

개선된 특징차 비교 방법을 이용한 컷 검출 알고리즘에 관한 연구

정회원 최인호*, 이대영**

Region-based H.263 Video Codec with Effective Rate Control Algorithm for Low VBR Video

In-ho Choi*, Dae-young Lee** *Regular Members*

요 약

동영상 데이터를 내용기반 검색을 하기 위해 비디오 시퀀스를 계층적 분할해야 한다. 컷 검출 알고리즘은 샷을 분류하는 중요한 처리이다. 일반적인 컷 검출 알고리즘은 이전 프레임과 현재 프레임을 비교하기 때문에 카메라의 움직임이나 물체의 움직임에 의해 화면변화가 크면 컷을 오검출 할 수도 있다. 본 논문에서는 컷 예상지점을 먼저 선정하고 컷 예상지점이 실제의 컷인가를 판별한다. 컷 예상지점의 프레임과 이웃 프레임의 특징차를 비교하므로써 실제 컷을 검출하므로 기존 알고리즘보다 오검출을 줄일 수 있다. MPEG비디오 시퀀스에서 DC영상을 추출하여 비교 프레임으로 사용하였다. 실험 결과, 제안한 알고리즘이 기존의 알고리즘보다 컷 검출율이 높은 것을 확인하였다.

ABSTRACT

Video sequence should be hierachically classified for the content-based retrieval. Cut detection algorithm is an essential process to classify shots. It is generally difficult for cut detection algorithms to detect cut points since a current frame is compared with a previous one, because movement of camera or object made abrupt scene change. We reduce ratio of failed cut detection so that compare the difference between frames of predicted cut point and their neighbors.

In this paper, first we get predicted cut point, then we judge that the predicted cut point is true point or not. And we extracted DC images in MPEG video sequence for comparison. As a result of experiments, We confirmed that the cut detection ratio of the proposed algorithm is higher than that of any other algorithms.

I. 서론

최근 들어 주문형 비디오(VOD), 디지털 라이브러리, 전자상거래, 의료 비디오 검색 시스템, 방송프로그램 편집 시스템 등 멀티 미디어 서비스 요구가 증대되어 가고 있다. 이는 방대한 데이터베이스 시스템 구축을 필요로 하는데 문서 이외 특히 비디오

의 색인 검색은 매우 어려운 일이다. 이에 최근 비디오 데이터의 내용을 분석하여 인덱스를 구성하고 질의를 주는 내용기반 검색에 대한 연구가 활발히 진행중이다^{[1][2][3]}.

특히 동영상 데이터에 대한 내용기반검색은 MPEG-7등에서도 제안되는 등 큰 관심사이다.

* 김포대학 전자정보계열 조교수
논문번호 : 99323-0814, 접수일자 : 1999년 8월 14일

** 경희대학교 전자공학과 교수

동영상 데이터를 내용기반의 검색을 하기 위해 제일 먼저 동영상데이터를 계층적으로 내용별 분할해야 한다. 그리고 분할된 동영상 데이터에 색인을 다는 절차가 필요하다.

내용별 분할의 가장 중요한 처리는 연속 비디오 시퀀스에서 한 장면에서 다른 장면으로 변환되는 점, 즉 컷(cut)을 검출하는 것이다.

컷 검출 알고리즘은 처리 대상관점에서 비압축된 비디오 데이터와 압축된 비디오 데이터의 두 가지로 대별된다.

비압축된 비디오 데이터의 컷 검출방법은 처리단위 관점에서 화소 단위에 의한 검출방법, 부분영역 단위에 의한 검출방법 및 프레임단위에 의한 검출방법 등으로 나뉜다. 이중 영역별 검출 방법이 비교적 우수하다^[4]. 비압축된 비디오 데이터의 처리는 현재 저장과 통신을 위한 비디오 데이터가 압축되어 있는 상태가 대부분이어서 압축을 신장해야 하는 부가 절차가 필요하므로 미래지향적으로 볼 때 비효율적이라 하겠다.

압축된 비디오 데이터의 컷 검출방법은 MJPEG, MPEG, Wavelet, VQ 등의 압축종류에 따라 검출방법이 틀리며 압축된 비디오 특성을 살려 압축을 신장하지 않고 압축된 비디오 시퀀스에서 직접 특징을 추출하여 컷을 검출한다.

본 논문에서는 MPEG으로 압축된 비디오 시퀀스에 대하여 컷을 검출한다. 먼저 압축된 비디오 시퀀스에서 DCT성분 중 DC성분만을 추출하여 해당 픽처(picture)의 DC영상을 만든다.

이 DC영상에 대하여 각 프레임을 비교하되 기존에는 앞, 뒤 프레임의 특징 차로써만 컷을 검출하지만 본 논문에서는 컷 예상지점의 프레임과 두개의 이웃 프레임의 특징차를 비교하여 보다 정확한 컷 검출을 수행 한다. 카메라의 움직임과 대상 물체의 움직임에 의한 오검출에 강하기 때문에 컷 검출율을 높일 수 있다.

본 논문의 구성은 다음과 같다.

I장 서론에 이어 II장은 MPEG 비디오 시퀀스로부터 DC영상을 추출하는 방법에 대해 설명하고 III장에서는 추출된 DC영상을 이용하여 장면 전환을 검출하는 새로운 비교 방법에 대해 설명 한다. IV장에서는 실험한 결과와 이에 대한 고찰을 하였고 V장에서 결론을 맺었다.

II. MPEG 비디오 시퀀스로부터 DC영상 추출

2-1 MPEG 비디오 시퀀스의 계층적 구조

MPEG 비디오 영상은 6개의 계층으로 분류 되는데 그것은 시퀀스(sequence)층, GOP(Group of Picture)층, 픽처(picture)층, 슬라이스(slice)층, 매크로블럭(macro block)층, 블럭(block)층이다^[5]. 시퀀스층은 다수개의 GOP로 되어있다. 또 각 GOP는 그림 1과 같이 다수개의 픽처로, 픽처층은 다수개의 슬라이스로, 슬라이스층은 다수개의 매크로 블럭으로, 매크로 블럭은 휘도 신호 블럭과 색차 신호 블럭으로 구성되어 있다.

GOP층의 각 픽처는 크게 세종류로 나뉘어지는데, 프레임내 정보만을 보호화하는 프레임내 예측부호화인 I픽처와 프레임간 단 방향 예측부호화인 P픽처, 프레임간 쌍방향 예측부호화인 B픽처로 나뉜다. 보통 한 GOP당 15개 이하의 픽처들로 구성된다.

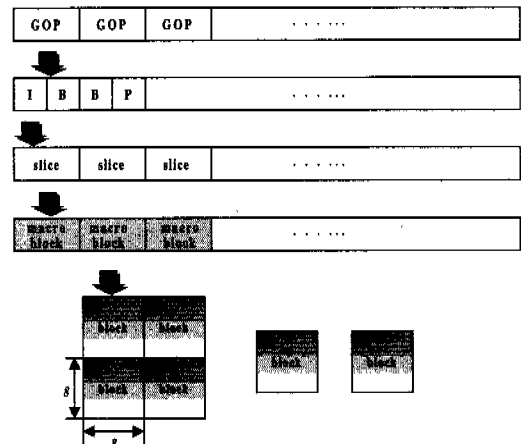


그림 1. MPEG 비디오 시퀀스의 계층적 구조

이때 인트라 부호화인 I픽처에서 각 슬라이스를 추출하고 슬라이스에 있는 매크로블럭에서 각 8 × 8 DCT영역의 DC성분을 추출한다. 이때 한 I 픽처당 1개의 DC영상이 생성된다.

DCT 변환 계수 중 8 × 8 블럭 당 변환 좌표 (0,0)의 값이 DC 계수이므로 다음 식을 이용해서 이를 추출한다.

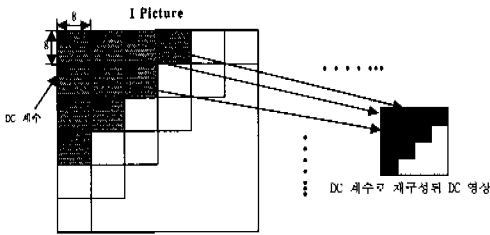


그림 2. DC 영상으로 재구성

$$G(u,v) = \frac{2c(u)c(v)}{N} \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} g(n,m) \cos \frac{(2n+1)u\pi}{2N} \cos \frac{(2m+1)v\pi}{2N}$$

$$g(n,m) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} c(u)c(v)G(u,v) \cos \frac{(2n+1)u\pi}{2N} \cos \frac{(2m+1)v\pi}{2N}$$

$$c(u), c(v) = \begin{cases} \frac{1}{\sqrt{2}} & (u,v=0) \\ 1 & (u,v=1,2,\dots,N-1) \\ 0 & (etc) \end{cases} \quad (1)$$

이때 8×8 블록의 좌상귀의 값인 DC계수는 앞 블록과의 DC 차분치이므로 앞 블록의 최초의 계수에 차분치를 더해가면 각 블록의 DC값이 생성된다. 보통 TV 프로그램 등의 비디오 시퀀스에서 컷의 발생 빈도는 5초 정도이다. 초당 29.97프레임(NTSC 용)인 MPEG의 한 GOP에는 15개 이하의 픽처가 있고 한 GOP에서 I픽처가 적어도 1번은 출현하고 각 GOP의 대표 프레임이라 할 수 있기 때문에 I 픽처만으로 컷을 검출하여도 크게 무리는 없다^[6].

2-2 MPEG 비디오 시퀀스에서 I픽처를 이용한 기존의 컷 검출 알고리즘

Sethi는 MPEG 비디오 시퀀스의 I픽처에서 DC성분을 추출하여 앞, 뒤 프레임간 추출된 DC영상에 대해 Yakimovsky Likelihood Ratio Test, Chi-square Test, Kolmogrov-smirnov Test를 실행하여 비교 분석하였다^[6].

Yakimovsky Likelihood Ratio Test 식은 다음과 같다.

$$Y = \frac{(\sigma_0^2)^{m+n}}{(\sigma_1^2)^m \cdot (\sigma_2^2)^n} \quad (2)$$

여기서 σ_1^2, σ_2^2 는 각각 현 프레임과 지나간 전

프레임의 분산이고 σ_0^2 는 두 히스토그램의 공통부분으로부터 얻어진 분산이다.

m과 n은 두 히스토그램의 요소의 갯수이다.

모든 I픽처는 같은 크기이므로 m과 n은 같다.

만약 Y의 값이 임계치를 초과하면 컷으로 판정 한다.

Chi-square Test는 다음과 같다.

$$X^2 = \sum_j \frac{(HP_j - HC_j)^2}{(HP_j + HC_j)^2} \quad (3)$$

HP_j, HC_j는 bin j에서 그레이레벨 블록의 수이다.

Chi-square Test는 가장 많이 사용되는 Test이다. X²이 임계치 보다 클때 컷이 발생된 것이다.

Kolmogorov-smirnov Test는 다음식으로 주어진다.

$$D = \max_j |CHP_j - CHC_j| \quad (4)$$

CHP_j와 CHC_j는 각각 전 프레임과 현재 프레임의 히스토그램차의 누적 합이다. D의 값이 임계치 보다 크면 컷이 발생된 것이다.

기존의 알고리즘은 그림 3과 같이 I픽처의 DC영상을 위의 세가지 Test식을 이용하여 이전 프레임과 현재프레임을 비교해가며 비디오 시퀀스의 끝까지 실행하여 임계치를 초과하는 DC영상을 추출하였다.

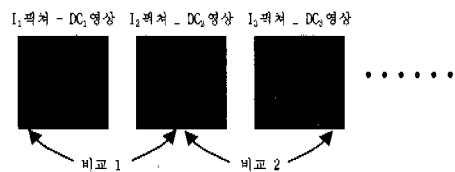


그림 3. DC 영상간 유사성 비교 방법

III. 제안 알고리즘

MPEG 비디오 시퀀스로부터 그림 2와 같이 I 픽처로 부터 DC영상을 각각 추출한다.

이때 비교법은 기존 논문에서 제시한 Yakimovsky Likelihood Ratio Test, Chi-square Test, Kolmogrov-smirnov Test 중 가장 일반적인 Chi-

square Test를 이용한다⁷⁾. Chi-square Test는 X^2 test 라고도 불리운다. 그림 3과 같이 비디오 시퀀스를 순차적으로 비교하여 구한 측정값이 임계치, T 를 초과하면 다음과 같이 컷 예상지점으로 지정한다.

$$C_{i,i+1} \geq T : \text{컷 예상 지점} \quad (5)$$

$$C_{i,i+1} < T : \text{컷이 아님}$$

여기서 $C_{i,i+1}$ 는 i 와 $i+1$ 의 프레임 사이의 X^2 의 값을 의미한다.

이때 컷 예상 지점으로 검출된 두 DC영상의 컷 지점인가를 판별해야 한다.

이는 컷 예상 지점의 이웃 프레임들의 특징차의 평균을 컷 예상지점의 특징량과 비교함으로써 이루어진다.

카메라 정지 등에 의해 컷이 일어나는 지점은 급격한 화면의 변화가 일어나므로 높은 $C_{i,i+1}$ 의 값을 갖는다. 하지만 샷(shot) 내의 프레임 중에서도 물체의 빠른 움직임, 큰 물체의 움직임, 카메라의 움직임, 카메라 후레쉬의 선풍 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 컷으로 오검출할 수 있다.

본 논문의 알고리즘은 이를 해결하기 위해 기존에 컷으로 판정된 컷을 컷 예상 지점으로 지정하고 실제의 컷인지 아닌지를 이웃 프레임의 DC영상을 사용하여 재판정한다.

실제 컷이 일어나는 지점은 컷을 기준으로 이전 프레임들과 이후 프레임들은 크게 차이가 난다. 하지만 컷 기준으로 이전의 프레임들과 이후의 프레임들은 각각 비슷하다.

또 실제 컷이 아닌 지점은 각 프레임의 특징차는 조금 크다 하더라도 이웃의 프레임들과 특정 부분을 제외하고는 비슷하다.

그러므로 기존의 두 비교 프레임의 전과 후의 프레임을 선정하여 총 네 개의 프레임을 이용해, 컷 예상지점의 이웃 프레임들과의 특징차의 평균을 구하여 컷 예상 지점의 특징차와 비교하여 실제의 컷인지 아닌지를 판별한다.

두 프레임이 아닌 한 프레임 전이나 한 프레임 후의 프레임과 특징차를 각각 비교하여 평균값을 취하면 해당 프레임들의 평균 특징차가 된다. 이때 이 평균값보다 컷 예상 지점의 특징차가 크다면 실제의 컷으로 판정한다.

컷 예상 지점의 두 비교 프레임의 이전 프레임과 이후 프레임의 비교는 그림 4와 같다.

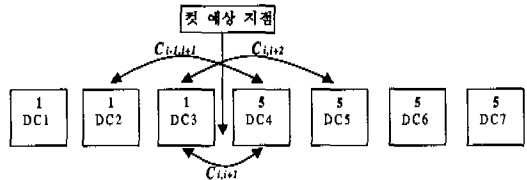


그림 4. 제안된 영상 비교 방법

만약 $C_{i,i+1}$ 의 비교 측정 값이 임계치를 초과하여 컷의 예상점으로 지정 되었다면 $C_{i-1,i+1}$ 와 $C_{i,i+2}$ 의 비교 측정치를 이용하여 $M_{i,i+1}$ 을 구해서 $C_{i,i+1}$ 와 $M_{i,i+1}$ 을 비교한다. $C_{i,i+1}$ 이 평균치, $M_{i,i+1}$ 이상이라면 컷 지점으로 판정한다.

판정식은 다음과 같다.

$$C_{i,i+1} \geq M_{i,i+1} : \text{컷} \quad (6)$$

$$C_{i,i+1} < M_{i,i+1} : \text{컷이 아님}$$

여기서 $M_{i,i+1}$ 은 $\frac{C_{i-1,i+1} + C_{i,i+2}}{2}$ 이고 i 는 비교할 DC영상의 순차번호이다.

본 알고리즘의 순서 흐름도는 그림 5와 같다.

IV. 실험 및 고찰

본 논문에서는 Sethi의 MPEG 비디오 시퀀스의 DC추출방법을 이용하였고 기존 논문들의 매체(parameter) 비교 알고리즘인 Chi-square Test를 사용하였다.

특징차 비교는 다수개의 프레임을 같이 참조하는 새로운 방법을 제안하여 실행하였다. 실험 대상 비디오 시퀀스는 뉴스와 축구경기와 드라마가 녹화된 비디오 1과 영화가 녹화된 비디오 2이다. 그리고 비

디오 1과 2는 해상도가 320 x240이고 각각3880개와 946개의 프레임을 갖는 MPEG-1 비디오이다. 비디오 1의 시퀀스에는 18개의 급격한 장면 변화와 1개의 점진적인 장면 변화를 가지고 있고 비디오 2의 시퀀스에는 13개의 급격한 장면 변화와 1개의 점진적인 장면 변화가 포함되어 있다.

본 연구에서 제안한 알고리즘을 MicroSoft Windows98환경에서 Visual C++6.0 컴파일러를 사용하여 실험하였다.

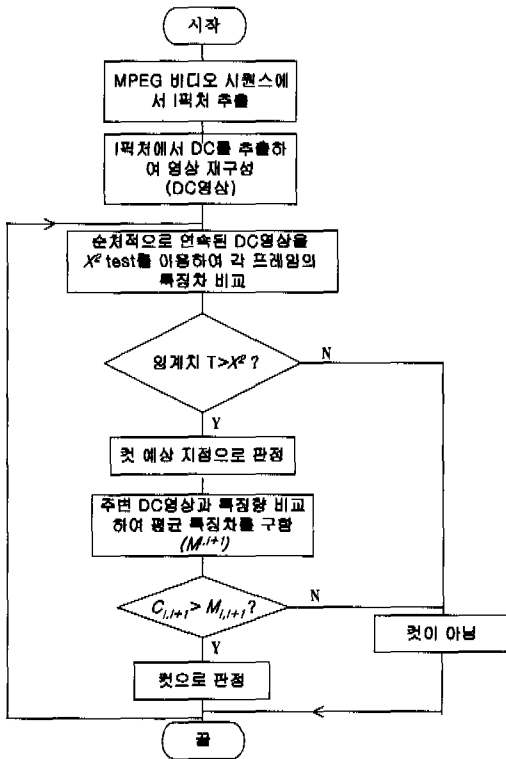


그림 5. 제안 알고리즘의 처리 순서도

기존의 MPEG 비디오 시퀀스의 DC영상을 전프레임과 현프레임의 1:1 비교방법과 제안한 비교방법의 결과가 각각 표1과 그림 6에서 그림 11까지 제시되었다. 그림 6과 그림 9는 각각 비디오 1과 비디오 2의 X^2

test 결과이다. 그림 7과 그림 10은 각각 비디오 1과 비디오 2를 기존의 비교 방법을 이용한 컷 검출 결과이다. 그림 8과 그림 11은 각각 비디오 1과 비디오 2를 제안된 비교 방법을 이용하여 컷 검출한 결과이다.

그림의 검출 구분 중 c, f, m 은 각각 컷 검출 프레임, 오검출 프레임, 검출하지 못한 프레임을 나타낸다.

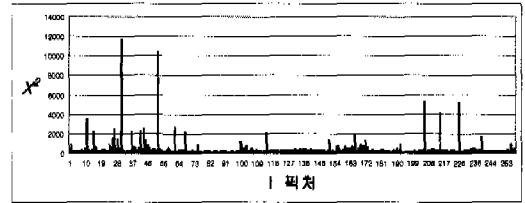


그림 6. X^2 알고리즘을 이용한 비디오 1의 컷 검출 결과

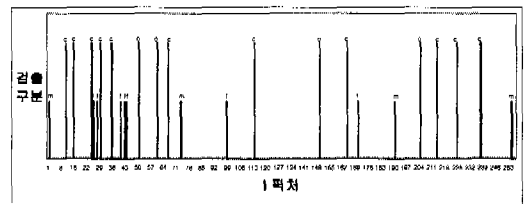


그림 7. 기존 알고리즘을 이용한 비디오 1의 컷 검출 결과

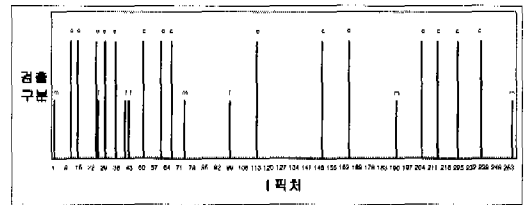


그림 8. 제안 알고리즘을 이용한 비디오 1의 컷 검출 결과

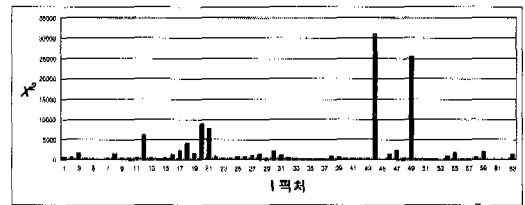


그림 9. X^2 알고리즘을 이용한 비디오 2의 컷 검출 결과

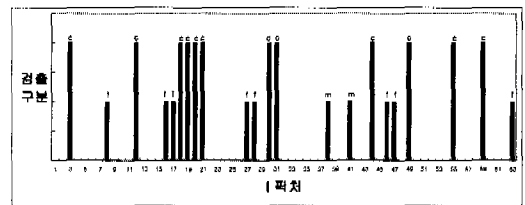


그림 10. 기존 알고리즘을 이용한 비디오 2의 컷 검출 결과

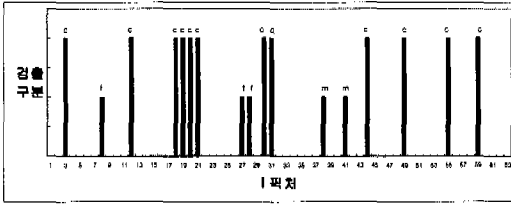


그림 11. 제안 알고리즘을 이용한 비디오 2의 컷 검출 결과

표 1. 기존 알고리즘과 제안 알고리즘의 검출 결과 비교

	기존 알고리즘							제안 알고리즘					
	실계의 컷수	검출한 컷수	Nc	Nf	Nm	Recall	Precision	검출한 컷수	Nc	Nf	Nm	Recall	Precision
비디오 1	19	22	15	7	4	0.68	0.78	19	15	4	4	0.78	0.78
비디오 2	14	20	12	8	2	0.6	0.86	15	12	3	2	0.8	0.86

여기서 비디오 1과 2의 기존 방법의 Recall은 각각 0.68과 0.6이고 Precision은 각각 0.78과 0.86이다. 그리고 제안한 알고리즘의 Recall은 각각 0.78과 0.8이고 Precision은 각각 0.78과 0.86이다. 그러므로 Recall면에서 제안한 방법이 우수한 것을 알 수 있다.

Recall과 Precision의 표현식은 식 7과 같다.

$$Recall = \frac{n_c}{n_c + n_f} \quad (7)$$

$$Precision = \frac{n_c}{n_c + n_m}$$

n_c : 올바르게 검출된 컷 수

n_m : 검출하지 못한 컷 수

n_f : 잘못 검출한 컷 수

V. 결론

비디오 시퀀스를 계층적으로 분할하기 위한 컷 검출 연구는 다양하지만 MPEG 비디오 시퀀스에서 컷 검출 연구는 아직 미흡한 단계이다. 본 논문에서는 MPEG 비디오의 DC 성분을 이용하여 컷 검출을 시행하여 기존의 현재 프레임과 이전 프레임

임만 비교하는 방법보다 우수한 결과를 보였다. 기존의 특징차 측정법은 현재 프레임과 이전 프레임만 고려하므로 물체의 움직임이 빠르거나 큰 물체의 움직임 등을 컷으로 오검출을 하기 쉬운데 본 논문의 비교방법은 이웃 프레임과의 평균값을 이용해 기존 방법의 오검출을 줄여서 컷 검출율을 높였다. 또한 처리속도 문제를 해결한다면 비압축 동영상 상에도 적용 할 수 있다.

향후 연구는 다수개의 DC영상을 비교함으로써 일어나는 처리 속도저하를 해결하는 방법을 연구해야 하며 디졸브와 fade와 같은 점진적인 장면변화와 panning, zooming과 같은 카메라의 특수효과를 구별하여 검출하는 연구가 필요하다.

참고 문헌

- [1] J. A. R. Center, *Query by Image and Video Content*; The Qbic system. IEEE computer, 28(9):23~32, Sept, 1995
- [2] T. Chiueh, *Content-Based Image Indexing*, Pro. Of the 20st VLDB conf, Santiago Chile, pp. 582~593, 1994
- [3] S. Smoliar and H. Zhang, *Content-Based Video Indexing and Retrieval*, IEEE Multimedia, vol. 1, No. 2, pp. 62~72, 1994
- [4] John S. Boreczky, Lawrence A. Rowe, *Comparison of Video shot boundary Detection Techniques*, in Storage & Retrieval for Image and Video Database IV, Proc. Of SPIE 2670, pp. 170~179, 1996
- [5] ISO/IEC 11172-2 Committee Draft (MPEG)
- [6] Nilesh V. Patel and Ishwar K. Sethi, *Compressed Video Processing for cut Detection*, IEE Proc. Video, Image and Signal Processing, 1996
- [7] Akio Nagasaki and Yuzuru Tanaka, *Automatic Video Indexing and Full-Video 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

