

# GPU 가속화를 통한 이미지 특징점 기반 RGB-D 3차원 SLAM

## Image Feature-Based Real-Time RGB-D 3D SLAM with GPU Acceleration

이 동 화, 김 형 진, 명 현\*  
(Donghwa Lee<sup>1</sup>, Hyongjin Kim<sup>1,2</sup>, and Hyun Myung<sup>1,2</sup>)

<sup>1</sup>Dept. of Civil and Environmental Engineering, KAIST

<sup>2</sup>Robotics Program, KAIST

**Abstract:** This paper proposes an image feature-based real-time RGB-D (Red-Green-Blue Depth) 3D SLAM (Simultaneous Localization and Mapping) system. RGB-D data from Kinect style sensors contain a 2D image and per-pixel depth information. 6-DOF (Degree-of-Freedom) visual odometry is obtained through the 3D-RANSAC (RANDOM SAmple Consensus) algorithm with 2D image features and depth data. For speed up extraction of features, parallel computation is performed with GPU acceleration. After a feature manager detects a loop closure, a graph-based SLAM algorithm optimizes trajectory of the sensor and builds a 3D point cloud based map.

**Keywords:** SLAM, 3D SLAM, RGB-D, image feature

### I. 서론

지난 십 수년간 위치 인식과 지도 생성을 동시에 수행하는 SLAM (Simultaneous Localization and Mapping) 알고리즘에 대한 많은 연구가 진행되어 왔다[1-5]. 초기 SLAM 연구들은 2차원 공간상에서의 로봇의 움직임에 집중되었으나, 최근에는 3차원 공간상에서의 6자유도 움직임에 대한 SLAM 알고리즘이 많이 개발되고 있으며, 이는 비행 및 수중 로봇에 적용되고 있다[6-8].

초기의 3차원 SLAM에서는 레이저 거리 센서 혹은 스테레오 비전이 많이 사용되었다. 그러나 레이저 거리 센서는 정확도가 높은 반면 가격이 비싼 단점이 있으며, 스테레오 비전의 경우 가격적인 부담이 크지 않으나 측정된 결과값에 잡음이 많다는 단점이 있다. 최근 Microsoft 사에서는 Kinect 센서를 발표하였고, 이를 이용한 3차원 SLAM이 많이 연구되고 있다[9-15]. Kinect 센서는 적외선 패턴을 이용하여 거리를 측정하는 방식이며, 실내에서 사용이 가능하다. 가격 또한 레이저 센서에 비하여 상당히 저렴하며, 스테레오 비전에 비해 정확한 비교적 거리 값을 얻을 수 있고, 2차원 이미지 및 각 이미지 픽셀(pixel)에 대한 거리 정보를 30Hz의 속도로 제공한다[16]. 이러한 Kinect 형태의 센서들을 RGB-D (Red-Green-Blue Depth) 카메라 센서라 칭하고 있다. Kinect 센서의 외관 및 획득된 데이터는 그림 1과 같다.

P. Henry 등은 FAST (Features from Accelerated Segment Test) 특징점 추출 알고리즘과 ICP (Iterative Closest Point) 정합을 이



(a) Microsoft Kinect sensor.



(b) RGB color image.



(c) Per-pixel depth data.

그림 1. RGB-D 센서 시스템.

Fig. 1. RGB-D sensor system.

용하여 Kinect 센서와 함께 3차원 SLAM을 구현하였다[12]. 이때 SLAM 코어 알고리즘으로는 TORO (Tree-based netwORk Optimizer) 알고리즘을 사용하였다. F. Endres 등은 ICP 정합 및 g2o (General framework for Graph Optimization) SLAM 알고리즘을 이용하여 3차원 SLAM을 구현하였다[11]. 하지만 이들 알고리즘들은 연산량의 제한으로 초당 3 프레임 정도의 속도를 보여주고 있으며, 실시간 처리에 있어 한계점이 있다.

본 논문에서는 GPU를 이용한 병렬연산을 통해 3D SLAM 알고리즘을 가속화하여 초당 20 프레임 이상의 처리 속도를 가지는 RGB-D 3차원 SLAM 알고리즘을 제안한다. 비주얼 오도메트리(visual odometry) 추정 및 루프 클로저(loop closure)을 위해 2차원 이미지의 특징점이 사용되었다. 그리고 3차원 RANSAC (RANDOM SAmple Consensus) 알고리즘을

\* 책임저자(Corresponding Author)

Manuscript received February 28, 2013 / revised March 8, 2013 / accepted April 2, 2013

이동화: KAIST 건설 및 환경공학과(ledonghwa@kaist.ac.kr)

김형진, 명현: KAIST 건설 및 환경공학과, 로봇공학 학계전공

(hjkim86@kaist.ac.kr/hmyung@kaist.ac.kr)

※ 본 연구는 지식경제부의 지원 (실용화환경에 강인한 도로 기반 저가형 자율주행기술 개발, 10035354) 및 국토해양부 U-City 석·박사 과정 지원 사업에 의하여 연구되었음.

통해 6자유도 변위를 추정하였다. SLAM 알고리즘으로는 MIT에서 개발된 그래프 기반 iSAM (Incremental Smoothing And Mapping) 알고리즘을 사용하였고[6], 방 크기의 공간에서 실내 실험을 통해 제안된 시스템을 검증하였다.

II. 제안된 3차원 SLAM 시스템

본 논문에서 제안하는 3D SLAM 시스템의 전체 구성은 그림 2와 같다. 먼저 RGB-D 센서로부터 2차원 이미지와 거리 정보가 입력된다. 거리 정보의 해상도 또한 2차원 이미지의 해상도와 동일하다. 다음으로 2차원 이미지로부터 특징점이 추출되며, 이 특징점은 바로 전 RGB-D 센서 데이터와의 비교를 통해 비주얼 오도메트리 추정에 사용되며, 또한 과거의 특징점 정보들을 이용하여 로봇이 이전에 존재했던 위치와의 연관성을 확인한다. 과거 위치와의 연관성이 확인될 시 이를 루프 클로저라하며, 이는 그래프 기반 SLAM 알고리즘의 구속조건으로 이용된다. 비주얼 오도메트리 및 구속 조건을 통해 SLAM 알고리즘이 수행되며, 최종적으로 보정된 로봇의 궤적과 3차원 지도를 얻게 된다. 제안되는 3D SLAM 시스템의 각 단계별 자세한 사항은 아래와 같다.

1. 이미지 특징점 추출 및 매칭, 비주얼 오도메트리 추정

센서의 자세 변화 추정을 위해 본 시스템에서는 추가적인 센서의 사용 없이 RGB-D 센서만을 이용하였다. 비주얼 오도메트리 추정을 위해 이미지에서 특징점을 추출하였으며, 특징점의 매칭 및 매칭된 특징점들의 이미지상의 위치 변화를 이용하여 센서의 최종 자세 변위를 추정하게 된다.

특징점 추출 및 매칭에 많이 사용되는 알고리즘으로 SIFT (Scale-Invariant Feature Transform) 및 SURF (Speeded-Up Robust Features) 알고리즘이 있다. 정확도 면에서 SIFT 알고리즘의 성능이 우수하나 연산 속도에 있어 SURF 알고리즘에 비해 불리하다. 본 연구에서는 실시간 구현을 위해 SURF 알고리



(a) Previous image frame. (b) Current image frame.

그림 3. 이미지 특징점 추출 및 매칭 결과.

Fig. 3. Image feature extraction and matching.

즘을 사용하였다. SURF 알고리즘을 통한 이미지 매칭 결과는 그림 3과 같다. 좌측 화면은 이전에 센서로 들어온 화면이며 우측 화면은 현재 센서로 입력되는 화면이다. 각 화면에서 추출된 특징점은 점으로 표시되었으며, 특징점 기술자 벡터의 유클리디언 거리 비교를 통해 서로 매칭된 특징점들은 선으로 연결하였다. 최근 GPU를 통한 병렬연산이 특징점 추출 알고리즘에 적용되고 있으며, CPU 연산에 비해 20배 정도의 연산 속도 향상을 보여준다. 본 연구에서 사용되는 SURF 알고리즘 또한 GPU 연산을 통해 가속화 되었다.

일반적으로 2차원 이미지상의 특징점 비교를 통해 변위를 추정할 시 3차원 공간상의 6자유도 변위 추정에는 정확도에 한계가 있다. RGB-D 센서에서는 이미지 각 픽셀에 해당하는 거리 값을 알 수 있으므로 매칭된 특징점들의 3차원 좌표 정보를 얻을 수 있다. 특징점의 3차원 좌표  $(x, y, z)$  중  $z$  값은 센서에서 얻어지는 거리 값과 같으며  $x, y$  값은 아래 식과 같이 구할 수 있다.

$$x = \frac{uD(u, v)}{f} \tag{1}$$

$$y = \frac{vD(u, v)}{f} \tag{2}$$

$u, v$  는 이미지상의 픽셀좌표를 나타내며,  $D(u, v)$ 는 각 픽셀의 거리 값을 나타낸다.  $f$  는 카메라의 초점 거리(focal length) 값이다. 특징점들의 3차원 좌표 획득 후, 3차원 RANSAC 알고리즘을 이용하여 매칭된 특징점들의 3차원 좌표 비교를 통해 가외치(outlier)를 제거한 후, 6자유도 변위를 추정한다. 이를 통해 각기 다른 위치에서 측정된 센서 데이터간의 6자유도 변위를 추정하게 된다. 결국 현재 센서 데이터와 바로 전 센서 데이터와의 특징점 비교를 통해 6자유도 변위값을 가지는 비주얼 오도메트리가 추정된다.

2. 특징점 관리자 및 루프 클로저

비주얼 오도메트리를 통해 추정된 센서의 위치와 6자유도 정보는 그래프(graph) 형태로 저장된다(그림 4). 각 노드(node)는 센서의 위치를 나타내며, 노드 사이의 에지(edge)는 노드들 간의 6자유도 변위를 나타낸다. 앞서 구해진 비주얼 오도메트리 정보는 시간상으로 바로 전 노드와 현재 노드를 연결하는 에지 정보가 된다.

각 노드에서 추출된 특징점들은 우선 특징점 관리자에 모두 보내진다. 특징점 관리자는 기준에 등록된 특징점들의 기술자 비교를 통해 기준치 이상의 유클리디언 거리값을 가지는 새로운 특징점만을 등록한다. 등록된 모든 특징점들은 각

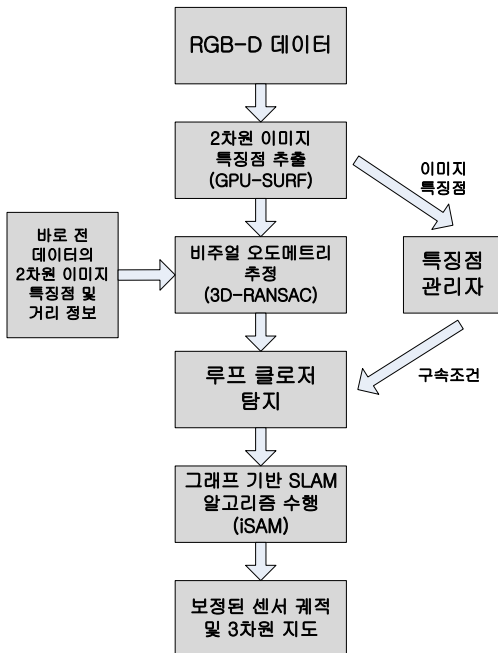


그림 2. 제안된 3차원 SLAM 시스템 구성도.  
Fig. 2. Overview of the proposed RGB-D 3D SLAM system.

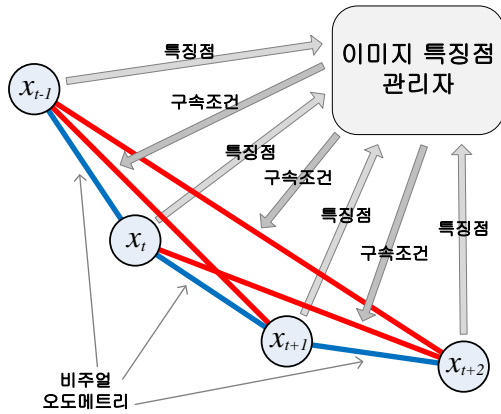


그림 4. 이미지 특징점 관리자와 루프 클로저.

Fig. 4. Image feature manager and loop closure.

특징점들이 속한 노드들의 정보가 기록되어 있으며, 하나의 특징점은 다수의 노드에 속할 수 있다. 또한 각 특징점들은 각 노드에서 위치한 3차원좌표 정보를 가지고 있으며, 이는 루프 클로저 탐지 노드간의 6자유도 변위 추정을 위해 사용된다.

특징점 관리자는 새로운 노드가 추가되면 해당 노드의 특징점 정보를 이용하여, 새로운 특징점을 등록 후 과거 노드들과 현재 노드간의 공통된 특징점들을 찾아준다. 일정 수 이상의 공통된 특징점들을 가질 경우 RANSAC 알고리즘을 이용하여 두 노드간의 6자유도 변위를 추정하게 되며, 이를 노드간의 구속 조건이라 부른다. 추정된 변위는 그래프상의 새로운 에지정보로 추가되며, 이러한 과정을 루프 클로저라 칭한다.

결론적으로 각 노드 사이를 연결하는 에지들은 초기 오도메트리와 특징점 관리자를 통한 구속 조건이며, 이를 통해 각 센서 위치에 대한 그래프 형태가 만들어진다. 특징점 관리자 및 그래프에 대한 상세한 그림은 그림 4와 같다.

### 3. 그래프 기반 SLAM 알고리즘

그래프 기반 SLAM 알고리즘에서는 센서의 위치를 나타내는 노드와 노드 사이의 구속 조건을 나타내는 에지로 그래프를 구성하며, SLAM 알고리즘을 통하여 노드들의 위치가 최적화된다. 기존 EKF SLAM 알고리즘과 같은 필터링 기반의 알고리즘들은 현재 센서 위치만을 보정한다[1]. 전체 센서의 위치 궤적을 보정하기 위해서는 기존 연구들에서는 오프라인 SLAM 알고리즘이 사용되었다[17]. 그러나 최근 그래프 기반의 전체 궤적을 보정하는 실시간 SLAM 알고리즘이 개발되고 있다. 본 연구에서는 최종 얻어지는 3차원 모델의 최적화를 위해 전체 궤적을 보정하였다. 이를 위해 사용된 그래프 기반 SLAM 알고리즘은 M. Kaess에 의해 개발된 iSAM 알고리즘이다. iSAM 알고리즘은 그래프 기반 SLAM 문제를 희소 선형 대수 (sparse linear algebra) 기법을 이용하여 최적화시켰으며, 희소 선형 대수의 특징을 이용하여 속도를 가속화시켰다[6]. 본 시스템에 적용 결과 실시간성을 충분히 보장해 주었다.

### III. 실험 결과

본 연구에서 제안된 3차원 SLAM 알고리즘은 일반 방 크기의 공간에서 실험을 통하여 검증되었다.

RGB-D 데이터를 위하여 Microsoft 사의 Kinect 센서가 이용되었다. Kinect 센서는 640×480의 해상도를 가지며 30Hz 속도로 동작한다. 각각의 이미지 픽셀은 적외선 패턴을 이용하여 얻어진 거리 정보를 가진다. 거리 값의 유효 범위는 대략 0.5~5m로 알려져 있다[16].

본 시스템은 Intel i7 CPU와 GPU 연산을 위한 Nvidia GT 560 Ti 그래픽 카드가 장착된 데스크탑 PC에서 구현되었다. OS는 Ubuntu 12.04를 사용하였고, Nvidia 그래픽 카드의 병렬 연산을 위해 CUDA (Compute Unified Device Architecture) 프로그래밍 언어를 이용하였다. 이미지 특징점 추출을 위해 OpenCV 2.4.0이 사용되었다[18]. 3차원 RANSAC 알고리즘을 위해서는 PCL (Point Cloud Library) 이 사용되었다[19]. 또한 PCL은 RGB-D 센서 데이터의 표시 및 3차원 지도 생성 등에도 이용되었다.

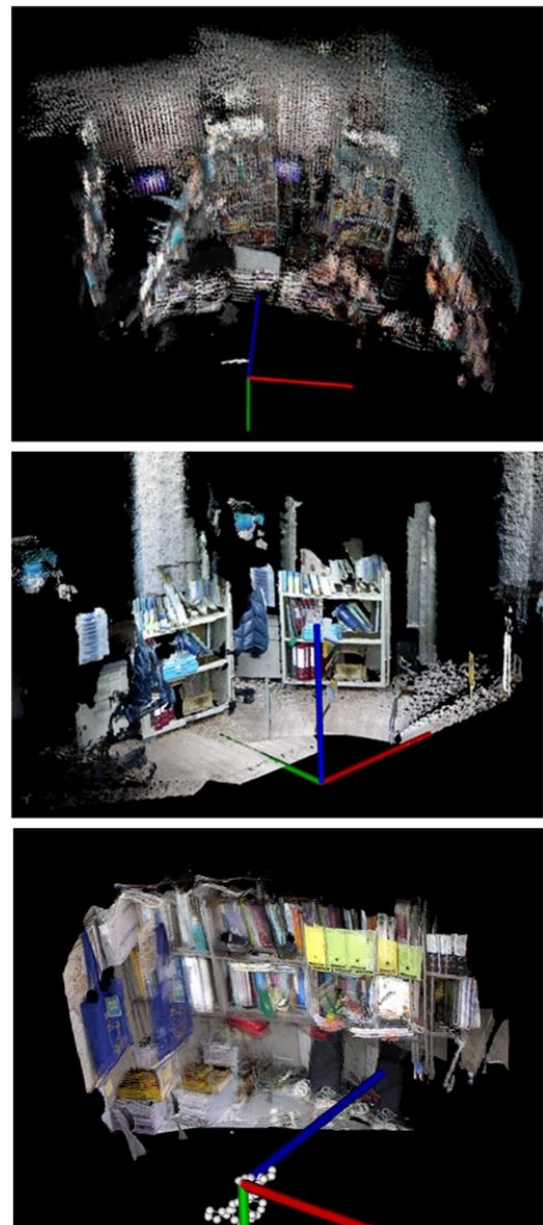


그림 5. 비주얼 오도메트리 만을 통해 생성된 3차원 지도.

Fig. 5. 3D map by visual odometry only.

방 하나의 공간에서 제안된 3차원 SLAM 시스템의 실험이 수행되었다. Kinect 센서를 손으로 이동시키며 센서 데이터를 획득하였으며, 이와 동시에 데스크탑 PC에서 실시간으로 알고리즘이 수행되었다. 각각의 노드에서 획득된 센서 데이터들은 2차원 이미지 정보 및 거리 정보를 이용하여 각 픽셀이 3차원 상에서 컬러 정보를 가지는 하나의 점으로 표현된다. 이러한 점들의 집합, 즉 점군 데이터(point cloud)들은 각 노드가 가지는 6자유도 변위에 의해 좌표계가 바뀌어 3차원 공간에 뿌려진다. 이러한 점군 데이터 집합의 결과는 실험이 진행된 공간의 3차원 지도로 나타나게 된다. 실험을 통해 생성된 3차원 지도는 그림 5 및 6과 같다. 그림 5는 비주얼 오도메트리 정보만을 이용하여 생성된 3차원 지도 결과이다. 오도메트리의 미끌림(drift) 현상으로 인하여 3차원 지도가 정합

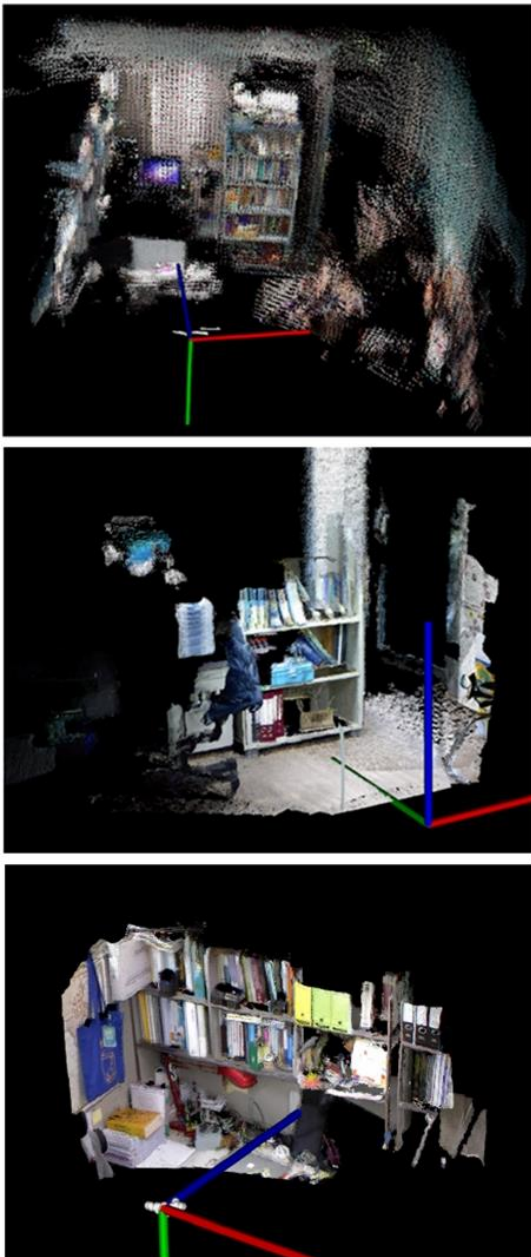


그림 6. 제안된 알고리즘을 통해 보정된 3차원 지도.  
Fig. 6. Optimized 3D map by the proposed algorithm.

표 1. 시스템 구성 요소들의 평균 처리 시간.

Table 1. Average processing time of the system components.

시스템 구성 요소	처리 시간(ms)
SURF 알고리즘 특징점 추출	17.2
비주얼 오도메트리 추정	5.9
루프 클로저	8.1
iSAM SLAM 알고리즘	5.4
총합	36.6

되지 않은 모습을 보여준다. 그림 6은 제안된 알고리즘을 통해 보정된 3차원 지도 결과이다. 센서의 전체 궤적이 보정되었으며, 이를 통해 정합된 3차원 지도 결과를 보여준다.

시스템 구성 요소들의 수행 시간은 위의 표 1과 같다. 이미지 특징점 추출에 가장 많은 시간이 소요됨을 알 수 있다. 그 다음으로 특징점 관리자를 통한 루프 클로저 탐지에 많은 시간이 소요되었다. 제안된 3차원 SLAM 시스템의 1회 수행에 있어 평균적인 전체 수행 시간은 36.6 ms로 측정되었다. 결과적으로 제안된 시스템은 20Hz 이상의 수행 속도를 보여준다.

#### IV. 결론

본 연구에서는 RGB-D 센서를 이용한 실시간 3D SLAM 알고리즘을 제안하였다. SURF 이미지 특징점 추출 알고리즘이 사용되었으며, GPU를 이용한 병렬 연산을 통해 가속화되었다. 각 센서 위치의 6자유도 변위 추정을 위해 3차원 RANSAC 알고리즘이 사용되었으며, 이미지 특징점 관리자 그래프 생성 및 루프 클로저 탐지가 이루어졌다. 그래프 기반 SLAM 알고리즘은 센서의 전체 궤적을 최적화 시켰으며, 이를 통해 점군 형태의 보정된 3차원 지도를 얻을 수 있었다. 실시간성에 있어서는 데스크탑 PC에서 초당 20회 이상의 알고리즘 수행 속도를 보여주어, 충분히 실시간 수행이 가능하였다.

향후 연구에서는 보다 낮은 성능의 임베디드 프로세서에서 본 시스템을 구현할 예정이다. 또한, 특징점을 추출하기 어려운 공간에서도 거리 정보만을 이용한 ICP 알고리즘 등을 통해 성능을 향상시킬 계획이다.

#### 참고문헌

- [1] M. W. M. G. Dissanayake, P. Newman, S. Clark, H. F. Durrant-Whyte, and M. Csorba, "A solution to the simultaneous localization and map building (SLAM) problem," *Robotics and Automation, IEEE Transactions on*, vol. 17, no. 3, pp. 229-241, Jun. 2001.
- [2] D. Hahnel, W. Burgard, D. Fox, and S. Thrun, "An efficient FastSLAM algorithm for generating maps of large-scale cyclic environments from raw laser range measurements," *Proc. Intelligent Robots and Systems, 2003. (IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on*, pp. 206-211, Oct. 2003.
- [3] M. Montemerlo, S. Thrun, D. Koller, and B. Wegbreit, "FastSLAM 2.0: An improved particle filtering algorithm for simultaneous localization and mapping that provably converges," *Proc. of the 18th International Joint Conference on Artificial Intelligence*, pp. 1151-1156, 2003.

- [4] S.-Y. Hwang and J.-B. Song, "Monocular vision and odometry-based SLAM using position and orientation of ceiling lamps," *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 17, no. 2, pp. 164-170, Feb. 2011.
- [5] K.-Y. Yoon, S.-W. Choi, and C.-H. Lee, "An approach for localization around indoor corridors based on visual attention model," *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 17, no. 2, pp. 93-101, Feb. 2011.
- [6] M. Kaess, A. Ranganathan, and F. Dellaert, "iSAM: Incremental smoothing and mapping," *Robotics, IEEE Transactions on*, vol. 24, no. 6, pp. 1365-1378, Dec. 2008.
- [7] A. S. Huang, A. Bachrach, P. Henry, M. Krainin, D. Maturana, D. Fox, and N. Roy, "Visual odometry and mapping for autonomous flight using an RGB-D camera," *Proc. of The 15th International Symposium on Robotics Research (ISRR)*, 2011.
- [8] R. Kummerle, G. Grisetti, H. Strasdat, K. Konolige, and W. Burgard, "g2o: A general framework for graph optimization," *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, pp. 3607-3613, May 2011.
- [9] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers, "A benchmark for the evaluation of RGB-D SLAM systems," *Proc. Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, pp. 573-580, Oct. 2012.
- [10] F. Endres, J. Hess, N. Engelhard, J. Sturm, D. Cremers, and W. Burgard, "An evaluation of the RGB-D SLAM system," *Proc. Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pp. 1691-1696, May 2012.
- [11] N. Engelhard, F. Endres, J. Hess, J. Sturm, and W. Burgard, "Real-time 3D visual SLAM with a hand-held RGB-D camera," *Proc. of the RGB-D Workshop on 3D Perception in Robotics at the European Robotics Forum, Vasteras, Sweden, 2011*.
- [12] P. Henry, M. Krainin, E. Herbst, X. Ren, and D. Fox, "RGB-D mapping: Using Kinect-style depth cameras for dense 3D modeling of indoor environments," *The International Journal of Robotics Research*, vol. 31, no. 5, pp. 647-663, Apr. 2012.
- [13] D. Lee, H. Kim, and H. Myung, "Real-time RGB-D 3D SLAM with GPU acceleration," *Proc. of Daejeon & Chungcheong Regional Conference of Institute of Control, Robotics and Systems (in Korean)*, pp. 179-182, Dec. 2012.
- [14] D. Lee, H. Kim, and H. Myung, "2D image feature-based real-time RGB-D 3D SLAM," *Proc. of International Conference on Robot Intelligence Technology and Applications 2012 (RiTA 2012)*, pp. 485-492, Gwangju, Korea, Dec. 2012.
- [15] D. Lee, H. Kim, and H. Myung, "GPU-based real-time RGB-D 3D SLAM," *Proc. of International Conference on Ubiquitous Robots and Ambient Intelligence 2012 (URAI 2012)*, pp. 46-48, Daejeon, Korea, Nov. 26-29, 2012.
- [16] Z. Zhang, "Microsoft Kinect sensor and its effect," *IEEE Multimedia*, vol. 19, no. 2, pp. 4-10, 2012.
- [17] S. Thrun and M. Montemerlo, "The GraphSLAM algorithm with applications to large-scale mapping of urban structures," *The International Journal of Robotics Research*, vol. 25, no. 5-6, pp. 403-429, May-Jun. 2006.
- [18] OpenCV (Open source Computer Vision): <http://opencv.org/>
- [19] PCL (Point Cloud Library): <http://pointclouds.org/>



### 이 동 환

2005년 경북대학교 전자전기공학부 학사. 2005년~2008년 (주)비앤디 선임연구원. 2010년 경북대학교 대학원 전자전기컴퓨터학부 석사. 2010년~현재 KAIST 건설 및 환경공학과 박사과정. 관심분야는 로봇 네비게이션, 영상 처리, 수중/수상 로봇, 수중음향신호처리.



### 김 형 진

2012년 경희대학교 전자공학과 학사. 2012년~현재 KAIST 건설 및 환경공학과 석사과정. 관심분야는 로봇 네비게이션, 영상 처리, 수중/수상 로봇.



### 명 현

1992년 KAIST 전자전산학 학사. 1994년 KAIST 전기 및 전자공학 석사. 1998년 KAIST 전기 및 전자공학 박사. 1998년~2002년 한국전자통신연구원(ETRI) 선임연구원. 2002년~2003년 (주)머시스 연구소장/기술이사. 2003년~2008년 삼성 전자종합기술원 전문연구원. 2008년~현재 KAIST 건설 및 환경공학과 교수. 관심분야는 로봇 네비게이션, 건설로봇, 수중/수상 로봇, 소프트웨어컴퓨팅.