

논문 2004-41SP-5-4

급격한 밝기 변화가 있는 비디오 시퀀스에서 효율적인 비디오 색인 알고리즘

(An Efficient Video Indexing Algorithm for Video Sequences
with Abrupt Brightness Variation)

김 상 현*, 박 래 홍*

(Sang Hyun Kim and Rae-Hong Park)

요 약

디지털 미디어 데이터의 증가에 따라 디지털 미디어를 효과적으로 관리하고 사용하기 위하여 다양한 비디오 색인 및 비디오 시퀀스 정합 알고리즘이 제안되었다. 본 논문에서는 급격한 밝기 변화를 가지는 비디오 시퀀스에 대해서 효율적인 비디오 색인 알고리즘과 비디오 시퀀스 질의에 대한 비디오 시퀀스 정합 알고리즘을 제안한다. 급격한 밝기 변화를 고려한 비디오 색인의 정확도를 향상시키고 계산량을 줄이기 위해 제안한 알고리즘은 칼라 특성뿐만 아니라 에지 특성도 함께 사용하였으며 기존의 알고리즘에 비해 현저한 성능 향상을 보였다. 제안한 알고리즘은 먼저 칼라 히스토그램을 사용하여 후보 샷경계 지점을 추출하고 에지 정합과 밝기 보상을 이용하여 후보점들이 샷경계인지 밝기 변화인지를 결정한다. 장면내의 밝기 변화가 작은 경우 에지 정합과 밝기 보상은 샷경계에서만 일어난다. 실험 결과 제안한 방법은 기존의 방법에 비해 비슷한 계산량으로 현저히 향상된 성능과 효율을 보였다

Abstract

With increase in digitalmedia data, various video indexing and video sequence matching algorithms have been proposed to efficiently manage and utilize digital media. In this paper, we propose a robust video indexing algorithm to detect scene changes for video sequences with abrupt luminance variations and an efficient video sequence matching algorithm for video sequence query. To improve the accuracy and to reduce the computational complexity for video indexing with abrupt luminance variations, the proposed algorithm utilizes edge features as well as color features, which yields a remarkably better performance than conventional algorithms. In the proposed algorithm, first we extract the candidate shot boundaries using color histograms and then determine using edge matching and luminance compensation if they are shot boundaries or luminance changes. If the scene contains trivial brightness variations, the edge matching and luminance compensation are performed only for shot boundaries. In experimental results, the proposed method gives remarkably a high performance and efficiency than the conventional methods with the similar computational complexity.

Keywords : video indexing, scene change detection, brightness variation, luminance compensation, edge matching

I. 서 론

디지털미디어의 증가에 따라 효율적인 비디오 색인 및 검색을 위해 다양한 알고리즘이 제안되었다. 본 논문에서는 급격한 조명변화를 가지는 비디오 시퀀스에

대해서 효과적인 장면 변화 검출이 가능한 알고리즘을 제안한다.

디지털 비디오를 효율적으로 사용하기 위하여 비디오 시퀀스를 색인하고 비디오 콘텐츠를 효과적으로 표현하는 것이 중요하다. 효율적인 비디오 색인을 위해서는 비디오 시퀀스를 샷 (shot) 단위로 분할하는 작업이 필요하다. 비디오 시퀀스에서 샷을 검출하는 것을 장면 전환 검출이라 하고 이러한 샷들의 경계에 해당하는 부분을 샷경계 (shot boundary)라 한다. 대용량의 비디오

* 정회원, 서강대학교 전자공학과
(Dept. of Electronic Eng, Sogang University)

※ 본 연구는 서강대학교 산업기술연구소에서 지원되었음.
접수일자: 2003년3월5일, 수정완료일: 2004년8월18일

시퀀스에 대해 추출된 샷경계는 비디오 데이터의 전체적인 내용을 요약해서 나타내고 브라우징하는 데에도 사용될 수 있다. 일반적으로 샷내에는 팬 (pan), 틸트 (tilt) 또는 줌 (zoom)이 포함될 수 있으며 카메라의 녹화에서 정지까지의 단위가 하나의 샷이 된다. 샷 단위로 나뉘어진 시퀀스는 사용자에게 색인이나 내용검색 등에 매우 유용하게 사용될 수 있다^[1-2].

샷경계 검출에 관한 연구는 칼라 비디오 시퀀스에 대해 칼라 히스토그램을 사용하여 시간적으로 분할하는 연구가 진행되었다. 연속되는 프레임의 히스토그램간의 차이를 계산하여 프레임간 차이가 큰 경우를 샷경계로 검출한다. 저장된 디지털 비디오는 압축되어 있으므로 압축된 비디오의 특성을 이용하여 직접 샷경계를 검출할 수 있다. MPEG 표준이 객체 단위의 코딩으로 발전함에 따라 압축 특성을 이용한 비디오 색인이 가능하도록 연구되고 있다. 압축 영역에서의 비디오 색인은 블록 단위의 움직임 예측에 의한 움직임 정보에 의해서도 가능하다.

지금까지의 샷경계 검출 알고리즘은 샷내의 밝기 변화가 일정하다는 전제 하에 성능을 향상시키는 알고리즘이 제안되었으나 대부분의 알고리즘들은 밝기 변화가 큰 비디오 시퀀스에 대해서 잘못된 검출 (false detection)이나 누락된 검출 (miss detection)에 의해서 성능 저하가 크게 나타난다^[3-6]. 이러한 문제를 해결하기 위해 몇가지 알고리즘이 제안되었으나 대부분 후처리 과정이 필요하게 된다^[7-8]. 급격한 밝기 변화가 있는 경우 장면전환 검출을 위해 필요한 후처리 과정의 계산량을 줄이고 정확도를 향상시키기 위해 제안한 알고리즘에서는 기존의 칼라 특성뿐만 아니라 에지 특성을 사용하여 기존의 방법에 비해 현저히 향상된 결과를 얻었다.

비디오 시퀀스에서 샷경계를 추출하는 것은 시퀀스의 시간적인 내용에 현저한 변화가 일어나는 것을 검출하는 것이다. 그러므로 연속되는 프레임으로 구성되는 시퀀스에서의 비연속적 경계를 찾아내게 된다. 이를 찾기 위해서는 먼저 특징 (feature)의 추출이 필요하며 여기서 특징은 비디오 시퀀스의 특성을 잘 묘사할 수 있어야 한다. 이렇게 프레임별로 추출된 영상의 특징을 비교함으로써 샷경계를 추출할 수 있다. 그러나 급격한 밝기 변화를 가지는 경우 대부분의 알고리즘에서 적용되고 있는 칼라특성이 밝기 변화가 큰 지점에서 불연속적이며 같은 샷내의 칼라 특성과 큰 차이를 나타낸다. 밝기 변화에 의한 동일한 샷내의 프레임간에 추출된 특징들의 유사도 차이는 비디오 색인의 성능과 정확도를

저하시키므로 동일한 샷내에서는 밝기나 조명 변화에 강건한 특징의 추출과 결합이 필요하다.

제안한 알고리즘에서는 먼저 압축영역 (compressed domain)에서 칼라 히스토그램 (color histogram)을 사용하여 샷경계로 추출되는 후보점을 추출한다. 추출된 후보점에 대해서는 비압축 영역 (uncompressed domain)에서 에지 정합과 밝기 보상 (luminance compensation)을 수행함으로써 추출된 후보점이 프레임간의 급격한 밝기 변화인지 샷경계인지를 구분하게 된다. 칼라 특성과 에지특성을 계층적으로 사용함으로써 계산량을 감소시키고 후처리 과정 없이도 플래쉬라이트 (flash-light)나 스포트라이트 (spotlight) 등의 급격한 밝기 변화와 샷경계를 효율적으로 추출할 수 있다

II. 기존의 비디오 색인 알고리즘

기존의 비디오 색인 알고리즘은 크게 비압축 영역에서의 알고리즘과 압축영역에서 비디오 압축의 특성을 이용한 알고리즘으로 나눌 수 있다. 비압축 영역에서의 비디오 색인 알고리즘은 연속되는 프레임으로 구성된 비디오 시퀀스에서 동일한 샷내에서의 인접 프레임간에는 동일성이 강하고 샷경계 부근에서는 프레임간 유사도가 떨어진다는 데 기반을 두고 있으며 화소나 히스토그램 기반의 방법이 주로 사용된다. 이 중 히스토그램 (histogram)을 이용한 방법이 많이 사용되는데 히스토그램은 동일한 샷 내에서는 프레임간의 변화에 대해 덜 민감한 특성을 나타낼 뿐만 아니라 각 프레임에서 히스토그램을 추출하는 데 필요한 계산량도 작다^[9-12].

1. 히스토그램 차(Histogram Difference)

히스토그램 차 방식은

$$\sum_j H_{t+1}(j) - H_t(j) \quad (1)$$

로 정의할 수 있는데 여기서 $H_t(j)$ 는 j 번째 빈 (bin)의 히스토그램을 나타내고 t 는 t 번째 프레임을 나타낸다. 샷경계는 히스토그램 차의 합이 주어진 기준치 보다 높을 때로 간주된다^[7].

2. Chi-Square 테스트

두개의 히스토그램 분포를 비교하기 위하여

$$\sum_j \frac{(H_{t+1}(j) - H_t(j))^2}{H_{t+1}(j) + H_t(j)} \quad (2)$$

와 같이 정의된 차이 스퀘어 테스트를 사용한다^[9]. 압축 영역에서의 알고리즘은 MPEG과 같은 압축 표준에서 압축 특성을 이용하여 비디오 색인이 가능하다.

III. 제안한 비디오 색인 알고리즘

실제 비디오 시퀀스 색인시 플래쉬라이트와 같은 급격한 밝기 변화로 인한 샷경계 검출의 오류를 줄이기 위하여 몇가지 알고리즘이 제안되었으나 밝기 변화정도가 심할 경우 정확한 샷경계 추출이 어렵고^[3-4], 별도의 후처리 과정을 거쳐야 하는 단점을 가진다^[8].

제안한 알고리즘은 심각한 밝기 변화를 포함하는 비디오 시퀀스에 대해서도 높은 정확도를 보이고 후처리 과정없이 직접 샷경계 지점과 밝기 변화지점을 구분할 수 있다. 제안한 알고리즘은 크게 두 부분으로 나눌 수 있다. 첫번째 부분은 칼라 히스토그램을 사용하여 비디오 시퀀스내에 샷경계로 추정되는 후보점을 추출하고 두번째 부분은 에지 정합과 밝기 보상을 수행하여 이전 단계에서 추출된 후보점이 샷경계인지 아니면 급격한 밝기 변화인지를 결정한다.

1. 첫번째 단: 칼라 히스토그램을 이용한 샷경계 후보점의 추출

샷경계의 후보점을 추출하기 위하여 먼저 입력시퀀스에 대해 DC영상을 추출하고 추출된 DC 시퀀스에 대하여 히스토그램 차 방식을 적용한다.

DC 영상은 원영상을 공간적으로 (spatially) 축소한 영상에 해당된다. 추출된 DC 영상은 비디오 샷의 비교나 비디오 데이터의 구조 추출 또는 브라우징에 효과적으로 사용될 수 있다^[13-15]. 원영상의 (i,j)위치에서의 화소값을 $I(i, j, t)$ 라 하면 DC영상 $\bar{I}(u, v, t)$ 는

$$\bar{I}(u, v, t) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I((u-1)M+m, (v-1)N+n, t) \tag{3}$$

와 같이 얻을 수 있고 u, v 는 $0 \leq u \leq U-1, 0 \leq v \leq V-1$ 의 범위를 가진다. 여기서 $I((u-1)M+m, (v-1)N+n, t)$ 는 t번째 입력 프레임에서 서브블록 (u, v)에서의 (m, n) 화소의 값을 나타낸다.

샷경계 검출의 효율성은 연속되는 프레임간의 유사도 측정 방식에 따라 많은 영향을 받는다. 본 논문에서는 샷경계의 후보점을 추출하기 위해서 칼라 히스토그램 차 방식

$$\sum_j (|H_{t+1}^Y(j) - H_t^Y(j)| + |H_{t+1}^U(j) - H_t^U(j)| + |H_{t+1}^V(j) - H_t^V(j)|) \tag{4}$$

을 사용하였다. 여기서 $H_t^Y(j), H_t^U(j), H_t^V(j)$ 는 t 프레임에서의 j번째에 레벨에 해당하는 히스토그램을 나타내고 Y, U, V는 각각의 칼라요소 (color component)를 나타낸다. 이렇게 구해진 칼라 히스토그램의 차이가 정해진 문턱치 (threshold) 이상일 때 샷경계의 후보점으로 결정된다.

2. 두번째 단: 에지 정합과 밝기보상을 이용한 샷경계와 밝기 변화의 추출

첫번째 단계에서 추출된 샷경계의 후보점이 급격한 밝기 변화인지 샷경계인 지를 결정하기 위하여 에지 정합과 밝기 보상을 수행한다. 에지 특성을 효율적으로 추출하기 위해 Marr-Hildreth 에지 검출 방법을 사용하였다^[16]. 프레임 I(i,j)내의 에지들은

$$\nabla^2 G * I(i, j) = 0 \tag{5}$$

를 만족하는데 여기서 G는 가우시안 함수 (Gaussian Function)를 나타내고

$$G(r) = \exp\left(\frac{-r^2}{2\sigma^2}\right) \tag{6}$$

와 같이 정의할 수 있다. 여기서 $r = \sqrt{i^2 + j^2}$ 과 σ 는 각각 가우시안 확장 파라미터 (Gaussian Spread Parameter)와 표준편차를 나타낸다. $\nabla^2 G$ 은

$$\nabla^2 G(r) = \frac{r^2 - 2\sigma^2}{\sigma^4} \exp\left(\frac{-r^2}{2\sigma^2}\right) \tag{7}$$

가우시안 (Gaussian)의 라플라시안 (Laplacian)을 나타낸다. 에지밀도와 굵기는 σ 를 변화시켜 조정할 수 있다.

본 논문에서는 에지를 효과적으로 정합하기 위하여 EMR (edge matching rate)을 이용한 방법을 제안한다. 에지 특성은 프레임간의 밝기 변화에 대해서도 강건한 특성을 나타내고 Y 신호 성분만을 사용하여 간단히 추출해 낼 수 있다. 제안하는 알고리즘은 움직이는 에지 (moving edge)에 의해 발생할 수 있는 오류를 피하기 위해 밝기 보상도 함께 수행한다. 에지 정합과 밝기 보상을 적용한 제안한 알고리즘은 급격한 밝기 변화에 대해서도 기존의 방법에서 나타나는 심각한 성능저하 없

이 견고한 성능을 나타낸다^[17].

에지 정합과 밝기 보상의 수행과정은 다음과 같다.

1) 이전 프레임 $I_p(i, j)$ 와 현재 프레임 $I_c(i, j)$ 로부터 각각의 에지 영상 $E_p(i, j)$ 와 $E_c(i, j)$ 를 추출한다.

2) $E_p(i, j)$ 와 $E_c(i, j)$ 로부터 'AND' 연산 (operation)을 사용하여 공통에지 영상 (common edge image)를 얻는다.

3) EMR (edge matching rate)

$$EMR = \frac{NEP_{in\ the\ common\ edge\ image}}{\min\{NEP_{in\ E_p(i, j)}, NEP_{in\ E_c(i, j)}\}} \quad (8)$$

를 계산한다. 여기서 NEP (number of edge pixels)는 에지 화소들의 수를 나타낸다.

4) 밝기 보상된 프레임은 $I_{lc}(i, j) = I_p(i, j) + Avg$ 에 의해 얻을 수 있는데 여기서 $I_{lc}(i, j)$ 는 밝기 보상된 프레임을 나타내고 Avg 는 $I_p(i, j)$ 와 $I_c(i, j)$ 간의 차영상 (difference image)의 밝기 평균을 나타낸다. 밝기 보상된 프레임을 추출한 후에 현재 프레임과 밝기 보상된 프레임간의 차를 나타내는 FDV (frame difference value)

$$FDV = \sum_i \sum_j I_c(i, j) - I_{lc}(i, j) \quad (9)$$

를 계산한다.

5) 계산된 EMR과 FDV가 문턱치 보다 높으면 샷경계 후보점은 같은 샷내에서 급격한 밝기 변화가 일어난 지점으로 간주된다. 즉 EMR과 FDV가 문턱치 보다 낮은 경우만 실제 샷경계로 추출된다.

실제 급격한 밝기 변화가 있는 비디오 시퀀스의 경우 히스토그램차는 크게 나타나므로 에지 정합을 통해 밝기 변화와 샷경계를 구분하는 것은 효과적이다. 또한 카메라의 움직임이나 프레임내의 심한 움직임으로 인해 움직인 에지의 수가 많은 경우 발생할 수 있는 에지 정합의 오류를 줄이기 위해 에지 정합과 더불어 밝기 보상을 함께 사용하였다. 심각한 밝기 변화가 있는 비디오 시퀀스는 프레임차가 크게 나타나고 이러한 밝기 변화는 플래쉬라이트나 스포트라이트외에 다양한 조명의 영향에 의해 발생된다. 이 경우 밝기 보상을 수행한 후의 프레임차는 밝기 보상이 없는 경우의 프레임차에 비해 현저히 작은 값을 나타낸다. 그러므로 FDV는 프레임간의 밝기 변화에 대해서는 작은 값을 나타내고 샷경

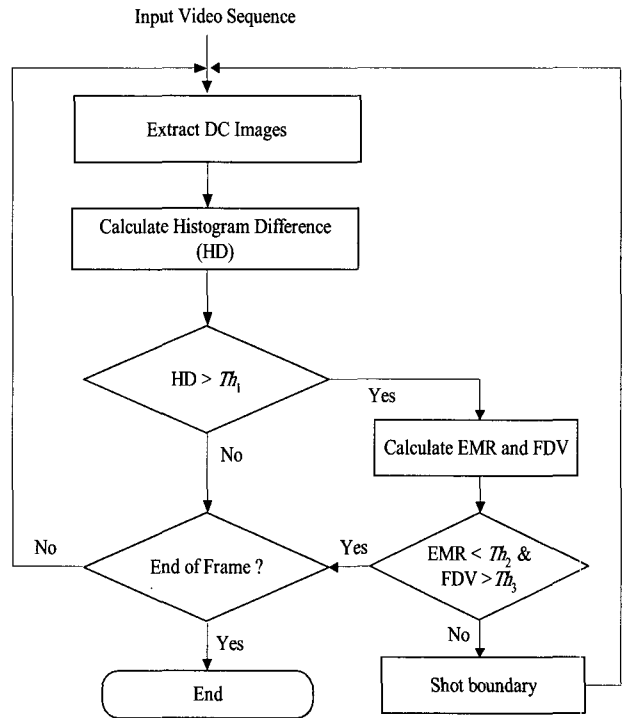


그림 1. 제안한 알고리즘의 흐름도.
Fig. 1. Flowchart of the proposed algorithm.

계에 대해서는 큰값은 나타낸다.

그림 1은 제안한 알고리즘의 흐름도를 나타낸다. 여기서 Th_1, Th_2, Th_3 는 각각 0.15, 0.25, 0.2로 실험에 의해 튜닝되었다. 그림 1에서와 같이 입력 시퀀스에 대해서 먼저DC영상을 추출하고 칼라 히스토그램의 차이를 계산하여 샷경계 후보점인지 아닌지를 판별한다. 히스토그램의 차이가 작은 경우는 계속해서 프레임간의 차이만을 계산하고 히스토그램의 차이가 큰 경우는 에지 정합 (EMR)과 밝기 보상에 의한 프레임간 차이 (FDV)를 계산하여 이로부터 샷경계를 추출한다.

에지정합과 밝기 보상은 샷경계로 추출되는 후보점에 대해서만 수행하므로 샷의 길이가 길고 샷내의 밝기 변화가 미미한 경우는 히스토그램만을 사용하는 경우에 비해 큰 계산량의 증가 없이 알고리즘을 수행한다. 밝기 변화가 심한 프레임이 많은 시퀀스에 대해서는 계산량은 증가하나 기존의 알고리즘에서 대부분 오류 샷검출로 인한 성능저하를 줄일 수 있다. 시퀀스의 밝기 변화정도가 커질수록 기존의 방법과 제안한 방법간의 성능 차이는 커지게 된다. 또한 기존의 알고리즘의 경우에는 칼라 히스토그램의 차이에 의한 샷경계 검출의 기준값에 따라 성능 변화가 심하다. 즉 밝기 변화가 심한 시퀀스의 경우 오류 검출이 많이 발생함으로 샷경계 검출의 기준값을 높여도 그만큼의 누락되는 샷경계 검출

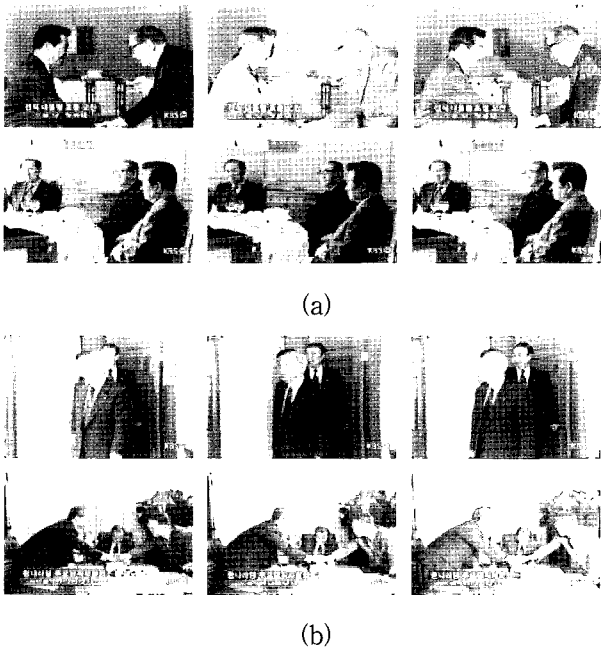


그림 2. 급격한 밝기 변화를 포함하는 뉴스 비디오 시퀀스. (a) 'News Article 1' 시퀀스 (2808 프레임). (b) 'News Article 2' 시퀀스 (3386 프레임).
 Fig. 2. News video sequences with abrupt brightness variations. (a) 'News Article 1' sequence (2808 frames). (b) 'News Article 2' sequence (3386 frames).

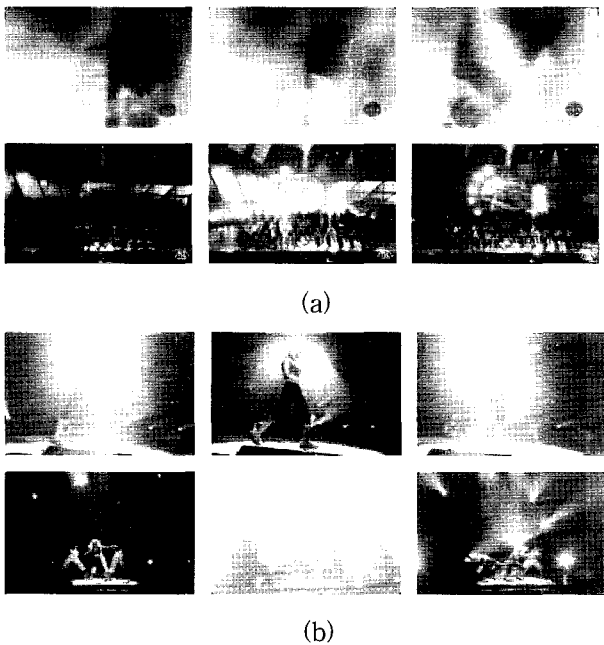


그림 3. 급격한 밝기 변화를 포함하는 뮤직 비디오 시퀀스. (a) 'Music Video 1' 시퀀스 (Britney, 4914 프레임). (b) 'Music Video 2' 시퀀스 (Christina, 7029 프레임).
 Fig. 3. Music video sequences with abrupt brightness variations. (a) 'Music Video 1' sequence (Britney, 4914 frames). (b) 'Music video 2' sequence (Christina, 7029 frames).

표 1. 뉴스 비디오 시퀀스에 대한 샷경계 검출방법의 성능비교.

Table 1. Performance comparison of shot boundary detection for news video sequences.

테스트 시퀀스	방법	Recall (%)	Precision (%)
News Article 1 (2808 프레임)	Histogram Difference	39.13	12.16
	Chi-Square Test	30.43	10.45
	Proposed Method	82.61	100.00
News Article 2 (3386 프레임)	Histogram Difference	67.74	52.50
	Chi-Square Test	38.71	54.55
	Proposed Method	96.77	96.77

표 2. 뮤직 비디오 시퀀스에 대한 샷경계 검출방법의 성능비교.

Table 2. Performance comparison of shot boundary detection for music video sequences.

테스트 시퀀스	방법	Recall (%)	Precision (%)
Britney (4914 프레임)	Histogram Difference	61.11	47.83
	Chi-Square Test	55.56	45.45
	Proposed Method	85.71	100.00
Christina (7029 프레임)	Histogram Difference	60.98	64.10
	Chi-Square Test	59.76	61.25
	Proposed Method	95.12	93.98

의 수가 많아진다. 반면 제안한 알고리즘에서는 첫번째 후보 샷경계 검출단에서 입력 시퀀스에 관계없이 히스토그램 차이에 대한 일정한 기준값을 가져도 다음 단계에서 예지특성등을 고려하여 재검증하므로 범용 비디오 색인 시스템에 더욱 효과적이다.

기존의 알고리즘에서는 밝기 변화가 심한 비디오 시퀀스에 대해서 후처리 과정을 통해 밝기 변화 성분을 검출하므로 계산량 증가가 많고 일반적인 시퀀스에 대해서는 불필요한 계산이 요구되었다. 제안한 알고리즘은 두 단계의 색인 알고리즘을 적용하여 심각한 밝기 변화를 가지고 있는 비디오 시퀀스 뿐만 아니라 그렇지 않은 시퀀스에 대해서도 효율적으로 사용될 수 있다.

IV. 실험 결과 및 토의

제안한 알고리즘과 기존의 알고리즘에 대해 비교분석하기 위해 2808프레임과 23개의 샷으로 구성된 비디오 시퀀스와 3386프레임과 31개의 샷으로 구성된 비디오 시퀀스를 사용하여 실험하였다. 실험에 사용된 시퀀스는 그림 2에서 보여주듯이 밝기 변화가 심하고 플래쉬라이트가 자주 나타나는 'News Article 1'과 'News Article 2'의 뉴스 시퀀스이다.

표 1은 세 가지 샷경계 검출 방법의 성능과 정확도를 리콜 (recall)과 프리시전 (precision)

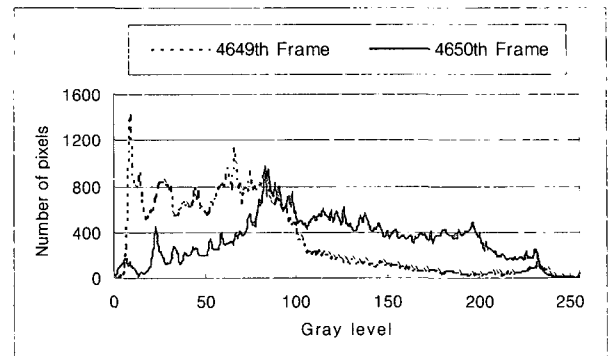
$$Recall = \frac{NCD}{NCD + NMD} \quad (10)$$

$$Precision = \frac{NCD}{NCD + NID} \quad (11)$$

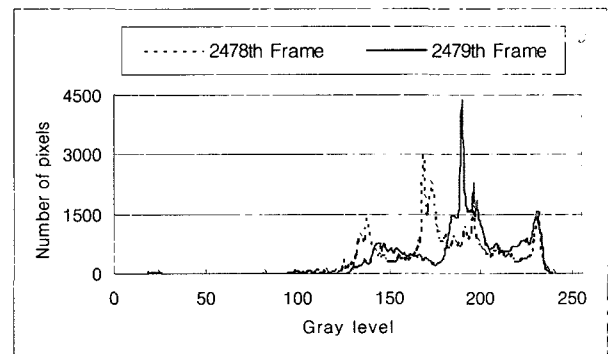
에 의해 나타내고 있다^[7]. 여기서 NCD (number of correct detection), NMD (number of missed detection), NID (number of incorrect detection)는 각각 올바른 샷경계 검출 횟수, 누락된 샷경계 검출 횟수, 잘못된 샷경계 검출 횟수를 나타낸다. 식에서도 알 수 있는 바와 같이 recall은 누락된 샷경계 검출에 많은 영향을 받으며 precision은 잘못된 샷경계 검출 정도에 따라 달라진다. 두 가지 성능 평가에서 누락된 샷경계와 잘못된 샷경계 검출이 작을수록 recall과 precision이 커지며 성능이 높아진다.

표 1은 실험에 사용한 뉴스 비디오 시퀀스에 대해 히스토그램 차, Chi-Square 테스트 및 제안한 방법을 사용하여 얻은 샷경계 검출결과와 성능을 비교한 것이다.

표 1에서 'News Article 1'에 대해서 히스토그램 차 방법과 Chi-Square 테스트는 매우 낮은 recall과 precision 성능을 보이고 있다. 특히 precision이 많이 떨어져 있는데 이는 심한 밝기 변화 성분에 의하여 샷경계 검출 오류가 많이 발생하였기 때문이다. 반면에 제안한 알고리즘에서는 밝기 변화를 고려한 색인 알고리즘으로 recall과 precision이 크게 향상됨을 알 수 있다. 특히 샷경계 검출의 오류가 거의 발생하지 않아 높은 precision결과를 보였다. 'News Article 2'에 대한 실험에서도 'News Article 1'과 비슷한 성능변화를 보이고 있으며 'News Article 1'에 비해 'News Article 2'에서 나타나는 밝기 변화가 적기 때문에 기존의 알고리즘에서는 약간 향상된 성능을 보이고 있다. 제안한 알고리즘에서는 'News Article 1'과 'News Article 2'에서 나



(a)



(b)

그림 4. 밝기 변화가 있는 시퀀스에 대한 히스토그램. (a) 'Music Video 1' 시퀀스의 히스토그램 (4649 번째와 4650 번째 프레임). (b) 'Music Video 2' 시퀀스의 히스토그램 (2478 번째와 2479 번째 프레임).

Fig. 4. Histograms of brightness variation sequences (a) Histograms of 'Music Video 1' sequence (4649th and 4650th frames) (b) Histograms of 'Music Video 2' sequence (2478th and 2479th frames).

타나는 밝기 변화성분의 차이에 의한 성능변화가 크게 나타나지 않고 일정한 recall과 precision을 보이고 있다. 표 1에서 나타나는 바와 같이 기존의 알고리즘에서는 시퀀스내의 밝기 변화 성분의 정도에 따라 성능 변화가 크게 나타나며 밝기 변화가 심할수록 기존의 방법과 제안한 방법의 성능차이는 더욱 커지는 것을 알 수 있다.

표 1에서 나타나는 바와 같이 에지 정합과 밝기 보상을 수행한 제안한 알고리즘이 기존의 방법에 비해 잘못된 샷경계 검출과 누락된 샷경계 검출의 횟수를 크게 줄임으로써 향상된 성능과 높은 정확도를 나타내었다. 제안한 방법에서 사용한 에지 정합은 밝기 변화 정도가 심한 경우에도 에지 성분의 변화가 칼라 성분의 변화에 비해 상대적으로 덜 민감하므로 비디오 색인에서의 오

류를 줄일 수 있다.

기존의 알고리즘의 경우 밝기 변화에 의한 잘못된 샷 경계 검출을 줄이기 위해 문턱치를 높여도 누락된 샷 경계 검출이 크게 증가하여 심각한 성능 저하를 가져오게 된다. 즉 밝기 변화에 의한 성능저하는 recall과 precision의 성능 저하를 동시에 가져올 수 있으며 이를 해결하기 위해서는 다른 특징을 고려하여야 한다. 그러므로 급격한 밝기 변화를 가지는 비디오 시퀀스에 대해 효과적인 비디오 색인이 가능하게 하기 위해서는 제안한 방법에서 사용된 밝기 변화를 고려한 알고리즘이 필수적이라 할 수 있다.

다양한 영상특성을 가지는 실험 시퀀스에 대한 제안한 알고리즘의 효율성을 검증하기 위해 칼라 뮤직비디오 시퀀스 (352×240)에 대해 실험하였다. 그림 3에 나타나 있는 바와 같이 뮤직비디오 시퀀스는 4914 프레임의 18개 샷으로 구성된 'Britney' 시퀀스와 7029프레임의 82개 샷으로 구성된 'Christina' 시퀀스를 사용하였으며 두 시퀀스 모두 다양한 조명 환경하에서 복잡한 밝기 변화를 가진다. 또한 뮤직비디오 특성상 움직임 크고 카메라 조작이나 이동이 심하다. 뮤직 비디오 시퀀스의 경우 심각한 조명변화로 인해 밝기 변화가 심한 많은 프레임에서 실제 샷경계인 지점보다 히스토그램차가 더 크게 나타날 수 있다. 그림에서 나타나는 바와 같이 동일한 샷내에서도 밝기 성분의 변화에 의해 영상의 밝기 특성이 크게 차이 나는 것을 알 수 있다.

실제 밝기 변화가 있는 연속된 프레임에서의 히스토그램이 그림 4에 나타나 있다. 그림 4(a)는 'Music Video 1' 시퀀스중 밝기 변화가 심한 프레임간의 히스토그램을 나타내고 그림 4(b)는 'Music Video 2' 내의 프레임간 히스토그램을 보여주고 있다. 그림에서 보여주듯이 같은 샷내의 인접한 프레임간의 히스토그램의 분포는 많은 차이를 나타낸다. 그림 4(a)의 'Music Video 1'에 대한 히스토그램에서는 프레임간 밝기 변화(brightness offset) 성분과 대조(contrast) 성분이 모두 많은 변화를 나타내고 그림 4(b)의 'Music Video 2'에 대한 히스토그램에서는 프레임간 대조 성분보다는 밝기 변화 성분이 크게 나타나고 있음을 알 수 있다.

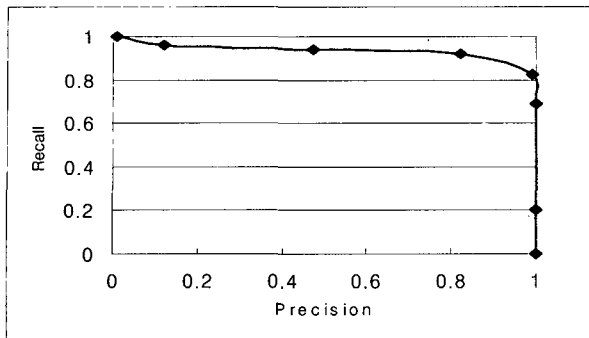
그림 4에 나타난 것과 같이 연속되는 프레임간의 히스토그램 차이가 큰 경우 기존의 칼라 히스토그램만을 사용한 알고리즘에서는 동일한 샷 내의 프레임들이지만 장면 전환이 일어난 것으로 판단하게 되는 것이다. 반면에 에지정합(EMR)과 밝기보상(FDV)을 사용한 제안한 알고리즘은 히스토그램 분포의 차이가 크더라도

에지 정합과 밝기 보상을 수행하므로 급격한 밝기 변화에 의한 히스토그램 분포의 차이를 구별하여 보다 견고한 성능을 보이게 된다.

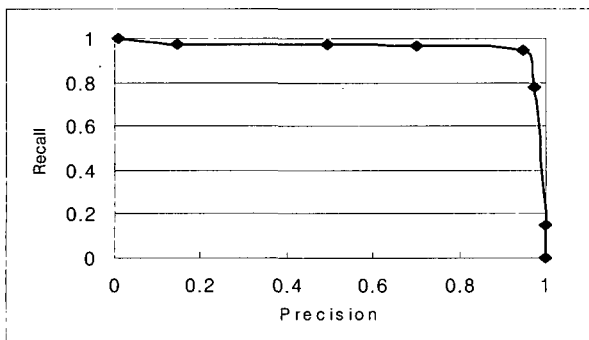
표 2는 실험에 사용한 뮤직 비디오 시퀀스에 대해 기존의 히스토그램 차, Chi-Square 테스트와 제안한 방법에 관해 실험한 결과를 나타내고 있다. 표 2에서 'Britney' 시퀀스에 대해서 기존의 히스토그램 차 방법과 Chi-Square 테스트는 낮은 recall과 precision 성능을 보이고 있는 반면 제안한 알고리즘에서는 향상된 recall과 precision 성능을 보여주고 있다. 'Christina' 시퀀스에 대한 비디오 색인 실험에서 'Christina' 시퀀스가 'Britney' 시퀀스에 비해 상대적으로 조명과 같은 밝기 변화 성분이 적기 때문에 'Christina' 시퀀스에 대한 기존 알고리즘의 성능은 'Britney' 시퀀스에 대한 성능에 비해 비교적 향상된 결과를 보이고 있고 제안한 알고리즘에서는 실험에 사용한 두가지 뮤직 비디오 시퀀스에서 일정한 성능을 나타내었다.

뮤직 비디오를 실험영상으로 사용한 비디오 색인 실험에서도 제안한 알고리즘은 기존의 방법에 비해 향상된 성능을 보였으며 뉴스 비디오 시퀀스에 대한 실험 결과와 비교하여 비슷한 성능을 나타내었다. 뉴스 비디오에서의 밝기 변화 특성은 플래쉬라이트등의 급격한 밝기 성분의 증가가 많이 나타나고 뮤직 비디오에서는 조명의 변화에 의한 밝기 변화로서 대조 성분에도 많은 변화를 나타낸다. 두 종류의 시퀀스에 대한 실험 결과에서도 나타나듯이 기존의 비디오 색인 알고리즘은 밝기 변화의 특성보다는 밝기 변화 정도의 차이에 의해서 큰 성능차이를 보였다. 즉 프레임간의 밝기 특성이 일정한 간격을 두고 서서히 변화하는 경우는 성능저하가 비교적 적은 반면 밝기 변화가 급격한 경우는 밝기 특성과 관계 없이 심한 성능 저하를 나타내었다. 제안한 알고리즘에서는 에지특성을 고려하면서 밝기 보상에 의한 차이를 동시에 고려하므로써 에지정합만을 수행한 경우에 발생할 수 있는 오류나 누락검출을 감소시킬 수 있다.

표 2에서 제안한 알고리즘은 기존의 알고리즘에 비해 급격한 조명변화나 밝기변화에 의한 잘못된 샷경계 검출이나 누락된 검출을 크게 줄임으로써 recall이나 precision 면에서 모두 현저히 향상된 성능을 보였다. 테스트 시퀀스들에 대한 실험 결과에서 제안한 알고리즘은 기존의 알고리즘에 비해 30%에서 50% 이상의 향상된 성능을 보였다. 실험결과에서도 알 수 있는 바와 같이 밝기 성분의 보상과 칼라와 에지 특성을 결합한 알



(a)



(b)

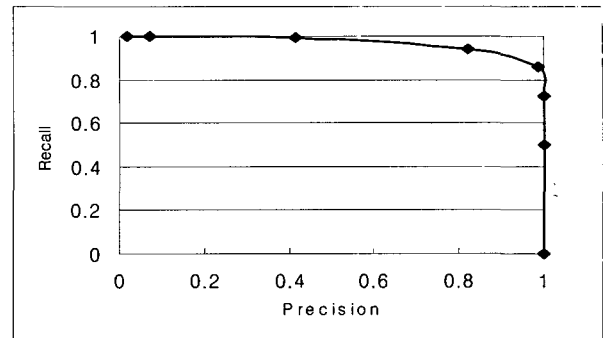
그림 5. 뉴스 시퀀스에 대한 제안한 알고리즘의 성능. (a) 'News Article 1' 시퀀스. (b) 'News Article 2' 시퀀스.

Fig. 5. Performance of proposed algorithm for news sequence. (a) 'News Article 1' sequence. (b) 'News Article 2' sequence.

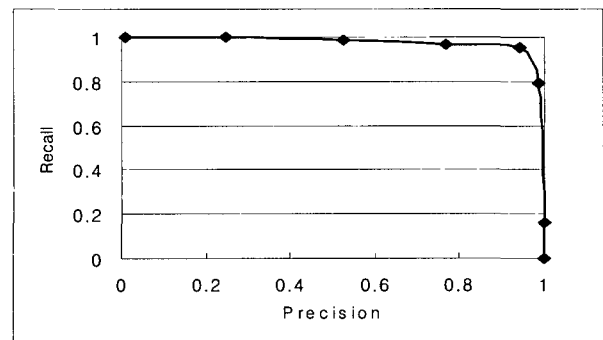
고리즘은 급격한 밝기 변화를 가지는 비디오 시퀀스에 대한 비디오 인덱싱의 성능을 크게 향상시킬 수 있다.

그림 5는 뉴스 시퀀스에 대한 제안한 알고리즘의 성능비교를 나타내고 있다. 그림 5에서 수평축은 precision을 수직축은 recall을 나타내고 있다. 제안한 알고리즘의 성능 변화에 가장 큰 역할은 하는 에지 정합의 문턱치를 변화시키면서 이에 따른 성능 변화를 나타내었고 precision과 recall은 tradeoff 관계를 보이고 있다. 에지 정합의 문턱치가 증가하면 누락 검출의 수는 줄어드나 오류 검출의 수가 늘어나게 되므로 recall은 높아지고 precision은 떨어지게 된다. 제안한 알고리즘에서는 recall과 precision의 성능을 동시에 만족시킬 수 있도록 문턱치를 설정하였으며 응용분야에 따라 문턱치를 조정하여 recall 또는 precision 어느 한 쪽의 성능이 크게 좋아지게 할 수 있다.

그림 6은 뮤직 비디오 시퀀스에 대한 제안한 알고리즘의 성능비교를 나타내고 있다. 그림 6에서도 에지 정합의 문턱치를 변화시키면서 이에 따른 성능 변화를 나



(a)



(b)

그림 6. 뮤직 비디오 시퀀스에 대한 제안한 알고리즘의 성능. (a) 'Britney' 시퀀스. (b) 'Christina' 시퀀스.

Fig. 6. Performance of proposed algorithm for music sequence. (a) 'Britney' sequence. (b) 'Christina' sequence.

타내었다. 뉴스 시퀀스에 대한 실험에 비해 recall이 약간 높은 수치를 보이고 있는데 이는 뮤직 비디오 시퀀스가ダイナミック한 장면 전환이 많이 일어나므로 누락 검출의 수가 감소하기 때문이다.

일반적으로 영화, 뉴스, 광고, 뮤직비디오 등의 각종 비디오 시퀀스에서 다양한 조명의 변화로 프레임간의 급격한 밝기 변화는 자주 발생할 수 있다. 이런 경우 기존의 방법에서는 급격한 성능저하를 보이는 반면 제안한 알고리즘에서는 밝기 변화가 크지 않은 시퀀스 뿐만 아니라 다양한 밝기성분의 변화를 가지는 시퀀스에 대해서도 장면전환 검출이나 비디오 색인에 효율적이다.

V. 결 론

본 논문에서는 급격한 밝기 변화가 있는 비디오 시퀀스에 대해서도 칼라와 에지 특성을 이용하여 효과적인 비디오 색인이 가능한 알고리즘을 제안하였다. 기존의 비디오 색인 알고리즘들은 밝기 변화가 큰 비디오 시퀀스에서 샷경계의 오류 검출등으로 인해 정확도가 떨어

지게 된다. 기존의 칼라 히스토그램만을 사용한 알고리즘의 한계를 극복하기 위하여 에지 정합과 밝기 보상을 수행함으로써 같은 장면내의 밝기 변화에도 견고한 성능을 가질 수 있도록 하였다.

제안한 알고리즘은 크게 두 단계로 구성되며 첫번째 단계에서 칼라 히스토그램을 사용하여 후보 샷경계 지점을 추출하고 두번째 단계에서 에지 정합과 밝기 보상을 이용하여 후보점들이 샷경계인지 밝기 변화인지를 결정한다. 샷내의 밝기 변화가 작은 경우 에지 정합과 밝기 보상은 샷경계에서만 일어나게 되므로 제안한 알고리즘은 밝기 변화를 가지는 시퀀스 뿐만 아니라 그렇지 않은 시퀀스에 대해서도 밝기 변화 고려에 따른계산량 증가를 현저히 줄일 수 있다. 효율적인 에지 정합과 밝기 보상이 가능하게 하기 위하여 EMR과 FDV를 제안하였다. 실제 비디오 시퀀스를 테스트 영상으로 실험한 결과 제안한 알고리즘이 기존의 방법에 비해 현저히 향상된 성능과 효율성을 나타내었다. 향후 복잡한 밝기 변화 성분을 가지는 다양한 비디오 시퀀스에 대한 알고리즘 성능 분석에 연구가 진행될 것이다.

참 고 문 헌

- [1] N. D. Doulamis, A. D. Doulamis, Y. S. Avrithis, K. S. Ntalianis, and S. D. Kollias, "Efficient summarization of stereoscopic video sequences," *IEEE Trans. Circuits and Systems for Video Technology*, vol. CSVT-10, no. 4, pp. 501-517, June 2000.
- [2] X. Gao and X. Tang, "Unsupervised video-shot segmentation and model-free anchorperson detection for news video story parsing," *IEEE Trans. Circuits and Systems for Video Technology*, vol. CSVT-12, no. 9, pp. 765-776, Sep. 2002.
- [3] N. V. Patel and I. K. Sethi, "Video shot detection and characterization for video database," *Pattern Recognition*, vol. 30, no. 4, pp. 583-592, Apr. 1997.
- [4] V. Kobla, D. Doermann, and K. I. Lin, "Archiving, indexing, and retrieval of video in compressed domain," in *Proc. SPIE Conf. Multimedia Storage and Archiving Systems*, vol. 2916, pp. 78-89, Boston, MA, USA, Nov. 1996.
- [5] J. Meng, Y. Juan, and S. F. Chang, "Scene change detection in a MPEG compressed video sequence," in *Proc. SPIE Symposium Digital Video Compression: Algorithms and Technologies*, vol. 2419, pp. 14-25, San Jose, CA, USA, Feb. 1995.
- [6] C.-L. Huang and B.-Y. Liao, "A robust scene change detection method for video segmentation," *IEEE Trans. Circuits and Systems for Video Technology*, vol. CSVT-11, no. 12, pp. 1281-1288, Dec. 2001.
- [7] B.-L. Yeo and B. Liu, "Rapid scene analysis on compressed video," *IEEE Trans. Circuits and Systems for Video Technology*, vol. CSVT-5, no. 6, pp. 533-544, Dec. 1995.
- [8] W. J. Heng and K. N. Ngan, "Post shot boundary detection technique: Flashlight scene determination," in *Proc. Fifth Int. Symposium on Signal Processing and Its Applications*, pp. 447-450, Brisbane, Australia, Aug. 1999.
- [9] R. Brunelli, O. Mich, and C. M. Modena, "A survey on the automatic indexing of video data," *Journal Visual Communication and Image Representation*, vol. 10, no. 2, pp. 78-112, June 1999.
- [10] U. Gargi, R. Kasturi, and S. H. Strayer, "Performance characterization of video-shot-change detection methods," *IEEE Trans. Circuits and Systems for Video Technology*, vol. CSVT-10, no. 1, pp. 533-544, Feb. 2000.
- [11] J. Hafner, H. S. Sawhney, W. Equitz, M. Flickner, and W. Niblack, "Efficient color histogram indexing for quadratic form distance functions," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. PAMI-17, no. 7, pp. 729-736, July 1995.
- [12] A. M. Ferman, A. M. Tekalp, and R. Mehrotra, "Robust color histogram descriptors for video segment retrieval and identification," *IEEE Trans. Image Processing*, vol. IP-11, no. 5, pp. 497-508, May 2002.
- [13] S. H. Kim and R.-H. Park, "A novel approach to scene change detection using a cross entropy," in *Proc. 2000 IEEE Int. Conf. Image Processing*, vol. 3, pp. 937-940, Vancouver, Canada, Sept. 2000.
- [14] J. Song and B.-L. Yeo, "Fast extraction of spatially reduced image sequences from MPEG-2 compressed video," *IEEE Trans. Circuits and Systems for Video Technology*, vol. CSVT-9, no.

- 7, pp. 1100-1114, Oct. 1999.
- [15] S.-W. Lee, Y.-M. Kim, and S. W. Choi, "Fast scene change detection using direct feature extraction from MPEG compressed videos," *IEEE Trans. Multimedia*, vol. 2, no. 4, pp. 240-254, Dec. 2000.
- [16] R. Kasturi and R. C. Jain, *Computer Vision: Principles*. IEEE Computer Society Press, Los Alamitos, CA, 1991.
- [17] S. H. Kim and R.-H. Park, "Robust video indexing for video sequences with complex brightness variations," in *Proc. IASTED Int. Conf. Signal and Image Processing*, pp. 410-414, Kauai, HI, USA, Aug. 2002.

 저자 소개



김 상 현(정회원)
 1997년 한국외국어대학교
 전자공학과 학사 졸업
 1999년 한국외국어대학교
 전자공학과 석사 졸업
 2003년 서강대학교 전자공학과
 박사 졸업

2003년 1월 ~ 2004년 5월 LG전자 DM연구소

선임연구원

2004년 6월 ~ 현재 삼성종합기술원 Digital연구소

전문연구원

<주관심분야: 비디오 색인 및 검색, 비디오 코딩,
 영상처리, 컴퓨터 비전>



박 래 홍(정회원)
 1976년 서울대학교 전자공학과
 학사 졸업
 1979년 서울대학교 전자공학과
 석사 졸업
 1981년 Stanford University
 (Stanford, CA)
 전기공학과 석사 졸업

1984년 Stanford University (Stanford, CA) 전기
 공학과 박사 졸업

1984년~현재 서강대학교 전자공학과 교수

1990년 University of Maryland (College Park,
 MD) 방문교수

<주관심 분야 : 영상통신, 컴퓨터 비전 및 패턴인
 식>