

이벤트 기반 물체 추적 시스템 개발

김상준^{a,b} 이현경^{a,b} 이승아^a 김대연^{b,†}^a연세대학교^b한국과학기술연구원

†daeyeon@kist.re.kr

Development of Event-based Object Tracking System

Sang-Jun, Kim^{a,b} Hyunkyung Lee^{a,b} Seung Ah Lee^a Dae-Yeon, Kim^{b,†}^aOptical Imaging Systems Laboratory, Department of Electrical and Electronic Engineering,
Yonsei University, Seoul 03722, Republic of Korea^bCenter for Opto-Electronic Materials and Devices, Korea Institute of Science and
Technology, Seoul 02792, Republic of Korea

요약

동적 비전 센서(Dynamic Vision Sensor)라고도 알려진 이벤트 카메라는 생체에서 영감을 받은 새로운 시각 센서이다. 고정된 속도로 이미지를 생성하는 기존 카메라와 달리 이벤트 기반 카메라의 픽셀은 독립적이고 비동기적으로 작동한다. 기존 프레임 기반 카메라보다 이벤트 기반 카메라가 움직임을 포착하는데 더 적합하며 모션 블러(Motion Blur)가 없고 시간 해상도가 높다는 이점을 통해 고속카메라로 활용할 수 있다. 본 논문은 이벤트 카메라의 높은 시간 해상도와 동적 범위, 낮은 지연시간, 전력 소비량의 이점을 활용하여 움직이는 물체를 모션 블러 없이 포착하는 이벤트 기반 물체 추적 시스템을 제안한다. 실험을 통해 전체 영상을 포착하는 기존 프레임 기반 카메라에 비해 밝기 변화에 따른 동적 변화만을 추적하는 이벤트 기반 카메라는 모션 블러가 없다는 점을 검증하였다.

1. 서론

뉴로모픽 카메라(Neuromorphic Camera) 또는 동적 비전 센서(Dynamic Vision Sensor)라고도 불리는 이벤트 카메라는 생체에서 영감을 받은 새로운 시각 센서로 기존 카메라와 달리 비동기 이벤트 형태의 강도 변화만 감지할 수 있다. 따라서 사람의 홍채가 정보를 받아들이는 것과 유사하며 프레임 없이 밝기 변화에 반응하여 '이벤트'가 발생한다. 이벤트가 발생하는 과정은 픽셀이 독립적이며 비동기적으로 작동하면서 밝기 변화가 임계값을 초과하면 픽셀은 x, y 좌표, 타임스탬프, 극성을 포함하는 이벤트를 방출한다. 이벤트 카메라는 기존 카메라보다 시간 해상도가 높고 프레임이 높은 빠른 움직임(초당 수천 프레임)의 비디오를 재구성할 수 있다. 동기적으로 작동하는 기존 프레임 기반 카메라에 비해 낮은 지연시간, 전력 소비량, 그리고 움직이는 물체에 대해 모션 블러가 거의 없다는 이점을 제공한다. 최근 이벤트 카메라를 이용한 국내 연구로는 라인 특징 추출에 의한 데이터 처리[1], 시간차 평면을 이용한 오픈티컬 플로우 추정[2], 공간적 스파이킹 어텐션 매커니즘을 통한 이벤트 시퀀스 인식[3] 등이 있다. 해외에서도 이벤트 카메라 기반의 연구[4-8]가 활발히 진행되고 있다.

본 논문에서는 실험을 통해 기존의 카메라에 비해 시간 해상도와 동적 범위가 높은 이벤트 카메라가 모션 블러가 없음을 검증하며, 이러한 이벤트 카메라의 특성을 활용한 모션 블러 없이 물체를 포착할 수 있는 이벤트 기반 물체 추적 시스템을 제안한다.

2. 본론

기존의 카메라 센서는 사진들을 연속으로 겹쳐 동영상을 만드는 프레임 형식인 반면에 이벤트 카메라 센서는 각각의 픽셀마다 같은 노출 시간을 갖고 전체 영상을 일정한 간격으로 기록하지 않고 장면의 변화만 기록하는 동적 시각 정보를 획득한다. 각각 픽셀들은 시간에 대한 빛의 세기에 로그를 취한 값을 기록하며 빛의 급격한 변화가 일어나 일정한 임계값을 넘어가게 된다면 이벤트가 발생하였다고 신호를 보낸다. 이벤트가 발생하는 신호를 받으면 이벤트 카메라 픽셀의 x, y 좌표와 타임스탬프, 극성으로 이벤트가 기록된다. 타임스탬프는 유닉스 타임스탬프 기반 마이크로초 단위로 기록되며, 날짜가 지정된 정보를 추적하고 정렬하는 컴퓨터 시스템에 유용하다. 극성은 픽셀이 임계값 기준으로 더 밝아

지거나 어두워지는 것을 나타내는 것이다. 본 논문에서 사용된 이벤트 카메라는 DAVIS346-COLOR로 마이크로초 단위의 높은 시간 해상도와 120dB의 높은 동적 범위를 제공하며, 346×260 픽셀로 이루어져 있다. 이벤트 카메라는 동시에 일반 프레임 이미지와 이벤트 기반 이미지를 보여줄 수 있고, 이벤트 기반 컬러 비디오로 재구성이 가능하다.

아래의 그림 1은 본 논문에서 설계한 실험 개요도이다. 이벤트 카메라를 통해 촬영하면 이벤트가 발생하고 ROS(Robot Operating System) 시스템을 이용하여 이벤트 기록, 추출 및 데이터화를 한다. 데이터화된 이벤트들은 알고리즘을 통해 재구성할 수 있으며, 이를 본 실험에 적용한다면 높은 시간 해상도와 동적 범위의 특성으로 모션 블러 없는 물체 추적 카메라로 활용할 수 있다. ROS 시스템은 로봇 운영 체제로 로봇 공학 애플리케이션을 위한 오픈 소스 소프트웨어 개발 키트로 실시간 움직임을 감지 및 추적하는 작업에 유용하다. 본 논문에서는 ROS 시스템을 이벤트 카메라에서 이벤트를 기록하는 용도로 사용하며, 이벤트를 기록하면 “.bag”라는 ROSBAG 파일 형식으로 데이터 세트가 생성된다. 생성된 ROSBAG 파일은 알고리즘 모델을 통해 이벤트 기반 이미지로 복원할 수 있다. 본 논문에서의 알고리즘은 마이크로초 시간 해상도를 가진 연속적인 이벤트 스트림을 이미지, 비디오로 변환하는 E2VID 네트워크를 이용한다. E2VID 모델은 순환 신경망을 사용하여 이벤트 데이터에서 비디오 재구성하는 알고리즘 모델로 안정적인 동작이 이루어지며 다양한 이벤트 수를 처리하는 완전한 컨볼루션 네트워크이다.

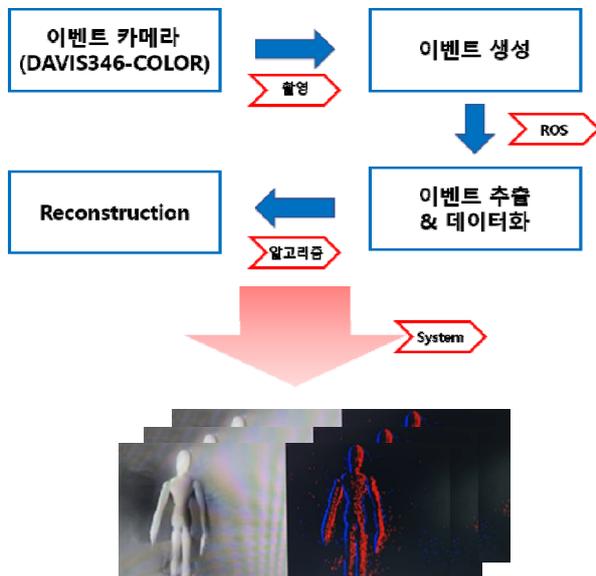


그림 1. 이벤트 기반 물체 시스템 개요

본 실험에서의 촬영 물체는 진자운동 하는 그네 모형을 사용하였다. 아래의 그림 2는 재구성되기 전 이벤트 카메라로 촬영한 원래의 이벤트 영상이다. 임계값 대비 극성에 따라 빨간색은 양수 이벤트, 파란색은 음수 이벤트로 지정된다. 그림 2의 이벤트 기반 이미지는 움직이는 물체에 대해서만 이벤트가 발생하며, 생성 및 추출되는 이벤트 수는 영상에서의 물체의 움직임에 영향을 받는다. 물체의 움직임이 빠르게 움직일수록 생성되는 이벤트의 수는 많아지며 추출되는 이벤트의 수 또한 많아진다. 이와 같이 프레임 속도 제한이 있는 기존 카메라에 비해 더 많은 물체의

움직임을 포착할 수 있기 때문에 초고속카메라로 활용할 수 있다. 아래의 그림 3의 (a)는 일반 카메라로 촬영한 영상이다. 15초의 영상을 촬영하였고, 이를 30프레임 기준으로 450개의 이미지가 생성되었다. 생성된 이미지들은 대부분 모션 블러가 심하게 나타나며 이미지 속의 물체를 정확히 인식하기 어렵다. 기존 카메라에서는 모션 블러로 인해 특정 환경에서의 선명한 물체 인식은 불가능하다.

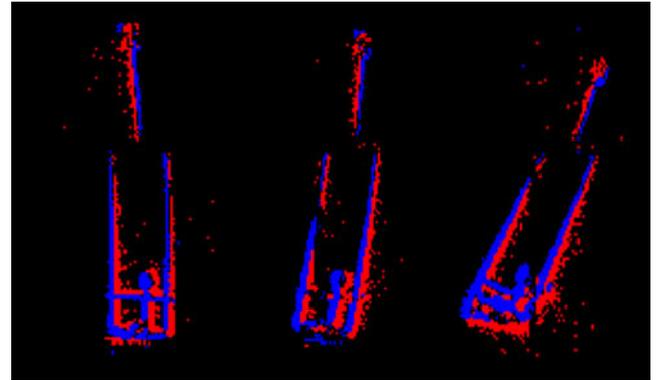


그림 2. 재구성되기 전 이벤트 카메라로 촬영한 영상

그림 3의 (b)는 그림 2의 기존 이벤트를 재구성하여 복원된 영상이다. 일반 RGB 이미지와 같이 15초 영상을 촬영하여 이벤트가 생성되었고, 35만 개의 이벤트가 추출되었다. 일반 RGB 카메라로 획득한 영상은 프레임 속도 한계(30프레임 기준 450개 이미지)로 인해 모션 블러 및 동적 범위 문제를 이벤트 기반 카메라에서는 해결할 수 있음을 검증하였다. 이벤트 기반 이미지는 기존 프레임 기반 카메라에서 나오는 이미지보다 더 효율적이고 짧은 대기 시간으로 처리할 수 있는 이벤트 스트림 생성 및 초고속 비전 기반 제어가 가능하다. 이는 모션 블러 없는 물체 추적 시스템을 개발할 수 있을 뿐 아니라 저지연, 저전력으로 실시간 움직임을 감지 및 추적하는 시스템으로 확장하여 개발할 수 있다.

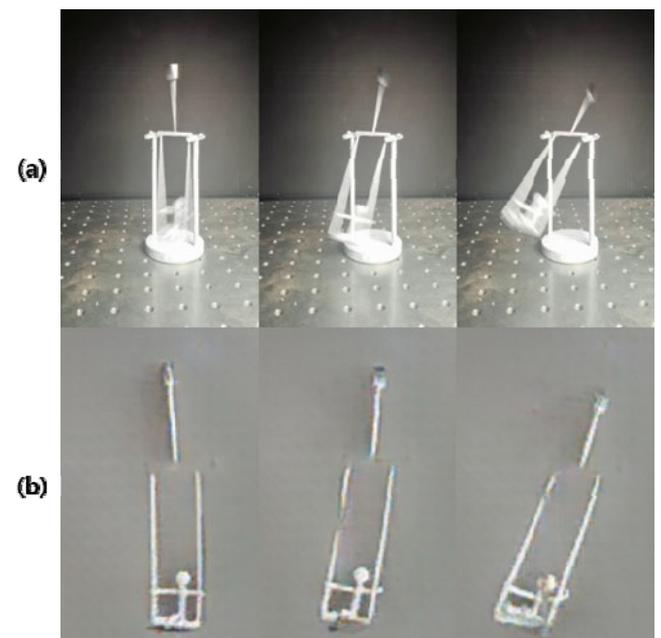


그림 3. (a) 일반 RGB 이미지 (b) 이벤트 기반 재구성된 이미지

3. 결론

이벤트 카메라의 특징은 짧은 대기 시간, 낮은 전력 소비, 높은 동적 범위 및 모션 블러가 없다는 점이며, 이를 활용하여 이벤트 기반 물체 추적 시스템에 적용할 수 있다는 것을 제안하였다. 특히 이벤트 카메라가 높은 동적 범위를 가지는 것은 이벤트 카메라의 픽셀이 절대 강도 대신 강도의 로그를 사용하여 밝기의 변화를 찾고 이벤트 발생을 결정하기 때문이다. 따라서 빠른 물체 인식 및 감지를 할 수 있으며 이는 이벤트 기반 물체 추적 시스템뿐만 아니라 자율주행 차량과 같이 매우 높은 동적 범위가 필요한 분야에 매우 적합하다. 또한 드론, 로봇에 접목하면 밝기의 변화가 심하거나 초고속 움직임을 포착하는데 필요한 환경에서의 물체 인식, 추적, 모니터링 등에 대한 활용이 비전 연구 기반 활발히 이루어질 것을 예상하며, 이러한 연구들은 여러 분야에 접목되어 다양하게 쓰일 것으로 기대한다.

감사의 글

이 논문은 2022년도 정부(경찰청)의 재원으로 지원받아 수행된 연구 결과임 [내역사업명: 확장현실(XR) 기반 복합테러 대응 교육·훈련 테스트 베드 구축 / 연구개발과제번호: PR08-04-000-21]

참고문헌

- [1] Y. Hwang, Soo Jeon Lee, J. Yang, U. Choi, D. Song, S. Paek, J.I. Park, "Data Processing and Positioning based on Line Feature Extraction of Event Camera," *The Korean Society for Aeronautical and Space Sciences*, pp.976-977, November 2021.
- [2] S.I. Lee, H.R. Kim, C.H. Kim, H.J. Kim, "Optical Flow Estimation for Event Cameras using Surface of Time Difference Correspondence" *Institute of Control, Robotics and Systems*, pp438-439, 2020.
- [3] Y.E. Kim, S.H. Lee, W.J. Lee, Y.S. Kim, H. Myung, "Event Sequence Recognition with Spatial Spiking Attention Mechanism," *Institute of Control, Robotics and Systems*, pp.462-463, 2020.
- [4] H. Rebecq, R. Ranftl, V. Koltun and D. Scaramuzza, "High Speed and High Dynamic Range Video with an Event Camera," *IEEE transactions on pattern analysis and machine intelligence*, vol. 43, no. 6, pp. 1964-1980, June 2019.
- [5] M. Gehrig, S. B. Shrestha, D. Mouritzen and D. Scaramuzza, "Event-Based Angular Velocity Regression with Spiking Networks," *2020 IEEE International Conference on Robotics and Automation (ICRA)*, Vurtual Conference, pp. 4195-4202, August 2020.
- [6] P. Duan, Z. W. Wang, X. Zhou, Y. Ma and B. Shi, "EventZoom: Learning to Denoise and Super Resolve Neuromorphic Events," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Virtual Conference*, pp. 12824-12833, June 2021.
- [7] C. Scheerlinck, H. Rebecq, T. Stoffregen, N. Barnes, R. Mahony and D. Scaramuzza, "CED: Color Event Camera Dataset," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, Seoul, South Korea, pp. 0-0, June 2019.
- [8] D. Gehrig, A. Loquercio, K. Derpanis and D. Scaramuzza, "End-to-End Learning of Representations for Asynchronous Event-Based Data," *Proceedings of the IEEE/CVF International Conference on Computer Vision*, Seoul, South Korea, pp. 5633-5643, November 2019.