

수중에서의 특징점 매칭을 위한 CNN기반 Opti-Acoustic변환

CNN-based Opti-Acoustic Transformation for Underwater Feature Matching

장혜수¹·이영준²·김기섭³·김아영[†]

Hyesu Jang¹, Yeongjun Lee², Giseop Kim³, Ayoung Kim[†]

Abstract: In this paper, we introduce the methodology that utilizes deep learning-based front-end to enhance underwater feature matching. Both optical camera and sonar are widely applicable sensors in underwater research, however, each sensor has its own weaknesses, such as light condition and turbidity for the optic camera, and noise for sonar. To overcome the problems, we proposed the opti-acoustic transformation method. Since feature detection in sonar image is challenging, we converted the sonar image to an optic style image. Maintaining the main contents in the sonar image, CNN-based style transfer method changed the style of the image that facilitates feature detection. Finally, we verified our result using cosine similarity comparison and feature matching against the original optic image.

Keywords: Sonar, Deep Learning, Underwater, Feature Matching

1. 서 론

수중 비전 연구에 있어 광학카메라와 소나는 광범위하게 사용되는 이미징 센서이다. 광학카메라의 경우 사람이 보기에 직관적인 이미지를 만들기 때문에 지상과 수중 관계없이 많이 사용되며, 소나의 경우 광조건에 상관없이 이미지를 만들기 때문에 수중 환경에 특화되어 있다. 현재, 컴퓨터 비전 연구들이 많이 발전하였으나 기존의 기술을 수중에서 그대로 사용하기에는 한계가 있다. 광학 이미지의 경우 수중에서 광조건과

탁도의 영향을 많이 받고, 소나 이미지는 노이즈가 많이 발생하기 때문이다.

각 센서들의 한계점을 극복하기 위해서 다양한 연구들이 활발히 이루어 지고 있다. 그 예로, [1], [2], [3]에서는 수중 광학 이미지의 탁도를 개선하고 복원하는 모습을 보여주었고, [4]와 [5]에서는 소나 이미지의 복원과 철극성 추출방법에 대한 연구를 다루었다. 우리는 다양한 센서들을 결합시키는 것에 큰 비중을 두었고, [6]에서 소개된 광학 센서와 음향 센서의 스테레오 이미징에서 아이디어를 착안하여 연구를 진행하기로 결정하였다. 앞서 언급하였듯, 소나 센서는 탁도에 무관하게 동작하나 취득한 이미지가 작업자에게 직관적이지 않고, 영상 센서는 풍부한 정보를 제공하나 가시거리에 제한을 받는다는 점에서 두 센서의 상보적 사용은 수중 환경의 인식 문제에 있어서 중요하다. 이를 위하여 확연한 차이가 있는 두 이미지 간의 매칭 문제를 해결하고자 한다.

인공 신경망의 첫 발견 이래로 컴퓨터 비전 연구는 매우 광범위하게 발전하였다. 본 논문은 그 중 신경망 기반 스타일 변환 기법에서 상기 기술된 두 개의 다른 이미지 간의 매칭 문제를 해결하고자 한다. 스타일 변환 기법은 [Fig. 1]의 오른쪽 사진과 같이 기존의 이미지에 새로운 스타일을 입힌 영상을

Received : Dec. 10. 2019; Revised : Jan. 7. 2020; Accepted : Jan. 7. 2020

※ This research was supported by a grand from Endowment Project of "Development of Basic Technologies of 3D Object Reconstruction and Robot-Manipulator Motion Compensation Control for Autonomous Underwater Working" funded by Korea Research Institute of Ships and Ocean Engineering (PES3170)

1. Master student, Dept. of Civil and Environmental Engineering, KAIST, Daejeon, Korea (iriter@kaist.ac.kr)
2. Junior Engineer, Korea Research Institute Ship and Ocean engineering (KRISO), Daejeon, Korea (leeyoungjun@kriso.re.kr)
3. Ph.D. student, Dept. of Civil and Environmental Engineering, KAIST, Daejeon, Korea (paulgkim@kaist.ac.kr)

† Assistant professor, Corresponding author: Dept. of Civil and Environmental Engineering, KAIST, 291 Daehak-ro, Yuseong-gu, Daejeon, Korea (ayoungk@kaist.ac.kr)



[Fig. 1] Optical image, sonar image, and style-transferred sonar image

생성할 수 있어, 본 논문에서 해결하고자 하는 두 센서 도메인 (광학과 소나) 사이의 간극을 줄이는 데 사용할 수 있다.

본 논문에서는 신경망 기반의 광학-소나(opti-acoustic) 이미지 간 특징점 매칭 문제 해결에 대해 다룰 것이다. 광학 이미지는 선명한 것에 비해 소나 이미지는 노이즈를 많이 가지고 있는 것이 특징이다. 같은 물체를 촬영하더라도 검출되는 특징점의 개수에서는 차이가 있다. 본 논문의 목적은 신경망을 기반으로 변환된 이미지를 이용하여 광학 이미지와 소나 이미지 간의 상관관계를 찾아내는 것이다. 신경망 기반의 변환 기법을 통해 노이즈를 줄이고, 소나 이미지를 보다 선명하게 변환시킬 것으로 예상된다. 다음 항목들이 본 논문에서 제시하고자 하는 논점들이다.

- 노이즈의 영향을 줄이기 위해서 이미지 전처리방법을 사용하여 신경망이 소나가 촬영한 물체에만 집중할 수 있도록 하였다.
- 소나 이미지의 변환을 통해 광학-음향 이미지 간의 스테레오 정보를 구성할 수 있도록 하였다.
- Cosine 유사도와 특징점 검출 방법을 이용하여 제시한 방법으로 생성한 이미지의 효용성을 검증하였다.

2. 선행 연구 조사

2.1 광-음향학(Opti-acoustics)

광학 이미지와 소나 이미지 간의 특징점 매칭과 이미지 캘리브레이션 방법론이 [6]에서 소개되었다. 사전에 측정된 다섯 개의 광학-소나 이미지 간 정보를 이용하여 소나 이미지의 에피폴라 곡선을 만들었다. 광학-소나 정보와 에피폴라 기하학을 바탕으로 3D 복원을 진행하는 것을 논문에서 보여주었다. 광-음향 투영 모델을 기반으로 한 번들 조정이 [7]에서 소개되었다. SIFT (Scale Invariant Feature Transform), BRISK (Binary Robust Invariant Scalable Keypoints), A-KAZE (Accelerated-KAZE), DASC (Dense Adaptive Self-Correlation)과 같은 다양한 특징점 검출 방법들을 이용하여 광학 이미지와 소나 이미지 간의 매칭을 시도하였다. 결과적으로 DASC가 특징점 검출에 있어 가장 효과적이라는 것을 보여주었다. 하지만 이러한 광-음향학 논문들이 가정하고 있는 부분은 광학 카메라와 소나

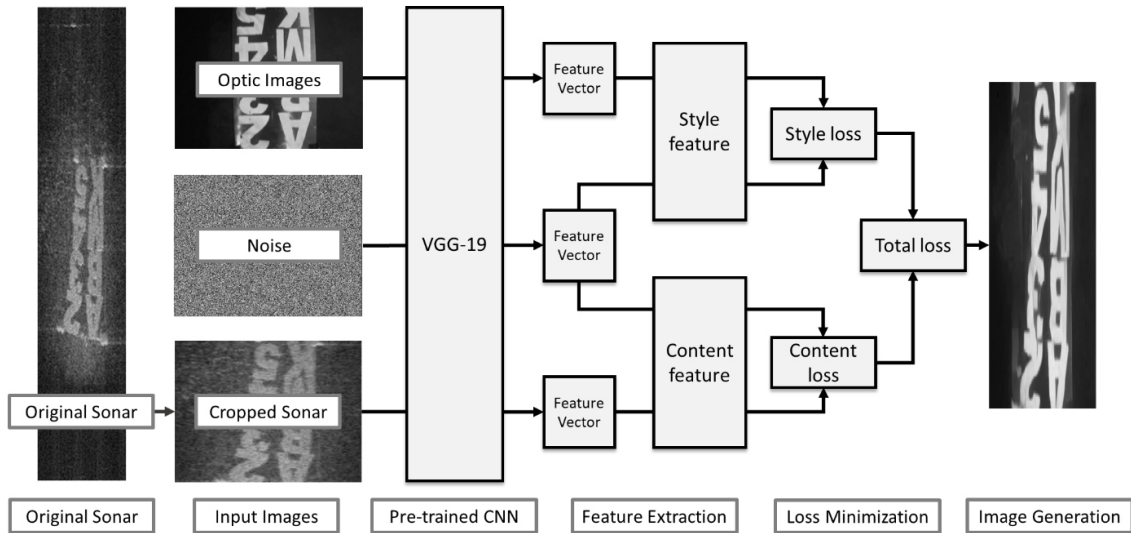
이미지가 같은 장면을 촬영하였다는 것이다. 기존 컴퓨터 비전 연구에 활용하기 위해서는 맞춰진 환경이 아닌 곳에서 특징점 검출 및 매칭이 이루어져야 한다. [7]에서 언급했듯이, 광학 이미지와 소나 이미지 간의 특징점 매칭 문제는 소나 이미지의 심한 노이즈 때문에 매우 도전적인 과제이다.

2.2 신경망 기반 스타일 변환

신경망 기반의 이미지 스타일 변환 기법은 최근 컴퓨터 비전과 로봇틱스 연구에서 떠오르고 있다. 이미지의 콘텐츠를 유지하면서 새로운 스타일을 부여하는 것은 다양한 방법으로 활용할 수 있는 기술이기 때문이다. 신경망 기반 이미지 스타일 변환은 크게 두 가지 방법으로 나뉘며, 사전 학습된 CNN을 활용하는 방법과 GAN (Generative Adversarial Network)를 활용하는 방법이 있다. [8]에서는 CNN과 그라프 행렬을 이용하여 이미지의 스타일을 추출하는 방법을 소개하였다. 메인 네트워크로는 [9]에서 소개한 VGG-19를 사용하였으며, 레이어들 간의 차이를 이용하여 이미지 스타일을 추출해냈다. 최종적으로 [10]에서 CNN 기반의 스타일 변환 알고리즘을 확립하였으며, 스타일 특징과 콘텐츠 특징을 합쳐 변환된 이미지를 생성해냈다. GAN을 이용한 스타일 변환은 [11]과 [12]에 소개되었다. Pix2pix와 CycleGAN이라는 이름으로 유명한 두 방법은 생성자와 구분자를 동시에 학습시킴으로써 현실적인 이미지를 생성하도록 설계된다. GAN으로 생성된 이미지는 매우 좋은 퀄리티를 지녔으나, 학습에 필요한 데이터셋으로 매우 직관적인 이미지들을 요구한다는 한계점이 있다.

2.3 수중 이미지에의 적용

수중 연구에서는, [3]이 GAN을 기반으로 한 수중 광학 이미지 복원 연구를 진행하였다. 총 6500장의 광학 이미지들을 이용하여 네트워크를 학습시켰고, 매우 뛰어난 퀄리티의 결과물을 만들어냈다. 소나를 활용한 연구는 [5]에서 진행되었으며, CNN을 활용하여 소나 이미지의 철극성에 대한 점수를 부여하였다. 부여된 점수를 이용하여 수중 SLAM (Simultaneous Localization and Mapping) 성능을 향상시켰다. 다른 연구들에서 알 수 있듯이, 이전의 신경망 기반 수중 연구는 단일 센서에 대하여 주로 진행되었다. 이와 차별되게 본 연구는 광학 이미지와 소나 이미지를 동시에 고려하는 모습을 보여주었다. GAN을 이용하기 위해서는 충분한 이미지 데이터셋이 뒷받침되어야 한다. 그러나 광학-소나 스테레오 이미지에 대한 공개 데이터셋이 충분하지 않기 때문에 현재로서는 광-음향 이미지 데이터 처리에 GAN을 이용하는 것은 한계가 있다. 따라서 본 연구에서는 CNN기반의 스타일 변환 기법을 발전시켜 활용하였다.



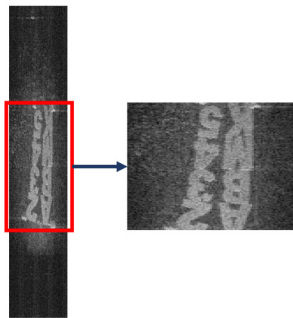
[Fig. 2] Style transfer algorithm for optic image and sonar image. Pre-processed sonar images are utilized as input sonar images. Pre-trained CNN VGG-19 extracts the feature information for each input images. Finally calculate the loss with style and content feature to generate a minimum loss image

3. 연구 방법

본 논문에서 제시한 방법론은 크게 세가지로 나눌 수 있다. 첫째는 이미지 전처리 단계이며, 이미지에 발생하는 노이즈가 이미지의 주 콘텐츠가 되지 않도록 만들어주는 역할을 한다. 그 뒤에 이루어지는 작업은 신경망 기반 특징 추출로, 사전 학습된 CNN 네트워크를 이용하여 이미지의 스타일과 콘텐츠 정보를 추출해낸다. 마지막으로 학습에서의 손실 최소화를 통해 제시한 스타일과 콘텐츠 모두를 만족하는 이미지를 생성해낸다. 전체 과정은 [Fig. 2]에 설명되어 있다.

3.1 이미지 전처리

[Fig. 3]에서 소나리를 이용하여 얻은 이미지를 볼 수 있다. 이미지의 윗부분과 아랫부분은 촬영된 물체없이 노이즈만 존재함을 확인할 수 있다. 이렇듯 특징 분석에 방해가 되는 요소를



[Fig. 3] Original sonar images and pre-processed sonar image

사전에 제거함으로써 노이즈가 물체로 인식되는 현상을 방지하였다. 다양한 물체가 존재할 경우 개별 물체에 대한 속성을 잃을 가능성이 있으므로, 본 연구에서는 단일 물체에 대한 영역을 잘라내어 학습을 진행하였다.

3.2 CNN을 이용한 이미지 특징 추출

앞서 소개했듯이, 이미지 스타일 변환에 사용되는 신경망은 사전 학습된 CNN과 GAN이 있으며, 각각의 방법들은 장단점이 확실하다는 점에서 둘 다 가치가 있다고 말할 수 있다. 먼저, CNN을 활용한 스타일 변환의 경우 스타일을 가져올 이미지와 콘텐츠를 가져올 이미지 총 2장의 이미지만 있어도 변환이 가능하며, 학습시간이 비교적 짧다는 장점이 있다. 그러나 다양한 상황에 적용시키기 위해서는 다양한 스타일 표본들이 필요하다는 단점이 있다. 이에 반해 GAN을 활용한 방법의 경우 학습을 위해 매우 많은 데이터셋들과 시간을 필요로 하는 반면, 다양한 상황에 걸맞는 결과물을 도출해내는 장점을 가지고 있다. 지상의 경우 ImageNet과 CIFAR-10과 같은 다양한 공개 데이터셋이 존재하지만, 수중 환경을 배경으로 한 공개 소나 데이터셋은 아직까지 존재하지 않는다. 또한, GAN을 이용한 이미지 생성은 심미적인 효과의 극대화가 목적이기 때문에 특징점 추출 및 매칭을 필수로 하는 로봇 비전에 필요한 이미지를 도출해내지 않을 가능성이 크다. 따라서 현재로서는 GAN을 이용한 방법에 비해 사전 학습된 CNN을 이용한 스타일 변환 방법이 경쟁력이 있다고 판단하여, 이미지를 본 방법으로 생성하였다.

사진 학습된 CNN으로는 [13]에서 소개된 것과 같이 VGG-19를 사용하였다. 이미지에서 추출해야 하는 특징은 스타일과 콘텐츠로, 각각의 특징은 다음과 같이 계산할 수 있다. 먼저 콘텐츠 특징의 경우 각각의 픽셀 정보보다 시맨틱 정보가 더 큰 가치를 지녔으므로, 시맨틱 정보만 가지고 있는 CNN의 가장 깊은 층 데이터를 콘텐츠 특징 C^l 로 정의한다. 스타일 특징 S^l 의 경우 [8]에 제시된 방법과 같이 그람 행렬(1)을 사용하여 특징 정보 지도 F^l 들 간의 상관관계로 정의하였다.

$$G^l_{ij} = \sum_k F^l_{ik} F^l_{jk} \quad (1)$$

그람 행렬 G^l_{ij} 은 l layer의 특징 정보 지도 i, j 간의 내적으로 정의되며, 각각의 콘텐츠 특징과 스타일 특징은 손실 함수로 사용되었다.

3.3 손실 최소화를 통한 이미지 생성

앞서 소개된 스타일 및 콘텐츠 특징을 이용하여 우리는 [13]과 같이 두 손실 함수 $L_{content}$ 와 L_{style} 을 정의하였다. L_{style} 은 생성된 이미지와 베이스로 제시된 이미지 간의 스타일 특징 차이를 나타낸 손실 함수이며, 그 값이 작아질수록 무작위하게 생성된 이미지의 스타일이 제시된 이미지와 유사하다는 것을 의미한다. 식 (2)와 (3)을 통해 스타일 손실 함수의 정의를 확인할 수 있으며, 생성된 이미지와 제시된 이미지의 그람 행렬 G^l, S^l 이 계산에 있어 핵심 요소로 작용한다. N_l 은 특징 정보 지도의 수를 의미하며, M_l 은 l layer에서의 특징 정보 지도 크기를 의미한다. 또한, 가중치 w_l 를 사용하여 각 층별 정보를 취합하였다.

$$E_l = \frac{1}{N_l^2 M_l^2} \sum_{i,j} (G^l_{ij} - S^l_{ij})^2 \quad (2)$$

$$L_{style} = \sum_l w_l E_l \quad (3)$$

콘텐츠 손실 함수 $L_{content}$ 는 제시된 이미지와 생성된 이미지의 콘텐츠 특징 차이를 이용하여 계산되었다. 콘텐츠 손실 함수가 최소화 될 경우 생성된 이미지와 제시된 이미지가 같은 콘텐츠를 가지고 있다고 해석할 수 있다. C^l 은 제시된 이미지의 콘텐츠 특징 정보이며, I^l 은 생성된 이미지의 콘텐츠 특징 정보이다. 각각은 VGG-19 네트워크의 l layer에서의 정보를 의미하며, 우리는 l 값을 가장 깊은 층인 5로 정의하여 콘텐츠 특징을 계산하였다.

$$L_{content} = \sum_{i,j} (I^l_{ij} - C^l_{ij})^2 \quad (4)$$

전체 네트워크의 목적은 두 손실 함수를 최소화 시키는 것이다. 하지만 과적합 문제와 한 가지 손실 함수만 최종 결과에 영향을 끼치는 것을 방지해야 하므로 식 (5)와 같이 두 손실 함수의 값을 조절하여 최종 손실 함수를 만들었다.

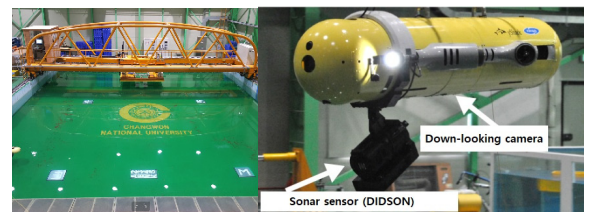
$$L_{total} = \alpha L_{style} + \beta L_{content} \quad (5)$$

4. 연구 결과

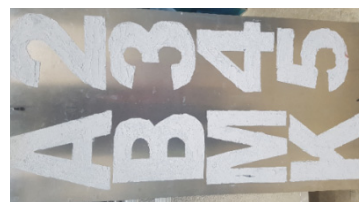
제시한 방법론을 평가하기 위해 광-음향 스테레오 이미징 장비가 사용되었다. 취득한 이미지를 이용하여 우리는 정량적 및 정성적 분석을 진행하였다. 정량적 분석으로는 각 이미지에서 추출된 특징 벡터 간의 cosine 유사도와 특징점 매칭에 대한 검증을 진행하였으며, 정성적으로는 생성된 이미지가 소나로 촬영된 물체를 가지고 있으면서 광학 카메라의 스타일을 가지고 있음을 확인하였다.

4.1 수조 실험 환경

실험은 [Fig. 4]에 나와있는 AUV와 창원대학교의 수조에서 진행되었다. 소나 장비인 DIDSON이 AUV의 앞쪽에 설치되었고, 광학 카메라는 수조 바닥을 촬영할 수 있도록 AUV하단에 설치되었다. DIDSON은 96×512 크기의 흑백 이미지를 생성하며, 동시에 광학 카메라는 1024×1024 크기의 RGB 이미지를 생성한다. 실험에 사용된 목표물은 [Fig. 5]에 나타내어져 있으며, 숫자와 알파벳의 조합을 사용하였다. AUV는 수면을 따라 이동하며, DIDSON과 광학 카메라 둘다 목표물을 촬영할 수 있도록 하였다.



[Fig. 4] Experimental environment and AUV for image acquisition



[Fig. 5] Target object for the experiment. Alphabets and numbers are printed on the board



[Fig. 6] Input images for neural network. (a): original optic image, (b): noise data for image generation



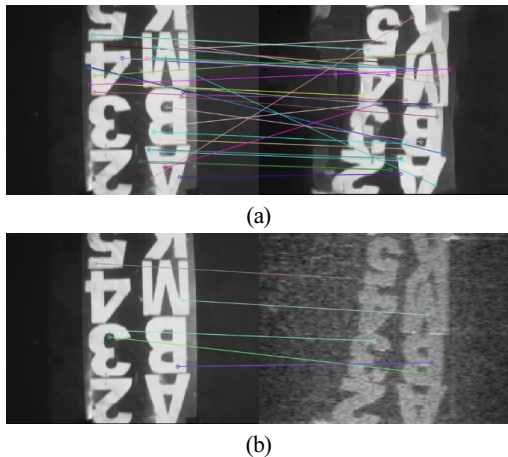
[Fig. 7] Generated images from the neural network. Sonar images are converted as optic style image

4.2 신경망 기반 스타일 변환

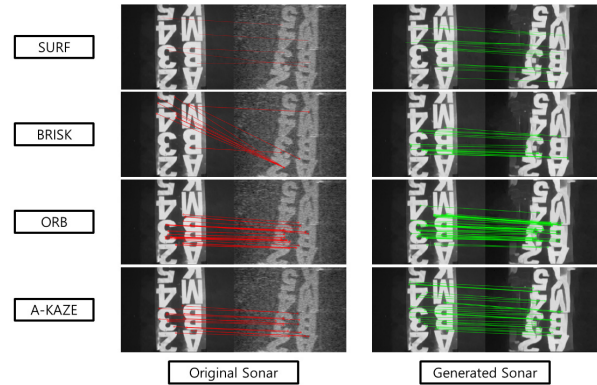
본 연구에서 사용된 베이스 스타일은 광학 카메라 이미지의 스타일이며, [Fig. 6(a)]에 제시된 것과 같다. 이미지 생성을 위해 최초로 만들어지는 이미지는 [Fig. 6(b)]와 같은 노이즈 이미지이다. 신경망을 이용하여 만들어진 이미지 샘플은 [Fig. 7]에서 확인 할 수 있다. 생성된 이미지는 소나 이미지의 콘텐츠를 담고 있음과 동시에 광학 카메라의 스타일을 지녔다.

4.3 특징점 매칭 평가

특징점 매칭 결과는 [Fig. 8]과 [Fig. 9]에서 확인할 수 있다. 특징점 검출은 SIFT, BRISK, ORB, A-KAZE를 활용하였으며,



[Fig. 8] All SIFT-based feature detection and matches between the optic image and (a) Generated sonar, (b) Original sonar

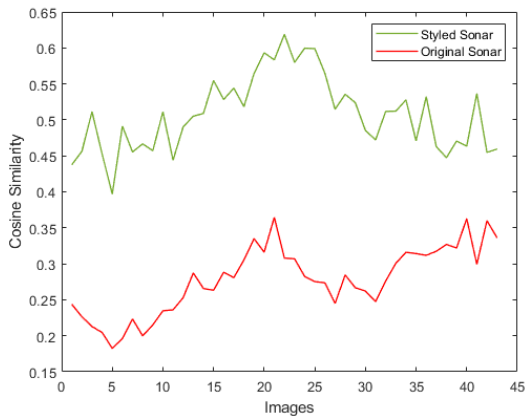


[Fig. 9] Feature matching results with various methods (SURF, BRISK, ORB, A-KAZE). The generated image depicts more precise matching than the original sonar in all matching methods

[Table 1] Feature matching result (Average) Inlier is excluded for SURF since the provided function in Matlab does not support inliers

Image	Methods	Matches	Inliers	Inliers Ratio
Raw	SURF	3.05	-	-
	BRISK	1.95	0	0
	ORB	4.6	1.25	0.272
	A-KAZE	4	0.65	0.163
Generated	SURF	14.6	-	-
	BRISK	19.95	15.5	0.777
	ORB	43.3	32.7	0.755
	A-KAZE	25.5	20.6	0.808

매칭은 FLANN (Fast Library for Approximate Nearest Neighbors) 방법을 사용하였다. [Fig. 8(a)]와 [Fig. 8(b)]는 광학 이미지와 소나 이미지, 생성된 이미지 간의 SURF를 사용한 전체 매칭 결과를 나타내고 있으며, 기존 소나 이미지의 경우 광학 이미지와 매칭된 특징점이 5개만 존재함을 확인할 수 있다. 그에 반해 생성된 이미지와 광학 이미지 사이에는 25개의 매칭이 발생함을 확인할 수 있다. [Fig. 9]에서는 잘못 매치된 값들을 제거하기 위하여 RANSAC (Random Sample Consensus) 기법을 사용한 결과물들을 보여주고 있다. SURF, BRISK, ORB, A-KAZE의 다양한 특징점 검출 방법들에 대해 RANSAC을 적용하였으며, 스타일 변환을 한 뒤에 안정적인 특징점 매칭이 이뤄짐을 확인할 수 있다. 이전 문헌[14]에서 매칭이 잘 이루어지지 않는다고 알려진 SURF, BRISK, ORB에서도 스타일 변환시 매칭 성능이 향상되었고, A-KAZE 또한 매칭 정확도가 증가한 것을 확인할 수 있었다. [Table 1]을 통해 각 특징점 매칭 방법 별 매칭 결과 값들을 상세히 확인할 수 있다. 각각의 결과들은 40쌍의 이미지들에 대한 평균 매칭 값이며, 기존 소나 이미지를 이용한 것 보다 생성된 이미지를 활용하여 매칭을 시도하는 것이 훨씬 안정적인 결과를 나타냄을 확인할 수 있다.



[Fig. 10] Cosine similarity comparison between styled sonar and original sonar. Similarities are calculated with optic images

[Table 2] Average cosine similarity

Image	Raw Sonar	Styled Sonar
Cosine Similarity	0.277	0.517

4.4 특징 벡터의 유사성 평가

이미지 간의 특징점 매칭은 주어진 후보들 중 가장 유사도가 높은 특징점을 선정하는 과정이다. 즉, 전체 이미지의 유사도를 평가하는 것이 일반적인 특징점 매칭보다 더 엄밀한 분석이 될 수 있다. 따라서, 다음과 같은 특징 벡터 간의 유사도 분석을 추가적으로 수행하여 변환된 이미지에서 추출한 특징 벡터가 타겟과 얼마나 유사한지를 분석하였다. Cosine 유사도에 사용된 벡터들은 VGG-19로부터 계산하였으며, 전범위적 분석을 위해 이미지 모든 픽셀에 대하여 유사도를 계산하였다. [Fig. 10]을 통해 Cosine 유사도 비교 값을 확인할 수 있다. Cosine 유사도는 0에서 1 사이값을 가지는데, 1에 가까울수록 실험군이 대조군과 유사함을 의미한다. [Table 2]에서 평균 cosine 유사도 비교 값을 확인할 수 있다.

5. 결 론

본 논문에서는 신경망을 이용한 광-음향 이미지 간 변환에 대한 방법론을 제시하였다. 예상한 바와 같이 광학 이미지 스타일로 변환된 소나 이미지는 심미적으로 매우 높은 질을 가짐과 동시에 정량적 분석에도 안정적인 모습을 보였다. 원본 소나 이미지가 특징점 검출에 매우 취약한 것과 반대로, 변환된 이미지는 강건한 모습을 보였다.

결과물이 뛰어난 성능을 가졌음에도 불구하고 우리는 두 가지의 문제점을 확인할 수 있었다. [Fig. 7(b)]에서 확인할 수 있듯이, 원본 광학 이미지에는 아무것도 없는 부분이지만 생

성된 이미지에서는 물체로 채워져 있는 현상이 발생하였다. 이는 노이즈 값들이 콘텐츠로 인식되어 변환된 것으로 보이며, 이러한 노이즈의 무분별한 변환을 막기 위한 노력이 추후에 필요할 것으로 보인다. 두번째로는 광 조건이 변화할 때 마다 새로운 광학 이미지가 필요하다는 단점이 있다. 이는 앞서 말한 GAN의 장점과 대비되는 것으로, 비록 학습 시간은 짧지만 매번 스타일을 변환시키는 것은 분명히 개선해야 할 점이다.

추후에 진행될 연구로, 다양한 수중 상황에 대처할 수 있도록 광-음향 데이터셋을 추가로 취득하여 본 연구를 일반화할 예정이다. 또한, 변환 강건성을 위해 이미지 전처리 단계에서 [15]의 학습 전략과 같이 노이즈를 최소화 하고 목표물만을 인식할 수 있도록 신경망을 추가로 활용할 것이다. 나아가 기존의 음향만을 사용한 항법[16]이 아니라 광-음향을 모두 사용한 항법이 가능하다.

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, CA, USA, pp. 1-7, 2016.
- [2] Y. Cho and A. Kim, "Visibility enhancement for underwater visual slam based on underwater light scattering model," *2017 IEEE International Conference on Robotics and Automation (ICRA)*, Singapore, pp. 710-717, 2017.
- [3] J. Li, K. A. Skinner, R. M. Eustice, and M. Johnson-Roberson, "Watergan: unsupervised generative network to enable real-time color correction of monocular underwater images," *IEEE Robotics and Automation Letters*, vol. 3, no. 1, pp. 387-394, Jan., 2018.
- [4] N. Hurtós, X. Cufí, Y. Petillot, and J. Salvi, "Fourier-based registrations for two-dimensional forward-looking sonar image mosaicing," *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vilamoura, Portugal, pp. 5298-5305, 2012.
- [5] J. Li, M. Kaess, R. M. Eustice, and M. Johnson-Roberson, "Pose-graph slam using forward-looking sonar," *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 2330-2337, Jul., 2018.
- [6] S. Negahdaripour, H. Sekkati, and H. Pirsiavash, "Opticoustic stereo imaging: On system calibration and 3-d target reconstruction," *IEEE Transactions on Image Processing*, vol. 18, no. 6, pp. 1203-1214, Jun., 2009.
- [7] D.-H. Gwon, Y.-S. Shin, Y. Kim, A. Kim, Y. Lee, and H.-T. Choi, "Nontemporal relative pose estimation for opti-acoustic bundle adjustment," *OCEANS 2016 MTS/IEEE Monterey*, Monterey, CA, USA, pp. 1-5, 2016.
- [8] L. A. Gatys, A. S. Ecker, and M. Bethge, "Texture synthesis using convolutional neural networks," *Computer Vision and Pattern Recognition*, pp. 262-270, 2015.

- [9] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv:1409.1556*, 2014.
- [10] L. Gatys, A. Ecker, and M. Bethge, "A neural algorithm of artistic style," *Journal of Vision*, 2015.
- [11] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Imageto- image translation with conditional adversarial networks," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, pp. 1125-1134, 2017.
- [12] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, pp. 2223-2232, 2017.
- [13] L. A. Gatys, A. S. Ecker, and M. Bethge, "Image style transfer using convolutional neural networks," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016.
- [14] Y.-S. Shin, Y. Cho, Y. Lee, H.-T. Choi, and A. Kim, "Comparative Study of Sonar Image Processing for Underwater Navigation," *Journal of Ocean Engineering and Technology*, vol. 30, no. 3, pp. 214-220, 2016.
- [15] M. H. Jeon, Y. Lee, Y.-S. Shin, H. Jang, T. Yeu, and A. Kim "Synthesizing Image and Automated Annotation Tool for CNN based Under Water Object Detection," *Journal of Korea Robotics Society*, vol. 14, no. 2, pp. 139-149, Jun., 2019.
- [16] Y.-S. Shin, Y.-J. Lee, H-T. Choi, and A. Kim, "Bundle Adjustment and 3D Reconstruction Method for Underwater Sonar Image," *Journal of Korea Robotics Society*, vol. 11, no. 2, pp. 051-059, Jun., 2016.



장혜수

2018 KAIST 건설 및 환경공학과(학사)
2018~현재 KAIST 건설 및 환경공학과 석사
과정

관심분야: 수중 로봇, 로봇 비전, SLAM



이영준

2009 충남대학교 메카트로닉스공학과(학사)
2014 충남대학교 메카트로닉스공학과(석사)
2011~현재 한국해양과학기술원 부설 선박
해양플랜트연구소 기술원

관심분야: 수중 로봇, 자율 탐사, 소나기반물체인식



김기섭

2017 KAIST 건설 및 환경공학과(학사)
2019 KAIST 건설 및 환경공학과(석사)
2019~현재 KAIST 건설 및 환경공학과 박사
과정

관심분야: SLAM, 로봇 비전, 수중 로봇



김아영

2005 서울대학교 기계항공공학부(공학사)
2007 서울대학교 기계항공공학전공(공학석사)
2012 미시간대학교 기계공학전공(공학박사)
2014~현재 한국과학기술원 건설 및 환경공
학과 조교수

관심분야: 영상기반 SLAM