

구면 PTAM의 구현을 위한 카메라 모델 설계

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

Design of Camera Model for Implementation of Spherical PTAM

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

요 약

시각적 환경 인식을 위하여 PTAM 연구가 활발히 이루어지고 있다. 최근 모든 방향의 시야각을 제공하는 구면 비디오를 위한 연구로 확장되고 있다. 기존의 구면 SLAM 방법은 Unified Sphere Model을 사용하며 앞면 시야각만 제공할 수 있는 한계가 있다. 본 논문에서는 구면 비디오를 위한 PTAM의 구현을 위한 카메라 모델을 제시한다. 제안된 카메라 모델은 핀홀 투영 카메라에 기반한 듀얼 영상 평면을 사용한다. 제안 방법은 앞면 시야각에 제약되지 않으며 전체 시야각을 지원한다. 또한 구면 비디오의 PTAM 적용 과정에서 평면 연산식을 직접 적용할 수 있는 장점이 있다.

1. 서론

구면 파노라마 이미지는 전 방향에 대한 정보를 한 장의 이미지로 표현한다. 이 이미지는 가로세로 2대 1의 비율을 가지는 이미지 평면에 구면 좌표계를 기반으로 이미지를 표현하기 때문에 이미지의 적도 라인에서 멀어질수록 급격하게 왜곡이 늘어난다. 구면 파노라마 이미지는 왜곡량이 많을 뿐더러 균일하지 않기 때문에 일반적으로 다른 형태의 카메라 모델을 설계하여 재투영하는 과정을 거친다.

SLAM(Simultaneous Localization and Mapping)은 미지의 영역에 대한 이미지의 특징점을 활용하여 지속적으로 맵을 그려나가 실시간으로 지형을 파악하고 카메라의 위치를 추정하는 기술이다. 구면 파노라마 이미지는 한 장의 이미지로도 주변의 거의 모든 정보를 분석할 수 있는 장점 때문에 과거부터 구면 파노라마 이미지를 입력으로 하는 SLAM에 대한 연구가 활발히 이어졌다.

David Caruso등[1]은 Fisheye Camera를 이용하여 Unified Omnidirectional Model에 대한 직접적인 공식화를 통해 Large-Scale Direct SLAM을 구현하였다. David Valiente등[2]은 이동형 로봇에 Fisheye Camera를 부착하여 로봇의 시스템 정보 매트릭스를 활용하여 불안정한 실측 정보를 보완

SLAM을 구현하였다. Alejandro Rituerto등[3]은 EKF(Extended Kalman Filter)를 활용하여 전 방향 카메라를 이용한 SLAM을 구현하였다. Jianfeng Li 등[4]은 구면 좌표계를 사용하는 구면 투영 모델을 설계하여 실내 환경에 적합한 SLAM 기능을 구현하였다.

PTAM[5]은 SLAM에서 파생된 기술로, Tracking과 Mapping을 병렬적으로 수행한다. PTAM은 상대적으로 수행 시간이 짧은 Tracking은 매 프레임마다 수행하여 정확한 카메라 위치 추정 데이터를 제공하고, 수행 시간이 긴 Mapping은 특정한 조건을 만족시키는 좋은 Frame에 한하여 수행한다. PTAM은 이와 같은 방법으로 수행 시간을 대폭 감소시켜 AR Workspace에 적합한 SLAM 기술이다.

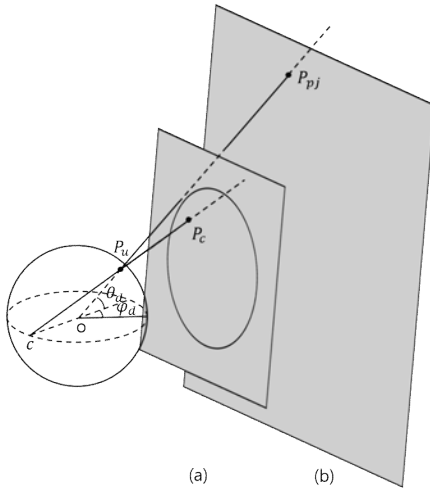
기존의 구면 SLAM의 구현을 위한 방법은 Unified Sphere Model을 사용하며 앞면 시야각만 제공할 수 있는 한계가 있다. 본 논문은 Spherical Panorama Video를 입력 이미지로 하는 PTAM의 구현을 위해 전 방향을 모두 활용할 수 있는 카메라 모델의 설계 방법을 제안한다. 본 논문에서 제안하는 카메라 모델은 구형의 이미지의 전면과 후면 반구를 각각 투영시킨 듀얼 영상 평면을 사용한다. 이는 전방과 후방 모두 시야각에 제약이 없기 때문에

전체 시야각을 지원한다. 또한 기존의 PTAM과 동일한 핀홀 모델을 기반으로 설계함으로써 PTAM에 적용하는 과정에서 평면 연산식을 직접 사용할 수 있는 장점이 있다.

2. 구면 PTAM을 위한 카메라 모델

핀홀 모델은 가장 잘 알려졌으며, 가장 많은 연구가 이루어지고 있는 카메라 모델이다. 따라서 핀홀 모델에 기반한 투영 모델은 다양한 기존의 시스템과 호환을 이룰 수 있다. 본 장에서는 구면 파노라마 이미지를 핀홀 모델에 투영하고, 최대한의 영역을 투영할 수 있는 내부 파라미터를 설계하는 방법을 제안한다.

2.1 카메라 모델 설계



(그림 1) 카메라 투영 모델: (a) Unified Sphere Model, (b) Pinhole Model

실제 세계의 한 점 $P_w = [x_w, y_w, z_w]^T$ 을 깊이 z_{pj} 인 핀홀 투영 평면상의 한 점 $P_{pj} = [x_{pj}, y_{pj}, z_{pj}]^T$ 로 투영하기 위한 공식은 식 (1)과 같다.

$$P_{pj} = AC \begin{bmatrix} P_w \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & scf_x & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (1)$$

여기서 A 는 카메라의 내부 파라미터이며, f_x 와 f_y 는 초점거리, c_x 와 c_y 는 주점, scf_x 는 비대칭계수이다. C 은 카메라의 외부 파라미터로 회전벡터 행렬과 평행이동 변환 행렬의 조합이다. 이는 실제 세

계의 한 점과의 변환 공식이므로, 이미 얻어진 구면 파노라마 이미지에서의 한 점 $P_s = [x_s, y_s]^T$ 을 핀홀 투영 평면상의 한 점 P_{pj} 로 변환하는 방법을 알아야 한다. 이 변환을 위해서는 데카르트 좌표계에서 벗어나 극 좌표계에서의 계산이 필요하다. 우선, 구면 파노라마 이미지의 정 중앙 지점을 경도(ϕ)와 위도(θ)가 0인 지점이라 하고 이미지의 크기를 S 라 할 때 P_s 을 극 좌표계에서의 점 $P_l = [\phi_l, \theta_l]^T$ 로 변환해야 한다.

$$P_l = \begin{bmatrix} x_s 2\pi / S_x - \pi \\ y_s \pi / S_y - \pi/2 \end{bmatrix}, \quad (2)$$

그 후 P_l 은 유클리드 공간에서 3차원 좌표를 갖는 한 점 $P_u = [x_u, y_u, z_u]^T$ 로 변환이 가능하다.

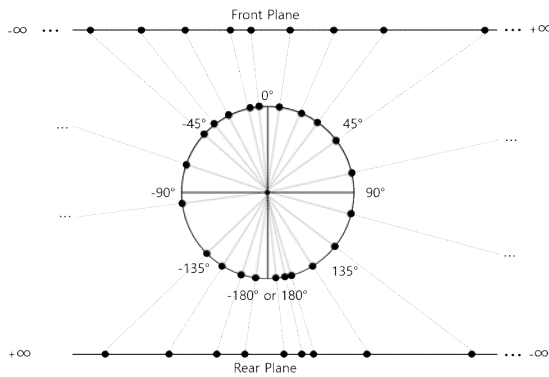
$$P_u = \begin{bmatrix} \cos \theta_d \cos \phi_d \\ \cos \theta_d \sin \phi_d \\ \sin \theta_d \end{bmatrix}, \quad (3)$$

P_u 은 반지름이 1인 3차원 구에서의 좌표이다. 구면 파노라마 이미지상의 모든 점을 위의 과정으로 투영하여 원하는 각도로 선형 회전 변환을 하고 한 축을 제거한 후 적절한 배율을 곱해준다면 Fisheye Image가 된다. 따라서 Fisheye Image의 경우 여기서부터 투영 과정을 수행하면 된다. 핀홀 모델에서는 구면 파노라마 이미지의 모든 영역을 투영할 수 없으므로 투영하고자 하는 부분의 경도와 위도를 정해 부분적인 투영을 해야 한다. 투영하고자 하는 부분의 경도와 위도를 각각 ϕ_r , θ_r 라 할 때 식 (4), (5)와 같은 방법으로 핀홀 투영 이미지상의 점 $P_{pj} = [x_{pj}, y_{pj}, z_{pj}]^T$ 을 구할 수 있다.

$$P_{ru} = \begin{bmatrix} \cos \theta_r & 0 & \sin \theta_r \\ 0 & 1 & 0 \\ -\sin \theta_r & 0 & \cos \theta_r \end{bmatrix} \begin{bmatrix} \cos \phi_r - \sin \phi_r & 0 \\ \sin \phi_r & \cos \phi_r & 0 \\ 0 & 0 & 1 \end{bmatrix} P_u, \quad (4)$$

$$P_{pj} = s \left[\frac{y_{ru}}{x_{ru}}, \frac{z_{ru}}{x_{ru}}, 1 \right]^T. \quad (5)$$

이렇게 구해진 P_{pj} 은 적절한 Z축 깊이 s 를 지정함으로써 육안으로 확인할 수 있는 이미지로 사용이 가능하다.

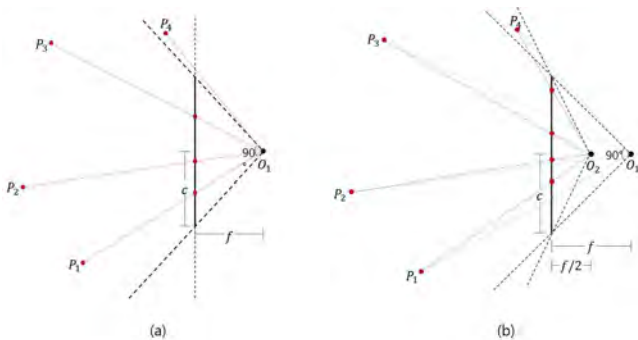


(그림 2) 듀얼 영상 평면 투영

본 논문에서 제안하는 시스템은 구의 전면 반구와 후면 반구를 투영시킨 두 장의 핀홀 투영 이미지를 획득할 것이므로 식 (4)에서 (ϕ_r, θ_r) 에 각각 $(0, 0)$, $(\pi, 0)$ 을 대입한 투영 이미지를 두 장 획득한다.

2.2 내부 파라미터 설계

일반적인 핀홀 투영 모델에서의 투영 구조는 식 (1)과 같다. 이 중 내부 파라미터 A 은 캘리브레이션을 통해 구하게 된다. 우선 2.1과 같은 방법으로 구면 파노라마 이미지를 화각이 90° 인 핀홀 카메라에 투영하여 얻어낸 평면 영상이 있다고 가정하자. 해당 영상의 캘리브레이션을 통해 내부 파라미터를 얻는다면 (그림 3)과 같은 투영 법칙이 성립할 것이다.



(그림 3) 초점거리에 따른 투영 법칙의 변화 : (a) 90° 화각 투영, (b) 초점거리가 변했을 때의 투영

(그림 3)의 (a)에서 점 P_1, P_2, P_3 은 이미지 평면에 도달해 픽셀로 자리 잡지만, P_4 의 경우는 90° 의 각에 들어오지 않아 이미지 평면에 자리 잡지 못한다. 하지만, 만약 모든 내부 파라미터가 같은 채 초점거리 f 가 짧아진다면 다른 결과를 가져온다. (그림 3)의 (b)는 다른 내부 파라미터는 그대로 둔 채 초점거리를 두 배 줄였을 때의 모습이다. 점 $P_1,$

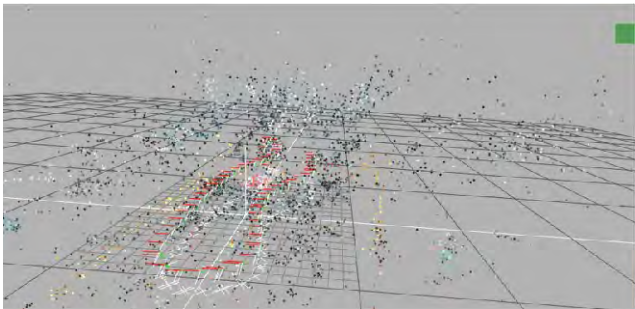
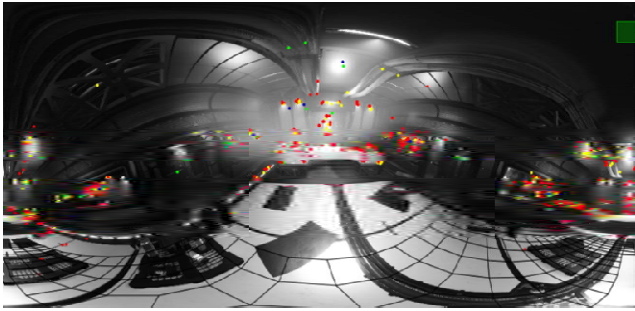
P_2, P_3 의 투영 지점은 중심에 가까워졌으며, 초점거리를 f 로 설정하였을 때 이미지 평면 안에 들어오지 않았던 P_4 역시 이미지 평면에 들어온 것을 볼 수 있다. 이는 화각이 더욱 넓어진 것을 뜻한다. 이때, 카메라 화각을 θ 라 할 때, $\cos(\theta/2) = f/c$ 이므로, θ 와 초점거리 f 사이에 다음과 같은 식이 성립한다.

$$\theta = 2\arccos\left(\frac{f}{c}\right). \quad (6)$$

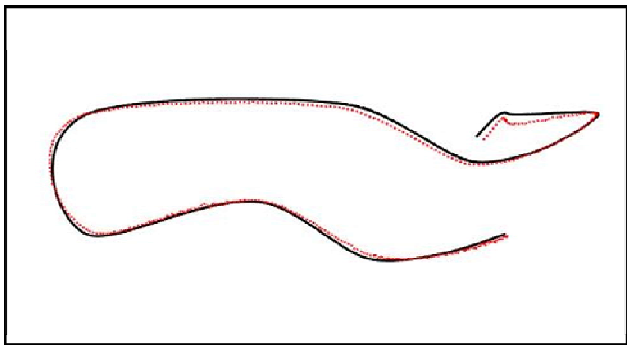
따라서, f 가 짧아지면 카메라 화각은 늘어나게 되고, f 가 0에 수렴하면 카메라 화각은 180° 가 된다. 만약 $f = n$ 일 때 화각이 90° 라면, $f = n/10$ 일 때 화각은 약 172° 가 나오고, $f = n/20$ 일 때 화각은 약 176° 가 나온다.

3. 실험 결과

카메라 투영 모델이 올바르게 적용되는지 확인하기 위해 PTAM에 적용시켜 Tracking 상태를 점검하였다. 만약 좋은 Tracking 결과가 나온다면 해당 투영 모델은 올바르게 적용되어 동작하는 것이다. 실험은 3.40GHz Intel(R) Core i7-6700 CPU와 32.0GB RAM의 사양을 가지는 Desktop PC를 통해 진행했으며, Windows 운영체제 환경에서 진행하였다. PTAM은 매 프레임 프레임마다 영상 정보를 포함한 다양한 정보를 함축한 Keyframe 구조체를 생성하고 이것을 기반으로 동작하는데, 이 때 한 개의 Keyframe은 한 장의 프레임 이미지만을 사용한다. 본 논문에서 제안하는 카메라 투영 모델은 한 프레임 당 전면과 후면 두 장의 이미지를 가지기 때문에 이에 맞게 한 프레임 당 전면과 후면에 관한 두 번의 처리를 할 수 있도록 PTAM의 내부 코드를 일부 변형하였다. 또한, PTAM 내부에 들어가는 많은 연산은 본 논문에서 제안하는 투영을 진행한 후 하였지만, 초광각의 핀홀 평면을 이미지로 표현하는 것은 가시성이 떨어지며 그 사이즈가 매우 크기 때문에 최종적인 출력 이미지와 그 위에 표시되는 특징점 등은 다시 구면 파노라마 형태로 투영하여 보여주도록 설계하였다. 실험은 정확한 Ground Truth를 획득하기 위해 언리얼 엔진 상에서 제작한 3D 가상 환경에서 진행하였다.



(그림 4) 가상 세계의 프레임과 3D 맵



(그림 5) Ground Truth(검은 선)와 추정 경로(빨간 선)

(그림 4)는 3D 가상 환경에서 촬영한 비디오를 통해 실측 경로와 구면 PTAM이 추정한 경로를 비교한 것이다. 실험을 위한 시스템은 62개의 키프레임과 5802개의 Map Point를 가졌다. 실측 경로는 언리얼 엔진에서의 길이 단위를 기준으로 측정하였다. 실측 경로와의 RMSE는 0.0770m이다. 오차는 1.38cm/m로 좋은 위치 추정 상태를 보였다. 이는 카메라 투영 모델이 올바르게 설계되어 동작함을 보여주는 결과이다.

4. 결론

본 논문에서는 AR 시스템에 적합한 구면 PTAM 구현을 위한 카메라 모델 설계 방법을 제안하였다. 이 과정에서 높은 수행 속도를 유지하였으며, 정확한 동작을 통해 PTAM의 위치 추정 정확도를 유지

하였다. 본 논문의 아이디어는 기존의 PTAM과 같은 투영 모델로 변환함으로써 호환성이 우수하고 정확한 법칙에 의해 동작하는 모델을 설계했다는 점에서 독창성이 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] 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.
- [2] David Valiente, Arturo Gil, Luis Paya, Jose M. Sebastian, Oscar Reinoso “Robust Visual Localization with Dynamic Uncertainty Management in Omnidirectional SLAM” Applied Science, 1294, 2017.
- [3] Alejandro Rituerto, Luis Puig, J. J. Guerrero “Comparison of omnidirectional and conventional monocular systems for visual SLAM” 10th OMNIVIS with RSS, 2010.
- [4] Jianfeng Li, Xiaowei Wang, Shigang Li “Spherical-Model-Based SLAM on Full-View Images for Indoor Environments” Applied Science, 2268, 2018.
- [5] Georg Klein, David Murray “Parallel Tracking and Mapping for Small AR Workspaces” In Proceedings of the 6th IEEE and ACM ISMAR, Vol. 17, No. 4, pp.225-234, 13-16 November 2007.