

심층 강화학습 기반 적응적 SAR 이미지 필터링 알고리즘 설계

김태윤*, 정소이*, 김재현^o

Deep Reinforcement Learning Based Adaptive Synthetic Aperture Radar Image Filtering Algorithm

Tae-Yoon Kim*, Soyi Jung*, Jae-Hyun Kim^o

요약

Low Earth orbit (LEO) 위성 synthetic aperture radar (SAR)는 지구관측의 효율적인 기술로 최근 많은 연구가 진행 중이다. 하지만 LEO 위성의 시간 제한적 특성상, SAR 원본 이미지의 효율적인 처리에 대한 연구는 필수적이다. 본 논문에서는 LEO SAR 이미지의 효율적인 처리를 위한 심층 강화학습 기반 적응형 스펙클 잡음 필터링 알고리즘을 제안한다. 제안하는 알고리즘은 버퍼 상태에 따라 필터 크기를 적응적으로 선택하여 제한된 시간에 최대한의 이미지 해상도를 도출한다. 시뮬레이션 결과, 제안하는 알고리즘은 다른 알고리즘과 비교하였을 때 버퍼 상태를 고려하여 최적의 필터 선택을 하는 것을 확인하였다.

Key Words : Low Earth Orbit (LEO), Synthetic Aperture Radar (SAR), Speckle Image Filtering, Deep Reinforcement Learning (DRL), Buffer State

ABSTRACT

Low Earth orbit (LEO) satellite synthetic aperture radar (SAR) is an efficient technology for Earth

observation, and a lot of research is in progress. However, due to the time-limited property of LEO satellites, research on time-efficient SAR image processing is essential. In this paper, we propose an adaptive speckle noise filtering algorithm based on deep reinforcement learning for efficient processing of LEO SAR images. The proposed algorithm adaptively selects the filter size according to the buffer state to derive the maximum image resolution in a limited time. As a result of the simulation, the proposed algorithm selects the filter size more efficiently according to the buffer state than the method of conventional algorithm.

I. 서론

Low Earth orbit (LEO) 위성 synthetic aperture radar (SAR)는 기상 조건과 관계없이 24시간 작동할 수 있어 객체 탐지, 재난 관측 등 다양한 분야에서 활용된다^[1]. 하지만 LEO 위성과 지상국 사이의 통신 시간은 제한적이고 SAR 원본 이미지에는 물체 감지를 어렵게 만드는 스펙클 잡음이 있다. 이러한 스펙클 잡음을 줄이는 이미지 처리는 많은 연구가 진행되고 있지만^[2] LEO 위성과 지상국 사이의 제한적인 통신 시간을 고려한 효율적인 이미지 처리에 관한 연구도 필요하다.

따라서, 본 논문에서는 LEO 위성의 제한적인 통신 시간을 고려한 강화학습 기반 적응형 SAR 이미지 스펙클 잡음 필터링 알고리즘을 제안한다. 제안하는 알고리즘은 SAR 이미지 처리에 크기가 큰 필터를 사용할수록 더욱 정확한 이미지 처리가 가능하지만, 처리 시간이 길어진다는 점에 주목한다. 강화학습을 기반으로 제한된 시간 동안 버퍼 상태에 따라 적절한 필터 크기를 선정하도록 학습하여 효율적인 이미지 처리를 가능하게 한다. 성능분석 결과, 제안하는 강화학습 기반 알고리즘이 기존 다른 두 가지 알고리즘과 비교하여 더욱 효과적인 이미지 처리가 가능하다는 결과를 보였다.

* 본 연구는 한국연구재단의 지원을 받아 수행된 기초연구사업(No. 2021R1A4A1030775)과 정보통신기획평가원(IITP-2022-2018-0-01424)의 지원을 받아 수행된 연구입니다.

• First Author : (ORCID:0000-0003-1559-9514) Ajou University Department of Artificial Intelligence Convergence Network, xodbsxogjs@ajou.ac.kr, 학생(석박통합과정), 정회원

◦ Corresponding Author : (ORCID:0000-0003-4716-6916) Ajou University Department of Electrical and Computer Engineering, jkim@ajou.ac.kr, 정교수, 중신회원

* (ORCID:0000-0001-8435-0646) Ajou University Department of Electrical and Computer Engineering, sjung@ajou.ac.kr, 조교수, 정회원

논문번호 : 202210-262-A-LU, Received October 31, 2022; Revised November 11, 2022; Accepted November 11, 2022

II. 시스템 모델

본 논문에서 고려하는 시스템 모델은 그림 1과 같이 하나의 LEO 위성과 1개의 지상국으로 구성되며 특정 시간 동안 LEO 위성은 지상국에 매초마다 λ 만큼의 SAR 원본 이미지를 전송한다. 지상국은 유한한 크기의 버퍼를 가지고 LEO 위성으로부터 받은 SAR 원본 이미지를 버퍼에 저장하여 순서대로 이미지 필터링을 거친다. 지상국에서 SAR 원본 이미지를 처리하는 속도는 이미지 처리를 위해 사용한 필터의 크기에 따라 달라진다. 따라서 제한된 시간 동안 효과적인 SAR 이미지 처리를 위해 지상국은 버퍼의 상태를 고려한 필터 선택이 필요하다.

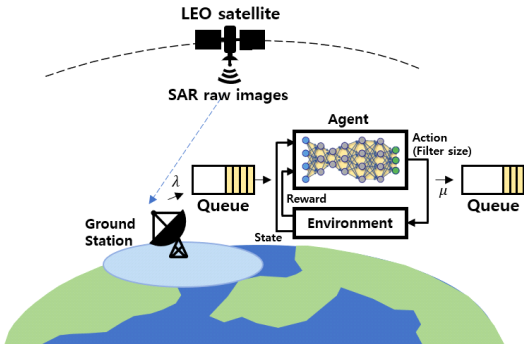


그림 1. 시스템 모델
Fig. 1. System Model

III. MDP 정형화 및 심층 강화학습 프레임워크

지상국의 버퍼 활용을 극대화하기 위해서는 현재 버퍼 상태를 기반으로 필터 크기를 결정해야 한다. 이러한 문제는 Markov Decision Process (MDP)로 공식화될 수 있다. MDP는 일반적으로 상태 (S), 행동 (A), 보상 (R), 천이 확률 (P), 감가율 (γ) 5개의 요소로 구성되는데, 본 논문에서는 상태 천이 확률을 알지 못하는 partially observable MDP (POMDP)로 다음과 같이 구성된다.

시간 t 에서의 S 는 다음과 같이 구성된다.

$$S_t = (GS_t, L_t), \quad (1)$$

GS_t 는 지상국의 버퍼 상태, L_t 는 LEO 위성과 지국이 통신 가능한 시간을 의미한다.

시간 t 에서의 A 는 다음과 같이 구성된다.

알고리즘 1. 제안하는 DQN 기반 최적 필터 선택 알고리즘
Algorithm 1. The proposed optimal filter select algorithm based on DQN

Algorithm 1 제안하는 DQN 기반 최적 필터 선택 알고리즘

- 1: 초기값: 행동-가치 함수 Q 의 랜덤 가중치 θ , 목표 행동-가치 함수 Q 의 가중치 $\theta^- = \theta$, 리플레이 버퍼 B 의 용량 C , minibatch 크기 M , 총 에피소드 수 E ,
- 2: 입력값: 상태 S , 행동 A , 감가율 γ
- 3: 출력값: Median 필터 크기 MF
- 4: for Episode = 1 to E do
- 5: GS 버퍼 상태 초기화
- 6: for timestep : $t = 1$ to T do
- 7: Decaying ϵ -greedy 방법을 통한 행동 a_t 선택
- 8: $= \begin{cases} \operatorname{argmax}_a Q(s_t, a; \theta) & , \text{ 확률 } 1 - \epsilon \\ \text{임의의 행동} & , \text{ 확률 } \epsilon \end{cases}$
- 9: 샘플 (s_t, a_t, r_t, s_{t+1}) 를 B 에 저장
- 10: if $B > C$ then
- 11: B 로 부터 랜덤한 minibatch (s_j, a_j, r_j, s_{j+1})
- 12: 샘플 선택
- 13: for $j = 1$ in M do
- 14: $y_j = r_j + \gamma \max_{a'} \hat{Q}(s_{j+1}, a'; \theta^-)$
- 15: end for
- 16: 가중치 θ 의 관점에서 경사 하강 단계
- 17: $(y_j - Q(s_j, a_j; \theta))^2$ 수행
- 18: 매 스텝마다 Q 의 파라미터를 \hat{Q} 로 복사
- 19: end if
- 20: end for
- 21: end for

$$A_t = (MF_t), \quad (2)$$

MF_t 는 지상국이 SAR 원본 이미지 처리를 위해 선택한 median 필터의 크기이다. median 필터는 SAR 원본 이미지의 품질을 저하하는 스�কে를 잡음을 제거하는 필터로 다음과 같이 표현된다.

$$\hat{f}(x, y) = \frac{1}{N^2} \sum_{s=0}^{N-1} \sum_{t=0}^{N-1} g(s, v), \quad (3)$$

\hat{f} 와 g 는 각각 복원된 이미지와 스�কে를 잡음 이미지이고 $N \in \{3, 5, 7, 11, 13\}$ 은 선택할 수 있는 필터 크기를 의미한다. x, y 는 원본 이미지의 좌표, s, v 는 필터의 좌표를 나타낸다.

시간 t 에서의 버퍼의 활용을 최대화하기 위해 지상국이 버퍼 상태에 따라 적절한 필터 크기를 선정하도록 다음과 같이 R 을 정의한다.

$$R_t = QI_t - \alpha(GS_t/BS) + \zeta \quad (4)$$

QI_t 는 선택된 필터에 따른 이미지 품질을 의미하고, α 는 현재 버퍼 상태와 이미지 품질 사이의 계수이다. BS 는 지국 버퍼의 정해진 크기를 의미하고, ζ

는 처벌 함수로 지상국의 최대 버퍼 크기를 넘어가면 처벌하도록 설정한다.

γ 는 0에서 1 사이의 값으로 미래에 받은 보상을 현재 시점에서 고려할 때 감가하는 비율을 의미한다. 앞서 정의한 POMDP의 효과적인 최적화를 위해 본 논문에서는 deep Q-Network (DQN) 모델³⁾을 기반으로 한 심층 강화학습 알고리즘을 제안한다. 기존 알고리즘에서 사용한 convolutional neural network (CNN) 구조의 학습 방식을 3개의 은닉층과 rectified linear unit (ReLU) 활성화 함수로 구성된 deep neural network (DNN)으로 수정하였으며 시간이 지남에 따라 점차 epsilon 값이 감소하는 decaying epsilon-greedy 방법을 사용한다. 제안하는 알고리즘은 알고리즘 1과 같이 동작한다.

IV. 성능 평가

4.1 시뮬레이션 환경

본 논문에서는 MATLAB과 PYTHON을 활용하여 시뮬레이션을 수행하였다. 에피소드는 10,000, replay buffer의 크기 B는 1,000,000, buffer capacity C는 100,000, mini batch 크기는 32, time step은 LEO 위성 통신 가능 최대 시간을 고려하여 600초로 설정하였다. α 는 각 필터 크기마다 처리 시간, 품질 지표에 따라 설정하였으며⁴⁾, λ 는 정규분포를 따라 SAR 원본 이미지를 생성하도록 설정하였다. 성능분석을 위해 두 가지 필터 선택 방식을 시뮬레이션에 고려하였다.

- Random-Selection: 지상국의 버퍼 상태와 관계없이 항상 랜덤한 필터 크기를 선택
- Max-Selection: 지상국 버퍼가 꽉 차기 전까지 가장 큰 크기의 필터를 우선순위로 선택

4.2 시뮬레이션 결과

그림 2 (a), (b)는 각각 지상국 버퍼의 크기가 2,000, 4,000일 때 시간에 따라 지상국이 선택하는 필터 크기를 나타내고, (c)는 지상국 버퍼의 크기가 2,000과 4,000일 때, 시간에 따른 버퍼 상태를 나타낸다. Random-Selection의 경우 항상 랜덤한 필터 크기를 선택하기 때문에 버퍼에 데이터가 불안정하게 쌓이고, 버퍼의 포화 상태에서도 크기가 큰 필터를 선택하는 비효율적인 모습을 보였다. Max-Selection의 경우 Random-Selection보다 안정적인 모습을 보이지만 버퍼가 꽉 차기 전까지 가장 큰 크기인 13×13 필터만 선택하기 때문에 빠르게 버퍼가 포화 된다. 제안하는 방식의 경우 버퍼 상태에 따라 유동적으로 필터 크

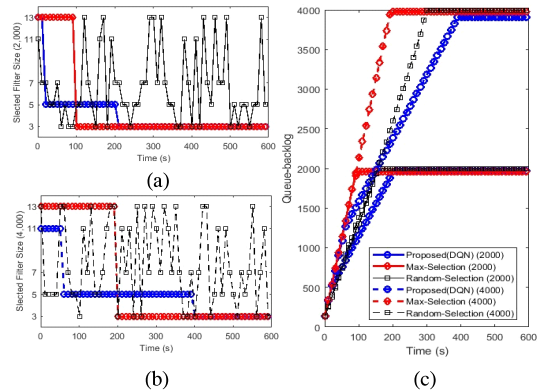


그림 2. (a) 버퍼 크기 2,000, (b) 4,000일 때, 시간에 따른 선택된 필터 크기, (c) 버퍼 크기 2,000과 4,000일 때 시간에 따른 버퍼 상태
Fig. 2. (a) buffer size 2,000, (b) 4000, selected filter size based on time, (c) buffer state based on time, when buffer size is 2,000 and 4,000

기를 선택하는 모습을 볼 수 있었다. 상대적으로 버퍼가 여유 있는 상황에서는 최대한 큰 필터를 선택하고, 점점 버퍼가 포화 되어감에 따라, 그림 2. (a) 약 10초, (b) 약 50초일 때 좀 더 작은 크기인 5×5 필터를 고르고, 그림 2. (a) 약 200초, (b) 약 400초, 포화 직전의 상황에서는 버퍼가 오버 되지 않게 최소한의 필터 크기 3×3을 선택하여 버퍼를 효율적으로 사용하는 것을 확인할 수 있었다. 즉, 지상국 버퍼의 최대 크기가 증가하여도 상황에 맞게 최적의 필터를 선택하는 결과를 확인할 수 있었다.

V. 결론

본 논문에서는 LEO 위성의 제한적인 통신 시간을 고려한 심층 강화학습 기반 적응형 SAR 이미지 스펙클 잡음 필터링 알고리즘을 제안하였다. 제안하는 방식은 심층 강화학습 알고리즘을 통해 버퍼 상태를 기반으로 최적의 필터 크기를 선택하여 다른 두 가지 방식과 비교하여 더 안정적인 버퍼 관리가 가능했다. 또한 심층 강화학습을 통해 버퍼 최대 크기에 알맞은 학습을 진행하여 최적화된 버퍼 관리가 되는 것을 확인하였다.

References

[1] S. Paek, S. Balasubramanian, S. Kim, and O. Weck, "Small-satellite synthetic aperture radar for continuous global biospheric monitoring: A

- review,” *Remote Sensing*, vol. 12, no. 2546, Aug. 2020. (<https://doi.org/10.3390/rs12162546>)
- [2] F. Lattari, B. Gonzalez Leon, F. Asaro, A. Rucci, C. Prati, and M. Matteucci, “Deep learning for sar image despeckling,” *Remote Sensing*, vol. 11, no. 1532, Jun. 2019. (<https://doi.org/10.3390/rs11131532>)
- [3] V. Mnih et al., “Human-level control through deep reinforcement learning,” *Nature*, vol. 518, pp. 529-533, Feb. 2015. (<https://doi.org/10.1038/nature14236>)
- [4] K. Kim, S. Jung, and J. Kim, “Adaptive speckle filtering for real-time computing in low earth orbit satellite synthetic aperture radar,” *ICT Express*, vol. 7, pp. 187- 190, Jun. 2021. (<https://doi.org/10.1016/j.ict.2021.02.003>)