

위너 최적화 기법을 이용한 영상기반 보행자 키 추정

전상희[†], 송종관^{**}, 박장식^{***}, 윤병우^{****}

Video Based Pedestrian Height Estimation Using Winer Optimization

Sang Hee Jeon[†], Jong Kwan Song^{**}, Jang Sik Park^{***}, Byung Woo Yoon^{****}

ABSTRACT

In this paper, we proposed a method which can detect pedestrians from CCTV video and estimate the height of the detected objects. We separate the foreground using Gaussian mixture model and the pedestrian is detected using the conditions such as the width-height ratio and the size of the candidate objects. In order to obtain the optimal model for estimating the height of pedestrian, we get many training data from the pedestrian whose height is known. Using these training data, we designed optimal Wiener height estimator and used to estimate the height of pedestrians. The height of the pedestrian at various distance is estimated and the accuracy is evaluated. In the experimental results, proposed method shows that it can estimate the height of pedestrian for various positions effectively.

Key words: Height Estimation, Gaussian Mixture Model, Wiener Estimation

1. 서 론

최근 영상보안 시스템(video surveillance system)은 영상분석(video analysis), 패턴인식(pattern recognition) 등의 기술을 적용하여 자동으로 사람을 인식하거나 화재를 감지하는 등 지능형으로 발전하고 있다. 영상보안 시스템에서 신원 확인을 위한 신체 정보(biometry)로는 얼굴, 걸음걸이, 키 등이 있고, 옷, 모자 등의 정보를 추가로 이용하는 방법들이 연구되고 있다. 사람의 신체정보 중 하나인 키(height)는 대상의 신원을 확인하는데 중요한 정보이다[1-6]. 키를 추정하기 위하여 소실점을 이용하는 방법 등이 제안되었으나 보행자 또는 사람의 측면을

촬영하여 키를 추정하였다. 방법 및 방재를 위하여 약 3에서 5 m 높이에 설치된 CCTV 카메라로 촬영한 영상에 대하여 키를 추정하기 위하여 영상 내에서 특정 물체의 높이를 기준으로 키를 추정하는 방법[7]이 연구되었다. 특정 물체의 높이를 기준으로 키를 추정하는 방법은 카메라로부터 원거리, 중거리, 근거리로 나누어 각각의 기준 높이와 화소의 비를 선형 추정하여 키를 추정한다.

본 논문에서는 키를 알고 있는 보행자가 CCTV 카메라 시야 내에서 이동하면서 영상 내에서의 화소 데이터를 수집하고, 영상 내에서 실제 키와 화소의 비를 계산하여 보행자들의 키를 추정하는 방법을 제안한다. 영상 좌표와 실제 키에 대한 화소의 비를 추

※ Corresponding Author: Jang Sik Park, Address: (608-736) 309, Suyeong-ro (Daeyeon-dong), Nam-gu, Busan, Korea, TEL: +82-51-663-4768, FAX: +82-51-624-5980, E-mail: jsipark@ks.ac.kr

Receipt date: Jan. 23, 2016, Approval date: Feb. 4, 2016

[†] Dept. of Electronics Eng., Graduate School, Kyung-sung University (E-mail: sanghee2757@ks.ac.kr)

^{**} Dept. of Electronics Eng., Graduate School, Kyung-sung University (E-mail: jsong@ks.ac.kr)

^{***} Dept. of Electronics Eng., Graduate School, Kyung-sung University

^{****} Dept. of Electronics Eng., Graduate School, Kyung-sung University (E-mail: bwoon@ks.ac.kr)

※ This work was supported by the industrial strategic technology development program, 10045260, Development of Context Awareness Monitoring and Search System Based on High Definition Multi-video funded By the Ministry of Science, ICT and Future Planning (Korea) and Busan City, BB21 program

정하기 위하여 최적 위너 필터(optimal Wiener filter)를 적용한다. 카메라 영상에서 보행자를 배경으로부터 분리하기 가우시안 혼합 모델(GMM, Gaussian Mixture Model) 이용한다. 가우시안 혼합 모델은 시간의 경과에 따라 변화되는 배경 정보를 반영하는데 효과적이다. 가우시안 혼합 모델을 적용한 배경 추정을 통하여 보행자 후보를 검출하고[7-14], 후보 영역의 높이 화소에 대하여 위치별 키에 대한 화소 최적비를 적용하여 보행자의 키를 추정한다. CCTV 카메라 영상에 대하여 제안하는 키 추정 방법이 효과적임을 실험을 통하여 보인다.

2. 키 추정 방법

제안하는 키 추정 방법의 흐름도는 Fig. 1과 같다. 키 추정 과정을 수행하기 전에 보행자 검출 과정을 수행한다. 보행자 검출 과정은 입력영상으로부터 배경과 움직이는 객체를 분리하고 분리된 후보 객체들 중에서 보행자를 판별한다[7]. Fig. 1의 좌하단 ‘Optimal Estimator Design’ 과정에서는 키가 알려진 보행자로부터 데이터를 획득하고 최적의 모델을 추정하게 되며, 우하단의 ‘Running Mode’에서는 ‘Optimal Estimator Design’ 과정에서 구한 추정값을 이용하여 입력 영상으로부터 키가 알려지지 않은 보행자의

키를 계산하는 과정이다. 실제의 경우 CCTV 카메라가 신규로 설치되거나 설치 각도, 확대와 축소 조건 등의 설치 환경이 변화된 경우 한번만 ‘Optimal Estimator Design’ 과정을 거치면 그 이후로는 ‘Running Mode’로 동작시켜서 자동적으로 보행자의 키를 추정한다.

2.1 보행자 검출

보행자를 검출하기 위해서 CCTV 카메라를 통하여 입력된 영상에서 가우시안 혼합 모델을 이용하여 전경을 분리하게 된다. 배경 추정은 최근 프레임들의 화소 값을 버퍼에 저장하고 일정 프레임마다 버퍼에 저장된 값들을 이용하여 배경 모델을 갱신한다. 그리고 매 프레임마다 화소들을 아래의 수식과 같이 배경 모델과 비교하여 배경과 전경을 분리한다.

$$Distance(I_i(x,y), B_t(x,y)) = \max(|I_i(x,y) \cdot c - B_t(x,y) \cdot c|) \tag{1}$$

I 는 입력 이미지이고 B 는 배경 모델 그리고 c 는 R, G, B값을 의미한다. 위의 값이 일정 임계값보다 작으면 해당 화소를 전경으로 분리한다. 이 방법은 배경 모델의 안정성이 높다는 장점이 있지만, 최근 화소들을 저장할 버퍼를 필요로 한다는 단점이 있다.

분리된 전경 이미지로부터 모폴로지 필터(morphology filter)를 적용하여 잡음을 제거하고 미세하게 끊어진 전경들을 이어주게 된다. 모폴로지 필터링을 통하여 객체로 추정되는 몇 개의 영역만 남게 되고 이들 후보 객체들에 대해 레이블링(labeling)을 수행하여 객체의 특징 데이터를 수집한다. Fig. 2는 입력영상에서 가우시안 혼합모델과 모폴로지 필터링 그리고 레이블링 수행 결과이다. 본 논문에서는 보행자 판별을 위한 임계값을 아래와 같이 정의하였다.

$$180 < R < 400 \tag{2}$$

$$13 < W < 100 \tag{3}$$

여기서 W 와 H 를 후보 객체의 수평 및 수직 화소 수이며 $R = (100 * H) / W$ 로 정의되는 후보 객체의 중횡비(ratio of width and height)이다. 위의 두 조건을 만족하면 최종적으로 보행자로 판별된다. Fig. 3은 보행자를 검출한 결과이다.

2.2 키 추정 방법

본 논문에서는 키 추정을 위하여 비율 k 를 식 (4)

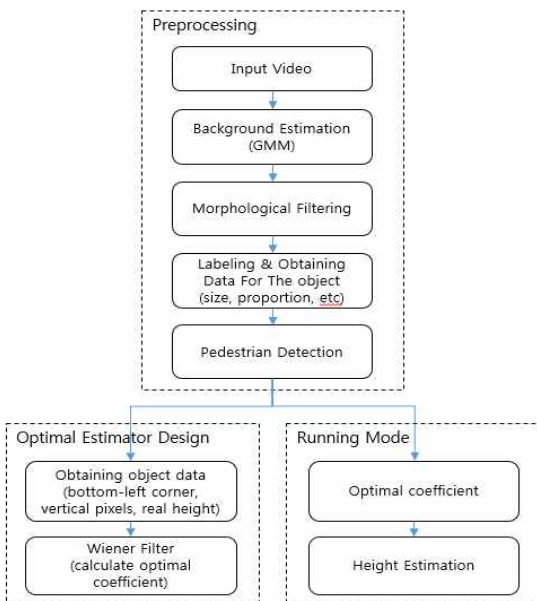


Fig. 1. Flow chart of the proposed height estimation.

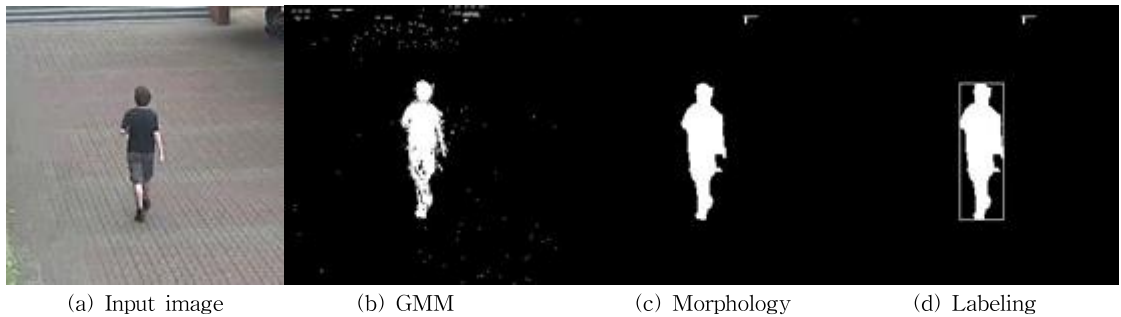


Fig. 2. Pre-processing of pedestrian detection.



Fig. 3. Results of pedestrian detection.

와 같이 정의하고 이를 활용하여 키를 추정한다. 식 (4)에서 $height_{pixel}$ 은 검출된 보행자의 수직 화소 수이며 $height_{cm}$ 는 보행자의 실제 키이다. 이 비율을 보행자의 좌하단 좌표와 비교하면 Fig. 4와 같은 결과를 얻을 수 있다.

$$k = \frac{height_{pixel}}{height_{cm}} * 100 \tag{4}$$

Fig. 4에서 k 값이 x, y 좌표와 비례하는 것을 볼 수 있다. 객체의 좌표에 따른 최적의 k 를 계산하면 정확한 키 추정을 수행할 수 있다. 본 논문에서는 k 와 x, y 좌표의 관계를 1차 선형함수로 모델링하고 최적 모델을 구하기 위하여 워너 추정 방법을 이용하였으며, 평균제곱오차(MSE, Mean Square Error)가 최소가 되는 최적 모델을 구하고, 이를 이용하여 키를 추정하게 된다.

입력 영상에 따라 사전에 정보를 수집하기 위하여 키를 알고 있는 보행자가 영상 내의 다양한 위치에 이동하면서 매 프레임마다 보행자의 좌하단 좌표, 수직 화소 수 그리고 실제 키(cm)를 같이 저장하게 된다. 이 과정이 끝난 후 저장된 정보들을 다음의 워너 추정 방법에 적용하게 되며, 이를 통해 최적의 모델을 도출해내는 과정은 다음과 같다. $\mathbf{x} = [x \ y \ 1]$ 는 입력 벡터로 정의하였고, 이 때 (x, y) 는 보행자의 좌하단 x, y 좌표이다. 그리고 $\mathbf{h} = [h_x \ h_y \ h_c]$ 는 반응벡터이다. 평균제곱오차는 다음과 같이 정의하였다.

$$E\{\epsilon^2\} = \{[k - \hat{k}]^2\} \tag{5}$$

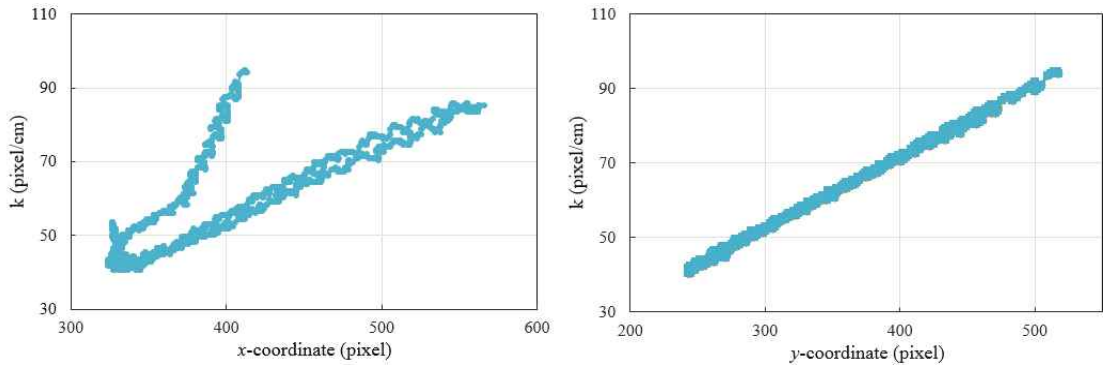


Fig. 4. The x, y coordinate of bottom-left corner and k for each position.

여기서 k 는 실제값이고 \hat{k} 는 측정값이다. 이 평균 제곱오차를 최소화하기 위해, $\frac{\partial E\{e^2\}}{\partial k} = 0$ 을 만족하는 \mathbf{h} 를 구해야한다. 그러므로 \mathbf{h}^* 로 정의하는 최적의 \mathbf{h} 는 위너-호프 식(Winer-Hoff equation)에 의해 다음과 같이 정의한다.

$$\mathbf{h}^* = r_{xx}^{-1} * r_{kx} \tag{6}$$

$r_{xx} = E[\mathbf{x}^T \cdot \mathbf{x}]$ 는 자기상관행렬(auto-correlation matrix)이고, $r_{kx} = E[k \cdot \mathbf{x}]$ 는 상호상관행렬(cross-correlation matrix)이다. 본 논문에서는 보행자 키 추정을 위하여 \mathbf{h}^* 를 활용한다. 이를 이용하여 \hat{k} 를 다음과 같이 정의할 수 있다.

$$\hat{k} = \mathbf{h}^* * \mathbf{x}^T \tag{7}$$

위의 \hat{k} 값과 식 (8)을 이용하여 다음과 같이 보행자의 키를 계산할 수 있다.

$$height_{cm} = \frac{height_{pixel}}{\hat{k}} * 100 \tag{8}$$

3. 실험 결과 및 고찰

본 논문에서는 Microsoft Visual Studio 2013과 영상처리 공개 라이브러린 OpenCV 2.4.9를 활용하여 초당 20 프레임을 처리한다. CCTV 카메라 영상 내에서 각각 다른 위치에서의 키를 측정하였다. 실험은 근거리와 원거리로 구분하여 진행하였으며, 카메라로부터 10 m 거리는 근거리, 10 m 보다 먼 경우를 원거리로 정의하였다. 그리고 기존 키 추정 방법[7]과 결과를 비교하였다.

Fig. 5는 4명의 보행자 검출 및 키 추정 결과 이미지를 보여준다. Fig. 6은 제안하는 방법을 이용하여 183cm의 키를 가지는 보행자 한 명에 대한 매 프레임에서의 키 추정 오차 결과를 나타낸 것으로써, 왼쪽 그래프는 근거리에서의 결과, 오른쪽 그래프는 원거리에서의 결과이다. Table. 1은 근거리에서의 기존 키 추정 방법과 본 논문에서 제안하는 키 추정 방법의 결과로써 4명의 보행자에 대한 결과를 나타낸다. 기존 방법의 결과에서는 평균 오차가 각각 8.66cm, 7.76cm, 8.54cm, 9.36cm이고 평균 오차율은 5.38%, 4.61%, 4.73%, 5.41%이다. 제안하는 방법의 평균 오



Fig. 5. The results of the applied algorithm. (a) object 1, (b) object 2, (c) object 3 and (d) object 4.

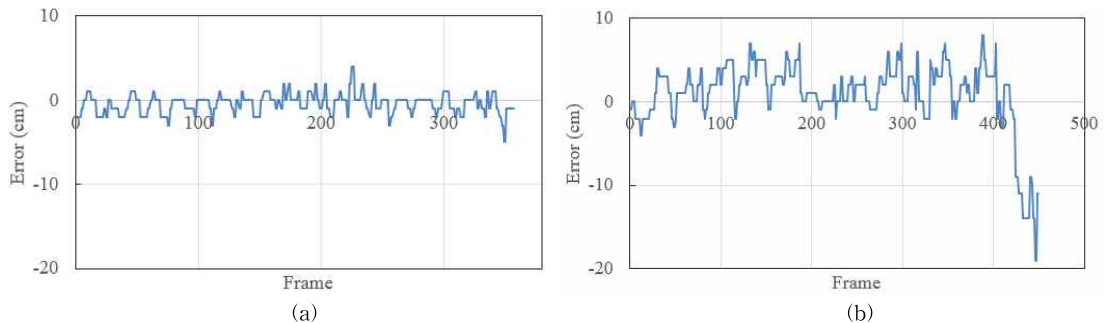


Fig. 6. The height estimation results of object 1. (a) near-distance, (b) long-distance.

Table 1. The results of near-distance height estimation

Object	Real Height (cm)	Conventional method [7]		Proposed height estimation	
		Average Error (cm)	Average Error Rate (%)	Average Error (cm)	Average Error Rate (%)
1	183	8.66	5.38	0.88	0.48
2	161	7.76	4.61	2.65	1.65
3	168	8.54	4.73	1.14	0.68
4	173	9.36	5.41	1.04	0.59

Table 2. The results of long-distance height estimation

Object	Real Height (cm)	Conventional [7]		Proposed height estimation	
		Average Error (cm)	Average Error Rate (%)	Average Error (cm)	Average Error Rate (%)
1	183	11.09	6.89	2.91	1.59
2	161	13.00	7.73	4.54	2.82
3	168	15.99	8.74	2.95	1.76
4	173	10.57	6.11	2.93	1.68

차는 0.88cm, 2.65cm, 1.14cm, 1.04cm이고 평균 오차율은 0.48%, 1.65%, 0.68%, 0.59%이다. Table. 2는 원거리에서의 결과로써 기존 방법의 결과에서는 평균 오차가 11.09cm, 13cm, 15.99cm, 10.57cm이고, 평균 오차율은 6.89%, 7.73%, 8.74%, 6.11%이다. 제안하는 방법의 평균 오차는 각각 2.91cm, 4.54cm, 2.95cm, 2.93cm이고 평균 오차율은 1.59%, 2.82%, 1.76%, 1.68%이다. 최종적으로 근거리에서보다 원거리에서의 오차가 큰 것을 확인할 수 있었고, 기존의 키 추정 방법은 91% 이상의 정확도를 보여주는 반면에, 제안하는 보행자 키 추정 방법은 97% 이상의 정확도를 보여주었다. 이를 통해 제안하는 방법이 기존 방법보다 효율적인 것을 확인하였다.

4. 결 론

본 논문에서는 CCTV 영상을 기반으로 보행자의 키를 추정하는 방법을 제안하였다. 객체를 배경으로부터 정확히 분리하기 위하여 가우시안 혼합 모델을 사용하였고, 보행자 객체로부터 수집한 정보들을 워너 추정 방법에 적용하여 최적의 모델을 구하였다. 영상 내에서 보행자의 위치에 대응하는 보정 계수를 구하고 이를 이용하여 보행자의 위치에 제약 없이 키를 추정할 수 있었다. 실험을 통하여 기존 방법보

다 제안하는 키 추정 방법이 효과적으로 키를 추정할 수 있음을 보였다.

REFERENCE

- [1] H. Moon and S. Pan, "The Human Identification Method in Video Surveillance System," *Journal of Korea Institute of Information Technology*, Vol. 8, No. 5, pp. 199-206, 2010.
- [2] H. Moon and S. Pan, "The Analysis of De-identification for Privacy Protection in Intelligent Video Surveillance System," *Journal of The Korea Institute of Information Technology*, Vol. 9, No. 7, pp. 189-200, 2011.
- [3] Y. Choi, K. Kim, Y. Nam, and W. Cho, "Retrieval of Identical Clothing Images Based on Local Color Histograms," *Proceeding of the International Conference on Convergence and Hybrid Information Technology*, Vol. 1, pp. 818-823, 2008.
- [4] A. Bovyryn and K. Rodyushkin, "Human Height Prediction and Roads Estimation for Advanced Video Surveillance Systems," *Proceeding of IEEE Conference on Advanced*

- Video and Signal Based Surveillance*, Vol. 15-16, pp. 219-223, 2005.
- [5] E. Jeaes, I. Kispal, and Z. Hornak, "Measuring Human Height Using Calibrated Cameras," *Proceeding of IEEE Conference on Human System Interactions*, Vol. 25-27, pp. 755-760, 2008.
- [6] S. Jung, J. Kim, D. Lee, S. Lee, and G. Kim, "Intelligent Video Event Detection System Used by Image Object Identification Technique," *Journal of Korea Multimedia Society*, Vol. 13, No. 2, pp. 171-178, 2010.
- [7] S. Kim, J. Song, B. Yoon, and J. Park, "Height Estimation of Pedestrian Based on Image," *Journal of The Korea Institute of Electronic Communication Sciences*, Vol. 9, No. 9, pp. 1035-1042, 2014.
- [8] L. Bazzani, D. Bloisi, and V. Murino, "A Comparison of Multi Hypothesis Kalman Filter and Particle Filter for Multi-target Tracking" *Proceeding of Eleventh IEEE International Workshop on Performance Evaluation of Tracking and Surveillance*, pp. 47-54, 2009.
- [9] H. Kim, G. Lee, J. Park, and Y. Yu, "Vehicle Detection in Tunnel Using Gaussian Mixture Model and Mathematical Morphological Processing," *Journal of The Korea Institute of Electronic Communication Sciences*, Vol. 7, No. 5, pp. 967-974, 2012.
- [10] M. Kim, C. Oh, D. Aurrahman, Y. Ahn, and C. Lee, "The Virtual Screen Using Skin Tone and GMM Foreground Segmentation," *Proceeding of the Spring Conference of the Korea Information Processing Society*, Vol. 15 No. 1, pp. 179-181, 2008.
- [11] D. Koller, J. Weber, T. Huang, J. Malik, G. Ogasawara, B. Rao, et al., "Towards Robust Automatic Traffic Scene Analysis in Real-time," *Proceeding of International Conference on International Association of Pattern Recognition*, Vol. 1, pp. 126-131, 1994.
- [12] C. Stauffer and W.E.L. Grimson, "Adaptive Background Mixture Models for Real-time Tracking," *Proceeding of IEEE Conference on Computer Society Computer Vision and Pattern Recognition*, Vol. 2, pp. 246-252, 1999.
- [13] A. Elgammal, D. Harwood, and L.S. Davis, "Non-parametric Model for Background Subtraction," *Proceeding of European Conference on Computer Vision*, Vol. 1843, pp. 751-767, 2000.
- [14] R. Cucchiara, C. Grana, M. Piccardi, and A. Prati, "Detecting Objects, Shadows and Ghosts in Video Streams by Exploiting Color and Motion Information," *Proceeding of International Conference on Image Analysis and Processing*, pp. 360-365, 2001.



전 상 희

2015년 동의대학교 멀티미디어공학과 졸업 (학사)
2016년 3월~현재 경성대학교 대학원 전자공학과 석사과정
관심분야: 영상처리 및 이해, GPU 병렬처리



박 장 식

1992년 부산대학교 전자공학과 졸업 (공학사)
1994년 부산대학교 대학원 전자공학과 졸업 (공학석사)
1999년 부산대학교 대학원 전자공학과 졸업(공학박사)

1997년 3월~2011년 2월 동의과학대학 전자과 교수
2011년 3월~현재 경성대학교 전자공학과 교수
관심분야: 영상처리 및 이해, 신호처리, 음성 및 음향신호처리, 임베디드 시스템, GPU 병렬처리



송 종 관

1989년 부산대학교 전자공학과 졸업 (공학사)
1991년 KAIST 전기 및 전자공학과 졸업 (공학석사)
1995년 8월 KAIST 전기 및 전자공학과 졸업 (공학박사)

1995년 9월~1997년2월 SK텔레콤 중앙연구소 선임 연구원
1997년 3월~현재 경성대학교 전자공학과 교수
관심분야: 영상처리, 디지털신호처리, 디지털신호처리 응용



윤 병 우

1987년 부산대학교 전자공학과 졸업 (공학사)
1989년 부산대학교 대학원 전자공학과 졸업 (공학석사)
1992년 부산대학교 대학원 전자공학과 졸업 (공학박사)

1993년~1995년 한국전자통신연구원 선임 연구원
1995년~현재 경성대학교 전자공학과 교수
관심분야: 신호처리, 영상처리, VLSI설계