

동적임계값을 이용한 인덱싱 알고리즘 (Indexing Algorithm Using Dynamic Threshold)

이 문 우* 박 종 운** 장 종 환***
(Moon-Woo Lee) (Jong-Woon Park) (Jong-Whan Jang)

요 약

압축된 동영상에서 인덱싱을 위한 장면전환 검출기법에서 기존의 방법들은 실험에 의한 고정 임계값을 설정하여 임계값 보다 크면 장면전환이라고 판단해왔다. 기존의 고정 임계값을 적용시켰을 때는 플래쉬나 카메라 움직임 등에 의한 오검출이 많은 문제점이 있었다. 본 논문에서는 장면 전환 검출을 위한 임계값을 동영상 특성 중, 장면전환점간격을 이용하여 임계값을 동적으로 변화시키는 방법이며, 고정 임계값을 사용하는 경우보다 오검출을 줄이는 향상된 장면전환 검출기법을 제안한다. 실험에서는 동영상 특성을 통계적으로 분석하여, 기존의 고정임계값과 제안한 동적임계값을 사용한 결과값을 비교분석 하였다. 제안한 방법은 기존의 방법보다 30%정도 오검출이 줄었다.

ABSTRACT

In detection of a scene change of the moving pictures which has massive information capacity, the temporal sampling method has a faster searching speed and lower missing scene change detection than the sequential searching method for the whole moving pictures, yet employed searching algorithm and detection interval greatly affect missing frame and searching precision. In this study, the whole moving pictures were primarily retrieved threshold by the temporal difference of histogram scene change detection method. We suggested a dynamic threshold algorithm using cut detection interval and derived an equation formula to determine optimal primary retrieval threshold which can cut detection interval computation. Experimental results show that the proposed dynamic threshold algorithm using cut detection interval method works up about 30 percent in precision of performance than the sequential searching method.

1. 서론

현대는 다양한 미디어의 데이터를 사용하고있지만, 이 중, 가장 효과적인 정보전달 매체는 동영상이다. 멀티미디어 정보를 효과적으로 사용 또는 서비스하기 위해서는 인덱싱이 필요하게 된다. 이를 위해 내용기반으로 검색, 저장하여 볼 수 있도록 하는 기술이 필요하다.

동영상 정보로부터 내용기반 정보검색을 하기 위한 중요한 기술의 하나는 video indexing이다[1-4]. Nagasaka[2]는 내용기반 정보검색을 위한 video indexing을 하기 위해 히스토그램 차, 필셀차를 이용하는 방법 등 다양한 방법이 연구 되고있다.

* 종신회원 : 서울정수기능대학 정보통신과 전임강사
** 종신회원 : 육군 전술 C41 개발단 검증장교
*** 종신회원 : 배재대학교 컴퓨터전자정보공학과 교수

논문접수 : 2001. 3. 27.
심사완료 : 2001. 4. 12.

인텍싱을 위한 여러 방법 중 장면전환 검출 기법에는 픽셀차, 히스토그램 차등 여러 방법이 있지만 기본적인 개념은 임계값을 설정하여 전화면과 현재 화면의 특징을 비교하여 임계값 이상이면 장면전환으로 판단하는 것이다. 고속검색과 정확한 검색을 위한 연구가 활발하다. 고속 검색을 위해 수직과 수평 블록을 이용한 검출방법[5]과 이진검색 알고리즘[6] 등은 검색 속도 면에서 뛰어 나다. 그리고 최근에는 압축된 비트스트림이 갖고 있는 내용을 직접 분석하여 장면변환에 필요한 의미 있는 특징 정보를 추출하여 동영상의 내용기반 장면변환 및 검색에 활용하려고 노력하고 있다[7-9]. 그리고 압축 데이터에서 블록모션 특성을 이용하여 장면전환 검출을 하는 기법도 있다[10]. 그러나 이런 방법들은 검색 속도는 향상되지만 고정임계값의 오검출에 대한 한계를 극복하지 못한다. 이러한 장면전환검출 방법들의 기본 원리는 임계값을 사용하는 것이다. 컷 검출을 하기 위해 프레임간의 특징 값의 차이를 구해 이 값이 미리 정해진 임계값보다 클 경우 장면전환이 일어난 것으로 판단하게 된다. 따라서 정확한 컷을 검출하기 위해서는 적절한 임계값이 선택 되어야 한다. 기존의 컷 검출 방식에서는 실험을 통한 적절한 고정된 임계값을 사용해 왔다. 그러나 여러 종류의 동영상에 대하여 고정된 임계값을 사용하는 것은 문제가 있다. 그 이유는 히스토그램 분포도와 같은 특징 값들이 동영상마다 다르기 때문이다. 또한 하나의 동영상이라도 많은 장면과 색상의 변화가 있기 때문에 같은 임계값을 지속적으로 적용하는 것은 문제가 될 수 있다. 이러한 문제점을 해결하고 장면전환 검출의 효율성을 향상시키기 위한 방법으로 장면전환 검출 과정 검출간격을 고려하여 임계값이 동적으로 변화하는 동적 임계값 기법을 제안하였다. 동적 임계값 기법은 기존의 임계값을 사용하는 대부분의 장면전환 검출 기법에 적용하여 사용할 수 있다.

본 논문에서는 컷 검출 횟수와 검출간격으로부터 동적 임계값을 추출하는 알고리즘을 제시하였다. 그리고 고정임계값과 동적 임계값을 적용하여 압축된 동영상 파일 중에서 광고, 뉴스, 드라마를 실험하여 검색횟수와 오검출 횟수 등을 비교하였다.

제2장에서는 컷 검출 횟수와 검출간격을 이용하여 동적 임계값을 추출하는 알고리즘을 제시하였다. 제3장에서는 고정임계값과 제안한 동적 임계값 추출방

법을 적용한 실험 및 결과를 기술하며 제 4장에서 결론 및 향후 과제를 제시하였다.

2. 제안된 동적 임계값 설정 알고리즘

2.1 히스토그램 비교법과 임계값 설정

동영상의 장면변환 검출을 위해 위에서 언급한 바와 같이 여러 가지 방법들이 제안되어 있다. 이들 중 구현하기 간단하고 널리 사용되는 프레임간의 히스토그램 차이를 이용하는 방법이 식(1)에 주어진다.

$$\sum_{j=0}^n |H_i(j) - H_{i+1}(j)| > T \quad (1)$$

여기서, n 은 그레이 레벨의 개수, i 는 프레임 번호, j 는 그레이 레벨, $H_i(j)$ 는 i 번째 프레임의 j 번째 그레이 레벨에서의 히스토그램 값이다. T 는 장면변환의 고정 임계값이다. 이 값이 적절한 임계값 T 보다 크면 장면변환이 있는 것으로 결정한다. T 값은 정확한 컷을 검출하기 위해서는 적절한 임계값을 선택해야 한다.

기존의 컷 검출 방식에서는 실험을 통한 적절한 고정된 임계값을 사용해 왔다. 고정임계값 사용은 다양한 내용의 동영상에 대하여 같은 값으로 고정시킨다는 것은 문제가 될 수 있다. 하나의 같은 동영상이라도 많은 장면과 색상의 변화가 있기 때문에 특정 고정 임계값을 적용하게 되면 명암값의 변화에 의하여 오검출이 증가할 수 있다.

2.2 동적 임계값 설정

동영상의 특성은 몇 초 가량의 광고부터 한시간 이상 되는 뉴스 및 드라마 등 다양하다. 그러나 광고, 뉴스, 드라마등 각 종류의 동영상들은 영상의 특성상 각각 일정한 프레임 간격으로 장면 전환이 일어난을 통계적으로 알 수 있다. 그러나 광고나 뉴스 등 각 특징과 총 프레임 수와 무관한 일반적인 규칙은 장면전환 평균 간격이다. 일반적으로 평균 장면전환 프레임 간격은 총 프레임 수를 N , 총 장면 전

환 검출 횟수를 K 라고 할 때, 식 (2)로 나타낸다.

$$D = \lim_{N \rightarrow \infty} \frac{N}{K} \quad (2)$$

광고의 빠른 장면전환이 이루어지며, 뉴스의 경우 기사의 전달로 인하여 장면전환이 광고에 비하여 천천히 일어남을 볼 수 있다. 장면 변환의 경우는 표 1 과 같이 총 프레임 수에 상관없이 광고나 뉴스 등 같은 종류의 동영상에 평균 장면전환간격은 일정 프레임 구간에 존재함을 볼 수 있다.

<표 1>평균 검출 간격

<Table 1> Average of Cut Detection Interval

a. 광고

	N	K	D
광고1	449	10	44.90
광고2	669	16	41.81
광고3	440	12	36.67
광고4	449	12	37.40
평균			40.20

b. 뉴스

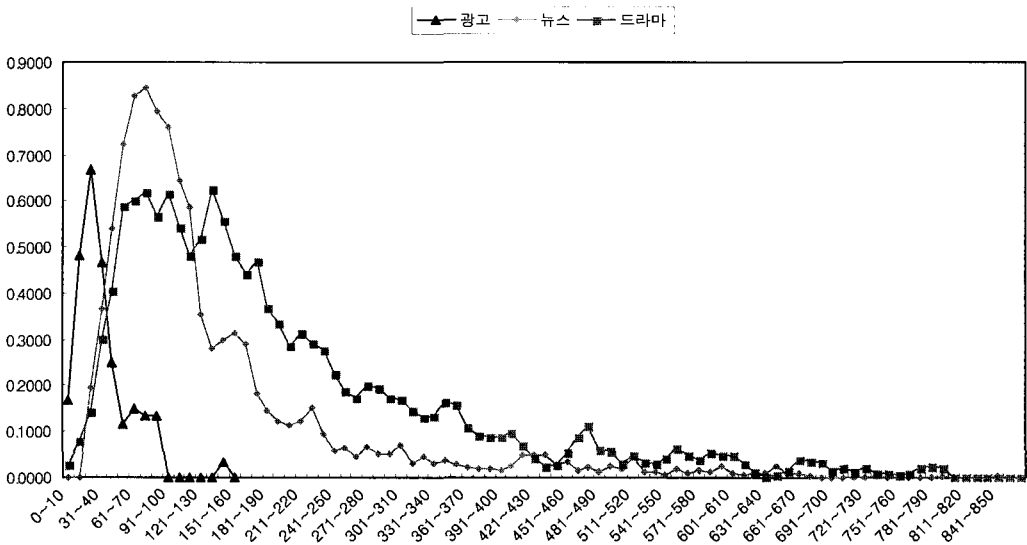
	N	K	D
뉴스1	85530	635	134.69
뉴스2	9060	68	133.23
뉴스3	20400	156	130.76
평균			132.89

c. 드라마

	N	K	D
드라마1	45297	346	130.9
드라마2	39378	309	127.4
드라마3	33084	167	198.1
평균			132.89

[그림 1]은 동영상의 특징에 따라 컷이 검출되는 간격을 통계적 조사로 나타낸 분포도이다. 장면전환 검출이 일어날 확률은 장면 전환이 일어난 다음 프레임보다는 프레임이 많이 지나갈수록 크다.

본 논문에서는 동영상에서 장면전환 검출이 된 직후에는 장면전환 검출이 되지 않다가 일정한 프레임이 지나가야 장면전환 검출이 나타나는 통계틀 이



[그림 1] 동영상 특성에 따른 컷 검출 분포

[Fig. 1] Distribution of Cut Detection in MPEG Video

용하였다. 장면전환이 검출된 직후에는 임계값을 높였다가 장면이 경과할 수록, 임계값이 낮아지게 하는 방법으로 동적 임계값을 정의하였다.

장면전환이 일어난 직후에 임계값이 $Th_{high(n,d)}$ 이 되며, $Th_{(n,d)} \geq Th_{ini}$ 라는 조건을 만족하는 임계값을 구하면 동적 임계값 Th 는 식 (3)으로 정의된다.

$$Th_{(n,d)} = Th_{ini} + (Th_{high(n,d)} - Th_{ini}) \cdot e^{-\frac{n}{d}} \quad (3)$$

여기서, 최저 임계값인 Th_{ini} 은 일반적으로 우리가 실험에서 적용하는 고정임계값을 적용해 주었다. n 은 컷과 컷 사이의 프레임 수, d 는 동영상특성에 따라 동적으로 변화되는 실시간 장면전환 검출 지점간 평균 간격 값(프레임수)이다.

$Th_{high(n,d)}$ 는 장면전환이 발생한 프레임과 전프레임의 히스토그램 그레이레벨 차 값의 평균이다. 히스토그램 평균 값은 식 (1)을 적용하여 장면 전환이 된 부분의 합과 검출된 개수를 적용하면 식 (4)와 같이 정의된다.

$$Th_{high} = \frac{\sum_{j=b}^a |H_{cut}(j) - H_{cut+1}(j)|}{k} \quad (4)$$

여기서, $H_{cut}(j)$ 는 식 (1)을 적용하여 장면전환 검출이 된 프레임의 히스토그램 차이 값이고 k 는 장면전환이 검출된 프레임 개수이다. 동영상의 내용기반에 동적으로 적용하기 위하여 3개의 장면전환 검출값을 사용하여 사용한다. 뉴스의 앵커 프레임같이 긴 것을 평균하여 $Th_{high(n,d)}$ 를 결정한다는 것은 평균값 자체의 의미를 잃게 되어 내용기반에 적용적으로 대처할 수 없게 된다. 이 문제점을 극복하기 위해 3개의 값 중 편차가 가장 큰 것을 평균산출에 제외시켜 사용한다.

d 는 평균 장면전환 검출 간격이며, 장면전환 검출을 하고 다음 장면전환 검출이 발생 될 때까지의 프레임 수에서 장면전환 검출을 한 회수와의 관계식은 식 (5)같은 방법으로 구해진다.

$$D_{avg} = \frac{\sum_{i=0}^N |d_{cuti} - d_{cuti+1}|}{k} \quad (5)$$

여기서, i 는 프레임 번호이며, d_{cuti} 컷이 일어난 프레임 번호의 간격을 나타내며, k 는 장면전환이 검출된 프레임 개수이다. 일반적인 동영상의 특성별 평균 장면전환 지점간 거리는, 각 특성별 동영상에서 검출횟수와 검출된 프레임간의 간격을 통계 조사에 의해 구해지며, 광고와 뮤직비디오는 12프레임, 뉴스 32프레임, 드라마 48 프레임마다 장면전환 검출이 된다.

3. 실험결과

3.1 실험환경

실험 환경은 윈도우2000, Visual C++, 개발된 MPEG라이브러리(배재대학) 사용하였다. 실험대상 동영상은 352x240, 30frame/sec의 MPEG 영상이다. 뉴스 영상, 15초 가량의 광고 영상 그리고 드라마를 선정하였다.

동적 임계값 기법을 조사하기 위해 본 연구에서는 히스토그램 컷 검출 기법을 사용하였다. 히스토그램 컷 검출 기법은 연속적인 두 영상 사이의 히스토그램 변화를 측정하여 변화 값이 주어진 임계값보다 큰 경우 컷으로 판단하는 방법으로 잡음이나 바기의 미세한 변화와 카메라 움직임 등에 비교적 강한 장점을 갖고 있으며 가장 보편적인 방법이라서 선택하였다. 식(4)에서 $Th_{high(n,d)}$ 는 프레임 전체를 평균하면 내용기반에 적용하는 어려움이 있어 최근 3개를 취하여 평균으로 사용하였다. 식(5)에서 평균간격은 뉴스 앵커 프레임과 같이 편차가 큰 경우 미검출의 문제가 발생할 수 있어 내용기반에 적용적일 수 없었으며, 실험에서는 장면전환 간격을 최근 3개의 간격을 사용하였다.

3.2 실험 및 결과

실험은 같은 동영상에 대하여, 기존의 고정 임계값과 제안한 방법인 동적 임계값을 사용했을 때의 결과 값인 장면전환 검출 수, 오검출 수, 미검출 수를 비교 평가하고, 제안한 알고리즘에 대한 성능을 분석하였다. 실험에 데이터로 사용된 MPEG데이터는

직접 사람의 눈에 의해 정확한 컷의 위치를 발견하여, 실제 장면전환 개수로 했다.

고정 임계값과 초기 임계값 Th_{ini} 는 일반적으로 많이 사용되고 있는 값을 임계값으로 설정하였다. 특수 영상 효과인 점진적 장면전환, 디졸브(dissolve), 페이드인/아웃(fade-in/out)등과 카메라 움직임에 의한 팬(pan), 줌(zoom), 스크롤(scroll)등에 대해서는 고려하지 않고 급작스런 장면전환에 대해서만 조사하였다.

<표 2>의 컷 횟수는 시간적으로 장면전환점의 수를 조사한 값이다. 총 검출횟수는 방법에 따라 실험하여 컷 검출된 수를 나타낸다. 실험 결과 실제 값을 비교해 보면, 고정 임계값을 적용한 경우보다 동적 임계값을 적용했을 때 오검출 수와 미검출 수가 많이 줄어들게 되는 것을 볼 수 있다. [그림 2]는 고정임계값과 동적임계값의 장면전환 검출 특성을 나타낸다.

고정임계값을 사용한 기법은 플래시(flash)부분과 고정임계값에 따라 오검출 수나, 미검출 수가 증가되는 문제점이 있다. 제안방법에서는 장면전환 검출이 일어난 직후 임계값이 높아져 오검출 수가 줄어들게 되고, 임계값이 점점 낮아져 장면전환 검출이 예측되는 지점에 동적 임계값이 낮아져 미검출 수를 줄일 수 있었다. 그림-3은 드라마(a), 뉴스(b), 광고(c), 스포츠(d) 비디오의 일반적인 장면전환 검출 간

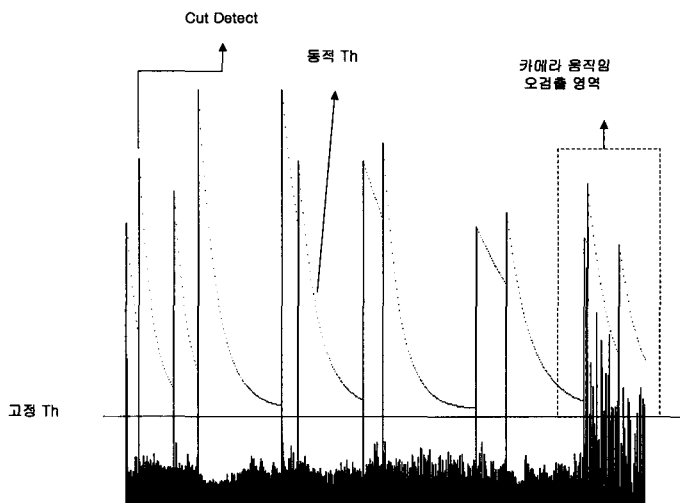
격을 알 수 있다. X축은 MPEG 비트스트림의 프레임번호 이고, Y축은 현재 프레임과 전프레임의 히스토그램 차이값을 나타낸다. 광고, 스포츠와 같이 움직임이 많은 비디오 데이터의 경우에 일반적인 고정 임계값으로 장면전환 검출을 할 때 보다 오검출 부분이 많이 나타난다. 그림-3에서 뉴스(b), 광고(c), 스포츠(d)의 점선의 사각형은 히스토그램 차 값이 고정 임계값보다 큰 경우이며, 이 때 일반적인 히스토그램 비교법을 적용하면 오검출이 증가하는 부분이다.

<표 2> 고정임계값과 동적 임계값 검출 결과

<Table 2> Results of Comparison with Thresholds

a. 광고

동영상	광고 1		광고 2		광고 2	
컷횟수	10		14		8	
기법	고정 임계값	제안된 방법	고정 임계값	제안된 방법	고정 임계값	제안된 방법
총 검출 횟수	21	9	18	12	10	9
컷 검출	10	9	12	12	8	8
오검출	11	0	6	0	2	1
미검출	0	1	2	2	0	0

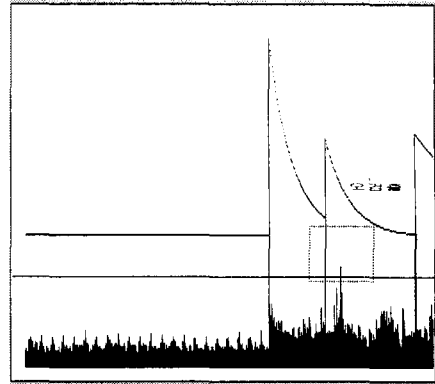


[그림 2] 고정임계값과 동적임계값 비교

[Fig. 2] Comparison of Global Threshold and Dynamic Threshold

b. 뉴스

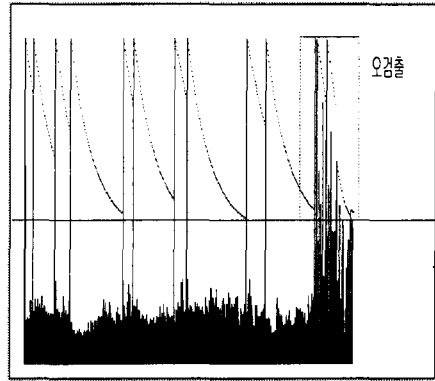
동영상	뉴스1 (40분)		뉴스2 (15분)	
컷횟수	626		150	
임계값 기법	고정 임계값	동적 임계값	고정 임계값	동적 임계값
총검출 횟수	1184	640	263	157
컷검출	613	618	147	149
오검출	571	16	116	7
미검출	13	6	3	1



(b) 뉴스

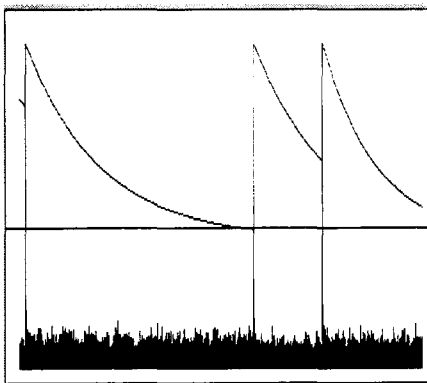
c. 드라마

동영상	드라마1		드라마2	
컷횟수	364		157	
임계값 기법	고정 임계값	동적 임계값	고정 임계값	동적 임계값
총검출 횟수	427	368	165	150
컷검출	359	363	156	150
오검출	68	4	9	0
미검출	5	1	1	0

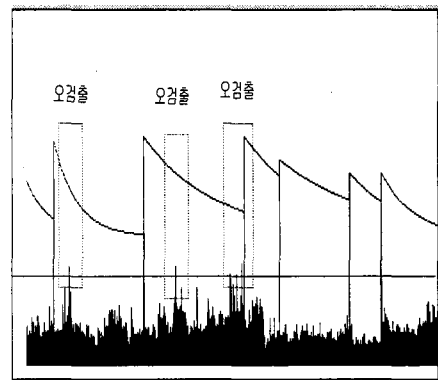


(c) 광고

실험 결과 동적 임계값을 적용하여 장면전환을 검출할 경우 고정임계값을 적용했을 때보다 후래쉬나 카메라의 움직임에 강건함을 보였고, 총 장면전환 검출 횟수와 잘못된 검출 횟수가 30% 줄어드는 결과를 얻을 수 있었다.



(a) 드라마



(d) 스포츠

[그림 3] MPEG 특성에 따른 오검출
[Fig. 3] Error Detection in MPEG Video

4. 결론

장면전환 검출 기법에서 임계값은 장면전환부분을 찾는 중요한 부분이면서도 제안된 특별한 방법이 없이 장면에 대한 전체적인 특징 값을 조사해서 임의로 그 값을 정하여 적용하여 왔다. 그러나 이렇게 정한 고정임계값의 우 장면전환을 검출할 때 움직임의 변화, 카메라의 움직임 등에 의해 장면변환으로 잘못 검출하는 경우가 많다. 오검출 수나 미검출 수를 줄이는 결과이지만, 제안한 방법은 카메라의 특수효과에서 대해서는 아직 명확하게 문제점을 풀수 없는 문제점이 있었다. 그러나 특수 효과 중 카메라의 이동 및 플래시(flash) 부분에서 강건함이 나타났다. 점진적인 장면전환, 디졸브(dissolve), 페이드인/아웃(fade-in/out) 등에서는 고정임계값을 사용했을 경우와 대등하게 나타났다.

향후 연구과제는 동적 임계값을 적용했을 경우에 일어나는 문제점을 해결해야되고, 방대한 동영상의 인덱싱을 위해 고속검색 방법을 고려한 알고리즘 개발이 과제이다.

제안한 방법은 픽셀차, 히스토그램 차 등 장면전환 검출기법 적용 가능할 것으로 기대된다.

※ 참고문헌

[1] M. Flickner et al. Query by image and video content: The QBIB system, *IEEE Computer*, 28(9): 23-32, September, 1995.

[2] A. Nagasaka and Y. Tanaka, Automatic Video Indexing and Full Motion Search for Object Appearance, *Proc. Of IFIP on Visual Database System*, pp.113-127, Sep. 1991.

[3] V. Ogle and M. Stonebraker, Chabot: Retrieval from a relational database of images, *IEEE Computer*, 28(9): 40-48, September 1995.

[4] A. Pentland, R. Picard, and S. Sclaroff, Photobook: Content-based manipulation of image database, *International Journal of Computer Vision*, 18(3): 233-254, June 1996.

[5] 이민섭, 안병철, "수평과 수직 블록을 이용한 MPEG-I 비디오 장면전환 검출", *한국정보처리학회 논문지*, 제7권, 제2호, pp.629-630, 2. 2000.

[6] 김성철, 오일균, 장중환, "동영상의 고속 장면분할을 위한 이진검색 알고리즘" *한국정보처리학회 논문지*, 제7권, 제4호, pp. 1044-1049, 4. 2000.

[7] V. Kobla and D. Doerman, *Compressed Domain Video Indexing Techniques Using DCT and Motion Vector Information in MPEG Video*, SPIE, Vol. 3022, pp.200-211, 1997.

[8] Jianhao Meng, Y. Juan, S. F. Chang, Scene Chang Detection in a MPEG Compressed Video Sequences, *Proc. of SPIE*, Vol. 2419, pp.14-25, 1995.

[9] Boon-Lock Yeo and Bede Liu, Rapid Scene Analysis on Compressed Video, *IEEE Transaction on Circuit and System for Video Technology*, Vol. 5, No. 6, pp. 533-544, December 1995.

[10] Ullas Gargi, Rangachar Kasturi, and Susan H. Strayer, "Performance Characterization of Video-Shot-Change Detection Methods", *IEEE Transaction on Circuit and System for Video Technology*, Vol. 10, No. 1, pp. 1-13, February 2000 .

이문우



1989년~대전산업대학교
전자공학과 졸업(학사)
1993년~승실대학교 대학원
전자공학과 졸업 (공학석사)
1999년~현재 배재대학교
정보통신공학과 박사과정
1983년~1998 한국전자통신
연구원 선임기술원
1998년~2000년 가톨릭상지대학
전기전자계열 초빙전임강사
2000년~현재 서울정수기능대학
정보통신과 전임강사
관심분야 : 멀티미디어검색,
컴퓨터네트워크, 영상처리

장종환



1979년~한양대학교 공과대학
전자통신공학과 졸 (학사)
1986년~미국 North Carolina
주립대학, 전기 및
컴퓨터공학과 졸(석사)
1990년~미국 North Carolina
주립대학, 전기 및
컴퓨터공학과 졸 (박사)
1990년~현재 배재대학교
컴퓨터전자정보공학과 교수
2001년~현재
한국소프트웨어진흥원
전문위원
2000년~현재 배재대학교
공과대학장
2000년~현재 한국컴퓨터산업
교육학회 부회장
2000년~현재
과학기술특허포럼 이사
2000년~현재 충청체신청
고객대표자 운영위원회 의장
2000년~현재 대덕실무기획단
의원
1999년~현재 대전광역시
최고정보화 책임관 (CIO)
1999년~현재 대전광역시
과학기술위원회 전문위원
1998년~현재 정보통신부 지정
정보통신창업지원센터 소장
1992년~1995년 배재대학교
전자계산소 소장
관심분야: 멀티미디어 콘텐츠,
멀티미디어정보처리

박종운



1998년~우송대학교
컴퓨터과학과 졸 (학사)
2000년~배재대학교
정보통신대학원
정보통신학과 (석사)
2001년~배재대학교 대학원
정보통신공학과 박사과정
2000년~현재 육군 전술 C4I
개발단 검증장교
1996년~1999년 육군전산소
제도분석 장교 및
전산망운영과장
관심분야 : 멀티미디어 콘텐츠,
멀티미디어정보처리