SIFT와 MOPS를 사용한 물체의 3차원 윤곽선 구성

박상현[O1] 이정욱[2] 임성규[2] 백두권[†1]
고려대학교[1] 건국대학교[2]
windscar2@gmail.com, ljo7@konkuk.ac.kr, hyub98@naver.com, baikdookwon@gmail.com

Reconstruction of 3D outlines of objects using SIFT and MOPS

Sang-Heon Park[O1] Jeong-Oog Lee[2] Sung-Gyu Im[2] Doo-Kwon Baik[†1]
Korea University[1] Konkuk University[2]

1. 서론

최근 3차원 물체의 인식에 대한 연구가 많이 이루어지고 있다. 지금까지 3차원 물체 인식에 관한 연구는 주로 단일 카메라나 스테레오 카메라, Laser 센서를 병행하여 사용하는 방법으로 이루어 지고 있다[1][2]. 스테레오 카메라와 Laser 센서의 경우 고가이고, 무게가 많이 나가는 단점이 있기 때문에 특정한 목적에 사용되거나 하중에 구해 받지 않는 로봇 등에서만 주로 사용된다. 단일 카메라의 경우 3차원 정보를 처리하는데 계산량이 많아 지는 단점이 있는 반면 가격이 낮고 무게가 가볍기 때문에 최근 이슈가 되고 있는 MAV(Micro Air Vehicle) 등 소형 비행 로봇의 연구에서 많이 사용되고 있다.

지금까지 단일 카메라를 사용하여 물체를 인식하는 논문은 대부분 데이터베이스에 존재하는 물체를 비교하여 형태를 재구성하거나 3차원 점 성분만으로 표현하여 물체의 대체적인 윤곽을 인식하는 논문이 많았다[3][4]. 또한 데이터베이스 상에 존재하지 않는 물체를 3차원적으로 재구성한 논문[5]은 물체를 정확하게 표현 가능하긴 하지만 계산량이 많은 단점이 있었다.

따라서 본 논문에서는 단일 카메라를 사용하여 데이터베이스에 존재하지 않는 물체를 빠르고 정확하게 인식하는데 목적을 갖고 다음 두 가지 목표에 중점을 둔다. 첫 번째, 물체의 빠른 인식을 위해 영상에서 물체들의 대략적인 윤곽선만 추출하여 외형을 파악한다. 두 번째, 윤곽선만으로 구성할 때 손실된 물체 내부의 정보를 SIFT 특징점의 거리정보와 윤곽선의 거리정보를 비교하여 물체 내부가 열린 공간인지 닫힌 공간인지 파악하여 복원한다.

2. SIFT와 MOP를 이용한 3차원 윤곽선 구성

본 논문에서 제한하는 방법은 두 가지 데이터 처리 과정을 병행하여 처리 한 후 두 처리결과를 병합하여 영상 전체 물체들의 3차원 윤곽선을 구성하는 것이다.

첫 번째 처리 과정은 영상의 물체들의 윤곽선을 빠르게 구성하는 방법이다. MOPS(Multi-Scale Oriented Patches)[6]를 사용하여 물체의 코너점의 MOPS 특징점을 얻는다. 이렇게 얻어진 두 영상의 MOPS 특징점을 매칭해 3차원 공간 정보를 포함하는 3차원 MOPS 특징점을 추출한다. 그 후 물체들의 대략적인 윤곽선 정보를 구성하기 위해 각 특징점을 엣지 정보와 병합하여 영상의 전체 물체들의 대략적인 윤곽선 정보를 구성한다. 이렇게 얻어진 정보는 윤곽선 사이의 내부 정보가 손실되었기 때문에 실제 내부 공간이 열린 공간인지 닫힌 공간인지 분간할 수 있는 정보가 존재하지 않는다. 따라서 다음 처리 과정을 통해서 내부 정보를 보완할 수 있는 정보를 얻을 수 있다.

두 번째 처리 과정은 윤곽선 사이의 내부 공간 정보를 얻기 위해 SIFT(Scale Invariant Feature Transform)[5] 특징점을 사용하는 방법이다. 두 영상을 SIFT로 특징점을 추출한다. SIFT 특징점은 물체 뿐만 아니라 벽이나 바닥 등 색이나 빛의 변화가 확실한 지점에서 추출되기 때문에 영상 전체 영역에 걸쳐 추출되기 때문에 전체적인 3차원 공간좌표를 점 성분으로 구성할 수 있다. 이렇게 얻어진 두 영상의 SIFT 특징점을 매칭하여 3차원 공간정보가 포함된 3차원 SIFT 특징점을 추출할 수 있다.

마지막 처리 과정은 첫 번째 처리과정을 통해서 얻어진 대략적인 윤곽선의 내부 정보를 3차원 SIFT 특징점으로 보완하여 상세한 정보를 얻는 과정이다. 한 면은 구성하는 윤곽선 내부에 3차원 SIFT

이 논문은 '2 단계 BK21 사업'에 의해 지원 되었음을 밝힙니다.

† 교신저자 (Corresponding Author) †

특징점 정보가 존재할 경우, 윤곽선의 거리정보와 SIFT 특징점 거리정보를 비교할 때 한 물체의 앞 면에 존재하는 임의의 3차원 MOPS 특징점을 M으로 표현하며 임의의 한 윤곽선 상에 존재한다. 그리고 임의의 3차원 SIFT 특징점을 S, 거리정보를 D, Threshold를 T라고 하면 다음 3가지 식으로 표현할 수 있다.

$$|D_M - D_S| \le T_1 \quad (1)$$

$$T_2 \ge |D_M - D_S| > T_1 \quad (2)$$

$$|D_M - D_S| > T_2 \quad (3)$$

Threshold T_1은 1~2cm 정도의 거리이고, T_2는 10~30cm 정도의 비교적 가까운 거리이다.

그림 1 S와 M의 거리정보가 비슷할 경우

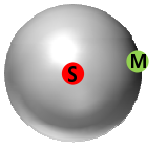

그림 3 M보다 S의 거리정보가 조금 더 가까울 경우

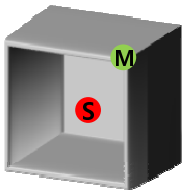

그림 2 M보다 S의 거리정보가 조금 더 멀 경우

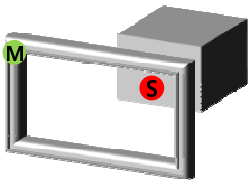

그림 4 M보다 S의 거리정보가 많이 멀 경우

식 (1)은 윤곽선 거리정보와 SIFT 특징점 거리정보 차의 절대값이 Threshold T_1안에 드는 경우이다. 그림 1처럼 그 면이 닫혀 있고, 평평한 면으로 판단할 수 있다. 식 (2)는 윤곽선 거리정보보다 SIFT 특징점 거리정보가 큰 차이가 나지 않는 경우이다. 식 (2)의 경우, 그림 2처럼 그 면은 볼록한 타원형이나 원 뿔 형태로 존재하거나, 그림 3과 같이 서랍장 같은 형태의 물체나 같이 열린 공간 뒤에 바로 다른 물체 존재하여 겹쳐진 경우로 판단 할 수 있다. 식 (3)은 윤곽선의 거리정보보다 SIFT 특징점의 거리정보가 확실하게 멀 경우이다. 이 때, 그림 4와 같이 해당 면은 열린 공간으로 판단할 수 있다. 열린 공간 뒤로 다른 물체의 특징점이 추출된 경우이다.

3. 결론

본 논문은 단일 카메라를 사용하여 물체의 윤곽선 추출 및 손실된 내부 정보를 보완하는 프로세스를 소개하였다. 물체의 윤곽선 추출에 걸린 시간은 평균 총 0.540s로 물체의 자세한 3차원 정보를 복원하는게 아니기 때문에 계산 시간이 적고 윤곽선 내 손실된 내부 정보를 3차원 SIFT 특징점을 사용하여 보완하기 때문에 물체의 인식률을 크게 향상 시킬 수 있다.

본 연구는 물체의 윤곽선에 위치한 MOPS 특징점의 거리정보와 그 내부에 위치한 SIFT 특징점을 비교하여 물체의 볼륨 및 각 물체간의 거리를 표현하지만 3차원 형태로 표현 및 저장을 못하는 한계가 있다. 현재 진행하고 있는 연구는 각 물체에 해당하는 윤곽선과 SIFT 특징점을 그룹화하여 3차원 맵을 구성하는 연구가 진행 중이며 실제 MAV에 적용하여 후 전방에 놓인 물체 회피에 효과적인 판단을 할 수 있도록 할 예정이다. 또한, 수행시간에 다소 지장을 줄 수 있는 SIFT 알고리즘을 개선하여 MAV에 탑재 시 실시간 판단이 가능하도록 할 연구를 계획하고 있다.

4. 참고문헌

[1] Andrew J. Davison, Nobuyuki Kita, “3D Simultaneous Localisation and Map-Building Using Active Vision for a Robot Moving on Undulating Terrain”, IEEE Conference on Computer Vision and Pattern Recognition, December 2001.
[2] M. Garcia, A. Solanas, “3D simultaneous localization and modeling from stereo vision”, in IEEE Conference on Robotics and Automation, pp. 847-853, 2004.
[3] D. Lowe, “Distinctive image features from scale-invariant keypoints”, In International Journal of Computer Vision, vol. 20, pp. 91-10, 2003.
[4] Se, S., Lowe, D. and Little, J., “Mobile robot localization and mapping with uncertainty using scale-invariant visual landmarks”, International Journal of Robotics Research, vol. 21 i8, pp. 735-758, 2002.
[5] Ankur Agarwal, Bill Triggs, “Recovering 3D Human Pose from Monocular Images”, IEEE Transactions on pattern analysis and machine intelligence, vol. 28, no. 1, January 2006.
[6] M. Brown, R. Szeliski, and S. Winder, "Multi-image matching using multi-scale oriented patches", Technical Report MSR-TR-2004-133, Microsoft Research, December 2004.