

논문-98-3-2-03

## 장면 전환 검출을 위한 MPEG 비디오 시퀀스로부터의 특징 요소 추출

최윤석\*, 광영경\*, 고성제\*

### Feature Extraction for Scene Change Detection in an MPEG Video Sequence

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

#### 요약

본 논문에서는 효율적인 장면 전환 검출을 위하여 압축된 MPEG 비디오 시퀀스로부터 특징 요소를 추출하는 방법을 제안한다. 제안한 방법은 압축된 MPEG 비디오 시퀀스로부터 일부만을 복원하여 각 DCT 블록의 주요 5개의 AC 계수를 추출하고, 이로부터 각 블록의 에지 정보를 구하여 에지 영상을 얻는다. 제안된 에지 추출 방법에서는 기존의 DC 계수를 이용하여 에지를 검출하는 방법에서 생기는 블러링(blurring) 문제를 해결할 수 있다. 또한 제안된 방법에 의해 얻어진 에지 영상은 각 프레임의 내용 정보를 포함하고 있으므로 장면 전환 검출뿐만 아니라 내용 기반(content-based) 검색 등에도 효과적으로 이용될 수 있다. 실험 결과, AC 계수를 이용하여 구한 에지 영상은 DC 계수를 이용하여 구한 에지 영상보다 원 영상의 에지 정보를 더 잘 나타내었으며 장면 전환 검출에 효과적으로 적용될 수 있음을 확인하였다.

#### Abstract

In this paper, we propose the method of extracting edge information from MPEG video sequences for the detection of scene changes. In a the proposed method, five significant AC coefficients of each MPEG block are utilized to obtain edge images from the MPEG video. AC edge images obtained by the proposed scheme not only produce better object boundary information than conventional methods using only DC coefficients, but also can reduce the boundary effects produced by DC-based. Since the AC edge image contains the content information of each frame, it can be effectively utilized for the detection of scene change as well as the content-based video query. Experimental results show that the proposed method can be effectively utilized for the detection of scene changes.

#### I. 서론

디지털화된 비디오 정보를 데이터베이스로 구축하고자 할 때에 선행되어야 할 것은 각각의 비디오 정보들을 내용별로 인덱싱(indexing) 하여 사용자가 원하는 비디오 정보에 대해 비순차적으로 접근할 수 있게 하고 내용 기반(content-based) 탐색을 가능하게 하는 것이다. 이를

위해서는 한편의 비디오에서 장면이 전환되는 시점 단위로 분할하여야 한다. 따라서 장면 전환 검출(scene change detection)은 비디오 정보의 데이터베이스화에 있어서 필수적으로 요구되는 기술이다.

장면 전환 검출 알고리즘은 크게 두 가지 형태로 분류할 수 있다. 하나는 비 압축된 상태의 비디오 데이터를 이용하는 방법<sup>[1],[2]</sup>이고 다른 하나는 MPEG 파일과 같이 압축된 상태의 비디오 데이터로부터 복원과정을 거치지 않거나 최소한의 복원처리만을 통하여 바로 장면 전환의 경계를 검출하는 방법<sup>[3],[4]</sup>이다. 비디오 데이터는 양이 많아서 효과적인 전송이나 저장을 위하여 대부분 압축된 상태

\* 고려대학교 전자공학과  
Department of Electronic Engineering

로 되어 있다. 따라서 비 압축된 상태의 데이터를 이용하여 장면 전환 검출을 하기 위해서는 압축된 비디오 데이터 전부를 복원해야 한다. 이는 처리 속도의 측면이나 여러 응용 분야에 적용할 때에 상당한 제약이 된다. 따라서 압축된 상태의 비디오 데이터로부터 복원과정을 거치지 않거나 일부만을 복원하여 한 대의 카메라가 연속해서 찍은 단위인 장면 전환을 검출하는 방법에 대한 연구가 더 활발하게 진행되고 있다<sup>[3],[7]</sup>. 장면 전환을 검출하기 위해서는 프레임간의 차이를 정량적으로 나타낼 수 있는 요소들이 필요하다. 압축된 비디오 시퀀스에서는 압축 표준에 의해 정해진 각각의 패킷 내의 데이터들이 주요한 특징 요소가 될 수 있다. 대표적인 특징 요소들로는 DC 계수와 매크로블록 (macroblock) 타입이 있다<sup>[3],[4]</sup>. 이를 이용하여 장면 전환을 검출하는 대표적인 방법들로는 DC 계수를 이용하여 구한 축소된 영상으로부터 인접한 두 프레임간의 MAD (mean absolute difference) 를 구하거나 영상의 휘도와 색차 성분의 분포를 나타내는 히스토그램의 차이를 이용한 방법들이 있다<sup>[3],[5]</sup>.

MAD를 이용한 방법은 구현하기는 쉬우나 움직임이 많이 존재하는 영상에 대해서는 비효율적이다. 또한 히스토그램을 이용한 방법은 움직임이 크더라도 그 분포가 변하지 않으므로 효과적이나 데이터의 지역적 정보가 없으므로 조명의 변화가 다양한 영상에서는 그 분포가 크게 변하므로 비효율적이다. 따라서 이러한 문제점들을 개선하기 위해서 영상의 형태를 잘 나타내는 에지 영상을 각 프레임으로부터 추출하여 장면 전환 검출에 적용하는 방법들이 연구되고 있다<sup>[8],[10]</sup>.

기존의 압축된 비디오 데이터에서의 DC 계수를 이용하여 에지를 검출하는 방법은 MPEG 비디오 시퀀스로부터 일부만을 복원하여 얻은 DC 영상으로부터 에지 영상을 얻는다<sup>[10]</sup>. 이러한 에지 영상은 DC 영상의 블록 간의 에지만을 나타내므로 원 영상에 있어서  $8 \times 8$  블록의 평균값들간의 차이를 이용하여 구한 에지와 같다. 블록 단위의 평균값을 이용하여 구한 에지 영상은 블러링에 의해 원 영상에 포함된 상당한 에지 정보가 손실되므로 성능이 떨어진다. 이를 개선하기 위해서는 부가적인 처리가 필요하게 된다.

본 논문에서는 프레임 내의 에지 정보를 추출하기 위하여 MPEG 비디오 시퀀스로부터 주요 5개의 AC 계수만을 이용하는 방법을 제안하였다. AC 계수 크기는 각 블록 내에 포함된 에지 성분의 정도와 비례하므로 이를 이용하여 블록 내의 에지의 강도를 조사하고 임계값을 이용하여 해당 블록을 에지 블록으로 결정하여 에지 영상을 얻을 수 있다. 실험 결과, AC 계수를 이용하여 구한 에지 영상은 DC 계수를 이용하여 구한 에지 영상보다 더 좋은 성능을 나타내었다. 또한 제안된 방법에 의해 구한 에지 영상을 장면 전환 검출에 적용하여 제안된 방법의 효용성을 입증하였다.

본 논문의 구성은 다음과 같다. I 장 서론에 이어 II 장

에서는 MPEG 비디오 시퀀스로부터 에지 영상을 추출하는 방법에 대해 설명하였고 III 장에서는 추출된 에지 영상을 이용하여 장면 전환을 검출하는 방법에 대해 설명하였다. IV 장에서는 실험한 결과와 이에 대한 고찰을 하였으며, 마지막으로 V 장에서 결론을 맺었다.

## II. MPEG 비디오 시퀀스로부터의 에지 영상 추출

MPEG 비디오 시퀀스에서 DCT의 AC 계수는 영상의 블록 내에 포함된 공간 주파수 성분의 크기와 비례한다<sup>[11]</sup>. 공간 주파수는 영상의 복잡도와 관련이 있으므로 AC 계수를 이용하면 각 블록에서의 복잡도를 구할 수 있고 이를 이용하여 에지 영상을 얻을 수 있다. MPEG 비디오 영상은 DCT에 의해 인트라 (intra) 부호화된 I 프레임과 움직임 예측에 의해 인터 (inter) 부호화된 P와 B 프레임으로 구성되어 있다. 따라서 MPEG 비디오 시퀀스로부터 에지 영상을 얻기 위해서는 인트라 부호화된 I 프레임에서는 DCT의 AC 계수를 이용하여 영상의 복잡도를 구하고 움직임 예측 부호화된 P와 B 프레임에서는 움직임 보상을 통하여 영상의 복잡도를 예측하여야 한다. 다음은 I 프레임 및 P와 B 프레임에서 에지 정보를 추출하는 방법들에 대하여 설명한다.

### 1. AC 계수를 이용한 에지 영상 추출

MPEG 비디오 시퀀스의 I 프레임은 DCT에 의해  $8 \times 8$  블록 단위로 부호화되어 있다. DCT는 픽셀 값을 공간 주파수 성분으로 변환하므로 DCT 계수는 해당 블록 내에 포함되어 있는 공간 주파수 성분의 크기를 나타낸다. 다음은  $N \times N$  블록 단위의 DCT를 나타낸 것이다.

$$F_{uv} = C_u C_v \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \cos \frac{(2i+1)u\pi}{2N} \cos \frac{(2j+1)v\pi}{2N} f(i, j),$$

$$C_u, C_v = \begin{cases} \sqrt{\frac{1}{N}} & \text{for } u, v = 0 \\ \sqrt{\frac{2}{N}} & \text{otherwise} \end{cases} \quad (1)$$

여기서,  $f(i, j)$ 는  $(i, j)$  위치의 픽셀 값이고,  $F_{uv}$ 는  $(u, v)$ 의 공간 주파수에 해당하는 DCT 계수이다. 공간 주파수 성분은 그림 1과 같은 기저 영상들( $A_m$ )로 표현될 수 있으며, 이를 이용하여 에지 정보를 얻을 수 있다. 예를 들어,  $N=4$ 인 블록 단위의 16개의 DCT 계수 ( $F_{00}, F_{01}, \dots, F_{33}$ )는 그림 1의 영상 내에 각각에 해당하는 기저 영상의 성분이 얼마나 포함되어 있는지를 나타낸다.

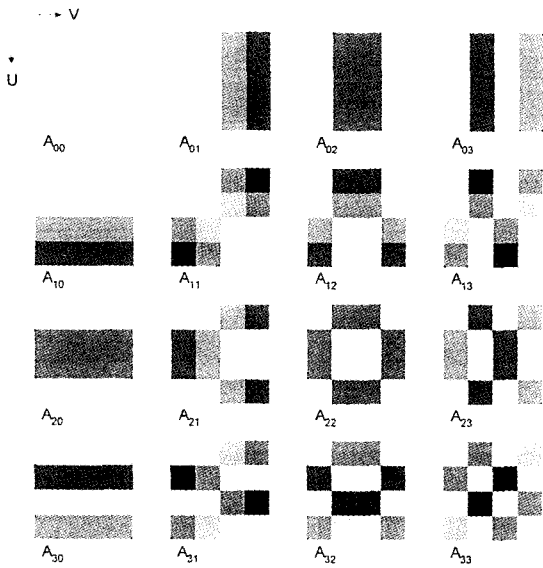


그림 1. DCT 변환의 기저 영상  
Fig. 1. Discrete cosine transform basis image.

16개의 DCT 계수 중에서 5개의 AC 계수,  $F_{01}, F_{10}, F_{02}, F_{20}, F_{11}$ 는 저주파에 해당하는 기저 영상을 나타내는 성분으로 수직, 수평, 그리고 대각선 방향으로 눈에 띄게 변하는 에지 성분을 나타낸다. 반면에 이러한 5개의 AC 계수를 제외한 나머지 AC 계수들은 고주파에 해당하는 기저 영상을 나타내는 성분으로 시각적으로 볼 때 그 변화 정도가 복잡하여 에지 정보에 크게 영향을 주지 못한다. 즉, 고주파에 해당하는 AC 계수는 물체의 전체 윤곽을 나타내기보다는 물체 내의 복잡한 무늬를 나타내는 경우가 많기 때문에 모든 AC 계수를 이용한다면 원치 않는 부분에서의 에지를 추출하는 경우가 많게 된다. 따라서 본 논문에서는 저주파에 해당하는 5개의 AC 계수만을 사용하여 에지 영상을 구하였다.

이러한 영상의 에지는 블록의 경계에서 약간만 어긋나 있어도 블록 내의 AC 계수의 값들이 커지므로 블록 내의 에지 성분이 커지게 되어 에지 검출을 할 수 있게 된다. 즉 블록 간의 경계를 이루는 에지들도 결국은 블록 내에 포함된다. 즉, 블록 내의 에지는 블록 간의 에지를 포함하는 형태가 된다.

블록 내에 포함된 에지 성분의 정도는 다음과 같이 AC 계수들의 제곱의 합인 에지 강도 (edge power)  $P_E$ 로 나타낼 수 있다.

$$P_E = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F_{uv}^2 - F_{00}^2 \quad (2)$$

따라서 주요한 에지 정보의 양을 나타내는 저주파 에지 강도  $P_L$ 은 다음과 같이 5개의 AC 계수의 제곱의 합으로 구할 수 있다.

$$P_L = F_{01}^2 + F_{10}^2 + F_{02}^2 + F_{20}^2 + F_{11}^2 \quad (3)$$

제안된 에지 검출 방법에서, I 프레임은 블록 단위로 분할되어 부호화되므로 각 블록마다의 저주파 에지 강도를 구하고, 이 에지 강도로부터 임계값  $T$ 를 이용하여 에지 블록인지 아닌지를 결정하면 블록 단위의 에지 영상을 얻을 수 있다. 즉 원 영상의 각 블록에서 블록 내에 주요한 에지 성분이 많이 존재하면 그 블록은 에지 블록으로 간주된다. 그리하여 각각의 블록들은 에지 블록 또는 비 에지 블록으로 분류되며 각각의 블록들을 하나의 픽셀 값으로 표현함으로써 결국, 원 영상보다  $N$ 분의 1 크기로 축소된 이진 에지 영상을 얻게 된다. 예를 들어, 각 블록에서의 저주파 에지 강도  $P_L$ 이 그림 2(a)와 같을 때, 임계값  $T=80$ 을 적용하여 그림 2(b)와 같은 블록 단위의 이진 에지 영상을 얻을 수 있다.

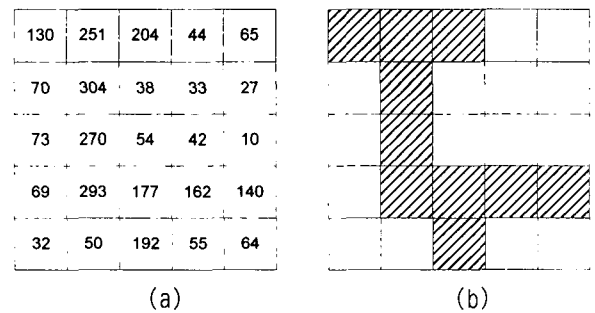


그림 2. 에지 영상 추출,  
(a) 블록 단위의 에지 강도, (b) 블록 단위의 이진 에지 영상.  
Fig. 2. Extraction of edge image,  
(a) block-based edge power, (b) block-based binary edge image.

기존의 DC 계수를 이용한 에지 영상 추출 방법에서는 각 블록의 평균값, 즉 DC 계수를 사용하여 인접한 블록의 DC 계수와 비교하여 그 차이가 임계값 이상일 경우에는 에지가 두 블록 사이에 존재한다고 가정하여 에지 영상을 추출한다. 이러한 방법에서는 블록 단위의 평균값을 이용하여 에지 영상을 구하므로 블록 내의 블러링 현상에 의해 원 영상에 포함된 상당한 에지 정보가 손실된다. 그러나 AC 계수를 이용한 방법은 각 블록 내에 포함된 주요한 에지 성분의 정도에 따라 에지 영상을 추출하므로 블록 내의 에지 성분이 보존된 에지 영상을 구할 수 있다. 또한 5개의 AC 계수만을 사용하므로 모든 AC 계수를 복원하지 않아도 되므로 시간적으로도 효율적으로 에지 영상을 구할 수 있다.

이처럼 DCT에 의해 부호화된 I 프레임에서는 각 블록의 DCT 계수로부터 주요 AC 계수를 이용하여 에지 강도를 구하고 이를 이용하여 에지 영상을 얻을 수 있다. 그러나 움직임 예측 부호화된 P나 B 프레임에서는 각 블록에서 바로 AC 계수를 구할 수 없다. 따라서 P와 B 프레임에서는 에지 강도 예측을 통하여 에지 정보를 추출하여야 한다. 다음은 움직임 예측 부호화된 블록에서 에지 강도를 예측하고 이를 이용하여 에지 정보를 추출하는 방법에 대하여 설명한다.

2. P와 B 프레임에서의 에지 정보 추출

P와 B 프레임에서의 움직임 예측 부호화된 블록은 참조 프레임으로부터 해당 픽셀 값을 참조한다. 따라서 움직임 예측 부호화된 블록의 에지 정보는 참조 프레임의 에지 정보와 밀접한 관련이 있다. 가장 간단한 방법으로 움직임 예측 부호화된 부분의 값을 4개의 참조 블록 중 가장 많이 겹쳐진 블록의 값으로 대체하는 방법이 있다. 이 방법은 정확한 에지 강도의 추출보다는 에지 강도 값을 이전 프레임과 똑같이 유지하는 방법으로 계산량이 적어서 빠르게 에지를 검출한다는 장점이 있다. 그러나 겹쳐진 네 블록 모두를 고려하는 방법으로 좀 더 정확한 에지 강도를 검출할 수 있다. 다음은 4개의 참조 블록 모두를 고려한 경우의 에지 강도 예측 방법이다.

앞 절에서도 설명하였듯이 에지 정보는 AC 계수의 제곱의 합인 에지 강도를 이용하여 구할 수 있으므로 움직임 예측 부호화된 블록에서 에지 강도를 예측하여 P와 B 프레임에서의 에지 정보를 구할 수 있다. 다음은 분산값을 이용하여 에지 강도를 예측하는 방법을 설명한 것이다.

$N \times N$  블록 내의 샘플들의 분산값  $\sigma_b^2$ 은 다음과 같이 구할 수 있다.

$$\sigma_b^2 = \frac{1}{N^2} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f^2(x, y) - \mu_b^2 \tag{4}$$

여기에서,  $f(x, y)$ 는  $(x, y)$  위치의 픽셀 값,  $N^2$ 은 블록의 총 픽셀 수,  $\mu_b$ 은 블록 내의 샘플들의 평균값이다.

이는 Parseval의 정리

$$\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f^2(x, y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F_{uv}^2 \tag{5}$$

와 식 (1), (4)를 이용하여 정리하면 다음과 같이 정리된다.

$$\sigma_b^2 = \frac{1}{N^2} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F_{uv}^2 - \mu_b^2 \tag{6a}$$

$$\mu_b = \frac{1}{N^2} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) = \frac{F_{00}}{N} \tag{6b}$$

여기에서, 분산값  $\sigma_b^2$ 은 DCT 계수만으로 구할 수 있음을 알 수 있다. 따라서 식 (6)을 이용하여 식 (2)의 AC 계수의 제곱의 합인 에지 강도  $P_E$ 를 다음과 같이 분산값과 블록 내의 샘플들의 총 개수의 곱으로 나타낼 수 있다.

$$P_E = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} F_{uv}^2 - F_{00}^2 = N^2 \sigma_b^2 \tag{7}$$

식 (7)에서  $P_E$ 는 블록 내의 샘플들의 분산값에 비례하므로 움직임 예측 부호화된 블록에서의 에지 강도  $P_E$ 는 분산값 예측을 통하여 구할 수 있다. 움직임 예측 부호화 블록의 분산값은 움직임 벡터에 의해 참조되는 블록들의 평균 및 분산값들과 움직임 예측 부호화 블록과 참조 블록이 겹치는 부분의 면적을 이용하여 예측할 수 있다<sup>[12]</sup>. 그림 3은 움직임 예측 블록  $B_p$ , 움직임 예측 부호화 블록의 참조 블록(reference block)  $B_1, B_2, B_3, B_4$ , 그리고 움직임 벡터와의 관계를 나타내고 있다.

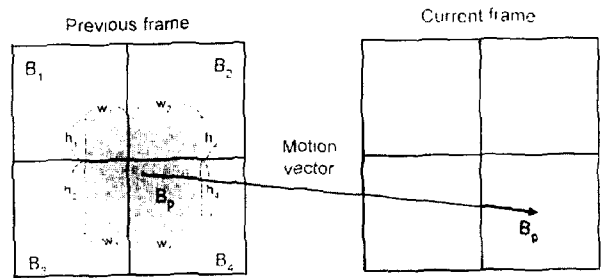


그림 3. 분산값 예측  
Fig. 3. Prediction of the variance

그림 3에서 각각의 해당 블록의 평균과 분산이 각각  $E[B_i] = \mu_i$ ,  $Var[B_i] = \sigma_i^2$ ,  $i = 1, 2, 3, 4$ 이고 움직임 벡터에 의해 참조되는 블록  $B_i$ 와 겹치는 부분의 면적이  $h_i w_i$ 라 할 때,  $B_p$ 의 표본들은  $X(1), X(2), X(3), \dots, X(N^2)$ ,  $N^2 = h_1 w_1 + h_2 w_2 + h_3 w_3 + h_4 w_4$  이며 각각이 참조하는 블록이 아래와 같다면

$$\begin{aligned} B_1 &= \{X(1), X(2), \dots, X(h_1 w_1)\} \\ B_2 &= \{X(h_1 w_1 + 1), \dots, X(h_1 w_1 + h_2 w_2)\} \\ B_3 &= \{X(h_1 w_1 + h_2 w_2 + 1), \dots, X(h_1 w_1 + h_2 w_2 + h_3 w_3)\} \\ B_4 &= \{X(h_1 w_1 + h_2 w_2 + h_3 w_3 + 1), \dots, X(h_1 w_1 + h_2 w_2 + h_3 w_3 + h_4 w_4)\} \end{aligned} \tag{8}$$

$N=8$ 인 움직임 예측 블록  $B_p$ 의 평균값  $E[B_p]$ 는 다음과 같이 겹쳐진 블록의 면적과 평균값으로 나타낼 수 있다.

$$E[B_p] = \frac{X(1) + X(2) + X(3) + \dots + X(64)}{64} \approx \frac{h_1 w_1 \mu_1 + h_2 w_2 \mu_2 + h_3 w_3 \mu_3 + h_4 w_4 \mu_4}{64} \quad (9)$$

$$= \sum_{i=1}^4 \frac{h_i w_i}{64} \mu_i$$

또한,  $B_p$ 의 분산값  $Var[B_p]$ 는 다음과 같이 구할 수 있다.

$$Var[B_p] = E[B_p^2] - E^2[B_p] = \frac{X^2(1) + X^2(2) + X^2(3) + \dots + X^2(64)}{64} - E^2[B_p]$$

$$= \frac{h_1 w_1 (\sigma_1^2 + \mu_1^2) + h_2 w_2 (\sigma_2^2 + \mu_2^2) + h_3 w_3 (\sigma_3^2 + \mu_3^2) + h_4 w_4 (\sigma_4^2 + \mu_4^2)}{64} - E^2[B_p]$$

$$= \sum_{i=1}^4 \frac{h_i w_i}{64} (\sigma_i^2 + \mu_i^2) - \left( \sum_{i=1}^4 \frac{h_i w_i}{64} \mu_i \right)^2 \quad (10)$$

식 (7)에서 예지 강도는 분산값과 블록 내의 샘플의 총 개수의 곱으로 나타낼 수 있으므로 움직임 예측된 블록에서의 예지 강도  $\tilde{P}_E$ 는 다음과 같이 나타낼 수 있다.

$$\tilde{P}_E = 64 Var[B_p] = 64 \left[ \sum_{i=1}^4 \frac{h_i w_i}{64} (\sigma_i^2 + \mu_i^2) - \left( \sum_{i=1}^4 \frac{h_i w_i}{64} \mu_i \right)^2 \right] \quad (11)$$

위의 움직임 예측 부호화된 블록의 예지 강도  $\tilde{P}_E$ 를 식 (7)을 이용하여 저주파 예지 강도  $\tilde{P}_L$ 과 고주파 예지 강도  $\tilde{P}_H$ 로 분리하면 다음과 같다.

$$\tilde{P}_E = \sum_{i=1}^4 h_i w_i \left( \frac{P_{L_i} + P_{H_i}}{64} + \mu_i^2 \right) - 64 \left( \sum_{i=1}^4 \frac{h_i w_i}{64} \mu_i \right)^2$$

$$= \sum_{i=1}^4 \frac{h_i w_i}{64} P_{L_i} + \sum_{i=1}^4 \frac{h_i w_i}{64} P_{H_i} + \sum_{i=1}^4 (h_i w_i \mu_i^2) - 64 \left( \sum_{i=1}^4 \frac{h_i w_i}{64} \mu_i \right)^2$$

$$= \tilde{P}_L + \tilde{P}_H + P_M \quad (12)$$

여기에서,  $P_{L_i}, P_{H_i}$ 는 각각  $B_i$  블록에서의 저주파 예지 강도와 고주파 예지 강도를 나타내며  $P_M$ 은 움직임 예측 부호화된 블록  $B_p$  내에 포함되는 참조 블록  $B_i$ 들 사이의 평균값의 차이에 의해 생기는 예지 성분의 강도이다.

여기서  $\tilde{P}_M$ 는 고주파 예지 강도를 나타낸다.

$$\tilde{P}_L = \sum_{i=1}^4 \frac{h_i w_i}{64} P_{L_i}$$

$$\tilde{P}_H = \sum_{i=1}^4 \frac{h_i w_i}{64} P_{H_i} \quad (13)$$

$$P_M = \sum_{i=1}^4 (h_i w_i \mu_i^2) - 64 \left( \sum_{i=1}^4 \frac{h_i w_i}{64} \mu_i \right)^2$$

움직임 예측 부호화된 블록의 주요한 예지 성분의 크기는 식 (13)으로부터 저주파 예지 강도  $\tilde{P}_L$ 과 블록 간의 예지 강도  $P_M$ 를 더하여 얻을 수 있으며, 이러한 방법으로 예지 강도 영상을 구할 수가 있게 된다. 또한 이 예지 강도 영상에 I 프레임에서와 같은 임계값  $T$ 를 이용하여 예지 블록을 결정하면 이로부터 예지 영상을 구할 수 있다.

다음 장에서는 I, P, 그리고 B 프레임으로부터 구한 예지 영상을 이용하여 장면 전환을 검출하는 방법에 대하여 설명한다.

### III. 예지 영상을 이용한 장면 전환 검출

본 장에서는 AC 계수와 분산값 예측을 통하여 얻은 예지 영상을 이용하여 장면 전환을 검출하는 방법에 대해 설명한다. 예지 영상을 이용하여 장면 전환 검출을 위한 인접 프레임간의 유사도를 측정하는 방법으로는 픽셀 단위로 예지 영상을 매칭하여 비교하는 방법과 프레임의 수평, 수직 방향으로의 히스토그램을 구한 후에 각각의 방향에 대한 히스토그램의 유사도를 비교하는 방법이 있다. 단순히 예지를 비교하는 방법은 예지가 조금만 어긋나도 커다란 차이를 나타낸다. 따라서 움직임에 덜 민감한 수평, 수직 방향의 히스토그램을 이용한 방법이 더 효과적이다.

이를 좀 더 개선하기 위하여 본 논문에서는 움직임 보상을 하여 예지 히스토그램의 유사도를 비교하였다. 이는 움직임이 큰 영상에 대해서도 효과적으로 장면 전환을 검출할 수 있게 한다. 그림 4는 수평, 수직 방향으로의 히스토그램을 구하는 과정과 움직임 보상을 하는 과정을 나타내고 있다. 그림 4에서 수평 방향의 히스토그램인  $P_H$ 는 현재 프레임의 0부터  $M-1$ 행까지 세로 방향으로의 예지 블록의 개수를 나타내는 히스토그램이고, 수직 방향의 히스토그램인  $P_V$ 는 현재 프레임의 0부터  $L-1$ 행까지 가로 방향으로의 예지 블록의 개수를 나타내는 히스토그램이다. 또한  $P_H, P_V$ 는 각각 이전 프레임에서의 수평, 수직 방향의 히스토그램이다. 현재 프레임과 이전 프레임의 유사도를 측정하기 위해서는 이러한 수평, 수직 방향의 히스토그램을 이용하여 움직임 보상을 하여야 하므로 움직임 벡터를 구해야 한다.

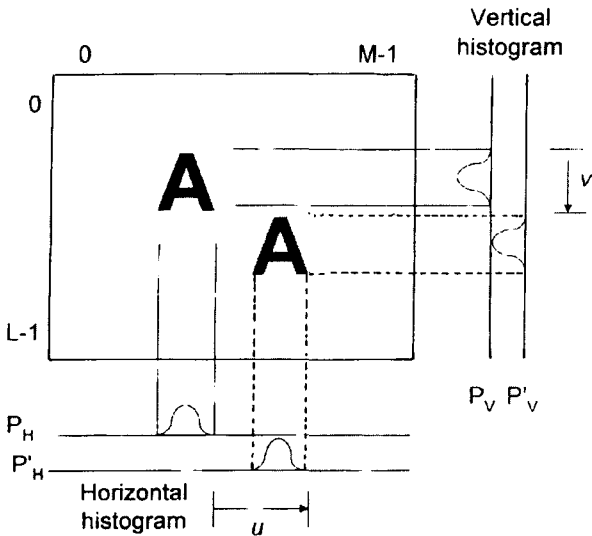


그림 4. 방향성 에지 히스토그램을 이용한 움직임 보상  
 Fig. 4 Motion compensation using directional edge histogram.

그림 4에서  $(u, v)$ 가 각각 에지 영상의 수평, 수직 방향의 움직임 벡터를 나타낸다. 따라서 움직임 보상된 두 프레임의 수평, 수직 방향의 유사도  $C_H(u), C_V(v)$ 는 다음과 같이 에지 히스토그램의 차이로 나타낼 수 있다.

$$C_H(u) = \sum_{x=0}^{M-1} |P'_H(x+u) - P_H(x)|, C_V(v) = \sum_{y=0}^{L-1} |P'_V(y+v) - P_V(y)| \quad (14)$$

이를 이용하여 두 프레임의 전체 유사도  $S$ 는 다음과 같이 구할 수 있다.

$$S = \frac{1}{C_H(u)} + \frac{1}{C_V(v)} \quad (15)$$

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

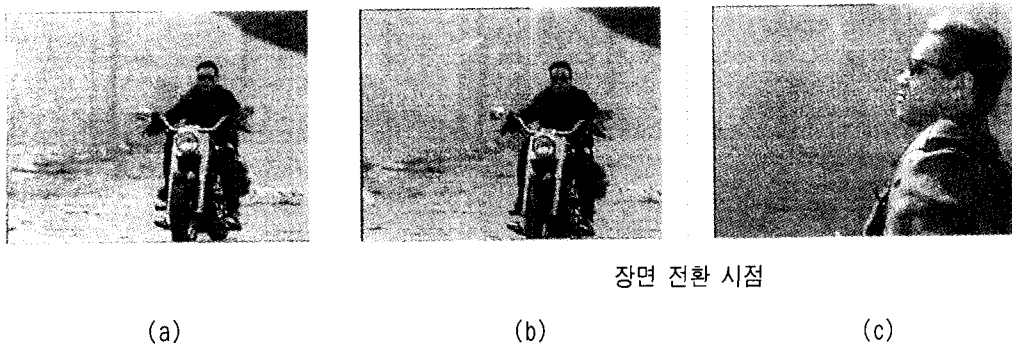
#### IV. 실험 결과 및 고찰

본 장에서는 AC 계수로부터 에지를 추출한 결과와 이를 이용하여 장면 전환 검출에 적용한 결과 보이고, 기존의 DC 계수로부터 구한 에지와 이를 이용하여 장면 전환을 검출한 결과와의 비교를 통하여 제안된 기법을 평가하였다. 장면의 전환 시점은 연속되는 두 장면이 급격하게 달라지는 시점으로 정하였으며 실험은 움직임과 조명의 변화가 많은 영상을 위주로 하였다. 그림 5는 150번째 프레임에서 급격한 장면 전환을 갖는 bike.mpg 영상에서 장면 전환이 일어나는 부근의 프레임들을 나타낸 것이다.

bike.mpg의 비디오 시퀀스로부터 부분 복원하여 DC 계수를 구하여 DC 영상을 얻을 수 있다. 그림 6는 bike.mpg 영상으로부터 구한 장면 전환 시점 부근에서의 DC 영상을 8배 확대하여 나타낸 것이다.

MPEG 비디오 시퀀스로부터 부분 복원하여 구한 DC 영상으로부터 인접 픽셀 값과의 차이를 이용하면 에지 영상을 얻을 수 있다. 다음 그림 7은 DC 영상으로부터 임계값을 정하여 구한 에지 영상을 8배로 확대하여 나타낸 것이다.

임계값은 MPEG 신택스 중 DC 계수 차의 비트수를 나타내는 dct\_dc\_size를 사용한 경우와 같도록 설정하였으며[10], 실험에서는 dct\_dc\_size가 3 이하인 것과 그렇지 않은 것을 구분하도록 임계값을 7로 하였으며, 이 값은 실험적으로 에지가 가장 잘 나타나는 것으로 결정한 것이다.



장면 전환 시점

- (a)
- (b)
- (c)

그림 5. 장면 전환이 일어나는 MPEG 비디오 파일의 예,  
 (a) 147번째 프레임(P 타입), (b) 149번째 프레임(B 타입), (c) 150번째 프레임(I 타입).  
 Fig. 5. The example of MPEG video file containing scene change,  
 (a) 147th frame (P type), (b) 149th frame (B type), (c) 150th frame (I type).

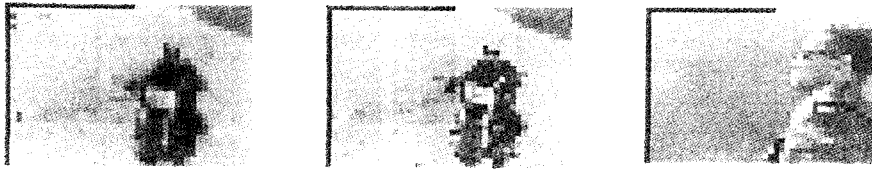


그림 6. MPEG 비디오 시퀀스로부터 부분 복원하여 구한 DC 영상.  
Fig. 6. DC images of bike.mpg.



그림 7. DC 계수로부터 구한 에지 영상.  
Fig. 7. Edge images obtained from DC coefficient.

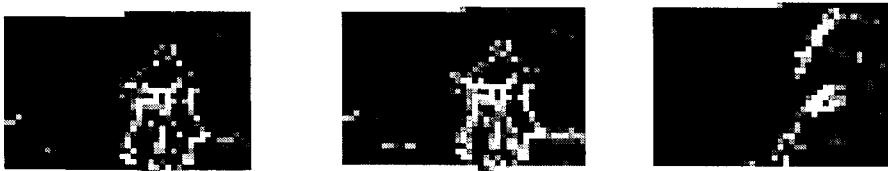


그림 8. AC 계수와 분산값을 이용하여 구한 에지 강도 영상.  
Fig. 8. Edge power images obtained from AC coefficient and variance.



그림 9. AC 계수를 이용하여 구한 에지 영상.  
Fig. 9. Edge images obtained from AC coefficient.

그림 7에서의 DC 영상으로부터 구한 에지 영상은 원 영상의 에지를 그리 효과적으로 표현하지는 못하였다. 그림 8은 본 논문에서 제안한 AC 계수와 블록 내의 분산값을 이용하여 구한 에지 강도를 나타내는 영상이다. 이로부터 임계값을 이용하면 그림 9와 같은 에지 영상을 얻을 수 있다. 이처럼 AC 계수로부터 구한 에지 영상은 DC 계수를 이용하여 구한 에지 영상인 그림 7과 비교하여 알 수 있듯이 원 영상에서의 물체의 형태를 더 잘 나타내고 있다.

이는 DC 성분 자체가 원 영상의 8×8 블록의 평균값을 나타내는 값이므로 블러링 효과가 크기 때문이다. 따라서 이를 이용하여 구한 에지 영상은 영상의 복잡도를 표현하고 있는 AC 계수를 이용하여 구한 에지보다 그 성능이

떨가 많을 뿐만 아니라 움직임이 큰 물체를 많이 포함하고 어진다.

다음으로는 실제 비디오 영상에 DC 계수와 AC 계수를 이용하여 구한 에지 영상을 장면 전환을 검출하는데 적용하여 그 성능을 비교하여 보았다. 장면 전환의 검출의 임계값으로는 가장 간단한 방법인 결과값 중에서 최대의 차이를 나타내는 결과값의 절반에 해당하는 값으로 정하였다. 실험 영상으로는 airwolf.mpg 파일 (412 프레임, 24×18), DebbieHarry.mpg 파일 (100 프레임, 20×16), 그리고 jackson.mpg 파일 (155 프레임, 20×16)을 사용하였다. airwolf.mpg는 형태 변화가 많은 비디오 영상이고, DebbieHarry.mpg와 jackson.mpg는 뮤직 비디오로써 조명 변화있는 비디오 영상이다. 이러한 영상에서 에지 영상을

구한 후, 에지 히스토그램을 구하였다. 그리고 나서, 마지막으로 움직임 예측을 통하여 에지 히스토그램을 움직임 보상하고, 이 결과와 식 (14)를 이용하여 장면 전환 검출을 하였다.

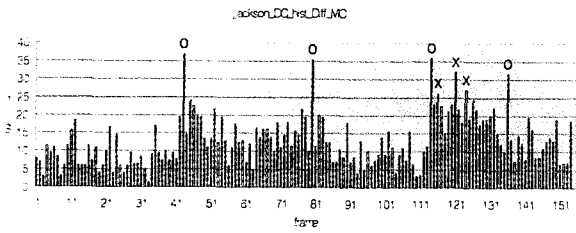


그림 10. DC 성분으로부터 구한 에지 영상을 이용한 에지 히스토그램의 차, 실험 영상: jackson.mpg.  
Fig. 10. Difference of edge histogram using DC edge image, test image: jackson.mpg.

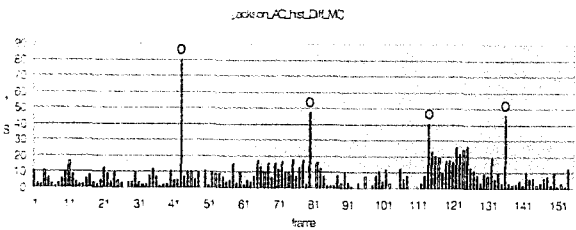


그림 11. AC 계수로부터 구한 에지 영상을 이용한 에지 히스토그램의 차, 실험 영상: jackson.mpg.  
Fig. 11. Difference of edge histogram using AC edge image, test image: jackson.mpg.

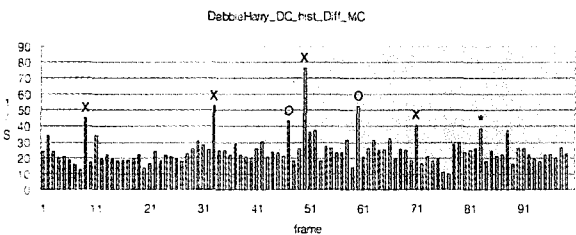


그림 12. DC 성분으로부터 구한 에지 영상을 이용한 에지 히스토그램의 차, 실험 영상: DebbieHarry.mpg.  
Fig. 12. Difference of edge histogram using DC edge image, test image: DebbieHarry.mpg.

그림 10~15는 jackson.mpg, DebbieHarry.mpg 및 airwolf.mpg 영상에 대해 각각 DC 계수와 AC 계수를 사용하여 구한 에지 영상의 프레임간의 유사도의 역을 취한

결과이다.

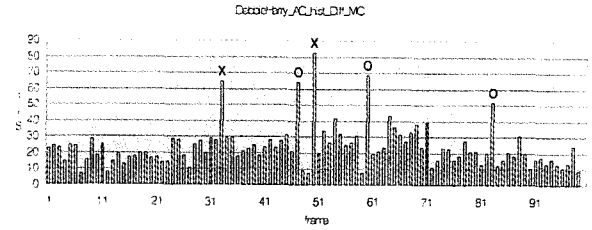


그림 13. AC 계수로부터 구한 에지 영상을 이용한 에지 히스토그램의 차, 실험 영상: DebbieHarry.mpg.  
Fig. 13. Difference of edge histogram using AC edge image, test image: DebbieHarry.mpg.

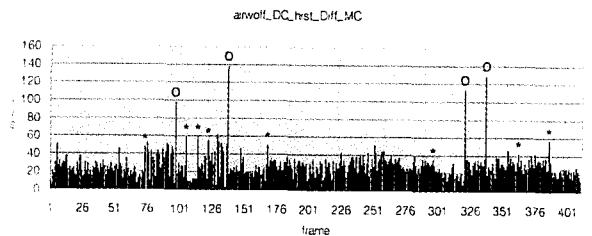


그림 14. DC 성분으로부터 구한 에지 영상을 이용한 에지 히스토그램의 차, 실험 영상: airwolf.mpg.  
Fig. 14. Difference of edge histogram using DC edge image, test image: airwolf.mpg.

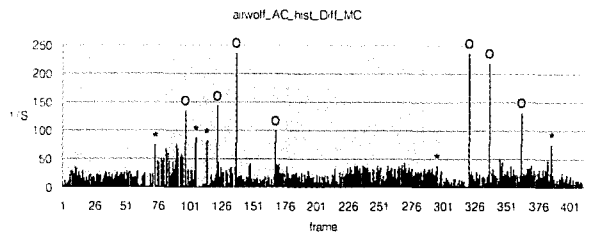


그림 15. AC 계수로부터 구한 에지 영상을 이용한 에지 히스토그램의 차, 실험 영상: airwolf.mpg.  
Fig. 15. Difference of edge histogram using AC edge image, test image: airwolf.mpg.

이 때의 결과값이 임계값을 넘으면 해당 프레임을 장면 전환이 발생한 곳이라고 간주하고, 장면 전환을 빠르게 찾았는지를 O, X로 표시하였으며, 찾지 못한 장면 전환 부분에 대해서는 \*로 표시하였다.







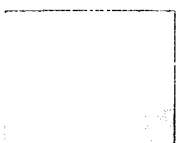

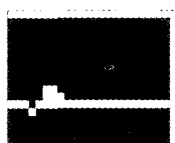



표 1을 보면 DC 계수 및 AC 계수를 이용하여 장면 전환을 검출한 결과, AC 에지 영상을 이용한 경우에 있어서 장면 전환을 놓친 경우와 잘못 검출한 장면 전환의 수가 적다. 즉, AC 계수를 이용한 방법이 장면전환을 더 정확



표 1. DC 에지 및 AC 에지 영상을 이용하여 장면 전환 검출을 한 결과.  
Table 1. The result of scene change detection using DC and AC edge image.

실험 영상	실제 장면 전환	사용한 에지 영상	제대로 검출한 장면 전환	잘못 검출한 장면 전환	검출하지 못한 장면 전환
jjackson	4	DC	4	3	0
		AC	4	0	0
DebbieHarry	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

표 2. DC 및 AC 에지 영상을 이용하여 실험한 결과 영상 비교.  
Table 2. Comparison of the result images of experiments using DC and AC edge imagees.

영상 종류	원영상	DC 에지 영상	AC 에지 영상	비고
DebbieHarry 8 프레임				플래쉬. DC 에지 영상에서 잘못된 검출을 함
DebbieHarry 9 프레임				
airwolf 168 프레임				평이한 영상에서 복잡한 영상으로의 장면 전환. DC 에지 영상에서 못찾음.
airwolf 169 프레임				

하게 찾는다는 것을 알 수 있다. DC 에지 영상으로 실험한 경우, jjackson 영상에서는 빠른 동작에 의한 빛의 변화로 인해서 장면 전환을 잘못 검출하였으며, DebbieHarry 영상에서는 플래쉬 부분에서 잘못 검출하였고 또한 어두

운 장면에서 검출하지 못한 장면도 있었다. airwolf 영상에서는 전체적으로 평활한 영상에서 아주 복잡한 영상으로 바뀌는 부분에서의 장면 전환을 검출하지 못하였다. 다음 표 2에는 DC 및 AC 계수를 이용하여 실험하여 다른

결과가 나온 부분에 대한 예를 나타내었다.

그림 10~15를 보면 DC 에지 영상 및 AC 에지 영상의 히스토그램을 이용하여 장면 전환을 검출한 결과를 볼 수가 있으며, 표 1에 정리하여 나타내었다.

실험 결과, DC 계수를 이용한 방법에서는 블러링 현상에 의해서 그 에지 정보가 손실되거나 가늘고 복잡한 에지와 굵고 선명한 에지를 구분하지 못하여 장면 전환을 검출하기가 어려웠다. 그러나 AC 계수를 이용한 방법에서는 블록 내의 복잡도로부터 에지 정보를 얻으며 굵고 선명한 에지에 더욱 중점을 두기 때문에 물체의 형태 변화에 의한 장면 전환 검출이 더욱 용이함을 알 수 있었다. 따라서 AC 계수로부터 구한 에지 영상을 이용한 장면 전환 검출은 조명의 변화에 의한 잘못된 장면 전환을 검출하지 않으며, 비슷한 배경 하에서의 형태 변화에 의한 장면 전환 검출을 할 수 있으므로 DC 계수를 이용한 방법에 비해 우수한 성능을 나타낸다는 사실을 확인할 수 있었다.

또한 AC 계수를 이용한 방법의 히스토그램을 보면 장면 전환이 검출되는 프레임에서의 값들이 다른 값들에 비해 뚜렷하게 큰 값을 가지는 경향이 있음을 알 수가 있다. 따라서 장면 전환 검출의 임계값을 단순히 최고값의 절반이 아닌 구간에 따른 값으로 조정하거나 기존의 장면 전환 검출 방법과 함께 쓴다면 더욱 정확한 장면 전환을 검출할 수 있을 것이다. 그러나 AC 계수로 구하는 에지 영상은 DC 계수뿐만이 아니라 5개의 AC 계수를 복원해야 하며, 움직임 예측된 블록에 대해서는 분산값 예측을 이용하여 에지를 찾기 때문에 기존의 DC 계수만을 이용하는 경우보다 계산량이 많아지게 된다.

## V. 결론

기존의 장면 전환 검출 알고리즘은 대부분 영상의 DC 계수 또는 움직임 벡터의 타입을 이용하여 장면 전환을 검출하는 방법이었다. 이러한 방법에서는 각 비디오 프레임 내에 포함되어 있는 영상의 내용 변화에 의한 장면 전환을 검출하기가 어렵고 조명 변화에 너무 민감하다는 문제점이 있었다. 따라서 이러한 문제점을 해결하기 위해서 본 논문에서는 MPEG 비디오 시퀀스로부터 AC 계수를 이용하여 에지를 추출하고 이를 이용하여 장면 전환을 검출하는 방법을 제안하였다. DC 계수에 의한 에지 정보 추출 방법은 DC 계수 자체가 블록의 평균값을 나타내므로 블러링 현상에 의해서 원 영상에서 상당한 에지 정보가 손실되므로 그리 효과적 에지 정보를 얻지 못하였다. 반면에 AC 계수를 이용하여 영상의 복잡도를 측정하고 이를 이용하여 에지를 구하는 방법은 블록 내의 복잡도를 이용하여 바로 에지 정보를 추출하였으므로 블러링 현상이 나타나지 않을 뿐만 아니라 전 데이터를 복원하지 않고도 타당한 에지 정보를 추출할 수 있었다. 또한 DC 계수만을

이용한 에지 영상보다 더 명료하여 장면 전환 검출에 더 효과적으로 적용될 수 있음을 확인하였다. 본 논문에서 제안한 기술은 내용 기반 검색과 같은 응용분야에 적용될 수 있으리라 기대된다.

## 참고 문헌

- [1] 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.
- [2] 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.
- [3] 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.
- [4] 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.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] 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.
- [9] 박영경, 최윤석, 지은석, 김운경, 고성제, "에지 영상 추출을 이용한 MPEG 비디오의 장면 전환 검출," *제 11 회 신호처리합동학술대회*, 제11권, 제1호, pp. 169-172, Oct. 1998.
- [10] 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.

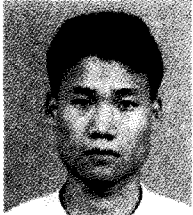
[11] B. Shen and I. K. Sethi, "Direct feature extraction from compressed images," *Storage and Retrieval for Image and Video Database IV*, vol. SPIE-2670, pp. 404-414, 1996.

[12] 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.

---

저 자 소 개

---



**최 윤 석**

1996년 2월 고려대학교 공과대학 전자공학과 졸업(학사)  
 1998년 8월 고려대학교 공과대학 전자공학과 졸업(공학석사)  
 현재 LG정보통신 차세대 통신연구소 연구원  
 주관심분야 : 영상통신, 영상압축, 이동통신



**곽 영 경**

1997년 2월 고려대학교 공과대학 전자공학과 졸업(학사)  
 현재 고려대학교 공과대학 전자공학과 석사과정  
 주관심분야 : 영상압축, 영상개선, 영상통신



**고 성 제**

1980년 2월 고려대학교 공과대학 전자공학과 졸업(학사)  
 1986년 5월 State Univ. of New York at Buffalo, 전기 및 컴퓨터 공학과(공학석사)  
 1988년 8월 State Univ. of New York at Buffalo, 전기 및 컴퓨터 공학과(공학박사)  
 1981년 8월 ~ 1983년 12월 대한전선 중앙연구소 연구원  
 1988년 8월 ~ 1992년 5월 The University of Michigan-Dearborn  
 전기 및 컴퓨터공학과 조교수  
 1996년 11월 IEEE APCCAS best paper award  
 1997년 12월 대한전자공학회 해동 논문상 수상  
 1997년 ~ 현재 IEEE Senior member  
 현재 고려대학교 전기전자전파공학부 부교수  
 주관심분야 : 신호 및 영상처리, 영상압축 및 통신, 멀티미디어 통신 등