Journal of Institute of Control, Robotics and Systems (2016) 22(11):952-959
http://dx.doi.org/10.5302/J.ICROS.2016.16.0079 ISSN:1976-5622 eISSN:2233-4335

실내 환경에서의 이동로봇의 위치추정을 위한 카메라 센서 네트워크 기반의 실내 위치 확인 시스템

Indoor Positioning System Based on Camera Sensor Network for Mobile Robot Localization in Indoor Environments

지 용 훈*, Atsushi Yamashita, Hajime Asama
(Yonghoon Ji[1,*], Atsushi Yamashita[1], and Hajime Asama[1])
[1]Department of Precision Engineering, The University of Tokyo

Abstract: This paper proposes a novel indoor positioning system (IPS) that uses a calibrated camera sensor network and dense 3D map information. The proposed IPS information is obtained by generating a bird's-eye image from multiple camera images; thus, our proposed IPS can provide accurate position information when objects (e.g., the mobile robot or pedestrians) are detected from multiple camera views. We evaluate the proposed IPS in a real environment with moving objects in a wireless camera sensor network. The results demonstrate that the proposed IPS can provide accurate position information for moving objects. This can improve the localization performance for mobile robot operation.

Keywords: GPS (Global Positioning System), IPS (Indoor Positioning System), camera network, mobile robot localization

I. 서론

본 논문은 실내 환경에서의 이동로봇 등의 위치추정 성능을 향상시키기 위한 분산 카메라 네트워크와 3차원 지도 정보 기반의 실내 위치 확인 시스템(IPS: Indoor Positioning System)을 제안한다. 실내 환경에서 이동로봇을 운용하기 위한 대부분의 주행 기술은 SLAM (Simultaneous Localization and Mapping)을 통해 작성된 환경의 지도 정보를 기반으로 구현된다[1]. 그러나 최근에 인간과의 공존 환경에서의 로봇을 운용하는 경우가 늘어남에 따라 환경의 동적인 변화를 반영하지 못하는 정적인 지도 정보만으로는 로봇의 운용이 어려운 경우가 많이 발생하고 있다.

한편, 환경 내에 분산 카메라 센서 인프라를 구축한다면 환경의 동적 변화에 대한 실시간 모니터링이 가능해지며, 이러한 센서 네트워크의 자동 캘리브레이션 시스템에 대한 연구가 다수 제안되었다[2-4]. 즉, 분산 카메라 네트워크로부터의 영상 정보를 처리함으로써 실시간으로 환경의 동적 물체(이동로봇 및 보행자 등)를 검출할 수 있다. 따라서 정적인 지도 정보와 캘리브레이션된 카메라 네트워크 시스템으로부터의 동적 정보를 융합함으로써 이동로봇의 위치추정 성능을 크게 향상시킬 수 있다.

환경의 분산 센서 네트워크 정보를 이용하여 이동로봇의 위치추정의 성능을 향상시키는 연구가 다수 제안되었다[5,6]. 그러나 이러한 연구들에서는 단순한 센서 네트워크의 구축뿐만 아니라 구축된 센서 네트워크 시스템의 최적화를 위하여 환경 및 이동로봇에 대한 추가적인 구조물 설치 등 대상 공간의 구조에 대한 큰 변형이 요구된다. 또한, 카메라 네트워크로부터의 2차원 영상 정보만으로는 3차원 위치 정보의 취득이 불가능하기 때문에 추가적인 레이저 거리 센서 기반의 네트워크 구축이 필요하다.

실외 환경에서 위성 위치 확인 시스템(GPS: Global Positioning System)은 가장 쉽게 위치 정보를 획득할 수 있는 장치로써 이동로봇의 위치추정의 성능 향상에 큰 도움을 준다[7,8]. 그러나 위성 신호의 수신이 불가능한 실내 환경에서는 동작하지 않으며, 이러한 문제를 해결하기 위해 Wi-Fi, 전자장 등 다양한 신호를 활용한 실내 위치 확인 시스템들이 제안되었다[9-11]. 그러나 신호의 난반사 및 정확도 등에 있어서 많은 문제점들이 남아있다. 따라서 이동로봇의 위치추정의 성능 향상을 위해 본 연구에서는 분산 카메라 센서 네트워크와 3차원 지도 정보를 활용한 실내 위치 확인 시스템을 제안한다. 전체 환경에 대한 모든 점유 좌표 정보를 포함하고 있는 3차원 지도 정보는 건물 설계 시의 CAD 정보 및 SLAM 기술 등을 통해 사전에 얻을 수 있다.

본 논문에서는 복수의 카메라 영상으로부터 조감도 영상(bird's-eye image)을 생성하여 이동 물체의 위치 신호를 생성한다. 조감도 영상은 환경을 공중에서 수직으로 바라본 시점의 영상이며, 일반적으로 실내 환경의 바닥 면은 평평하고 높이 정보는 무시 가능하기 때문에 이동 물체의 위치 검출을 위한 매우 효과적인 정보를 제시할 수 있다. 따라서 본 논문에서는 복수의 카메라 영상에서 동시에 검출된 이동 물체의 영역을 조감도 영상을 통해 검출하여 실내 환경에서의 신뢰성 있는 위치 신호를 생성한다.

본 논문의 구성은 다음과 같다. II장에서는 실내 위치 신호

* Corresponding Author
Manuscript received April 18, 2016 / revised August 13, 2016 / accepted October 5, 2016
지용훈, Atsushi Yamashita, Hajime Asama: 도쿄대학 정밀공학전공
(ji@robot.t.u-tokyo.ac.jp/yamashita@robot.t.u-tokyo.ac.jp/asama@robot.t.u-tokyo.ac.jp)
※ This work was funded by Tough Robotics Challenge, ImPACT Program of the Council for Science, Technology and Innovation (Cabinet Office, Government of Japan).
※ 이 논문은 2016 제 31회 제어로봇시스템학회 학술대회에 초안이 발표되었음.

의 생성을 위한 카메라 네트워크 시스템 기반의 전역 조감도 영상의 생성법에 대해 자세히 설명하고, III장에서는 생성된 조감도 영상을 활용하여 실내 위치 신호를 생성하는 방법에 대해 설명한다. IV장에서는 이동로봇의 위치추정에 대한 실험 결과를 통해 제안한 실내 위치 확인 시스템의 신뢰성에 대한 검증을 수행한 후, 마지막 V장에서 전체 내용에 대한 결론을 맺는다.

II. 조감도 영상의 생성

본 논문에서는 전체 환경에 대한 모니터링을 용이하게 하기 위하여, 캘리브레이션된 카메라 네트워크로부터 촬영된 각 영상 정보는 사용자가 정의한 가상의 카메라로부터 촬영된 지역 조감도 영상 정보로 변환된 후 전역 조감도 영상 정보로 통합된다. 그림 1은 전역 조감도 영상의 생성을 위한 전체 과정을 나타낸다. 본 장에서는 3대의 카메라로 구성된 센서 네트워크로부터의 영상 정보를 예로써 설명한다. 우선, 실제 카메라 영상 내의 바닥 면 검출을 수행한 후, 캘리브레이션된 카메라의 외부 파라미터 즉, 카메라의 6자유도 상태 변수와 주어진 3차원 지도 정보를 이용하여 예측되는 거리 영상을 생성한다. 다음으로 각각의 생성된 바닥 면이 검출된 영상은 역투영 변환(IPM: Inverse Perspective Mapping)을 통하여 지역 조감도 영상으로 변환된 후, 하나의 전역 조감도 영상으로 통합된다. 여기서 주목할 특징은, 일반적인 2차원 영상 정보만으로는 3차원 정보의 제공이 불가능하지만, 3차원 지도 정보 기반의 캘리브레이션이 완료된 분산 카메라 네트워크 시스템을 활용함으로써 3차원 정보를 제공 가능한 거리 영상과 RGB 컬러 영상의 획득이 동시에 가능하다는 점이다. 3차원 지도 정보는 정적인 모델이기 때문이 이를 통해 생성된 거리 영상은 환경의 동적 정보를 실시간으로 반영할 수 없지만, RGB 컬러 영상은 실시간으로 동적인 정보가 갱신된다. 따라서 제안한 시스템은 실내 환경의 바닥 면이 평평하다는 가정 하에 각기 다른 시점으로부터 촬영된 복수의 RGB 컬러 영상에서 중복 검출된 이동 물체의 영역을 통합함으로써 검출된 이동 물체에 대한 신뢰성 있는 실내 위치 신호의 생성이 가능하다.

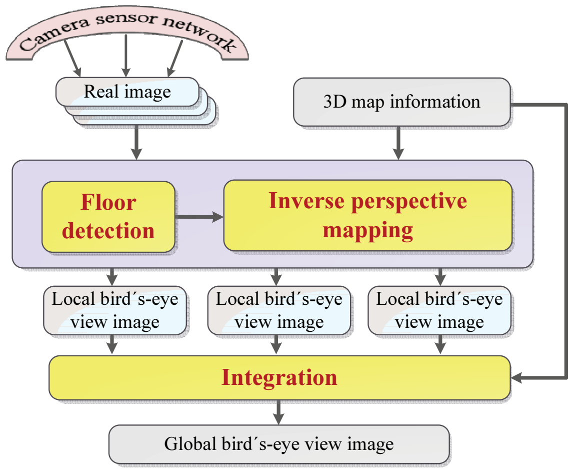

그림 1. 전역 조감도 영상 생성을 위한 전체 과정.
Fig. 1. Overall process to generate global bird's-eye image.

1. 바닥 면 검출

본 절에서는 카메라 네트워크로부터 획득한 각각의 영상 정보로부터 바닥 면을 검출하기 위한 과정에 대하여 설명한다. 우선 캘리브레이션이 완료된 카메라의 6자유도 상태 변수 $\mathbf{w} = [x_c\ y_c\ z_c\ \psi_c\ \theta_c\ \varphi_c]^T$와 주어진 3차원 지도 정보로부터 높이 영상 $\boldsymbol{I}^H_{R(\mathbf{w})}$을 생성한다. 다음으로 각각의 카메라 영상 정보 $\boldsymbol{I}_{R(\mathbf{w})}$ (그림 2(a))는 다음 식과 같이 바닥 면 영상 $\boldsymbol{I}^F_{R(\mathbf{w})}$ (그림 2(d))으로 변환된다. 이때 생성된 높이 영상 $\boldsymbol{I}^H_{R(\mathbf{w})}$ (그림 2(c))정보는 바닥 면의 높이를 구분해내는 경계값으로써 활용된다.

$$\boldsymbol{I}^F_{R(\mathbf{w}^{(k)})}(u,v) = \begin{cases} \boldsymbol{I}_{R(\mathbf{w}^{(k)})}(u,v), & \boldsymbol{I}^H_{R(\mathbf{w}^{(k)})}(u,v) < 0+\varepsilon \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

여기서 k는 카메라의 인덱스를 의미하며, ε는 각 카메라의 캘리브레이션 오차 등을 고려하기 위한 상수이다. 따라서 식 (1)은 실내 환경의 바닥 면의 높이가 0이라 가정할 때 카메라 영상 정보 $\boldsymbol{I}_{R(\mathbf{w})}$ 에서 바닥 면 이외에 해당하는 영역을 제거한다. 실내 환경에서의 동적 물체는 일반적으로 바닥 면 위를 주행한다는 점에서, 위와 같이 동적 물체의 주행 영역과 무관한 바닥 면 이외의 영역을 제거하는 것은 타당하다.

실제 카메라 영상에서 바닥 면을 검출하는 자세한 과정은 다음과 같다. 첫 번째로 그림 2(b)와 같이 각 카메라의 6자유도 상태 변수와 3차원 지도 정보로부터 각 카메라에 대응하는 거리 영상 $\boldsymbol{I}^D_{R(\mathbf{w})}$을 생성한다. 이와 같이 일반적인 광학 카메라만으로는 거리 정보를 직접 획득할 수 없지만, 캘리브레이션이 완료된 분산 카메라 네트워크 시스템과 3차원 지도

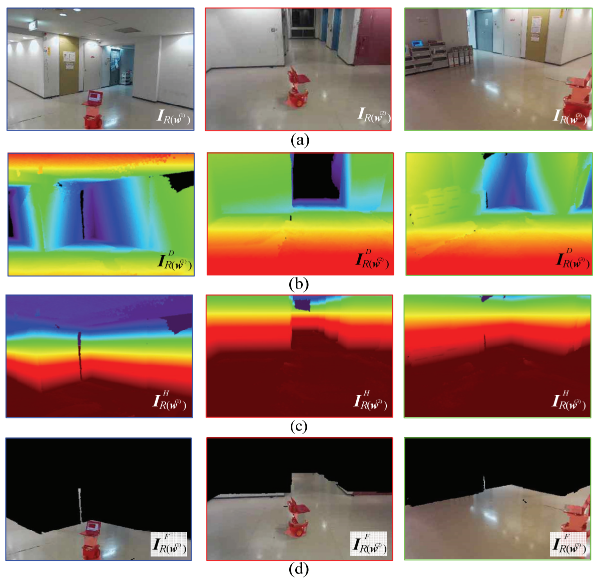

그림 2. 바닥 면 검출 과정: (a) 카메라 센서 네트워크로부터의 실제 카메라 영상 $\boldsymbol{I}_{R(\mathbf{w})}$, (b) 3차원 지도 정보로부터 예측된 거리 영상 $\boldsymbol{I}^D_{R(\mathbf{w})}$, (c) 높이 영상 $\boldsymbol{I}^H_{R(\mathbf{w})}$, (d) 바닥 면 이외의 영역이 제거된 바닥 면 영상 $\boldsymbol{I}^F_{R(\mathbf{w})}$.
Fig. 2. Floor detection process: (a) real camera images $\boldsymbol{I}_{R(\mathbf{w})}$ captured from camera sensor networks, (b) predicted depth images $\boldsymbol{I}^D_{R(\mathbf{w})}$ generated by 3D map information, (c) height images $\boldsymbol{I}^H_{R(\mathbf{w})}$, and (d) floor images $\boldsymbol{I}^F_{R(\mathbf{w})}$ where regions other than the floor have been removed.

정보를 활용함으로써 RGB-D 센서와 유사한 기능을 구현 가능하다. 다음으로 실제 카메라 영상에 대응하는 높이 정보를 각 화소값에 포함하고 있는 높이 영상 $\boldsymbol{I}^H_{R(\mathbf{w})}$ (그림 2(c))을 다음 식과 같이 생성한다.

$$\boldsymbol{I}^H_{R(\mathbf{w})}(u,v) = z_c - {}^C x\sin\theta_c + {}^C y\cos\theta_c\sin\psi_c + {}^C z\cos\theta_c\sin\psi_c \quad (2)$$

$${}^C x = \boldsymbol{I}^D_{R(\mathbf{w})}(u,v) \quad (3)$$

$${}^C y = \frac{1}{f_u}(u-c_u)\boldsymbol{I}^D_{R(\mathbf{w})}(u,v) \quad (4)$$

$${}^C z = \frac{1}{f_v}(c_v-v)\boldsymbol{I}^D_{R(\mathbf{w})}(u,v) \quad (5)$$

식 (3)~(5)는 카메라의 초점거리 (f_u, f_v) 및 주점 (c_u, c_v) 정보를 활용하여 거리 영상 정보 $\boldsymbol{I}^D_{R(\mathbf{w})}$의 각 2차원 화소값 ($u$, v)를 지역 카메라 좌표계 기준의 3차원 좌표값 (${}^Cx$, ${}^Cy$, ${}^Cz$)로 변환하는 역할을 수행한다. 여기서 윗첨자 C는 지역 카메라 좌표계를 의미하며 $\mathbf{w} = [x_c\ y_c\ z_c\ \psi_c\ \theta_c\ \varphi_c]^{\mathrm{T}}$는 전역 좌표계 $\{\boldsymbol{W}\}$ 기준의 카메라에 대한 6자유도 상태 변수를 의미한다. 식 (2)는 카메라 좌표계 $\{\boldsymbol{C}\}$ 기준의 환경에 대한 3차원 정보 (${}^Cx$, ${}^Cy$, ${}^Cz$)를 전역 좌표계 $\{\boldsymbol{W}\}$ 기준의 높이 정보로 변환하여 높이 영상 $\boldsymbol{I}^H_{R(\mathbf{w})}$을 생성하는 과정을 나타낸다. 마지막으로 식 (1)을 통해 각 카메라 영상 $\boldsymbol{I}_{R(\mathbf{w})}$에 대한 바닥면 영상 $\boldsymbol{I}^F_{R(\mathbf{w})}$ (그림 2(d))을 생성할 수 있다. 본 논문에서는 카메라의 광축을 지역 카메라 좌표계 $\{\boldsymbol{C}\}$의 x축으로 정의하였으며, 적용한 각 좌표계의 정의는 그림 3에 나타나 있다.

2. 역투영 변환

카메라 네트워크에 대한 각각의 지역 조감도 영상을 생성하기 위하여 본 논문에서는 역투영 변환을 적용한다. 역투영 변환은 바닥면이 평평하다는 가정 하여 원근 효과를 제거해줌으로써 환경에 대한 모니터링을 용이하게 해준다[12]. 이동로봇 등의 동적 물체는 일반적인 실내 환경에서 바닥 면을 주행한다고 가정할 수 있기 때문에 본 논문에서는 생성된 각각의 바닥 면 영상 $\boldsymbol{I}^F_{R(\mathbf{w})}$에 대해 역투영 변환을 적용한다.

조감도 영상은 기하학적 투시 변환(geometrical perspective transformation) 원리에 기반한 영상 처리에 의해 생성 가능하다. 실제 카메라 영상 정보를 가상의 공중 수직 시점에서 바라본 조감도 영상으로 변환하기 위한 원리는 다음과 같다.

그림 3은 전역 좌표계 $\{\boldsymbol{W}\}$, 카메라 영상 좌표계 $\{\boldsymbol{R}\}$, 그리고 가상의 조감도 영상 좌표계 $\{\boldsymbol{B}\}$ 사이의 관계를 표현하며, 각각의 실제 카메라와 사용자가 정의한 가상의 카메라 사이의 투시 변환(perspective transformation)을 나타내고 있다. 각각의 관계는 변환 행렬을 사용하여 표현 가능하다. 우선 ${}^W\boldsymbol{x} = [{}^Wx\ {}^Wy\ {}^Wz\ 1]^{\mathrm{T}}$와 ${}^R\boldsymbol{u} = [{}^Ru\ {}^Rv\ s]^{\mathrm{T}}$ 사이의 관계식은 다음과 같이 표현된다.

$${}^R\boldsymbol{u} = {}^R\boldsymbol{Q}_W\, {}^W\boldsymbol{x}$$
$$\begin{bmatrix} {}^Ru \\ {}^Rv \\ s \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \\ h_{31} & h_{32} & h_{33} & h_{34} \end{bmatrix} \begin{bmatrix} {}^Wx \\ {}^Wy \\ {}^Wz \\ 1 \end{bmatrix} \quad (6)$$

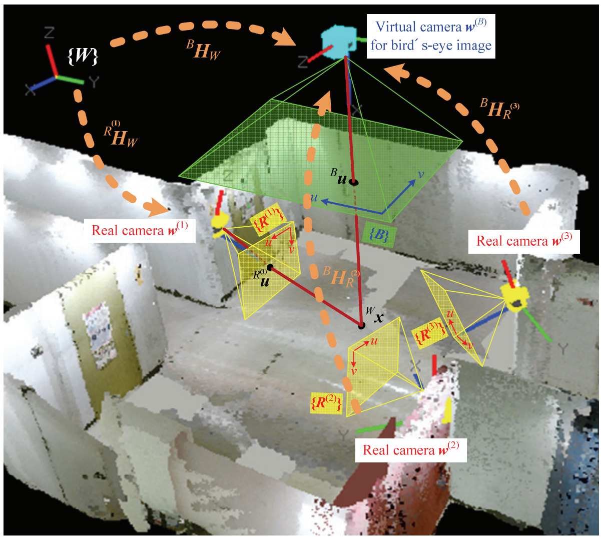

그림 3. 바닥 면, 실제 카메라 영상, 가상의 카메라로부터의 조감도 영상 사이의 투시 변환.

Fig. 3. Perspective transformations between floor plane, real camera images, and bird's-eye image for virtual camera.

여기서 s는 스케일 팩터를 의미한다. 식 (6)은 바닥 면의 높이가 0이라는 가정(${}^Wz$ = 0) 하에 다음과 같이 간략화시킬 수 있다.

$${}^R\boldsymbol{u} = {}^R\boldsymbol{H}_W\, {}^W\tilde{\boldsymbol{x}}$$
$$\begin{bmatrix} {}^Ru \\ {}^Rv \\ s \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{14} \\ h_{21} & h_{22} & h_{24} \\ h_{31} & h_{32} & h_{34} \end{bmatrix} \begin{bmatrix} {}^Wx \\ {}^Wy \\ 1 \end{bmatrix} \quad (7)$$

여기서 ${}^W\tilde{\boldsymbol{x}} = [{}^Wx\ {}^Wy\ 1]^{\mathrm{T}}$는 전역 좌표계 기준의 z요소를 제외한 위치 벡터를 의미한다. 기하학 분야에서 호모그래피 행렬(homography matrix)로 불리는 ${}^R\boldsymbol{H}_W$는 하나의 평면을 다른 평면으로 사상시키는 투영 변환을 수행한다.

동일한 원리로, ${}^W\tilde{\boldsymbol{x}} = [{}^Wx\ {}^Wy\ 1]^{\mathrm{T}}$와 ${}^B\boldsymbol{u} = [{}^Bu\ {}^Bv\ s]^{\mathrm{T}}$사이의 관계식은 다음과 같이 정의된다.

$${}^B\boldsymbol{u} = {}^B\boldsymbol{H}_W\, {}^W\tilde{\boldsymbol{x}} \quad (8)$$

여기서 ${}^B\boldsymbol{H}_W$는 투영 면의 높이가 0이라고 가정할 때, 전역 좌표계 $\{\boldsymbol{W}\}$와 조감도 영상 좌표계 $\{\boldsymbol{B}\}$ 사이의 호모그래피 행렬을 의미한다.

마지막으로 식 (8)을 식 (7)에 대입함으로써 실제 카메라 영상 좌표계 $\{\boldsymbol{R}\}$와 조감도 영상 좌표계 $\{\boldsymbol{B}\}$ 사이의 관계를 다음과 같이 산출할 수 있다.

$${}^R\boldsymbol{u} = {}^R\boldsymbol{H}_W \left({}^B\boldsymbol{H}_W\right)^{-1} {}^B\boldsymbol{u} = {}^R\boldsymbol{H}_W\, {}^W\boldsymbol{H}_B\, {}^B\boldsymbol{u} \quad (9)$$

$${}^B\boldsymbol{u} = \left({}^R\boldsymbol{H}_B\right)^{-1} {}^R\boldsymbol{u} = {}^B\boldsymbol{H}_R\, {}^R\boldsymbol{u} \quad (10)$$

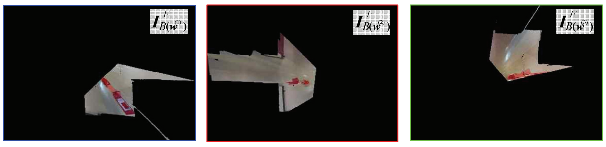

그림 4. 각 바닥면 영상 $\boldsymbol{I}^F_{R(\mathbf{w})}$에 대응하는 지역 조감도 영상 $\boldsymbol{I}^F_{B(\mathbf{w})}$.

Fig. 4. Local bird's-eye images $\boldsymbol{I}^F_{B(\mathbf{w})}$ corresponding to floor images $\boldsymbol{I}^F_{R(\mathbf{w})}$.

식 (10)에 나타난 바와 같이, 호모그래피 행렬 ${}^B\boldsymbol{H}_R$은 실제 카메라 영상을 지역 조감도 영상으로 변환하는 역투영 변환을 수행한다. 여기서 3×3 호모그래피 행렬 ${}^B\boldsymbol{H}_R$은 실제 카메라 영상 $\{\boldsymbol{R}\}$과 조감도 영상 $\{\boldsymbol{B}\}$에서 공통으로 관측되는 최소 4개 이상의 점을 임의로 선택 후, 환경 내의 모든 점유 좌표 정보를 포함하고 있는 3차원 지도 정보를 활용하여 캘리브레이션이 가능하다.

그림 4는 캘리브레이션된 호모그래피 행렬 ${}^B\boldsymbol{H}_R$을 각 바닥면 영상 $\boldsymbol{I}^F_{R(\mathbf{w})}$에 적용하여 생성한 지역 조감도 영상 $\boldsymbol{I}^F_{B(\mathbf{w})}$을 나타낸다. 여기서 조감도 영상의 촬영 위치에 해당하는가상의 카메라에 대한 상태변수 $\mathbf{w}^{(B)}$는 전체 환경 중심의 높은 위치에서 바닥 면을 수직으로 바라보도록 설정하였다.

3. 지역 조감도 영상의 통합

역투영 변환을 적용하여 생성한 각각의 지역 조감도 영상 $\boldsymbol{I}^F_{B(\mathbf{w})}$은 전역 조감도 영상 ${}^G\boldsymbol{I}^F_{B(\mathbf{w})}$으로 다음 식을 적용하여 통합된다.

$$ {}^G\boldsymbol{I}^F_B(u,v) = {}^G\boldsymbol{I}^F_{B(\mathbf{w}^{(k^*)})}({}^Bu^*, {}^Bv^*) \tag{11} $$

$$ \begin{bmatrix} {}^Bu^* \\ {}^Bv^* \\ 1 \end{bmatrix} = {}^B\boldsymbol{H}_{R^{(k^*)}} \begin{bmatrix} {}^Ru^* \\ {}^Rv^* \\ 1 \end{bmatrix} \tag{12} $$

$$ (k^*, {}^Ru^*, {}^Rv^*) = \arg\min_{k,u,v}[\boldsymbol{I}^D_{R(\mathbf{w}^{(k)})}(u,v)] \tag{13} $$

식 (11)~(13)은 동일한 지점이 복수의 카메라 영상에 의해 촬영되었을 경우, 최단 거리에 해당하는 카메라 영상의 화소값이 통합에 사용된다는 것을 의미한다. 여기서 k는 카메라의 인덱스(본 논문의 예시에서는 1~3)를 의미한다. 여기서 최단거리를 계산하기 위해 각 카메라에서 예측되는 거리영상 $\boldsymbol{I}^D_{R(\mathbf{w})}$을 다시 활용한다. 카메라에서 물리적인 거리가 짧은 지점일수록 조감도 영상에서의 왜곡이 작아지기 때문에 위와 같은 통합 과정은 타당하다.

그림 5는 식 (11)~(13)을 적용하여 통합한 전역 조감도 영상 ${}^G\boldsymbol{I}^F_{B(\mathbf{w})}$을 나타낸다. 이와 같은 전역 조감도 영상을 활용함으로써, 카메라 네트워크 시스템으로부터 촬영되는 전체 환경의 바닥 면에 대한 직관적인 모니터링이 가능하며, 이동로봇 등의 동적 물체에 대한 위치 신호의 생성을 매우 용이하게 해준다.

그림 5의 통합된 조감도 영상은 경계선 등이 고르지 못하고 오차를 포함하고 있다. 이는 호모그래피 행렬 ${}^B\boldsymbol{H}_R$에 대한 캘리브레이션 오차로 인한 현상이며, 카메라의 외부 파라미터 즉, 각 카메라의 6자유도 상태 변수 $\mathbf{w} = [x_c\ y_c\ z_c\ \psi_c\ \theta_c\ \varphi_c]^T$에 대한 캘리브레이션 오차에서 기인한다. 본 논문에서는 본 연구팀이 선행 연구로써 제안한 3차원 지도 정보 기반의 기법을 통해 각각의 카메라 파라미터 $\mathbf{w}$에 대한 자동 캘리브레이션을 수행하였다[2]. 즉, 3차원 지도 정보의 정확도는 통합된 조감도 영상의 정확도에 영향을 끼치며, 이는 결국 실내 위치 신호에 대한 오차 발생의 요인이 된다. 따라서 본 논문에서는 최종적인 실내 위치 신호에 대한 오차를 확률 밀도 함수를 사용한 표현을 통해 고려한다. 이에 대하여 다음 장에서 자세히 설명한다.

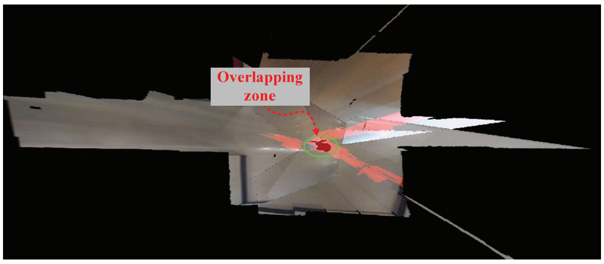

그림 5. 통합된 전역 조감도 영상 ${}^G\boldsymbol{I}^F_{B(\mathbf{w})}$.

Fig. 5. Integrated global bird's-eye images ${}^G\boldsymbol{I}^F_{B(\mathbf{w})}$.

III. 위치 신호의 생성

본 장에서는 카메라 네트워트로부터의 영상 정보로부터 생성된 전역 조감도 영상을 활용한 위치 신호의 생성법에 대하여 설명한다. 본 논문에서는 영상 정보 내의 이동 물체를 검출하기 위하여 혼합 정규 분포 모델에 기반한 배경차분법을 사용한다[13]. 이동 물체가 정지 상태에 있거나 지속으로 이동하는 경우 배경차분법을 이용한 검출에는 한계가 있다. 또한, 환경 내에 복수의 이동 물체가 존재할 경우, 각 이동 물체간의 독립된 실내 위치 신호의 수신이 필요하다. 이와 같은 경우, 실내 위치 확인 시스템의 수신기로써 이동 물체의 몸체에 표식을 부착하여 검출하는 방법을 고려할 수 있다.

그림 2(a), 2(d), 4, 5에서의 붉은색 음영 영역은 배경차분법을 적용함으로써 검출된 이동 물체의 영역을 의미하며, 그림 5의 통합된 전역 조감도 영상에서의 진한 붉은색 음영 영역은 복수의 카메라 영상에서 동시에 검출된 이동 물체의 영역을 의미한다. 본 연구에서는 이와 같은 중복 영역을 활용하여 왜곡을 제거한다. 즉, 하나의 카메라로부터 검출된 이동 물체는 그림 4와 같이 조감도 영상에서 크게 왜곡되어 표현되지만, 각기 다른 시점에서 중복 검출된 이동 물체의 영역을 그림 5와 같이 통합함으로써 왜곡이 제거되고 정확한 위치의 검출이 가능하다.

그림 6에 이와 같이 검출된 이동 물체에 대한 영상 내의 왜곡을 제거하는 원리가 나타나 있다. 스테레오 카메라와 달리 단안 카메라는 촬영된 물체까지의 거리를 직접적으로 측정할 수 없다. 그러나 각기 다른 시점의 복수의 카메라로부터 동일한 물체가 검출되었을 경우, 물체까지의 거리에 대한 불확실성(존재 영역에 대한 불확실성)을 크게 감소시킬 수 있다. 즉, 본 논문에서 제안한 방법은 복잡한 스테레오 계측과 같은 계산을 하지 않더라도 실내 환경의 바닥 면의 높이가 일정하다는 가정 하에 매우 간단한 원리로 위치 신호의 생성이 가능하다. 이러한 원리는 잘 알려진 단안 카메라를

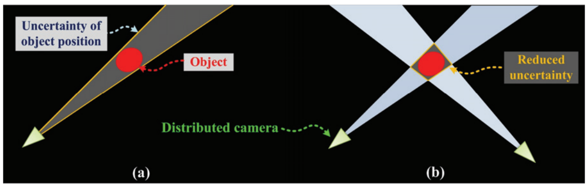

그림 6. 각기 다른 시점의 복수의 카메라 영상으로부터 왜곡을 제거하는 원리: (a) 하나의 카메라로부터 검출된 물체, (b) 2대의 카메라로부터 검출된 물체.

Fig. 6. Principle of distortion removal based on multiple camera observations from different viewpoints: (a) object detected from one camera, (b) object detected from two cameras.

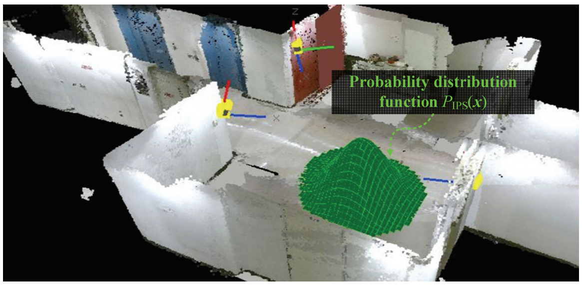

그림 7. 제안한 실내 위치 확인 시스템으로부터 검출된 이동 물체에 대한 존재 확률 분포를 의미하는 확률 밀도 함수 $P_{\text{IPS}}(\boldsymbol{x})$.

Fig. 7. Probability distribution function $P_{\text{IPS}}(\boldsymbol{x})$ for the existence of moving object detected by proposed IPS system.

이용한 SLAM과 밀접한 관련이 있다[14]. 단안 카메라를 이용한 SLAM 기법은 매우 복잡한 확률적인 계산을 필요로 하는데에 반하여, 본 논문에서 제안하는 전역 조감도 영상에 기반한 방법은 검출된 물체에 대한 왜곡을 매우 직관적으로 제거 가능하다.

제안한 방법을 통한 실내 위치 신호의 신뢰성 있는 성능을 유지하기 위해서는 이동 물체가 항상, 떨어진 2대 이상의 카메라로부터 동시에 관측될 필요가 있다. 따라서 제한된 카메라 대수로 대상 실내 환경의 전 영역을 커버하기 위해서는 최적의 카메라 배치를 사전에 고려할 필요가 있다. 이는 계산기하학 분야의 art gallery problem으로 불리는 고전적인 문제로서 연구되었으며, 이에 대한 자세한 내용은 참고문헌에 설명되어 있다[15]. 즉, 카메라 네트워크의 배치가 환경의 전 영역을 적절하게 커버한다는 가정 하에, 제안한 방법은 기존의 실내 위치 확인 시스템에서 자주 발생하는 신호의 난반사 등의 문제가 존재하지 않으므로 매우 안정적인 위치 신호의 수신이 가능하며, 이를 실내 이동로봇의 전역 위치추정 문제에 적용할 경우 성능을 크게 향상시킬 수 있다.

제안한 방법을 통해 검출한 이동 물체는 조감도 영상 좌표계 $\{\boldsymbol{B}\}$ 기반의 화소 좌표값 ${}^{B}\boldsymbol{u}_{\text{IPS}}$로 표현된다. 따라서 다음과 같이 호모그래피 행렬 ${}^{W}\boldsymbol{H}_{B}$을 통해 실제의 좌표값 ${}^{W}\tilde{\boldsymbol{x}}_{\text{IPS}}$로 변환할 필요가 있다.

$$\begin{aligned} {}^{W}\tilde{\boldsymbol{x}}_{\text{IPS}} &= \left({}^{B}\boldsymbol{H}_{W}\right)^{-1} {}^{B}\boldsymbol{u}_{\text{IPS}} \\ &= {}^{W}\boldsymbol{H}_{B}\, {}^{B}\boldsymbol{u}_{\text{IPS}} \end{aligned} \tag{14}$$

여기서 호모그래피 행렬 ${}^{B}\boldsymbol{H}_{W}$는 조감도 영상을 생성하기 위해 정의한 가상의 카메라에 대한 6자유도 상태를 정의하는 외부 파라미터 행렬과 내부 파라미터 행렬의 곱으로 정의되며, 실내 환경의 바닥 면의 높이가 0이라는 가정 하에 3열 벡터의 요소는 모두 제외된다. 다음으로 검출된 이동 물체에 대한 위치 정보 ${}^{W}\tilde{\boldsymbol{x}}_{\text{IPS}}$는 다음 식과 같이 커널 밀도 추정을 통해 확률 밀도 함수로 표현되는 위치 신호 $P_{\text{IPS}}(\boldsymbol{x})$로 변환된다.

$$P_{\text{IPS}}(\mathbf{x}) = \frac{1}{Nh^{D}} \sum_{i=1}^{N} \boldsymbol{K}\left(\frac{\boldsymbol{x}-\boldsymbol{x}_{\text{IPS}}^{(i)}}{h}\right) \tag{15}$$

$$\boldsymbol{K}(\boldsymbol{x}) = \frac{1}{(2\pi)^{D/2}} \exp\left(-\frac{1}{2}\boldsymbol{x}^{\text{T}}\boldsymbol{x}\right) \tag{16}$$

식 (15)에서 $\boldsymbol{x}^{(i)}_{\text{IPS}} = [x^{(i)}_{\text{IPS}}\, y^{(i)}_{\text{IPS}}]^{\text{T}}$와 N은 각각 검출된 이동 물체에 대한 영역의 좌표와 통합된 조감도 영상에서 검출된 화소 ${}^{B}\boldsymbol{u}_{\text{IPS}}$의 개수를 의미한다. 여기서 전역 좌표계를 의미하는 윗첨자 W는 생략하였다. h와 D는 각각 평활 파라미터와 좌표값의 차원을 의미한다(본 연구의 경우는 2). 식 (16)은 일반적인 평균 $\boldsymbol{O}$, 공분산 $\boldsymbol{I}$의 가우시안 커널 함수를 의미하며, 카메라 캘리브레이션 및 배경차분법 등에서 발생할 수 있는 노이즈를 고려하기 위해 적용한다. 위와 같은 과정을 통해 최종적으로 생성된 이동 물체의 위치 신호 $P_{\text{IPS}}(\boldsymbol{x})$가 그림 7에 나타나 있다. 생성된 위치 신호 $P_{\text{IPS}}(\boldsymbol{x})$는 카메라 네트워크를 통해 검출된 이동 물체에 대한 존재 확률 분포로서 해석 가능하다.

위와 같은 방법을 통해 생성한 이동 로봇에 대한 실내 위치 신호는, 위치추정의 수행에 있어서 센서 정보에 대한 관측 모델로써 활용할 수 있다. 이에 대하여 다음장의 실험 내용과 함께 자세히 설명한다.

IV. 이동로봇의 위치추정 실험 결과

본 장에서는 본 논문에서 제안한 실내 위치 신호를 활용한 이동로봇의 위치추정에 대한 실험 결과에 대해 설명한다. 제안한 실내 위치 신호를 활용함으로써 이동로봇의 전역 위치추정에 대한 성능을 비약적으로 향상시킬 수 있다. 본 논문에서는 위치추정을 수행하기 위한 알고리즘으로써 파티클 필터에 기반한 몬테카를로 위치추정(MCL: Monte Carlo Localization)을 사용하였다. 몬테카를로 위치추정에 대한 더 자세한 내용은 참고문헌에 설명되어 있다[16].

생성한 실내 위치 신호를 몬테카를로 위치추정에 적용하기 위해서는 실내 위치 신호에 대한 관측 모델을 정의할 필요가 있다. 여기서 식 (15)의 커널 밀도 추정을 통해 생성한 확률 밀도 함수 $P_{\text{IPS}}(\boldsymbol{x})$를 실내 위치 신호에 대한 관측 모델 $p({}^{\text{IPS}}\boldsymbol{z} \mid \boldsymbol{x}_r)$로서 다음 식과 같이 베이즈 룰에 의해 직접 사용할 수 있다.

$$\begin{aligned} p({}^{\text{IPS}}\boldsymbol{z} \mid \boldsymbol{x}_r) &= \frac{p(\boldsymbol{x}_r \mid {}^{\text{IPS}}\boldsymbol{z})\, p({}^{\text{IPS}}\boldsymbol{z})}{p(\boldsymbol{x}_r)} \\ &\approx \alpha P_{\text{IPS}}(\breve{\boldsymbol{x}}_r) \end{aligned} \tag{17}$$

여기서 $\boldsymbol{x}_r = [x_r\, y_r\, \theta_r]^{\text{T}}$과 $\breve{\boldsymbol{x}}_r = [x_r\, y_r]^{\text{T}}$는 각각 이동로봇의 3자유도 상태 변수와 방향 정보를 제외한 위치 벡터를 의미한다. 실내 위치 신호 ${}^{\text{IPS}}\boldsymbol{z}$는 로봇의 상태 변수와 관측 정보에 대한

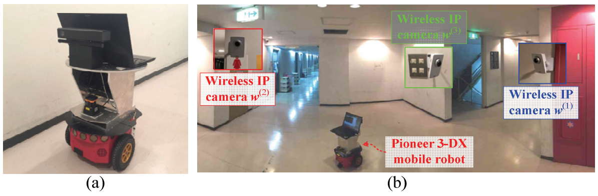

그림 8. 실험 장치: (a) Pioneer 3-DX 이동 로봇, (b) 카메라 센서 네트워크를 구축한 실험 환경.

Fig. 8. Experimental setup: (a) Pioneer 3-DX mobile robot and (b) experimental environment with camera sensor network.

어떠한 사전 정보도 없이 관측된 값이기 때문에, 식 (17)에서 로봇의 상태 변수에 x_r에 대한 전확률 $p(x_r)$과 실내 위치 신호에 대한 사전확률 $p(^{IPS}z)$은 균일 분포로 간주할 수 있다. 따라서 이 항들은 소거할 수 있으며, 결국 $P_{IPS}(\breve{x}_r)$ 는 관측 정보 ^{IPS}z가 주어졌을 때 이동로봇의 상태 x_r에 대한 확률 분포 $p(x_r|^{IPS}z)$로서 직접 근사할 수 있다. 여기서, 실내 위치 신호에 대한 불확실성은 식 (15)에서의 평활 파라미터 h에 의해 조절 가능하다. α는 확률 밀도 함수를 몬테카를로 위치추정에 사용하기 위한 관측 모델로 변환하기 위한 상수이다. 즉, α를 높은 수치로 설정할수록 실내 위치 신호가 몬테 카를로 위치추정 과정에 큰 영향을 주게 된다. 몬테카를로 위치추정에서, 각 파티클에 대한 확률값은 식 (17)의 근사된 관측 모델 $\alpha P_{IPS}(\breve{x}_r)$ 에 의해 추가적으로 갱신된다.

제안한 실내 위치 확인 시스템을 이동로봇의 전역 위치 추정 과정에 적용하였을 시의 효과를 검증하기 위해, 그림 8(a)와 같은 Pioneer-3DX 이동로봇을 사용하여 실제 환경 내에서 비교 실험을 수행하였다. 로봇은 그림 8(b)와 같은 환경 내를 주행하였다. 카메라 네트워크 시스템은 그림 8(b)와 같이 3대의 무선 IP 카메라 AXIS M1004-W를 사용하여 구축하였으며, 각 카메라에 대한 캘리브레이션은 본 연구팀이 기존에 제안한 방법을 통해 사전에 수행되었다[2].

그림 9는 몬테카를로 위치추정에 사용된 파티클들의 수렴

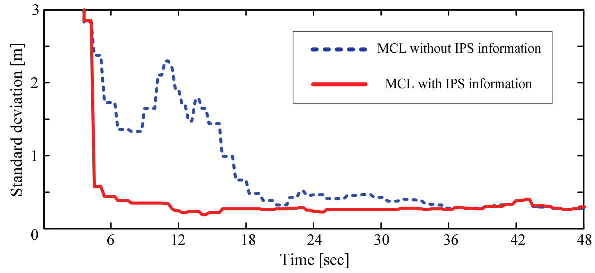

그림 9. 몬테카를로 위치추정 과정에 따른 파티클의 위치 분포에 대한 표준편차.

Fig. 9. Standard deviation of particles' positions as function of MCL process.

경향을 파티클들의 위치 분포에 대한 표준 편차로서 나타낸 그래프이다. 전역 위치추정에 사용한 파티클의 개수는 환경 내의 파티클의 위치 분포에 대한 분산값에 따라 최소 1,000개에서 최대 10,000개(초기 파티클의 개수)까지 적절히 조절하였다. 각 파티클의 확률값 갱신은 이동로봇의 병진 이동량이 0.15 m 혹은 회전 이동량이 5° 이상 누적되는 시점에 수행하였다. 또한 각 센서 데이터의 획득 주기는 매 300 ms 마다로 동기화시켰다. 본 논문에서 제안한 실내 위치 신호를 각 파티클의 확률 갱신에 적용한 경우 알고리즘의 평균 수행 시간은 매 주기 당 107 ms, 적용하지 않은 경우는 21 ms 이었다(쿼드코어 4.0 GHz CPU). 실내 위치 신호를 생성하기 위해서는 본 논문에서 설명한 화상 처리 등의 과정이 요구되므로 추가적인 수행 시간이 소요되지만 센서 데이터의 획득 주기 이내에 처리가 가능하므로 알고리즘의 실시간성은 보장된다.

제안한 실내 위치 확인 시스템으로부터의 위치 신호 정보를 활용한 경우, 추가적인 파티클들의 확률 갱신이 수행되기 때문에 파티클들이 실제 로봇의 위치 근처에 수렴하는데에 훨씬 짧은 시간이 소요되었다.

그림 10은 제안한 실내 위치 확인 시스템으로부터의 위치 신호를 활용한 이동로봇의 몬테카를로 위치추정 과정을 나

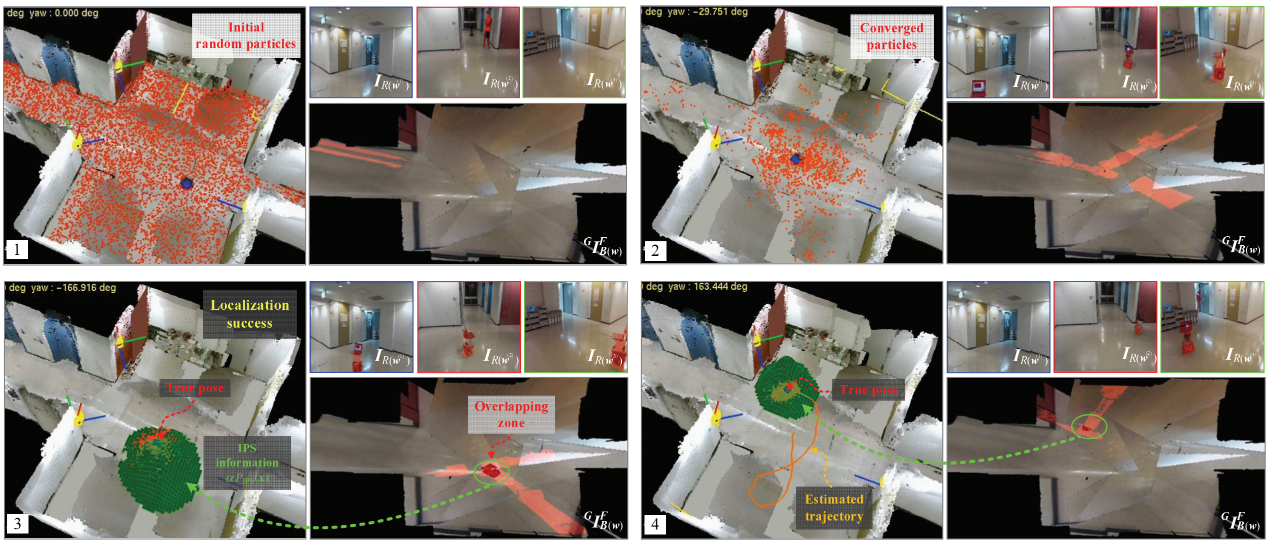

그림 10. 전역 위치추정 실험 결과(실내 위치 확인 시스템 신호를 추가적인 관측 모델로서 적용한 몬테카를로 위치추정).

Fig. 10. Experimental result for global localization (MCL with additional measurements from the IPS information).

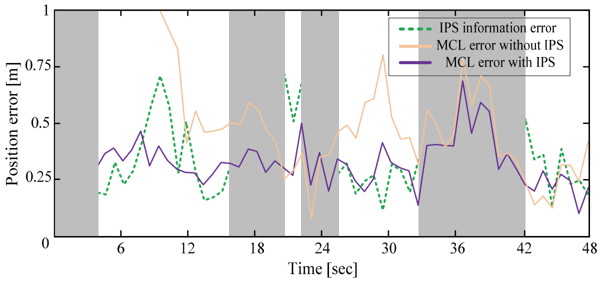

그림 11. 실내 위치 신호 및 몬테카를로 위치추정 과정에 따른 이동로봇의 위치 오차.

Fig. 11. Position error for IPS information and MCL results of mobile robot.

타낸다. 모든 파티클이 이동로봇의 실제 위치 주변에 표준편차 0.5 m 이내로 수렴한 시점 즉, 그림 9의 종축 값이 0.5 m 이하가 되는 시점을 전역 위치추정의 성공 시점으로 판단하였다. 제안한 실내 위치 확인 시스템으로부터의 위치 신호가 수신된 경우 매우 빠르게 파티클이 실제 로봇의 위치로 수렴하여, 성공적으로 전역 위치추정이 수행됨을 그림 10의 실험 결과로부터 알 수 있다.

그림 11의 점선은 수신된 실내 위치 신호 자체의 위치 오차를 나타내며 검출된 이동 물체의 좌표값 ${}^{W}\tilde{\boldsymbol{x}}_{\text{IPS}}$ 에 대한 무게 중심점을 기준으로 계산하였다. 여기서 음영 부분은 실내 위치 신호가 수신되지 않은 구간 즉, 2대 미만의 카메라에서만 이동로봇이 관측된 경우를 의미한다. 옅은 실선과 진한 실선은 각각 제안한 실내 위치 신호 정보를 적용하지 않은 경우와 적용한 경우에 대한 로봇의 위치추정 오차를 나타낸다. 대부분 0.5 m 이내의 오차를 가진 실내 위치 신호가 수신됨을 확인할 수 있으며, 오차는 II장에서 언급한 바와 같이 3차원 지도 정보의 정확도 및 카메라 캘리브레이션 과정에서 발생하는 오차에서 기인한다. 실내 위치 신호를 적용한 경우에 대한 이동로봇의 위치추정 결과는 적용하지 않은 경우에 비해 전반적으로 작은 오차를 유지하는 것을 알 수 있으며, 오도메트리 등에 대한 확률 모델이 동시에 고려되기 때문에 완전히 일치하지는 않지만, 실내 위치 신호의 거동을 비슷하게 따르는 것을 확인할 수 있다.

V. 결론

본 논문은 분산 카메라 센서 네트워크를 활용한 실내 위치 확인 시스템을 제안하였다. 카메라 네트워크로부터의 신뢰성 있는 실내 위치 신호의 생성을 위해 3차원 지도 정보를 사용하여 전역 조감도 영상을 생성하는 방법을 활용하였다. 카메라 네트워크로부터의 각 영상 정보를 조감도 영상으로 변환하기 위한 호모그래피 행렬의 캘리브레이션은 전체 환경의 모든 점유 영역 정보를 포함하고 있는 3차원 지도 정보를 통해 자동적으로 수행된다. 제안한 실내 위치 확인 시스템은 환경 내의 이동 물체에 대한 신뢰성있는 위치 정보를 제공한다. 따라서 이동로봇에 대한 위치 신호가 검출될 경우, 그에 대한 추가적인 관측 모델을 정의함으로써 전역 위치추정의 성능을 크게 향상시킬 수 있다.

REFERENCES

[1] J. B. Song and S. Y. Hwang, "Past and state-of-the-art SLAM technologies," *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 20, no. 3, pp. 372-379, Mar. 2014.

[2] Y. Ji, A. Yamashita, and H. Asama, "Automatic calibration of camera sensor networks based on 3D texture map information," *Robotics and Autonomous Systems*, 2016, doi:10.1016/j.robot. 2016.09.015.

[3] Y. Ji, A. Yamashita, and H. Asama, "Automatic calibration and trajectory reconstruction of mobile robot in camera sensor network," *Proc. of 11th Annual IEEE International Conference on Automation Science and Engineering (CASE2015)*, pp. 206-211, Aug. 2015.

[4] A. Rahimi, B. Dunagan, and T. Darrell, "Simultaneous calibration and tracking with a network of non-overlapping sensors," *Proc. of 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR2004)*, pp. I-287, Jul. 2004.

[5] D. Brscic and H. Hashimoto, "Model based robot localization using onboard and distributed laser range finders," *Proc. of 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2008)*, pp. 1154-1159, Sep. 2008.

[6] P. G. Jayasekara, H. Hashimoto, and T. Kubota, "Simultaneous localization assistance for mobile robot navigation in real, populated environment," *SICE Journal of Control, Measurement, and System Integration*, vol. 5, no. 6, pp. 349-358, Nov. 2012.

[7] S. H. Choi, G. J. Kim, Y. K, Kim, and J. M. Lee, "Outdoor positioning estimation of multi-GPS / INS integrated system by EKF/UPF filter conversion," *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 20, no. 12, pp. 1284-1289, Dec. 2014.

[8] B. H. Lee, S. H. Im, M. B. Heo, and G. I. Jee, "Curve-modeled lane detection based GPS lateral error correction enhancement," *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 21, no. 2, pp. 81-86, Feb. 2015.

[9] N. Chang, R. Rashidzadeh, and M. Ahmadi, "Robust indoor positioning using differential Wi-Fi access points," *IEEE Transactions on Consumer Electronics*, vol. 56, no. 3, pp. 1860-1867, Aug. 2010.

[10] K. Curran, E. Furey, T. Lunney, J. Santos, D. Woods, and A. McCaughey, "An evaluation of indoor location determination technologies," *Journal of Location Based Services*, vol. 5, no. 2, pp. 61-78, Apr. 2011.

[11] A. W. Reza and T. K. Geok, "Investigation of indoor location sensing via RFID reader network utilizing grid covering algorithm," *Wireless Personal Communications*, vol. 49, no. 1, pp. 67-80, May 2009.

[12] M. Bertozz, A. Broggi, and A. Fascioli, "Stereo inverse perspective mapping: theory and applications," *Image and Vision Computing*, vol. 16, no. 8, pp. 585-590, Jun. 1998.

[13] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," *Proc. of 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR1999)*, pp. 246-252, Jun. 1999.

[14] A. J. Davison, I. D. Reid, N. D. Molton, and O. Stasse, "MonoSLAM: real-time single camera SLAM," *IEEE*

Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 6, pp. 1052-1067, Jun. 2007.
[15] J. O'Rourke, *Art Gallery Theorems and Algorithms*, Oxford University Press, Oxford, 1987.
[16] F. Dellaert, D. Fox, W. Burgard, and S. Thrun, "Monte carlo localization for mobile robots," *Proc. of 1999 IEEE International Conference on Robotics and Automation (ICRA1999)*, pp. 1322-1328, May 1999.
[17] Y. Ji, A. Yamashita, and H. Asama, "Indoor positioning system using camera network based on 3D map information," *Proc. of 2016 31th ICROS Annual Conference (in Korean)*, pp. 263-264, Mar. 2016.

지 용 훈

2010년 경희대학교 기계공학과(공학사). 2010년 경희대학교 컴퓨터공학과(공학사). 2012년 고려대학교 메카트로닉스학과(공학석사). 2016년 도쿄대학 정밀공학전공(공학박사). 2013년 4월~2016년 3월 일본 대사관 추천 문부과학성 국비장학생(연구유학생). 2016년 4월~현재 도쿄대학 일본학술진흥회 특별연구원 (Overseas researcher under Postdoctoral Fellowship of Japan Society for the Promotion of Science). 관심분야는 환경지능화, 이동로봇의 주행, 무인차량 기술, 음향 카메라 수중 계측, 기계학습.

Atsushi Yamashita

1996년 도쿄대학 정밀기계공학과(공학사). 1998년 도쿄대학 정밀기계공학전공(공학석사). 2001년 도쿄대학 정밀기계공학전공(공학박사). 2001년~2011년 시즈오카대학 기계공학과 조수, 조교, 준교수. 2006년 11월~2007년 4월 캘리포니아공과대학 객원연구원. 2011년 10월~현재 도쿄대학 정밀공학전공 준교수. 관심분야는 로봇비전, 영상처리, 로봇의 운동제어.

Hajime Asama

1982년 도쿄대학 정밀기계공학과(공학사). 1984년 도쿄대학 정밀기계공학전공(공학석사). 1989년 도쿄대학 정밀기계공학(공학박사). 1986년~2002년 이화학연구소(RIKEN) 연구원보조, 연구원, 부주임연구원. 2002년 11월 도쿄대학 인공물공학연구센터 교수. 2009년 11월~현재 도쿄대학 정밀공학전공 교수. 관심분야는 환경지능화, 자율분산시스템, 이동지, 서비스공학.