

# 투영된 모션과 히스토그램 인터섹션을 이용한 강건한 물체추적

이 봉 석<sup>†</sup> · 문 영 식<sup>\*\*</sup>

## 요 약

기존의 물체추적기법은 템플릿 매칭, 물체의 경계선 재 검출, 물체의 움직임 정보 등을 사용하여 수행되었다. 그러나, 템플릿 매칭의 경우 많은 계산 시간을 요구하고, 경계선을 재 검출하는 경우 윤곽선이 잘못 설정되는 경우가 있으며, 물체의 움직임을 사용하는 경우에는 움직이는 카메라에서 움직이는 물체만을 추적하기가 쉽지 않은 단점이 있다. 본 논문에서는 투영된 모션과 히스토그램 인터섹션을 이용한 강건한 물체추적 방법을 제안한다. 초기 객체추출은 영상분할 후 영역선택을 통하여 구성하고 선택된 객체를 가로 및 세로의 밝기 값을 1차원 신호로 투영하여 객체의 개략적인 평행이동 벡터를 추정한다. 추정된 변위를 기준으로 하여 객체의 가능한 회전 및 스케일에 대한 템플릿을 구성하고, 이들에 대하여 개선된 히스토그램 인터섹션을 사용하여 물체 추적을 수행한다. 제안한 알고리즘의 강건한 물체추적 성능을 실험에 의하여 확인하였다.

## Robust object tracking using projected motion and histogram intersection

Bong Seok Lee<sup>†</sup> · Young Shik Moon<sup>\*\*</sup>

### ABSTRACT

Existing methods of object tracking use template matching, re-detection of object boundaries or motion information. The template matching method requires very long computation time. The re-detection of object boundaries may produce false edges. The method using motion information shows poor tracking performance in moving camera. In this paper, a robust object tracking algorithm is proposed, using projected motion and histogram intersection. The initial object image is constructed by selecting the regions of interest after image segmentation. From the selected object, the approximate displacement of the object is computed by using 1-dimensional intensity projection in horizontal and vertical direction. Based on the estimated displacement, various template masks are constructed for possible orientations and scales of the object. The best template is selected by using the modified histogram intersection method. The robustness of the proposed tracking algorithm has been verified by experimental results.

키워드 : 물체추적(object tracking), 투영모션(projected motion), 히스토그램 인터섹션(histogram intersection)

### 1. 서 론

컴퓨터의 대중화, 통신기술의 향상과 영상기기의 급속한 발달로 우리 주변에서는 많은 양의 멀티미디어 자료들을 접하게 되었다. 그 결과로 현재는 인터넷을 통해 동영상을 실시간으로 전달받아 시청할 수 있다. 이로 인해 동영상을 통해 많은 정보를 전달하고자 하는 방법들이 개발되고 있으며, 동영상에서 관심 있는 객체를 추적하는 것은 많은 응용분야에 적용될 수 있는 필수 기술중 하나이다. MPEG-4에서는 객체 기반의 영상 압축 기법을 국제 표준으로 확정하였으나[1], 현

재의 기술로는 의미 있는 객체의 자동적인 분할은 많은 어려움이 있다. 본 논문에서는 인간의 개입을 최소화시켜 객체를 추출하는 방법과 빠른 시간에 객체의 움직임을 추적하여 회전, 스케일에 적응적이고 노이즈에 강건한 물체추적 알고리즘을 제안한다. 영상분할에 관한 연구는 불연속성, 에지 연결과 윤곽선, thresholding 등을 이용한 분할기법으로 나눌 수 있으며[2], 본 논문에서는 수리형태학적(morphological) 연산자를 이용한 영상분할 기법을 이용한다. 물체추적에 관한 연구는 템플릿 매칭 기법[2], affine 변환 기법[3], 블록정합 기법[4], 투영된 모션 기법[5] 등으로 나눌 수 있다. 템플릿 매칭은 이동, 회전, 스케일 된 템플릿으로 재구성하여 비교하면 어느 정도 나은 결과를 얻을 수 있지만 모든 영역을 탐색하기 때문에 계산량이 매우 많다는 단점이 있다. Affine 변환을 통한 움직임 추정 기법은 비교적 정확한 결과를 가져올 수

\* 본 연구는 한국과학재단 목적기초연구(과제번호 : 2000-2-303-005-3) 지원으로 수행되었음.

† 정 회 원 : 삼성전자 DCT 연구원

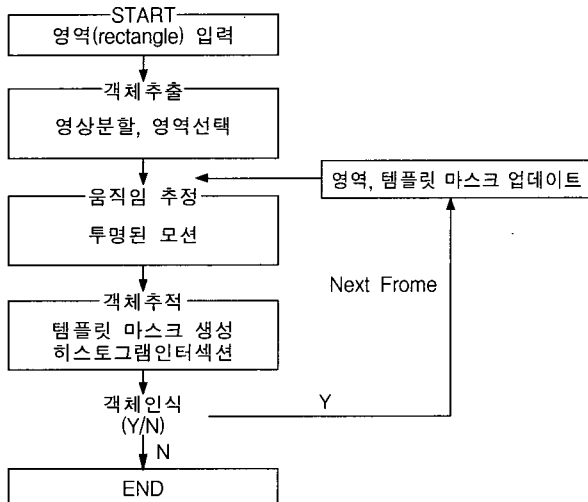
\*\* 정 회 원 : 한양대학교 컴퓨터공학과 부교수

논문접수 : 2001년 2월 26일, 심사완료 : 2002년 1월 3일

있지만 객체 내의 각 영역에 대하여 움직임 추정에 많은 연산 및 시간을 요구한다. 블록 정합 기법은 물체의 평행 이동에 대해서 최적의 움직임 벡터를 찾는 장점이 있는 반면에 물체가 회전, 스케일 변화되는 경우에는 추적하기 어려운 단점이 있다. 투영된 모션 기법은 가로 및 세로의 밝기 값을 1차원 신호로 투영하여 개략적인 움직임을 감지한다. 이 기법의 경우 상당히 빠른 시간에 물체의 움직임 위치를 찾아 낼 수 있으며 약간의 에러는 있지만 템플릿 마스크를 적용하여 히스토그램을 비교하면 에러를 보정할 수 있다. 객체인식의 경우 컬러특징 값을 이용한 방법, 모양 특징 값을 이용한 방법, 컬러와 텍스처 특징 값을 혼용하는 방법 등이 있다[6-8]. 본 논문에서는 영상 전체 면적으로 정규화 한 확률 값을 이용한 변형된 색상 히스토그램 인터섹션 기법을 사용한다.

## 2. 제안하는 알고리즘

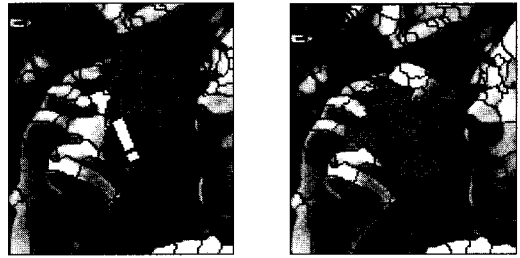
제안된 알고리즘은 영상분할, 객체추출, 투영된 모션, 템플릿 마스크 생성, 히스토그램 인터섹션을 사용한 물체인식으로 순으로 구성된다. (그림 1)은 전체적인 순서도이다. 일단 사용자로부터 객체가 존재하는 관심영역을 입력받고 입력된 영역을 영상분할 하여 영역을 선택하는 방법으로 객체를 추출한다. 추출된 객체에 대하여 가로, 세로로 1차원으로 투영하고, 밝기 값을 축적하여 얻어진 값으로 비교하는 투영된 모션을 사용하여 개략적인 물체의 움직임을 추정한다. 추정된 움직임 벡터를 기준으로 하여 물체의 가능한 회전과 스케일에 대한 템플릿 마스크를 생성하고 각각에 대하여 변경된 히스토그램 인터섹션 기법을 적용하여 물체의 회전, 스케일을 감지한다. 객체를 인식했을 경우 객체가 존재하는 새로운 영역과 템플릿 마스크를 업데이트하여 움직임 추정과 객체 추적 방법을 반복하며 객체를 인식 못했을 경우 제안된 방법은 종료된다.



(그림 1) 제안하는 방법의 전체 순서도

### 2.1 객체추출

관심있는 객체의 추출은 인간이 최소의 개입으로 빠른 시간에 에러 발생의 소지가 없고 비디오의 종류에 관계없이 객체를 추출할 수 있는 기법을 사용한다. 먼저 Demin Wang이 제안한 기울기와 워터셰드(Watershed)를 이용한 기법을 사용하여 영상을 분할하고[9], 분할된 영상에서 영역을 선택함으로써 객체를 결정한다. (그림 2)는 분할된 영상으로부터 마우스를 이용하여 원하는 객체를 추출하는 것을 보여준다.



(그림 2) 객체추출

### 2.2 투영된 모션을 이용한 움직임 추정

추출된 객체 영역에 대하여 투영된 모션을 이용하여 객체의 이동 변위를 계산한다. (그림 3)은 2차원 영상에 대하여 종, 횡으로 각각 투영된 값들을 나타낸 것이다. 투영의 방법은 각 방향에 따라 화소의 밝기 값을 누적하는 것으로 가로(x), 세로(y) 방향으로 각각 1차원 신호의 투영정보를 만들 수 있다. 투영은 주어진 영상 전체에 대하여 투영을 하는 것이 아니라 사용자가 추적하고자 하는 객체(영역)에 대해서 가로, 세로 길이에  $+a$  값을 주어서 투영을 하고 그 투영된 정보들이 다음 영상의 투영정보와 잘 정합 되는 x축 및 y축의 이동변위 값을 구함으로써 움직임 정보를 얻을 수 있다. 여기서  $+a$  값을 주는 이유는 객체 영상이 매우 작을 경우 큰 에러가 생길 수 있기 때문에 에러 보정을 위해 사용하였다.



(그림 3) 1차원 신호로 투영된 영상

k 번째 영상과 k+1 번째 영상에 대하여 i 번째 행(column)을 세로 방향으로 누적한 값을 각각  $X_k(i), X_{k+1}(i)$ 라하고, i 번째 열(row)을 가로 방향으로 누적한 값을 각각  $Y_k(i), Y_{k+1}(i)$

라 할 경우 수평과 수직에 대하여 객체의 이동변위  $dx$ ,  $dy$ 는 식 (1), 식 (2)와 같이 구할 수 있다[5].

$$dx = \arg \min_{\tau} \left\{ \left| \sum_{i=x_1}^{x_2} |X_k(i) - X_{k+1}(i + \tau)| \right| \right\}, -SR \leq \tau \leq SR \quad (1)$$

$$dy = \arg \min_{\tau} \left\{ \left| \sum_{i=y_1}^{y_2} |Y_k(i) - Y_{k+1}(i + \tau)| \right| \right\}, -SR \leq \tau \leq SR \quad (2)$$

여기서 추적하고자 하는 객체를 포함하는 직사각형 영역의 좌상, 우하의 좌표를  $(x_1, y_1)$ ,  $(x_2, y_2)$ 로 표현하면 식 (1), 식 (2)에서 이동변위의 탐색 범위를 가로, 세로 각각  $-SR \leq \tau \leq SR$  구간 내로 한정하여 1차원 투영 정보가 가장 잘 정합하는 부분을 해당 방향의 변위로 결정한다.

### 2.3 객체추적

투영된 모션으로부터 추정된 움직임 벡터를 기준으로 하여 객체의 가능한 회전, 스케일의 템플릿을 구성한 후, 이중 최적의 템플릿을 선택한다. (그림 4)는 템플릿 마스크를 얻어오는 과정을 보여준다. 얻어진 템플릿 마스크를 스케일과 회전의 변화에 적용하기 위해 템플릿 마스크를 변화시킨다. (그림 5)는 스케일된 영상과 회전된 영상을 보여준다. 본 논문에서는 템플릿 마스크 회전 각도와 스케일 변화는 각각  $\pm 10^\circ$ 와  $\pm 10\%$ 를 적용하며 업데이트된 템플릿 마스크는 다음 프레임에서 그대로 적용된다.



(그림 4) 템플릿 마스크 추출



(그림 5) 스케일 및 회전된 템플릿 마스크

객체인식을 위하여 템플릿 마스크로 정의된 질의 영상과 프레임 영상 비교를 RGB 컬러 모델을 가지고 영상 전체면적

으로 정규화한 확률 값을 가지는 면적히스토그램 인터섹션 기법을 사용한다[6, 11]. 히스토그램 인터섹션 방법은 1991년 Swain과 Ballard가 처음으로 제안하였다[11]. 이 기법은 두 영상에 대하여 컬러 히스토그램을 구한 다음 두 히스토그램이 일치하는 정도를 나타내는 이론이며 정의는 식 (3)과 같다.

$$H(I, M) = \sum_{j=1}^n \min(I_j, M_j) \quad (3)$$

여기서  $I$ 와  $M$ 은 각각 질의 영상과 데이터베이스 영상이며  $I_j$ ,  $M_j$ 는 영상에 대한 히스토그램을 나타낸다. 식 (3)의 결과는 영상들간에 존재하는 유사한 컬러들의 픽셀 수의 합을 나타내며 0에서 1사이로 정규화하기 위해서 식 (4)로 표현한다.

$$H(I, M) = \frac{\sum_{j=1}^n \min(I_j, M_j)}{\sum_{j=1}^n M_j} \quad (4)$$

그러나, 이 방법은 한 영상의 전체적인 스케일이 서로 다른 경우에는 같은 영상이라도 비교 값이 작아지는 문제점이 있다. 예를 들면, 같은 그림이 2배로 커지면 면적은 4배로 증가되므로 비교값이  $\frac{1}{4}$ 로 낮아진다. 수정된 히스토그램 인터섹션 방법은 기존의 히스토그램 인터섹션 방법에서 각각의 빈들에 있는 픽셀 개수 대신에 영상 전체 면적으로 정규화한 확률 값으로 표현되어진다. 식 (5)는 이것을 표현한 것이며  $S_I$ ,  $S_M$ 은 영상의 면적을 뜻한다.

$$H(I, M) = \sum_{k=1}^n \left( \frac{I_j}{S_I}, \frac{M_j}{S_M} \right) \quad (5)$$

### 3. 실험 결과 및 분석

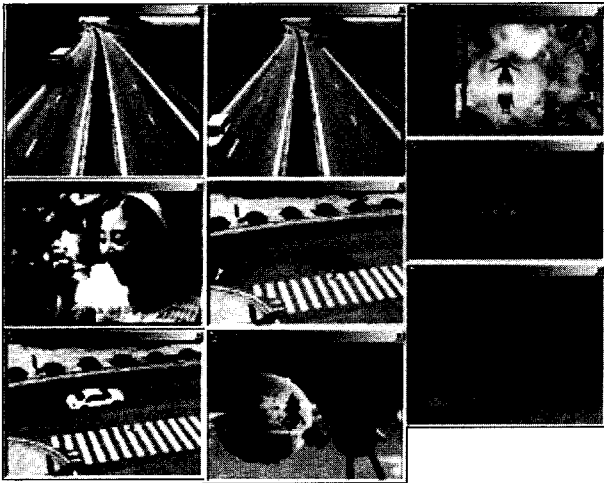
본 논문에서 제안한 방법과 기존의 방법의 성능을 분석하기 위하여 다양한 동영상에 대하여 실험을 수행하였다.

#### 3.1 실험 환경

시스템은 IBM PC Pentium III 600 MHz 컴퓨터를 사용하였고 프로그램 언어는 Visual C++ 6.0을 사용하였다. 실험에 사용되어진 동영상 종류는 <표 1>과 같고 총 동영상의 개수는 약 100여 개이다. (그림 6)은 실험에 사용된 동영상 샘플을 보여준다.

<표 1> 실험에 사용된 동영상 종류

동영상 종류	개 수
광 고	50
고속도로	20
도 로	10
기 타	20



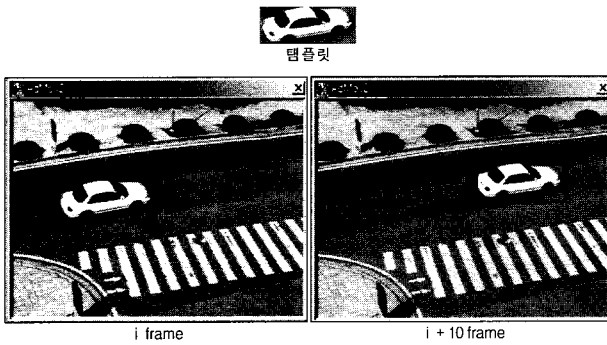
(그림 6) 동영상 샘플

### 3.2 실험결과

본 논문에서 제안한 알고리즘을 적용하여 실험용 동영상 샘플에 적용해 보았다.

#### 3.2.1 투영모션 기법

추출된 객체영역에 투영된 모션 기법을 적용하여 개략적인 물체의 움직임을 감지한다. (그림 7)은 투영모션 기법을 사용한 움직임 추정 결과를 보여준다.



(a) 자동차 영상



(b) 광고 영상

(그림 7) 투영모션 기법을 사용한 움직임 추정

#### 3.2.2 객체추적

투영모션 기법을 사용하여 물체가 움직인 위치를 대략적으로

알아낸 뒤 객체추적을 통하여 다시  $\pm 1$  or 2 pixel 정도 물체의 움직임 에러를 보정한 후 템플릿마스크를 회전과 스케일을 각각  $\pm 10^\circ$ ,  $\pm 10\%$  씩 적용하여 테스트하였다. (그림 8)은 템플릿마스크와 히스토그램 인터섹션 기법을 적용하여 객체를 추적한 결과를 보여준다.



-10도 회전, 비교값 0.87, 0.9배 축소

20도 회전, 비교값 0.85, 1.0배 확대

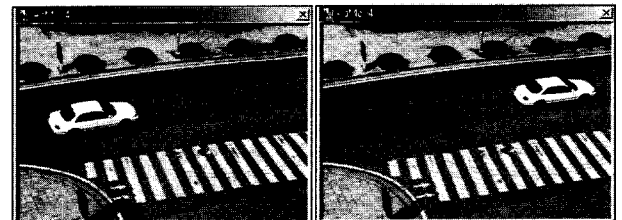
(a) 광고영상 1

(b) 광고영상 2

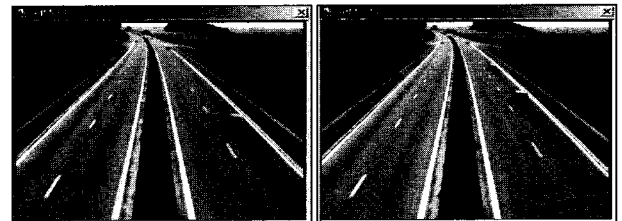
(그림 8) 객체추적 결과 영상

### 3.3 성능평가

물체추적은 Color Histogram Matching(CHM) 기법, Active Contour Model(ACM) 기법[10, 12]과 본 논문에서 제안한 알고리즘을 비교하였다. (그림 9)는 성능평가 실험에 사용된 동영상이다. 물체의 평행이동, 스케일 변화, 회전을 포함하는 동영상이다.



(a) 객체 이동



(b) 객체 이동, 스케일 변화



(c) 객체 이동, 스케일 변화, 객체 회전

(그림 9) 성능평가 실험에 사용된 동영상

주위 검색영역(Search Area)은 모두  $\pm 30$ 픽셀(pixel)로 정의하였고 20프레임씩 비교하였다. <표 2>는 실험동영상 (a), (b), (c)에 대한 수행시간을 비교한 결과이다. 비교시간은 객체의 크기에 따라 변한다.

<표 2> 물체추적 수행시간 비교

동영상	제안된 방법	CHM 기법	ACM 기법
(a)	1.07	11.08	2.13
(b)	0.21	2.98	2.19
(c)	1.16	15.48	3.41

<표 3>은 실험동영상 (a), (b), (c)에 대한 물체추적 성공 여부 결과이다. 그 기준은 물체가 사라질 때까지 추적이 가능하면 성공, 그렇지 않으면 실패로 정하였다.

<표 3> 물체추적 성공여부

동영상	제안된 방법	CHM 방법	ACM 방법
(a)	성공	성공	성공
(b)	성공	실패	성공
(c)	성공	실패	실패

<표 4>는 크기가  $352 \times 240$ 인 동영상에서 물체추적 기법들에 대한 성능분석을 보여준다. 전체적으로 제안된 방법이 다른 기법들 보다 우수함을 알 수 있다. 확대, 축소, 회전에 대해서는 CHM 방법이 좋지 않은 결과를 보였으며 수행시간에 대해서도 제안된 방법이 우수함을 알 수 있다. 전체적으로는 ACM방법이 가장 정확하게 객체를 추적하였으나 카메라 모션이 존재하는 경우, 추적 성능이 떨어지게 된다.

<표 4> 물체추적 기법들 성능 분석

비교대상	제안된 방법	CHM 방법	ACM 방법
확대 및 축소	O	X	O
회전	O	X	O
수행시간(sec)	0.2 ... 1.5	2.0 ... 20.0	2.0 ... 5.0
카메라모션	O	O	X

#### 4. 결론 및 향후과제

본 논문에서는 투영된 모션과 히스토그램 인터섹션 기법을 이용한 노이즈에 강건한 물체추적 알고리즘을 제안하였다. 초기영상은 영상분할 후 영역 선택을 통하여 구성하고, 움직임 추정은 가로 및 세로의 밝기 값을 1차원 신호로 투영하여 대략적인 움직임을 감지하며, 물체추적은 질의영상의 템플릿 마스크를 이동하여 회전과 스케일에 맞게 변경한 후 감지하였다.

Color Histogram Matching 기법, Active Contour Model 기법과 비교해 본 결과 검색 시간에서 우수한 결과를 보였고 객체의 회전 및 스케일 변화, 카메라 모션에도 적응적임을 확인하였다.

향후 연구과제는 배경과 객체를 자동으로 분리하는 방법과 템플릿 마스크를 사용하지 않고 질의영상의 특징을 보다 빠르고 효과적으로 나타낼 수 있는 방법의 연구가 가장 큰 과제이다. 동일한 배경에 동일한 객체가 여러 개 존재하는 경우, 객체의 한 부분이 다른 객체에 의해 가려졌을 경우, 객체가 화면 옆으로 사라지는 경우에도 추적이 가능하도록 연구되어야 한다.

#### 참고 문헌

- [1] "MPEG-4 Video Verification Model Version 13.0," ISO/IEC JTC1/SC29/WG11 MPEG99/N2687, 1999.
- [2] R. C. Gonzalez and R. E. Woods, Digital Image Processing, 1992.
- [3] R. Crane, "A Simplified Approach to Image Processing," Prentice Hall, pp.206-211, 1997.
- [4] 서재수, 남재열, 박진석, 이명호, "블럭 정합 움직임 추정을 위한 적응적 예측 방향성 탐색 알고리즘", 제12회 영상처리 및 이해에 관한 워크샵 발표논문집, pp.415-420, 2000.
- [5] 박동권, 윤호석, 전우성, 원치선, "투영된 모션을 이용한 반자동 객체 추적", 제12회 영상처리 및 이해에 관한 워크샵 발표 논문집, pp.139-144, 2000.
- [6] J. B. Oh and Y. S. Moon, "Content-Based Image Retrieval Based on Scale-Space Theory," IEICE Trans. Fundamental, June, 1999.
- [7] A. K. Jain and A. Vailaya, "Image Retrieval Using Color and Shape," Pattern Recognition, Vol.29, No.8, pp.1233-1244, 1996.
- [8] C. Carson, S. Belongie, H. Greenspan, and J. Malik, "Region-Based Image Querying," Proc. Workshop on Content-Based Access of Image and Video Libraries, 1997.
- [9] D. Wang, "Unsupervised Video Segmentation Based on Watersheds and Temporal Tracking," IEEE Trans. Circuits and System for Video Technology, Vol.8, No.5, pp.539-546, Sep. 1998.
- [10] N. Paragios and R. Deriche, "Geodesic Active Contours and Level Sets for the Detection and Tracking of Moving Objects," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.22, No.3, pp.266-280, Mar. 2000.
- [11] M. J. Swain and D. H. Ballard, "Color Indexing," International Journal of Computer Vision, Vol.7, No.1, pp.11-32, 1991.
- [12] J. A. Sethian, Level Set Methods and Fast Marching Methods, 1999.



### 이 봉 석

e-mail : superlee@samsung.co.kr  
1999년 한양대학교 전자계산학과 졸업(학사)  
2001년 한양대학교 대학원 컴퓨터공학과  
(공학석사)  
2002년~현재 삼성전자 DCT 연구원  
관심분야 : Face detection & recognition,  
Content based Image retrieval



### 문 영 식

e-mail : ysmoon@cse.hanyang.ac.kr  
1980년 서울대학교 공과대학 전자공학과  
졸업(학사)  
1982년 한국과학기술원 전기 및 전자공학과  
졸업(석사)  
1990년 University of California at Irvine  
Dept. of Electrical and Computer  
Engr.(박사)  
1982년~1985년 한국전자통신연구소 연구원  
1989년~1990년 Inno Vision Medical 선임연구원  
1990년~1992년 생산기술연구원 선임연구원  
1992년~현재 한양대학교 컴퓨터공학과 부교수  
관심분야 : 컴퓨터 비전, 패턴인식 등