

뉴로모픽 시스템을 위한 인코딩 및 디코딩 인터페이스 구축

탁 동 혁, 임 재 한*
광운대학교, *광운대학교

tdh8316@kw.ac.kr, *ljhar@kw.ac.kr

Implementation of encoding and decoding interfaces for neuromorphic systems

Donghyeok Tak, Jae-Han Lim*
Kwangwoon Univ., *Kwangwoon Univ.

요 약

뉴로모픽 시스템은 뇌가 정보를 처리하는 방식을 모방한 컴퓨팅 시스템이다. Spiking Neural Networks(SNNs)는 뇌의 정보 처리 방식에 관한 모델을 제공하므로 뉴로모픽 시스템에서 중요한 위치를 차지한다. 하지만 SNN은 기존의 인공 신경망과는 다르게 스파이크 형태로 데이터를 처리하고 저장한다. 이러한 특성 때문에 SNN에 기반하여 구현된 뉴로모픽 시스템은 처리할 데이터를 스파이크 형태로 변환하는 뉴럴 인코딩 과정이 필요하다. 특히 하드웨어로 구현된 뉴로모픽 시스템에서는 해당 과정을 실시간으로 수행하는 별도의 프로세서가 필요하다. 따라서 본 연구에서는 실시간으로 데이터를 스파이크로 변환하고, 스파이크 형태의 출력을 해석하는 뉴럴 인코딩 및 디코딩 인터페이스를 구축하였다. 그리고 이를 사용해 사각파와 톱니파를 분류하는 실험을 수행하였다. 이를 통해 하드웨어로 구현된 뉴로모픽 시스템과 인터페이스간의 상호작용이 가능함을 확인함으로써 본 연구에서 구축한 인터페이스가 뉴로모픽 시스템에 사용될 수 있음을 보인다.

I. 서 론

최근 Spiking Neural Networks(SNNs)에 기반한 뉴로모픽 시스템이 활발하게 연구되고 있다. SNN은 뇌가 작동하는 방식을 모방한 인공신경망으로, 가장 주목받는 특징은 에너지 효율성이다. 퍼셉트론에 기초한 기존의 인공신경망은 데이터를 입력받을 때마다 가중치가 변경되어 학습이 이루어진다. 하지만 이와는 대조적으로 SNN은 뉴런의 막 전위가 임계치에 도달해야 뉴런이 스파이크를 방출하고 시냅스 가중치가 변경된다. 그러므로 데이터가 입력될 때마다 매번 가중치가 갱신되지는 않는다. 따라서 SNN은 기존의 인공신경망에 비하여 에너지 효율성이 높은 것으로 알려져 있다[1]. 또한 SNN은 뇌의 효율적인 정보 처리 방식에 관한 모델을 제공하므로 뇌의 작동 방식을 모방하는 뉴로모픽 컴퓨팅에서 중요한 위치를 차지한다.

뉴로모픽 컴퓨팅과 관련해서, 기존의 폰 노이만 아키텍처나 FPGA에 기반하지 않는 하드웨어 또한 활발히 연구되고 있다[2]. 이러한 하드웨어로는 기존의 컴퓨팅 아키텍처에서보다 뉴로모픽 컴퓨팅을 더욱 효율적으로 구현할 수 있다. SNN에 에너지 효율성을 부여하는 것은 뉴런이 이산적인 스파이크 형태로 데이터를 처리하고, 뉴런의 막 전위가 임계치를 넘어야 데이터가 전달된다는 특성 때문이다. 이는 기존의 인공신경망이 실수 데이터를 처리하며 데이터 입력 시마다 출력이 발생하는 것과는 대조된다. 그러나 이러한 SNN의 특성 때문에, 실수 데이터를 SNN에서 사용하는 스파이크 형태로 변환하는 과정이 필요하다. 또한 SNN의 출력도 스파이크 형태이므로 이를 해석할 수 있는 과정도 필요하다. 이를 각각 뉴럴 인코딩, 뉴럴 디코딩이라고 하며 이 과정은 SNN 기반의 뉴로모픽 시스템에서도 필요하다.

따라서 하드웨어로 구현된 뉴로모픽 시스템을 사용하기 위해서는 뉴럴 인코딩 및 디코딩을 수행하여 데이터를 변환하는 인터페이스 장치가 필요하다. 이에 본

연구에서는 이를 수행하는 인터페이스의 작동 과정을 정의하고, 뉴로모픽 칩과 해당 인터페이스 간의 연동 실험을 진행한다. 그리고 이 실험을 통해 하드웨어로 구현된 뉴로모픽 시스템이 높은 정확도로 사각파와 톱니파를 구분하는 것을 보인다. 이는 뉴로모픽 하드웨어가 SNN을 기반으로 제대로 작동함을 의미한다. 더해서 향후 뉴로모픽 시스템에 적용할 수 있는 인터페이스 기술을 제안한다.

II. 본론

1. 인터페이스 구현

본 연구에서 사용한 인터페이스는 인코더와 디코더로 구성되는데, 모두 라즈베리파이를 사용해 구현되었다. 스파이크의 시간차는 뉴로모픽 시스템에서 굉장히 중요한 요소이므로, 스파이크를 정밀하게 생성해야 한다. 그러므로 인터페이스의 소프트웨어는 실시간성을 보장받을 필요가 있는데, 이를 위해 실시간 리눅스 커널을 사용했다. 인터페이스에서 사용할 데이터셋은 Universal Software Radio Peripheral(USRP)에서 생성한 Radio Frequency(RF)대역의 사각파와 톱니파를 양자화해서 생성했다.

인터페이스는 파동의 진폭 값을 한 주기당 4개 샘플로 샘플링하여 인코딩한다. 먼저 샘플링하여 얻은 0 이상 1 이하의 진폭값에 5를 곱한 값을 전압으로 취한다. 그리고 500us 동안 해당 전압을 인가하고, 다시 500us 동안 전압을 인가하지 않는다. 또한 인터페이스는 각 출력 뉴런이 하나의 데이터 군집을 대표하는 것으로 간주한다. 따라서 뉴로모픽 칩의 각 출력 뉴런에 데이터의 종류를 의미하는 라벨을 부여한다. 라벨을 부여하는 기준은 뉴런의 발화 정도이다. 어떠한 입력 데이터에 관하여 가장 많이 발화한 출력 뉴런에 그 데이터의 라벨을 부여한다. [그림 1]은 구현된 인터페이스 장치의 모습이다.

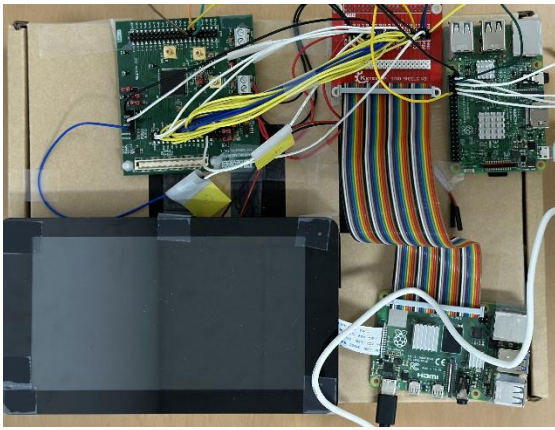


그림 1 - 인터페이스 장치

2. 실험 설정

본 연구에서는 두 가지 실험을 진행하였다. 첫 번째로 인터페이스의 인코더에서 출력 신호를 받은 하드웨어 뉴런이 발화하는지 확인하였다. 두 번째로 뉴로모픽 하드웨어가 인터페이스에서 받은 신호로부터 입력 스파이크를 생성하고, 이를 통해 사각파와 톱니파를 올바르게 구분할 수 있는지 실험하였다.

RF 대역의 사각파와 톱니파를 구분하는 실험의 경우 각각 4 개의 Leaky Integrate-and-Fire(LIF) 모델을 따르는 하드웨어 뉴런을 사용하였다. 그리고 소프트웨어로 구현한 SNN 에서 동일한 데이터를 학습하여 얻은 시냅스 가중치를 하드웨어 시냅스에 적용하였다. 또한 출력 뉴런에 라벨을 부여하기 위해서, 정확도를 측정하기 전 무작위로 사각파와 톱니파를 선택하여 이를 나타내는 라벨을 디코더에 전송하였다. 이후 인터페이스는 해당 라벨에 관한 실수 데이터를 스파이크로 인코딩하고, 동시에 디코더에서 1ms 동안 출력 뉴런의 발화를 측정하여 가장 많이 발화한 뉴런에 전달받은 라벨을 부여하였다. 정확도 측정 시에는 인터페이스가 데이터를 인코딩하기 전에 먼저 해당 데이터의 라벨을 디코더에 전송한다. 이후 디코더는 가장 많이 발화한 출력 뉴런에 부여된 라벨을 읽어 두 라벨이 일치할 경우 뉴로모픽 시스템이 올바른 분류를 수행한 것으로 판단한다.

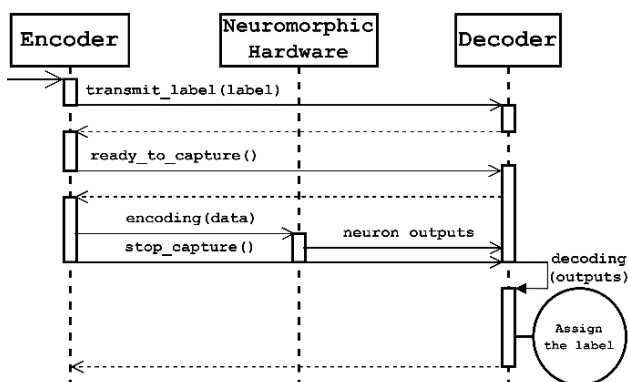


그림 2 - 뉴로모픽 하드웨어와 인터페이스간 Message Sequence Chart

3. 실험 결과

[그림 3]은 인터페이스에서 생성한 신호(노란색)를 받은 뉴런의 막 전위(청록색)와 발화(빨간색)를 나타낸 것이다. 인터페이스의 출력을 받은 뉴런의 막 전위가 증가하다 임계전위에 다다르면 발화하였다. 또한 발화와 동시에 뉴런의 막 전위는 휴지전위로 낮아진다.

[그림 4]는 전 뉴런과 후 뉴런 사이의 상호작용을 나타낸 것이다. 인터페이스의 신호(노란색)를 받은 전

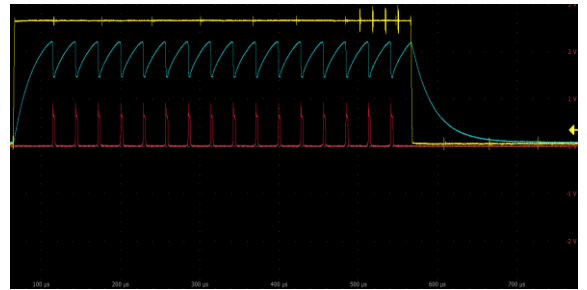


그림 3 - 인터페이스의 출력(노란색)과 이에 반응하는 뉴로모픽 칩의 뉴런 막 전위(청록색) 및 뉴런의 발화(빨간색)를 오실로스코프에서 확인한 모습

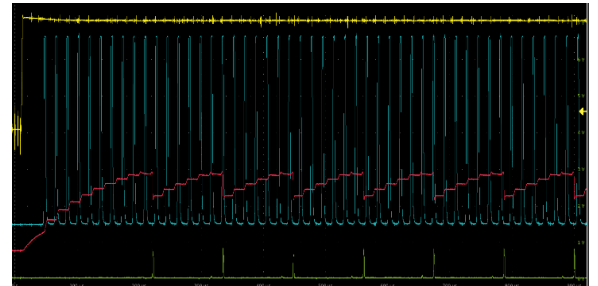


그림 4 - 인터페이스의 출력(노란색), 전 뉴런의 발화(청록색), 후 뉴런의 막 전위(빨간색), 후 뉴런의 발화(초록색) 간의 시간적 관계를 오실로스코프로 확인한 모습

뉴런이 발화(청록색)하면 후 뉴런의 막 전위가 증가한다. 또한 후 뉴런의 막 전위가 임계 전위에 다다르면 후 뉴런이 올바르게 발화하였다.

또한 동일한 인터페이스 장치를 사용해 뉴로모픽 시스템을 구축하여 100개의 RF대역의 사인파와 사각파를 분류하는 실험을 진행하였고, 96%의 정확도로 두 개의 파형을 분류하였다.

III. 결론

본 연구에서는 뉴로모픽 칩을 위한 뉴런 인코딩 및 디코딩을 수행하는 인터페이스를 구현하고, 인터페이스와 뉴로모픽 하드웨어 사이의 연동 실험을 진행하였다. 해당 실험에서 뉴로모픽 하드웨어와 본 연구에서 구축한 인터페이스가 연동되어 96%의 높은 정확도로 파형을 분류하였다. 이는 본 연구에서 구축한 인코딩 및 디코딩 인터페이스가 뉴로모픽 하드웨어와 성공적으로 연동되었음을 의미한다. 또한, 뉴로모픽 시스템의 개발에 본 논문에서 제시한 방식의 인터페이스가 적용될 수 있음을 확인하였다.

ACKNOWLEDGMENT

본 연구는 한국연구재단 신소자원과학기술개발사업 (grant no. NRF-2021M3F3A2A01037962)의 지원을 받아 수행된 연구입니다.

참 고 문 헌

- [1] Jingren, W. Jingjing, Y. Jingwei, W. Chunmao and P. Shiliang, "Deep Spiking Neural Network for High-Accuracy and Energy-Efficient Face Action Unit Recognition," 2021 International Joint Conference on Neural Networks (IJCNN), 2021, pp. 1-7, doi: 10.1109/IJCNN52387.2021.9533451.
- [2] M. Davies, et al., "Loihi: A neuromorphic manycore processor with on-chip learning," IEEE Micro, vol. 38, no. 1, pp. 82-99, 20