프레임 특징 비교를 이용한 압축비디오 분할
MPEG Video Segmentation Using Frame Feature Comparison

김영호, 강대성
Young Ho Kim, Dae-Seong Kang

동아대학교 전자공학과
Dept. of Electronic Engineering, Dong-A University
e-mail : dskang@daunet.donga.ac.kr

요 약

최근 들어 디지털 기술의 발전은 문자, 음성, 화상, 비디오 등과 같은 멀티미디어 정보가 매우 큰 비중을 차지하고 있다. 이들 중에서 비디오와 관련된 연구로는 비디오 색인 및 검색에 관한 연구가 많이 진행되고 있다. 본 논문에서는 MPEG으로 압축된 비디오 분할을 위한 새로운 알고리즘을 제안한다. Shot 검출 즉 장면전환 검출은 MPEG 비디오 시퀀스에서 분할하는 가장 기본적이면서 중요한 작업이며 비디오 색인 및 검색을 위한 첫 번째 단계이다. 일반적으로 많이 사용되는 분할 알고리즘은 이전 프레임과 현재 프레임을 비교하기 때문에 물체의 빠른 움직임이나 카메라의 움직임, 카메라 후레쉬의 섬광 등 화면 변화에 따라 오검출이 생기는 단점이 있었다. 따라서 본 논문에서는 기존 알고리즘에 검출한 장면전환 지점을 사용하여 이웃 프레임들과의 특징차 비교를 통하여 한번 더 판별하였다. 실험 결과 기존 알고리즘 보다 정확한 장면전환 검출을 할 수 있었다.

Abstract

Recently, development of digital technology is occupying a large part of multimedia information like character, voice, image, video, etc. Research about video indexing and retrieval progresses especially in research relative to video. In this paper, we propose new algorithm(Frame Feature Comparison) for MPEG video segmentation. Shot, Scene Change detection is basic and important works that segment it in MPEG video sequence. Generally, the segmentation algorithm that uses much has defect that occurs an error detection according to a flash of camera, movement of camera and fast movement of an object, because of comparing former frames with present frames. Therefore, we distinguish a scene change one more time using a scene change point detected in the conventional algorithm through comparing its mean value with abutted frames. In the result, we could detect more corrective scene change than the conventional algorithm.

Key words : Frame Feature Comparison, MPEG Video Segmentation.

I. 서 론

최근 들어 디지털 기술의 발전은 문자, 음성, 화상, 비디오 등과 같은 멀티미디어 정보가 매우 큰 비중을 차지하고 있다. 이들 중에서 비디오 와 관련된 연구로는 비디오 색인 및 검색에 관한 연구가 많이 진행되고 있다.[1][2] 비디오 데이터는 매우 복잡한 계층 구조를 가지는데 이를 효과적으로 표현하기 위해서는 여러 단계의 계층으로 나눌 필요가 있다. 비디오 데이터의 가장 기본이 되는 단위는 프레임(frame)이며 그 다음 상위는 필름이 끊기지 않고 시간적으로나 공간적으로 연속해서 촬영된 일정구간의 프레임들로 구성된 Shot 이다. 비디오 데이터의 Shot 검출 즉, 장면전환 검출은 연속되는 두 개의 Shot 사이의 경계를 검출하는 것으로서 비디오 색인 및 검색의 첫 번째 단계로서 매우 중요한 역할을 한다. 최근 몇 년 동안 비디오 데이터의 장면전환 검출에 관한 연구가 많이 수행되어 왔다.[3][4][5] 대부분의 연구는 pixel 차에 의한 검출, 히스토그램을 이용한 검출, 에지를 이용한 검출, 비디오 형성 모델에 근거한 검출, 움직임 벡터를 이용한 검출등 비압축 영역에서의 검출 방법과 DCT의

DC를 이용한 검출, I, P, B를 이용한 검출, 비트열 간의 정보를 이용한 검출등 압축 영역에서의 검출 방법을 사용했다. 본 논문에서는 MPEG 비디오의 색인 및 검색을 위한 비디오의 최소 단위인 프레임들을 Shot 단위로 분류하기 위한 새로운 알고리즘을 제안하였다. 기존의 압축 영역에서 이용한 DCT의 DC를 이용하여 DC 영상을 구한 후, pixel간의 차분값, 히스토그램에 대한 chi-square 값, 히스토그램의 분산의 차분값, 양자화한 영상의 히스토그램의 각 bin 값들의 열과 행의 위치에 대한 분산의 chi-square 값을 이용하여 구해진 장면전환 프레임을 이웃 프레임간의 특징차 비교를 통하여 한번 더 판별하였다.

본 논문의 구성은 다음과 같다. Ⅰ장 서론에 이어 Ⅱ장은 배경이론 및 전체 알고리즘, Ⅲ장에서는 장면 전환 검출 알고리즘을 제안하고 Ⅳ장에서는 실험한 결과와 이에 대한 고찰을 하였고 Ⅴ장에서 결론을 맺었다

Ⅱ. 배경 이론 및 전체알고리즘

MPEG 비디오 시퀀스에서는 I(Intra), P(Predictive), B(Bidirectionally predictive) picture가 있다[6]. I picture는 inter frame 예측을 사용하지 않고 해당 화면 정보만으로 부호화하는 화면으로서 모든 MB(Macroblock) type은 intra frame이며, 채널 전환 시의 원영상 복구와 오류의 전파를 막기 위해 GOP(Group of Picture)내에 최저 한 장의 I picture를 필요로 한다. P picture는 I 혹은 P picture로 부터의 순방향 움직임 보상 예측 수행으로 생기는 화면으로서 MB type은 intra frame, forward inter frame, backward inter frame, interpolative inter frame 예측부호화가 수행된다. B picture가 삽입됨으로써 화면 처리 순서가 원 화면 순서와 달라져서 부호기에서는 B picture를 건너 뛰어 다음의 I, P picture를 우선 부호화하고, 그 후 사이에 있는 B picture를 부호화 한다. MPEG 비디오의 시퀀스층은 다수개의 GOP로 되어있다. 각 GOP는 다수개의 picture로, picture층은 다수개의 슬라이스로, 슬라이스층은 다수개의 MB로, MB는 휘도 신호 블록과 색차 신호 블록으로 구성되어 있다. 보통 한 GOP당 15개 이하의 picture들로 구성된다. 이때 인트라 부호화인 I picture에서 각 슬라이스를 추출하고 슬라이스의 MB에서 각 8 × 8 DCT영역의 DC성분을 추출한다. 이때 한 I picture당 1개의 DC영상이 생성된다.

본 논문에서는 I picture에 대해서 DCT DC 계수를 구하고 이를 정규화 시킨 후 DC영상을 구성하였다. 그림 1은 전체알고리즘 나타낸 것이다. 압축비디오에서 DC영상을 검출한 후 기존의 4개의 feature를 사용하여 구한 장면전환 프레임들을 제안알고리즘을 통해 한번 더 판별함으로서 개선된 장면전환 프레임들을 검출할 수 있다.

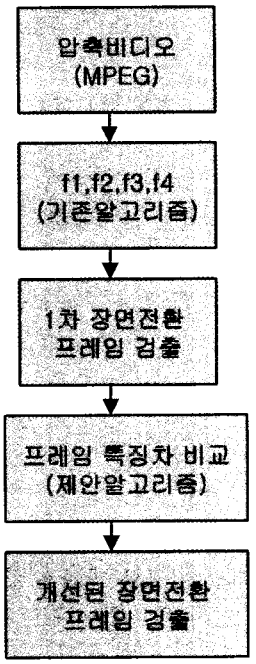


그림 1. 전체 알고리즘
Fig 1. A General Algorithm

Ⅲ. 장면 전환 검출 알고리즘

1. 기존의 장면 전환 검출 알고리즘[7][8]

MPEG 비디오 시퀀스의 I picture에서 DC영상을 추출하여 다음 4개의 feature들을 이용하여 장면 전환 검출을 하였다. 첫 번째 feature는 이전 영상과의 pixel간의 차분값으로서 서로 다른 영상의 유사도를 측정하는데 가장 기본이 되며, 전체적인 휘도 변화를 나타낸다. 이 feature를 구하는 수식은 아래와 같다.

$$DiffImg_i = \frac{\sum_{x=0,y=0}^{M-1,N-1} |I_{i-1}(x,y) - I_i(x,y)|}{NM} \quad (1)$$

여기서 M은 DC영상의 열의 개수이고, N은 행의 개수이다.

두 번째 feature는 일반적으로 사용되어 지고 있는 DC영상의 히스토그램에 대한 chi-square 값이다. 이전 영상에 대한 히스토그램의 변화를 나타낸다. 이 feature를 구하는 수식은 아래와 같다.

$$X_i^2 = \sum_{k=0}^{n-1} \frac{(H_{i-1}(k) - H_i(k))^2}{H_i(k)} \quad (2)$$

여기서, $H_i(k)$는 i 번째 DC영상 히스토그램의 k 번째 bin의 값을 나타낸다. n은 8bit 영상을 사용할 경우 256이다.

세 번째 feature는 이전 DC영상과 현재 DC영상과의 히스토그램 분산의 차분값이다. 이 feature는 히스토그램의 전체적인 분포에 대한 변화를 나타낸다. feature를 구하는 수식은 아래와 같다.

$$DiffD_i^2 = \frac{(D_{i-1} - D_i)^2}{D_i} \tag{3}$$

여기서, D_i는 i 번째 DC영상의 히스토그램 분산 값이다.

네 번째 feature는 양자화한 영상의 히스토그램의 각 bin값들의 열과 행의 위치에 대한 분산의 chi-square 값이다. 이 feature는 히스토그램의 bin 값들을 이용함으로써 객체의 움직임에 강인하고, 각 bin 값들의 열과 행의 위치에 대한 분산의 chi-square 값을 구함으로서 칼라의 변화에 둔감하다. 이 feature를 구하는 수식은 아래와 같다.

$$PX_i^2 = \sum_{k=0}^{n-1} \frac{(\rho X_{i-1}(k) - \rho X_i(k))^2}{\rho X_i(k)} \tag{4}$$

$$PY_i^2 = \sum_{k=0}^{n-1} \frac{(\rho Y_{i-1}(k) - \rho Y_i(k))^2}{\rho Y_i(k)} \tag{5}$$

$$\rho X_i(n)^2 = E(|BX_i(n) - \overline{BX_i(n)}|^2) \tag{6}$$

$$\rho Y_i(n)^2 = E(|BY_i(n) - \overline{BY_i(n)}|^2) \tag{7}$$

이상과 같이 구해진 DC영상의 feature들로부터 아래와 같은 단계로 장면전환 프레임을 검출한다.

Step 1. 각 파라메터들의 전체 프레임에 대한 평균을 구한다.

$\overline{DiffImg}$, $\overline{X}$, $\overline{DiffD}$, $\overline{PX}$, $\overline{PY}$

Step 2. DC영상의 각 feature 값이 아래의 조건식에서 condition 1과 2를 동시에 만족하면서 condition 3 혹은 4가 만족할 경우 장면전환 프레임으로 검출한다.

condition1. $DiffImg_i > \alpha\,\overline{DiffImg}$

condition 2. $X_i > \beta\,\overline{X}$

condition 3. $DiffD_i > \gamma\,\overline{DiffD}$

condition 4. $PX_i > \delta\,\overline{PX}$ and $PY_i > \delta\,\overline{PY}$

(※ 각 상수 $\alpha, \beta, \gamma, \delta$는 실험을 통해서 2.0, 2.0, 2.0, 2.0으로 검출)

2. 제안 알고리즘

I picture로부터 DC영상을 각각 추출한다. 기존의 4개의 feature를 순차적으로 비교하여 구한 두 DC영상이 장면 전환 지점인가를 판별해야한다. 이는 장면 전환 지점의 이웃 프레임들의 특징차의 평균을 장면 전환 지점의 특징량과 비교함으로써 이루어진다. 카메라 정지 등에 의해 장면 전환이 일어나는 지점은 급격한 화면의 변화가 일어나므로 높은 평균값을 갖는다. 하지만 장면 전환 내의 프레임 중에서도 물체의 빠른 움직임, 큰 물체의 움직임, 카메라의 움직임, 카메라의 후레쉬의 섬광 등은 프레임 내 화소값에 크게 영향을 주므로 DC영상도 많이 변하게 된다. 이때 기존 알고리즘은 위와 같이 큰 물체의 움직임 또는 물체의 빠른 움직임 등의 프레임을 장면 전환으로 오검출 할 수 있다. 따라서 제안한 알고리즘은 이를 해결하기 위해 기존에 장면 전환 지점으로 판별된 장면 전환 지점을 실제의 장면 전환 지점인지 아닌지 이웃 프레임의 DC영상과 비교하여 재판별한다. 실제 장면 전환이 일어나는 지점은 이전 프레임들과 이후 프레임들은 크게 차이가 난다. 하지만 장면 전환 지점을 기준으로 이전의 프레임들과 이후의 프레임들은 각각 비슷하다. 또 실제로 장면 전환 지점이 아닌 지점은 각 프레임의 특징차가 조금 크다 하더라도 이웃의 프레임들과 특징 부분을 제외하고는 비슷하다. 그러므로 기존의 두 비교 프레임의 전과 후의 프레임을 선정하여 총 네 개의 프레임을 이용해 장면 전환 지점의 이웃 프레임들과의 특징차의 평균을 구하여 장면 전환 지점의 특징차와 비교하여 실제 장면 전환 지점인지 아닌지 판별한다. 두 프레임이 아닌 한 프레임 전이나 한 프레임 후의 프레임과 특징차를 각각 비교하여 평균값을 취하면 해당 프레임들의 평균 특징차가 된다. 이때 이 평균값보다 장면 전환 지점의 특징차가 크다면 실제의 장면 전환 지점으로 판별한다.

$Shot_{i,i+1} \geq ShotM_{i,i+1}$: Shot 경계 지점

$Shot_{i,i+1} < ShotM_{i,i+1}$: Shot 경계 지점 아님

$Shot_{i-1,i+1}$와 $Shot_{i,i+2}$의 비교 측정치를 이용하여 $ShotM_{i,i+1}$을 구해서 $Shot_{i,i+1}$와 $ShotM_{i,i+1}$을 비교한다. $Shot_{i,i+1}$이 평균치, $ShotM_{i,i+1}$ 이상이라면 장면 전환 지점으로 판정한다.

Ⅳ. 실험결과 및 고찰

실험에서 사용된 대상 영상은 움직임이 강한 실험영상 1("likethat")과 카메라의 후레쉬의 섬광 효과가 많은 실험영상2("respect the Power of Love")를 사용하였다. 각각 100개의 프레임을 이용하여 실험하였고, 각각의 프레임들은 320 X 240 사이즈에 8비트로 구성이 되어있다. 본 논문에서의 실험은 MPEG 비디오의 시퀀스로부터 DC영상을 검출하여 일반적인 장면전환 검출에 사용되는

pixel간 차분값 feature와 히스토그램을 사용하는 feature, 히스토그램 분산의 차분값 feature, 히스토그램 분산과 히스토그램 bin의 위치에 따른 분산 값 feature에서 얻은 결과를 이웃 프레임들간의 특징차 비교를 통해 한번 더 판별함으로서 개선된 결과를 보였다. 그림 2,3은 기존의 4개의 feature를 이용하여 구한 장면 전환 프레임들의 결과이고, 그림 4,5는 프레임 특징차 비교를 이용해서 구한 장면전환 프레임들의 결과이다. 그림 2의 85개의 프레임 중에서 17, 20, 22, 24, 31, 32, 34, 45, 49, 52, 53, 55, 59, 60, 62, 63, 64, 67, 68, 69, 71, 75, 78, 81, 82, 83, 84번째 프레임에서 오검출이 발생하였다. 그림 3에서는 7, 8, 9, 11, 16, 19, 20, 21, 22, 23, 24, 29, 37, 38, 39, 40번째 프레임에서 오검출이 발생하였다. 제안알고리즘에서 구한 장면전환 프레임 그림 4에서는 21, 28, 45, 51, 54, 59번째 프레임에서 오검출이 발생하였고, 그림 5에서는 15, 16번째 프레임에서 오검출이 발생하였다. 표 1은 장면전환 프레임 검출 결과를 나타내었다.

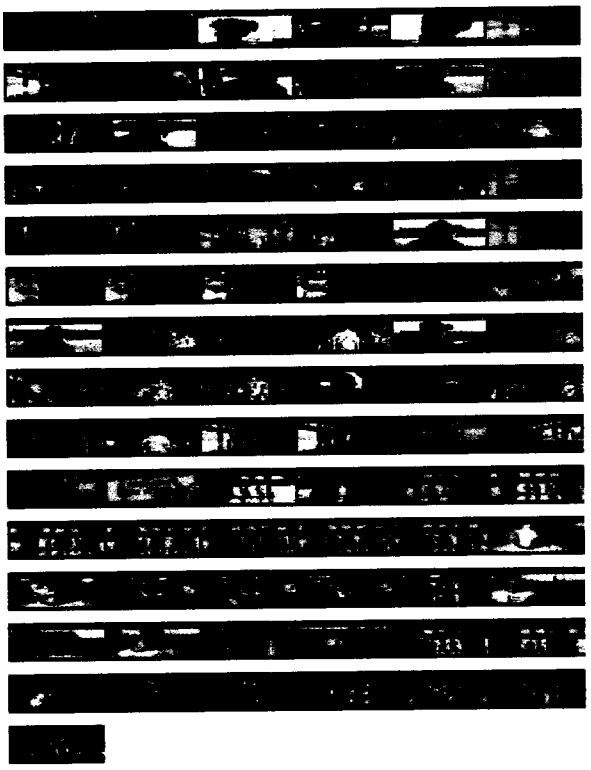


그림 2. 4개의 feature를 이용한 장면전환 프레임들

Fig 2. Scene change frames using 4 feature

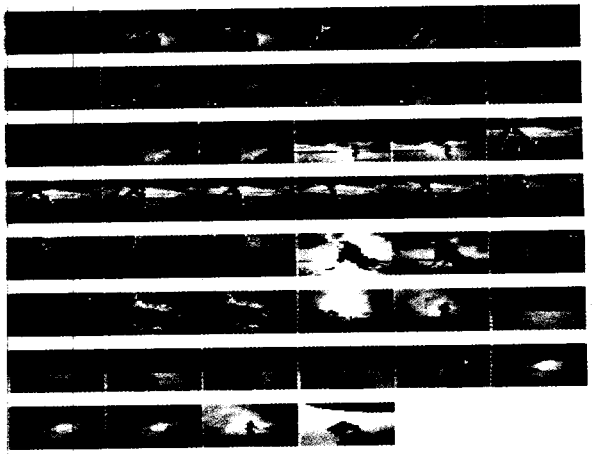


그림 3. 4개의 feature를 이용한 장면전환 프레임들

Fig 3. Scene change frames using 4 feature

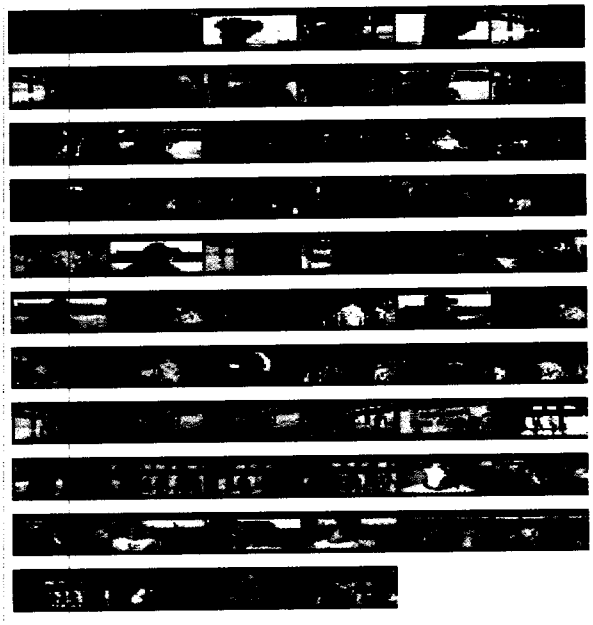


그림 4. 프레임 특징차 비교를 이용한 장면전환 프레임들

Fig 4. Scene change frames using Frame Feature Comparison

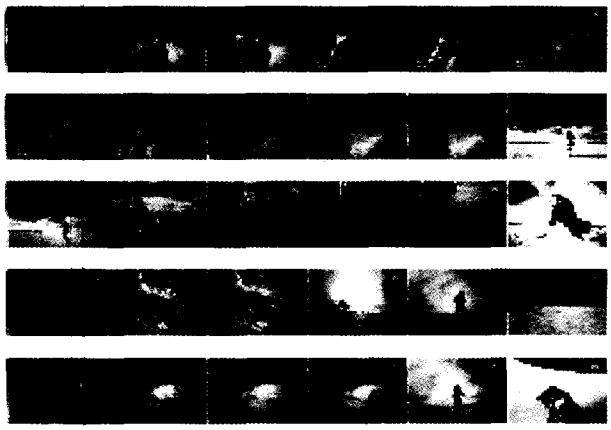


그림 5. 프레임 특징차 비교를 이용한 장면전환 프레임들

Fig 5. Scene change frames using Frame Feature Comparison

표 1. 장면전환 프레임 검출 결과

Table 1. Experiment results of scene change frame detection

영상1 / 영상2	장면전환 총개수	검출 개수	오 검출	오 검출율 (%)
기존 알고리즘	58/28	85/46	27/18	31.8/39.1
제안 알고리즘	58/28	64/30	6/2	9.3/6.7

V. 결론

실험영상1에서는 실제 장면전환이 발생하는 프레임 개수는 58개이고 실험영상2에서는 28개였다. 기존 알고리즘에서 검출한 장면 전환 프레임은 실험영상1에서는 85개, 실험영상2에서는 46개이고, 새로이 제안한 알고리즘에서는 실험영상1에서는 64개, 실험영상2에서는 30개였다. 오검출은 실험영상1에서는 27개에서 6개, 실험영상2에서는 18개에서 2개였다. 오검출율은 실험영상1에서는 31.8%에서 9.3%로 실험영상2에서는 39.1%에서 6.7%로 오검출을 줄일 수 있었다. 기존 연구에서는 물체의 빠른 움직임, 큰 물체의 움직임, 카메라의 움직임, 카메라의 후레쉬의 섬광 등은 프레임 내 화소값이 큰 영상에 대해서는 장면전환 검출의 오검출율이 높았다. 그리고 실험한 영상의 특징만을 위주로 실험한 결과 다른 영상을 실험할 경우 오검출율 또한 높았다. 따라서 본 연구에서는 기존 알고리즘에서 구한 장면전환 프레임들을 이용 새로운 알고리즘에 한번 더 적용함으로서 어느 정도의 오 검출율을 줄일 수 있었고 여러 종류의 실험영상에 대해서도 일단 검출되어진 프레임들을 한번 더 판별함으로서 여러 종류의 영상에 대해 강한 특징을 보였다. 향후 이러한 장면전환 프레임 검출로 구해진 shot 단위 프레임들을 이용해서 MPEG 비디오의 색인 및 검색에 이용하는 연구가 필요하다.

접수일자 : 2003. 2. 18 수정완료 : 2003. 3. 29

이 논문은 2002년도 정보통신 IT학과 장비지원사업의 동아대학교 대응자금에 의해 연구되었음.

참고문헌

[1] "Efficient Scene Change Detection and Camera Motion Annotation for Video Classification," Wei Xiong and Chung-Mong Lee, Computer Vision and Image Understanding Vol.71, Nos 2/2, pp. 166-181, August, 1998.

[2] "A hierarchical Multiresolution Video shot Transition Detection Scheme," Hong Heather Yu, Computer Vision and Image Understanding Vol.75, Nos 1/2, pp. 196-213, July/August, 1999

[3] S. W. Smith and H. J. Zhang, "Content-Based Video Indexing and Retrieval", IEEE Mulimedia, summer 1994, pp 62-72, 1994.

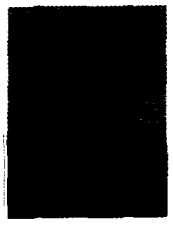
[4] Boon-Lock Yeo and Bed Liu, "Rapid Scene Analysis on Compressed Video", IEEE Transaction on Circuits and Systems for Video Technology, Vol. 5, No. 6, pp 533-544, 1995.

[5] Shih-Fu Chang, "Compressed-Domain Techniques for Image/Video Indexing and Manipulation", IEEE ICIP-95, Vol. 1, pp 314-317, 1995.

[6] 후지와라 히로시 "그림으로 보는 최신 MPEG," 교보문고, 2001

[7] John S. Boreczky and Lawrence A. Rowe, "Comparison of Video Shot Boundary Detection Techniques," Storage and Retrieval Image and Video Database IV, Proc. of IS&T/SPIE 1996 Symp. on Elec. Imaging: Science and Technology, February 1996

[8] Arun Hampapur, Ramesh Jain and Terry Weymouth, "Digital Video Segmentation," Proc. Second Annual ACM Multimedia Conference, October. 1994



김 영 호(Young Ho Kim)
正會員
1999년 2월 동아대학교 전자공학과 학사. 2002년 동아대학교 전자공학과 석사. 2003년 동아대학교 전자공학과 박사과정. 주관심분야 영상처리, 멀티미디어 색인 및 검색, 영상통신 등.

강 대 성(Dae-Seong Kang)
正會員
1984년 2월 경북대학교 전자공학과 학사. 1991년 8월 Texas A&M Univ., Electric Eng. 석사. 1994년 5월 Texas A&M Univ., Electric Eng. 박사. 1984년 3월 ~ 1989년 6월 국방과학연구소 연구원. 1994년 7월 ~ 1995년 8월 한국전자통신연구소 선임연구원. 1995년 9월 ~ 현재 동아대학교 전기전자컴퓨터 공학부 부교수. 주관심분야는 영상처리, 패턴인식, 영상코딩, 통신시스템 등.