

DBSCAN 방법을 이용한 패널 글라스 영상의 에지 직선 추정 방법

이승목*, 김영형**, 임재권***

A Method of Edge Line Estimation for Panel Glass Images using DBSCAN Algorithm

Seung-Mok Lee*, Young-Hyung Kim** and Jae-Kwon Eem***

이 연구는 금오공과대학교 학술연구비로 지원되었음(과제번호: 202001810001).

요약

본 논문은 왜곡이 심한 패널 글라스 영상에 대해 DBSCAN 알고리즘을 이용하여 에지 직선을 추정하는 방법을 제안하였다. 패널 글라스를 촬영하면 패널 글라스 에지의 단차, 에지 오염 및 이물 침범으로 인한 에지의 왜곡이 발생된다. 이러한 왜곡을 배제하기 위하여, 에지 검출 후에 검출된 에지 위치와 에지 위치의 변화를 데이터로 하여 이의 히스토그램을 얻고 빈도가 낮은 데이터를 제거한다. 그다음 DBSCAN 알고리즘을 통하여 클러스터링 후, 최대 크기의 클러스터에 대하여 직선 적합을 적용하여 에지 직선 추정을 하였다. 제안된 알고리즘의 성능을 평가하기 위하여 실제 설비에서 나타나는 다양한 영상을 사용하였고 일반적으로 많이 사용되는 아웃라이어를 배제하여 최적화하는 기법인 RANSAC 방법과 비교하여 제안된 알고리즘의 성능을 평가하였다.

Abstract

This paper proposes a method of the edge line estimation using DBSCAN algorithm for panel glass images with heavy distortions. When a panel glass image is taken, edge distortions are generated due to step of the panel glass edge, edge contamination and foreign matter invasion. In order to exclude such distortions, after edge detection, a histogram is obtained using the detected edge position and the edge position variation as data set, and low frequency data is removed. After this process, the edge line was estimated by clustering through the DBSCAN algorithm and applying line-fitting to the cluster of the maximum size. For evaluating performance of the proposed algorithm, various images appearing in actual facilities were used, and the performance was evaluated by comparing it with the RANSAC method, a technique that optimizes by excluding outliers that are commonly used.

Keywords

edge detection, line detection, RANSAC, DBSCAN, line fitting

* (주)신세계엔지니어링

- ORCID: <http://orcid.org/0000-0003-3703-5998>

** 금오공과대학교 IT 융합학과

- ORCID: <http://orcid.org/0000-0001-5708-854X>

*** 금오공과대학교 전자공학과(교신저자)

- ORCID: <http://orcid.org/0000-0001-8506-9434>

• Received: Feb. 16, 2021, Revised: May 10, 2021, Accepted: May 13, 2021

• Corresponding Author: Jae-Kwon Eem

Dept. of Electronics Engineering, Kumoh National Institute of Technology, 61 Daehak-ro(yangho-dong), Gumi, Gyeongbuk, 39177, Korea,

Tel.: +82-54-478-7428, Email: jkeem@kumoh.ac.kr

I. 서 론

최근에는 여러 산업 분야에서 수집한 데이터들을 분석하여 이를 적극적으로 활용하는 추세이다. 클러스터링은 데이터를 동종 그룹으로 나누는 중요한 분석 작업으로 클러스터 내의 유사성을 최대화하여 데이터 집합을 그룹화하는 방법이다. 이러한 데이터 그룹에 잡음이 있는 경우, 데이터의 잡음을 제거하기 위하여 일반적으로 많이 활용되는 DBSCAN (Density Based Spatial Clustering of Applications with Noise) 알고리즘은 밀도를 기반으로 그룹화하는 방법이다[1].

DBSCAN 알고리즘은 클러스터 내의 밀도 변화를 고려하여 클러스터를 확장하는 방법이나 데이터의 확률 분포 모델을 사용하여 밀도 수준을 이용하여 클러스터링하는 방법, 각 클러스터의 병렬 탐색을 수행하여 클러스터링하는 방법 등이 연구되어왔다 [2]-[4]. 또한 군집된 잡음을 가진 데이터에서 잡음을 필터링하는 방법이나 DBSCAN 알고리즘의 고속화에 대한 연구도 이루어졌다[5][6]. 나아가서 계층적 클러스터링이나 그리드 기반 클러스터링 등의 다양한 클러스터링 연구도 활발하게 이루어지고 있다[7][8].

본 논문에서는 영상 위치 보정시스템이나 영상 검사 시스템에 활용되는 에지 직선 추정에 관한 연구이고, 패널 글라스의 에지를 검출하는 경우, 영상에 나타나는 왜곡을 효과적으로 배제하여 에지 직선을 추정하는 방법을 제안하였다. 검출된 에지 위치와 에지 위치의 변화를 데이터로 하여 히스토그램을 얻고 낮은 빈도의 데이터를 제거한다. 그다음 DBSCAN을 통하여 클러스터링한 후, 최대 크기의 클러스터의 데이터로 에지 직선을 추정하는 방법이다. 2장에서는 DBSCAN 방법을 통한 에지 직선 측정에 대해 서술하고, 3장에서는 제안된 알고리즘에 대하여 기술하고, 일반적으로 데이터의 아웃라이어 (Outlier) 배제를 통한 최적화하는 RANSAC (RANdom SAmple Consensus) 알고리즘과 비교하였다. 4장에서는 실제 생산 설비에서 촬영한 영상에 적용하여 에지 직선을 추정함으로써 RANSAC 방법의 결과와 제안된 알고리즘의 결과를 비교하여 알

고리즘의 성능을 확인하였다. 이어서 5장에서는 제안된 알고리즘의 대한 성능을 평가하고 결론을 내렸다.

II. DBSCAN 알고리즘

DBSCAN 알고리즘은 밀도 기반 클러스터링 비모수(Non-Parametric) 알고리즘으로 일정 개수 이상의 모여있는 점들을 모아 그룹을 만들고, 이 그룹의 점들과 가까운 점들을 이들 그룹에 포함시키는 방법으로 반복하여 가까운 점이 없을 때까지 수행한다. 그리고 클러스터링 알고리즘 중에 많이 사용되는 일반적인 알고리즘이다. DBSCAN 알고리즘을 그림 1에 나타내었다.

그림 1에 나타낸 알고리즘을 살펴보면 점들 간의 가까운 정도를 평가하기 위한 T값이 사용되고 그룹이 될 만한 정도의 점들이 모였는지를 확인하기 위한 M값이 사용되어진다. 우선 알고리즘은 데이터에서 임의의 점 P를 가지고 와서 P와 가까운 점들을 찾아서 다른 클러스터로 레이블링되지 않았는지 확인한다. 그리고 클러스터의 생성을 위한 최소 점들 이상을 가지게 되면 새 클러스터를 위한 레이블 값을 정하고, 점 P에 레이블링하며 새 클러스터 그룹을 만든다.

새 클러스터 그룹의 점 Q에 대하여 클러스터의 레이블 값으로 레이블링하고, 점 Q와 가까운 점들을 찾아 일정 수 이상이 되면 클러스터 그룹에 추가한다. 이 클러스터 그룹의 모든 점들에 대하여 이 과정을 수행함으로써 클러스터에 속한 모든 점들에 대해 레이블링을 완료하게 된다. 그리고 결정되지 않은 데이터의 점들로부터 다시 새로운 클러스터 그룹을 찾아 레이블링하는 과정을 반복함으로써 데이터의 모든 점에 대하여 클러스터링을 수행하게 된다.

III. 제안된 직선 추정 알고리즘

본 논문에서 제안한 알고리즘은 DBSCAN 방법을 활용한 직선 추정 방법이며, 검출된 에지의 위치와 에지 위치의 변화를 데이터 세트로 구성한다. 오염

이나 이물이 가지는 에지 데이터의 빈도가 낮기 때문에 이를 구분하여 제거함으로써, 남은 데이터에 대한 DBSCAN을 통하여 보다 세분화된 클러스터링을 수행한다. 이어서 가장 큰 클러스터의 데이터를 사용하여 에지 직선을 추정하였다.

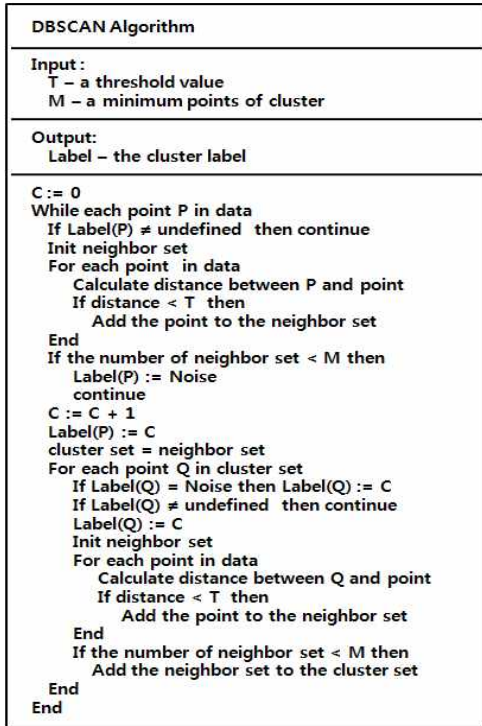


그림 1. DBSCAN 알고리즘
Fig. 1. DBSCAN algorithm

그림 2는 제안된 알고리즘을 나타낸 것으로 우선 영상 이진화를 위한 문턱치는 Otsu 방법을 사용하였고, 에지 위치 검출은 이진화된 영상의 이진 값의 처음 변화 지점을 에지 위치로 검출하였다. Otsu 방법은 영상의 히스토그램에서 2개의 클래스로 나누는 경우, 클래스 간 분산을 최대로 하는 값을 문턱치로 찾는 방법이다[9].

에지 위치를 검출한 후, 에지 위치의 변화는 특정 위치 n 에서 특정 거리 사이의 변화를 얻기 위하여 식 (1)로 계산하였다.

$$v(n) = x(n+k) - x(n-k) \tag{1}$$

여기서 $v(n)$ 은 n 위치에서의 에지 위치의 변화이

고, $x(n)$ 의 n 위치에서의 에지 위치이다. k 는 거리를 의미한다.

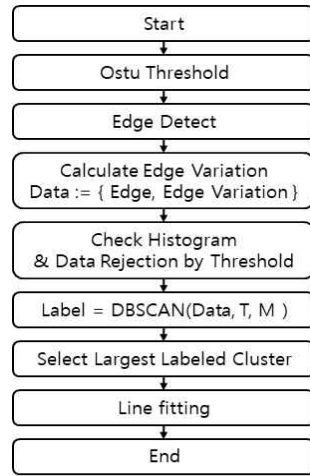


그림 2. 제안된 알고리즘
Fig. 2. Proposed algorithm

에지 위치와 에지 위치의 변화로 데이터를 구성하였으며, DBSCAN 알고리즘을 수행하기 전에 왜곡과 같이 빈도가 낮은 정보를 제거하기 위하여 데이터에 대한 히스토그램을 얻어 낮은 빈도의 데이터를 제거하였다. 그리고 DBSCAN 알고리즘을 사용하여 데이터를 클러스터링하고 가장 많은 수의 점들로 구성된 클러스터의 데이터로 직선 적합을 수행하여 에지 직선을 추정하였다.

제안된 알고리즘에서 에지 위치(x)와 에지 위치의 변화(v)를 사용하여 얻은 히스토그램은 그림 3(a)에 나타내었다. 그림 3(b)는 데이터를 제거하기 전의 데이터를 $x-v$ 평면에 표시하였고, 그림 3(c)는 문턱치를 사용하여 데이터 제거를 수행한 결과를 나타내었다.

일반적으로 잡음으로 인해 아웃라이어 많이 포함된 데이터에 대하여 최적화하는 경우, 아웃라이어를 배제하기 위하여 RANSAC 알고리즘을 많이 사용한다[10][11]. RANSAC 알고리즘은 N 번 반복하여 랜덤 샘플링을 수행하고 최적화하는 방법으로 아웃라이어가 포함되지 않은 랜덤 샘플 데이터를 얻을 확률 p 는 식 (2)와 같다.

$$p = 1 - (1 - d^N) \tag{2}$$

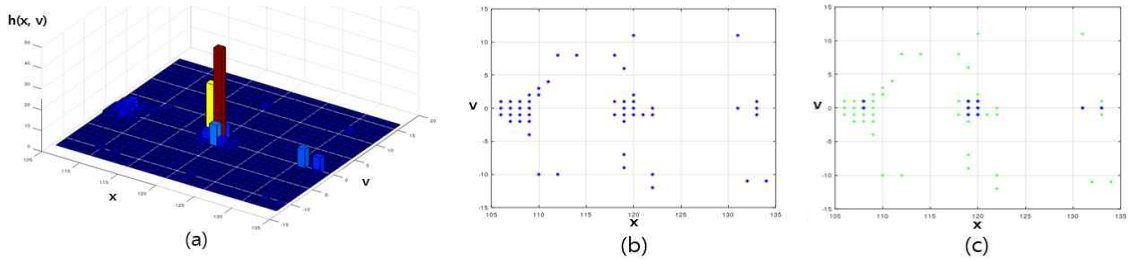


그림 3. 히스토그램과 데이터 제거
Fig. 3. Histogram and data-rejection

여기서 q 는 랜덤 선택한 데이터가 인라이어일 경우의 확률이고 l 은 랜덤 샘플링될 데이터 수이다.

그림 4은 일반적으로 많이 사용되는 잡음에 강한 RANSAC 방법과 제안된 알고리즘의 처리과정의 결과들을 나타낸 것으로 (a)는 왜곡이 포함된 에지 영상을 나타낸 것이다. 그리고 (b)는 검출된 에지를 표시하였고 (c)는 제안된 알고리즘에서 DBSCAN 후에 가장 큰 클러스터에 속한 에지를 표시한 것이다. (d)와 (e)는 RANSAC 방법을 사용한 경우, 선 추정이 아웃라이어를 배제하고 이루어지는 경우와 아웃라이어의 영향을 받아서 이루어지는 경우를 함께 표시하였다. (f)는 제안된 알고리즘에서 직선 적합으로 직선 추정한 결과를 나타낸 것이다.

IV. 실험 및 고찰

제안된 선 추정 알고리즘은 실제 설비에서 획득한 왜곡을 가진 여러 영상에 대하여 적용하였으며, 성능 평가를 위하여 일반적으로 많이 사용되는 잡음에 강한 RANSAC 알고리즘을 함께 적용하여 제안된 알고리즘의 성능을 평가하였다. 영상의 FOV는 $8.4[\text{mm}] \times 7.0[\text{mm}]$ 에 대해 $5\text{M}[\text{pixel}]$ 해상도 카메라로 획득하였다. RANSAC 방법을 적용하기 위하여 반복 횟수 N 은 1000회, 샘플 수는 20개, 인라이어 판단을 위한 문턱치는 위치 데이터의 표준편차로 설정하였다. 그리고 제안된 알고리즘은 히스토그램 문턱치는 5이고, DBSCAN 알고리즘의 최소 거리 문턱치와 최소 점수는 각각 5로 설정하였다.

샘플 영상은 그림 5에 나타내었고, 선 추정을 사용하기 위하여 영역을 설정하는 경우 나타날 수 있

는 다양한 왜곡을 가진 영상들로 구성하였다. (a)는 돌출부를 가진 영상, (b)는 오염된 영상, 그리고 (c)는 이물이 간섭한 영상이다. 이에 대한 에지 검출 결과, RANSAC 적용 결과 및 제안된 알고리즘의 결과를 그림 6에 나타내었다.

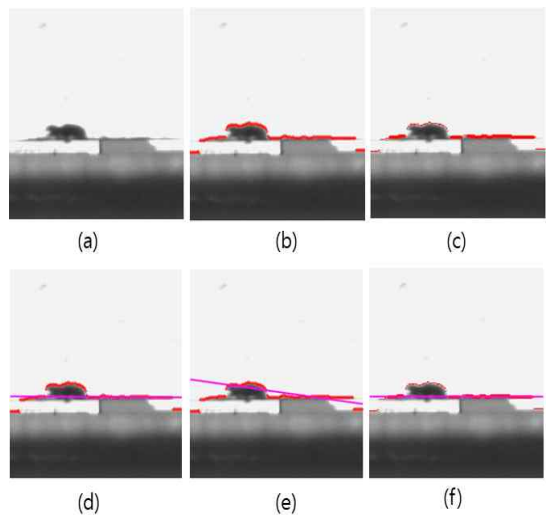


그림 4. RANSAC 방법과 제안된 알고리즘의 결과들
Fig. 4. The results of RANSAC method and proposed algorithm.

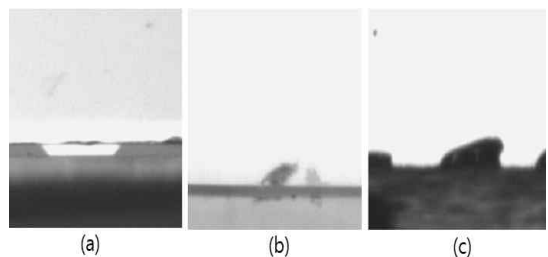


그림 5. 샘플 영상들
Fig. 5. Sample images

그림 6은 샘플 영상들에 대하여 에지 검출 결과, RANSAC 알고리즘으로 얻은 추정 선과 제안된 알고리즘으로 수행한 추정 선을 순서대로 나타내었다. 그림 6의 (a)와 (b)를 살펴보면 RANSAC 알고리즘의 경우, 아웃라이어를 배제하고 에지 직선을 추정하였지만 아웃라이어의 영향이 있음을 확인할 수 있다. 반면에 제안된 알고리즘의 경우, DBSCAN에 의하여 왜곡 영역이 제거되어 왜곡의 영향을 배제하고 에지 직선 추정이 이루어짐을 알 수 있다. 그림 6(c)는 RANSAC 알고리즘으로 에지 직선 추정이 정확히 이루어지지 않음을 볼 수 있으나 제안된 알고리즘은 정확하게 에지 직선을 추정하였다.

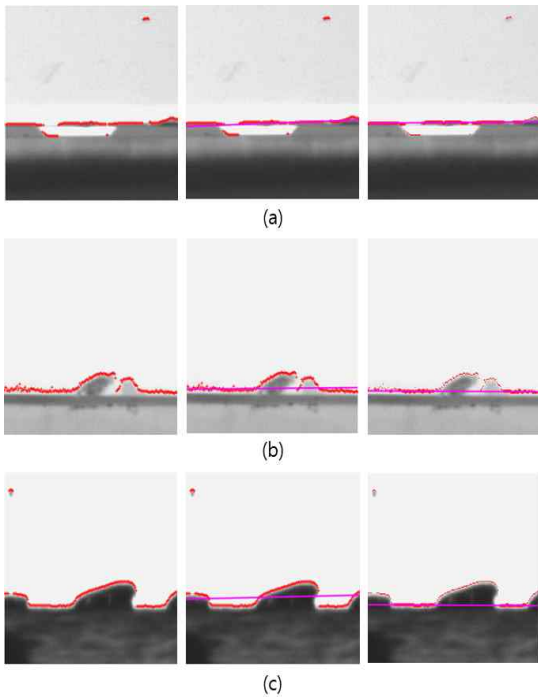


그림 6. 에지 검출, RANSAC 방법 및 제안된 알고리즘의 결과물들

Fig. 6. Results of edge detection, RANSAC method and proposed algorithm

표 1. 알고리즘에 따른 샘플 영상의 평균제곱 오차
Table 1. MSE of algorithms on sample images

method	Sample Image		
	(a)	(b)	(c)
RANSAC	1.268	0.982	75.688
Proposed	0.189	0.302	0.168

샘플 영상들에 적용한 제안된 알고리즘의 성능 평가를 위하여, 에지 직선을 추정하기 위한 데이터와 추정된 선과의 평균제곱 오차를 구하였고, 이를 표 1에 나타내었다. RANSAC 알고리즘의 경우, 10회 반복 측정한 평균제곱 오차를 평균하여 나타내었다. 샘플 영상 (a)와 (b)의 경우, RANSAC 알고리즘의 평균제곱오차가 1.268[pixel], 0.982[pixel]로 에지 직선 추정이 비교적 잘되는 것을 볼 수 있으나 샘플 영상 (c)의 경우는 상당한 오차를 가지는 것을 볼 수 있다. 그러나 제안된 알고리즘의 경우, 샘플 영상 (a)~(c)에 대한 평균제곱 오차는 각각 0.189[pixel], 0.302 [pixel] 및 0.168[pixel]로 RANSAC에 비하여 우수한 성능을 나타내었다.

V. 결 론

본 논문은 왜곡이 심한 영상의 에지 직선을 추정하기 위한 방법을 제안한 것으로 실제 설비 영상을 통해 알고리즘을 실험하였다. 알고리즘의 성능 평가를 위하여, 아웃라이어를 가진 데이터의 추정에 많이 사용되는 RANSAC 방법과 비교함으로써 제안된 알고리즘의 성능을 확인하였다.

제안된 알고리즘이 실제 나타나는 여러 가지 왜곡에 대한 직선 추정이 안정적인 성능을 나타낸 것을 볼 수 있었다. 특히, 샘플 영상 (c)의 심한 이물 간섭에 대해서도 잘 대응되는 것을 확인하였다. 제안된 알고리즘의 평균제곱 오차를 샘플영상 (a)~(c)에 대하여 실제 치수로 변환하는 경우, 각각 0.66[μm], 1.06[μm] 및 0.59[μm] 정도로 나타났으며 심한 왜곡에 대하여도 강인한 성능을 보였다.

References

- [1] M. Easter, H. P. Kriegel, J. Sander, and X. Xu, "A Density-Based Algorithm for Discovering Clusters in large Spatial Databases with Noise", In Proceedings of KDD, pp. 226-231, 1996.
- [2] A. Ram, S. Jalal, A. S. Jalal, and M. Kumar, "A Density based Algorithm for Discovering Density Varied Clusters in Large Spatial Databases", Int'l

Journal of Computer Applications, Vol. 3, No. 6, Jun. 2010.

[3] X. Li, P. Zhang, and G. Zhu, "DBSCAN Clustering Algorithms for Non-Uniform Density Data and Its Application in Urban Rail Passenger Aggregation Distribution", *Energies*, Vol. 12, No. 19, 2019.

[4] D. Arlia, and M. Coppola, "Experiments in Parallel Clustering with DBSCAN", *Proceedings of the 7th Int'l Euro-Par Conference Manchester on Parallel Processing*, pp. 326-331, Aug. 2001.

[5] S. Kim, and S. Lee, "Clustered Noise Filtering Based Trajectory Refinement Method", *Journal of KIIT*, Vol. 18, No. 3, pp. 11-20, Mar. 2020.

[6] A. Gunawan, "A faster algorithm for DBSCAN", *Computer Science*, 2013.

[7] B. Perret, J. Cousty, S. J. Ferzoli Guimarães, Y. Kenmochi, and L. Najman, "Removing non-significant regions in hierarchical clustering and segmentation", *Pattern Recognition Letters*, Vol. 128, pp. 433-439, Dec. 2019.

[8] T. Boonchoo, X. Ao, Y. Liu, W. Zhao, F. Zhuang, and Q. He, "Grid-based DBSCAN: indexing and inference", *Pattern Recognition*, Vol. 90, pp. 271-284, Jun. 2019.

[9] M. Sezgin, and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation", *Journal of Electronic Imaging*, Vol. 13, No. 1, Jan. 2004.

[10] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for cartography", *Communications of the ACM*, Vol. 24, No. 6, pp. 381-395, Jun. 1981.

[11] K. G. Derpanis, "Overview of the RANSAC Algorithm", *Image Rochester, N.Y.*, Vol. 4, No. 1, pp. 2-3, 2010.

저자소개

이 승 목 (Seung-Mok Lee)



1996년 2월 : 금오공과대학교
전자공학과(공학사)
1998년 2월 : 금오공과대학교
대학원 전자공학과(공학석사)
2016년 2월 : 금오공과대학교
대학원 전자공학과 (공학박사)
2017년 8월 ~ 현재 :

신세계엔지니어링 연구소장

관심분야 : 컴퓨터비전, 영상처리, 임베디드, 자동차

김 영 형 (Young-Hyung Kim)



1992년 2월 : 금오공과대학교
전자공학과(공학사)
1996년 2월 : 금오공과대학교
대학원 산업경영학과(경영학석사)
2010년 2월 : 금오공과대학교
대학원 산업경영학과(경영학박사)
2014년 ~ 현재 : 금오공과대학교

IT융합학과 교수

2016년 ~ 2017년 : 3D 프린팅 산업협회장

관심분야 : 산업조직, HRD, 리더십, 3D 프린팅

임 재 권 (Jae-Kwon Eem)



1977년 2월 : 경북대학교
전자공학과(공학사)
1979년 2월 : 경북대학교 대학원
전자공학과(공학석사)
1991년 5월 : 미국 Texas A&M 대
전기및전자공학과(공학박사)
1979년 ~ 1986년 : 국방과학연구소

선임연구원

1991년 ~ 현재 : 금오공과대학교 전자공학부 교수

관심분야 : 컴퓨터비전, 영상처리, 칼라재현