SVM-Based Shot Boundary Detection with a Novel Feature," IEEE Int. Conf. on Multimedia and Expo, pp. 1837-1840, July 2006.

[13] Shin S. Y., Baik S. E., Pyo S. B., Rhee Y. W., "Scene Change Detection Using Local  $\chi^2$ ," Journal of The Korean Society of Computer and Information, Vol. 15, No. 1, pp. 203-207, 2007 June.

**저 자 소개**



**이 준 구**  
 2012: 단국대학교 컴퓨터과학과 공학사  
 현 재: 단국대학교  
 전자계산학과 석사과정  
 관심분야: Image Processing,  
 Parallel Processing  
 Email : leejg01679@gmail.com



**한 기 선**  
 1988: 충남대학교 계산통계학과 이학사  
 1990: 충남대학교  
 계산통계학과 이학석사  
 2004: 충남대학교  
 정보통신공학과 박사수료  
 현 재: 강동대학교  
 방송영상미디어과 교수  
 관심분야: Image Processing,  
 Info-Communications  
 Email : ksdream1@hanmail.net



**유 병 문**  
 1987: 경북대학교 전자계산학과 공학사  
 1992: 충남대학교 전산과 이학석사  
 2000: Wayne State University  
 컴퓨터공학과 공학박사  
 현 재: (주)엘앤에이비전 대표이사  
 관심분야: Motion Analysis,  
 3D/2D image registration,  
 (medical) image processing  
 Email : byoungmoon.you@gmail.com



**황 두 성**  
 1986: 충남대학교 이학사  
 1990: 충남대학교 이학 석사  
 2003: Wayne State University, 박사  
 현 재: 단국대학교 컴퓨터과학과 부교수  
 관심분야: Machine Learning,  
 Parallel Processing,  
 Semantic Web  
 Email : dshwang@dankook.ac.kr