

논문-99-4-2-03

MPEG 비트스트림과 구간 복호 영상을 사용한 장면 전환 검출

나윤정*, 하명환*, 이상길*

Scene Change Detection Using MPEG Bitstream and Sectionally Decoded Video

Yoon-Jeong Na*, Myung-Hwan Ha*, and Sanggil Lee*

요 약

동영상에서 장면이 전환되는 지점을 빠르고 정확하게 검출하는 방법을 설계하였다. 이 방법은 MPEG 압축 영상에 대하여 시간적 표본화를 통하여 추출된 압축 영역의 데이터를 사용하여 장면 전환의 후보 구간들을 정하는 첫 번째 단계와, 이들 구간 안에서 각 프레임의 화소값을 얻고, 이를 사용하여 정확한 장면 전환 지점을 찾아내는 두 번째 단계로 구성된다. 두 번째 단계에서는 명암과 윤곽선 변화를 결합하여 장면 전환을 검출하였다. 또한 카메라 플래시 때문에 장면 전환으로 잘못 검출되는 것을 방지할 수 있는 방법을 연구하였다. 이상의 방법들을 통합함으로써 장면 전환을 빠르고 정확하게 검출할 수 있는 구조를 제안한다.

Abstract

We proposed an algorithm which detects scene changes in video with speediness and accuracy. It is a two-step approach. In the first step, we decide potential scene change segments using the compressed domain data extracted by temporal sampling of MPEG compressed video. In the second step, we determine the exact scene change positions using the pixel values of each frame in those segments by means of combining the intensity and edge changes. In addition we discuss the method to remove false detection generated from camera flash. Integrating the above methods, we introduce a structure that can detect scene changes speedily and accurately.

I. 서 론

방송국과 같이 많은 동영상을 보유하고 있는 곳에서는 그것을 효과적으로 관리하고 검색할 수 있는 환경을 구축하는 것이 필요하다. 디지털 동영상이 일반적으로 사용되면서 종래의 텍스트를 중심으로 한 검색 환경에서 동영상의 내용에 바탕을 둔 검색 환경으로 발전하고 있다. 따라서 동영상을 그 내용에 따라 색인하는 작업이 필요하게 되었다. 그것을 위해서 동영상 내용의 기본 단위가 되는 연속된 장면 단위로 동영상을 분할하는 작업이 선행되어

야 한다. 이것을 일일이 사람이 한다면 매우 번거롭고 많은 시간이 소요될 것이다. 따라서 이러한 동영상의 분할을 위해 장면이 전환된 지점을 자동으로 찾는 기술이 연구되었다^[1]. 이러한 기술은 동영상 데이터베이스의 구축, 영상 감시 같은 여러 분야에 응용될 수 있다.

장면 전환에는 인접하는 두 프레임 사이에서 갑자기 장면이 변하는 컷(cut)과 여러 프레임에 걸쳐 서서히 장면이 변하는 디졸브(dissolve), 와이프(wipe), 페이드 인(fade in), 페이드 아웃(fade out) 등과 같은 점진적 장면 전환이 있다. 컷에서는 프레임 사이의 차이가 크기 때문에 명암 또는 색 히스토그램의 차이를 계산하는 방법 등을 사용하여 비교적 간단하게 검출할 수 있으나, 점진적 장면 전환에서는 인접한 프레임 사이의 차이가 작기 때문에 컷

*한국방송공사 기술연구소
Korean Broadcasting System, Technical Research Institute

을 검출하는 방법을 그대로 적용할 수 없다. 점진적 장면 전환을 검출하기 위해 제안된 여러 방법들은 다소 이론적이기 때문에 그것을 실제 영상에 적용할 때 많은 오류를 낳는다. 컷을 검출할 때에도 영상의 움직임이 큰 부분, 조명이 변화하는 부분 또는 카메라 플래시가 터지는 부분에서는 컷으로 잘못 판정할 가능성이 높다.

장면 전환을 검출하는 많은 방법들은^[2] 압축 영역에서 하는 방법과 화소 영역에서 하는 방법으로 구분될 수 있다. 압축 영역에서는 MPEG과 같은 압축 영상으로부터 추출된 DCT(Discrete Cosine Transform) 계수 또는 움직임 벡터를 사용하여 장면 전환을 검출하는 것이다. 화소 영역에서는 압축 영상을 디코딩하여 얻은 화소값을 사용하여 장면 전환을 검출한다. 일반적으로 압축 영역에서는 검출 속도를 빠르게 할 수 있지만 다룰 수 있는 정보의 한계 때문에 검출 성능을 높이는 데 많은 제한이 있다. 반면에 화소 영역에서는 화소값을 얻기 위해 디코딩을 해야 하므로 처리 속도가 느리지만 영상에 관한 모든 정보를 다룰 수 있으므로 검출 성능을 높이는 데 제한이 적다.

이 논문에서는 압축 영역과 화소 영역을 결합하여 양쪽의 장점을 살림으로써 정확성을 유지하면서 검출 속도를 빠르게 할 수 있는 두 단계로 이루어진 장면 전환 검출 구조를 설계하였다. 화소 영역에서는 명암 변화와 윤곽선 변화를 효율적으로 결합하여 정확성을 유지하면서 검출 속도를 향상시켰다. 아울러 화소의 밝기 변화를 이용하여 카메라 플래시 때문에 장면 전환으로 잘못 검출되는 것을 방지할 수 있는 방법을 설명한다. 이러한 방법들을 이용하여 컷 검출, 점진적 장면 전환 검출, 카메라 플래시 검출을 하나의 통합된 알고리즘으로 구성한 것을 보인다. 실험을 통하여 제안된 방법의 성능을 확인하였다.

II. MPEG 비트스트림과 구간 복호 영상을 사용한 장면 전환 검출

MPEG과 같은 압축 영상의 비트스트림으로부터 직접 추출할 수 있는 DCT 계수 또는 움직임 벡터를 사용하여 장면 전환을 검출하는 여러 방법들이 제안되었다.^{[3][4][5][6]} 예를 들면, DC 계수의 히스토그램 차이를 이용하는 방법, 각 프레임에서 표본으로 추출한 DCT 계수로 벡터를 구성하고 이 벡터들의 내적을 이용하는 방법^[4], B 프레임의 경우에는 순방향 예측 움직임 벡터와 역방향 예측 움직임

벡터의 갯수를 비교하고, P 프레임에 대해서는 움직임 보상이 있는 매크로 블록과 움직임 보상이 없는 매크로 블록의 갯수를 비교하는 방법^[3] 등이 있다. 이렇게 압축 영상의 비트스트림 데이터를 직접 이용하면 처리 시간을 줄일 수 있다. 그러나 이러한 방법으로는 다룰 수 있는 정보가 한정되고 윤곽선과 같은 영상의 특징에 관련된 정보를 얻기 어렵기 때문에 검출 성능을 높이는 데 많은 제한이 있다. 반면에 화소값을 사용하여 장면 전환을 검출하면 디코딩으로 인해 처리 시간이 많이 소요되지만 정확성을 높일 수 있다. 따라서 압축 영상 비트스트림의 데이터를 직접 이용하는 방법과 영상의 화소값을 이용하는 방법을 적당히 결합하면 두 방법의 장점을 모두 취할 수 있다.

MPEG 영상은 GOP(Group of Pictures) 단위로 압축된다. GOP의 첫 번째 프레임은 I 프레임이다. 예를 들면 GOP안에 있는 프레임 수가 6이고 I 또는 P 프레임이 나타나는 주기가 3인 경우, MPEG 비트스트림의 프레임 순서는 IBBPBBIBBPBBI...와 같다. 따라서 MPEG 영상에서 GOP 간격으로 GOP의 첫 번째 프레임을 추출하여 그 프레임들을 비교한다면 시간적 표본화를 하면서 장면 전환을 검출하는 효과를 낳는다.

압축 영역에서 I 프레임에 대해서는 DCT 계수를 직접 얻을 수 있지만, P와 B 프레임에 대해서는 본래 영상의 DCT 계수를 직접 얻을 수 없어서 그 값을 추정하는 방법이 이용된다^[7]. 곧, 압축 영상을 디코딩하지 않고는 P와 B 프레임에 대한 정확한 DCT 계수를 직접 구할 수 없다. 이 논문에서는 압축 영역에서 정확한 값을 얻을 수 없는 P와 B 프레임의 DCT 계수를 사용하지 않고, 정확한 값을 얻을 수 있는 I 프레임의 DCT 계수만을 사용하였다. 곧, 그림 1에 보인 것과 같이 GOP의 첫 프레임만을 선택하고, 그 프레임에 포함된 DCT 블록들의 첫 번째 계수(DC 계수)만을 추출한다. 이 값을 사용하여 인접한 두개의 I 프레임을 비교한다. 이와 같은 구조는 일정 간격으로 시간적 표본화를 하면서 장면 전환을 검출하는 것이다. DC 계수를 사용하여 두 프레임을 비교하기 위해 DC 계수의 히스토그램 차이를 구하여 이 값이 만일 일정한 임계값 보다 작으면 두 프레임 사이에 장면 전환이 없는 것으로 보고, 그 값이 임계값을 넘으면 두 프레임 사이에 장면 전환이 있을 가능성이 있는 것으로 보고, 이 구간을 장면 전환 후보 구간으로 정한다. 다음 단계인 화소 영역에서는 이 후보 구간에 대해서만 부분 디코딩을 하여 화소값을 얻고 한 프레임씩 이동하면서 인접한 프레임들을 비교하여 그 구간 안에서 정확한 장면 전환 지점을 찾는다.

이와 같이 압축 영역과 화소 영역을 결합하면 압축 영역에서 시간적 표본화를 통하여 빠르게 후보 구간을 정하고, 화소 영역에서 정확한 장면 전환 지점을 검출할 수 있으므로 전체적으로 빠르면서도 정확한 장면 전환 검출을 구현할 수 있다. 화소 영역에서 장면 전환을 검출하는 구체적 방법에 관해서는 다음 장에서 설명한다.

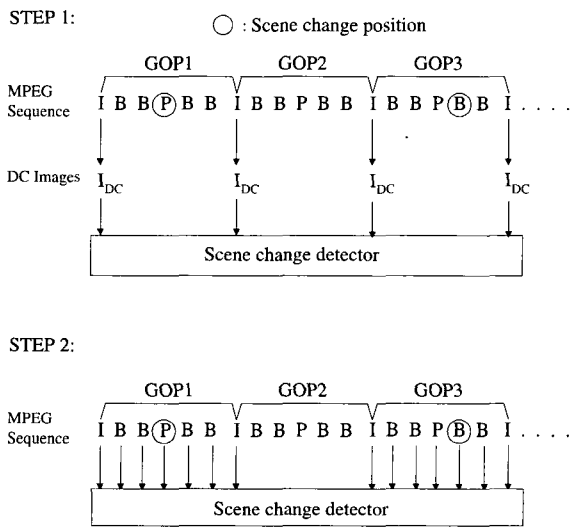


그림 1. 고속 장면 전환 검출 구조
Fig 1. High-speed scene change detection structure

Ⅲ. 화소 영역에서 장면 전환 검출

1. 명암과 윤곽선 변화를 결합한 장면 전환 검출

장면 전환을 검출하는 방법 가운데 화소 비교법, 히스토그램 비교법과 같이 영상의 명암 변화를 이용한 방법들이 있다. 이러한 방법들은 비교적 간단하여 구현하기 쉽기 때문에 많이 사용되고 있지만, 조명의 변화에 민감하고, 압축 영상을 복호화하여 그 방법들을 적용하는 경우에는 양자화 잡음의 영향을 많이 받는 단점이 있다. 이러한 문제점을 해결하기 위하여 영상의 특징을 이용하여 장면 전환을 검출하려는 방법이 연구되고 있다. 그 가운데 영상의 윤곽선 변화를 이용하는 방법이 제안되었다^[8]. 이 방법은 비록 많은 처리 시간을 요구하지만, 명암 변화에 바탕을 둔 방법들보다 정확성이 높고 점진적 장면 전환 검출에서

도 우수한 성능을 보인다.

여기서 윤곽선 변화만을 사용하면 많은 시간이 걸리지만, 명암 변화와 윤곽선 변화를 결합하면 검출 시간을 단축시킬 수 있다. 곧, 인접한 두 프레임 사이를 순차적으로 비교할 때, 먼저 명암 변화를 계산하여 두 프레임 사이의 차이가 특정 임계값을 넘을 때에만 윤곽선을 구하여 두 프레임을 비교한다. 이와 같은 방법을 사용하면 윤곽선 변화만을 사용한 방법에 견주어 볼 때 검출의 정확성을 그대로 유지하면서 처리 시간을 줄일 수 있다.

명암 변화를 이용한 방법 가운데 히스토그램 비교법을 사용한다면, 프레임 사이의 히스토그램 차이는 식 (1)과 같은 기준식에 의해 계산되고 비교된다. 여기서 G 는 히스토그램의 총 구간 수 이고, $H_i(j)$ 는 i 번째 프레임의 j 구간의 히스토그램 값이다.

$$HD_i = \sum_j^G |H_i(j) - H_{i-1}(j)| \quad (1)$$

윤곽선의 변화를 계산하기 위해서 기존의 윤곽선에 새로 추가되는 윤곽선의 양을 나타내는 ρ_{in} 과, 기존의 윤곽선 가운데서 사라지는 윤곽선의 양을 나타내는 ρ_{out} 을 아래 식 (2), (3)과 같이 각각 계산하고, 식 (4)와 같이 두 값 가운데 최대값을 취한다.

$$\rho_{in} = 1 - \frac{\sum_{x,y} \overline{E[x,y]} E'[x,y]}{\sum_{x,y} E'[x,y]} \quad (2)$$

$$\rho_{out} = 1 - \frac{\sum_{x,y} E[x,y] \overline{E'[x,y]}}{\sum_{x,y} E[x,y]} \quad (3)$$

$$\rho = \max(\rho_{in}, \rho_{out}) \quad (4)$$

여기서 x, y 는 각각 화소 평면의 수평축과 수직축의 좌표를 뜻하며, E 와 E' 은 각각 비교할 두 영상의 윤곽선 영상을 나타낸다. \overline{E} 는 E 를 일정한 반지름 r 만큼 팽창시켜 얻은 다이레이션된 영상이며, 다음과 같이 표현된다.

$$\overline{E}[x,y] = E[x,y] \oplus A_{mask}[x,y] \quad (5)$$

$$A_{mask}[x,y] = \begin{cases} 1 & |x| + |y| \leq r \\ 0 & \text{elsewhere} \end{cases}$$

컷에서는 ρ 가 큰 값으로 나타나고, 점진적 변화 부분

에서는 특수 효과에 따라 ρ 가 그 구간에서 증가하거나 감소한다. 곧, 디졸브의 경우에는 ρ 가 증가하다가 감소하는 모양을, 페이드의 경우에는 ρ 가 증가하는 모양을 갖는다. 윤곽선을 검출하는 여러 알고리즘 가운데 비교적 간단한 Sobel 연산자에 의한 방법을 사용하여 실험하였다.

2. 밝기 변화를 이용한 카메라 플래시 검출

뉴스 영상의 경우 카메라 플래시가 터지는 장면이 자주 나오는데, 그 순간 화면의 밝기가 화면의 모든 영역에 걸쳐 크게 증가하므로 화소 차이(pixel difference)나 히스토그램 차이(histogram difference)가 크게 나타난다. 따라서 연속된 영상임에도 불구하고 장면 전환 지점으로 잘못 검출되는 결과를 낳는다.

카메라 플래시 때문에 잘못 검출되는 것을 방지하기 위한 방법으로 프레임 사이의 차이가 큰 영상이 연속적으로 나타날 때 그것들을 모두 장면 전환이 아닌 것으로 처리하는 방법이 있으나, 이 방법으로는 실제 장면 전환 지점을 놓치는 경우가 발생한다. 보기를 들면, 컷이 발생한 영상 바로 다음 프레임에서 플래시가 터지는 경우, 컷이 발생한 직후 움직임이 큰 영상이 나타날 경우이다. 다른 방법으로 프레임 사이의 차이가 큰 영상이 연속적으로 나타날 때 그 가운데 한 개만 장면 전환으로 판단하는 방법이 있다. 앞의 방법보다 실제 장면 전환 지점을 놓치는 가능성만 줄어들 뿐이고, 여전히 검출하지 못할 논리적 가능성이 있고, 잘못 검출되는 경우도 발생한다.

이 논문에서는 이러한 문제점을 해결하기 위하여 화면의 밝기가 변하는 유형을 이용하였다. 카메라 플래시 한 개가 터지는 경우에는 화면 전체가 밝아졌다가 다음 프레임에서 어두워지면서 원래의 영상으로 되 돌아온다. 여러 개의 카메라 플래시가 한 프레임 내의 시간차를 두고 연속적으로 발생하는 경우에도 플래시가 터지는 시작점에서 화면이 밝아지고 몇 프레임 안에 본래의 영상으로 되 돌아오는 특성을 갖는다. 따라서 화면의 밝기가 일정 수준 이상 밝아지는 지점이 나타나면 이 지점을 표시해 두고, 그 이후 수 프레임 안에 화면의 밝기가 일정 수준 이상 어두워지는 장면이 나타나고, 이 장면이 밝아지기 직전의 본래의 영상과 큰 차이가 없으면, 밝아지기 시작한 지점과 어두워진 지점 사이에 있는 모든 예비 장면 전환 지점을 카메라 플래시 영상으로 판단할 수 있다. 다만 본래의 영상으로 되 돌아오기 전에 다른 장면으로 전환되는 경우에는 이 방법이 효과가 없지만 이러한 경우는 거의 없다고

볼 수 있다. 이상의 방법을 자세히 설명하면 다음과 같다.

- 1) 장면 전환 검출을 위해 순차적으로 프레임들을 비교하다가 화소의 밝기 변화(luminance difference)가 일정한 임계값(T_f)을 넘는 프레임(a)을 발견하면 이 지점을 표시해 두고, 이전 영상의 히스토그램($H_{a-1}[j]$)을 저장한다. 여기서 i 번째 프레임과 그 다음에 인접하는 프레임 사이의 화소의 밝기 변화 LD_i 는 다음과 같이 표시된다.

$$LD_i = \frac{SPD_i - SND_i}{M \times N} \times 100 \quad (6)$$

여기서, $M \times N$ 은 한 프레임에 포함된 화소의 갯수이고, SPD_i 와 SND_i 는 각각 밝기가 증가한 화소의 갯수, 밝기가 감소한 화소의 갯수를 나타낸다. SPD_i 는 다음과 같이 표시된다.

$$SPD_i = \sum_{x,y=1}^{M,N} PDP_i(x,y) \quad (7)$$

여기서 $P_i(x,y)$ 를 i 번째 프레임의 화소값, t 를 일정한 임계값이라고 하면,

$$PDP_i(x,y) = \begin{cases} 1 & \text{if } P_{i+1}(x,y) - P_i(x,y) > t \\ 0 & \text{otherwise} \end{cases}$$

마찬가지 방법으로 SND_i 는 다음과 같이 표시된다.

$$SND_i = \sum_{x,y=1}^{M,N} NDP_i(x,y) \quad (8)$$

여기서,

$$NDP_i(x,y) = \begin{cases} 1 & \text{if } P_i(x,y) - P_{i+1}(x,y) > t \\ 0 & \text{otherwise} \end{cases}$$

- 2) 검색을 계속 수행하면서 두세 프레임 안에 화소의 밝기 변화가 일정한 음의 임계값($-T_f$)보다 작은 프

레이미(b)이 나타나고, 이때 $H_{a-1}[j]$ 과 $H_b[j]$ 를 비교하여 이것이 일정 임계값보다 작다면, a 와 b 사이에 있는 모든 예비 장면 전환 지점들을 카메라 플래시 영상으로 판단한다.

3) 위의 조건을 만족하지 못하면 b 프레임 다음으로 장면 전환 검출을 계속한다.

IV. 장면 전환 검출의 통합 구현

지금까지 설명한 방법을 사용하여 첫, 점진적 변화 영상, 그리고 카메라 플래시 영상의 검출을 하나의 알고리즘으로 통합한 구조를 그림 2에 나타냈다. 압축 영역에 바탕을 둔 첫 번째 단계에서 MPEG 비트열을 분석하여 I 프레임에 대하여만 DC 계수를 추출하고 이것을 인접한 I 프레임과 서로 비교하여 그 차이가 큰 구간을 장면 전환이 예상되는 후보 구간으로 선정한다. 화소 영역에서 진행되는 두 번째 단계에서는 비디오열의 처음으로 돌아가 후보 구간에 대해서만 디코딩을 하여 화소값을 얻고, 이 값을 사용하여 정확한 장면 전환 지점을 찾기 위한 과정을 진행한다. 여기서는 히스토그램 비교법과 윤곽선 비교법을 함께 사용하였고, 윤곽선 변화를 사용하여 점진적 변화 영상을 검출하는 부분과 화소의 밝기 변화를 계산하여 카메라 플래시 영상을 검출하는 부분을 모두 통합하였다. 이러한 구조는 실제 장면 전환 검출 구현에 효과적으로 사용될 수 있다.

그림 2(b)에 보인 것과 같이 화소 영역에서는 히스토그램 차이(HD)를 계산하고, 이것을 임계값과 비교한다. 임계값으로는 것을 판별하기 위한 T_I 과, 이것보다 작은 값으로서 점진적 변화 영상을 판단하기 위한 T_s 가 사용된다. 이것은 점진적 변화 영상을 검출하기 위한 이중 비교법(twin comparison)에서 두 개의 임계값을 사용하는 방법과 유사하다[1]. 이중 비교법에서 작은 임계값은 점진적 변화가 시작되고 계속되는 것을 판단하는 기준으로 사용된다. HD 가 T_I 보다 큰 경우에는 윤곽선을 검출하고, 윤곽선 변화가 임계값보다 크면 것으로 판단한다. 아울러 화소의 밝기 변화를 계산해 두어 카메라 플래시 영상을 검출하는 데 사용한다. 다음으로 HD 가 T_s 와 T_I 사이에 있으면 점진적 변화가 시작되거나 점진적 변화 구간 안에 있는지 여부를 판단하기 위해, 윤곽선을 검출하여 윤곽선 변화의 패턴을 저장한다. HD 가 T_s 보다 작으면 장면 전환이 없는 것으로 판단하거나, 앞부분이 점진적 변화 구간이라면 그 구간이 끝나는 지점으로 판단한다. 점진적 변화

영상에 대한 판단은 그 구간의 시작과 종료를 위와 같이 판단하고 그 구간에서 얻은 윤곽선 변화의 패턴을 분석하면 된다. 그런데 점진적 변화가 일어나는 구간은 여러 GOP 구간에 걸쳐 있을 수 있으므로 점진적 변화가 시작되었는데도 불구하고 후보 구간의 끝까지 진행했을 때 점진적 변화가 종료되지 않은 것으로 판단되면, 다음 GOP 구간을 계속 확인할 필요가 있다. 실험에서는 GOP 길이가 15 프레임이고 점진적 변화가 15프레임 이상 계속되지 않는다는 가정하에 후보 구간의 바로 다음에 있는 GOP 구간까지만 디코딩을 하면서 점진적 변화 여부를 확인하도록 하였다.

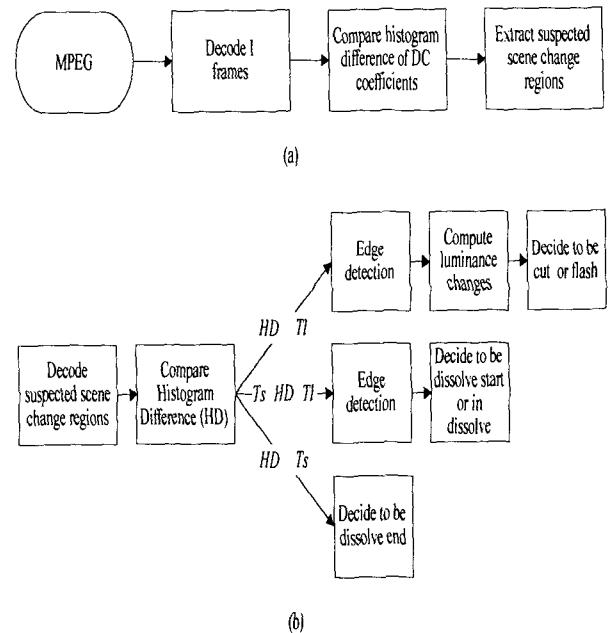


그림 2. 장면 전환 검출을 위한 통합 구조 (a) 첫 번째 단계, (b) 두 번째 단계
Fig. 2. Integrated structure for scene change detection (a) first step, (b) second step

V. 실험 결과

실험은 마이크로소프트 윈도 NT가 설치된 펜티엄 컴퓨터에서 수행되었다. KBS 뉴스, 드라마, 스포츠와 같은 동영상 파일을 MPEG-1 파일로 압축하고, 압축 영상에 대하여 장면 전환을 검출하였다. MPEG 영상의 크기는 352×240

으로 하였다. 표 1은 이 논문에서 제안한 압축 영역과 화소 영역으로 이루어진 두 단계 방법을 화소 영역만 사용한 방법과 비교한 것이다. 여기서 화소 영역만 사용한 방법이란 압축 영상을 처음부터 끝까지 디코딩하여 화소값을 얻고, 여기에 이 논문에서 제안한 화소 영역의 장면 전환 검출 알고리즘을 적용한 것이다. 제안된 결합 방법의 검출 시간(t_2)이 화소 영역만 사용한 방법의 검출 시간(t_1)보다 작다는 사실을 확인할 수 있다. 검출 속도가 향상되는 정도를 검출속도비(t_1/t_2)로 표현하였다. 여러 영상에 대해 실험한 결과 검출 속도가 평균적으로 2배 정도 향상되는 것으로 나타났다. 움직임이나 변화가 많은 영상에서는 장면 전환의 후보 구간이 많아지므로 검출 시간이 길어지게 된다. 따라서 제안된 방법은 움직임이나 변화가 적은 영상에 적용될 때 검출 속도가 향상되는 효과가 크다.

제안된 방법의 검출 성능을 알아보기 위해 실제 장면 전환 갯수, 맞게 검출한 장면 전환 갯수, 검출하지 못한 장면 전환 갯수, 잘못 검출한 장면 전환 갯수를 측정하였다. 표 2에 그 결과를 보였다. 화소 영역만 사용한 방법을 채택했을 때에도 검출 성능은 동일하게 나왔다. 따라서 압축 영역과 화소 영역의 결합을 통해 검출 성능 또는 정확성에 영향을 주지 않으면서 검출 속도를 향상시킬 수 있다. 명암만을 이용한 방법은 윤곽선을 이용한 방법보다 잘못 검출한 갯수가 많음을 알 수 있다. 이것은 명암만을 이용하면 조명의 변화나 노이즈의 영향을 많이 받기 때문이다. 명암과 윤곽선을 결합한 방법은 윤곽선만 이용한 방법과 검출 성능이 같지만 검출 속도가 더 빠른 것이다. 장면 전환을 검출하지 못한 경우는 같은 장면을 다른 카메라 각도에서 본 경우와 같이 장면 사이의 히스토그램 차이가 매우 작은 경우이다. 이런 오류를 줄이려면 히스토그램 비교에서 사용되는 임계값을 더 낮추면 되지만 그렇게 하면 검출 시간을 증가시킬 수 있으므로 적당한 수준에서 임계값을 결정해야 한다.

표 1. 검출 시간의 비교

Table 1. Comparison of detection time

동영상	화소영역만 사용한 방법 (t_1)	압축영역과 화소영역을 결합한 방법 (t_2)	검출속도비 (t_1/t_2)
뉴스(15분)	838초	368초	2.28
드라마(15분)	903초	337초	2.68
스포츠(15분)	897초	476초	1.88

표 2. 제안된 방법의 성능

Table 2. Performance of the proposed method

검출 방법	동영상	실제 장면전환 갯수	맞게 검출한 갯수	검출하지 못한 갯수	잘못 검출한 갯수
명암만을 이용한 방법	뉴스(15분)	214	212	2	15
	드라마(15분)	112	110	2	15
	스포츠(15분)	185	184	1	34
윤곽선만을 이용한 방법	뉴스(15분)	214	213	1	12
	드라마(15분)	112	110	2	10
	스포츠(15분)	185	184	1	29
명암과 윤곽선을 결합한 방법	뉴스(15분)	214	213	1	12
	드라마(15분)	112	110	2	10
	스포츠(15분)	185	184	1	29

그림 3은 카메라 플래시 영상의 보기를 나타낸다. 지도를 나타내는 본래의 영상 위에 카메라 플래시가 터졌을 때 화면 전체의 밝기가 크게 증가한 것을 알 수 있다. 뉴스 영상의 일부인 이 지도 영상은 프레임 번호 4640부터 4697까지 계속되면서 여러 대의 카메라가 임의의 순서로 지도 위에 플래시를 터뜨린다. 그림 4에 그 구간에서 프레임 사이의 히스토그램 차이를 보였다. 플래시가 터지거나 본래의 영상으로 돌아올 때, 히스토그램 차이가 매우 큰 것을 알 수 있다. 그림 5는 프레임 사이의 밝기 변화를 나타낸 것이다. 플래시가 터졌을 때에는 밝기가 크게 증가하고, 본래의 영상으로 돌아올 때에는 밝기가 크게 감소하는 것을 볼 수 있다. 이러한 특성을 이용하여 본 논문에서 설명한 카메라 플래시 검출 방법을 적용한 결과, 그 구간 안에 있는 모든 카메라 플래시를 검출할 수 있었다.



그림 3. 카메라 플래시 영상의 보기 (a) 본래의 영상, (b) 카메라 플래시 영상

Fig. 3. Integrated structure for scene change detection (a) original image, (b) camera flash image

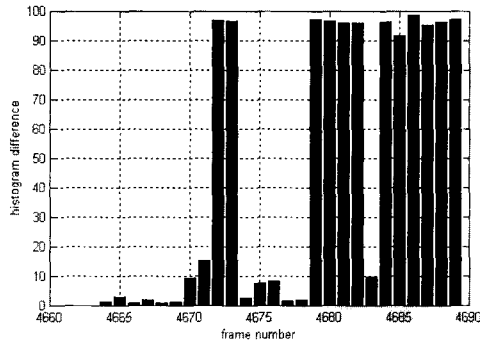


그림 4. 카메라 플래시 영상의 히스토그램 차이
Fig. 4. Histogram difference of camera flash images

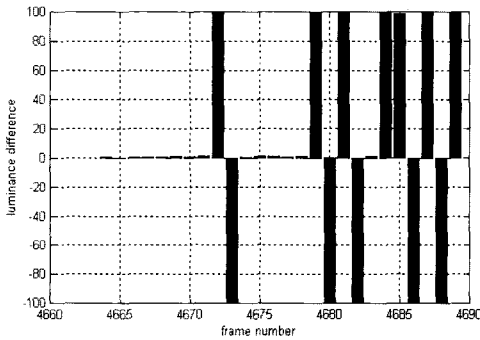


그림 5. 카메라 플래시 영상의 밝기 변화
Fig. 5. Luminance difference of camera flash images

VI. 맺음말

본 논문에서는 압축 영역과 화소 영역을 결합하여 장면 전환을 검출하는 구조를 설계하였다. 이러한 방법이 검출의 정확성을 유지하면서 검출 속도를 향상시킨다는 사실을 실험을 통해 확인하였다. 화소 영역에서는 명암 변화와 윤곽선 변화를 함께 사용하여 장면 전환을 검출하였다. 여기서 명암 변화가 일정한 값을 넘는 경우에만 윤곽선 검출을 하므로 처리 시간을 더욱 단축시킬 수 있다. 화소 영역에서 화소의 밝기 변화를 분석하여 카메라 플래시를 검출하는 방법을 소개하였다. 이 방법은 카메라 플래시가 자주 발생하는 뉴스 영상에 대해 장면 전환을 검출할 때 유용하게 사용될 수 있다. 마지막으로 컷 검출과 점진적 변화 검출, 그리고 카메라 플래시 검출을 통합하여 실제 장면 전환의 검출에 효과적으로 사용할 수 있는 구조를 소개하였다.

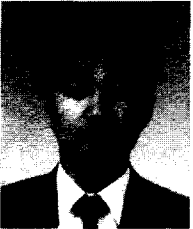
참고 문헌

- [1] H. J. Zhang, A. Kankanhalli, and S.W. Smoliar, "Automatic partitioning of full-motion video," *ACM Multimedia System*, vol. 1, no. 1, pp. 10-28, 1993.
- [2] G. Ahanger and T. D. C. Little, "A survey of technologies for parsing and indexing digital video," *J. Vis. Commun. Image Represent. (USA)*, vol. 7, no. 1, pp. 28-43, Mar. 1996.
- [3] J. Meng, Y. Juan, and S. F. Chang, "Scene change detection in a MPEG compressed video sequence," *Proc. IS&T/SPIE 2419*, pp. 14-25, Feb. 1995.
- [4] F. Arman, A. Hsu, and M. Y. Chiu, "Image processing on compressed data for large video databases," *Proceedings 1st ACM International Conference on Multimedia*, pp. 267-272, Aug. 1993.
- [5] H. J. Zhang, C. Y. Low, and S. W. Smoliar, "Video parsing and browsing using compressed data," *Multimedia Tools and Applications*, pp. 89-111, 1995.
- [6] V. Kobla and D. Doermann, "Compressed-domain video indexing techniques using DCT and motion vector information in MPEG video," *SPIE*, vol. 3022, pp. 200-211, 1997.
- [7] B. L. Yeo and B. Liu, "On the extraction of DC sequence from MPEG compressed video," *IEEE International Conference on Image Processing*, pp. 260-264, Oct. 1995.
- [8] Ramin Zabih, Justin Miller, and Kevin Mai, "A feature-based algorithm for detecting and classifying scene breaks," *Proceedings 4th ACM International Conference on Multimedia*, pp. 189-200, Nov. 1995.

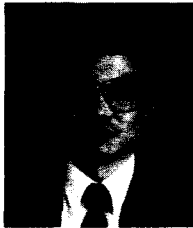
 저 자 소개

**나윤정**

1992년 2월 : 서울대학교 공과대학 전기공학과 졸업(공학사).
 1994년 2월 : 한국과학기술원 전기 및 전자공학과 졸업(공학석사).
 1994년 2월 ~ 현재 : 한국방송공사 기술연구소 연구원.
 주관심분야 : 멀티미디어 방송제작, 비디오 인덱싱, 영상 처리.

**하명환**

1993년 2월 : 경북대학교 공과대학 전자공학과 졸업(공학사).
 1995년 2월 : 한국과학기술원 전기 및 전자공학과 졸업(공학석사).
 1995년 2월 ~ 현재 : 한국방송공사 기술연구소 연구원.
 주관심분야 : 멀티미디어 방송제작, 비디오 인덱싱, 영상 처리.

**이상길**

1981년 2월 : 서울대학교 공과대학 전자공학과 졸업(공학사).
 1983년 2월 : 서울대학교 대학원 전자공학과 졸업(공학석사).
 1994년 2월 : 서울대학교 대학원 전자공학과 졸업(공학박사).
 1995년 : NHK 방송기술연구소 객원연구원.
 1983년 2월 ~ 현재 : 한국방송공사 기술연구소 근무(선임연구원).
 주관심분야 : 멀티미디어 방송, 영상 부호화 및 신호 처리.