

SLAM 과 소리 위치 탐색 알고리즘 통합 시스템

안인규¹, 손명배², 윤성의²

¹ 한국과학기술원 로봇공학학제전공, ² 한국과학기술원 전산학부

System integration of SLAM and Sound Localization algorithm

Inkyu An¹, Myungbae Son², Sung-Eui Yoon²

¹KAIST The Robotics Program, ²KAIST School of Computing

e-mail: inkyu_ahn@kaist.ac.kr, nedsociety@kaist.ac.kr, sungeui@kaist.edu

요 약

로봇이 실제 환경에서 구동 될 경우, 장애물 탐지, 목표물 추적 등 여러 물체들과의 상호작용은 필수적이다. 최근 연구들에서는 이와 더불어 로봇의 주변에서 감지되는 소리 정보를 또 하나의 중요한 정보로 인식하고 있다. 이는 소리를 내는 물체는 때때로 로봇이 피해야 할 장애물 또는 목적지에 대한 단서를 주기 때문이다. 뿐만 아니라 탐색된 소리의 위치와 로봇이 탐지한 지도와의 맵핑 역시 실용화를 위한 중요 요소이다. 본 논문에서는 SLAM(Simultaneous Localization and Mapping)을 통해 생성한 지도와 로봇의 위치 정보에 소리 위치 탐색 알고리즘을 이용해 획득한 소리 위치 정보를 통합한 시스템을 소개한다.

1. 서론

로봇이 실제 환경에서 안정적으로 동작하기 위해서는 수많은 상황에 대처할 수 있는 능력은 필수적이다. 그에 따라 다양한 센서 정보가 요구되며 소리 또한 그 중 하나다. 소리 정보는 로봇이 다른 물체와 상호작용이 일어나는 상황에서 중요도가 높은 요소이며, 특히 소리가 발생하는 위치에 대한 정보는 주변 환경을 파악하는 데 있어 상당한 정보를 제공하는 단서이다. 예를 들어, 로봇이 사람의 호출을 받아 이동할 경우 소리를 통해 사람의 위치를 측정할 수 있어야 한다. 또한 로봇이 이동 중 소리를 이용해 장애물을 추적하고 경로 생성 알고리즘에 반영할 수 있다. 이러한 이유로 소리의 위치를 찾는 알고리즘이 필요하며, 여러 관련 연구가 진행되고 있다. [1]

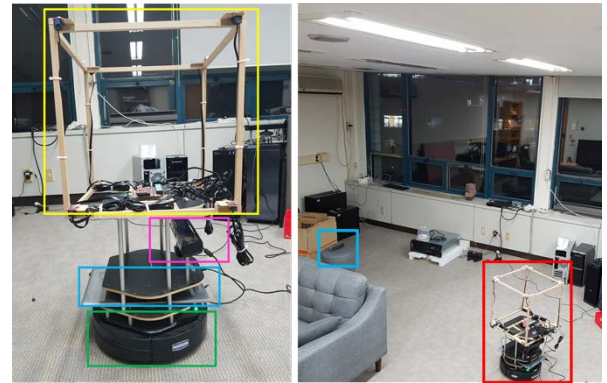
로봇 공학이 고도로 발전하며 더욱 다양한 종류의 센서가 로봇에 탑재되고 있다. 이러한 상황에서 여러 종류의 센서를 통합하여 새로운 정보를 제공하는 센서 융합(sensor fusion)이 각광받고 있다. 소리 또한 하나의 정보로는 사용이 제한될 수 있으나, 다른 센서와 통합을 통해 더욱 유용한 정보를 생산할 수 있다. 즉, 소리 위치 탐색 알고리즘 또한 지도와 로봇의 위치를 알고 있을 때 로봇의 상대적인 위치에 따른 소리의 방향을 정확히 측정할 수 있으며, 더욱 의미 있는 정보로 사용될 수 있다. 특히 카메라, 키넥트와 같이 특정 화각을 갖는 시각 센서의 경우, 시야 밖의 물체를 인식하기에 어려움이 있다. 이러한 경우 소리와 카메라의 융합을 통해 극복이 가능하다.

본 논문에서는 [1]에서 제안된 소리 위치 탐색 알고리즘을 사용해 소리의 위치를 탐색했으며, [2]에서 제안된 RGB-D SLAM을 이용하여 지도 정보인 Octomap[3]과 로봇의 위치 정보를 추정하였다. 소리의 위치와 지도 정보 통합은 ROS에 구현되었다.

결과에서는 카메라 시야 밖에서 소리가 발생했을 때 로봇이 탐지가 가능하다는 점을 보여준다.

2. 본론

2.1 소리 위치 탐색 알고리즘



[그림 1] 좌: 실험에 사용된 로봇(노랑: 마이크 큐브, 자주: 키넥트 센서, 파랑: 노트북, 초록: 터틀봇 v2), 우: 실험 환경(파랑: 블루투스 스피커(LG-NP7550), 빨강: 로봇 위치)

소리의 위치를 탐색하기 위해 두 마이크의 소리 도달 시간의 차이를 측정하는 TDOA(Time Difference Of Arrival)[1] 방법을 많이 사용한다. 마이크 M_i 과 마이크 M_j 에서 L 개의 소리를 샘플링하여 측정된 신호가 $X_i(k)$ 과 $X_j(k)$ 라 했을 때 아래 [식 (1)]을 통해 두 신호의 cross-correlation을 구한다.

$$R_{i,j}(\tau) \approx \sum_{k=0}^{L-1} X_i(k)X_j(k)^* e^{\frac{j2\pi k\tau}{L}} \dots (1)$$

식 (1)에서 $R_{i,j}(\tau)$ 이 최대가 되는 τ 를 구하고, 이 값을 두 신호의 도달 시간의 차이 $D_{i,j}$ 로 추정할 수 있다. [식 (2)]

$$D_{i,j} \approx \operatorname{argmax}_{\tau} (R_{i,j}(\tau)) \dots (2)$$

하지만 위 [식 (2)]는 2D 공간에서 전방 180도의 방향만을 측정할 수 있다. 3D 공간의 모든 방향을 처리하기 위해서는 최소 4개 이상의 마이크가

필요하며, 본 논문에서는 8개의 마이크가 장착된 큐브[그림 1, 좌, 노랑]를 사용하였다. [4]

마이크 큐브의 중앙을 기준으로 N개의 격자(grid)로 분할된 가상의 단위구를 생성한 후, 모든 격자에 대해 8개의 마이크로 가질 수 있는 모든 경우의 $D_{i,j}$ 를 계산해 저장한 후 함수화(lookup(n, i, j)) 한다. [식 (3)]

$$D'_{i,j} = \text{lookup}(n, i, j) \dots (3)$$

$$E_n^t = E_n^{t-1} + R_{i,j}(D'_{i,j}) \dots (4)$$

where, $\{n \in N\}$ and $\{(i, j) \in (\text{all microphone pair})\}$

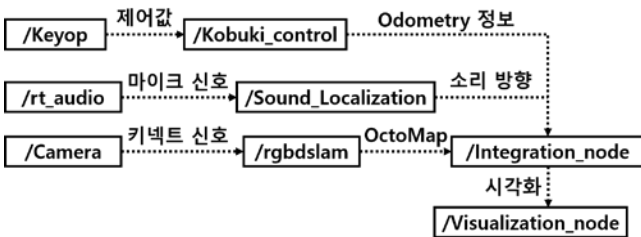
E_n 는 해당 격자에서 8개의 마이크로 가질 수 있는 모든 경우의 cross correlation의 합이다. 모든 마이크 쌍에 대해 [식 (3)]을 통해 미리 계산된 $D'_{i,j}$ 를 찾은 후, [식 (4)]를 통해 모든 마이크 쌍의 cross correlation의 합 E_n 을 구한다. 이때 E_n 가 최대가 되는 격자 n이 소리의 방향이 된다.

2.2 RGB-D SLAM

3D 공간 지도 정보인 Octomap[3]과 로봇이 이동하며 실시간으로 위치를 추정하기 위해서는 SLAM은 필수적이다. 본 논문에서는 [2]의 RGB-D SLAM을 사용하여 시스템을 구성했다.

RGB-D SLAM은 키넥트 센서의 RGB와 Depth 정보를 사용한 기술로 카메라로부터 SIFT 정보를 추출한 후, 각 프레임의 Geometry 값의 차이를 이용해 로봇의 위치를 추정한다.

[그림 3]에서 가시화된 Octomap을 확인할 수 있다.



[그림 2] ROS에서 동작하는 노드들과 메시지 통신 관계

소리 위치 탐색 알고리즘과 RGB-D SLAM의 결과는 [그림 2]와 같이 ROS에서 통합된다. 생성된 octomap과 추정한 소리의 방향은 integration node에서 통합 후, visualization node에서 시각화한다.

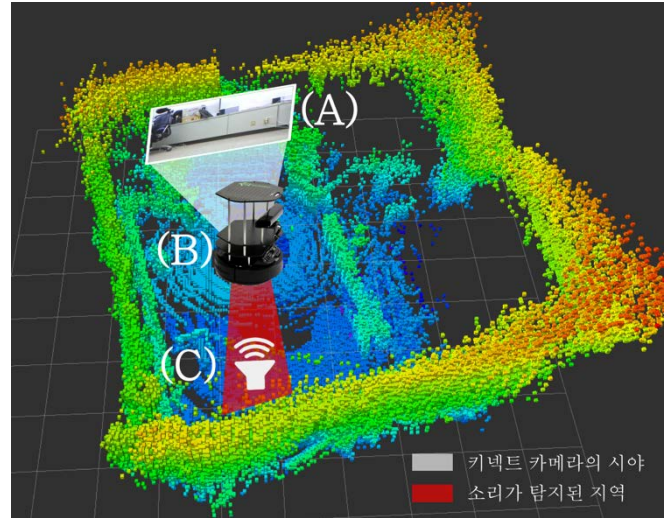
3. 결과

제안된 프레임워크를 검증하기 위해 소리 위치 탐색 알고리즘과 SLAM을 통합한 시스템을 구현하여, 로봇의 상대적인 위치에서 소리가 발생한 방향을 표시하게 하였다. [그림 3]와 같이 지도 정보(Octomap) 상에서 키넥트 카메라의 이미지가 (A) 일 때, 로봇의 상대 위치(B)와 탐지된 소리의 위치 정보(C)가 표시된다. 이는 키넥트 카메라의 화각(A)을 벗어난 지점에서 소리가 발생 시에도 로봇에서 탐색이 가능하다

는 사실을 확인할 수 있다(C). 이는 로봇 제어의 안정성을 높일 수 있을 것이라 기대된다.

본 논문에서 제시한 방법은 추후 사람과의 대화형 애플리케이션 또는 경로 생성 알고리즘에서 활용 가능할 것이라 생각된다.

현재 소리 위치를 방향 정보로만 찾는다는 한계가 존재한다. 향후 연구에서 지도에서 소리의 좌표 공간 위치를 탐색할 수 있는 방향으로 발전시키고자 한다.



[그림 3] 지도(octomap) 상에서 물체가 키넥트 카메라 이미지에서 보이지 않더라도 소리 위치 탐색 알고리즘을 이용해 탐지할 수 있다. (A): 키넥트 카메라 이미지, (B): 마이크 큐브와 키넥트가 장착된 터틀봇, (C): 소리의 발생지

후기

이 논문은 한국연구재단의 지원으로 수행되었습니다(NO.2013-067321).

참고문헌

- [1] Jean-Marc Valin, Francois Michaud, Jean Rouat, "Robust localization and tracking of simultaneous moving sound sources using beamforming and particle filtering," Robotics and Autonomous Systems, pp. 216-228, 2007.
- [2] Felix Endres, Jurgen Hess, Jurgen Sturm, Daniel Cremers, Wolfram Burgard, "3D Mapping with an RGB-D Camera," IEEE Transactions on Robotics, volume 30, Issue 1, pp. 177-187, 2014.
- [3] Armin Hornung, Kai M. Wurm, Maren Bennewitz, Cyrill Stachniss, Wolfram Burgard, "OctoMap: an efficient probabilistic 3D mapping framework based on octrees," Autonomous Robots, volume 34 issue 3, pp. 189-206, 2013.
- [4] Abran-Cote D., Bandou M., Beland A., Cayer G, Choquette S., Gosselin F., Robitaille F., Telly Kizio D., "USB Synchronous Multichannel Audio Acquisition", IntRoLab Technical Paper, 2011.