

어안 ORB-SLAM 알고리즘을 사용한 구면 비디오로부터의 3D 맵 생성

김기식, 박종승
인천대학교 컴퓨터공학과
wis1906@naver.com, jong@inu.ac.kr

3D Map Construction from Spherical Video using Fisheye ORB-SLAM Algorithm

Ki-Sik Kim, Jong-Seung Park
Dept. of Computer Science & Engineering, Incheon National University

요 약

본 논문에서는 구면 파노라마를 기반으로 하는 SLAM 시스템을 제안한다. Vision SLAM은 촬영하는 시야각이 넓을수록 적은 프레임으로도 주변을 빠르게 파악할 수 있고, 많은 양의 주변 데이터를 이용해 더욱 안정적인 추정이 가능하다. 구면 파노라마 비디오는 가장 화각이 넓은 영상으로, 모든 방향을 활용할 수 있기 때문에 Fisheye 영상보다 더욱 빠르게 3D 맵을 확장해나갈 수 있다. 기존의 시스템 중 Fisheye 영상을 기반으로 하는 시스템은 전면 광각만을 수용할 수 있기 때문에 구면 파노라마를 입력으로 하는 경우보다 적용 범위가 줄어들게 된다. 본 논문에서는 기존에 Fisheye 비디오를 기반으로 하는 SLAM 시스템을 구면 파노라마의 영역으로 확장하는 방법을 제안한다. 제안 방법은 카메라의 투영 모델이 요구하는 파라미터를 정확히 계산하고, Dual Fisheye Model을 통해 모든 시야각을 손실 없이 활용한다.

1. 서론

SLAM(Simultaneous Localization and Mapping) Robotics와 AR(Augmented Reality) 등의 분야에서 3D Map을 생성하거나 장치의 위치를 추정하기 위한 과거부터 연구가 활발히 이루어진 기술이다. 이 중, 단일 촬영 장치 하나만을 이용해 SLAM 기능을 수행하는 Monocular Vision SLAM 기술은 카메라라는 저렴하고 가볍고 작은 하드웨어로도 준수한 성능을 낼 수 있다는 강점이 있다. Vision SLAM은 카메라가 가지는 시야각의 넓이가 넓을수록 빠른 시간 안에 주변을 파악할 수 있고, 많은 양의 주변 데이터를 이용해 더욱 정확한 위치 추정이 가능하다 [1]. 이를 위해 방사왜곡을 포함한 Fisheye 비디오와 구면 파노라마 비디오 등을 입력 영상으로 활용하려는 Omnidirectional SLAM 연구가 이루어지고 있다.

David Caruso등[2]은 이전에 제안되었던 LSD-SLAM[3]을 활용하여 Fisheye 비디오를 이용하여 Unified Omnidirectional Model에 대한 공식화를 통해 SLAM 시스템을 구현하였다. Alejandro Rituerto등[4]은 EKF(Extended Kalman Filter)을

활용한 MonoSLAM을 기반으로 하여 Fisheye Camera를 이용한 SLAM 시스템을 구현하였다.

이러한 SLAM 시스템들은 Fisheye 비디오를 입력으로 하였기 때문에 약 180°의 화각에서 한계점을 보인다. 구면 파노라마 비디오는 화각이 가장 넓은 이미지로, 모든 방향에 대한 시각적 데이터를 매 프레임마다 활용하기 때문에 Fisheye 비디오보다 더욱 빠르게 3D 맵을 확장해나갈 수 있다.

Liu등[5]은 비교적 최근에 Omnidirectional SLAM 시스템을 제안하였다. Liu등의 시스템은 ORB-SLAM2[6]에 Enhanced Unified Camera Model(EUCM)[7]을 적용하여 Fisheye Video에서도 투영 과정을 정확하게 진행할 수 있는 시스템을 제안하였다. Liu등의 시스템도 Fisheye Video를 기반으로 동작하기 때문에 시야각에 대한 한계점이 있다. 이를 보완하기 위해서는 구면 파노라마를 입력 비디오로 하는 시스템의 변화가 필요하다.

본 논문에서는 Liu등의 시스템에 구면 파노라마를 입력 비디오로 하여 모든 방향에 대한 시각적 데이터를 활용하는 SLAM 시스템을 제안한다. 제안 방법은 구면 파노라마 비디오를 Dual Fisheye 비디

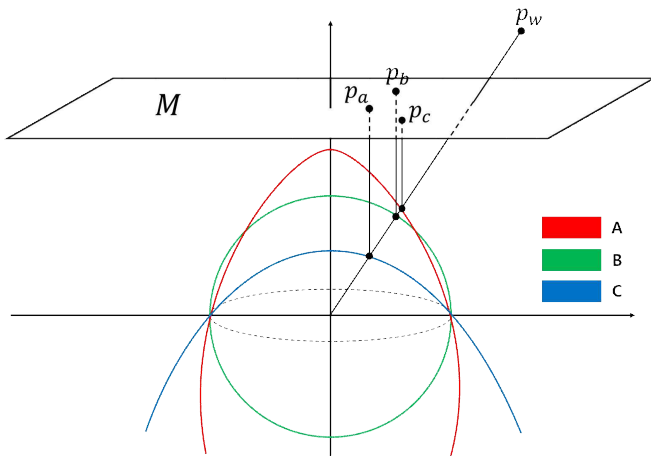
오의 형태로 다루고, 카메라 좌표의 전방과 후방에 두 개의 Fisheye 투영 평면을 생성하여 기능을 수행한다.

2. 제안 방법

Liu등의 시스템은 EUCM을 활용하여 195°의 시야각을 가지는 Fisheye 영상에 대해 기능을 수행할 수 있으며, EUCM의 고유한 내부 파라미터인 α, β 를 조정하여 다른 시야각의 Fisheye 영상을 수용할 수 있다. 본 논문에서 제안하는 방법은 구면 파노라마 영상을 Dual Fisheye 형태로 투영하여 사용한다. 이는 구형의 이미지를 반구로 나누어 각각 투영하는 것이기 때문에, 180°의 화각을 가지는 두 장의 Fisheye 영상이 된다. 따라서, 본 장에서는 180°의 화각에 적절한 투영 파라미터 α, β 를 구하는 방법에 대해 설명하고, 이어서 Dual Fisheye 비디오를 적용한 시스템에 대해 설명한다.

2.1 카메라 투영 파라미터 획득

EUCM은 핀홀 모델이 가지는 초점거리(f), 주점(c) 등의 내부 파라미터 외에도 α, β 파라미터를 가지고 있다.



(그림 1) 서로 다른 EUCM의 α, β 파라미터 값에 따른 투영하고자 하는 이미지의 모양 변화

EUCM에서는 월드 속 점을 왜곡이 존재하는 원본 이미지의 투영면에 한 번 투영하고, 평면의 투영면에 재투영한다. EUCM의 α, β 파라미터는 첫 번째 투영면의 모양을 결정하는 EUCM의 핵심 파라미터이다. 두 파라미터 값의 변화에 따라 (그림 1)의 A와 같은 뾰족한 형태가, B와 같은 원의 형태가, C와 같은 완만한 형태가 될 수 있다. α, β 의 값이 다

른 세 형태에 실제 세계의 점 P_w 가 투영된다면, $P_a \neq P_b \neq P_c$ 의 모두 다른 좌표에 투영되기 때문에, 정확한 투영을 위해서 α, β 파라미터의 값을 정확하게 구해야 한다. 본 논문에서는 구형의 이미지인 구면 파노라마 이미지를 입력으로 하기 때문에, (그림 2)의 B와 같은 형태가 나오도록 해야 한다.

EUCM에서, 투영 평면 M 상의 점 X 가 왜곡이 존재하는 원본 이미지의 투영면을 거쳐 평면 투영 평면 M 상의 점 P 로 재투영되기 위한 식은 다음과 같다.

$$P = \begin{bmatrix} x/(\alpha\rho + (1-\alpha)z) \\ y/(\alpha\rho + (1-\alpha)z) \\ 1 \end{bmatrix}, \alpha \in [0,1] \text{ and } \beta > 0, \quad (1)$$

$$\rho = \sqrt{\beta(x^2 + y^2 + z^2)}, \quad (2)$$

해당 식에 대해 구면 파노라마에서의 α 와 β 은 핀홀 평면에서 구면 영상의 투영 공식을 통해 알아낼 수 있다. $z=1$ 인 투영 평면 M 의 점 X' 가 구형 투영면 속 점 P' 로 투영되기 위한 식은 다음과 같다.

$$P' = \left[\frac{x}{r}, \frac{y}{r}, \frac{z}{r} \right]^T, r = \sqrt{x^2 + y^2 + 1} \quad (3)$$

식 (3)을 통해 구한 P' 의 z 축 값을 1로 바꾸면 투영 평면 M 상의 점이 되는데, 이 점은 식(1)의 P 와 같은 값을 가진다. 따라서, 다음과 같은 식이 성립된다.

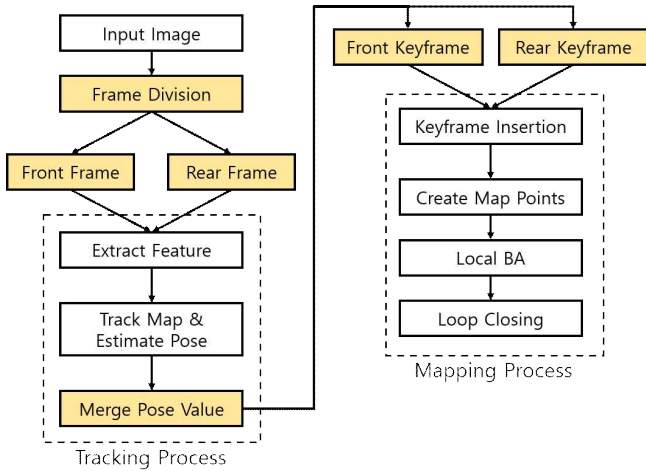
$$\alpha \sqrt{\beta(x^2 + y^2 + 1) + (1-\alpha)} = \sqrt{x^2 + y^2 + 1} \quad (4)$$

여기서, $\beta=1$ 을 대입하고, $\sqrt{x^2 + y^2 + 1}$ 을 k 로 치환한다면 $\alpha(k-1)=k-1$ 와 같이 정리할 수 있고, 이를 통해 $\alpha=1$ 임을 도출할 수 있다. 즉, EUCM에서 구면 영상을 투영하기 위해서는 $\alpha=1, \beta=1$ 의 파라미터 값을 대입해야 한다.

2.2 제안하는 시스템

기존의 시스템은 한 프레임 당 단일 입력 이미지에 대해서 동작하였다. 이를 Dual Fisheye 비디오에 적용시키기 위해서는 매 프레임마다 구면 파노라마

의 전면 반구에 대한 Fisheye 이미지, 후면 반구에 대한 Fisheye 이미지에 대해 각각 Tracking과 Mapping 과정을 수행해야 한다.



(그림 2) 제안하는 시스템의 흐름도

(그림 2)는 이를 위해 Liu등의 시스템을 변형한 시스템의 흐름도이다. 그림 속 노란색 박스는 결정적으로 추가한 부분이다. 제안하는 시스템은 구면 파노라마의 입력 이미지가 들어왔을 때, 해당 이미지를 전면 반구에 대한 Fisheye 이미지, 후면 반구에 대한 Fisheye 이미지로 재투영한다. 그리고 각각의 이미지 프레임에 대하여 특징점을 검출하고, 기존의 Map Point들과 대조하여 카메라 위치를 추정하는 과정을 거친다. 그리고 나서 전면 반구 프레임을 분석한 추정 좌표와 후면 반구 프레임을 분석한 추정 좌표를 통해 상호 보완하여 중간이 되는 추정 값을 구하는 과정을 거친다. 이에 대한 식은 다음과 같다.

$$X_m = X_f \frac{f_f}{f_f} + \begin{bmatrix} \cos\pi & 0 & \sin\pi \\ 0 & 1 & 0 \\ -\sin\pi & 0 & \cos\pi \end{bmatrix} X_r \frac{f_r}{f_t} \quad (5)$$

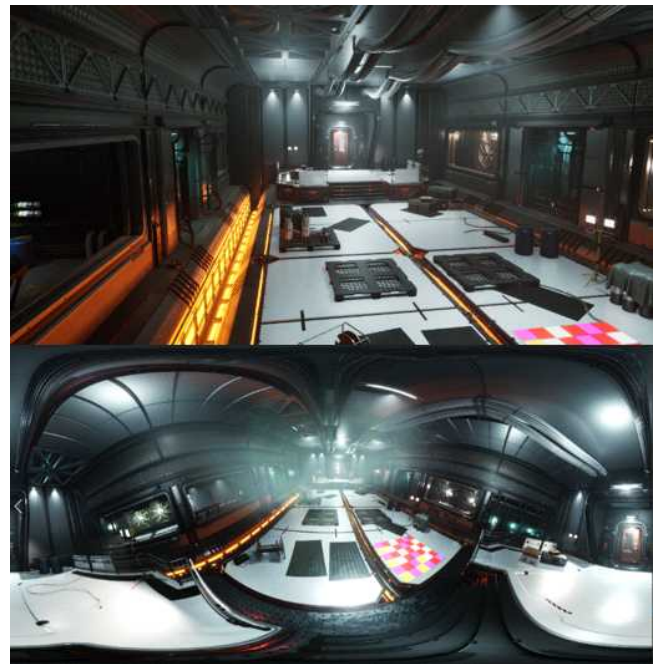
식 (5)은 전면 프레임과 후면 프레임에서의 검출된 특징점의 개수를 반영하여 보정된 카메라의 위치 추정 좌표 X_m 을 추정하는 식이다. 좋은 특징점의 개수가 많을수록 상대적으로 높은 정확도를 보이기 때문에, 상대적으로 많은 특징점을 가지는 경우 더욱 높은 가중치를 주는 방식으로 위치를 보정한다. 식 (5)의 X_f, X_r 은 각각 전면 프레임과 후면 프레임의 위치 추정 좌표이며 f_f, f_r, f_t 은 각각 전면 프레

임, 후면 프레임, 전후면 모두에서의 검출된 특징점의 개수이다. 본 시스템에서는 전면 반구에 대한 프레임인 X_f 을 기준이 되는 방향으로 삼았다. 따라서, X_r 의 경우는 반대 방향을 가리키기 때문에 y축을 기준으로 180° 회전시켜 좌표를 연산하였다.

Tracking 과정에 대한 모든 프로세스가 완료된 후에는 전면 프레임과 후면 프레임에 대하여 3D Map Point를 생성하는 과정인 Mapping 과정을 각각 수행한다. 마지막으로 Mapping이 완료된 후 Loop Closing 과정을 통해 최적화를 수행하여 더욱 정확한 위치를 추정한다.

3. 실험 결과

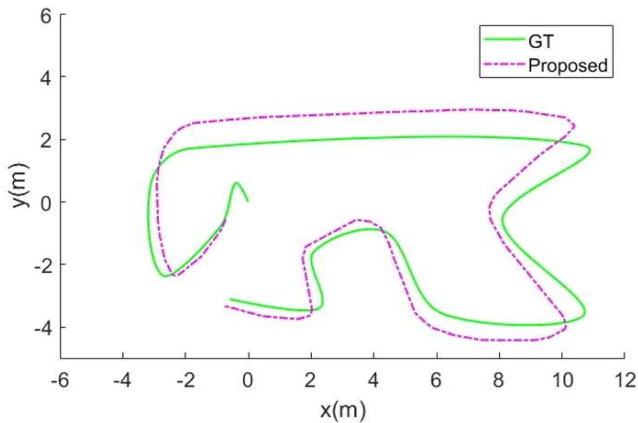
획득한 EUCM의 파라미터 값과 제안한 알고리즘이 올바르게 적용되는지 확인하기 위해 Tracking의 상태를 점검하였다. 만약 좋은 Tracking 결과가 나온다면 해당 시스템은 올바르게 설계되어 동작하는 것이다. 실험은 3.40GHz Intel(R) Core i7-6700 CPU와 32.0GB RAM의 사양을 가지는 Desktop PC를 통해 진행하였다.



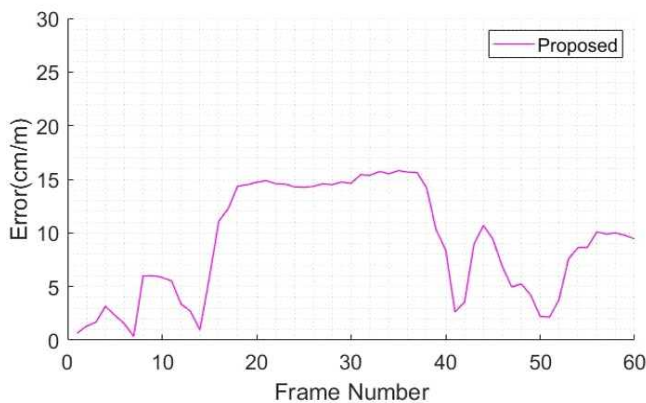
(그림 3) 실험을 위한 가상 환경

실험은 신뢰성 있는 검사를 위해 언리얼 엔진 상에서 설계한 3D 가상 환경에서 진행하였다. 언리얼 엔진 상에서 구면 파노라마 영상을 촬영할 수 있는 카메라를 원하는 경로로 이동시켜 구면 파노라마 비디오를 획득하고, 이동한 경로를 매 프레임마다 기

록하여 Ground Truth를 확보하였다.



(그림 4) Ground Truth와 추정 경로



(그림 5) 오차 그래프

(그림 4)는 가상 환경에서의 실측 경로와 제안하는 시스템의 경로를 비교한 것이다. 실측 경로는 언리얼 엔진에서의 길이 단위를 기준으로 하였다. 실험을 위한 시스템은 82개의 키프레임과 3235개의 Map Point를 가졌다. RMSE는 0.4448m이다. 오차는 7.99cm/m로 넓은 맵에 비해 준수한 추정 상태를 보였다. 이는 본 논문에서 계산한 파라미터 값과 제안하는 알고리즘이 올바르게 동작함을 보여주는 결과이다.

4. 결론

본 논문에서는 Liu등의 시스템에 구면 파노라마를 입력 비디오로 하여 모든 방향에 대한 시각적 데이터를 활용하는 SLAM 시스템을 제안하였다. 이 과정에서 구면 파노라마의 변형을 위한 Dual Fisheye Model을 활용하였고, 카메라 모델의 투영 파라미터의 공식화를 통해 정확한 계산을 보였다. 본 논문의 아이디어는 정확한 투영 법칙을 통해 투

영 파라미터를 획득하고, 시야각의 손실 없이 모든 방향의 시야각을 확보하는 시스템을 설계했다는 점에서 우수성이 있다.

Acknowledgments

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. NRF-2019R1F1A1060828).

참고문헌

- [1] Z. Zhang, H. Rebecq, C. Forster, D. Scaramuzza "Benefit of large field-of-view cameras for visual odometry" in Proc. IEEE Int. Conf. Robot. Automat., pp. 801 - 808, 2016.
- [2] David Caruso, Jakob Engel, Daniel Cremers "Large-Scale Direct SLAM for Omnidirectional Cameras" In Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp.141-148, 28 September-2 October 2015.
- [3] J. Engel, T. Schops, D. Cremers "LSD-SLAM: Large-scale direct monocular SLAM" In Proceedings of European Conference on Computer Vision, pp.834-849, September 2014.
- [4] Alejandro Rituerto, Luis Puig, J. J. Guerrero "Comparison of omnidirectional and conventional monocular systems for visual SLAM" 10th OMNIVIS with RSS, 2010.
- [5] Shuoyuan Liu, Peng Guo, Lihui Feng, Aiyang Yang "Accurate and Robust Monocular SLAM with Omnidirectional Cameras" Sensors, vol. 19(20), 2019.
- [6] Raul Mur-Artal, Juan D. Tardos "ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras" IEEE TRANSACTIONS ON ROBOTICS, vol. 33, no. 5, pp. 1255-1262, 2017.
- [7] Bogdan Khomutenko, Gaetan Garcia, Philippe Martinet "An enhanced unified camera model" IEEE Robotics and Automation Letters, vol. 1, no. 1, pp. 137-144, 2016.