

MBR을 이용한 실시간 영상추적 시스템 개발

김희숙^{1*}

A Development of Video Tracking System on Real Time Using MBR

Hee-Sook Kim^{1*}

요 약 실시간 영상에서 객체 추적은 지난 수년 동안 컴퓨터 비전과 많은 실제 응용 분야에서 관심있는 분야이다. 그러나 때때로 시스템들은 배경 잡음을 객체로 인식하여 객체를 찾지 못하였다. 이 논문에서는 실시간으로 적용하는 배경이미지를 이용하여 객체의 추출과 추적을 위한 새로운 방법을 개발하였다. 배경이미지의 잡음을 없애고 조도에 영향 받지 않는 객체를 추출하기 위하여 이 시스템은 실시간적으로 배경이미지를 갱신하여 적응적인 배경이미지를 생성한다. 이 시스템의 객체 추출은 배경이미지와 카메라로부터 입력된 이미지의 차를 이용한다. MBR(Minimum Bounding Rectangle)을 셋팅 한 후 추출된 객체의 내부점을 이용하고, 시스템은 이 MBR을 통하여 객체를 추적한다. 추가로 본 논문은 기존의 추적 알고리즘과 비교된 제안한 방법의 수행에 대한 결과를 평가했다.

Abstract Object tracking in a real time image is one of interesting subjects in computer vision and many practical application fields past couple of years. But sometimes existing systems cannot find object by recognize background noise as object. This paper proposes a method of object detection and tracking using adaptive background image in real time. To detect object which does not influenced by illumination and remove noise in background image, this system generates adaptive background image by real time background image updating. This system detects object using the difference between background image and input image from camera. After setting up MBR(minimum bounding rectangle) using the internal point of detected object, the system tracks object through this MBR. In addition, this paper evaluates the test result about performance of proposed method as compared with existing tracking algorithm.

1. 서 론

실시간 영상에서 객체 추적을 위해 컴퓨터 비전 기술을 이용한 구현은 매우 어려운 일임에도 불구하고, 영상 처리 기법의 발전과 더불어 객체 인식과 객체 추적에 관한 연구가 활발히 진행되고 있다. 객체 추적은 카메라로부터 입력된 영상에서 움직임을 보이는 객체를 인식하고, 그 움직임을 추정하여 추적하는 것이다. 객체를 추적하는 방법은 보안, 의료, 군사, 교통, 제어분야 등 여러 분야에 응용될 수 있어 그동안 많은 연구와 개발이 이루어지고 있다[1]. 강인한 실시간 객체 추적 시스템을 구현하기 위하여 고가의 장비가 필요하고, 움직임을 감지하기 위해 센서를 장착한 카메라나 팬/틸트로 움직이는 카메라가 요구된다. 객체

추적 방법은 각각 특정 환경에서 적합한 동작을 보이도록 몇 가지 제약 조건을 두어 설계되고 있다[2]. 본 논문에서는 보안 및 감시 시스템 분야에서 적용되어 질 수 있는 방법을 제안하기 위하여, 저가형 PC카메라의 움직임이 고정되어 있고 외부 환경이 아닌 실내 환경이며 배경영상의 변화가 거의 없다는 특수 환경으로 제약 조건을 가진다. 제안하는 시스템은 시간의 경과에 따라 배경영역에서 잡음생성을 줄이고자 객체 영역 이외의 영역을 배경영역으로 갱신하여 항상 최신의 배경영상을 유지하도록 하였다. 이때 연산의 양을 줄이기 위하여 그물식 탐색방법을 이용한다. MBR 내의 객체로부터 얼굴영역을 추출한다. 본 논문은 실시간으로 입력되는 영상으로부터 적응적 배경영상으로 잡음을 현저히 줄였고 실시간으로 얼굴 영역을 추출함으로써 보안 및 감시 시스템으로 효용성을 향상시켰다.

¹한국폴리텍IV 아산대학 멀티미디어과 조교수

*교신저자 : 김희숙(prima@kopo.ac.kr)

II. 관련연구

기존의 추적 알고리즘은 크게 특징기반방법(feature-based approaches)과 이미지기반방법(image-based approaches)으로 나누어진다. 특징기반방법은 저수준 해석(Low-Level Analysis), 특징 해석(Feature Analysis), 능동 형상 모델(Active Shape Models) 방법이 있고, 이미지기반방법에는 선형 부분 공간 방법(Linear Subspace Methods), 신경망(Neural Networks), 통계적 방법(Statistical Approaches) 등이 있다. 기존의 많은 방법들이 객체 추적 및 얼굴 추출을 위하여 제안되어 왔다. 발표된 논문의 결과와 구현을 통한 각각의 방법들을 비교한 결과가 <표 1>에 나타나 있다. 객체 추적과 얼굴 추출에 우수한 성능을 보이는 방법은 상대적으로 처리 속도가 떨어짐을 볼 수 있다. 반대로 처리 속도가 빨라 많은 프레임을 처리할 수 있는 방법은 객체 추출과 얼굴 추출에 뛰어난 성능을 보이지 못하고 있다.

표 1. 기존 알고리즘 성능 분석

구 분	FPS	Detection rate	Tracking rate
저수준 해석	◎	○	○
특징 해석	○	△	△
능동 형상 모델	○	◎	○
선형 부분 공간 방법	△	◎	○
신경망	△	◎	○
통계적 방법	△	○	◎

◎ : 우수, ○ : 양호, △ : 보통)

II. 실시간 객체 추적 및 얼굴영역 추출

2.1 시스템 구조

본 논문에서 효율적인 실시간 객체 추적 및 얼굴영역 추출을 위하여 다음과 같은 제약 조건을 가진다. 센서 장착이나 움직임이 가능한 고가의 카메라가 아닌 저가형 PC카메라를 이용하는데 이는 카메라의 움직임이 고정됨을 의미한다. 또한 공원이거나 거리등의 야외 환경이 아닌 사무실이나 연구실, 기자재실, 자재 창고, 주차장, 은행 365일 코너 등 배경영상의 변화가 미미한 특수 환경으로 환경을 제한한다. PC카메라를 이용하여 실시간으로 받아들여지는 영상 중에서 객체의 출현이 감지되지 않는 영상을 초기의 배경영상으로

로 설정하여 준다. 이 후, 연속적으로 입력되는 영상과 초기 배경영상의 비교를 통하여 강인한 객체 추출을 할 수 있는 적응적 배경영상을 생성한다. 그물식 탐색 방법을 이용하여 객체의 출현을 감지하고, 감지된 객체의 내부점을 추출한다. 추출된 객체 내부점을 이용하여 객체를 포함하는 최소 사각형인 MBR을 설정해 주고, 영상으로부터 일련의 처리를 반복하게 된다. 연속된 MBR의 설정을 추적함으로써 객체의 실시간 추적을 가능토록 하였다.

2.2 시스템 구현 방법

2.2.1 적응적 배경영상 생성

본 시스템에서 사용되는 카메라는 저가이며 고성능이 아닌 일반 PC카메라이기 때문에 빛에 대해 상당히 민감한 반응을 보인다. 배경영상과 입력영상의 차이가 균집을 이루고 있으면 배경영상임에도 불구하고 새로운 객체로 인식하는 오류를 범할 수 있다. 여기서 우리는 잡음을 제거하여 보다 정확한 객체의 추출을 위해 적응적 배경영상을 생성한다. 전체적인 배경영상의 갱신은 많은 연산량을 필요로 하기 때문에 실시간 객체 추적 성능을 저하시킨다. 제한하는 시스템에서는 객체영역인 $n * m$ 의 영역을 제외한 나머지 영역을 배경영상으로 갱신하면서 객체 추적을 동시에 이루어질 수 있도록 하였다.

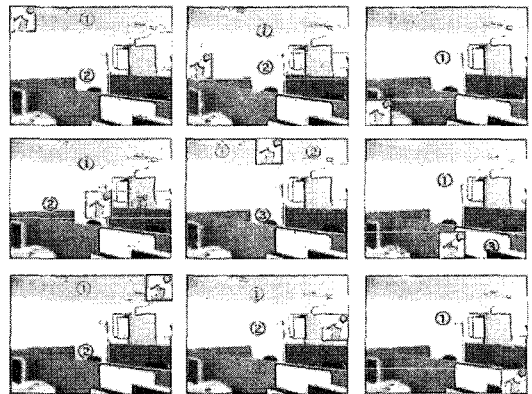


그림 1. 적응적 배경영상 생성

<그림 1>는 적응적 배경영상 생성 방법을 나타낸다. 이전의 배경영상과 입력영상에서 객체 추출 후 객체영역을 제외한 나머지 영역을 새로운 배경영상으로 대체한다. 입력영상에서 객체를 찾아내고 객체의 위치에 따라 객체부분이 아닌 사각영역을 배경영상으로 판단한다. 즉 객체의 위치에 따라 ①-④ 부분으로 분

류하고 입력영상을 배경영상으로 대체한다. 제안한 적응적 배경영상 방법을 이용하여 실험한 결과, 시변하는 배경영상내의 조명등에 의한 잡음을 객체 내부에 포함되어 있는 배경영역을 제외한 나머지 영역에서 95% 이상 제거할 수 있었다.

2.2.2 객체의 내부점 추출

객체의 추적을 위해서 객체의 위치 결정을 위한 객체 추출이 우선적이다. 완전 차영상 방식을 도입한 기존 시스템과는 달리, 제안하는 시스템에서는 그물식 탐색 방식을 이용한다. 그물식 탐색 방식은 영상을 위에서 아래로 δ 픽셀 간격으로 탐색해 나가는 방법으로 완전 차영상을 이용하는 시스템보다 처리속도의 높은 향상을 보인다.

<그림 2>은 그물식 탐색 방식 및 객체 내부점 추출이다. 객체의 내부점 추출은 그물식 탐색을 하면서 배경영상과 입력영상의 차영상을 이용하게 된다. 배경영상과 입력영상의 탐색라인상의 픽셀들을 각각 RGB 채널로 분할한다.

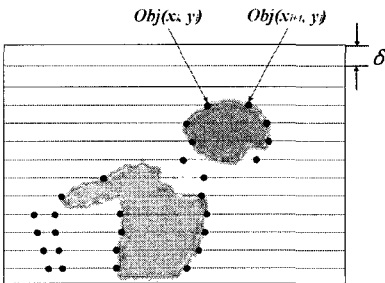


그림 2. 그물식 탐색 방식

2.2.3 MBR을 이용한 객체 추적

추출된 객체 내부점을 이용하여 MBR을 설정하여 준다. MBR은 객체 추적을 위한 객체를 포함하는 최소 사각 영역으로써, 얼굴 영역을 추출할 때 속도를 향상시키기 위한 범위를 제한해 주는 역할을 수행한다. 보다 정확한 MBR의 설정을 위한 전처리 과정으로 Median Filtering을 통해 잡음을 제거하였다. 제안하는 시스템에서는 완전 차영상이 아닌 그물식 탐색 방법을 이용하기 때문에 horizontal median filter window를 이용한다. 또한 컬러 영상에 적용되는 median filtering을 위한 절차를 따른다. (식 1)과 같이 각 화소 샘플에 대하여 각각의 RGB 채널에 대한 차들을 합한다. 이중 가장 작은 값을 가진 Distance_i는 필터의 출력인 x_i에 대응된다.

$$\sum_{i=1}^N |x_{med} - x_i| \leq \sum_{i=1}^N |y - x_i|$$

$$Distance_i = \sum_{j=1}^N (|red-red_j| + |green-green_j| + |blue-blue_j|) \tag{식1}$$

$x_{med} = x_i$ (단, i 는 $Distance_i$ 중 최소값에 대응하는 값)

* $Distance_i$: 각 컬러 요소들에 대한 차

* N : 필터 윈도우에 표현되는 샘플수

* i : 처리되는 화소

* j : 다른 화소 샘플

객체 내부점 $Obj(x_b, y_b)$ 의 x, y 좌표 중 (식 2)를 통하여 각각의 최대, 최소 좌표를 구하여 객체를 포함하는 최소한의 사각 영역을 설정해 준다.

$$Obj_left = \min [Obj(x_i)] - \delta$$

$$Obj_top = \min [Obj(y_j)] - \delta$$

$$Obj_right = \max [Obj(x_i)] + \delta$$

$$Obj_bottom = \max [Obj(y_j)] + \delta$$

$$MBR = \{Obj_left, Obj_top, Obj_right, Obj_bottom\} \tag{식 2}$$

* $Obj(x_b, y_b)$: 객체 내부점

* $Obj_left(top, right, bottom)$: MBR의 최대, 최소 좌표

2.2.4 얼굴 영역 추출

HSI 색상계는 색상(hue)과 채도(saturation) 만을 사용함으로써 조명의 영향을 흡수할 수 있다. HSI 색상계는 인간 시각 시스템의 색채 감지 특성에 기초한 영상처리 알고리즘 개발을 위한 이상적인 도구이다. HSI 색상계의 변환에서 가장 복잡한 부분은 색상의 계산 부분이다. 색상은 RGB의 혼합 비율을 빨강을 기준으로 하는 일반 전각으로 표현한다.

얼굴 영역 추출은 (식 3)에 따라 HSI 좌표계로 변환된 영상의 색상정보에서 피부영역으로 알려진 범위 값에서 추출된 부분을 군집화 하여 선택한다.

$$H = \cos^{-1} \left[\frac{\frac{1}{2} [(R-G) + (R-B)]}{\sqrt{(R-G)^2 + (R-B)(G-B)}} \right]$$

$$S = 1 - \frac{3}{(R+G+B)} [\min(R, G, B)]$$

$$I = \frac{1}{3} (R+G+B) \tag{식3}$$

색상을 통한 추출 방법은 수행속도가 빠르다는 장점이 있지만 조명을 통해 영상의 변화가 많이 일어나며 명도에 대한 정규화 과정이 필요하다. 이에 HSI 색상좌표계는 색상)과 채도 값만을 사용함으로 인해 조명의 영향을 감소시킬 수 있는 장점이 있다. HSI 모델 실험에서 피부색의 범위를 다음과 같이 사용하였다.

$$H_{\min} = 0^{\circ}, H_{\max} = 50^{\circ}$$

$$S_{\min} = 0.23, S_{\max} = 0.68$$

카메라를 통해 실시간으로 입력되는 영상의 픽셀당 RGB값을 HSI 좌표계로 변환한다. 변환값에서 피부색 값에 유사한 H, S 범위의 값을 임계값으로 설정하여 벗어나는 경우는 제거한다. 이렇게 추출된 값들 중 이웃하는 픽셀들을 연결하여 군집화하고 가장 큰 군집 영역을 1차 얼굴 영역으로 선택한다. 그러나 이 방법은 작은 영역에서도 다양한 군집을 나타내는 문제점을 가지고 있다. 이러한 단점을 해결하기 위해 median filtering 방법을 이용하여 잡음을 제거하여 필요영역만을 군집화한다. <그림 3>은 입력되어진 영상에서 HSI 색상계를 이용한 얼굴 영역 추출 결과를 보이고 있다.

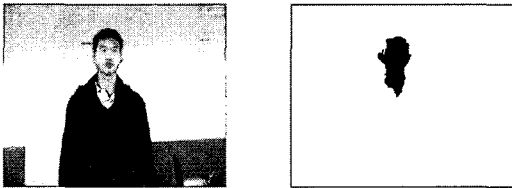


그림 3. HSI 색상계를 이용한 얼굴 영역 추출

HSI를 이용한 추출 방법은 조명에 덜 영향을 받으면서 얼굴 영역을 추출하는 장점이 있지만, 어두운 영역에서는 얼굴 영역의 추출에 단점을 가지고 있다. 이에 정규화 된 RGB영역을 이용하여 추출할 수 있다.

정규화된 RGB를 이용한 방법은 정규화한 값을 사용하여 HSI와 유사한 효과를 얻었다. 이때 $r'+g'+b'=1$ 이라는 성질을 이용하여 r', b' 의 두 채널 정보만을 사용하여 색을 구분하는 기준으로 사용할 수 있다. RGB 좌표계에서 색상의 정규화 식은 (식 4)과 같다.

$$(r', g', b') = \left(\frac{r}{r+g+b}, \frac{g}{r+g+b}, \frac{b}{r+g+b} \right) \quad (\text{식 4})$$

이 방법은 r, b의 값만을 이용하여 얼굴 영역 추출

에 이용할 수 있다. 정규화 된 RGB 모델 실험에서 피부색의 범위를 다음과 같이 사용하였다.

$$r'_{\min} = 0.3550, r'_{\max} = 0.4300$$

$$b'_{\min} = 0.2500, b'_{\max} = 0.3050$$

이 방식은 어두운 영역에서 후보 추출에 장점이 있으나 유사 색상영역에서는 구별이 어렵다. 이에 본 논문에서는 두 가지 방법을 혼합한 hybrid 방법론을 이용하였다. <그림 4>은 입력되어진 영상에서 HSI 색상계를 이용한 얼굴 영역 추출 결과를 보이고 있다.

추출된 얼굴 영역을 포함하는 최소 사각 영역 MBRF (Minimum Bounding Rectangle for Face Region)을 설정해 줌으로써 얼굴 영역을 추출한다. 이 영역은 추후 사람 얼굴 인식 분야와의 결합을 통해 보안 및 감시 시스템의 효용성을 높일 수 있다.

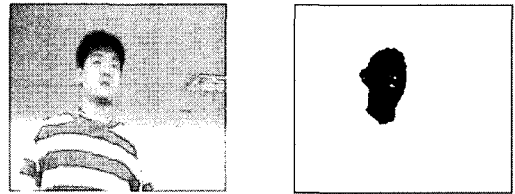


그림 4. 정규화된 RGB 색상계를 이용한 얼굴 영역추출

III. 구현 및 실험

3.1 실험 환경

본 실험은 카메라의 움직임이 고정되어 있고 배경 영상의 변화가 거의 없다는 환경으로 제약하고 입력 영상을 실시간으로 받아들여 객체의 위치를 탐지하고, 이동 객체의 추적을 실험하였다. 저가형 PC카메라로부터 받아들이는 배경영상과 입력영상의 크기는 320×240 의 24bit 컬러영상과 160×120 의 24bit 컬러영상을 이용하였다.

3.2 시스템 구현

<그림 5>는 구현된 실시간 객체 추적 및 얼굴영역 추출 시스템이다. 인터페이스의 상단부에는 제한한 시스템에서 사용되는 임계값과 변수에 대한 값을 입력하여 환경과 상황에 가장 적합한 객체 추적을 가능하도록 하였으며, 입력되어지는 동영상의 크기, 해상도 등을 선택적으로 입력되어지게 하였다. 배경영상 부분

은 계속적으로 갱신되고 있는 배경영상을 보이고, 입력영상 부분은 실시간으로 입력되는 영상을 보여준다. 왼쪽 하단부는 처리되어진 프레임 수와 초당 처리 속도(FPS, Frame Per Second)등을 나타내 준다.

객체추출 및 추적부분은 객체를 추출하고 추적하는 모습을 보이는데, 입력된 영상으로부터 객체를 잡음 없이 추출하고 실시간 객체 추적이 가능하다. 얼굴추출 부분은 작은 영상으로부터 선택되어져 얼굴영역을 정확히 추출할 수 있었다. 저장장치에 JPEG 파일형식으로 시간대별 실시간 객체 추적 과정을 저장함으로써 움직임이 추출된 후 객체의 움직임을 추후 검색이 가능하도록 하였다.

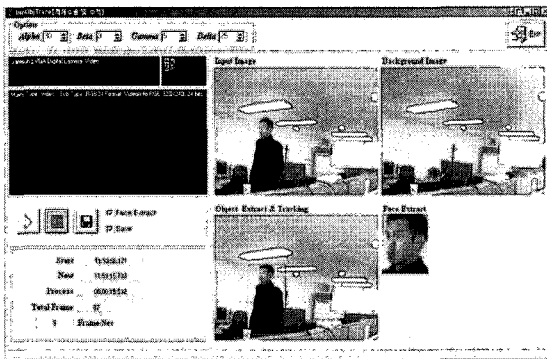


그림 5. 실시간 객체 추적 및 얼굴 영역 추출 시스템

3.3 실험결과

<표 2>는 본 논문에서 제안한 방법을 위한 임계값들과 변수에 대한 최적화 실험결과이다. α 가 25일 때 적응적 배경영상의 생성으로 많은 잡음을 제거할 수 있었고, β 가 30이었을 때 객체를 성공적으로 추출할 수 있었다. γ 가 5일 때에는 추출된 객체 내부점 후보군에서 많은 잡음 성공률을 보였고, δ 는 보다 조밀한 간격으로 탐색할 때 객체의 추출이 잘 이루어지나, 많은 연산량을 요구해 보다 빠른 탐색을 위하여 5를 선택하였다.

표 2. 임계값과 변수에 대한 최적화 실험 결과

임계값	1	2	...	5	...	25	...	30	...
α (적응적 배경영상 생성 임계값)	0.023	0.151	...	0.210	...	0.915	...	0.887	...
β (객체 내부점 후보군 추출 임계값)	0.101	0.256	...	0.494	...	0.856	...	0.934	...
γ (객체 내부점 추출 수)	0.495	0.531	...	0.926	...	0.349	...	0.324	...
δ (인접한 내부점 이진 수)	0.951	0.903	...	0.851	...	0.595	...	0.496	...

3.4 성능 평가

<그림 6>은 제안한 시스템과 기존 방법들과의 성능 평가이다. 추출율은 하나의 프레임에서 19,200(160 × 120)해상도 픽셀 중 실제 객체영역의 픽셀 수에 대한 추출된 객체영역의 픽셀 수를 나타낸다. 또한 추적율은 이전 프레임의 추출된 객체 영역 픽셀의 다음 프레임에서 추적된 성능을 나타낸다. 성능 평가에 사용된 영상은 160 × 120 해상도의 24 bit 컬러영상을 사용하였다.

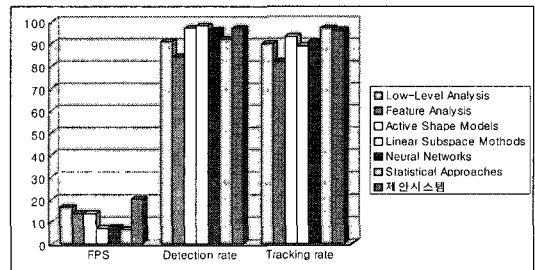


그림 6. 시스템 성능 비교

IV. 결론 및 향후 연구방향

본 논문에서는 입력되어지는 영상으로부터 배경영상의 실시간 갱신을 통해 객체의 위치를 탐지하여 객체를 추적하고, 객체의 내부점을 이용한 MBR에서의 얼굴영역을 추출하는 방법을 제안하였다. 실험은 고정 PC카메라와 배경영상의 변화가 거의 없다는 제한된 환경 조건에서 실시간으로 배경영상의 갱신과 객체의 추적이 안정적임을 보여주었다. 하지만 입력영상과 배경영상 각각의 R, G, B 값의 차를 이용, 객체의 부분인식을 인식할 때, 잡음과 조명의 영향으로 인하여 실험에 실패한 경우도 발생하여 이에 대한 보완이 필요하다. 그리고 레이블링을 이용한 다중객체의 추출과 추적이 실시간으로 이루어 질 수 있도록 연구가 진행 중이다.

참고문헌

[1]이희영, 최재영, 강동구, 김홍수, 차의영, 전태수, “배경영상을 이용한 목표물 추적에 관한 연구”, 한국멀티미디어학회 1999년도 춘계학술발표논문집 (학술발표), Vol.2, No.1, pp.386-390, 1999.

- [2] 황본우, 손형진, 이성환, "대화형 하이퍼 비디오 저작을 위한 객체 기반 추적 방법", 정보과학회 2001년 추계학술대회, Vol.28, No.2, pp.427-429, 2001.
- [3] E. Hjelm and B.K. Low, "Face Detection: A Survey." Computer Vision and Image Understanding, Vol.83, No.3, pp.236-274, 2001.
- [4] T.V. Pham, M. Worring, and A.W.M. Smeulders, "Face detection by aggregated Bayesian network classifiers." Technical Report 2001-04, Intelligent Sensory Information Systems Group, University of Amsterdam, 2001.
- [5] M. Rogers and J. Graham, "Robust active shape model search." Proceedings of the 7th European Conference on Computer Vision, No.4, pp.517-530, 2002.
- [6] K. Okada and C. von der Malsburg, "Analysis and synthesis of human faces with pose variations by a parametric piecewise linear subspace method." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, No.1, pp.761-768, 2001.
- [7] R. Feraud, O.J. Bernier, J.E. Viallet, and M. Collobert, "A Fast and Accurate Face Detection Based on Neural Network." IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.23, No.1, pp.42-53, 2001

김희숙(Hee-Sook Kim)

[정회원]



- 1991년 2월 : 원광대학교 전자계산공학과 졸업
- 1995년 2월 : 원광대학교 전자계산교육전공 (교육학석사)
- 2003년 2월 : 원광대학교 컴퓨터공학과 (공학박사)
- 2001년 10월 ~ 현재 : 한국폴리텍 IV 아산대학 멀티미디어과 조교수

<관심분야>

컴퓨터비전, 영상인식, 멀티미디어컨텐츠