

다중 카메라를 이용한 실시간 실내측위 방안

고영진*, 권성오°, 천성은*

Real-Time Indoor Positioning with Multiple Cameras

Young-jin Go*, Sungoh Kwon°, Seong-eun Cheon*

요약

본 연구에서는 다중 카메라를 이용한 실내 측위 방안을 제안한다. 스마트 공장 환경에서 작업관리와 안전관리 등을 위해 특정 물체의 정밀한 위치가 필요하다. CCTV(Closed Circuit TeleVision)나 웹캠 등의 다양한 비디오가 수집되고 있다. 복수의 카메라로부터 수집한 비디오 데이터를 이용해서 실내 3차원 위치를 정밀하게 측정하는 방안을 제안한다. 카메라의 특성과 배치 등의 다양한 환경에서 실험하여 성능을 검증하였다. 실험 결과 카메라와 물체의 거리가 2m인 환경에서 표준오차가 X축 9.2 mm, Y축 19.9mm, Z축 24.4mm 수준의 성능을 보였다.

Key Words : Indoor 3D positioning, Triangulation, Multiple camera, Image, Smart factory

ABSTRACT

In this paper, we propose an indoor positioning algorithm using video streams from multiple cameras. For safety and production management in a factory, a 3 dimensional location of a specific object is necessary. Video streams are gathered in a various ways such as CCTVs(Closed Circuit TeleVision) and webcams installed in the factory and the proposed algorithm estimates and traces locations of a specific object in a 3 dimension. We conducted experiments with various environments to verify performance. The experiments show root mean square errors of 9.2mm in the x-axis, 19.9 mm in the y-axis, and 24.4 mm in a z-axis.

1. 서론

근래의 실내 측위는 5G(the fifth generation) 기술의 발전에 따라 더 정밀하고 안정적인 기술이 필요하다. 실내 측위 기술은 지하철이나 공항에서 사람들의 안전 정보 활용으로 사고를 방지할 수 있다.^[1] 4차 산업혁명을 이끌게 되는 기술 중 IIoT(Industrial Internet of Things)에서는 특정 공간에서의 사용자 밀집도, 정보, 안전 관리가 필요하다.^[1] 특정 사용자의 위치와 공장 기계의 움직임을 실내에서 활용하기 위해서 정밀한 실내 3차원 측위가 필요하다.

실외 측위는 대표적으로 GPS(Global Positioning System)와 기지국을 활용한 측위로 구분된다. GPS는 위성으로부터 송신된 신호를 수신하여 좌표로 계산한다.^[2] GPS 기술은 정밀도가 높지만 실내에서는 GPS 신호를 수신하지 못하여 실내 측위로 적용하지 못한다. 기지국을 활용한 측위는 기지국으로부터 송신된 신호를 수신하여 좌표로 계산한다.^[3] 기지국을 활용한 측위는 실내에서 신호를 수신할 수 있지만 정밀도가 낮아서 실내 측위로 적합하지 않다.

다양한 기술들이 실내 측위를 위해 연구되고 있다. 각 실내 측위 기술은 장단점이 있고 상황에 맞는 기술

* 본 연구는 중소벤처기업부에서 지원하는 2018년도 산학연협력 기술개발사업(No. S2635310) 연구수행으로 인한 결과물임을 밝힙니다.

• First Author : School of Electrical Engineering, University of Ulsan, rhud999@naver.com, 학생(석사), 학생회원

° Corresponding Author : School of Electrical Engineering, University of Ulsan, sungoh@Ulsan.ac.kr, 정교수, 종신회원

* (주)스페이스코드, cse.dreamer@gmail.com, 정회원

논문번호 : 202004-074-C-RE, Received March 31, 2020; Revised June 9, 2020; Accepted July 19, 2020

이 필요하며, 필요한 방향성에 맞는 분야별로 기술들이 발전되었다.^[4] 실내 측위 기술은 크게 전파 기반, 센서 기반, 자기 시스템으로 구분된다.

전파 기반의 기술은 RF(Radio Frequency), UWB(Ultra Wide Band), US(Ultrasound), Lidar, RFID(Radio Frequency IDentification)가 있다.^[5] RF는 기존의 장비를 이용할 수 있는 장점에 비해 정밀도가 매우 낮은 단점이 있다. UWB는 침투성이 강하고 회절이 잘되 유연한 장점이 있지만 범위가 매우 작은 단점이 있다. US는 가격이 낮은 장점에 비하여 도플러 효과와 기타 요인에 영향을 많이 받는 단점 있다.^[6] Lidar는 높은 정밀도를 가지는 장점이 있지만 장비가 매우 비싼 단점이 있다. RFID는 회절이 되며 장비가 작은 장점이 있으나 사용범위가 매우 좁은 단점이 있다.

센서 기반의 기술로는 Floor Sensor, 광학 시스템이 있다. Floor Sensor의 경우 설치하면 편하게 이용이 가능한 장점에 반하여 정밀도가 낮고 다중 추적에 한계를 가지는 단점이 있다. 카메라를 이용하는 광학 시스템의 경우 높은 정밀도와 비 접촉적인 장점에 반하여 대상이 시야각 안에 존재해야 하는 단점이 있다.

마지막으로는 자기 시스템이 있다. 자기 시스템은 건축물 수준에서 회절과 투과성이 좋지만 자성을 포함한 물체에 영향을 받고 편차가 큰 단점이 있다. 표 1은 각 기술별로 장단점, 정밀도, 이용 가능한 범위까지 나타낸다.

자동화 공장에서 실내 측위를 위해 높은 정밀도와

같이 많은 장점을 가진 카메라를 이용하는 것이 적합하다. 카메라는 오차가 0.1mm에서 300mm 수준의 높은 정밀도를 가진다.^[5] 카메라는 다중 사용자를 식별할 수 있고 비접촉 방식이며 원거리에서도 식별용으로 이용할 수 있는 장점을 가진다.^[7] 카메라는 시야각 안에 물체가 존재해야 하는 단점이 있지만 공장에 존재하는 CCTV들을 연계하여 다중 카메라를 이용하면 단점을 해결할 수 있다. 실내 측위를 위한 광학시스템 이외의 기술 중 RF, RFID는 낮은 정밀도로 인하여 적합하지 않다. 자기장 시스템은 자동화 공장의 특성상 자기장에 영향을 주는 물질로 인해 적합하지 않다. Floor Sensor의 경우 추적이 독립적이라 적합하지 않다. UWB는 높은 투과성으로 인해 기타 장비들에서 나오는 신호에 간섭을 주고받을 수 있어서 적합하지 않다. US는 온도에 쉽게 영향을 받아서 오차 예측이 힘들기 때문에 적합하지 않다. Lidar는 비용이 많이 발생한다.

카메라를 이용한 측위 연구가 다방면으로 진행이 되고 있다. 2차원의 영역에서 단일 카메라를 이용해서 경계선을 식별을 하여 자신의 위치를 유도하고 로봇이나 차량의 경우 앞으로 나아가는 방향을 정하는 기술이 연구되고 있다.^[8] 하지만 곡선 환경과 큰 영역에서 활용하지 못하는 단점이 있다. 3차원 영역은 단일 카메라를 이용해 이미지의 초점에 따라 생기는 효과인 흐려지는 효과를 이용한 흐림도에 따른 3차원 깊이 추정기법이 있다.^[9] 하지만 단일 이미지에서 오류

표 1. 실내 측위 기술의 정확도와 범위 그리고 각 장단점[5]
Table 1. Accuracy and coverage of indoor positioning with each strength and drawback[5]

| Technology | Accuracy | Coverage | Strength | Drawback |
|---------------------------------|----------------|-----------------------|--|--|
| RF Base Measure | 2m ~ 20m | scalable | unmodified phones, PDAs | low reliability, pre calibration |
| UWB Base Measure | 50mm ~ 1m | 10m ~ 100m | penetration, multipath resistance | short range, computational cost |
| US Base Measure | 10mm ~ 50mm | 5m ~ scalable | low cost, low synchronization | temperature dependancy, multipath, Doppler shift |
| Lidar Base Measure | 1mm ~ 1m | 5m ~ 100m | high precision, low affect to weather | very high cost |
| RFID Base Measure | 100mm ~ 2m | 1m ~ 10m, scalable | penetration, unobtrusive | short range, elaborate deployment |
| Floor Sendor Base Measure | 100mm ~ 300mm | Room, scalable | unobtrusive, no tags on user | elaborate deployment, multi user tracking difficult |
| Optical System Base Measure | 0.01mm ~ 300mm | 0.1m ~ 10m | high precision, contactless tracking | no scale, Line of sight requirement |
| Magnetic System Base Measure | 1mm ~ 1m | 1m ~ 20m | penetration through building material | field disturbances, complexity of magnetic field |

가 크게 발생하고 두 개 이상의 이미지를 융합하면 오류가 감소하지만 동일 위치의 이미지가 필요해서 움직이는 물체의 추적은 한계가 있다. LED(Light Emitting Diode)의 불빛을 이용하여 4개점을 찍은 후 카메라로 찍어서 나오는 사진들을 이용하여 위치를 추정하는 기법이 있다.^[10] 하지만 LED 불빛이 보이는 장치를 카메라에 인식해야 하는 한계가 있다. 넓은 실내에서 정적인 물체를 기반으로 딥 러닝을 통한 실내 측위 방식이 있다.^[11] 하지만 기준이 되는 정적인 물체가 빛으로 인하여 딥 러닝을 위한 특이점 식별에 실패할 수 있다. 마지막으로 카메라의 속성과 다양한 기준 좌표를 이용해 실제 3차원 좌표계를 추정하는 방식이 있다.^[12] 다양한 요소를 이용해 실제 3차원 좌표계를 추정하는 방식은 높은 정밀도를 가지지만 좁은 영역의 한계와 정지상태의 요소들이 필요한 단점이 있다.

지금까지 연구에서 발견된 한계를 극복하기 위해 다중 카메라에서 추출된 동영상에서 특정 색상을 이용해 특정 물체를 추출한다. 중장거리 영역에서 활용 가능하고, 높은 정밀도와 외부의 영향을 적게 받고, 외부 센서들이 필요없는 3차원 위치에서 동적인 물체를 추적하는 방법이 필요하다. 본 연구를 통해 특정 물체의 정확한 3차원 위치 측위를 이용하여 활용할 방법을 연구한다. 특정 물체의 측위로 자동화 공장에서 특정 물체를 기준으로 움직임을 설정할 수 있다.

II. 시스템 모델

본 연구에서는 2개 이상인 N개의 카메라를 이용하여 특정 물체의 3D 위치를 찾는다. 카메라에서 추출되는 2차원의 이미지를 XY평면이라 하면 X축과 Y축 각각 화소수가 있다. X축의 화소는 X_{pix} , Y축의 화소가 Y_{pix} 일 때 본 연구에 사용된 N개의 카메라는 모두 같은 $X_{pix} \times Y_{pix}$ 의 화소를 이용한다. 카메라는 FOV(Field of View)인 시야각이 존재하며 모든 카메라의 시야각은 동일하게 한다. FOV는 카메라로 보이는 영역의 한쪽 대각선 끝에서 반대편 대각선 끝까지의 각도를 의미한다. 수평, 수직 방향의 시야각이 각각 FOV_x , FOV_y 라고 정의 한다.

n번째 카메라의 위치를 (X_n, Y_n, Z_n) 으로 정의한다. 모든 카메라들은 비틀림이 없고 Z좌표가 $Z_1 = \dots = Z_{n-1} = Z_n$ 와 같이 동일하다고 가정한다.

카메라와 물체의 거리를 이용하면 카메라에서 물체와의 각도를 알 수 있다. n번째 카메라와 물체 간의 각도를 각각 X, Y 축을 기준으로 알 수 있는데 이를

각각 $(\theta_{X(n)}, \theta_{Y(n)})$ 이라 한다. 예시는 그림 1에 변수는 표 2에 정리하였다.

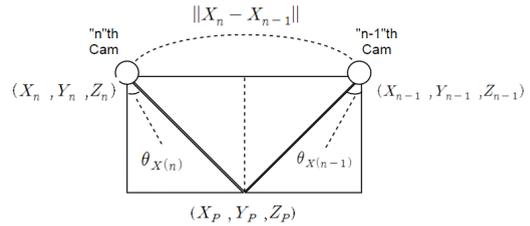


그림 1. 2중 카메라의 예시
Fig. 1. Example of two cameras

표 2. 사용된 변수 정리
Table 2. Parameter definitions

| Parameter | Definition |
|----------------------------------|---|
| n | number of cameras |
| (X_{pix}, Y_{pix}) | Value of used cameras 2D pixel |
| FOV | FOV of camera |
| FOV_x | Value of X-axis FOV |
| FOV_y | Value of Y-axis FOV |
| (X_n, Y_n, Z_n) | 3D position of n'th camera |
| (X_p, Y_p, Z_p) | 3D position of object |
| $(\theta_{X(n)}, \theta_{Y(n)})$ | x-axis, y-axis angle between object and n'th camera |
| $\ X_n - X_{n-1}\ $ | x-axis distance between n'th and n-1'th cameras |
| $\ Y_n - Y_{n-1}\ $ | y-axis distance between n'th and n-1'th cameras |

III. 제안 방법

본 장에서는 다중 카메라에서 추출된 비디오를 이용하여 특정 물체에 대한 3차원 위치 측정 알고리즘을 제안한다. 알고리즘은 다음 순서로 진행된다. 카메라의 재원과 위치 값을 설정하고, 각 카메라로부터 비디오를 수신한다. 특정 물체를 비디오로부터 추출하기 위해 색상의 기준을 변환한다. 카메라의 화소 오차에 따른 물체 인식 오차를 줄이기 위해 필터링(filtering)과 크기 제한을 통한 영상처리를 한다. 각 카메라로부터 수신된 비디오에서 추출한 이미지 상의 특정 물체 위치를 바탕으로 삼각 측위 법을 활용하여 3차원 위치를 측정한다. 그림 2는 알고리즘의 전체적인 개요이며, 1~4 절에서 각각 알고리즘을 단계별로 설명한다.

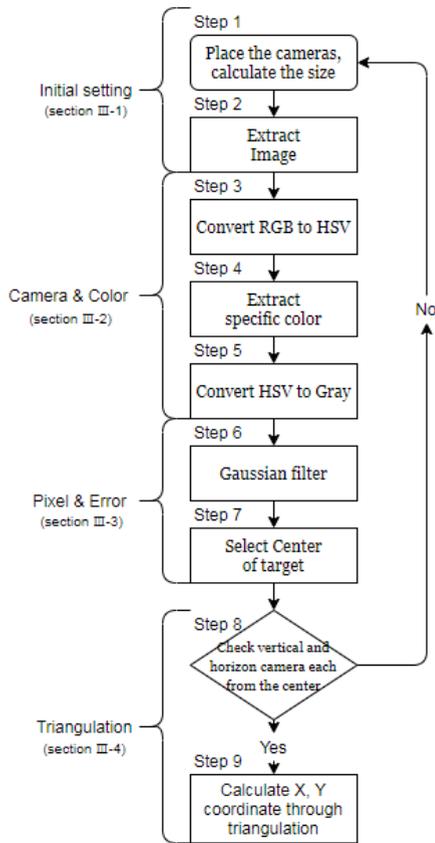


그림 2. 알고리즘
Fig. 2. Algorithm

3.1 카메라 실내 측위 기본 설정

첫 단계로 알고리즘에서 카메라 재원과 위치 값을 설정하고(Step 1), 각 카메라로부터 비디오 수신을 실행한다.(Step 2) 연구에 사용되는 카메라 재원은 화소, FOV와 카메라의 위치이다. 수신한 비디오는 프레임별로 분할하여 이미지화한다.

모든 카메라의 화소는 동일하게 설정한다. 만약 다를 경우 각 카메라가 갖는 화소별 거리가 다르고 정확도가 떨어진다. 그래서 모든 카메라의 화소가 동일하다고 가정한다.

FOV는 카메라가 제작될 때 설정되며 기본 설정의 비율이 존재하고 임의로 변경 시 오류가 발생한다. 연구에서 사용된 광학카메라는 FOV가 60~80°로 한정된다.^[13] 16:9의 화소 비율에서 $FOV = \theta_{16:9}^\circ$, $FOV_x = \theta_{16:9.x}^\circ$, $FOV_y = \theta_{16:9.y}^\circ$ 시야각으로 설계된 카메라와 4:3의 화소 비율에서 $FOV = \theta_{4:3}^\circ$, $FOV_x = \theta_{4:3.x}^\circ$, $FOV_y = \theta_{4:3.y}^\circ$ 시야각의 카메라

로 설계되어 있다. 물체까지의 z축 거리가 z라고 가정하면 아래의 수식과 같이 $z \tan(\theta_{16:9.x}/2) : z \tan(\theta_{16:9.y}/2) = 16 : 9$ 와 $z \tan(\theta_{4:3.x}/2) : z \tan(\theta_{4:3.y}/2) = 4 : 3$ 이 구해진다. 16:9의 비율로 제작된 카메라와 4:3의 비율로 제작된 카메라가 서로 다른 비율의 이미지로 추출이 가능하다. 하지만 임의로 FOV가 조정되어 영상에 빈 공간이 생기므로, 모든 카메라를 동일한 FOV 비율로 설정한다. 동일한 카메라가 아닐 경우 각 카메라에 맞는 FOV 비율로 설정한다.

비디오 수신은 수신되는 시간이 동일하다고 가정한다. 만약 수신 받는 시간이 동일하지 않으면 오차를 유발할 수 있다. 따라서 카메라의 프레임별 시간이 동기화된다고 가정한다.

3.2 카메라 실내 측위와 색상

두 번째 단계는 알고리즘으로부터 수신된 이미지에 들어있는 특정 물체를 이미지로부터 효율적으로 추출하기 위해 색상 변환을 한다. 예를 들어 붉은색을 추출할 때 붉은색은 대부분의 색에 연하거나 진하게 섞여 있는데 순수한 붉은색을 선택해야 한다.

색상표는 RGB, HSV 등으로 구분이 되어 있다. RGB는 붉은색(Red), 초록색(Green), 파란색(Blue)으로 이루어지는 3원색의 조합으로 특정 색을 추출할 때 범위 설정에 어려움이 있다. HSV는 고유의 색인 색상(Hue), 빛과 같은 외부 요인이 작용하는 채도(Saturation), 명도(Value)를 이용하는 방식이다. 그러므로 본 연구에서는 HSV 색상 체계를 이용한다.

수신된 이미지는 RGB 색상표를 기준으로 추출이 되며 HSV 색상표로 변환하면 색상(H)은 0~360°로 범위가 나누어지며 순수 색 정보를 가지고 있다. 따라서 원하는 물체의 색상 추출에 있어 정확도 향상을 위해 HSV 기반으로 변환한다.(Step 3)

색상표를 설정하고 색상을 기준으로 특정 물체를 분류하여 추출한다. HSV 변환으로 색을 결정하는 색상(H)의 설정 중 붉은색은 0~20°, 340~360°로 이루어져 있고 나머지는 연속된 값으로 이루어진다. 채도(S)는 0~255까지의 값으로 이루어져 있는데 수치가 높을수록 진한 색이며 0~60 수준의 수치는 색이 연해서 제외한다. 명도(V)도 0~255까지로 표기되며 높아질수록 밝아진다. 따라서 어두워서 색을 식별하기 힘든 0~60 수치를 제외한다.(Step 4)

물체를 추출하고 나면 목표 이외에도 유사색상의 다른 물체가 식별되기도 하며, 이러한 잡음(Noise)을 제거하기 위하여 그레이 스케일한다. 그레이 스케일은



그림 3a. RGB를 기준으로 붉은색 추출 예시
Fig. 3a. Example of red color extraction from RGB



그림 3b. HSV를 기준으로 붉은색 추출 예시
Fig. 3b. Example of red color extraction from HSV

구별 범위 별로 2진화, 4진화, 8진화 등등 채도 별로 구별하는 방법이 있다. 본 실험에서는 필요한 값과 필요 없는 값으로 나뉘어 2진화한다. 2진화 그레이 스케일로 객체를 검은색 이외를 흰색으로 구별해 알고리즘 계산을 위한 2진화를 한다.(Step 5)

3.3 카메라의 화소와 오류 검사

카메라의 화소 오차에 따른 물체 인식 오차를 감소시키기 위해서 가우시안 필터(Gaussian filter)를 이용한다. 오류로 판단이 되기 쉬운 고주파 신호들을 필터링하여 제거하고 추출한 물체들에서 적합하지 않은 크기의 물체들을 제거한다.

필터링을 이용한 화소의 오류 감소방식은 중심 화소를 기준으로 주변의 화소와 비교하여 평균을 내거나 비율이 높은 값을 선택을 하여 화소 별로 조금씩 발생할 수 있는 오차를 줄여나간다.^[14] 본 연구에서 사용한 가우시안 필터는 중심으로 갈수록 높아지는 가중치를 고려하여 평균을 낸다. $n \times n$ 의 범위를 정의하여 n^2 의 크기이다.

식 (1)은 가우시안 필터의 수식의 예시이며 중심과 가까울수록 가중치가 높다.^[15]



그림 4a. 필터로 잡음 제거 전 이미지 예시
Fig. 4a. Example of an image before noise filtering

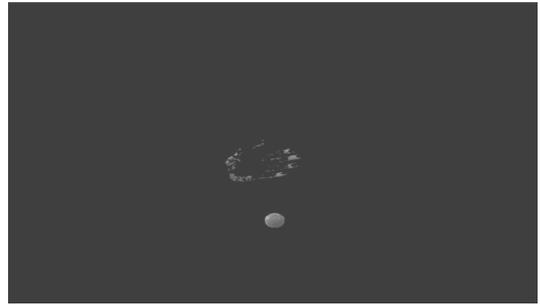


그림 4b. 필터로 잡음 제거 후 이미지 예시
Fig. 4b. Example of an image after noise filtering

$$1/273 \begin{pmatrix} 1 & 4 & 7 & 4 & 1 \\ 4 & 16 & 26 & 16 & 4 \\ 7 & 26 & 41 & 26 & 7 \\ 4 & 16 & 26 & 16 & 4 \\ 1 & 4 & 7 & 4 & 1 \end{pmatrix} \quad (1)$$

식 (2)는 평균 필터(Averaging filter) 수식의 예시이며 모든 영역의 가중치가 동일하다.

$$1/25 \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \quad (2)$$

위와 같은 필터를 이용하여 목표 이미지에 빛과 반사 등으로 발생한 잡음을 제거하여 3차원 측위를 계산할 수 있게 된다.(Step 6)

빛과 반사 등으로 발생한 잡음들을 필터로 제거한 후 목표의 중심을 구한다. 잡음 제거가 끝난 이미지를 보면 특정 색 추출로 실험을 했기 때문에 우리가 원하는 물체 이외의 다른 물체 또한 포함되기도 한다. 그래서 사용한 물체의 크기를 고려하여 가로세로 화소의 크기가 일정 화소 이상의 값인 것들만 식별되게 설

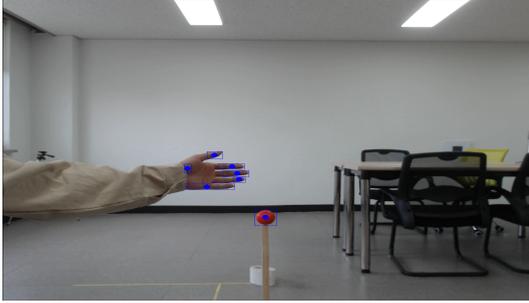


그림 5a. 목표 이외의 물체 제거 전 물체의 중심점
Fig. 5a. Center of the red target before removing other objects

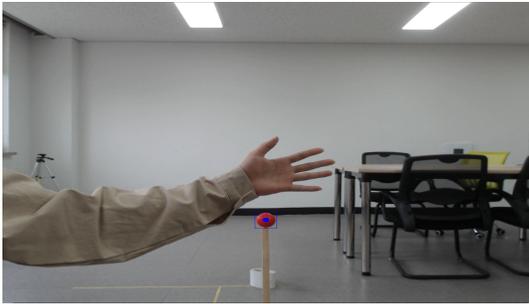


그림 5b. 목표 이외의 물체 제거 후 물체의 중심점
Fig. 5b. Center of the red target after removing other objects

정한다. 물체의 테두리를 측정하여 테두리를 기준으로 물체의 중심을 구한다.(Step 7)

3.4 카메라에서 삼각 측위 활용

각 카메라에서 추출한 이미지 상의 물체 위치를 바탕으로 삼각 측위 법을 활용하여 3차원 위치를 측정한다. 모든 카메라에 식별이 되었는지 확인한다.

추출 완료된 물체의 중심점, 카메라의 중심점, FOV, 카메라 간의 거리에서 오류를 확인한다. 오류가 있으면 Step 1로 돌아가고 오류가 없으면 삼각 측량을 진행한다.(Step 8)

화소로 나타내는 2차원 이미지와 실제 물체는 3차원에 있어서 다르다. 물체의 위치를 2차원 화소 좌표에서 (X', Y') 이 된다. 3차원으로 나타내기 위해 임의의 Z축을 추가하면 (X', Y', Z') 가 된다. 물체의 실제 3차원 좌표는 (X_p, Y_p, Z_p) 이다.

X축의 값을 계산하기 위하여 기준 카메라 1번과 X 좌표가 멀리 있는 2번 카메라로 측정한다. n번째 카메라의 중심에서 X축 방향 FOV_X 의 절반과 카메라의 X축 방향의 화소 값 X_{pix} 을 이용해서 카메라 Z축의 X축 대비 화소 값은 Z' 이며 식 (3)과 같다.

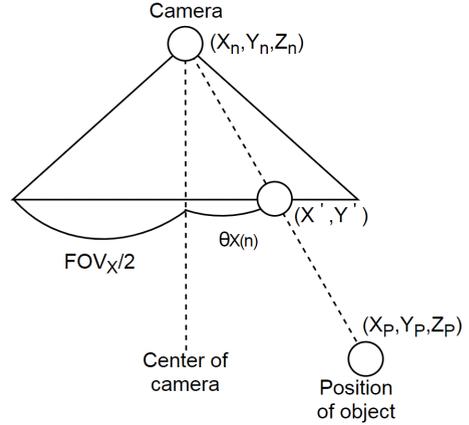


그림 6. 단일카메라에서 카메라 중심과 물체 위치
Fig. 6. Positions of an object and a single camera

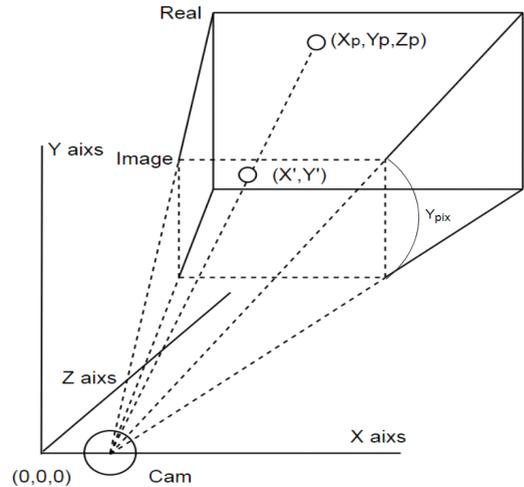


그림 7. 실제와 이미지상 목표의 좌표
Fig. 7. Real coordinate and coordinate projected on an image of a target

$$Z' = (X_{pix} / 2) \tan(FOV_X / 2) \quad (3)$$

식 (3)에서 Z' 과 카메라 중심에서 물체 중심까지의 X축 화소 수 X' 을 이용해 카메라 중심에서 물체의 중심까지의 각도를 구한다. 카메라 중심에서 물체의 중심까지의 각도 $\theta_{X(n)}$ 은 식 (4)와 같다.

$$\theta_{X(n)} = \tan^{-1}(X' / Z') \quad (4)$$

식 (4)를 통하여 카메라별로 물체까지의 각도는 $\theta_{X(1)}$ 와 $\theta_{X(2)}$ 이다. 물체의 중심과 카메라 간의 Z축의 실제 거리를 Z_p 라 하면 카메라 간의 X축 거리는

$\|X_2 - X_1\|$ 이며 식 (5)와 같다.

$$\|X_2 - X_1\| = Z_P \tan(\theta_{X(1)}) + Z_P \tan(\theta_{X(2)}) \quad (5)$$

1번 카메라가 중심 (0, 0, 0)이라고 가정을 하면 $Z_P \tan(\theta_{X(1)})$ 으로 실제 X좌표의 거리인 X_P 를 구할 수 있다. 식 (3), (4), (5)를 반복하여 기준 카메라 1번과 Y좌표가 멀리 있는 3번 카메라로 측정한다. Y_P 의 값을 계산하면 식 (5)를 통하여 Z_P 를 계산하면 (X_P, Y_P, Z_P)가 모두 계산된다.(Step 9)

IV. 시스템 모델

4.1 실험 환경

복수의 카메라로 비디오 추출을 통한 3차원 실내 측위를 실험한다. 식 (3)에 의하여 화소, FOV 가 식 (5)에 의하여 카메라 간의 거리, 목표의 Z축 좌표가 알고리즘에 영향을 준다. 해당 영향들이 결과의 오차에 미치는 영향을 실험을 통해 검증한다.

1번 카메라 위치를 원점 (0, 0, 0)으로 설정한다. 나머지 카메라를 $Z = 0$ 인 평면에 위치한다. 카메라는 모두 동일한 카메라를 사용하며 FOV 와 화소 설정을 동일하게 한다. 그림 8은 카메라가 3개일 때 그림 9는 카메라가 2개일 때의 실험 환경이다.

카메라의 수평방향 FOV 는 70.42° , 수직방향 FOV 는 43.3° 이다. FOV 로 결정된 비가시 거리는 카메라 간 거리의 절반에 $\tan(\pi/2 - FOV/2)$ 를 곱해 나오는 값으로 카메라 간의 거리가 1m일 때 0.60~0.86m 이후부터 모든 카메라에 이미지가 식별된다. 따라서 Z축이 1m 이후에서 실험한다.

화질의 경우 16:9의 비율인 720p, 1080p의 화질을 사용하고 목표물의 경우 지름 7.5cm 붉은색 구형의

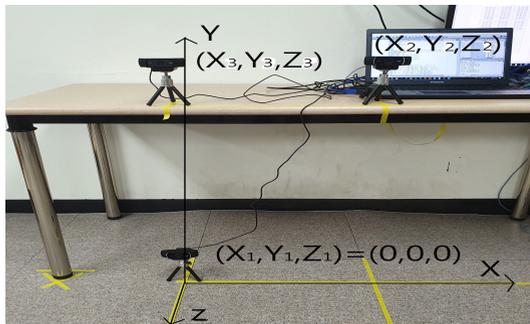


그림 8. 3중 카메라 실험 환경
Fig. 8. Experiment environment of three cameras

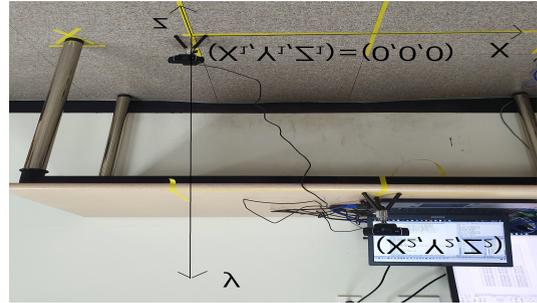


그림 9. 2중 카메라 실험 환경
Fig. 9. Experiment environment of two cameras

물체를 사용한다. 물체의 크기와 실험 환경을 고려하여 식별하는 물체를 50화소 이상의 물체가 식별되게 한다.

본 논문에서 제안하는 알고리즘은 1초마다 위치를 측정하게 설정하였으며, 하드웨어 성능에 따라 다르게 설정할 수 있다. 이러한 실시간 실내측위의 정확도를 확인하기 위한 실험을 수행하였고, 5개 데이터의 평균 값이다. 구형의 물체는 카메라들의 안쪽에서 3×3 의 지점 9개에 위치한다. Z좌표는 모든 카메라에서 식별되는 범위 내의 0.47m마다 측정하였다.

본 실험은 실험 환경을 바탕으로 목표물을 9개의 지점으로 이동하며 진행한다. 모든 실험값은 비디오의 프레임에서 5개의 값을 평균 내서 도출한다.

4.2 화소에 따른 비교 실험

화소를 제외한 실험 환경을 동일하게 하여 화소의 변화가 오차에 미치는 영향을 알아본다. 식 (3)에서 X_{pix} 이 증가하면 Z' 이 선형적으로 증가하여 정밀도가 향상된다. 실험 환경은 카메라를 그림 8과 같이 $(X_1, Y_1, Z_1) = (0, 0, 0)$, $(X_2, Y_2, Z_2) = (0, 745, 0)$, $(X_3, Y_3, Z_3) = (955, 745, 0)$ 인 기본 환경으로 배치한다. 측정하는 목표물은 Z 좌표가 2180mm에서 그림 10과 같이 9개의 지점의 결과를 도출한다. 높은 화소인 1080p와 낮은 화소인 720p를 그림 11과 같이 측정한다.

그림 12의 첫 번째와 두 번째의 값은 그림 11의 결과값에 기반해 오차를 표준오차(RMSE : Root mean square error)로 비교한다. 그림 12의 세 번째와 네 번째는 Z축이 2650mm에서 화소를 비교한다.

그림 12를 보면 Z축이 2180mm인 기본 환경에서는 표준오차가 12.1mm로 부터 11.4mm로 감소하고 Z축이 2650mm에서는 표준오차가 13.4mm로 부터 13.0mm로 감소해 화소가 증가하면 표준오차가 감소한다.

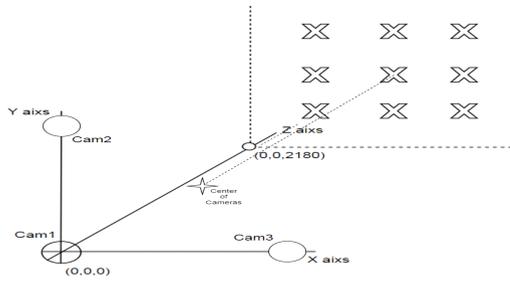


그림 10. 실험 환경 예시
Fig. 10. Illustration of an experiment environment

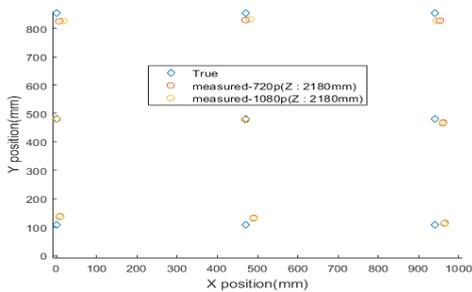


그림 11. 카메라 화소수에 따른 위치측정 결과 비교
Fig. 11. Measurement comparison according to the Number of pixels

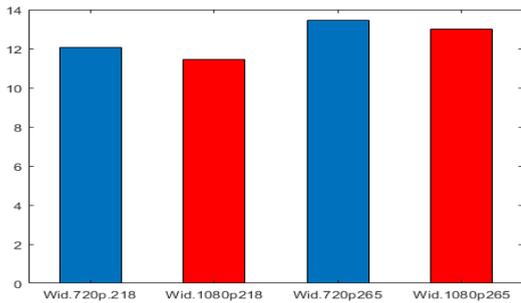


그림 12. 화소에 따른 표준오차 비교
Fig. 12. RMSE comparison according to the number of pixels

실내측위에서 정밀한 측정결과를 갖는 초음파를 이용한 실내측위 방법인 초음파/RF를 이용한 2차원 측위법⁶⁾과 정밀도를 비교하였다. 세 개의 초음파 센서를 이용한 기존의 논문은 2차원의 위치만을 측정이 가능한 반면, 본 논문에서 제안하는 위치측위 방안은 3차원까지 측위를 한다. 따라서, 두 측위기술 모두 2차원 (x, y)좌표에 대한 오차를 비교한다. 본 실험과 유사한 실험환경을 가진 초음파/RF의 실험⁶⁾에서 가운데 9개 지점에 대한 2차원의 오차평균이 57mm인

반면, 본 실험에서는 화소가 720p와 1080p인 경우 2차원에서 표준오차가 각각 12.1mm와 11.4mm으로 초음파/RF를 이용한 측위 기법보다 정밀도가 높다.

4.3 카메라의 FOV와 오차 관계 실험

카메라의 FOV는 알고리즘에서 이용되는 변수로 오차에 미치는 영향을 알아본다. 식 (3)에서 FOV_x 가 증가하면 Z' 이 증가하여 정밀도가 향상된다. 그림 8과 그림 9는 카메라가 가로로 위치한다. 따라서 카메라의 설정에 의해 Y축이 X축보다 FOV가 작다. 카메라를 세로로 위치하면 Y축이 X축보다 FOV가 크다. 기본 환경에서 카메라가 가로일 때(Case1)와 카메라가 세로일 때(Case2)를 비교한다.

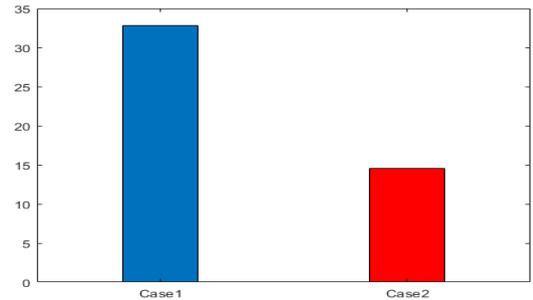


그림 13. FOV에 따른 표준오차 비교
Fig. 13. RMSE comparison according to FOV

4.4 카메라 간의 거리에 따른 비교 실험

카메라를 설치할 때 카메라 간의 거리가 정밀도에 미치는 영향을 알아본다. 식 (5)에서 $\|X_2 - X_1\|$ 이 감소하면 Z_p 은 고정된 값이기 때문에 $\theta_{X(1)}, \theta_{X(2)}$ 가 감소한다. $\theta_{X(1)}, \theta_{X(2)}$ 은 물체의 좌표를 구하는 요소이다. 그림 14는 기본 환경에서 카메라 간의 거리가 좁은 환경이 되게 $(X_3, Y_3, Z_3) = (445, 745, 0)$ 로 변경한다. 2절에서 확인한 1080p의 화소에서 기준 목표점을 실험한다. 그림 15의 경우 첫 번째 값과 두 번째 값은 그림 14에 기반해 Z축이 2180mm 일 때 그리고 세 번째 네 번째 값은 Z축이 2650mm 일 때의 기본 환경과 좁은 환경의 표준오차를 이용하여 비교한다.

Z축이 2180mm인 좁은 환경일 때 39.5mm에서 기본 환경일 때 11.4mm로 감소하고 2650mm인 좁은 환경일 때 32.6mm에서 기본 환경일 때 13.0mm으로 감소한다. 기본 환경에서 표준오차가 작게 발생한다. 하지만 카메라 간의 거리를 넓히면 FOV로 인한 비가시 영역이 증가하는 단점이 있다. 그래서 최적의 카메라 간의 거리와 FOV가 필요하다.

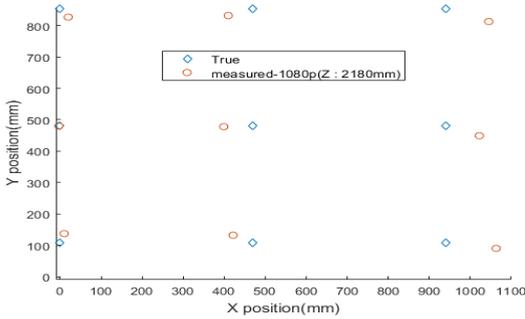


그림 14. 좁은 환경에서의 측정결과
Fig. 14. Measurements in a narrow area

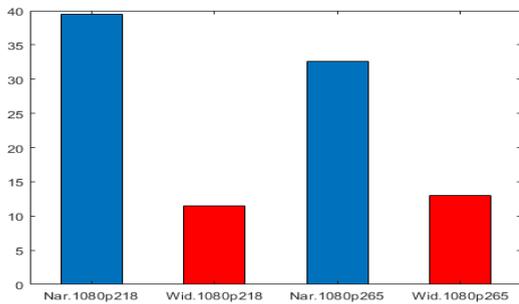


그림 15. 측정 환경에 따른 표준오차
Fig. 15. RMSE according to measurement environment

4.5 목표의 Z축 좌표가 멀리 있는 상황 실험

카메라의 위치에서 Z축 거리가 정밀도에 미치는 영향을 알아본다. 식 (5)에서 $\|X_2 - X_1\|$ 은 상수이면 $Z_p \tan(\theta_{X(1)})$, $Z_p \tan(\theta_{X(2)})$ 는 상수가 된다. $\tan(\theta_{X(1)}) = \text{상수}(const) / Z_p$ 와 식 (3)에서 상수 Z' 를 이용한 식 (4)에서 $\tan(\theta_{X(1)}) = X' / Z'$ 를 연립하면 상수 = $X' \times Z_p$ 가 되어 Z축 좌표의 거리의 증가로 인한 X' 의 감소가 Z_p 의 증가를 보인다. 2절에서 4절까지 실험은 Z축이 2180mm 일 때와 2650mm 일 때만 비교했으나 1080p의 기본 환경일 때 그림 16은 2180mm에서 5000mm까지 Z값을 비교한다.

그림 16과 같이 Z축의 값이 증가함에 따라 오차가 증가한다. $X_p/Z_p = \tan(\theta_{X(1)})$ 에서 Z축의 오차가 $\tan(\theta_{X(1)})$ 의 변화에 영향을 받는다.

4.6 카메라 개수 차이에 따른 비교 실험

카메라의 개수를 변경해 알고리즘으로 구해지는 평균값의 정밀도를 비교한다. 카메라의 개수는 비용에 직접적으로 영향을 주기 때문에 적당한 수량의 제한이 필요하다. 화소는 적합한 화소인 1080p를 이용한

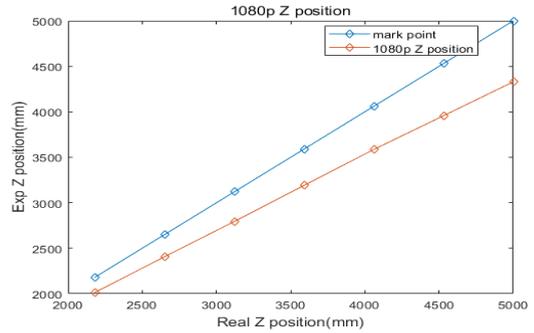


그림 16. Z좌표 거리별 결과
Fig. 16. Measured vs. real Z positions

다. 카메라는 그림 8과 같이 기본 환경으로 한다. 3번 카메라를 제거하여 그림 9와 같이 만든 환경과 1,2번 카메라 사이에 4번 카메라를 설치한 카메라 4개일 때의 표준오차로 비교한다.

Z축이 2180mm 일 때 카메라 2개는 20.08mm에서 카메라 3개가 되면 11.47mm으로 4개가 되면 8.0mm으로 감소하고 2650mm 일 때 카메라 2개는 26.95mm에서 카메라 3개가 되면 13.01mm로 4개가 되면 11.8mm로 감소한다. 카메라 개수가 2개일 때는 X축과 Y축이 모두 왜곡률이 높은 위치에 물체가 존재하지만 카메라 개수가 3개가 되면 X축과 Y축 모두 왜곡률이 적은 위치에 물체가 존재한다. 카메라가 3개에서 4개로 증가해도 감소되지만 감소폭의 효율상 카메라는 3개까지 이용한다.

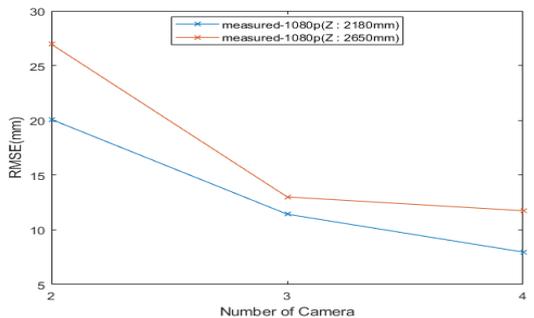


그림 17. 카메라 개수에 따른 표준오차
Fig. 17. RMSE vs. the number of cameras

V. 결론

복수의 카메라로부터 수신한 동영상으로 특정 물체의 3D 위치를 측정하는 알고리즘을 제안하였다. 화소, 카메라 간의 거리, 카메라 개수, 물체의 Z축 거리의

변화에 따른 실험을 통해 성능을 검증하였다. 실험 결과 화소가 높을수록 기준이 되는 카메라 간의 거리를 세부적으로 분할되고, 카메라 간의 거리가 증가하면 기준거리의 증가한다. 카메라 개수의 증가로 왜곡률이 감소하고, Z축 좌표가 가까울수록 카메라와 물체까지 각도 감소와 카메라 중심과 물체중심의 거리가 감소하여 정밀도가 증가한다. 카메라 3개를 이용하고 1080p인 화소로 기본 환경인 $Z_p = 2180\text{mm}$ 에서 특정 물체의 표준오차가 X축은 9.2mm Y축은 19.9mm Z축은 24.4mm가 나왔다.

기존의 근거리 실내 측위에서 이점을 보이는 본 연구와 유사환경의 초음파/RF를 이용한 기술에서 오차 평균이 57mm이다.^[6] 반면, 본 연구에서는 11.4mm로 더 오차가 감소되었다. Z축의 정밀도는 다른 축과 비교하면 오차가 크며 추가적인 연구가 필요하다. 본 연구에서는 자동화 공장에 중심을 가지고 연구했으나 공장의 자동화 기술 외에도 자율 주행 드론과 같이 특수 물체를 추적할 수 있다.

제안한 알고리즘은 왜곡률과 비틀림에 취약하다. 왜곡률과 비틀림을 감소하기 위하여 세부적인 요소를 관리할 수 있는 캘리브레이션 기술의 발전으로 카메라 개수를 감소할 수 있다. 카메라 렌즈에 의한 이미지의 비틀림과 왜곡률을 고려한 위치측위 방안에 대한 연구는 향후에 계속 수행예정이다.

References

[1] J. M. Lim, *Trends in indoor positioning*(2014), Retrieved Jan. 16, 2020, from <https://mycoordinates.org/trends-in-indoor-positioning/>

[2] M. M. Din, N. Jamil, J. Maniam, and M. A. Mohamed, "Indoor positioning: Technology comparison analysis," *Int. J. Eng. & Technol.*, vol. 7, no. 2.14, pp. 133-137, Feb. 2018.

[3] C. Laoudias, R. Pich, and C. G. Panayioyou, "Device self-calibration in location systems using signal strength histograms," *J. Location Based Services*, vol. 7, no. 3, pp. 165-181, Aug. 2013.

[4] B. Ray, *How An Indoor Positioning System Works*(2018), Retrieved Jan. 17, 2020, from <https://www.airfi-nder.com/blog/indoor-positioning-system>

[5] R. Mautz, "Indoor Positioning Technologies,"

Ph.D. Thesis, ETH Zurich, Zurich, Switzerland, 2012.

[6] B. D. Zho, S. Kwon, and S. E. Cheon, "Indoor positioning system using ultrasonic and RF," *J. KICS*, vol. 42, no. 2, pp. 413-423, Feb. 2017.

[7] M. C. Kim and K. R. Park, "A survey on face recognition in surveillance camera system," in *Proc. KICS ICC 2017*, pp. 530-531, Gangwondo, Korea, Jan. 2017.

[8] K. Lee, H. Kim, J. Kho, S. Lee, and M. Park, "Position estimation and navigation control of mobile robot using mono vision," *Automation and Syst. Eng.*, vol. 5, no. 5, pp. 529-539, Jul. 1999.

[9] S. Kim, Y. M. Ko, C. Bae, and D. J. Kim, "3D depth estimation by a single camera," *JBE*, vol. 24, no. 2, pp. 281-291, Mar. 2019.

[10] S. H. Jung, M. W. Lee, K. Y. Kim, and J. S. Cha, "Position information acquisition method based on LED lights and smart device camera using 3-Axis moving distance measurement," *J. KICS*, vol. 40, no. 01, pp. 226-232, Jan. 2015.

[11] A. Xiao, R. Chen, D. Li, Y. Chen, and D. Wu, *An indoor positioning system based on static objects in large indoor scenes by using smartphone cameras*, Retrieved Mar. 30, 2020, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6-069275/>

[12] K. Choi, H. Lee, S. Cheon, and S. Kwon, "3 dimensional location estimation using image process," *J. KICS*, vol. 2019, no. 6, pp. 588-589, Jun. 2019.

[13] M. Park, *Sigma A 30mm F1.4 DC HSM for Canon*(2014), Retrieved Mar. 27, 2020, from <https://canoninverse.tistory.com/34>

[14] R. Chandel and G. Gupta, "Image filtering algorithms and techniques: A review," *IJARCSSE*, vol. 3, no. 10, pp. 198-202, Oct. 2013.

[15] E. S. Gedraite and M. Hadad, "Investigation on the effect of a Gaussian Blur in image filtering and segmentation," in *Proc. IEEE ELMAR-2011*, pp. 393-396, 2011.

고 영 진 (Young-jin Go)



2020년 2월 : 울산대학교 전기
공학부 졸업
2020년 3월~현재 : 울산대학교
전기공학부 석사과정
<관심분야> 위치 측위, 센서
네트워크, 이미지 프로세스,
AI, 추정 이론, 통신 공학

[ORCID:0000-0003-4082-1658]

천 성 은 (Seong-eun Cheon)



1995년 2월 : 고려대학교 전자
공학과 졸업
1995년~1997년 : (주)엘지반도체
기술연구소 연구원
2000년~2003년 : (주)지메이트 정
보통신연구소 선임연구원
2003년~2006년 : (주)현대디지털
테크 부설연구소 책임연구원

2006년~2011년 : (주)시소시스템 대표이사

2017년~현재 : (주)스페이스스코드 대표이사

<관심분야> 실내측위, 센서네트워크, 자동제어, 통
신시스템

[ORCID:0000-0001-8727-6656]

권 성 오 (Sungoh Kwon)



1994년 2월 : KAIST 전기및전
자공학과 학사
1996년 2월 : KAIST 전기및전
자공학과 석사
1996년 3월~2001년 1월 : 신세
기통신 기술연구소
2007년 8월 : Purdue University
ECE 박사

2007년 8월~2010년 2월 : 삼성전자 DMC연구소 수석
연구원

2010년 3월~2014년 3월 : 울산대학교 전기공학부 조교수

2014년 4월~2019년 2월 : 울산대학교 전기공학부 부교수

2019년 3월~현재 : 울산대학교 전기공학부 교수

<관심분야> 이동통신 네트워크, MAC 설계, 간섭제어,
최적화, 신호처리, 뇌공학

[ORCID:0000-0001-5265-862X]