

극성 분할 Successive Subspace Learning 기반 이벤트 객체 분류 모델 Event-VoxelHop 연구

최재원 / 홍익대학교 Image Understanding Lab.

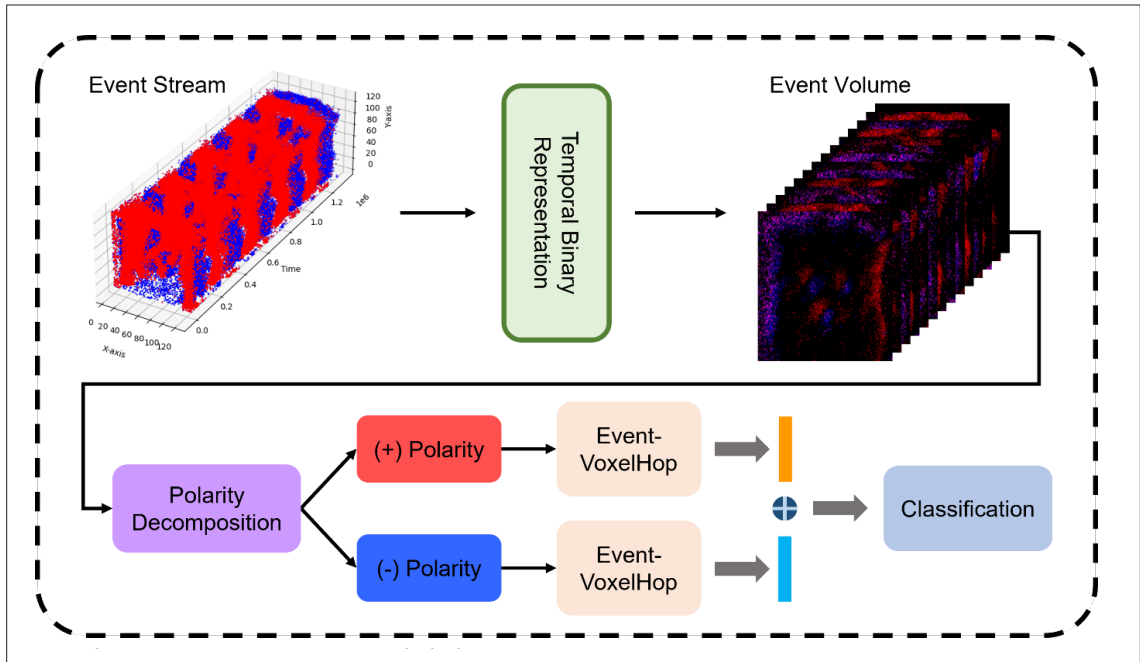
이벤트 카메라는 차세대 뉴로모픽 센서로 기존의 카메라와 비교하여 매우 빠른 속도와 저전력이라는 장점을 갖는다. 하지만 x, y, t, p 의 개별 이벤트 형태의 데이터 특성 때문에 일반적인 프레임 기반 네트워크에 직접적으로 적용하기 어렵다는 단점이 있다. 또한 이벤트 데이터를 프레임으로 변환하는 과정에서 적절한 방식을 사용하지 않으면 불필요한 정보를 과도하게 포함하거나, 필요한 정보를 과도하게 압축하여 데이터 품질이 떨어진다. 본 논문에서는 이 문제를 해결하고, 매우 낮은 연산량을 가지면서도 경쟁력 있는 이벤트 카메라 객체 분류 모델을 제안한다.

이벤트 스트림 데이터를 극성에 따라 Temporal Binary Representation(TBR) 방식을 적용하여 일정 시간 동안 세부 구간 내 이벤트 발생 여부를 이진화했고, 이를 십진수로 변환한 후 정규화했다. 이 과정을 통해 이벤트 스트림을 프레임으로 변환한 뒤 여러 장을 쌓아 3차원 이벤트 볼륨을 구성하였다. 이러한 볼륨은 공간과 시간 정보를 동시에 포함하여 기존의 방식보다 풍부한 정보를 가진다.

본 연구에서는 Successive Subspace Learning(SSL) 기법을 기반으로 한 경량화 모델 구조를 적용했다. SSL

은 back-propagation 기반 학습 대신 PCA 변형인 Saab Transform을 사용하여 단계적으로 특징을 추출하는 feedforward 방식으로, GPU 없이도 학습과 추론이 가능하다. 더 효과적인 특징 추출을 위해 이벤트 프레임의 극성 정보를 분리하여 독립적인 VoxelHop 유닛을 통과한다. 이후 Label-Assisted reGression(LAG) 유닛을 통해 특징을 보정한다. 여러 단계의 유닛별 출력을 합쳐 XGBoost 분류기로 최종 객체를 판별한다.

본 연구는 공개 데이터셋 CIFAR10-DVS와 N-MNIST를 이용해 알고리즘 성능을 검증하였다. 기존 CNN이나 SNN 기반 접근보다 안정적이고 효율적인 성능을 보였다. 특히 각각의 단일 극성만을 사용하여 객체 분류를 진행한 결과와 비교하였을 때, 극성 정보를 분리하여 처리한 후 결합하는 구조가 정확도 향상에 기여한다는 것을 증명했다. 이는 Event-VoxelHop 구조가 이벤트 데이터의 시공간적 특성을 효과적으로 학습할 수 있음을 의미한다. 또한 실험을 통해 이벤트 누적 시간(Δt)이 짧을수록 공간 정보가 부족해지고, 너무 길면 시간 정보가 손실된다는 점을 확인하였다. 80ms를 기준으로 설정했을 때 공간적 선명도와 시간적 변화를 모두 보존할 수 있었으며, 이 설정을



<그림 1> Event-VoxelHop 프레임워크

기반으로 한 이벤트 볼륨이 가장 안정적인 결과를 보였다.

연산 복잡도 측면에서 Event-VoxelHop은 매우 경량화 되어 있다. 약 0.12M의 파라미터와 1.22M FLOPs만으로도 높은 분류 성능을 보였으며, 전체 학습 과정은 20분 전후로 완료되었다.

본 연구는 이벤트 카메라 데이터의 시공간적 특성과 극성 정보를 효과적으로 처리할 수 있는 새로운 경량화 모델을 제시하였다. GPU 없이도 학습과 추론이 가능해, 저전력, 저지연 환경에서 실시간 인공지능 비전 시스템을 구현할 수 있는 기반 기술로서의 잠재력을 보여준다.



최재원

- 2023년 8월 : 홍익대학교 전자전기공학부 학사
- 2025년 8월 : 홍익대학교 전자전기공학부 석사
- 주관심분야 : 영상처리, 이벤트 카메라, 이상탐지