

드론 탐재 수중 라이다 기반 해양 생물 분포 탐지 클러스터링 기법 비교 분석

김현우, 이명훈*

국립순천대학교, *국립순천대학교

kimhw51793@naver.com, *leemh777@scnu.ac.kr

Comparative analysis of drone-borne underwater lidar-based clustering techniques for marine life distribution detection

Kim Hyeon Woo, MeongHun Lee*

Sunchon National Univ., *Sunchon National Univ.

요약

최근 기후변화와 연안 개발로 인해 해양 생물 다양성이 위협받고 있으며, 이에 따라 생물 분포를 정밀하게 탐지할 수 있는 기술의 수요가 증가하고 있다. 본 연구에서는 수중 라이다 포인트 클라우드 데이터를 기반으로, K-Means, DBSCAN, HDBSCAN 세 가지 클러스터링 알고리즘의 성능을 비교·분석하였다. 실험은 시뮬레이션 기반의 3차원 좌표 및 반사율 데이터를 활용하여, 이상치 제거, 정규화, 반사율 필터링의 전처리 과정을 거쳐 알고리즘을 적용하여 실험 결과를 분석하고, 세 개의 알고리즘이 어떤 알고리즘이 수중 라이다에 적용하기 적합한지 비교·분석한다. 본 연구를 통해 수중 생물 군집 탐지의 정밀도를 향상시키고, 향후 해양 생태계 자동 모니터링 시스템 개발에 활용될 수 있을 것으로 기대된다.

I. 서 론

최근 기후변화 및 연안 개발 등으로 인해 해양 생물의 다양성이 위협받고 있으며, 생물 군집의 분포 변화를 조기에 파악해 생태계 이상 신호를 감지에 대한 수요가 증가하고 있다. 이에 따라 해양 생태계 보존과 자원 관리를 위한 정밀한 생물 분포 탐지 기술의 중요성이 더욱 강조되고 있다. 기존의 수중 생물 조사 방식은 수중 카메라, 음향센서, 수작업으로 관찰하는 방식 등에 의존해왔지만, 이러한 방법은 조사 범위가 제한적이고 반복성과 정량성이 낮다는 한계를 가진다[1][2]. 이러한 제약을 해결하기 위한 대안으로 드론에 수중 라이다(LiDAR: Light Detection and Ranging)를 탐재하여 해저 지형 및 생물 분포를 3차원 포인트 클라우드 형태로 정밀하게 스캔하는 방식이 주목받고 있다[3]. 드론 기반 수중 탐사는 기동성과 자동화 기능을 바탕으로, 인력이 접근하기 어려운 연안 해역이나 위험 지역에서도 정밀한 생태계 조사가 가능하다는 장점이 있다. 수중 라이다는 반사율(intensity)과 공간 좌표(x, y, z)를 함께 수집함으로써 생물체와 지형의 구분 가능성을 높이며, 간접적인 생물 군집 탐지에 활용될 수 있다[4].

그러나 수중 라이다 포인트 클라우드 데이터는 노이즈가 많고 밀도 분포가 불균형하며 구조가 비정형적인 단점을 가지고 있다. 이를 효과적으로 분석하기 위해서는 비지도 학습 기반의 클러스터링 기법이 필요하다.[5] 특히 밀도 기반 알고리즘은 라벨이 없는 환경에서도 군집 구조를 자동 탐지 할 수 있어 수중 생물 분포 탐지에 적합한 방식으로 주목받고 있다.

본 연구에서는 수중 라이다 센서로부터 획득 가능한 환경과 유사한 시뮬레이션 포인트 클라우드 데이터를 구성하고, 대표적인 클러스터링 알고리즘인 K-Means, DBSCAN, HDBSCAN을 적용하여 그 성능을 비교 분석한다. 이를 통해 수중 환경에서의 군집 분포 구조를 탐지하는 데 적합한 기법을 도출하고, 향후 자동화된 해양 생물 모니터링 시스템 개발을 위한 기초 자료를 도출하고자 한다. 본 논문의 순서는 다음과 같다: 2장에서는 실험 데이터 구성, 전처리 과정 및 알고리즘 설정과 같이 연구 방법에 대해 설명하고, 3장에서는 알고리즘별 실험 결과를 비교·분석한다. 마지막으로 4장에서는 본 연구의 결론 및 향후 연구 방향을 제시하고자 한다.

II. 연구방법

본 연구에서는 수중 라이다 센서로 관측 가능한 수중 환경과 비슷한 시뮬레이션 기반 포인트 클라우드 데이터를 구성하고, 이를 기반으로 대표적인 클러스터링 알고리즘의 성능을 비교·분석하였다. 실험에 사용된 데이터는 3차원 공간 좌표(x, y, z)와 반사율(intensity) 값을 포함하고 있으며 수중 환경의 비정형성과 짐음 특성을 반영하여 구성되었다. 클러스터링 기법은 포인트 간 거리 또는 밀도 기반 유사도를 계산하여, 유사한 성질을 가진 포인트들을 자동으로 묶는 방식으로 작동한다. 특히 DBSCAN과 HDBSCAN은 일정 반경 내의 포인트 밀도를 기준으로 군집을 형성하고, 밀도가 낮은 포인트는 이상치로 분류한다. 수중 라이다 데이터는 반사율 분포와 밀도 불균형이 존재하므로, 이러한 밀도 기반 짐음 방식이 실제 생물 군집 탐지에 효과적이다.

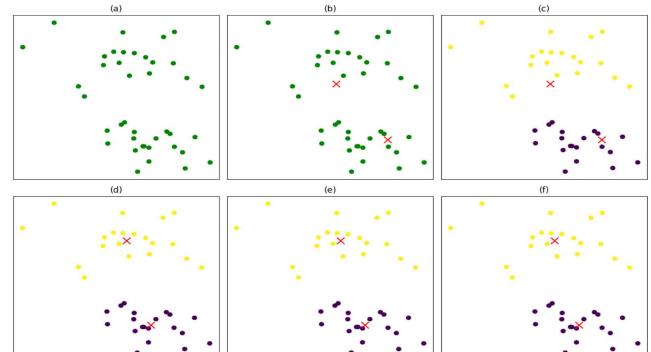


그림 1 K-means 알고리즘 작동 순서 예시

fig 1 Example K-means algorithm operation sequence

데이터 전처리 과정은 먼저 거리 기반 이상치 제거를 수행하고, 각 좌표는 0~1 범위로 정규화를 진행하였다. 반사율 값은 극단값을 필터링하여 생물체와 지형 간의 구분 가능성을 높였다. 이후 K-Means, DBSCAN, HDBSCAN 세 가지 클러스터링 알고리즘을 동일한 데이터셋에 적용하였다. K-Means는 클러스터 수 $k = 4$ 로 고정하고, DBSCAN은 반경 $\epsilon = 0.4$, 최소 포인트 수 $\text{minPts} = 6$ 으로 설정하였다. HDBSCAN은 클러스터 수를 자동을 추정하며, 최소 클러스터 크기 $\text{min_cluster_size} = 20$ 을 기준으로 적용하였다.

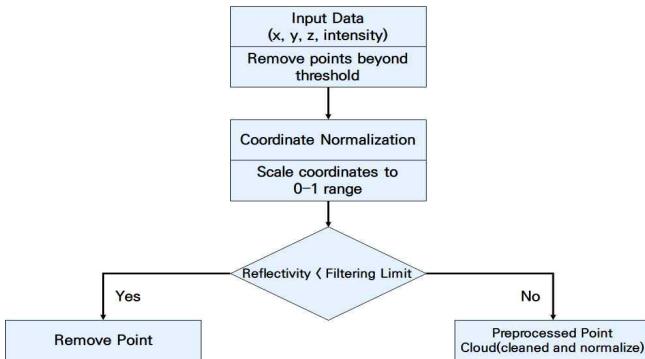


그림 2 데이터 전처리 과정 플로우차트

fig 2 Data preprocessing flowchart

알고리즘의 성능은 Silhouette Score와 Davies - Bouldin Index를 기준으로 평가하였다. Silhouette Score는 각 군집의 응집도와 분리도를 종합적으로 측정하며, 1에 가까울수록 군집 품질이 높다. Davies - Bouldin Index는 군집 간 거리 대비 내부 분산을 반영한 지표로, 값이 낮을수록 성능이 우수함을 의미한다.

III. 실험 및 결과 분석

본 장에서는 앞서 설계한 전처리 데이터를 기반으로, 각 클러스터링 알고리즘의 적용 결과를 비교·분석한다. 실험은 K-Means, DBSCAN, HDBSCAN 세 가지 클러스터링 알고리즘을 적용하여 해양 생물 분포 탐지에 적합한 군집화 기법을 비교·분석하였다. 실험 결과를 종합하여 다음과 같은 최종 결론을 도출하였다.

그림 3은 전처리된 포인트 클라우드에 대해 세 가지 클러스터링 알고리즘을 적용한 결과를 비교한 것이다. 각 그래프는 다음과 같은 특징을 보인다.

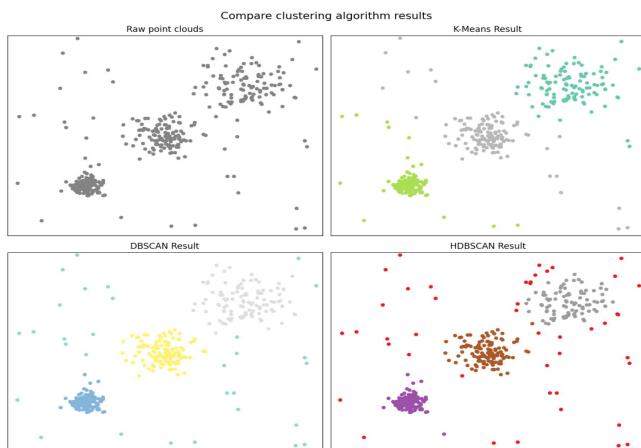


그림 3 클러스터링 결과 비교 시각화

fig 3 Visualize a comparison of clustering results

K-Means는 군집 수가 고정되어 있어 밀도나 형태에 관계없이 원형 분할된 결과를 보였으며, 노이즈에 취약한 경향을 나타냈다. DBSCAN은 복잡한 군집 경계를 잘 식별하였으며, 노이즈 포인트를 효과적으로 제외하였다. HDBSCAN은 DBSCAN과 유사하지만 자동으로 클러스터 수를 조정하며, 희소하거나 불규칙한 구조에서도 안정적으로 군집을 형성하였다.

표 1은 Silhouette Score 및 Davies - Bouldin Index(DBI)를 기준으로 각 알고리즘의 군집 품질 비교 결과를 정량적으로 나타낸 것이다.

| 알고리즘 | Silhouette Score | DBI |
|---------|------------------|------|
| K-Means | 0.56 | 0.84 |
| DBSCAN | 0.62 | 0.65 |
| HDBSCAN | 0.68 | 0.49 |

표 1 각 알고리즘 군집 품질 비교 정량적 결과

Table 1 Quantitative results comparing the quality of each algorithmic population

Silhouette Score 기준으로는 HDBSCAN이 가장 높은 군집 응집도와 분리도를 보였으며, DBI 역시 HDBSCAN이 가장 낮은 값을 기록하여, 군집 간 경계가 명확하고 내부 일관성이 높은 것으로 분석되었다.

실험 결과를 종합하면 다음과 같은 결론을 도출할 수 있다. K-Means는 간단하고 계산 속도가 빠르지만, 수중 포인트 클라우드처럼 군집 수가 불명확하고 밀도 편차가 존재하는 환경에서는 적절하지 않다. DBSCAN은 노이즈에 강하고 밀도 기반 탐지가 가능하나, 파라미터 설정에 민감하고 희소한 군집은 인식하지 못하는 경우가 있었다. HDBSCAN은 자동 군집 수 추정과 유연한 경계 처리가 가능하여, 수중 라이다와 같이 복잡한 공간 데이터에서 가장 우수한 성능을 보였다.

IV. 결론

본 연구에서는 수중 라이다 기반 포인트 클라우드 데이터를 대상으로, K-Means, DBSCAN, HDBSCAN 세 가지 클러스터링 알고리즘을 적용하여 해양 생물 분포 탐지에 적합한 군집화 기법을 비교·분석하였다. 실험 결과를 종합하여 다음과 같은 최종 결론을 도출하였다.

K-Means는 간단하고 계산 속도가 빠르다는 장점이 있으나, 군집 수가 고정되어 있어도 밀도 편차가 존재하는 수중 환경에서는 적합하지 않았다. DBSCAN은 노이즈에 강하고 복잡한 경계를 인식하는 데 효과적이었지만, 파라미터 설정에 민감하고 희소한 군집 탐지에는 한계가 있었다. HDBSCAN은 클러스터 수를 자동으로 추정하고, 불균일한 밀도 분포에서도 안정적으로 군집을 형성하여 가장 우수한 성능을 보였다. 이러한 결과를 통해, HDBSCAN이 해양 생물 분포 탐지에 가장 적합한 클러스터링 기법임을 확인할 수 있었다. 향후에는 실제 수중 라이다 센서를 통해 수집된 실환경 데이터에 본 알고리즘을 적용하고, 수심, 혼탁도, 반사율 변화 등 다양한 해양 조건에 대한 대응성을 높이기 위한 파라미터 자동 조정 기법 개발이 필요할 것이다. 또한, 실시간 모니터링 및 해양 생물 이상 탐지와 연계한 응용 연구도 추가적으로 가능하다.

ACKNOWLEDGMENT

본 연구는 산업통상자원부의 2023년 ‘남해안권 무인이동체 모니터링 및 실증 기반 구축 사업’ 과제로 (RS-2023-KI002744) 지원되어 수행되었습니다.○후

참 고 문 헌

- [1] 한국해양수산개발원, 「해양생태계 보전관리 기본계획 수립 연구」, 2016.
- [2] WWF, Living Blue Planet Report, 세계자연기금, 2015.
- [3] LiDAR Tech Pros, “Underwater Lidar Drones: A Dive into Subaqueous Mapping,” LiDAR Tech Pros, 2023.
- [4] Seabee, “Uncovering the Ocean’s Secrets with Drones and 3D Technology,” Seabee.no, 2022.
- [5] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, “PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 652 - 660.