

자율 수중 로봇을 위한 사실적인 실시간 고밀도 3차원 Mesh 지도 작성

Photorealistic Real-Time Dense 3D Mesh Mapping for AUV

이정우¹·조영근[†]
Jungwoo Lee¹, Younggun Cho[†]

Abstract: This paper proposes a photorealistic real-time dense 3D mapping system that utilizes a neural network-based image enhancement method and mesh-based map representation. Due to the characteristics of the underwater environment, where problems such as hazing and low contrast occur, it is hard to apply conventional simultaneous localization and mapping (SLAM) methods. At the same time, the behavior of Autonomous Underwater Vehicle (AUV) is computationally constrained. In this paper, we utilize a neural network-based image enhancement method to improve pose estimation and mapping quality and apply a sliding window-based mesh expansion method to enable lightweight, fast, and photorealistic mapping. To validate our results, we utilize real-world and indoor synthetic datasets. We performed qualitative validation with the real-world dataset and quantitative validation by modeling images from the indoor synthetic dataset as underwater scenes.

Keywords: Underwater Robotics, Dense Mapping, Autonomous Underwater Vehicle

1. 서 론

최근 해양 기술의 발전과 함께 수중 환경에서 로봇을 이용해 다양한 업무를 수행하기 위한 로봇의 위치 인식 및 지도 작성 기술인 SLAM 연구가 활발히 수행되고 있다. 이러한 연구는 해저 건설 및 구조물 관리, 해양 환경 분석, 해저 지도 작성 등에 다양하게 활용되고 있다. 특히 직접 인력을 투입하기 어려운 위험한 해양 지형과 안전을 위해 상시 관제가 필요한 구조물 등에 대해 자율 수중 로봇(AUV)의 활용이 강조된다. 본 논문에서는 수중 구조물의 균열을 검사하는 등의 정밀한 업무

수행을 위한 자율 수중 로봇의 강인한 위치 인식 및 사실적인 고밀도 지도 작성 시스템을 제안한다.

사람이 해석하기 쉬운 사실적인 지도 작성을 위해 영상 정보의 활용은 불가피하다. 하지만 수중 환경에서 영상 정보를 활용하는 것에는 어려움이 있다. 수중에서는 대기 환경과 다른 빛의 산란, 흡수 감쇠 현상에 의해 영상의 품질이 저하되는 문제가 발생한다. 색 왜곡과 영상의 대비가 감소하는 등의 영상 품질 저하 문제는 영상 정보를 사람이 해석하기 어렵게 할 뿐만 아니라 수중 로봇의 위치 인식과 지도 작성 성능에도 영향을 주게 된다. 이를 해결하기 위해 수중에서 촬영된 영상의 색 보정을 위한 다양한 방법들이 제안되었다^[1]. 최근에는 인공지능의 발전과 함께 신경망 학습을 통해 영상 개선을 수행하는 방법들이 많이 등장하였다^[2]. 본 논문에서 제안하는 시스템

Received : Jan. 19. 2024; Revised : Mar. 2. 2024; Accepted : Mar. 13. 2024

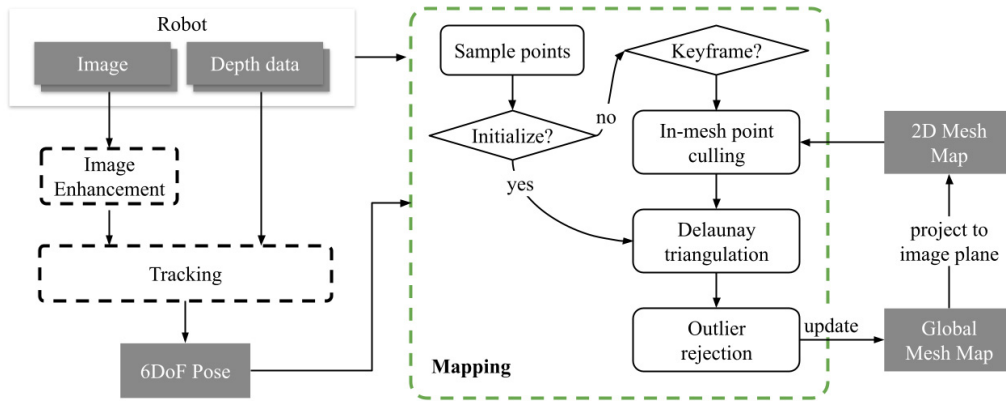
※ This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (RS-2023-00302589), Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No.2022-0-00448) and the Korea Institute of Marine Science and Technology Promotion (KIMST), funded by the Ministry of Oceans and Fisheries (20210562).

1. Master Student, Dept. Electrical and Computer Engineering, Inha University, Incheon, Korea (pihdsneirf@inha.edu)

† Associate Professor, Corresponding author: Dept. Electrical and Computer Engineering, Inha University, Incheon, Korea (yg.cho@inha.ac.kr)

[Table 1] Comparison of map representation

	Realistic	Speed	Difficulty
Pointcloud	Middle	Middle	Low
Voxel	Low	High	Middle
Mesh	High	High	High



[Fig. 1] System pipeline

은 신경망 기반의 영상 개선 방법을 활용하여 수중 로봇의 위치 인식 성능을 개선하고 동시에 사람이 해석하기 쉬운 지도 작성을 수행하였다.

사실적이고 사람이 해석하기 쉬운 지도 작성을 위해 고밀도 지도 작성이 필수적이고, 이를 위해 pointcloud^[3], voxel^[4], mesh^[5] 등의 지도 표현이 대표적으로 사용된다. Pointcloud를 이용해 지도를 표현하는 경우, 정보를 다루기 쉽고 특정 포인트에 대해서 사실적인 정보를 표현할 수 있지만, 연속적인 위치 정보를 표현하지 못한다는 단점이 있다. 반면 voxel을 활용하는 경우 미리 정의한 voxel의 크기에 따라서 구체적인 영역을 표현하는 정도를 조절할 수 있지만, 실시간 동작을 위해서는 사실적인 표현을 포기해야 되는 단점이 있다. 또 다른 접근으로 mesh 기반의 지도 표현 방법을 사용할 수 있다. Mesh는 정형화되지 않는 표현으로, 정보를 다루기 어렵지만 적은 수의 데이터를 사용하면서 동시에 texture mapping을 통해 사실적인 표현이 가능하다는 장점을 가지고 있다. [Table 1]을 참고하여 지도 표현 방법을 비교할 수 있다.

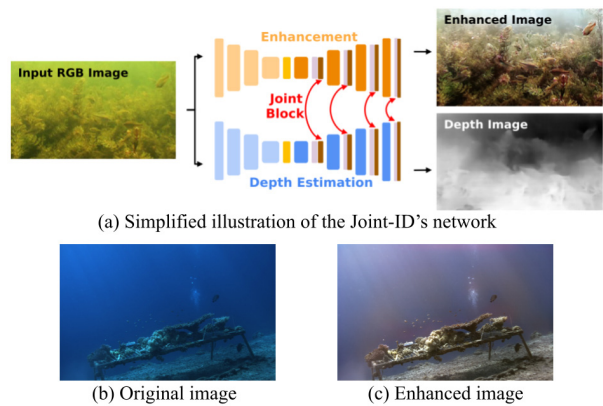
본 논문은 수중에서 강한 위치 인식 및 사실적인 지도 작성을 위한 시스템을 제안하고, 이는 다음과 같은 contribution을 가진다.

- 신경망 기반 수중 영상 개선 방법 적용을 통한 로봇의 위치 인식 성능 개선 및 지도 표현 개선.
- Mesh 기반의 표현 방법 사용을 통한 실시간 이미지 픽셀 단위 지도 작성.
- 실제 수중 환경을 촬영한 데이터셋을 통해 제안하는 방법의 정성적 검증 및 실내 합성 데이터셋을 이용한 정량적 검증.

2. 고밀도 3차원 Mesh 지도 작성

2.1 시스템 구성

[Fig. 1]은 제안하는 시스템의 흐름도를 보여준다. 먼저 로

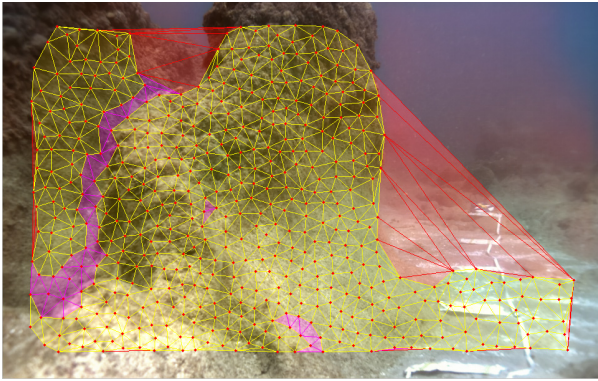


[Fig. 2] Illustration of Joint-ID and image enhancement result

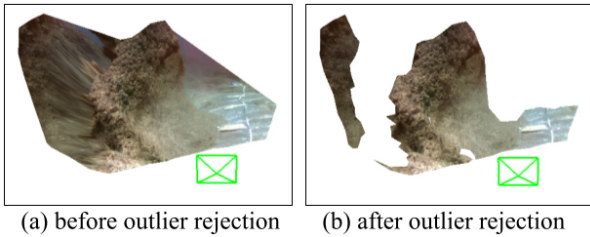
봇을 통해 취득한 수중 영상 개선을 수행한다. 본 논문에서는 수중 영상 개선을 위해 현재 가장 좋은 성능을 보이고 있는 Transformer 모델 기반 Joint-ID^[2]를 활용하였다. 영상 개선 결과는 [Fig. 2]의 (b), (c)를 통해 확인할 수 있다. 이후 개선된 영상 정보와 깊이 정보를 이용해 로봇의 위치를 추정한다. 위치 인식은 ORB-SLAM2^[6]의 방법을 기반으로 수행하였다. 최종적으로 개선된 영상, 깊이, 로봇의 위치와 자세 정보를 이용해 고밀도 3차원 mesh 지도 작성을 수행한다. 고밀도 3차원 mesh 지도 작성에 대해 자세한 내용은 2.3 절에서 설명할 것이다.

2.2 신경망 기반 수중 영상 개선

본 논문에서는 수중 영상 개선을 위해 Joint-ID^[2]를 활용하였다. Joint-ID는 Transformer 기반의 신경망으로, 영상 개선 혹은 깊이 정보 추정을 단일 목적으로 하는 기존의 다른 딥러닝 방법론들과 달리, 단일 수중 영상으로부터 영상 개선과 깊이 정보 추정을 동시에 수행하여 기존 방법론보다 훨씬 강력한 성능을 보임을 입증하였다. 신경망의 구조는 [Fig. 2]의 (a)에서 확인할 수 있다.



[Fig. 3] Point sampling and triangulation: The red points are extracted feature points. We performed Delaunay triangulation on the extracted points. The red triangle mesh is the one that is removed from the image plane. The purple triangle mesh is removed after the 3D projection. The yellow triangle mesh is the correct one in the map



(a) before outlier rejection (b) after outlier rejection

[Fig. 4] Generated 3D mesh map

2.3 고밀도 3차원 mesh 지도 생성

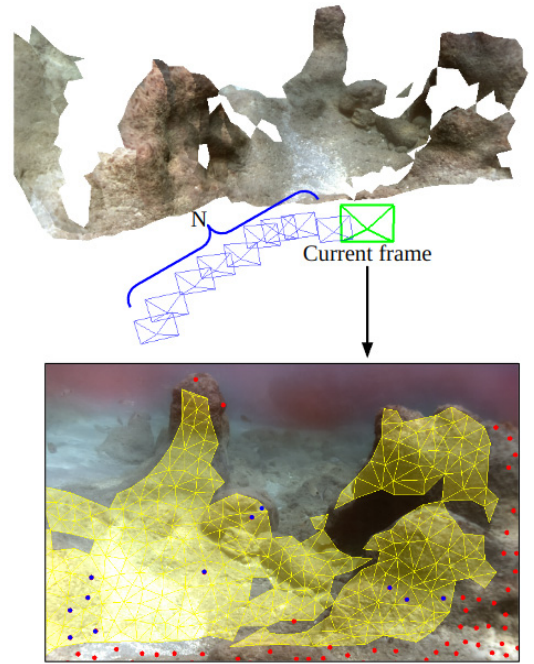
2.3.1 3차원 mesh 생성

제한하는 방법은 실시간 지도 작성을 위해 먼저 영상에서 포인트 샘플링을 수행한 후 이를 삼각형의 정점으로 하도록 Delaunay triangulation^[7]을 수행하여 2D mesh 정보를 생성한다. 본 논문의 실험에서는 포인트 샘플링의 방식으로 영상의 특징점을 추출^[8]하여 생성된 2D mesh 정보가 실제 환경의 평면정보를 최대한 반영할 수 있도록 하였다. 또한 Delaunay triangulation은 최대한 정삼각형에 가까운 삼각형을 생성하도록 하는 방법으로, 2D mesh의 정보가 3D로 투영했을 때에도 평면성과 영상의 텍스처를 최대한 유지할 수 있도록 하였다. [Fig. 3]에서 2D mesh 생성 결과를 확인할 수 있다.

이후 가지고 있는 깊이 정보를 활용해 2D mesh의 정점을 3차원으로 투영하여 3D mesh 지도를 생성한다. 결과는 [Fig. 4]의 (a)를 보면 확인할 수 있다.

2.3.2 이상치 제거

2D mesh를 생성하는 과정에서 실제세계의 3차원 평면성을 반영하지 않는 mesh가 생성될 수 있다. 본 논문에서는 2D



[Fig. 5] Mesh map expansion: (up) 3D mesh map generated from the previous N camera frame. (down) Yellow triangles are projected mesh from a global map. Red dots are sampled points found in the current frame that are outside the mesh. Blue points are sampled points that are above the existing mesh in the image plane but have a significant distance to the mesh in 3D

mesh와 3D mesh의 관점에서 이러한 이상치 제거를 수행한다. 먼저 2D mesh를 생성한 후 삼각형의 세 정점을 $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$ 라고 할 때, 아래의 식 (1)의 조건에 따라 삼각형 한 변의 길이가 특정한 값보다 긴 삼각형을 제거한다. 이는 [Fig. 3]에서 붉은 색 삼각형들을 의미한다.

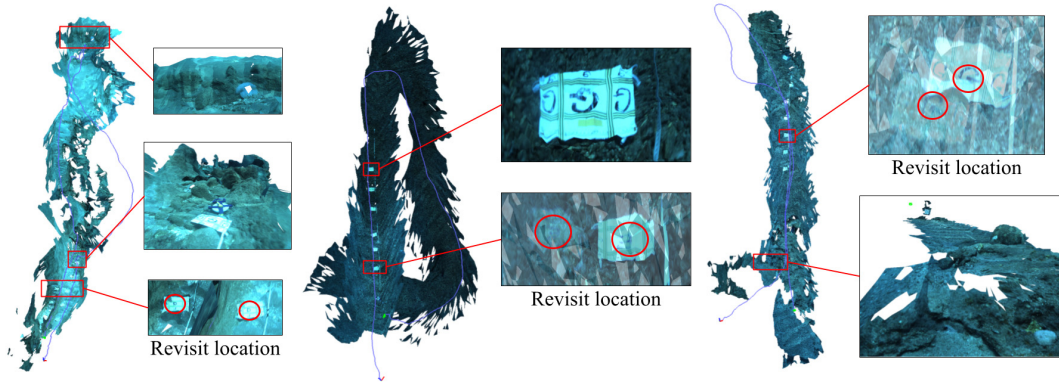
$$\max(|\mathbf{p}_2 - \mathbf{p}_1|, |\mathbf{p}_3 - \mathbf{p}_2|, |\mathbf{p}_1 - \mathbf{p}_3|) > l_p \quad (1)$$

여기서 l_p 는 2D mesh 삼각형 한 변의 최대 길이를 의미한다. 다음으로 mesh를 3차원으로 투영한 후 이상치 제거를 수행한다. 투영된 mesh의 세 정점을 $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ 라고 할 때, 먼저 2D mesh의 방식과 비슷하게 너무 큰 길이의 변을 가지는 삼각형을 제외 시켜준다. 이는 다음의 식 (2)를 활용한다.

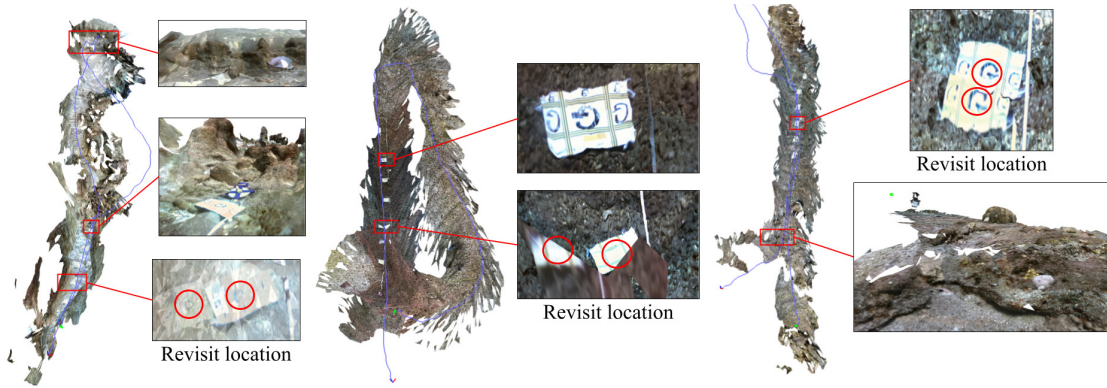
$$\max(|\mathbf{v}_2 - \mathbf{v}_1|, |\mathbf{v}_3 - \mathbf{v}_2|, |\mathbf{v}_1 - \mathbf{v}_3|) > l_v \quad (2)$$

다음으로 식 (3)을 사용하여 카메라에서부터 mesh의 중심까지 백터와 mesh의 법선 백터가 수직에 가까운 경우 제거 해준다. 이는 [Fig. 3]에서 보라색 삼각형들을 의미한다.

$$|\mathbf{v}_c \cdot \mathbf{n}| < d \quad (3)$$



[Fig. 6] Mapping result without image enhancement: (left) *FLSea-VI Canyons Flatiron*, (middle) *FLSea-VI Redseas Cross Pyramid Loop*, (right) *FLSea-VI Redseas Northeast Path*



[Fig. 7] Mapping result with image enhancement: (left) *FLSea-VI Canyons Flatiron*, (middle) *FLSea-VI Redseas Cross Pyramid Loop*, (right) *FLSea-VI Redseas Northeast Path*

여기서 $\mathbf{v}_c = (\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3) / |\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3|$ 이고, \mathbf{n} 은 mesh의 법선 벡터, l_v , d 는 임계 값을 의미한다. 이상치를 제거하기 전과 후의 3D mesh 지도는 [Fig. 4]의 (a)와 (b)를 통해 확인할 수 있다.

2.3.3 지도 확장

제안하는 방법은 지도를 실시간으로 확장하기 위해 먼저 이전 N 개 프레임에서 생성한 3D mesh 지도를 현재 로봇의 위치에서 이미지 평면에 투영한 후 지도를 확장하는 sliding window 방식을 사용한다. 이 때 투영된 정점에 대해 최소 거리 d_{min} 보다 먼 위치에 있도록 포인트 샘플링을 수행한다. 이때 얻은 점들은 다음의 2가지 경우 중 하나에 해당된다. (1) 기존 mesh 안에 있는 경우 (2) 기존 mesh 밖에 있는 경우. (2)의 경우에는 지도 확장을 위해 추출된 점의 정보를 저장해둔다. (1)의 경우 아래의 식 (4)와 같이 점과 평면사이의 거리를 계산해 특정 값 보다 차이가 많이 날 경우에만 값을 저장한다.

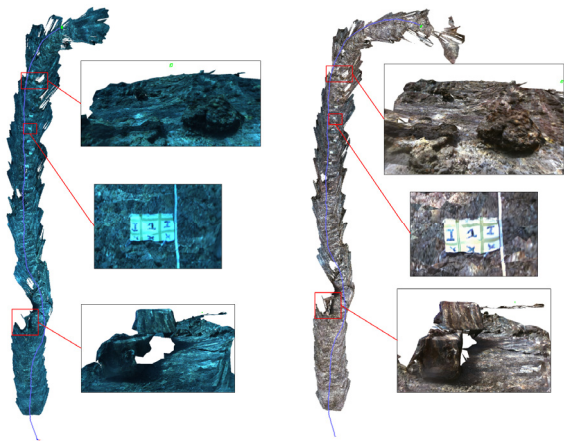
$$\frac{(\mathbf{v}_2 - \mathbf{v}_1) \times (\mathbf{v}_3 - \mathbf{v}_1)}{|\mathbf{v}_2 - \mathbf{v}_1 \times (\mathbf{v}_3 - \mathbf{v}_1)|} \cdot (\mathbf{P} - \mathbf{v}_1) < d_{min} \quad (4)$$

여기서 \mathbf{P} 는 특징 점의 3차원 위치를 의미한다. 최종적으로 얻은 기존의 정점들과 새로운 특징 점들을 포함하여 다시 Delaunay triangulation을 통해 지도 확장을 수행한다. [Fig. 5]를 통해 지도 확장 과정을 확인할 수 있다.

3. 실험

3.1 실험 환경

본 논문은 제안하는 방법의 정성적 검증을 위해 FLSea 데이터셋⁶⁾을 활용하였다. 이 데이터셋은 스테레오 비전으로 취득한 데이터와 카메라-관성 센서로 취득한 데이터를 제공한다. 대부분의 수중 데이터셋이 바닥을 바라보는 영상을 취득하는 반면 이 데이터셋은 전방을 바라보는 영상을 취득하였다. 또한 활용한 데이터셋은 상업 프로그램인 Agisoft Metashape을 이용해 Structure from Motion (SfM) 방식으로 생성한 3차원 정보를 깊이 영상의 형태로 제공한다. 본 논문에서는 데이터셋이 제공하는 영상과 SfM으로 구한 깊이 정보를 사용하여 제안하는 방법의 검증을 수행하였다. 추가적으로 제안하는 방법의 정량적



[Fig. 8] Mapping result for *FLSea-VI Redseas Big Dice Loop* sequence without image enhancement (left) and with image enhancement (right)

결과 비교를 수행하기 위해 실내 합성 데이터셋인 ICL-NUIM^[10]의 정보를 수중 영상으로 모델링하여 검증을 수행하였다.

실험에 사용된 CPU는 Intel Core i5-13400F이고, Ubuntu 20.04 환경에서 실험을 진행하였다.

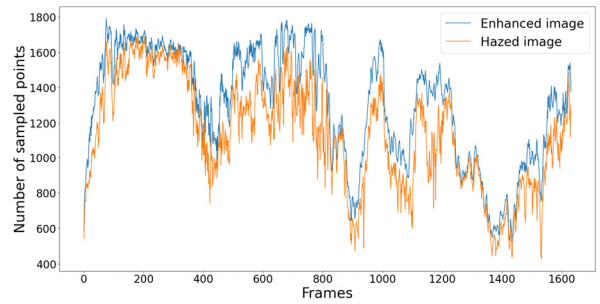
3.2 실험 결과

3.2.1 영상 개선에 따른 위치 인식 개선

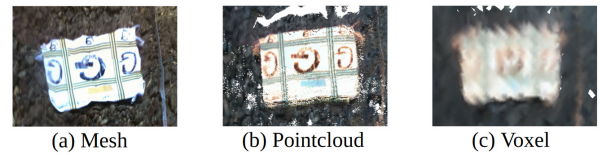
수중 환경의 특성 상 ground truth 경로를 구하기 힘들다. 활용하는 데이터셋에서도 ground truth 경로를 제공하고 있지 않기 때문에 영상 개선에 따른 위치 추정 개선에 대해 정성적인 검증을 진행하였다.

[Fig. 6], [Fig. 7]의 왼쪽 데이터 *FLSea-VI Canyons Flatiron*와 오른쪽 데이터 *FLSea-VI Redseas Northeast Path*는 역 루프 폐쇄(reverse loop closure)가 발생한다. [Fig. 6], [Fig. 7]의 “Revisit location”에서 붉은색 동그라미는 같은 표식물을 표현한 것으로, 각기 반대 방향을 바라보며 취득한 영상을 바탕으로 작성된 지도 결과이다. 결과를 보면 이미지 개선을 수행하지 않았을 때 작성된 지도에서 같은 표식물의 위치가 더 많이 차이나는 것을 확인할 수 있다. 이는 로봇의 위치 인식 과정에서 오차가 점차 누적되어 발생하는 drift error가 이미지 개선 후 줄어들었음을 의미한다.

[Fig. 6], [Fig. 7]의 가운데 데이터인 *FLSea-VI Redseas Cross Pyramid Loop*의 “Revisit location”을 확인해보면 인공 표식물의 위치가 크게 차이 나지 않는 것으로 확인된다. 이는 이미지 개선의 결과가 위치 인식 과정에서 특징점 추출 및 비교에 영향을 주지만 이미지 개선의 여부와 상관없이 같은 깊이 정보를 이용하였기 때문으로 생각된다. 만약 이미지 기반의 깊이 추정을 수행하는 경우에는 이미지 개선했을 때 위치 추정의



[Fig. 9] Number of sampled points for hazed and enhanced image



[Fig. 10] Comparison of qualitative results using artificial landmarks in the *FLSea-VI Redseas Cross Pyramid Loop* sequence. (a) Proposed method, (b) Pointcloud, (c) Voxblox

향상이 더 두드러지게 나타날 것이라고 기대할 수 있다.

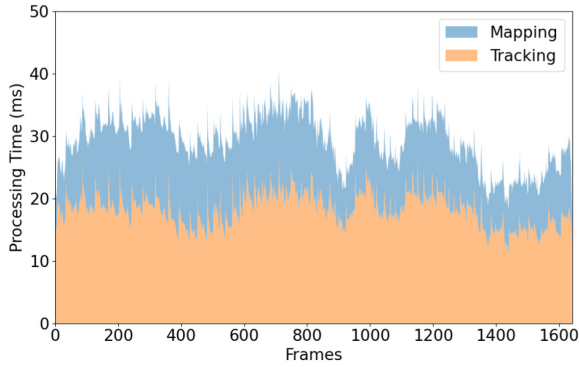
[Fig. 8]의 *FLSea-VI Redseas Big Dice Loop* 데이터에서는 중간에 로봇이 위치 인식을 실패하는 상황이 발생한다. 위치 인식을 실패하기 전까지 작성된 지도를 [Fig. 8]에서 확인할 수 있다. 로봇의 위치 인식 실패는 해당 데이터의 특정 지점에서 로봇의 빠른 회전 움직임으로 인해 특징점 추적을 실패했기 때문으로 예상된다. [Fig. 8]의 오른쪽 결과를 보면 이미지 개선을 수행하지 않았을 때보다 더 먼 위치까지 위치 인식을 성공한 모습을 확인할 수 있다. 이는 이미지 개선이 로봇의 강한 위치 인식에 도움을 준다는 것을 의미한다.

3.2.2 영상 개선에 따른 포인트 샘플링 성능

본 논문의 [Fig. 9]에서 영상 개선에 따른 포인트 샘플링의 성능 비교를 수행하였다. 제안하는 방법은 특징점 추출^[8] 방법을 기반으로 포인트 샘플링을 수행하기 때문에, 영상의 색 왜곡과 대비의 문제를 개선한 후 더 좋은 성능을 보여주게 된다. 포인트 샘플링의 성능을 비교하기 위해 *FLSea-VI Redseas Cross Pyramid Loop* 데이터의 영상에서 검출되는 특징점의 수를 영상 개선 전후로 검증하였다. 영상 개선 이후 검출되는 특징점의 수가 개선 전과 비교하여 약 16.1% 증가한 것을 확인하였다.

3.2.3 사실적인 지도작성

제안하는 방법은 mesh에 해당하는 텍스처를 입혀 사실적인 지도작성을 가능하도록 한다. 이는 자율 수중 로봇이 업무를 수행한 이후 사람이 결과를 해석하는 것에 있어 중요한 역할을 한다. 이미지 개선에 따른 사실적인 지도작성에 대해서 [Fig. 6], [Fig. 7], [Fig. 8]에서 확인할 수 있다.



[Fig. 11] Processing time for *FLSea-VI Redseas Cross Pyramid Loop* sequence: The orange area represents tracking time and the blue area represents mapping time

[Table 2] Number of point (voxel) in *FLSea-VI Redseas Cross Pyramid Loop* sequence

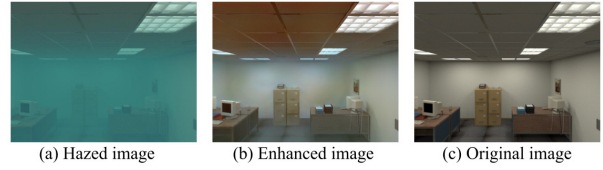
	Mesh	Pointcloud	Voxel (1 cm)
Number of point (voxel)	17,538	22,833,823	13,997

본 논문에서는 사람이 해석 가능한 수준의 지도 작성 여부를 확인하기 위해 *FLSea-VI Redseas Cross Pyramid Loop* 데이터에 존재하는 인공 표식물의 글씨를 충분히 인지할 수 있는 지를 기준으로 삼았다. Mesh를 이용한 지도 표현과 기존 방법론에서 많이 사용되는 pointcloud와 voxel을 기반으로 작성된 지도와 정성적인 비교를 진행하였다. Pointcloud 지도는 ORB-SLAM2^[6]의 keyframe 위치에서 깊이, 색 정보를 이용해 직접 생성하였고, voxel 기반 지도는 Voxblox^[4] 알고리즘을 이용해 1 cm 크기로 구한 TSDF 지도에 marching cube 알고리즘을 적용하여 얻은 mesh 지도를 통해 비교를 수행하였다.

[Fig. 10]의 결과를 보면, mesh와 pointcloud를 사용하였을 때 글씨를 인식할 수 있는 수준의 결과를 얻은 것을 확인할 수 있다. Voxel을 기반으로 지도를 표현했을 때는 1 cm 정도의 작은 크기의 voxel을 사용하였음에도 불구하고 글자를 인식하기 충분한 정도의 지도가 표현되지 않았다. Pointcloud의 결과가 좋게 나왔음에도 불구하고, pointcloud의 경우 1개의 점이 1개의 색 정보를 표현하기 때문에 사람이 해석할 수 있는 수준의 지도 표현을 위해서는 매우 큰 크기의 데이터가 요구된다. 이는 [Table 2]를 보면 알 수 있다.

3.2.4 데이터 크기와 실시간 동작

[Table 2]를 보면 지도 표현 방법에 따라 저장되는 point 및 voxel의 수를 확인할 수 있다. Mesh 표현을 사용하는 경우 pointcloud를 사용하는 것보다 훨씬 적은 수의 point를 저장하면서도 [Fig. 9]와 같이 높은 수준의 지도 작성 결과를 보여줄 수 있다. Voxel 기반의 표현을 사용하였을 경우, mesh



[Fig. 12] The original image of the dataset (c), the distorted image following the underwater light propagation model^[11] (a), and the enhanced image using Joint-ID^[2]

[Table 3] Absolute Trajectory Error (ATE) (m) for *of kt0* sequence

	Mean	Median	RMSE
without image enhancement	0.038324	0.020187	0.055224
with image enhancement	0.018117	0.012457	0.024910

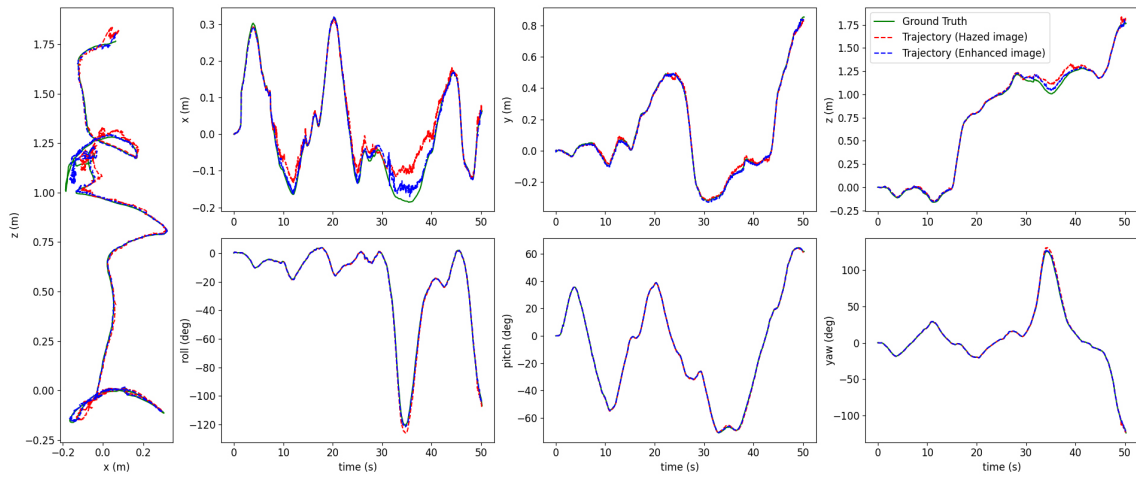
기반 지도보다 적은 수의 voxel block으로 지도 표현이 가능하지만 사람이 해석하기 충분할 정도의 사실적인 지도작성이 불가능하였다.

제안하는 방법은 mesh기반의 표현과 sliding window 방식의 지도 확장을 통해 실시간으로 지도 작성을 가능하도록 한다. 본 논문에서는 가로 720, 세로 456 해상도를 가지는 영상에 대해 N=25를 window 크기로 설정하여 실험을 진행하였다. [Fig. 11]의 결과를 보면 프레임 당 처리시간을 알 수 있다. 결과를 살펴보면 ORB-SLAM2 기반의 위치 인식은 평균 18.14 ms의 수행 시간이 소요되었고, 지도 작성을 하기위해서 평균 약 11.18 ms의 시간이 소요되었다. 전체 프레임당 처리 시간은 평균 약 29.32 ms으로, 이는 제안하는 방법이 위치 인식과 지도 작성의 관점에서 충분히 실시간으로 동작 가능함을 보여준다.

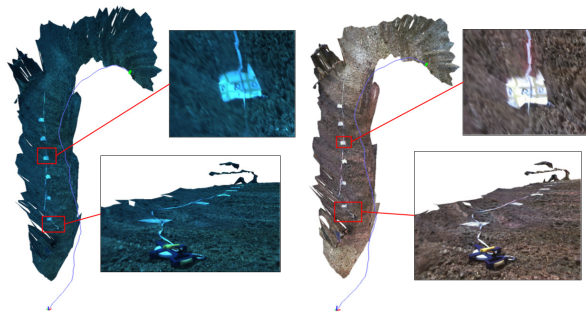
3.2.5 정량적 결과 비교

수중 환경의 특성 상 ground truth 경로 정보를 얻기 힘들고, 본 논문에서 사용한 데이터셋 또한 ground truth 경로 정보를 제공하고 있지 않다. 따라서 본 논문에서는 제안하는 방법의 정량적인 검증을 수행하기 위해 실내 합성 데이터셋인 ICL-NUIM^[10]의 영상 정보를 왜곡시켜 수중 환경을 가정하였다. 본 논문에서는 원본 이미지로부터 왜곡된 수중 이미지를 표현하기 위해 일반적으로 많이 사용되는 McGlamery의 수중 빛 전파 모델^[11]을 사용하였다. [Fig. 12]에서 원본 이미지와 왜곡시킨 수중 영상을 확인해볼 수 있다.

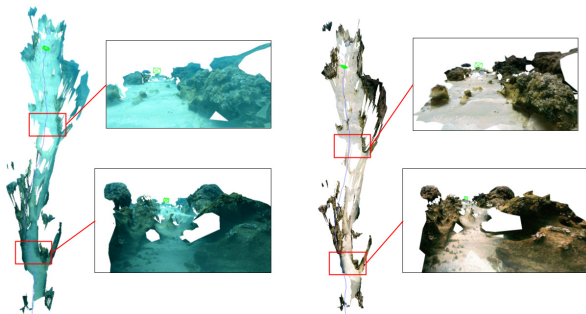
[Fig. 13]과 [Table 3]에서 *of kt0* 데이터에 대한 영상 개선 전후의 위치 인식 결과를 확인해볼 수 있다. 본 논문에서는 절대 경로 오차를 계산하기 위해 *evo*^[12]를 이용하였다. [Fig. 13]을 확인해보면 전체적으로 영상 개선 후에 위치 인식의 결과가



[Fig. 13] Trajectory, position, and orientation of ground truth compared to the results using the hazed and enhanced image for *of kt0* sequence



[Fig. 14] Mapping result for FLSea-VI Redseas Coral Table Loop sequence without image enhancement (left) and with image enhancement (right)



[Fig. 15] Mapping result for *FLSea-Stereo Rock Garden 1* sequence without image enhancement (left) and with image enhancement (right)

더 좋게 나온 것을 확인할 수 있다. *of kt0* 데이터는 중간에 천장을 바라보며 움직이는 구간이 있는데, 해당 구간에서는 질감이 부족한 영상의 특성으로 인해 모든 결과에서 위치 오차가 크게 나왔다. 다만 영상 개선 후에 오차가 더 작게 나온 것을 확인할 수 있다. [Table 3]의 내용을 보면 영상 개선 전후의 각각 절대 경로 오차의 평균이 0.038324 m, 0.018117 m로 영상 개선 후 오차가 약 2배 줄어든 것을 확인할 수 있다.

4. 결론

본 논문에서는 이미지 개선과 mesh 표현 방법을 사용해 위치 인식 정확도를 향상하고 사실적인 실시간 지도 작성을 위한 시스템을 제안하였다. 개선된 이미지를 사용하였을 때, 기존 이미지를 사용하는 것보다 로봇의 위치 인식 성능 및 지도 표현이 향상된 정성적으로 확인하였고, 작성된 지도가 사람이 이해하기 쉽도록 작성된 것을 알 수 있었다. 또한 [Fig. 14], [Fig. 15]에서 지도 작성의 정성적 검증에 위한 추가적인 결과를 제시한다. 향후에는 작성된 mesh 지도 정보를 이용해 위치 인식의 성능을 더욱 향상시키고 루프 폐쇄 감지를 통해 위치 보정을 할 수 있는 시스템을 설계할 예정이다.

References

- [1] Y. Cho, Y.-S. Shin, and A. Kim, "Online depth estimation and application to underwater image dehazing," *OCEANS 2016 MTS/IEEE Monterey*, Monterey, USA, pp. 1-7, 2016, DOI: 10.1109/OCEANS.2016.7761109.
- [2] G. Yang, G. Kang, J. Lee, and Y. Cho, "Joint-ID: Transformer-based Joint Image Enhancement and Depth Estimation for Underwater Environments," *IEEE Sensors Journal*, vol. 24, no. 3, pp. 3113-3122, Feb., 2023, DOI: 10.1109/JSEN.2023.3338730.
- [3] R. A. Newcombe, S. Izadi AU, O. Hilliges, D. Molyneaux, D. Kim, A. J. Davison, P. Kohi, J. Shotton, S. Hodges, and A. Fitzgibbon, "KinectFusion: Real-time dense surface mapping and tracking," *2011 10th IEEE International Symposium on Mixed and Augmented Reality*, Basel, Switzerland, pp. 127-136, 2011, DOI: 10.1109/ISMAR.2011.6092378.
- [4] H. Oleynikova, Z. Taylor, M. Fehr, R. Siegwart, and J. Nieto, "Voxblox: Incremental 3D Euclidean Signed Distance Fields for

on-board MAV planning,” *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Vancouver, Canada, pp. 1366-1373, 2017, DOI: 10.1109/IROS.2017.8202315.

[5] A. Rosinol, T. Sattler, M. Pollefeys, and L. Carlone, “Incremental Visual-Inertial 3D Mesh Generation with Structural Regularities,” *2019 International Conference on Robotics and Automation (ICRA)*, Montreal, Canada, pp. 8220-8226, 2019, DOI: 10.1109/ICRA.2019.8794456.

[6] R. Mur-Artal and J. D. Tardós, “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, Oct., 2017, DOI: 10.1109/TRO.2017.2705103.

[7] A. Bowyer, “Computing Dirichlet tessellations,” *The Computer Journal*, vol. 24, no. 2, pp. 162-166, Jan., 1981, DOI: 10.1093/comjnl/24.2.162.

[8] J. Shi and Tomasi, “Good features to track,” *1994 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Seattle, USA, pp. 593-600, 1994, DOI: 10.1109/CVPR.1994.323794.

[9] Y. Randall and T. Treibitz, “Flsea: Underwater visual-inertial and stereo-vision forward-looking datasets,” *arXiv:2302.12772*, 2023, DOI: 10.48550/arXiv.2302.12772.

[10] A. Handa, T. Whelan, J. McDonald and A. J. Davison, “A benchmark for RGB-D visual odometry, 3D reconstruction and SLAM,” *2014 IEEE International Conference on Robotics and Automation (ICRA)*, Hong Kong, China, pp. 1524-153, 2014, DOI: 10.1109/ICRA.2014.6907054.

[11] J. Y. Chiang and Y.-C. Chen, “Underwater image enhancement by wavelength compensation and dehazing,” *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1756-1769, Apr., 2012, DOI: 10.1109/TIP.2011.2179666.

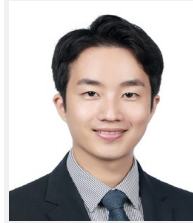
[12] MichaelGrupp, [Online], <https://github.com/MichaelGrupp/evo>, Accessed: Jun. 02, 2023.



이 정 우

2023 인하대학교 정보통신공학과(학사)
2023~현재 인하대학교 전기컴퓨터공학과 (석사)

관심분야: Visual SLAM, Mapping



조 영 군

2006~2012 인하대학교 전기공학과(학사)
2013~2015 KAIST 전기공학과(석사)
2013~2020 KAIST 건설및환경공학과/로봇틱스(박사)
2020~2021 영남대학교 로봇공학과 조교수
2021~현재 인하대학교 전기공학과 조교수

관심분야: SLAM, Spatial AI, Autonomous Robots