

MPEG 비디오의 특성 추출을 이용한 효과적인 장면 전환 검출 기법

정희원 박영경*, 최윤석**, 고성제**

Effective scene change detection methods using characteristics of MPEG video

Young-Gyoung Kwak*, Yoon-Seok Choi*, and Sung-Jea Ko** *Regular Members*

요 약

본 논문에서는 압축되어 있는 MPEG 비디오 시퀀스로부터 DCT의 AC 계수를 이용한 에지 영상을 구하여 급격한 장면의 전환을 검출하는 방법과 매크로블록 타입 정보를 이용하여 점진적 장면 전환인 디졸브(dissolve) 구간을 검출하는 기법을 제안하였다. 에지 추출에 기반한 장면 전환 검출 기법은 밝기의 변화에 덜 민감하며, AC 성분을 이용하면 DC 성분을 이용한 경우보다 원영상을 더욱 잘 표현하는 에지를 추출할 수 있으므로, 보다 정확한 장면 전환을 검출할 수 있다. 제안한 디졸브 검출 기법에서는 영상을 복원하지 않고, MPEG 비트스트림 내의 매크로블록 타입 정보로부터 계산된 인트라 매크로 블록의 개수를 이용하기 때문에 적은 계산량으로 디졸브를 검출할 수 있다. 제안한 장면 전환 검출 방법은 기존의 방법들에 비해서 성능이 우수함을 실험을 통해 입증하였다.

ABSTRACT

In this paper, we propose new methods to detect a scene cut and a dissolve region in compressed MPEG video sequences. The scene cut detection method uses edge images obtained using DCT AC coefficients and the dissolve detection method utilizes the macroblock type information of the MPEG stream. The proposed scene cut detection method is insensitive to brightness and can detect scene changes more precisely than the methods using DC coefficients since AC edge images can express original images more exactly than DC edge images do. The proposed dissolve detection method using the number of intra macroblocks(MBs) computationally efficient since it does not require the decoding process.

Experimental results show that the proposed methods perform better in detection scene changes than conventional other methods.

1. 서 론

디지털화된 비디오 정보를 데이터베이스로 구축하고자 할 때에 선행되어야 할 것은 한 편의 비디오를 내용별로 인덱싱하여 사용자가 원하는 비디오 정보에 대해 비순차적으로 접근할 수 있게 하고 내용 기반 탐색을 가능하게 하는 것이다. 한 편의 비

디오를 내용별로 인덱싱하기 위해서는 해당 비디오 데이터를 장면의 내용이 바뀌는 시점 단위로 분할하게 된다. 따라서 장면 전환 검출은 비디오 정보의 데이터베이스화에 있어서 필수적으로 요구되는 기술이다.

장면 전환이란 한 대의 카메라가 연속적으로 찍은 장면이 끝나고 다른 장면으로 바뀌는 것을 의미

* LG정보통신 전송연구소(hikyg@lgic.co.kr),

** LG정보통신 차세대통신연구소(can12@lgic.co.kr)

*** 고려대학교 전자공학과 (sjko@dali.korea.ac.kr)

논문번호 : 98495-1121, 접수일자 : 1998년 11월 21일

한다. 장면 전환의 종류는 크게 이웃한 두 프레임 내에서 장면 전환이 일어나는 급격한 장면 전환과 여러 프레임에 걸쳐서 일어나는 점진적인 장면 전환의 두 가지로 나뉘어진다.

급격한 장면 전환의 검출에 대해서는 특히 비압축 상태보다는 압축 상태에서의 연구가 활발하게 이루어져 왔는데^{[1]-[5]}, MPEG 신덱스에 포함된 매크로블록 타입(macroblocck type)을 이용하는 방법^[1], DCT의 DC 성분을 이용하여 MAD나 히스토그램 등 픽셀값을 이용하는 방법^{[5]-[7]}, DC 성분에서 추출된 에지를 이용하는 방법 등이 있다^[8]. 그러나 이러한 방법들 중, 픽셀값에 바탕을 둔 방법은 조명의 변화에 민감하고, 에지를 이용하는 방법은 뚜렷한 에지 검출이 어렵기 때문에 정확한 장면 전환을 검출하기가 어렵다.

점진적인 장면 전환의 종류에는 두 화면이 서로 겹쳐지면서 다음 화면으로 바뀌는 디졸브(dissolve), 화면이 점점 어두워져서 완전히 사라진 후에 다른 화면이 나타나는 페이드 인/아웃(fade-in/out), 그리고 새로운 화면이 옆에서부터 나타나는 와이프(wipe) 등으로 분류할 수 있다. 이들 중 디졸브는 점진적인 장면 전환의 대표적인 예라고 할 수 있으며, 이에 대한 연구도 최근들어 더욱 활발히 이루어지고 있다. Zabih는 복원 영상의 에지 픽셀의 분포를 이용하여 디졸브, 페이드 및 와이프를 검출하였으며^[9], Meng은 MPEG 스트림의 DC 영상으로부터 분산값을 구하여 U자형이 나타나는 구간으로 디졸브를 검출하였다^[11]. 그러나 Zabih의 방법은 에지를 찾기 위하여 영상을 모두 복원하는 과정에서, 그리고 Meng의 방법은 분산값을 구하는 과정에서 계산량이 많아지게 된다.

이러한 각각의 문제점들을 해결하기 위하여 본 논문에서는 먼저 급격한 장면 전환을 효과적으로 검출할 수 있는 방법을 제안하였다. I 프레임에서는 주요 에지와 관련 있는 DCT의 저주파의 5개 AC 계수를 이용하고, P, B 프레임에서는 예측을 이용하여 에지 영상을 만든 다음 에지 히스토그램 분포를 이용하여 급격한 장면 전환을 검출한다. 제안한 방법은 명확한 에지를 추출할 수 있으므로, 장면 전환 검출에도 더 좋은 성능을 나타낸다.

두 번째로 점진적인 장면 전환 검출 가운데 특히 디졸브 전환이 일어나는 구간에서의 특징을 살펴보고 이를 빠른 시간 내에 검출하는 기법을 제안한다. 디졸브 구간에서는 새로운 장면이 겹쳐질 때 에지의 수가 증가하고 이에 따라 양자화된 DCT 계수

중 0이 아닌 AC 계수의 수가 증가하며, P 프레임의 경우에는 인트라 부호화되는 매크로블록의 수가 많아지게 된다. 이러한 성질을 이용하여 P 프레임의 인트라 부호화된 매크로블록의 수를 비교하여 디졸브 구간을 검출할 수 있다. 또한 팬이나 줌과 같은 빠른 카메라 움직임에 대한 영향을 고려하므로 정확한 디졸브 구간 검출을 할 수 있게 된다.

본 논문의 구성은 다음과 같다. II장에서는 급격한 장면 전환 검출 방법을 III장에서는 점진적인 장면 전환 검출 방법을 제안하였다. IV장에서는 제안된 방법에 의하여 실험한 결과를 보여주고 이에 대한 고찰을 하며 V장에서 결론을 맺는다.

II. 급격한 장면 전환의 검출

기존의 급격한 장면 전환을 검출하는 기법들 가운데에서 픽셀값에 기반하는 방식은 조명에 민감하여 장면 전환이 일어나지 않은 경우에도 장면 전환으로 잘못 판단하는 경우가 있으며, 형태 기반의 에지 추출 방식 중 DCT의 DC 계수를 이용하는 경우에는 DC 영상을 구하는 과정에서 영상이 블러링되므로 정확한 에지 추출이 어렵다는 문제가 있었다. 이 장에서는 MPEG 비디오의 DCT 계수 중 AC 계수를 이용하여 에지 영상을 구하고 장면 전환을 검출하는 방법을 제안한다.

1. AC 계수를 이용한 에지 정보 추출

MPEG 비디오 시퀀스에서 DCT의 AC 계수는 영상 내에 포함된 공간 주파수 성분의 크기와 비례한다. 공간 주파수는 영상의 복잡도와 관련이 있으므로 AC 계수를 이용하면 각 블록에서의 복잡도를 구할 수 있고 이를 이용하여 에지 영상을 얻을 수 있다. MPEG 비디오 시퀀스로부터 에지 영상을 얻기 위해서 인트라 부호화된 I 프레임에서는 DCT의 AC 계수를 이용하여 영상의 복잡도를 구하고 움직임 예측 부호화된 P와 B 프레임에서는 움직임 보상을 통하여 영상의 복잡도를 예측하여야 한다.

1.1 I 프레임에서의 에지 정보 추출

MPEG 비디오의 I 프레임은 DCT에 의해 8×8 픽셀 단위로 블록 내 부호화되어 있다. 이 때 DCT 계수는 해당 블록 내에 포함되어 있는 공간 주파수 성분의 크기를 나타낸다.

공간 주파수 성분은 그림 1과 같은 기저 영상들로 표현될 수 있는데, 이에 대한 DCT 계수들 중에

서 $F_{0i}, F_{10}, F_{02}, F_{20}, F_{11}$ 은 저주파에 해당하는 계수들로 시간적으로 볼 때에 블록 내에서 눈에 띄게 변하는 수직, 수평, 대각선 방향의 에지 성분을 나타낸다.

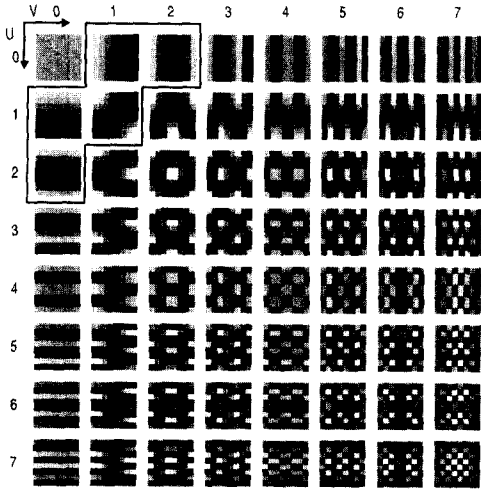


그림 1. DCT 변환의 기저 영상 (N=8).

반면에 이들 이외의 AC 계수들은 그 복잡도가 커서 블록 내의 외곽선 에지보다는 텍스처에 관련되는 경우가 더 많다. 그림 1의 좌측 상단의 표시된 5개 이미지를 보면 블록 내를 통과하는 수직, 수평, 대각선의 강한 에지 성분을 잘 나타내고 있음을 알 수 있다. F_{01}, F_{10} 은 블록을 이등분하는 수평, 수직 에지들, F_{20}, F_{02} 는 블록의 중심을 굽게 가로지르는 수평, 수직 에지를 나타내며, F_{11} 은 블록 내의 대각선 성분을 나타낸다. 특히 대각선 성분을 나타내는 F_{11} 을 포함하여 본 논문에서는 저주파에 해당하는 이들 5개의 AC 계수를 이용하였다^[10]. 이는 또한 모든 AC 성분을 복호하지 않아도 되어 시간적으로도 더 효율적이다.

블록 내에 포함된 에지 성분의 정도는 AC 계수들의 제곱의 합인 에지 강도(edge power)

$P_E = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F_{uv}^2 - F_{00}^2 = P_L + P_H$ 로 나타낼 수 있으며, 저주파 에지 강도 P_L 은 다음과 같이 5개의 AC 계수의 제곱의 합으로 나타낼 수 있다.

$$P_L = F_{01}^2 + F_{10}^2 + F_{02}^2 + F_{20}^2 + F_{11}^2 \quad (1)$$

제안된 에지 검출 방법에서, I 프레임에서는 각 블록마다의 저주파 에지 강도를 구할 수 있다. 원 영상의 각 블록에서 블록 내에 주요한 에지 성분이

많이 존재하면 그 블록은 큰 에지 강도의 값을 갖게 되며 이 에지 강도 영상에 임계값 T를 적용하면 원 영상에서 $\frac{1}{8}$ 로 축소된 크기의 이진 에지 영상을 얻을 수 있다.

1.2 P와 B 프레임에서의 에지 정보 추출

블록 내 부호화되는 I 프레임과는 달리 P와 B 프레임에서는 움직임 예측을 이용하기 때문에 에지 강도 역시 참조 프레임으로부터 예측을 통하여 추출하여야 한다.

가장 간단한 방법으로 움직임 예측 부호화된 부분의 값을 4개의 참조 블록 중 가장 많이 겹쳐진 블록의 값으로 대체하는 방법이 있다. 그러나 겹쳐진 네 블록 모두를 고려하는 방법으로 좀 더 정확한 에지 강도를 검출할 수 있다. Parseval의 정리에 의하여 $P_E = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} F_{uv}^2 - F_{00}^2 = N^2 \sigma_E^2$ 이므로, 4개의 참조 블록 모두를 고려한 경우의 예측된 에지 강도 \tilde{P}_E 는 분산값 예측을 통하여 구할 수 있으며, 참조 블록들의 평균값과 분산값 및 현재 블록과 겹치는 부분의 면적을 이용하여 계산할 수 있다^[11].

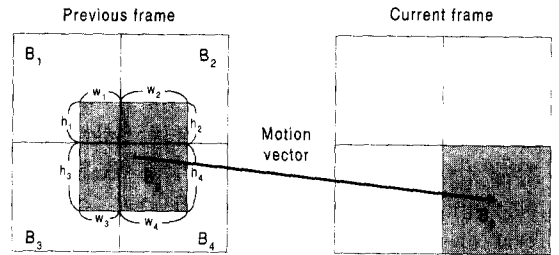


그림 2. 분산값 예측.

그림 2에서 각각의 해당 블록의 평균과 분산이 각각 $E[B_i] = \mu_i, \text{Var}[B_i] = \sigma_i^2, i=1,2,3,4$ 이고 움직임 벡터에 의해 참조되는 블록 B_i 와 겹치는 부분의 면적이 $h_i w_i$ 라 할 때, 예측된 에지 강도 \tilde{P}_E 는 다음과 같다^[12].

$$\begin{aligned} \tilde{P}_E &= 64 \text{Var}[\tilde{P}_E] \approx 64 \left[\sum_{i=1}^4 \frac{h_i w_i}{64} (\sigma_i^2 + \mu_i^2) - \left(\sum_{i=1}^4 \frac{h_i w_i}{64} \mu_i \right)^2 \right] \\ &= \sum_{i=1}^4 h_i w_i \left(\frac{P_i + P_H}{64} + \mu_i^2 \right) - 64 \left(\sum_{i=1}^4 \frac{h_i w_i}{64} \mu_i \right)^2 \\ &= \left[\sum_{i=1}^4 \frac{h_i w_i}{64} P_i \right] + \left[\sum_{i=1}^4 \frac{h_i w_i}{64} P_H \right] + \left[\sum_{i=1}^4 (h_i w_i \mu_i^2) - 64 \left(\sum_{i=1}^4 \frac{h_i w_i}{64} \mu_i \right)^2 \right] \\ &= \tilde{P}_L + \tilde{P}_H + P_H \end{aligned} \quad (2)$$

여기에서 \tilde{P}_L 과 \tilde{P}_H 는 각각 저주파 에지 강도와 고주파 에지 강도에 해당하며, P_L, P_H 는 B_i 블록에서의 저주파 에지 강도와 고주파 에지 강도이며 P_M 은 움직임 예측 부호화된 블록 내에 포함되는 참조 블록 B_i 들 사이의 평균값의 차이에 의해 생기는 에지 성분의 강도이다.

움직임 예측 부호화된 블록의 주요한 에지 성분의 크기는 식 (3)로부터 저주파 에지 강도 \tilde{P}_L 과 블록 간의 에지 강도 P_M 를 더하여 얻을 수 있으며, 이러한 방법으로 에지 강도 영상을 구할 수가 있게 된다.

$$\begin{aligned} \tilde{P}_L &= \sum_{i=1}^4 \frac{h_i w_i}{64} P_{L_i} \\ \tilde{P}_H &= \sum_{i=1}^4 \frac{h_i w_i}{64} P_{H_i} \\ P_M &= \sum_{i=1}^4 (h_i w_i \mu_i^2) - 64 \left(\sum_{i=1}^4 \frac{h_i w_i}{64} \mu_i \right)^2 \end{aligned} \quad (3)$$

또한 이 에지 강도 영상에 I 프레임에서와 같은 임계값 T 를 이용하여 에지 블록을 결정하면 이로부터 에지 영상을 구할 수 있다.

2. 에지 영상을 이용한 장면 전환 검출

본 절에서는 AC 계수와 분산값 예측을 통하여 얻은 에지 영상을 이용하여 장면 전환을 검출하는 방법에 대해 설명한다. 에지 영상을 이용하여 장면 전환 검출을 위한 인접 프레임간의 유사도를 측정하는 방법으로, 움직임에 덜 민감한 수평, 수직 방향의 히스토그램을 이용한 방법인 에지 히스토그램 비교 방법을 사용하였으며, 움직임이 큰 영상에 대해서도 효과적으로 장면 전환을 검출할 수 있도록 움직임 보상을 하였다. 그림 3은 수평, 수직 방향의 히스토그램을 구하는 과정과 움직임 보상을 하는 과정을 나타내고 있다.

그림 3에서 수평 방향의 히스토그램인 P_H 는 현재 프레임의 0부터 $M-1$ 행까지 세로 방향으로의 에지 블록의 개수를 나타내는 히스토그램이고, 수직 방향의 히스토그램인 P_V 는 현재 프레임의 0부터 $N-1$ 행까지 가로 방향으로의 에지 블록의 개수를 나타내는 히스토그램이다.

또한 P_H, P_V 는 각각 참조 프레임에서의 수평, 수직 방향의 히스토그램이다. 현재 프레임과 참조 프레임의 유사도를 측정하기 위해서 움직임 벡터

(u, v) 를 구한다.

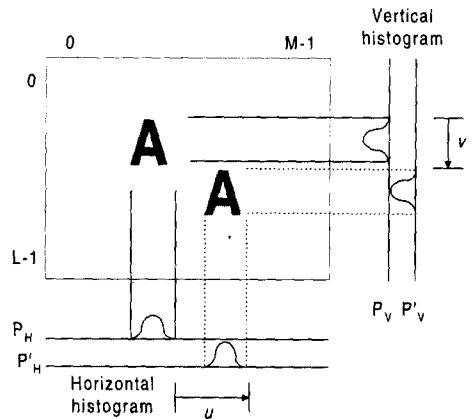


그림 3. 방향성 에지 히스토그램을 이용한 움직임 보상.

따라서 움직임 보상된 두 프레임의 수평, 수직 방향의 유사도 S 는 식 (4)와 같이 에지 히스토그램의 차로 나타낼 수 있다.

$$\begin{aligned} S^{-1} &= C_H(u) + C_V(v) \\ &= \sum_{x=0}^{M-1} |P'_H(x+u) - P_H(x)| + \sum_{y=0}^{N-1} |P'_V(y+v) - P_V(y)| \end{aligned} \quad (4)$$

유사도 S 가 크면 두 프레임은 비슷한 장면이므로 두 프레임 사이에 장면 전환이 발생하지 않고, 유사도 S 가 작으면 현재 프레임에서 장면 전환이 발생한다고 볼 수 있다.

III. 점진적인 장면 전환의 검출

기존의 디졸브를 검출하는 기법은 복원 상태에서 에지를 구하여 계산하거나 DC 영상에서 분산을 이용하여 U자 구간을 구하는 방법을 주로 사용하였다. 그러나 이러한 방법들은 완전 복원을 하거나 분산을 구하는 경우에 많은 계산량을 필요로 하므로 시간적 효율이 떨어지게 된다. 본 논문에서는 MPEG 비디오 스트림에서 완전 복원을 거치지 않고 빠른 계산 과정을 거쳐서 디졸브 구간을 찾는 방법을 제안하였다^[13].

1. 디졸브 전환과 DCT 계수와와의 관계

디졸브 구간에서의 프레임의 관계는 일반적으로 식 (5)와 같은 형태를 띠게 된다.

$$f(t) = f_1(t)[1 - a(t)] + f_2(t)a(t) \quad (5)$$

여기에서 $f_1(t)$ 와 $f_2(t)$ 는 t_1 부터 t_2 까지의 디졸브 구간 앞뒤의 프레임 픽셀값이며, $a(t) = (t - t_1) / (t_2 - t_1)$ 는 화면의 구성비에 따라 0에서부터 1까지 단조 증가하는 값이 된다. 그림 4는 4x4의 간단한 영상의 디졸브 전환을 나타낸 것이다.



그림 4. 간단한 디졸브 전환.

디졸브 구간 앞뒤의 프레임들은 에지의 수가 적고 뚜렷한 반면에 디졸브 구간 내에서는 점차 새로운 에지가 나타나면서 흐려지는 것을 볼 수가 있다. 즉 이러한 특징으로부터 영상을 DCT로 변환하였을 경우 미치는 영향을 생각해 볼 수가 있게 된다. 아래 그림 5는 위의 그림을 4x4 DCT를 취한 결과의 +/- 부호를 나타낸 것이다. 각 프레임의 좌측 상단의 빗금 부분은 DC 계수이며, 그 나머지 중 부호 표시가 되지 않은 부분은 0의 값을 가지는 것을 뜻한다.

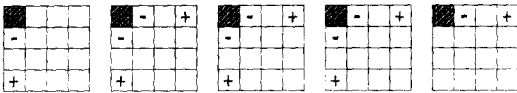


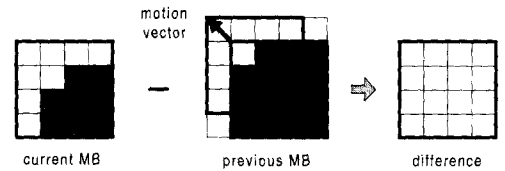
그림 5. 디졸브 영상의 DCT 결과.

그림 5의 양 끝에 있는 프레임에서는 0이 아닌 DCT의 AC 계수가 2개였지만, 디졸브 구간에서는 4개로 늘어나는 것을 알 수 있다. 즉 디졸브 구간에서는 에지의 수가 많아지게 되고 에지 수가 늘어나면 0이 아닌 DCT의 AC 계수를 갖는 경우의 수가 많아지게 된다. 따라서 일정 구간동안 0이 아닌 DCT의 AC 계수의 수가 많아지는 구역을 디졸브로 검출할 수 있게 된다. 따라서 기존의 DC 영상으로부터 분산값을 계산하는 것보다 훨씬 빠른 시간 내에 수행할 수 있게 된다.

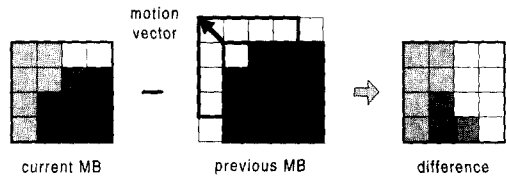
그러나 MPEG 스트림에서 P 프레임은 예측을 이용하여 inter 부호화 되므로 또 다른 알고리즘이 필요하게 된다. P 프레임에서는 움직임 벡터를 이용하여 움직임 보상된 매크로블록과 현재 자리에 위치한 매크로블록의 차영상만을 DCT 연산을 하게 된다. 장면 전환이 일어나지 않는 인접 프레임 간에는 현재 매크로블록과 움직임 보상된 참고 매크로블록

간의 차영상은 거의 0에 가까운 픽셀로 구성된다. 그러나 디졸브 구간의 인접 프레임에서는 새로 나타나는 영상이나 없어지는 영상에 의하여 두 매크로블록 간의 픽셀값이 완전히 일치하는 것이 아니라 전반적으로 조금씩 다른 색깔을 띠게 되므로 차영상은 희미하고 복잡한 형태의 영상이 된다. 그림 6은 이 두 가지 경우의 차영상의 예이다.

즉 디졸브 구간에서는 장면 전환이 일어나지 않는 경우에 비하여 많은 에지가 존재하므로 이에 대한 DCT 부호화 양이 많아지게 된다. 따라서 부호화 시에 inter 부호화 보다는 인트라 부호화 하는 경우가 많이 생겨나게 된다.



(a)



(b)

그림 6. 현재 매크로블록과 움직임 보상된 이전 매크로블록의 차영상, (a) 장면 전환이 일어나지 않는 경우, (b) 디졸브 전환이 일어나는 경우.

2. 제안한 디졸브 검출 기법

MPEG 스트림은 대부분 디스플레이 순서로 IBBPBBPBBIBBPBBPBB의 형태로 I, B, P 프레임이 모두 사용되어 부호화 되어진다. 여기에서 디졸브는 여러 프레임에 걸쳐서 나타나기 때문에 모든 프레임 하나하나를 조사하는 것이 아니라, 몇 개의 프레임만을 조사해도 디졸브 구간이 나타나는 곳을 알 수 있다. 디졸브 구간은 정확하게 시작하고 끝나는 부분을 한 프레임의 오차도 없이 가리는 것이 중요한 것이 아니라 디졸브 구간 자체를 찾아가 못 찾는가가 더 중요한 문제이다. 따라서 본 논문에서는 MPEG 스트림에 일정 간격으로 자주 등장하는 P 프레임을 이용하여 디졸브를 검출하는 방법을 제안하였다.

앞 절에서 설명한 바와 같이 디졸브 구간에서는 장면 전환이 없는 구간과는 달리 움직임 보상된 참고 프레임과 현재 프레임 간의 차 영상은 희미하고 복잡한 에지가 섞인 형태의 값들이 나온다. 이러한 값들은 DCT 부호화량을 증가시키므로 인트라 부호화되는 경향이 커지게 된다. 따라서 P 프레임 내의 인트라 매크로 블록의 수가 많아지는 부분을 디졸브 구간으로 검출할 수 있다.

P 프레임 내에서 인트라 부호화되는 매크로 블록의 수는 카메라의 움직임에 의해서 증가하기도 한다. 카메라가 옆으로 움직이는 경우(pan)나 점점 멀어지는 경우(zoom out)에는 이전의 화면에 없었던 부분이 나타나게 되므로 영상의 가장자리에 있는 매크로블록들은 인트라 부호화되는 경우가 발생한다. 이와 같은 매크로블록들의 영향을 배제하기 위하여 맨 가장자리를 제외한 곳에 위치한 매크로블록들을 대졸브 검출에 이용한다. 그림 7에서 회색으로 표시된 부분이 디졸브를 검출하기 위해서 이용한 매크로 블록들을 나타낸다.

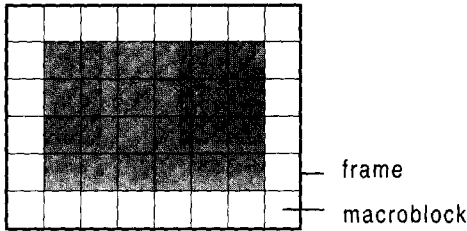


그림 7. 인트라 매크로블록의 수를 세는 구역 (회색부분).

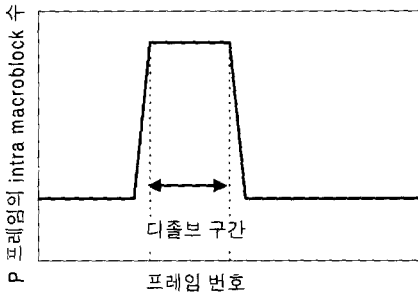


그림 8. P 프레임 내의 인트라 매크로블록의 수와 디졸브 구간과의 관계.

그림 8에서 볼 수 있는 바와 같이 P 프레임 내의 인트라 매크로블록은 디졸브 구간이 아닌 경우보다 디졸브 구간에서 월등히 많이 나타난다. 따라서 P 프레임 내의 인트라 매크로블록이 지속적으로

많이 나타나는 구간을 검색함으로써 디졸브를 효과적으로 검출할 수 있다.

DC 영상에서 분산을 이용하여 디졸브 구간을 검출하는 기존의 방법에 비하여 제안한 방법은 압축 스트림을 파싱하면서 단순히 매크로블록의 타입에 따라 수를 세어 나가는 것이므로 아주 빠른 시간에 디졸브 구간을 검출할 수 있다. 또한 P 프레임만을 이용하기 때문에 I 프레임이나 B 프레임의 파싱 시간도 상당히 단축되는 것을 알 수 있다.

IV. 실험 결과 및 고찰

본 장에서는 급격한 장면 전환 및 점진적인 장면 전환 검출의 두 가지 실험을 하고 제안된 기법과 기존의 기법을 비교 평가하였다.

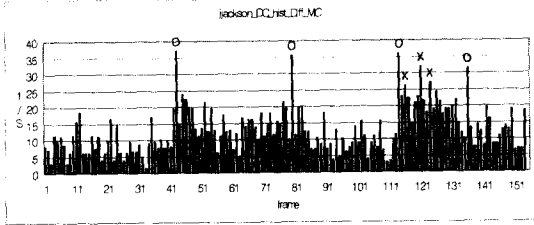
1. 급격한 장면 전환 검출 실험

본 절에서는 제안한 AC 계수로부터 구한 에지 영상과 기존의 DC 계수로부터 구한 에지 영상[8]을 이용하여 장면 전환을 검출한 결과를 비교함으로써 제안된 기법을 평가하였다. 장면 전환의 검출의 임계값으로는 가장 간단한 방법인 결과값 중에서 최대의 차이를 나타내는 결과값의 절반에 해당하는 값으로 정하였다. DC 영상의 임계값은 MPEG 신택스 중 DC 계수 차의 비트 수를 나타내는 dct_dc_size 3을 경계로 한 것과 같도록 7로 하였으며[8], 또한 이 값은 실험적으로 에지가 가장 잘 나타나는 것으로 결정된 것이다. 장면의 전환 시점은 연속되는 두 장면이 급격하게 달라지는 시점으로 정하였으며 실험 영상은 움직임과 조명의 변화가 많은 *airwolf.mpg*(412 프레임, 24×18), *Debbie-Harry.mpg*(100 프레임, 20×16), 그리고 *jjackson.mpg*(155 프레임, 20×16)을 사용하였다.

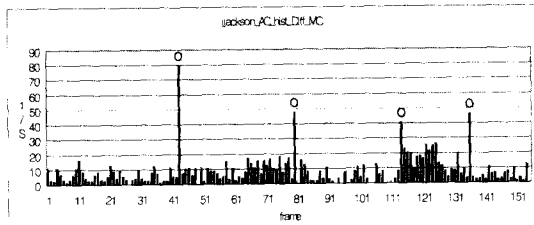
그림 9는 *jjackson.mpg*, *DebbieHarry.mpg* 및 *airwolf.mpg* 영상에 대해 각각 DC 계수와 AC 계수를 사용하여 구한 에지 영상의 프레임간의 유사도 S의 역을 취한 결과이다. 이 때의 결과값이 임계값을 넘으면 장면 전환을 검출하였으며, 장면 전환을 빠르게 찾았는지는 ○, ×로 표시하고, 찾지 못한 장면 전환 부분에 대해서는 *로 표시하였으며, 또한 그림 9의 장면 전환을 검출한 결과를 표 1에 정리하여 나타내었다.

그림 9의 실제로 장면 전환이 일어난 부분에서의 $\frac{1}{S}$ 의 값을 보면 장면 전환 부분에서 (a)보다 (b)에

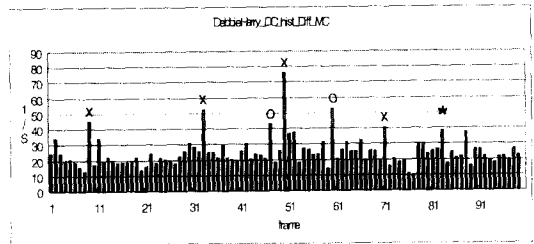
서 더 뚜렷한 차이를 나타낸다. 즉, AC 에지 영상이 보다 원 영상의 성질을 더욱 잘 나타낸다고 할 수 있다.



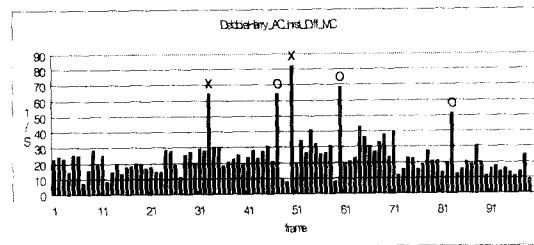
(a)



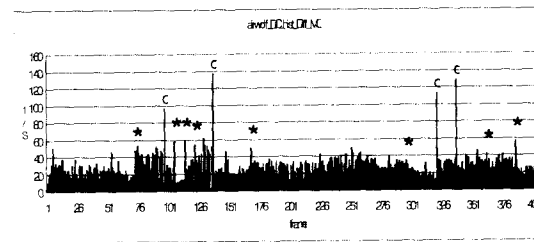
(b)



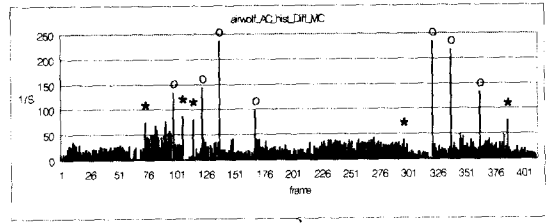
(c)



(d)



(e)



(f)

그림 9. 에지 영상을 이용한 히스토그램의 차,

- (a) jackson.mpg DC 에지,
- (b) jackson.mpg AC 에지,
- (c) DebbieHarry.mpg DC 에지,
- (d) DebbieHarry.mpg AC 에지,
- (e) airwolf.mpg DC 에지,
- (f) airwolf.mpg AC 에지

표 1을 보면 DC 계수 및 AC 계수를 이용하여 장면 전환을 검출한 결과, AC 에지 영상을 이용한 경우에 있어서 장면 전환을 놓친 경우와 잘못 검출한 장면 전환의 수가 적다. 즉, AC 계수를 이용한 방법이 장면전환을 더 정확하게 찾는다는 것을 알 수 있다. DC 에지 영상으로 실험한 경우, jackson 영상에서는 빠른 동작에 의한 빛의 변화로 인해서 장면 전환을 잘못 검출하였으며, DebbieHarry 영상에서는 플래쉬 부분에서 잘못 검출하였고 또한 어두운 장면에서 검출하지 못한 장면도 있었다. airwolf 영상에서는 전체적으로 평활한 영상에서 아주 복잡한 영상으로 바뀌는 부분에서의 장면 전환을 검출하지 못하였다.

표 1. DC 에지 및 AC 에지 영상을 이용하여 장면 전환 검출을 한 결과.

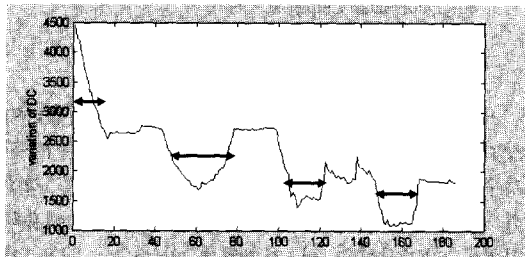
실험 영상	실제 장면 전환	사용한 에지 영상	바르게 검출한 장면 전환	잘못 검출한 장면 전환	검출하지 못한 장면 전환
jackson	4	DC	4	3	0
		AC	4	0	0
Debbie Harry	3	DC	2	4	1
		AC	3	2	0
airwolf	12	DC	4	0	8
		AC	7	0	5
계	19	DC	10	7	9
		AC	14	2	5

2. 점진적인 장면 전환 검출 실험

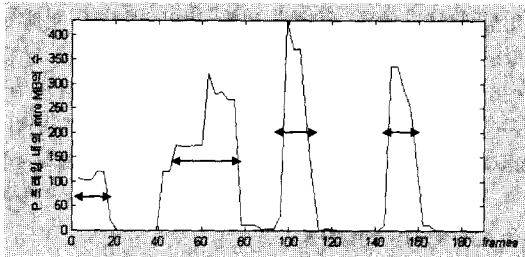
본 절에서는 제안된 기법과 DC 영상에서 분산값을 이용하는 기존의 기법을 점진적인 장면 전환 검출에 적용하여 비교 평가하였다. 점진적인 장면 전환 검출 실험은 디폴드가 포함되어 있는 beach.mpg (186 프레임, 352×240), autumn2.mpg(446 프레임,

352×240), sunrise.mpg(590 프레임, 352×240), clapton.mpg(151 프레임, 160×120)의 4가지 영상 시퀀스를 이용하였다. 각 영상 시퀀스의 GOP 구조는 clapton.mpg의 경우에는 IBBPBB와 같이 6개 프레임으로 구성되며, 나머지 세 시퀀스는 IBBPBBPBBPBBPBB 형태의 15개 프레임으로 구성되어 있다.

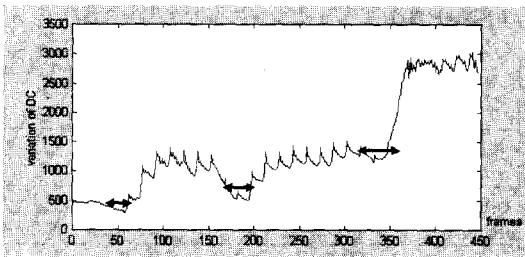
다음 그림 10은 각 영상에서 디졸브 구간을 분산을 이용하는 방법과 제안된 방법에 의하여 디졸브 구간을 검출한 결과이다.



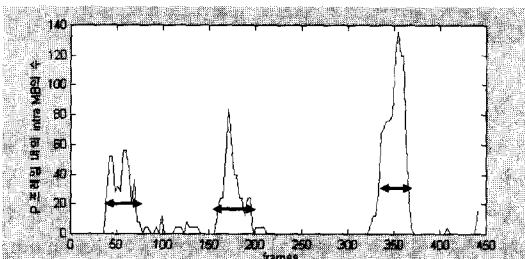
(a)



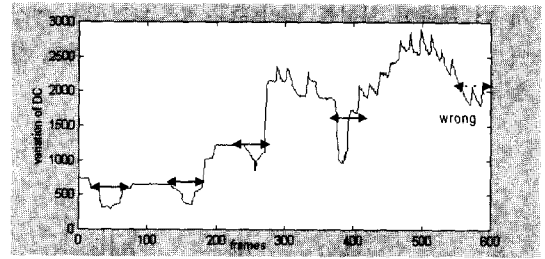
(b)



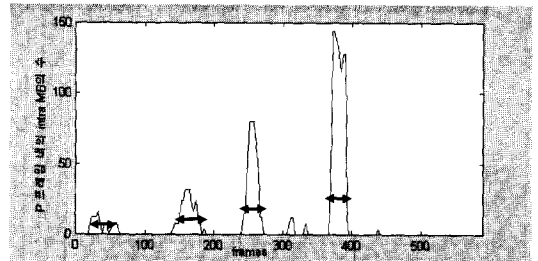
(c)



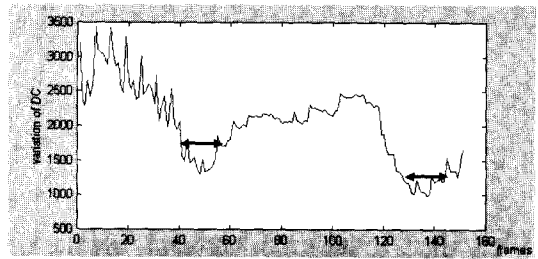
(d)



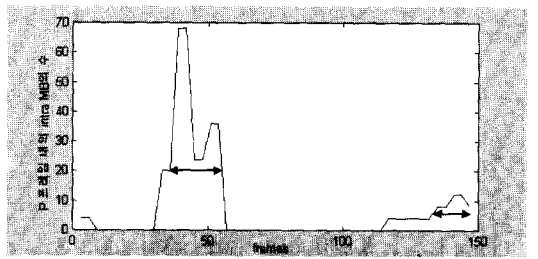
(e)



(f)



(g)



(h)

그림 10. 각 영상의 디졸브 구간 검출
 (a) 분산을 이용한 기법(beach.mpg),
 (b) 제안된 기법(beach.mpg),
 (c) 분산을 이용한 기법(autumn2.mpg),
 (d) 제안된 기법(autumn2.mpg),
 (e) 분산을 이용한 기법(sunrise.mpg),
 (f) 제안된 기법(sunrise.mpg),
 (g) 분산을 이용한 기법(clapton.mpg),
 (h) 제안된 기법(clapton.mpg).

분산을 이용한 경우에는 U자형이 뚜렷하게 나타나지 않아 디졸브를 검출하기 어려운 부분이 많았지만 제안한 기법에서는 디졸브 구간에서의 U자형

고원 부분이 명확히 나타났다. 특히 그림 10. (c)의 autumn2.mpg 영상에서는 분산을 이용한 방법에서 U자 구간을 구분하기가 어려웠다. 그리고, 그림 10.(e)의 sunrise.mpg 영상에서는 분산을 이용한 경우에 디졸브 구간을 잘못 찾은 부분이 있는데, 이 부분은 화면이 서서히 멀어지는 장면(zoom out)으로서 처음에 없었던 새로운 장면이 화면의 가장자리에 나타나면서 DC 영상 내의 분산값에 영상을 미쳤기 때문에 잘못된 결과를 초래하였다. 그러나 제안된 방법에서는 가장자리의 매크로블록을 계산하지 않았기 때문에 그 부분을 디졸브로 잘못 인식하는 일이 발생하지 않았다. 제안된 기법은 네 가지 실험 영상에서 디졸브 구간을 모두 제대로 찾았다는 것을 확인할 수 있었다. 기존의 디졸브 검출 방법에서 DC 영상 추출 및 분산값 계산에서 많은 계산량이 요구되었던 것에 반해 실험에서 제안된 기법은 MPEG 스트림 내에 포함된 형식인 macro-block type을 파싱하는 적은 계산량으로 네 가지 영상의 모든 디졸브를 바르게 검출하였으며, 또한 검출된 디졸브 구간은 뚜렷한 고원 구간을 가지므로 기존의 DC 영상의 분산값을 이용한 검출 방법보다 명료한 구간 검출이 가능하였다.

V. 결론

기존의 급격한 장면 전환 검출 알고리즘인 DC 계수 또는 움직임 벡터의 타입을 이용하여 장면 전환을 검출하는 방법은 각 비디오 프레임 내에 포함되어 있는 비슷한 배경 하에서의 형태 변화에 의한 장면 전환을 검출하기가 어렵고 조명 변화에 너무 민감하다는 문제점이 있었다. 또한 디졸브 검출 알고리즘 중에서 영상을 모두 복원하여 디졸브 전환을 검출하는 방법이나 DC 영상에서 분산을 구하여 디졸브 전환을 검출하는 기존의 방법들은 시간적으로 비효율적일 뿐만 아니라, 팬이나 줌과 같은 카메라 동작이 있을 경우에 잘못된 검출을 하는 경우가 있다. 따라서 이러한 문제점을 해결하기 위해서 본 논문에서는 MPEG 비디오로부터 AC 계수를 이용하여 에지를 추출하여 급격한 장면 전환을 검출하는 방법과 압축 상태에서 단순히 P 프레임의 매크로블록 타입만을 이용하여 디졸브 구간을 검출하는 방법을 제안하였다.

제안한 급격한 장면 전환 검출 방법은 블록 내의 복잡도를 이용하므로 기존의 DC 계수만을 이용하여 에지를 구하는 방법의 문제점인 블러링 현상이

나타나지 않아서 보다 더 명료하며, 장면 전환 검출에 더 효과적으로 적용될 수 있음을 확인하였으며 전 데이터를 복원하지 않고도 타당한 에지 정보를 추출할 수 있었다. 장면 전환 검출에서 검출률이 DC 에지를 이용한 경우보다 실험에서는 21% 향상되었으며, 잘못 검출한 경우도 AC 에지를 이용했을 경우 DC 에지를 이용한 경우의 29%로 뛰어난 성능을 갖는다는 것을 확인할 수 있었다.

또한 제안한 디졸브 검출 방법에서는 단순히 MPEG 스트림 중 P 프레임 내의 인트라 매크로블록의 개수를 이용하므로 복원 과정과 곱셈 연산 없이 빠른 시간에 수행할 수 있으며, 기존의 방법에 비해서 명료한 디졸브 구간을 검출할 수 있었다. 또한, 디졸브 구간을 검출하는데 있어서 프레임 안쪽에 위치한 매크로블록만을 고려함으로써 팬이나 줌에 의한 영향을 받지 않았다. 기존의 디졸브 검출 알고리즘은 복원 과정과 분산량 계산 등의 많은 계산량을 요구하였으나, 제안한 디졸브 검출 방법은 적은 계산량으로 디졸브 구간을 명확하게 검출하여 속도와 성능이 모두 뛰어난 것을 확인할 수 있었다.

참고 문헌

- [1] J. Meng, Y. Juan and S. F. Chang, "Scene Change Detection in a MPEG Compressed Video Sequence," *Digital Video Compression: Algorithms and Technologies*, vol. SPIE-2419, pp. 14-25, Feb. 1995.
- [2] I. K. Sethi and N. Patel, "A statistical approach to scene change detection," *Storage and Retrieval for Image and Video Database III*, vol. SPIE-2420, pp. 329-338, Feb. 1995.
- [3] H. C. Liu and G. L. Zick, "Scene decomposition of MPEG compressed video," *Digital Video Compression: Algorithms and Technologies*, vol. SPIE-2419, pp. 26-37, Feb. 1995.
- [4] H. J. Zhang, C. Y. Low, and S. W. Smoliar, "Video parsing and browsing using compressed data," *Multimedia Tools and Applications*, vol. 1, no. 1, pp. 89-111, Mar. 1995.
- [5] B. L. Yeo and B. Liu, "Rapid Scene Analysis on Compressed Video," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 5, No. 6, pp. 533-544, Dec. 1995.
- [6] A. Nagasaka and Y. Tanaka, "Automatic video

indexing and full-motion search for object appearances," in *Proc. IFIP TC2/WG2.6 Second Working Conf. on Visual Database Syst.*, pp. 113-127, Sept. 30-Oct. 3, 1991.

- [7] K. Otsuji, Y. Tonomura, and Y. Ohba, "Video browsing using brightness data," in *Visual Communication and Image Processing*, vol. SPIE-1606, pp. 980-989, 1991.
- [8] C. S. Won, D. K. Park, and S. J. Yoo, "Extracting image features from MPEG-2 compressed stream," *Storage and Retrieval for Image and Video Database VI*, vol. SPIE-3312, pp. 426-435, Jan. 1998.
- [9] R. Zabih, J. Miller, and K. Mai, "A feature-based algorithm for detecting and classifying scene breaks," *ACM International Conference on Multimedia*, pp. 189-200, 1995.
- [10] B. Shen and I. K. Sethi, "Direct feature extraction from compressed images," *Storage and Retrieval for Image and Video Database IV*, vol. SPIE-2670, pp. 404-414, 1996.
- [11] 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.
- [12] 광영경, 최윤석, 지은석, 김운경, 고성제, "에지 영상 추출을 이용한 MPEG 비디오의 장면 전환 검출," 제 11회 신호처리합동학술대회, 제 11권, 제 1호, pp. 169-172, Oct. 1998.
- [13] 광영경, 고성제, "압축 비디오에서의 효과적인 디졸브의 검출 기법," 제 11회 영상처리 및 이해에 관한 워크샵 발표 논문집, pp. 45-49, Feb. 1999.

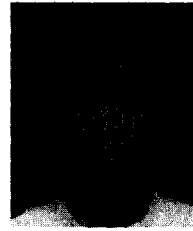
곽 영 경(Young-Gyoung Kwak)정회원



1997년 2월 : 고려대학교 전자공학과 졸업
 1999년 2월 : 고려대학교 전자공학과 석사
 1999년 3월~현재 : LG정보통신 전송연구소 연구원

<주관심 분야> 영상압축, 영상처리, 광대역 전송

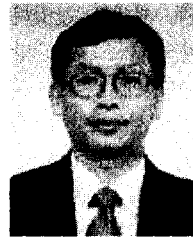
최 윤 석(Yoon-Seok Choi)정회원



1996년 2월 : 고려대학교 전자공학과 졸업
 1998년 8월 : 고려대학교 전자공학과 석사
 1999년 8월~현재 : LG정보통신 차세대 통신연구소 연구원

<주관심 분야> IMT-2000, 영상통신

고 성 제(Sung-Jea Ko)정회원



1980년 2월 : 고려대학교 전자공학과 졸업(학사)
 1986년 5월 : State Univ. of New York at Buffalo, 전기 및 컴퓨터공학과(공학석사)

1988년 8월 : State Univ. of New York at Buffalo, 전기 및 컴퓨터공학과(공학박사)

1981년 8월~1983년 12월 : 대한전선 중앙연구소 연구원

1988년 8월~1992년 5월 : The University of Michigan Dearbon, 전기 및 컴퓨터공학과 조교수

1992년 3월~현재 : 고려대학교 전자공학과 정교수

1996년 11월 : IEEE APCCAS best paper award

1997년 12월 : 대한전자공학회 해동 논문상 수상

<주관심 분야> 신호 및 영상처리, 영상 압축 및 통신, 멀티미디어 통신 등