

VAE-GMM 기반 생체모방 통신용 휘슬 시퀀스 생성 방법

박근호, 안종민, 김완진, 김인수, 이동훈

국방과학연구소

ghpark57935@add.re.kr, leedhun@add.re.kr

VAE-GMM-based Whistle Sequence Generation Method for Biomimetic Underwater Communication

Park Geun-Ho, Ahn Jongmin, Kim Wanjin, Kim In-Soo, Lee Donghun

Agency for Defense Development

요약

본 논문은 생체모방 수중통신용 인공 휘슬음의 시퀀스 생성 방안을 제시한다. 본 논문에서 제안하는 생체모방용 휘슬의 시퀀스 생성 방법은 크게 심층 생성 신경망 기반 휘슬의 시간-주파수 생성, 가우시안 컴포넌트 별 주파수 할당, 휘슬 간 간격 모델링의 세 단계로 구성된다. 휘슬 형태는 변이형 오토인코더와 혼합 가우시안 밀도를 학습하여 생성하도록 구현했다. 중심 주파수는 각 가우시안 컴포넌트를 생성하는 데에 기여한 휘슬의 주파수 중에서 단순임의추출 (Simple Random Sampling)하여 결정한다. 마지막으로 휘슬 간 간격은 휘슬 추출 레이블로부터 휘슬 간 간격 분포를 일반화된 파레토 분포 (Generalized Pareto Distribution)로 모델링한 후 샘플링하였다.

I. 서론

돌고래와 고래는 휘슬음과 광대역 펄스 형태의 클릭음을 활용하여, 의사소통 및 먹이탐색 등을 수행하는 것으로 알려져 있다. 특히 고래류는 고정된 주파수 변조 방법을 적용한 시그니처 휘슬 신호로 자신을 알리거나, 상황에 적합한 휘슬음을 발생시켜 의사소통에 활용하기도 한다.

생체모방 수중통신은 고래류가 발생시키는 휘슬음을 주로 모방한다[1]. 감청자는 휘슬음 모방 통신 신호를 통신 신호로 인식하기 어려우므로, 아군의 은밀 통신을 가능하게 한다.

고래의 휘슬음의 시간-주파수 형태는 매우 다양하므로, 결정론적이지 않고 확률 분포로 나타내는 것이 적절하다. 그러므로 결정론적 심볼을 사용하는 기존 디지털 통신과는 다른 접근이 필요하다.

이러한 접근 방법의 하나로, 생체모방 통신 신호의 송신기는 휘슬의 확률 분포에서 표본을 얻는 방법으로 송신 신호를 결정할 수 있다. 이 방법의 이점은 감청자가 해당 신호를 수집하더라도, 아군의 신호 송신 체계를 파악하기 위해 많은 자료를 수집하고 분석하도록 강요한다는 점이다.

휘슬음을 확률적으로 생성하는 방법은 심층 생성 신경망을 적용하는 것이다. 본 논문에서는 [2]에서 제안한 Variational Autoencoder (VAE)와 Gaussian Mixture Model (GMM)을 학습하여 휘슬 형태를 샘플링하도록 하였다.

본 논문의 목표는 단일 휘슬 형태를 모방하는 것을 넘어서 휘슬음을 연속적으로 생성한 결과도 데이터 분포에 근접하도록 설계하는 것이다. 이를 위해서는 휘슬의 중심 주파수와 휘슬음 간격까지 고려해야 한다.

하지만 단일 휘슬 형태만 생성하는 VAE-GMM은 중심 주파수와 연속적 휘슬 간격을 고려하지 않으므로, 주파수와 휘슬 간격을 생체 신호와 유사하게 샘플링할 수 있는 별도의 방안이 필요하다. 따라서 본 논문에서는 연속 휘슬 시퀀스 생성 문제에서 휘슬의 주파수는 단순임의추출 (Simple Random Sampling)하도록 설정했고, 휘슬음 간격의 확률 분포는 간격 데이터를 활용해 일반화된 파레토 분포 (Generalized Pareto Distribution)로 모델링하였다. 관련 세부내용은 II장부터 IV장에서 다룬다.

II. VAE-GMM 기반 휘슬음 생성 방법

VAE-GMM 기반 휘슬 생성 방법[2]은 휘슬음의 형태를 이미지 형태로 표현한 입력 데이터에 대해 잠재 변수를 추정하는 VAE와 잠재 변수의 주변 분포를 다수의 가우시안 분포로 모델링하는 GMM을 적용한 것이다. 휘슬을 이미지로 표현했을 때 대부분 0에 근접하는 픽셀 강도를 가지므로, 128×128 크기의 이미지를 64×1 잠재 벡터로 압축할 수 있었다. 잠재 변수의 분포, 즉 생성 모델에서의 사전 분포는 10,000개의 가우시안 분포로 모델링 하였다. 가우시안 분포의 매개변수는 Expectation-Maximization (EM) 알고리즘으로부터 추정하였다. 또한 VAE 학습 시, Wasserstein Generative Adversarial Network-Gradient Penalty (WGAN-GP)를 적용하여 데이터와 유사한 형태가 되도록 신경망을 학습했다. 자세한 내용은 [2]를 참고할 수 있다.

III. 휘슬 간격 및 주파수 샘플링 방법

휘슬 간 간격에 대한 데이터는 미국 해양대기청에서 공개하는 Detection, Classification, Localization and Density Estimation (DCLDE) Oahu [3]의 혹범고래 데이터에서 추출하였다. 먼저 DCLDE 2022 혹범고래 데이터의 레이블 작업을 통해 32,178개 휘슬의 시간-주파수 데이터를 확보한 후, 이 휘슬의 시작 시간 정보로부터 31,991개의 간격 샘플을 얻었다. 휘슬음 간격의 히스토그램은 그림 1에 나타내었다.

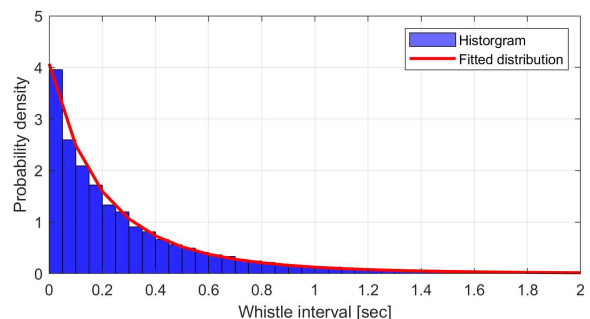


그림 1. 휘슬 간격 데이터 및 확률 분포 추정 결과.

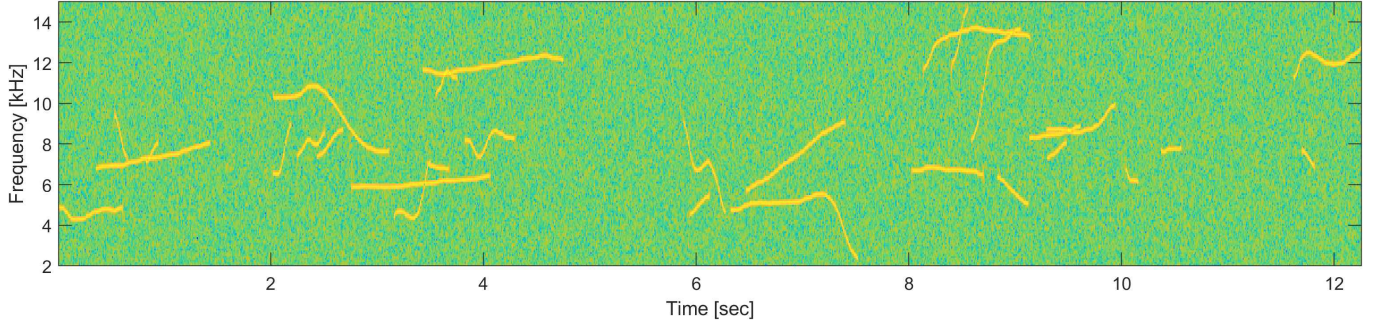


그림 2. VAE-GMM 기반 휘슬 시퀀스 생성 결과 1

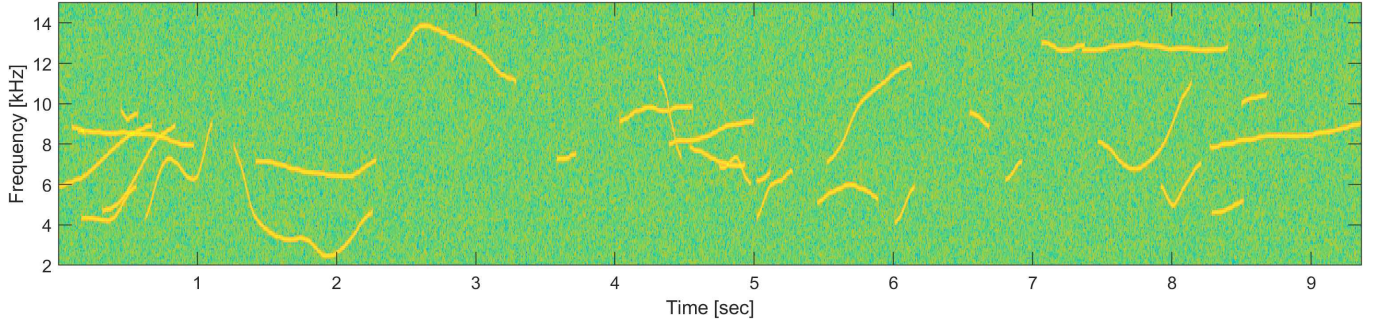


그림 3. VAE-GMM 기반 휘슬 시퀀스 생성 결과 2

본 논문은 그림 1의 휘슬 간격 데이터로부터 적색 실선의 확률 분포를 추정하였다. 확률 분포 모델은 일반화 파레토 분포 (Generalized Pareto Distribution)를 적용하였고, 이 분포의 shape, scale 및 threshold 파라미터 k , σ , θ 는 각각 0.2753, 0.2458, 0과 같다.

휘슬의 중심 주파수는 GMM의 각 가우시안 컴포넌트에 주파수 샘플을 할당한 후, 단순임의추출 하도록 설정하였다. 예를 들어, 5,113번째 가우시안 분포를 생성하는 데에 기여한 5개의 휘슬 데이터가 있고, 이 데이터의 중심 주파수가 각각 6.9, 7.0, 13.8, 14.0, 4.9 kHz라면, 5,113번째 가우시안 분포로부터 휘슬을 생성할 때에 중심 주파수는 이 5개 중 하나가 임의 추출되도록 하였다.

이와 같이 주파수 생성을 위해 단순임의추출을 적용한 이유는 각 가우시안 컴포넌트에 할당된 데이터의 수가 적은 경우가 매우 많다는 점, 그리고 주파수 데이터가 가우시안 분포에 적합하지 않다는 점 때문이다. 실제 64×1 크기의 잠재 변수에 주파수 데이터 차원을 추가하여 65×1 크기의 벡터를 GMM으로 학습하도록 했을 때, 중심 주파수에 대한 가우시안 분포의 분산이 매우 커지는 한계가 발생했다. 따라서 중심주파수에 대해 단순임의추출 방법이 가장 효과적이라고 판단하였다.

IV. 휘슬 시퀀스 생성 결과

그림 2와 3은 다음 세 단계를 거쳐 30개의 휘슬을 연속 생성한 결과를 나타낸 것으로 다음 세 단계를 거친다.

- (1) VAE-GMM 기반 휘슬 형태 샘플링
- (2) 단순임의추출 기반 주파수 샘플링
- (3) 일반화된 파레토 분포에서 휘슬 간격 샘플링

그림 2와 3에서 신호대잡음비는 -5 dB, 샘플링 주파수는 192 kHz, 스펙트로그램 FFT 크기는 2,048, 중첩 샘플 수는 1,024로 설정하였고, Hamming 윈도우를 적용하였다.

V. 결 론

본 논문에서는 VAE-GMM 기반 휘슬 형태 샘플링, 단순임의추출 기반 주파수 샘플링, 일반화된 파레토 분포 기반 휘슬 간격 샘플링의 세 단계를 통해 휘슬 시퀀스를 생성한 결과를 제시하였다. 연속적인 휘슬 생성 결과, 데이터와 유사한 형태를 갖는 것을 확인할 수 있었다. 향후에는 휘슬 합성 결과와 원 데이터 사이의 확률 분포 차이를 수치적으로 분석하여 도출할 계획이다.

ACKNOWLEDGMENT

이 논문은 2025년 정부의 재원으로 국방과학연구소에서 수행된 연구 결과임.

참 고 문 헌

- [1] Seungwhan Seol, *et. al.*, "Research trends of biomimetic covert underwater acoustic communication," *The Journal of the Acoust. Soc. of Korea*, vol. 41, no. 2, pp. 227-234, 2022.
- [2] G.-H. Park, J. Ahn, W. Kim, I. Kim, and D.-H. Lee, "A Generative Model for Cetacean Whistles Using Variational Autoencoder and Mixture of Gaussians," *IEEE Access* (to be submitted).
- [3] NOAA National Centers for Environmental Information (2017). *Passive Acoustic Data Collection*. Asheville, NC: NOAA National Centers for Environmental Information. doi: 10.25921/PF0H-SQ72.