

일반논문 (Regular Paper)

방송공학회논문지 제30권 제5호, 2025년 9월 (JBE Vol.30, No.5, September 2025)

<https://doi.org/10.5909/JBE.2025.30.5.795>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

이벤트 프레임 간 광류 추정을 활용한 인과적 RGB 영상 프레임 레이트 상향 변환 방법

김 영 육^{a)}, 주 대 연^{a)}, 박 장 수^{a)}, 박 주 현^{b)}, 정 해 룽^{b)}, 심 동 규^{a)†}

Causal RGB Video Frame Rate Up-Conversion Method Using Inter-Event Frame Optical Flow Estimation

Youngwook Kim^{a)}, DaeYeon Ju^{a)}, Jangsoo Park^{a)}, Juhyun Park^{b)}, Haeryong Chung^{b)}, and
Donggyu Sim^{a)†}

요 약

본 논문은 RGB 카메라와 이벤트 카메라를 결합한 RGB 영상 프레임 레이트 상향 변환 방법을 제안한다. 제안하는 방법은 이벤트 카메라로부터 취득된 이벤트를 다양한 필터를 통해 누적하여 이벤트 프레임을 생성하고, RGB 프레임으로부터 이벤트 프레임과 구조적 특징이 유사한 프레임을 생성하고, 생성된 프레임과 이벤트 프레임 간 호모그라피 추정으로 정합을 수행한다. 정합된 이벤트 프레임에 광류 추정 알고리즘을 기반으로 움직임 추정을 수행하며, RGB 영상에 대한 움직임 보상을 적용하여 프레임 레이트 상향 변환을 수행한다. 실험은 주행 환경에서 RGB 프레임과 이벤트가 취득된 DSEC 데이터셋을 사용하였으며, 제안 방법은 과거 시점의 RGB 프레임을 입력으로 하는 외삽 방법 대비 0.72dB 더 높은 PSNR 성능을 보였으며, 과거 시점의 RGB 프레임과 미래 시점의 RGB 프레임을 입력으로 하는 내삽 방법 0.28dB 더 낮은 PSNR 성능을 보였지만 인과성이 확보된다는 장점을 갖는다.

Abstract

This paper proposes an RGB video frame rate up-conversion method that combines RGB cameras and event cameras. The proposed method generates event frames by accumulating events acquired from the event camera through various filters, generates frames with similar structural features to event frames from RGB frames and performs registration between transformed RGB frames and event frames using Homography estimation. Motion estimation is performed on the registered event frames based on optical flow estimation algorithms, and motion compensation is applied to RGB video to achieve frame rate up-conversion. The experiments used the DSEC dataset, which contains RGB frames and events acquired in driving environments. The proposed method showed 0.72dB higher PSNR performance compared to extrapolation methods that use only past RGB frames, and 0.28dB lower PSNR performance compared to interpolation methods that use both past and future RGB frames. However, it has advantages of ensuring causality.

Keyword : Event-based Vision, Frame Rate Up-Conversion, Homography Estimation, Optical Flow Estimation,
RGB-Event Image Processing

Copyright © 2025 Korean Institute of Broadcast and Media Engineers. All rights reserved.

"This is an Open-Access article distributed under the terms of the Creative Commons BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and not altered."

I. 서 론

최근 RGB 카메라로부터 영상을 취득, 개선, 응용하는 기술이 발전한 것뿐만 아니라, 이벤트 카메라(Event camera), 라이다(LiDAR, Light Detection And Ranging), 관성 측정 장치(IMU, Inertial Measurement Unit) 등의 차세대 센서로부터 데이터를 취득, 개선, 응용하는 기술 또한 발전하였다^[1]. 특히, 각 센서 기술의 개별적 발전과 함께 센서 간 상호 보완적 특성을 활용한 다중 센서 기술에 대한 연구가 활발히 진행되고 있다. 이러한 다중 센서 기술은 단일 센서로는 탈성하기 어려운 성능 향상과 새로운 응용 분야의 창출을 가능하게 하고 있다. 자율주행, 스마트 팩토리, 드론 제어, 로봇 비전 등 다양한 분야에서 두 종류 이상의 센서를 결합하는 연구가 증가하고 있으며, 이러한 다중센서 기반 시스템은 복잡한 환경에서도 안정적이고 정밀한 인식 능력을 제공한다는 점에서, 최근 영상 처리 분야의 핵심 주제로 자리 잡고 있다. 특히 RGB 카메라와 이벤트 카메라를 결합하여 객체 탐지^{[2][3]}(Object detection), 영상 안정화(Stabilization), 깊이 추정(Depth estimation), 광류 추정(Optical Flow Estimation) 등의 작업을 고도화하는 연구가 증가하고 있다. RGB 카메라는 풍부한 색상 정보를 제공하지만 시간 해상도가 제한적이며, 반면 이벤트 카메라는 고속 동작이나 조도 변화에 우수한 반응성을 보이지만 출력되는 이벤트 데이터가 희소하여 풍부한 정보를 제공하기 어려운 한계가 존재한다. RGB 카메라의 낮은 시간 해상도와 이벤트 카메라의 제한적인 정보량을 보완하기 위해 두 카메라를

결합하는 방식에 대한 연구가 활발히 진행되고 있으며, 대표적으로 자율주행 시스템에서 두 카메라를 결합하여 저지연 처리와 높은 탐지 정확도를 동시에 달성하는 객체 탐지 방법들이 제안되고 있다^[4]. 그리고 두 카메라뿐만 아니라, 여러 DNN(Deep Neural Network) 모델을 이용하여 광류 추정(Optical Flow Estimation)의 작업을 고도화하는 방법들 또한 제안되고 있으며, 이러한 방법들은 매우 높은 성능을 달성하고 있다^{[5][14]}.

RGB 카메라를 사용하는 기존의 프레임 레이트 상향 변환(FRUC, Frame Rate Up-Conversion)은 프레임 내삽(Interpolation) 또는 외삽(Extrapolation)을 통해 임의 시점의 프레임을 생성한다. 그러나 이러한 방식은 RGB 카메라의 낮은 시간 해상도로 인하여, 고속 움직임이나 밝기 변화에 대한 대응이 어려우며, 광류 추정의 정확도가 저하되어 생성된 프레임의 품질이 떨어질 수 있다는 한계가 존재한다. 뿐만 아니라, 내삽 방법의 경우 미래 시점의 프레임을 요구한다는 한계가 있다. 그러나, 높은 시간 해상도를 갖는 이벤트 카메라를 RGB 카메라와 결합하여 활용한다면, 이러한 한계들을 극복할 수 있다.

이벤트 카메라는 프레임을 취득하는 RGB 카메라와는 달리 이벤트를 취득하는 카메라이다. RGB 카메라가 일정 주기로 프레임 전체를 취득하는 반면, 이벤트 카메라는 각 픽셀에서 밝기 변화가 감지될 때에만 비동기적으로 이벤트를 취득한다. 이러한 특성으로 이벤트 카메라는 1마이크로초 미만의 높은 시간 해상도, 140dB 이상의 넓은 동적 범위, 낮은 전력 소모라는 장점을 제공한다^[5]. RGB 카메라와 이벤트 카메라의 결합은 RGB 카메라의 풍부한 색상 정보와 이벤트 카메라의 높은 시간 해상도를 통합하여 우수한 화질을 갖는 영상을 생성할 수 있다는 잠재력이 있다. 그러나 이러한 결합의 실제 구현을 위하여 두 카메라의 서로 다른 데이터 취득 방식으로 인한 시간 동기화 문제, RGB 카메라의 상대적으로 느린 취득 속도로 인한 고속 움직임 발생 시 데이터 비교 문제, 희소한 이벤트와 공간적으로 밀집된 프레임 간의 비교 문제, 두 카메라의 서로 다른 노이즈 특성으로 인한 전처리 문제, 두 카메라의 렌즈 차이 및 물리적 위치 차이로 인한 정합 문제 등의 여러 문제를 해결해야 한다.

본 논문에서는 이벤트의 높은 시간 해상도를 활용하기

a) 광운대학교 컴퓨터공학과(Kwangwoon University)

b) 아르고(Argo)

‡ Corresponding Author : 심동규(Donggyu Sim)

E-mail: dgsim@kw.ac.kr

Tel: +82-2-941-6470

ORCID: <https://orcid.org/0000-0002-2794-9932>

* 이 논문은 2025년도 경기도의 재원으로 차세대융합기술연구원에서 관리하는 「경기도 테스트베드 활용 반도체 기술개발사업」(AICT-02-TB1)의 지원을 받아 수행된 연구임

* 이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원-대학ICT연구센터(ITRC)의 지원을 받아 수행된 연구임(IITP-2025-RS-2022-00156225)

* 이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원-학-석사연계ICT핵심인재양성의 지원(IITP-2025-RS-2022 - 00156215)을 받아 수행된 연구임

* 이 논문은 2024년도 교내학술연구비 지원을 받아 수행된 연구임

· Manuscript August 27, 2025; Revised September 12, 2025; Accepted September 12, 2025.

위하여 RGB 카메라와 이벤트 카메라를 결합하고, RGB 프레임의 시간 해상도를 향상하는 처리 방법을 제안한다. 제안하는 방법은 다음과 같은 주요 기여를 포함한다. 첫째, 이벤트로부터 다양한 필터 기반 이벤트 프레임을 생성하는 방법을 제안한다. 이는 희소한 이벤트 데이터를 공간적으로 일관된 프레임 형태로 생성하여 기존 영상 처리 알고리즘과의 호환성을 부여하는 단계이다. 둘째, RGB 프레임으로부터 이벤트 프레임과 구조적 특징이 유사한 RGB 이벤트 프레임을 생성하는 방법을 제안한다. 이는 RGB 프레임과 이벤트 프레임 간 효과적인 정합을 위한 단계이다. 셋째, 이벤트 프레임과 RGB 이벤트 프레임 간의 호모그래피(Homography) 기반 정합(Registration) 방법을 개발하여, 서로 다른 카메라 간의 공간적 일치성을 확보한다. 이는 두 카메라의 렌즈 차이나 물리적 위치 차이로 인한 공간적 불일치를 보정하는 단계이다. 넷째, 이벤트 프레임으로부터 광류를 추정하는 방법을 제안하며, 해당 방법은 TV-L1(Total Variance L1)^[6]을 기반으로 한다. 이는 RGB 프레임보다 높은 시간 해상도의 이벤트의 시간적 연속성을 활용하여 고속 움직임 정보를 추출하는 단계이다. 특히 제안하는 광류 추정 방식은 이벤트 데이터의 특성을 고려하여 밝기 불변 가정(Brightness constancy assumption)을 확장한 데이터 항을 포함하며, 이는 노이즈와 극성 반전에 강인한 특징을 갖는다. 마지막으로, 정합된 이벤트 프레임의 광류 정보를 이용하여 RGB 영상에 프레임 레이트 상향 변환을 적용하여, 높은 시간 해상도의 RGB 프레임을 생성한다. 특히 제안하는 프레임 레이트 상향 변환 방식은 미래 시점의 RGB 프레임 없이도 높은 시간 해상도의 RGB 영상 생성이 가능하다. 주행 환경에서 취득된 RGB 프레임과 이벤트로 구성된 DSEC 데이터셋을 사용한 실험에서, 제안 방법은 과거 시점의 RGB 프레임을 입력으로 하는 외삽 방법 대비 더 높은 성능을 보였다. 또한 과거 시점의 RGB 프레임과 미래 시점의 RGB 프레임을 모두 입력으로 하는 내삽 방법과 비교했을 때, 낮은 성능을 보였지만, 미래 시점의 RGB 프레임을 사용하지 않는다는 장점을 갖는다. 이는 자율주행, 드론 제어, 로봇 비전 저지연 응답이 필수적인 응용 분야에서 제안 방법의 우수성을 보여주며, 이벤트 카메라와 RGB 카메라의 결합이 차세대 영상 처리에서 중요한 역할을 할 수 있음을

실증적으로 증명한다.

본 논문의 구성은 다음과 같다. 2장에서는 이벤트 프레임 생성 방법, RGB 프레임과 이벤트 프레임 간 정합 방법, 이벤트 프레임 간 광류 추정 방법, RGB 영상 프레임 레이트 상향 변환 방법에 관한 관련 연구를 각각 소개한다. 3장에서는 제안하는 방법의 전체 구조와 각 단계별 알고리즘을 상세히 설명한다. 4장에서는 실험 결과를 통해 제안하는 방법의 성능을 분석한다. 마지막으로 5장에서는 결론 및 향후 연구 방향을 제시한다.

II. 관련 기술

1. 이벤트 취득 및 이벤트 프레임 생성

이벤트 카메라에서 취득되는 개별 이벤트는 감지된 시간을 의미하는 타임스탬프(t), 감지된 위치를 의미하는 수평 좌표(x) 및 수직 좌표(y), 감지된 밝기 변화 값을 의미하는 극성(p)으로 구성된다. 이벤트는 비동기적이고 희소한 특성이 있어, 데이터가 동기적이고 밀집한 프레임 데이터를 처리하는 기존 영상 처리 알고리즘에 적용하기 어렵다. 이러한 문제를 해결하기 위해 이벤트를 처리하는 별도의 데이터 처리 알고리즘을 적용하거나, 기존 알고리즘을 활용하기 위하여 이벤트를 프레임, 복셀(Voxel) 또는 포인트 클라우드(Point cloud) 형태로 생성하는 다양한 방법들이 제안되었다^[5]. 여기서 프레임 형태로 생성된 이벤트를 이벤트 프레임(Event frame)이라고 하며, 이벤트 프레임 생성 과정은 그림 1과 같은 구조를 갖는다.

이벤트 프레임 생성 과정은 시간 윈도우(Window) 설정, 이벤트 누적 과정으로 구성된다. 먼저 시간 윈도우 설정 과정에서는 이벤트 프레임을 생성할 시간 구간을 정의한다. 시간 구간은 고정 간격으로 또는 이벤트 발생량을 기준으로 설정한다. 그리고 이벤트 누적 과정에서는 정의된 시간 윈도우 내에서 각 픽셀 위치(Coordinate)에 발생한 이벤트들을 해당 픽셀에 누적한다. 누적 방식은 각 픽셀에서 발생한 이벤트의 개수를 합하여 계산하거나, 양(Positive)의 극성의 이벤트와 음(Negative)의 극성의 이벤트를 구분하여 합하여 계산할 수 있다. 이벤트 프레임은 이벤트를 프레임

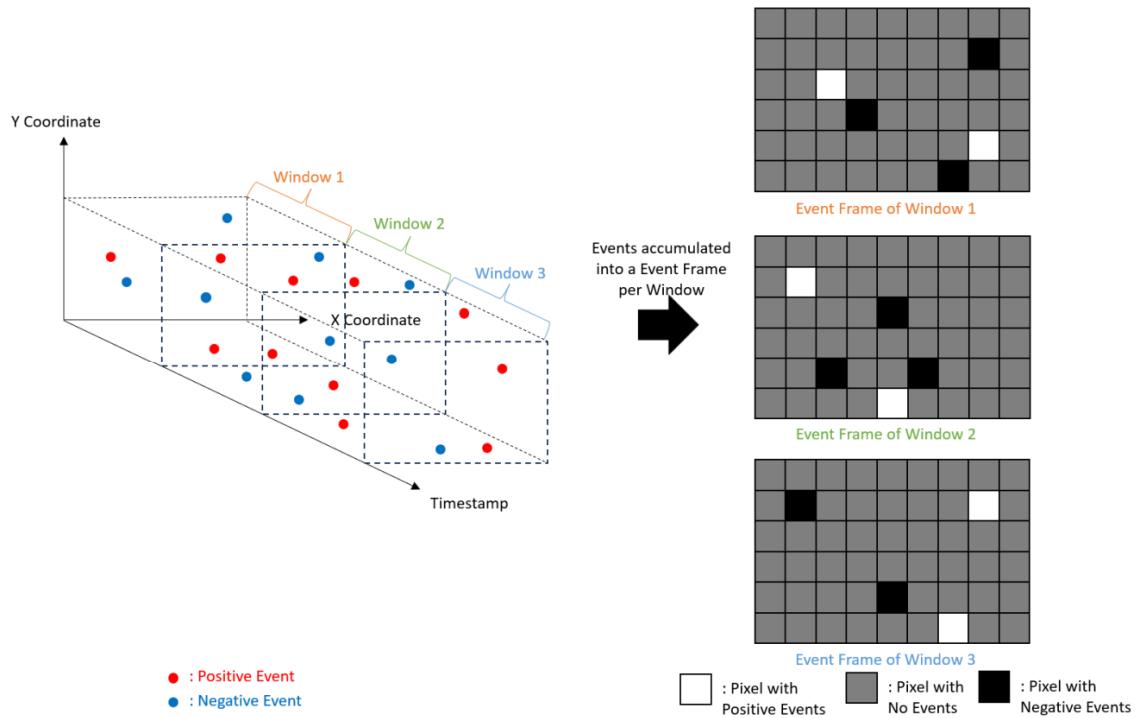


그림 1. 이벤트 프레임 생성 과정
Fig. 1. Event frame generation process

으로 생성하여 기존 컴퓨터 비전 알고리즘과의 높은 호환성을 제공한다는 점에서 큰 장점을 가진다. 그러나 타임 스탬프(Timestamp) 양자화로 인한 이벤트의 시간 해상도 손실, 프레임화로 인한 이벤트의 희소성 손실, 발생 이벤트 량에 따른 프레임 내 정보량의 불안정성 등의 한계가 존재한다. 따라서 이러한 한계를 극복하기 위한 이벤트 프레임을 효과적으로 누적하는 방법이나 이벤트 필터링 및 이벤트 프레임 필터링 방법에 대한 연구가 진행되고 있다.

2. RGB 프레임과 이벤트 프레임 간 정합

RGB 프레임과 이벤트 프레임은 물리적으로 서로 다른 카메라에서 생성되므로, 각 카메라의 초점거리(Focal length)와 각 픽셀의 물리적 크기 등의 물리적 특성 차이와, 카메라별 촬영 위치와 촬영 각도 등의 차이로 인해 두 데이터를 표현하는 좌표계 간 불일치가 발생한다. 따라서 RGB 카메라와 이벤트 카메라를 결합한 환경에서는 두 카메라로

부터 취득한 데이터 간의 정합이 필수적이다.

정합 문제를 해결하기 위해 다양한 방법들이 제안되어 왔으며, 명시적인 특징점(Keypoint) 사용 여부에 따라 특징 점 기반(Feature-based) 정합과 강도 기반(Intensity-based) 정합으로 분류하거나, 명시적인 기하 모델 여부에 따라 기하학적 변환 기반(Geometry-based) 정합과 비정형(Non-parametric) 정합으로 그 방법들을 분류할 수 있다^[7]. 명시적인 특징점의 사용 여부와 관련하여, 특징점 기반 정합은 RGB 프레임과 이벤트 프레임에서 각각 특징점을 추출하고 이를 매칭하여 정합을 수행하는 방법이다. 특징점 검출기로는 FAST(Features from Accelerated Segment Test)^[8], BRIEF(Binary Robust Independent Elementary Features)^[9] 등이 사용된다^[7]. 강도 기반 정합은 특징점 없이 프레임의 밝기나 픽셀 강도 분포를 비교하는 방법이다. 강도 기반 정합으로는 히스토그램 정합, 그래프 컷(Graph cut) 정합과, 암시적인 네트워크를 사용하는 합성곱 신경망(CNN, Convolutional Neural Network) 또는 트랜스포머(Transformer) 기반의 정합 등이 존재한다^[7].

명시적인 기하 모델 여부와 관련하여, 기하학적 변환 기반 정합은 두 카메라 간의 기하학적 관계를 호모그래피로 모델링하여 한 카메라에서의 좌표로부터 다른 카메라에서의 좌표로의 좌표 변환을 수행하는 방법이다^[7]. 비정형 정합은 기하 모델 없이 RGB 프레임과 이벤트 프레임 간의 복잡한 관계를 학습 또는 통계적 방식으로 추정하는 방법이다. 비정형 정합으로는 비모수적(Non-parametric) 정합^[10], 화률적 정합^[11] 등이 존재한다.

특징점 기반 정합과 기하학적 변환 기반 정합을 함께 사용할 경우, 특징점 매칭을 통해 명시적인 대응점을 확보하고, 기하 모델을 명시적인 파라미터를 통하여 추정할 수 있기 때문에 연산 효율성과 정합의 해석 가능성이 높아지는 장점이 있다. 그러나, 이벤트 프레임은 특징점 부족, 장면의 깊이 차이, 밝기 변화, 노이즈 등으로 인해 기하 모델의 정확도와 강인성이 저하되는 한계가 존재한다. 이러한 문제를 해결하기 위해 학습 기반 방법론^[12], 이벤트 최적화 특징점 추출 방법^[13] 등이 제안되었다.

3. 이벤트 프레임 간 광류 추정

이벤트를 이벤트 프레임으로 생성하여 RGB 프레임 간 광류 추정 방법을 이벤트 프레임 간 광류 추정에도 적용할 수 있다. 이벤트 프레임 간 광류 추정 방법은, 딥러닝을 사용하지 않는 최적화(Optimization) 기반 방법, 딥러닝을 사용하지만 라벨데이터가 필요 없는 자기지도학습(Self-Supervised learning) 기반 방법, 딥러닝을 사용하고 라벨 데이터가 필요한 지도학습(Supervised learning) 기반 방법이 존재한다^[14].

TV-L1^[6]은 광류 제약 조건(Optical flow constraint)에 전역 매끄러움 제약(Global smoothness constraint)을 추가한 RGB 프레임 간 광류 추정 방법이다. 이 방법은 L1 노름(L1 norm) 기반 데이터 항과 총 분산(Total variation) 정규화를 결합한 식 (1)과 같은 에너지 함수를 반복법(Iterative method)을 통하여 최소화한다. 여기서 움직임 발생 이후 프레임의 (x, y) 좌표의 점의 밝기값을 $I_{cur}(x, y)$ 로, 움직임 발생 이전 프레임의 (x, y) 좌표의 점의 밝기값을 $I_{ref}(x, y)$ 로 나타낸다. u_x 는 (x, y) 좌표의 수평 방향 움직임을, u_y 는

(x, y) 좌표의 수직 방향 움직임을 의미한다. 첫 번째 항은 L1 노름 기반 데이터 항(Data term)이며, 밝기 변화와 노이즈에 대한 강인성을 제공한다. 두 번째 항은 총 분산 정규화 항(Regularization term)이며, 전역적으로 매끄러운 광류를 유지하면서도 객체 경계에서의 불연속적인 움직임을 허용한다. λ 는 두 항 간의 균형을 조절하는 가중치이다. TV-L1은 정확도와 실행 시간 간의 트레이드 오프를 조절하기 위한 파라미터로, 각 반복별 에너지 함수의 수렴 여부 판정을 위한 에러의 임계값인 ϵ 을 사용하며, 반복 횟수가 너무 커지거나 에너지 함수 값이 ϵ 이하가 되면 반복을 중단한다. 즉, ϵ 이 감소할수록 정확도와 실행 시간은 모두 증가한다.

$$\int_{\Omega} |I_{cur}(x + u_x, y + u_y) - I_{ref}(x, y)| + \lambda(|\nabla u_x| + |\nabla u_y|) \quad (1)$$

그러나 이벤트 프레임은 밝기 변화가 없는 영역에서는 값을 가지지 않고, 이벤트의 극성에 따라 양수와 음수 값이 혼재되어 있고, 이벤트 카메라의 노이즈와 이벤트 프레임 생성 과정에서 발생하는 양자화 오차로 인한 아티팩트(Artifact)가 존재하는 등의 고유한 특성을 가지기 때문에 RGB 프레임 간 광류 추정 방법들을 이벤트 프레임 간 광류 추정에 적용하기에는 한계가 존재한다.

이러한 한계를 극복하기 위한 최적화 기반 방법들로는 광류 추정 알고리즘에 맞게 이벤트 프레임을 전처리하는 방법^[15]과 이벤트 프레임에 최적화된 광류 추정 알고리즘^[16]이 연구되고 있다. 전자는 Optical-flow-filter^[17]라는 프레임 기반 광류 추정 알고리즘을 그대로 사용하기 위하여 이벤트 프레임을 거리 변환을 통한 이벤트 프레임 생성 방법을 제안하였고^[15], 후자는 이벤트 프레임에서 에지를 따라 발생한 이벤트들을 정렬한 후, 프레임의 명암도에 대비 극대화(Contrast maximization)를 수행하는 프레임 기반 광류 추정 방법을 제안하였다^[16].

최근에는 딥러닝의 발전에 힘입어, 이러한 한계를 극복하기 위한 자기지도학습 기반의 방법들과 지도학습 기반의 방법들도 연구되고 있다. 자기지도학습 기반의 방법들로는 U-Net^[18] 등의 CNN 기반의 딥러닝 모델을 중추(Backbone)로 사용하여 비선형 궤적 사전 정보(Motion prior)와 대비

극대화를 결합하여 장기간 연속 시간 움직임을 추정하는 방법^[19], 순환 신경망(RNN, Recurrent Neural Network) 기반의 딥러닝 모델을 중추로 사용하여 반복적 이벤트 워핑(Iterative event warping)과 다중 시계열 대비 극대화(Multi-timescale contrast maximization) 손실을 도입하여 고주파수 움직임을 추정하는 방법^[20] 등이 제안되었다. 지도학습 기반의 방법들로는 스파이크 신경망(SNN, Spiking Neural Network) 기반의 딥러닝 모델을 중추로 사용하여 희소한 광류를 학습하여 밀집한 광류를 추정하는 방법^[21], 스파이크 신경망 기반의 딥러닝 모델과 트랜스포머 기반의 딥러닝 모델을 결합하여 광류를 추정하는 방법^[22] 등이 제안되었다.

4. RGB 영상 프레임 레이트 상향 변환

RGB 카메라와 이벤트 카메라를 결합한 환경에서는, 높은 시간 해상도를 갖는 이벤트를 통해 광류를 추정하고, 추정된 광류를 통하여 RGB 영상에 움직임 보상을 수행하여 RGB 영상의 시간 해상도를 향상하는 프레임 레이트 상향 변환이 수행 가능하다.

입력 데이터 유형과 관련하여, RGB 영상에 프레임 레이트 상향 변환을 수행하는 방법은 RGB 프레임 기반 보간(Frame-based interpolation), 다중 카메라 기반 보간(Multi-camera interpolation) 그리고 이벤트 기반 보간(Event-based interpolation)으로 세분화할 수 있다^[23]. 첫째, 프레임 기반 보간은 단일 RGB 카메라로부터 취득된 복수 개의 프레임을 입력으로 사용하여 목표 시각에 대응하는 프레임을 보간하는 방식이다. 둘째, 다중 카메라 기반 보간은 서로 다른 해상도, 프레임 레이트, 시점을 가진 복수 개의 RGB 카메라로부터 취득된 복수 개의 프레임을 융합하여 보간하는 방식이다^[24]. 셋째, 이벤트 기반 보간은 RGB 카메라와 이벤트 카메라 간의 영상 정합 후에 이벤트 카메라의 높은 시간 해상도를 활용하여 보다 정밀하게 RGB 프레임의 보간을 수행하는 방식이다.

광류를 모델링하는 방법과 관련하여, 프레임 기반 보간은 워핑 기반(Warping-based) 보간, 커널 기반(Kernel-based) 보간, 위상 기반(Phase-based) 보간으로 세분화할 수 있다^[23]. 워핑 기반 보간은 복수 개의 프레임 간의 광류를

명시적으로 추정한 후 이를 이용하여 프레임을 목표 시각에 대응하는 프레임으로 워핑한다. 커널 기반 방법은 광류를 명시적으로 추정하지 않고, 입력된 복수 개의 프레임으로부터 픽셀 단위의 지역적 커널을 예측하여 목표 시각에 대응하는 프레임을 합성한다. 위상 기반 방법은 복수 개의 프레임으로부터 목표 시간에 대응하는 프레임 생성을 위상 정보의 보간 문제로 접근하여, 광류를 명시적으로 추정하지 않고, 영상 신호를 주파수 도메인으로 변환한 후 목표 시각에서의 위상을 추정하여 재합성한다.

RGB 카메라와 이벤트 카메라를 결합한 환경에서 RGB 영상에 프레임 레이트 상향 변환을 수행하기 위하여 워핑 기반 보간을 사용할 경우, 광류가 높은 시간 해상도의 이벤트와 이벤트 프레임을 통해서 추정되어 있을 때, 과거 시점의 RGB 프레임으로부터 움직임을 보상하여 현재 시점의 RGB 프레임을 생성하는 방식으로 높은 시간 해상도의 RGB 영상을 생성할 수 있다. 그러나, 워핑 기반 보간은 광류를 추정하는 알고리즘의 성능의 영향을 받음으로, 이벤트만의 장점을 유지하면서 RGB 프레임과 높은 대응성을 유지하는 이벤트 프레임 생성법, 이벤트에 적합한 광류 추정 알고리즘, 이벤트와 RGB 영상의 정합 등이 모두 고려되어야 한다.

III. 제안 방법

본 절에서는 RGB 카메라와 이벤트 카메라를 결합하여 이벤트 프레임 간 광류 추정을 통해 RGB 영상의 프레임 레이트 상향 변환을 수행하는 방법을 제안한다. 제안하는 방법은 그림 2와 같은 구조를 가지며, 이벤트로부터 이벤트 프레임을 생성하는 첫 번째 단계, RGB 프레임으로부터 이벤트 프레임과 구조적 특징이 유사한 RGB 이벤트 프레임을 생성하는 두 번째 단계, RGB 이벤트 프레임과 이벤트 프레임 간 정합을 수행하는 세 번째 단계, 정합된 과거 및 현재 이벤트 프레임 간 광류를 추정하는 네 번째 단계, 정합된 이벤트 프레임 간 광류 추정 결과를 통해 과거 RGB 프레임으로부터 현재 RGB 프레임을 생성함으로써 프레임 레이트 상향 변환을 수행하는 마지막 단계 총 다섯 단계로 구성된다.

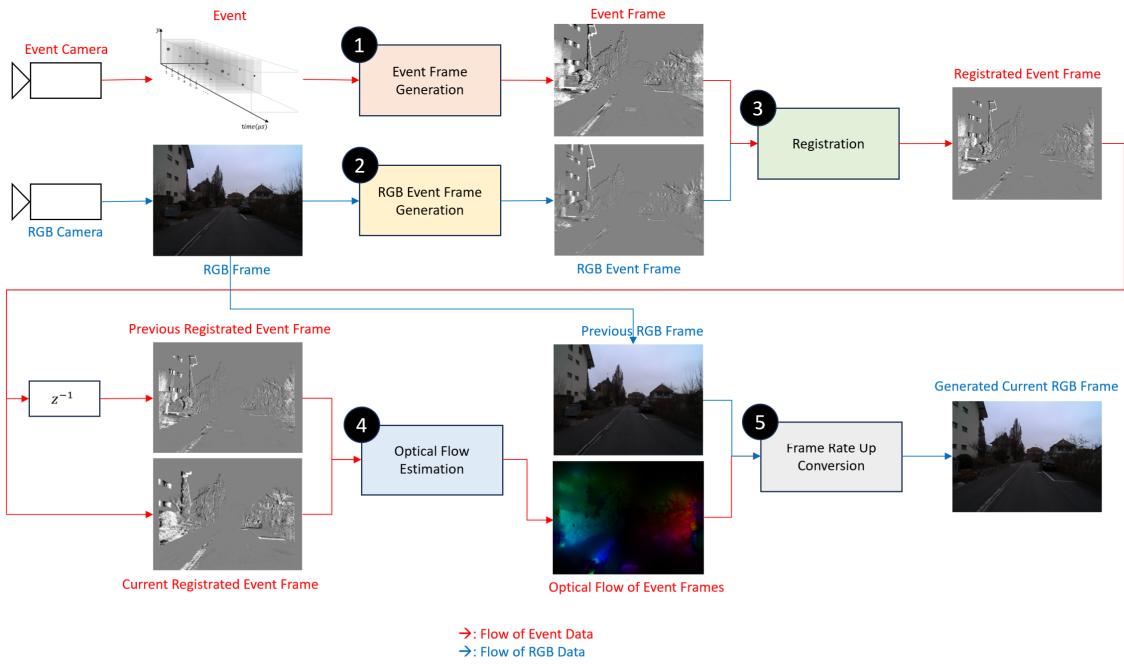


그림 2. 이벤트 프레임 간 광류 추정을 활용한 RGB 영상 프레임 레이트 상향 변환 방법

Fig. 2. RGB Video Frame Rate Up-Conversion Method Using Inter-Event Frame Optical Flow Estimation

1. 이벤트 프레임 생성

이벤트로부터 이벤트 프레임 생성 방법은, 시간 윈도우 설정 방법 및 이벤트 누적 방법으로 구성될 수 있다. 제안하는 방법에서는 이벤트 프레임과 RGB 프레임과의 시간 동기화를 위해 RGB 프레임의 취득 간격을 자연수로 나눈 시간 구간을 이벤트 프레임의 시간 윈도우로 설정한다. 제안하는 방법에서는 단순 누적 방식, 저주파수 성분 누적 방식, 고주파수 성분 누적 방식 총 세 가지 이벤트 누적 방법을 사용한다. 단순 누적 방식(Accumulation)은 현재 시간 윈도우 내 이벤트의 극성을 구분하되, 더 나중에 발생한 이벤트에 더 높은 가중치를 두어 누적하는 방식이다. 저주파수 성분 누적 방식(Low Pass Filtering)은 과거 시점의 이벤트 프레임과, 단순 누적 방식으로 생성된 현재 시점의 이벤트 프레임을 공간 축에 대하여 저주파수 통과 필터링한 결과물을 현재 시점의 이벤트 프레임으로 두는 방식이다. 고주파수 성분 누적 방식(High Pass Filtering)은 과거 시점의 이벤트 프레임과, 단순 누적 방식으로 생성된 현재 시점의 이벤트 프레임을, 공간 축에 대하여 고주파수 통과 필터링한 결

과물을 현재 시점의 이벤트 프레임으로 두는 방식이다. 저주파수 통과 필터링과 고주파수 통과 필터링은 연산 복잡도가 낮고 시간 복잡도가 적은, 각각 이동평균 필터(Moving average filter)와 차분 필터(Difference filter)를 사용하였다.

그림 3과 같이, 정합된 RGB 프레임과 이벤트 프레임에 움직이는 어두운 객체가 모두 존재할 때, 단순 누적 방식에서는 과거 시점의 프레임 대비, 현재 시점의 프레임에서 어두운 객체가 나타나는 부분에 대해서는 어두워지는 밝기 변화에 의한 음의 이벤트가 검은 픽셀로, 어두운 객체가 사라지는 부분에 대해서는 밝아지는 밝기 변화에 의한 양의 이벤트가 하얀 픽셀로 나타남을 확인할 수 있다. 따라서 객체에지를 따라 검은 픽셀과 하얀 픽셀들이 일정한 거리를 두고 쌍으로 생성되어 검은 객체와 하얀 객체, 총 2개의 객체가 나타난다. 저주파수 성분 누적 방식은 그 거리가 늘어나며, 두 객체 간 경계가 흐려진다. 고주파수 성분 누적 방식은 현재 시점의 프레임에서 객체가 존재하는 부분에 대해서는 검은 픽셀로, 객체의 이동 방향과 이동 반대 방향을 따라 하얀 픽셀들이 일정한 거리를 두고 쌍으로 생성되어, 하나의 검은 객체와 두 하얀 객체, 총 3개의 객체가 나타난다.

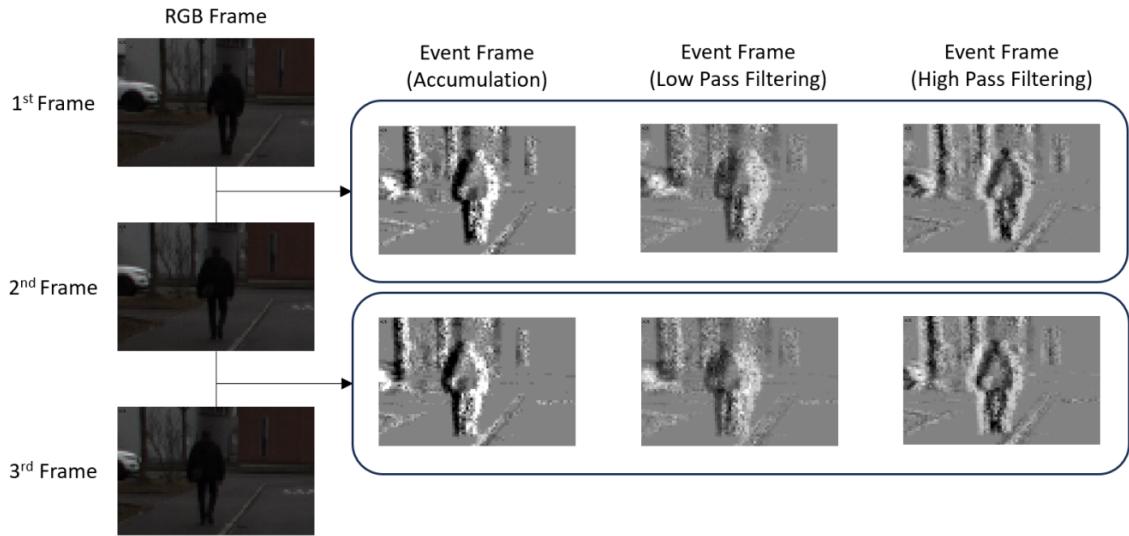


그림 3. 이벤트 누적 방법별 이벤트 프레임 비교 분석

Fig. 3. Event frame comparison analysis by event accumulation method

2. RGB 프레임으로부터 RGB 이벤트 프레임 생성

RGB 프레임과 이벤트 프레임 간 효과적인 정합을 위하여, RGB 프레임으로부터 이벤트 프레임과 구조적으로 유사한 형태의 프레임을 생성할 수 있다. 이 생성된 프레임을 RGB 이벤트 프레임(RGB event frame)이라고 한다. 해당 방법은 RGB 프레임으로부터 이벤트를 생성하는 연구^[25]를 기반으로 하나, 해당 연구가 Super-SloMo^[26]라는 딥러닝 모델을 통해서 이벤트를 생성하는 것과 달리, 제안하는 방법은 기존 연구를 개선하여 연산의 복잡도를 낮추고 이벤트 프레임을 생성한다. 제안하는 방법은 그림 4와 같다.

먼저 RGB 프레임을 식 (2)를 통해 회색 조로 변환한다. 다음으로 이벤트 카메라는 로그 스케일로 밝기 변화를 감지한다는 점에서, 이러한 특성을 RGB 프레임에도 모델링하기 위해 식 (3)을 통해 로그 스케일 밝기 변환을 적용한

다. 여기서 ε 는 로그 함수의 발산을 방지하는 작은 상수이다. 다음으로 시간적 변화를 모델링하기 위해 식 (4)를 통해 1차 무한 임펄스 응답(IIR, Infinite Impulse Response) 필터를 적용한다. 이는 저주파수 통과 필터링된 현재 시점의 RGB 이벤트 프레임과 저주파수 통과 필터링된 과거 시점의 RGB 이벤트 프레임을 픽셀 단위로 더하는 것으로 구현하였다. 여기서 α 는 시간 상수로, 이벤트 카메라의 응답 특성에 맞게 조정된다. 다음으로 각 픽셀에서의 밝기 변화량을 식 (5)를 통해 계산하여 이벤트 발생 여부를 결정한다. 이 때 발생한 이벤트를 RGB 이벤트라고 하며, RGB 이벤트는 식 (6)을 통해 임계값 θ_{pos} 와 θ_{neg} 를 이용하여 양극성 또는 음극성 값을 갖는다. 마지막으로 RGB 이벤트로부터 RGB 이벤트 프레임을 단순 누적 방식, 저주파수 통과 필터링 방식, 고주파수 통과 필터링 방식으로 생성한다.

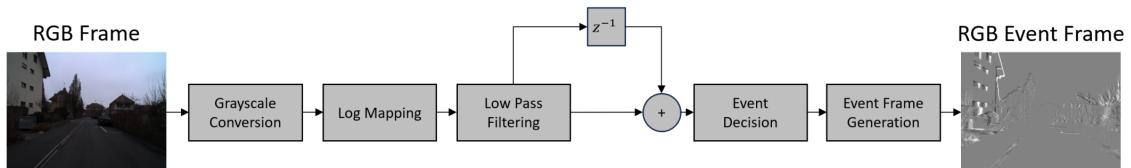


그림 4. RGB 이벤트 프레임 생성 과정

Fig. 4. RGB event frame generation process

$$I_{grav}(x, y, t) = 0.299 R(x, y, t) + 0.587 G(x, y, t) + 0.114 B(x, y, t) \quad (2)$$

$$L(x, y, t) = \log(I_{grar}(x, y, t) + \varepsilon) \quad (3)$$

$$L_{\text{filtered}}(x, y, t) = \alpha L(x, y, t) + (1 - \alpha)L_{\text{filtered}}(x, y, t - 1) \quad (4)$$

$$\Delta L(x, y, t) = L_{\text{filtered}}(x, y, t) - L_{\text{filtered}}(x, y, t - 1) \quad (5)$$

$$E_{RGB}(x, y, t) = \begin{cases} 1, & \text{if } \Delta L(x, y, t) > \theta_{pos} \\ -1, & \text{if } \Delta L(x, y, t) < -\theta_{neg} \end{cases} \quad (6)$$

3. RGB 이벤트 프레임과 이벤트 프레임 간 정합

서로 다른 원리로 동작하는 RGB 카메라와 이벤트 카메라의 공간적 정합을 위해, 제안하는 방법은 그림 5와 같은 절차를 따른다. 본 방법은 두 카메라의 내부 파라미터와 외부 파라미터는 주어져 있으나, 깊이 정보는 주어지지 않은 상황을 가정한다. 먼저 이벤트 프레임의 좌표를 RGB 프레임의 좌표계로 사상(Mapping)한다. 이후 RGB 좌표계로 투영된 이벤트 프레임과, RGB 프레임과 좌표계가 동일하고 구조적 특징이 이벤트 프레임과 유사한 RGB 이벤트 프레임의 특징점을 추출하여 매칭한다. 매칭된 특징점들을 이용해 두 프레임 간의 호모그래피를 추정하며, 이 과정에서 RANSAC을 이용하여 신뢰성 있는 호모그래피를 얻는다. 얻어진 호모그래피를 이용해 이벤트 프레임을 RGB 이벤트 프레임에 워핑(Warping)하여 공간 정합을 수행한다.

특징점 매칭의 성능을 향상하기 위해 이벤트 프레임의 좌표를 RGB 카메라의 좌표계로 사상한 후 특징점을 매칭

한다. 구체적으로, 우선 이벤트 카메라의 내부 파라미터를 이용하여 이벤트 프레임의 좌표를 정규화된 이벤트 프레임의 좌표계로 투영하고, 그다음으로 외부 파라미터를 이용하여 정규화된 RGB 프레임의 좌표계로 변환하고, 마지막으로 RGB 카메라의 내부 파라미터를 이용하여 RGB 프레임의 좌표계로 투영한다. 정규화된 이벤트 프레임의 좌표를 정규화된 RGB 프레임의 좌표계로 변환할 때에는 강체 변환(Rigid body transform) 행렬을 사용하였다.

좌표계 변환이 완료된 RGB 이벤트 프레임과 이벤트 프레임으로부터 특징점을 추출하기 위해 FAST 알고리즘을 사용한다. FAST는 픽셀 주변의 원형 패턴에서 연속된 픽셀들이 중심 픽셀보다 임계값 이상 밝거나 어두운 경우 코너로 판정하는 방법이다. 이벤트 프레임에서 신뢰성 있는 코너 검출을 위해, 중심 픽셀과의 강도 차이가 임계값을 초과하는 주변 픽셀의 개수가 상한선을 넘는 경우 노이즈로 판정하여 제외한다.

검출된 특징점에 대해 BRIEF 기술자를 적용하여 이진

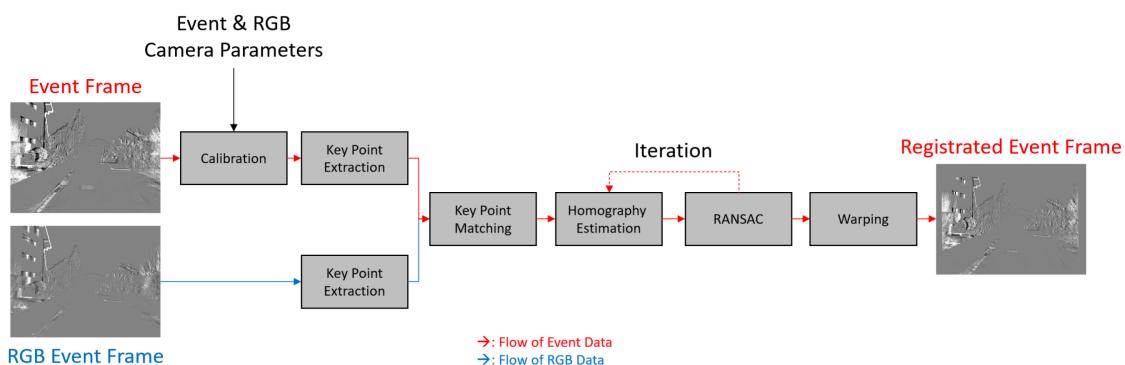


그림 5. RGB 이벤트 프레임과 이벤트 프레임 간 정합 과정

Fig. 5. Registration process between RGB event frame and event frame

기술자를 생성한다. BRIEF는 픽셀 쌍의 강도 비교를 통해 256비트 이진 기술자를 생성하며, 특징점 매칭은 해밍 거리를 이용하여 수행한다. 교차 검증을 통해 RGB 특징점 기준 매칭과 이벤트 특징점 기준 매칭이 서로 일치하는지 확인하여 잘못된 매칭을 제거한다.

매칭된 특징점 쌍으로부터 RGB 이벤트 프레임과 이벤트 프레임 간의 기하학적 변환을 추정한다. 이는 이벤트 프레임에서의 특징점과 RGB 이벤트 프레임에서의 특징점에 대해 매칭된 특징점 쌍으로부터 호모그래피를 추정하는 것이다. 호모그래피는 각 대응점 별 선형 방정식을 특잇값 분해(SVD, Singular Vector Decomposition)를 통하여 계산하였다.

이벤트 프레임에서는 노이즈와 이상치로 인해 정확한 호모그래피 추정이 어려우므로 RANSAC 알고리즘을 적용한다. RANSAC은 임의로 선택한 대응점들로부터 호모그래피를 추정하고, 재투사 오차를 기준으로 내부 점과 이상치를 분류하여 가장 많은 내부 점을 가진 호모그래피를 선택한다. 이 과정을 통해 이벤트 프레임과 RGB 이벤트 프레임 간의 정합이 가능해지며, 후속 광류 추정 결과를 RGB 프레임 레이트 상향 변환에 활용할 수 있다.

4. 이벤트 프레임 간 광류 추정

TV-L1^[6]은 광류 제약 조건(Optical flow constraint)에 전역 매끄러움 제약(Global smoothness constraint)을 추가한 RGB 프레임 간 광류 추정 방법이다. 이 방법은 L1 노름(L1 norm) 기반 데이터 항과 총 분산(Total variation) 정규화를 결합한 에너지 함수를 반복법(Iterative method)을 통하여 최소화한다. 여기서 움직임 발생 이후 프레임의 (x, y) 좌표의 점의 밝기값을 $I_{cur}(x, y)$ 로, 움직임 발생 이전 프레임의 (x, y) 좌표의 점의 밝기값을 $I_{ref}(x, y)$ 로 나타낸다. u_x 는 (x, y) 좌표의 수평 방향 움직임, u_y 는 (x, y) 좌표의 수직 방향 움직임을 의미한다. 첫 번째 항은 L1 노름 기반 데이터 항(Data term)이며, 밝기 변화와 노이즈에 대한 강인성을 제공한다. 두 번째 항은 총 분산 정규화 항(Regularization term)이며, 전역적으로 매끄러운 광류를 유지하면서도 객체 경계에서의 불연속적인 움직임을 허용한다. λ 는 두 항

간의 균형을 조절하는 가중치이다. TV-L1은 정확도와 실행 시간 간의 트레이드 오프를 조절하기 위한 파라미터로, 각 반복별 에너지 함수의 수렴 여부 판정을 위한 에러의 임계값인 ε 을 사용하며, 반복 횟수가 너무 커지거나 에너지 함수 값이 ε 이하가 되면 반복을 중단한다. 즉, ε 이 감소할 수록 정확도와 실행 시간은 모두 증가한다.

제안하는 방법은 기존 TV-L1 프레임 기반 광류 추정 알고리즘을 기반으로 하며, 해당 알고리즘에서 추정된 광류와 밝기 불변 조건과의 차이를 의미하는 척도인 데이터 항을 이벤트 프레임의 고유한 특성을 고려하여 세 가지 가능한 밝기 변화 가정 중 오차가 최소인 것을 선택하는 방식으로 수정하였다. 이벤트 프레임은 RGB 프레임과 달리 밝기 불변 가정을 직접적으로 적용하기 어렵다. 이벤트 발생 여부 및 발생한 이벤트의 극성은 객체의 움직임 정보와 무관하게 발생할 수 있는 특성을 반영하여, 식 (7)과 같이 과거 시점의 프레임 대비 각 픽셀에서 밝기값이 유지되는 불변 가정 $E_1(x, y)$, 밝기값이 아무런 이벤트가 발생하지 않았거나, 양의 이벤트와 음의 이벤트가 발생하였지만 상쇄되었음을 의미하는 중앙값인 중앙값 가정 $E_2(x, y)$, 밝기값이 중앙값을 중심으로 반전되는 반전 가정 $E_3(x, y)$ 총 세 가지의 밝기 변화 가정을 고려한다. 이때, 식 (8)과 같이 각 픽셀에서 세 가정의 오차를 계산하여 최소 오차를 갖는 가정을 선택한다. 이렇게 수정된 데이터 항과 기존 TV-L1의 정규화 항을 결합하여 식 (9)와 같이 에너지 함수를 구성한다. 여기서 움직임 발생 이후 프레임의 (x, y) 좌표의 점의 밝기값을 $I_{cur}(x, y)$ 로, 움직임 발생 이전 프레임의 (x, y) 좌표의 점의 밝기값을 $I_{ref}(x, y)$ 로 나타낸다. u_x 는 (x, y) 좌표의 수평 방향 움직임을, u_y 는 (x, y) 좌표의 수직 방향 움직임을 의미한다. λ 는 데이터 항과 정규화 항 간의 균형을 조절하는 가중치이며, 이 에너지 함수를 최소화하기 위해 에러의 임계값인 ε 을 사용하여 반복법을 통하여 최적화를 수행한다는 점은 기존 TV-L1 알고리즘과 동일하다.

$$\begin{aligned} E_1(x, y) &= |I_{cur}(x + u_x, y + u_y) - I_{ref}(x, y)| \\ E_2(x, y) &= |I_{cur}(x + u_x, y + u_y) - 0.5| \\ E_3(x, y) &= |I_{cur}(x + u_x, y + u_y) - (1 - I_{ref}(x, y))| \end{aligned} \quad (7)$$

$$E_{data}(x, y) = \min(E_1(x, y), E_2(x, y), E_3(x, y)) \quad (8)$$

$$E(x, y) = \int_{\Omega} E_{data}(x, y) + \lambda(|\nabla u_x| + |\nabla u_y|) \quad (9)$$

5. RGB 영상 프레임 레이트 상향 변환

정합된 과거 시점의 이벤트 프레임과 정합된 현재 시점의 이벤트 프레임 간 광류가 추정된 상태에서, 제안하는 방법은 과거 시점의 RGB 프레임으로부터 현재 시점의 RGB 프레임을 생성한다. 이를 위해, 먼저 정합된 과거 시점의 이벤트 프레임과 현재 시점의 이벤트 프레임 간 광류를 이용하여 움직임을 추정한다. 이후, 추정된 움직임을 과거 시점의 RGB 프레임에 적용하여, 움직임을 보상함으로써 현재 시점의 RGB 프레임을 생성한다. 이와 같은 움직임 보상

기반 프레임 생성은 광류 정보만으로 미래 시점의 RGB 프레임 없이 RGB 영상의 시간 해상도를 향상할 수 있다.

IV. 실험 결과

본 논문에서 제안하는 방법 및 각각의 세부 기능을 평가하기 위하여, 이벤트 카메라가 이벤트를 1마이크로초 미만의 시간 해상도로 비동기적으로 취득하며 이벤트 프레임을 20 FPS(Frame Per Second)의 주기로 생성하고, RGB 카메라는 RGB 프레임을 10 FPS로 취득하며 프레임 레이트 상향 변환의 목표 프레임 레이트가 20 FPS임을 대상으로 하였다. 그리고 두 카메라 각각의 내부 파라미터와 카메라 간 외부 파라미터가 존재함을 전제로 하였다. 실험 데이터셋은 DSEC(a Stereo Event Camera Dataset for driving sce-

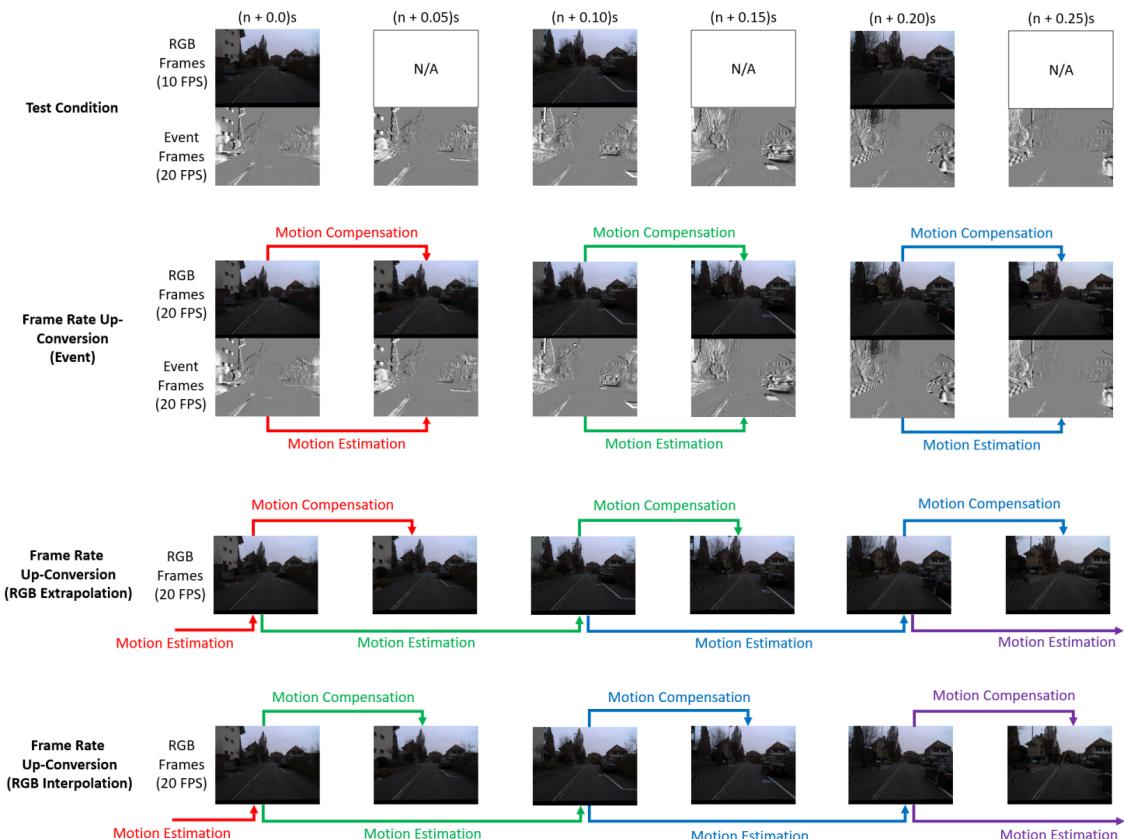


그림 6. 프레임 레이트 상향 변환 방법 평가를 위한 실험 조건
 Fig. 6. Test condition for frame rate up-conversion method

narios)를 사용하였다^[27]. 해당 데이터셋은 주행 환경에서 RGB 카메라와 이벤트 카메라를 결합하는 환경에서 취득된 8비트 심도, 1440x1080 해상도, 20 FPS의 RGB 프레임과 640x480 해상도, 1000 EPS(Event Per Second)의 이벤트 데이터가 쌍으로 구성된다. 해당 데이터셋은 추가적인 정답 데이터로 정합된 RGB 프레임을 20 FPS로, 광류 추정 성능 비교를 위한 정답 광류를 10 FPS로 제공한다. 따라서, 정답 RGB 프레임과의 프레임 레이트 상향 변환 성능 비교를 수행할 때에는 그림 6과 같이 입력 RGB 프레임의 프레임 레이트를 절반으로 서브 샘플링 후, 프레임 레이트를 다시 2 배로 복원하여 수행하였다. 그리고 정답 광류와의 광류 추정 성능 비교를 수행할 때에는 20 FPS가 아닌, 10 FPS로 취득되는 이벤트 프레임 간 광류를 추정하여 수행하였다.

실험은 AMD RYZEN5 5600 프로세서, 32G RAM, Windows 11 운영체제에서 가상 운영체제인 WSL Ubuntu 22.04에서 진행하였다. 제안하는 방법은 C언어로 구현되었으며, GCC 11.4로 컴파일이 가능하다. 또한, TV-L1 알고리즘은 OpenMP를 통한 CPU 상에서의 스레드 병렬화가 구현되었다. 또한, 이벤트로부터 복셀 그리드로 생성한 후 U-Net 딥러닝 모델을 사용하여 광류 추정을 하는 방법인 MotionPriorCM^[19]를 사용하는 경우, 학습은 Python 3.10, Pytorch-CUDA 12.1 환경의 RTX 3090 GPU에서 DSEC에서 정답 광류가 존재하는 모든 17개의 시퀀스를 9개의 시퀀스(thun_00_a, zurich_city_01_a ~ zurich_city_05_b)와 8개의 시퀀스(zurich_city_06_a ~ zurich_city_11_c)의 두 시퀀스 세트로 나눈 뒤, 하나의 시퀀스 세트를 배치를 1로, 에포크를 10으로 하여 학습한 뒤, 다른 시퀀스 세트에 대한 추론을 CPU 환경에서 진행하였다.

1. 이벤트 프레임 생성

제안하는 방법에서 이벤트 프레임은 정합을 위한 특징점 추출 및 특징점 매칭, 광류 추정의 대상이 되며, 프레임 레이트 상향 변환 성능이 고도화되도록 생성되는 것을 목적으로 한다. 본 절에서는 이벤트 프레임 생성 방법에 따른 RGB 프레임 레이트 상향 변환 성능을 비교한다.

표 1은 DSEC의 시퀀스 중 thun_00_a라는 시퀀스에 대해 8비트 심도를 갖는 이벤트 프레임 생성 시 각 이벤트가 꽉

표 1. 이벤트 프레임 생성 방법별 프레임 레이트 상향 변환 성능 비교 결과
Table 1. Frame rate up-conversion performance comparison results by event frame generation method

	Method	FRUC Result (PSNR (dB))
weight=1	Accumulation	28.06
	Low Pass Filtering	27.84
	High Pass Filtering	27.62
weight=2	Accumulation	28.07
	Low Pass Filtering	27.29
	High Pass Filtering	27.82
weight=4	Accumulation	28.22
	Low Pass Filtering	27.99
	High Pass Filtering	27.91
weight=8	Accumulation	28.44
	Low Pass Filtering	28.20
	High Pass Filtering	28.13
weight=16	Accumulation	28.61
	Low Pass Filtering	28.36
	High Pass Filtering	28.42
weight=32	Accumulation	28.39
	Low Pass Filtering	28.54
	High Pass Filtering	28.29
weight=64	Accumulation	28.08
	Low Pass Filtering	28.39
	High Pass Filtering	27.84
weight=128	Accumulation	27.61
	Low Pass Filtering	28.08
	High Pass Filtering	27.15
Overall (Mean/Standard Deviation)	Accumulation	28.14/0.293
	Low Pass Filtering	28.08/0.369
	High Pass Filtering	27.89/0.375

셀 값에 기여하는 가중치를 총 여덟 가지로 설정하였을 때, 단순 누적 방식(Accumulation), 저주파수 성분 누적 방식(Low Pass Filtering), 고주파수 성분 누적 방식(High Pass Filtering) 총 세 가지 이벤트 누적 방법별로 프레임 레이트 상향 변환 성능을 PSNR(Peak Signal-to-Noise Ratio)로 나타낸 표이다. 여기서 PSNR은 원본 영상과 복원된 영상 간의 화질 차이를 정량적으로 평가하는 지표로, 영상에서 가능한 최대 픽셀 값(MAX)의 제곱을 평균 제곱 오차(MSE, Mean Square Error)로 나눈 비율을 데시벨 단위로 표현하며, 값이 클수록 복원된 영상의 품질이 우수함을 의미한다^[7]. 입력 RGB 프레임은 정합된 RGB 프레임을 사용하였으며, 광류 추정 알고리즘은 TV-L1 알고리즘의 각 반복별 에너지 합수의 수렴 여부 판정을 위한 에러의 임계값 ϵ 을 0.1로 사용하였다.

실험 결과 분석을 통해 다음과 같은 특징을 확인할 수

있다. 가중치가 적절한 수준인 16 또는 32일 때 높은 프레임 레이트 상향 변환 성능을 보였다. 단순 누적 방식의 경우 가중치가 16일 때 28.61dB로 최고 성능을 보였으며, 저주파수 성분 누적 방식의 경우 가중치가 32일 때 28.54dB, 고주파수 성분 누적 방식의 경우 가중치가 16일 때 28.42dB의 높은 성능을 보였다. 단순 누적 방식의 PSNR 표준편차 0.293은 저주파수 성분 누적 방식의 PSNR 표준편차 0.369나 고주파수 성분 누적 방식의 PSNR 표준편차 0.375보다도 낮아, 가장 안정적인 성능을 보였다. 이러한 결과는 이벤트 프레임 생성에서 이벤트의 가중치와 필터링 방법이 프레임 레이트 상향 변환 성능 고도화 및 안정화에 핵심적인 역할을 함을 시사한다.

2. RGB 프레임과 이벤트 프레임 간 정합

제안하는 방법에서 정합은 이벤트 프레임 간 추정된 광류가 존재하고 과거 시점의 RGB 프레임으로부터 움직임을 보상할 때, 이벤트 프레임에서 픽셀별 움직임과 RGB 프레임에서 픽셀별 움직임이 동등해지는 것을 목표로 한다. 본 절에서는 정합 방법별 프레임 레이트 상향 변환 성능과 정합 성능을 비교한다.

표 2는 DSEC의 시퀀스 중 thun_00_a라는 시퀀스에 대해 이벤트를 RGB 좌표계로 정합 수행 후, 프레임 레이트 상향 변환 성능을 정답 RGB 프레임 대비 프레임 레이트 상향 변환이 수행된 RGB 프레임과의 PSNR로 나타내고, 정합 성능을 정합된 정답 RGB 프레임 대비 정합이 수행된 RGB 프레임과의 PSNR, 정규화된 상관관계(NCC, Normalized Cross Correlation)로 나타낸 표이다. 여기서 NCC는 두 영

상 간의 유사도를 측정하는 지표로, 각 영상의 평균을 빼고 표준편차로 나눈 값으로 정규화하여 계산되며 -1과 1 사이의 값으로 나타내어 1에 가까울수록 두 영상이 유사하고 0에 가까울수록 상관관계가 없음을 의미한다^[28]. RANSAC의 이상치 임계값이 낮으면 노이즈 판정 상한선이 낮아 특징점이 적게 검출됨을, 임계값이 높으면 상한선이 높아 특징점이 많이 검출됨을 의미한다. 또한, 호모그래피 추정 주기가 0이면 호모그래피를 미추정함을, 주기가 2/4/8이면 각각 이벤트 프레임 기준 2/4/8 프레임 간격으로 호모그래피를 추정함을 의미한다. 이벤트 프레임 생성 방법은 가중치가 16인 단순 누적 방식으로 하였으며, RGB 프레임으로부터 RGB 이벤트 프레임 생성 방법 또한 이벤트 생성 방법과 동일한 가중치가 16인 단순 누적 방식으로 하였다. 광류 추정 알고리즘은 TV-L1 알고리즘의 각 반복별 에너지 함수의 수렴 여부 판정을 위한 에러의 임계값 ε 을 0.1로 사용하였다.

실험 결과 분석을 통해 다음과 같은 특징을 확인할 수 있다. 호모그래피 추정 주기가 4로 적절한 수준일 때 높은 프레임 레이트 상향 변환 성능을 보였다. RANSAC의 이상치 임계값이 낮은 경우 26.81dB, 높은 경우 26.80dB로 거의 동일한 높은 성능을 보였다. 이는 프레임 레이트 상향 변환 성능은 광류 추정 기반 움직임 보상의 정확도에 좌우되므로, 호모그래피 추정 빈도가 너무 낮으면 카메라 움직임이나 장면 변화로 인한 호모그래피의 부정확성이 증가하여 성능이 저하되고, 반대로 추정 빈도가 너무 높으면 추정 오차가 빈번하게 발생하여 오히려 호모그래피의 불안정성이 증가하여 성능이 저하되기 때문이다. 정합 성능은 호모그래피 추정을 수행하지 않을 때 높은 성능을 보였다. 호모

표 2. 정합 방법별 프레임 레이트 상향 변환 성능 및 정합 성능 비교 결과

Table 2. Frame rate up-conversion and registration performance comparison results by registration method

Method	FRUC Result			Registration Result		
	PSNR (dB)	PSNR (dB)	NCC	PSNR (dB)	PSNR (dB)	NCC
RANSAC outlier threshold=low	Homography estimation period=2	26.75	19.04	0.9155		
	Homography estimation period=4	26.81	18.36	0.9010		
	Homography estimation period=8	26.69	18.68	0.9101		
	Homography estimation period=0	26.42	19.44	0.9276		
RANSAC outlier threshold=high	Homography estimation period=2	26.56	19.00	0.9124		
	Homography estimation period=4	26.80	19.42	0.9228		
	Homography estimation period=8	26.66	18.55	0.9077		
	Homography estimation period=0	26.42	19.44	0.9276		

그래피 추정을 수행하지 않을 때 정합 PSNR이 19.44dB로 최고 성능을 보였으며, NCC 역시 0.9276으로 최고 성능을 보였다. 이는 정합 성능이 특정 시점의 이벤트 및 RGB 프레임 간 변환의 정확도에 민감하게 변할 수 있는 경우, 호모그래피를 추정하지 않으면 정합의 일부 과정을 생략하지만 정합 안정성을 확보하여 지속적으로 높은 정합 성능을 유지할 수 있기 때문이다. RANSAC의 이상치 임계값 별 성능 차이는 미미하여, 노이즈 판정 상한선이 프레임 레이트 상향 변환 성능에는 결정적 영향을 미치지 않았다.

3. 이벤트 프레임 간 광류 추정

제안하는 방법에서 광류 추정은 두 이벤트 프레임 간 수행되며, 정합된 RGB 프레임에 움직임을 보상할 수 있는 광류를 생성하는 것을 목표로 한다. 본 절에서는 광류 추정 방법별 프레임 레이트 상향 변환 성능과 광류 추정 성능을 비교한다.

표 3은 DSEC의 시퀀스 중 thun_00_a 시퀀스에 대해 프레임 레이트 상향 변환 성능을 정답 RGB 프레임 대비 프레임 레이트 상향 변환이 수행된 RGB 프레임과의 PSNR로 나타내고, 광류 추정 성능을 정답 광류 대비 이벤트 프레임 간 추정된 광류와의 RAEE(Root Average Endpoint Error)로 나타낸 표이다. 그리고, 이벤트 프레임 간 광류 추정과 RGB 프레임 내삽 광류 추정 두 가지 방법에 대해 수정된 TV-L1과 기존 TV-L1 알고리즘의 성능을 비교하였다. 뿐만 아니라, 이벤트로부터 복셀 그리드를 생성한 후 U-Net 딥러닝 모델을 사용하여 광류 추정을 하는 방법인 MotionPriorCM^[19]과의 성능 비교를 수행하였다. RAEE는 광류 추정 성능을 평가하는 지표로, 추정된 광류 벡터와 정답 광류 벡터의 끝점 간 유clidean 거리의 평균제곱 오차(MSE)에 제곱근을 취한

값이다. RAEE 값이 작을수록 광류 추정 성능이 우수함을 의미한다^[29]. 실행 시간은 시퀀스 전체에서 광류 추정에 소모된 평균 시간을 초 단위로 나타낸 것이다. 인과성(Causality)은 미래 시점의 프레임을 사용하지 않는지를 나타낸 것이다. 정합된 RGB 프레임 간 광류를 추정하는 경우는 10 FPS로 취득되는 과거 시점의 RGB 프레임과 미래 시점의 RGB 프레임 두 프레임 간 광류를 추정하고 광류를 절반으로 나눈 뒤, 과거 시점의 RGB 프레임에 움직임을 보상하여 현재 시점의 RGB 프레임을 생성하는 내삽으로 프레임 레이트 상향 변환을 수행하였다. 이벤트 프레임 생성 방법은 가중치가 16인 단순 누적 방식으로 하였으며, 입력 RGB 프레임은 정합된 RGB 프레임을 사용하였다. 광류 추정 알고리즘은 TV-L1 알고리즘의 각 반복별 에너지 합수의 수렴 여부 판정을 위한 에러의 임계값 ε 을 0.1로 사용하였다.

실험 결과 분석을 통해 다음과 같은 특징을 확인할 수 있다. 이벤트 프레임 간 광류 추정을 수행하는 경우 수정된 TV-L1은 기존 TV-L1과 거의 동일한 프레임 레이트 상향 변환 성능과 광류 추정 성능을 보였다. RGB 프레임 간 광류 추정을 수행하는 경우는 이벤트 프레임 간 광류 추정을 수행하는 경우에 비해 실행 시간은 더 짧게 나타났으나, 더 낮은 프레임 레이트 상향 변환 성능과 더 낮은 광류 추정 성능을 보였고, 미래 시점의 프레임을 요구한다는 한계가 있었다. 뿐만 아니라, 이벤트로부터 복셀 그리드를 생성한 후 U-Net 딥러닝 모델을 사용하여 광류 추정을 하는 방법인 MotionPriorCM^[19]은 TV-L1을 사용하는 다른 방법들보다 우수한 광류 추정 성능을 보였지만, GPU를 사용하지 않는 CPU 실행 환경에서 실행 시간이 TV-L1을 사용하는 다른 방법들보다 훨씬 더 커서 연산 능력이 제한적인 환경에서 제안하는 방법이 효율이 있음을 시사한다.

표 3. 단일 시퀀스에 대한 광류 추정 방법별 프레임 레이트 상향 변환 성능 및 광류 추정 성능 비교 결과

Table 3. Frame rate up-conversion and optical flow estimation performance comparison results by optical flow estimation method for a single sequence

Method	PSNR (dB)	Optical Flow Estimation Result			
		RAEE Mean	RAEE Variance	Running Time Mean	Causality
TV-L1 on Event Frame	28.61	2.04	0.17	0.1883	O
Modified TV-L1 on Event Frame (Proposal)	28.60	2.05	0.18	0.1782	O
TV-L1 on RGB Frame (Interpolation)	28.45	2.39	0.26	0.1644	X
Modified TV-L1 on RGB Frame (Interpolation)	28.27	2.36	0.27	0.1696	X
MotionPriorCM on Event Frame	N/A	1.33	0.03	1.4656	O

표 4. 모든 시퀀스에 대한 광류 추정 방법별 프레임 레이트 상향 변환 성능 비교 결과

Table 4. Frame rate up-conversion performance comparison results by optical flow estimation method for all sequence

Sequence	FRUC Result (PSNR (dB))			
	Modified TV-L1 on Event Frame (Proposal)	TV-L1 on Event Frame	TV-L1 on RGB Frame (Extrapolation)	TV-L1 on RGB Frame (Interpolation)
thun_00_a	28.60	28.61	27.70	28.45
zurich_city_01_a	29.20	29.23	28.12	29.36
zurich_city_02_a	27.96	27.98	27.25	28.72
zurich_city_02_c	25.92	25.93	25.42	26.35
zurich_city_02_d	26.28	26.33	25.33	26.27
zurich_city_02_e	26.94	26.99	26.09	26.97
zurich_city_03_a	29.89	29.92	29.43	30.48
zurich_city_05_a	25.63	25.66	24.68	25.83
zurich_city_05_b	25.47	25.47	24.33	25.23
zurich_city_06_a	22.64	22.68	22.94	23.82
zurich_city_08_a	26.39	26.40	26.31	27.24
zurich_city_09_a	31.52	31.62	30.82	31.93
zurich_city_10_a	30.27	30.31	29.74	30.52
zurich_city_10_b	30.41	30.46	30.27	31.17
zurich_city_11_a	30.11	30.13	28.68	29.72
zurich_city_11_b	28.22	28.24	27.25	28.23
zurich_city_11_c	27.78	27.81	26.57	27.61
Overall (Mean/Standard Deviation)	27.83/2.23	27.86/2.23	27.11/2.15	28.11/2.17

표 4는 DSEC에서 정답 광류가 존재하는 모든 시퀀스를 대상으로, RGB 영상 프레임 레이트 상향 변환을 수행함에 있어서 이벤트 프레임 간 광류 추정에 대해 수정된 TV-L1과 기존 TV-L1 알고리즘의 성능과, RGB 프레임 내삽 광류 추정과 외삽 광류 추정에 대해 성능 비교 결과를 나타낸 표이다. 이벤트 프레임 생성 방법은 가중치가 16인 단순 누적 방식으로 하였으며, 입력 RGB 프레임은 정합된 RGB 프레임을 사용하였고, 광류 추정 알고리즘은 TV-L1 알고리즘의 각 반복별 에너지 함수의 수렴 여부 판정을 위한 에러의 임계값 ϵ 을 0.1로 사용하였다.

표 5는 DSEC에서 정답 광류가 존재하는 모든 시퀀스를 대상으로 광류 추정을 수행함에 있어서 이벤트 프레임 간 광류 추정에 대해 수정된 TV-L1과 기존 TV-L1 알고리즘의 성능과, RGB 프레임 간 광류 추정에 대해 기존 TV-L1 알고리즘의 성능과 이벤트로부터 복셀 그리드를 생성한 후 U-Net 딥러닝 모델을 사용하여 광류 추정을 하는 방법인 MotionPriorCM^[19]과의 성능 비교 결과를 나타낸 표이다. 여기서 RGB 프레임 간 광류 추정 성능은 내삽 방식과 외삽 방식이 동일하다. 이벤트 프레임 생성 방법은 가중치가 16인 단순 누적 방식으로 하였으며, 입력 RGB 프레임은

표 5. 모든 시퀀스에 대한 광류 추정 성능 비교 결과

Table 5. Optical flow estimation performance comparison results for all sequence

Sequence	Optical Flow Estimation Result (RAEE)			
	Modified TV-L1 on Event Frame (Proposal)	TV-L1 on Event Frame	TV-L1 on RGB Frame	Motion PriorCM on Event Frame
thun_00_a	2.05	2.05	2.39	1.33
zurich_city_01_a	2.64	2.64	2.84	1.18
zurich_city_02_a	1.88	1.84	2.61	1.22
zurich_city_02_c	2.80	2.80	3.06	1.88
zurich_city_02_d	3.03	3.03	3.20	1.24
zurich_city_02_e	2.76	2.77	2.91	1.98
zurich_city_03_a	1.38	1.37	1.92	2.16
zurich_city_05_a	3.32	3.32	3.35	1.38
zurich_city_05_b	2.99	2.99	3.12	1.52
zurich_city_06_a	2.97	2.89	2.82	1.77
zurich_city_08_a	2.12	2.11	2.42	2.02
zurich_city_09_a	2.43	2.42	2.60	1.77
zurich_city_10_a	3.25	3.25	3.25	1.74
zurich_city_10_b	2.40	2.39	2.60	1.80
zurich_city_11_a	1.79	1.79	2.06	2.20
zurich_city_11_b	2.01	2.01	2.40	2.19
zurich_city_11_c	2.23	2.23	2.53	2.18
Overall (Mean/Standard Deviation)	2.47/ 0.54	2.46/ 0.54	2.71/ 0.39	1.73/ 0.35

정합된 RGB 프레임을 사용하였고, 광류 추정 알고리즘은 TV-L1 알고리즘을 사용하는 경우에는 각 반복별 에너지 합수의 수렴 여부 판정을 위한 에러의 임계값 ε 을 0.1로 사용하였다.

실험 결과, 전반적으로 수정된 TV-L1이 기존 TV-L1과 광류 추정 및 프레임 레이트 상향 변환에 유사한 성능을 보였다. 프레임 레이트 상향 변환 측면에서는, RGB 프레임 내삽 방식이 가장 높은 PSNR을 보였지만, 미래 시점의 프레임을 요구한다는 한계가 있었다. 광류 추정 측면에서는 이벤트 프레임을 사용하는 방식이 대부분의 시퀀스에서 더 나은 성능을 보여주었다. 제안하는 방법인 이벤트 프레임 간 수정된 TV-L1을 사용하는 방식은 RGB 프레임 외삽 방식보다 프레임 레이트 상향 변환에서 0.69dB 높은 PSNR을, 광류 추정에서는 0.24 낮은 RAEE를 보였을 뿐만 아니

라 미래 시점의 프레임을 사용하지 않아, 이벤트 프레임 간 광류 추정이 제안하는 방법이 더 우수한 성능을 가짐을 시사한다. 뿐만 아니라, 이벤트로부터 복셀 그리드를 생성한 후 U-Net 딥러닝 모델을 사용하여 광류 추정을 하는 방법인 MotionPriorCM^[19]은 TV-L1을 사용하는 다른 방법들보다 RAEE에서 전체적으로 우수한 성능을 보였지만, 일부 시퀀스에서는 TV-L1을 사용하는 다른 방법들이 더 우수한 성능을 보이기도 하였다.

4. RGB 영상 프레임 레이트 상향 변환

표 6은 DSEC에서 일부 시퀀스를 대상으로 이벤트 프레임 간 광류 추정에 대해 수정된 TV-L1을 사용하는 방식과 기존 TV-L1을 사용하는 방식별로 프레임 레이트 상향 변

표 6. 프레임 레이트 상향 변환 방법별 화질 비교 분석

Table 6. Image quality comparison analysis by frame rate up-conversion method

Sequence	Ground Truth Frame	Modified TV-L1 on Event Frame (Proposal)		TV-L1 on Event Frame	
		Generated Frame	PSNR (dB)	Generated Frame	PSNR (dB)
zurich_city_02_d			19.93		19.31
zurich_city_02_e			21.27		17.98
zurich_city_03_a			25.90		23.72
zurich_city_09_a			24.47		22.94
zurich_city_10_b			22.18		20.61

한 결과물을 정답 RGB 프레임과의 비교를 통하여 보여준다. 이벤트 프레임 생성 방법은 가중치가 16인 단순 누적 방식으로 하였으며, 입력 RGB 프레임은 정합된 RGB 프레임을 사용하였고, 광류 추정 알고리즘은 TV-L1 알고리즘의 각 반복별 에너지 함수의 수렴 여부 판정을 위한 에러의 임계값 ε 을 0.1로 사용하였다.

주관적 화질을 비교 분석하면, 프레임의 경계 주위에서 급격한 밝기 변화가 발생하는 경우, 기존 TV-L1 방식으로 움직임을 추정하고, 움직임을 보상하여 프레임 레이트 상향 변환을 수행하는 경우는 실제로 밝기 변화가 발생했지만, 밝기 불변 가정을 적용하여 이를 반영하지 못해 신뢰도가 낮은 움직임을 추정하게 되고, 그 결과 움직임 보상을 시설제 광원은 하나이나 또 다른 광원이 산발적으로 보이는 현상(zurich_city_02_d, zurich_city_02_e), 광원 주위의 빛 변화가 불연속적으로 보이는 현상(zurich_city_03_a), 광원 주위에 링 모양으로 또 다른 광원이 보이는 현상(zurich_city_09_a, zurich_city_10_b) 등이 나타났다. 그러나, 밝기 불변 가정, 밝기 중앙값 가정, 밝기 반전 가정을 모두 적용하는 수정된 TV-L1 방식으로 움직임을 추정하는 경우에서는 이러한 현상이 발생하지 않았다. 뿐만 아니라, 표 4와 같이 시퀀스 전체로 보면 수정된 TV-L1은 객관적 화질이 기존 TV-L1과 거의 유사한 반면, 표 6과 같이 급격한 밝기 변화가 발생하는 일부 프레임에 대해서는 객관적 화질은 더 우수했다.

V. 결 론

본 논문에서는 이벤트의 높은 시간 해상도를 활용하여 RGB 프레임의 시간 해상도를 향상하는 프레임 레이트 상향 변환 방법을 제안하였다. 제안하는 방법은 이벤트 누적 방식에 기반한 이벤트 프레임 생성, RGB 프레임으로부터 RGB 이벤트 프레임 생성, RGB 프레임과 이벤트 프레임 간의 정합, TV-L1 알고리즘의 이벤트 프레임 간 광류 추정, 그리고 추정된 광류를 활용한 RGB 프레임 움직임 보상으로 구성된다. 이벤트 프레임 생성에서는 단순 누적 방식, 저주파수 필터링 방식, 고주파수 필터링 방식과 가중치 스케일 조절을 통해 프레임 레이트 상향 변환 성능 향상에

기여하는 최적 조합을 도출하였고, RGB 프레임으로부터 RGB 이벤트 프레임 생성에서는 RGB 프레임과 이벤트 프레임 간 효과적인 정합을 위한 연산량이 적은 방법을 제시하였다. 정합 과정에서는 FAST-BRIEF 기반 특징점 정합과 호모그래피 추정을 통해 정합 신뢰도를 확보하였다. 또한 TV-L1 알고리즘을 이벤트 프레임 특성에 맞게 3가지 밝기 가정 기반 데이터 항으로 수정함으로써, 극성 반전이나 이벤트 희소성 문제에 강건한 광류 추정이 가능하게 했다. 마지막으로, 추정된 광류를 통하여 움직임 보상을 수행하여 RGB 영상 프레임 레이트 상향 변환을 수행하였다. 그리고 이는 현재 시점의 RGB 프레임 또는 미래 시점의 RGB 프레임 없이, 과거 시점의 RGB 프레임만으로 이벤트 간의 관계를 통해서 수행하였다.

제안하는 방법은 DSEC 데이터셋을 기준으로 실험한 결과, 프레임 레이트 상향 변환을 과거 시점의 RGB 프레임을 이용하여 현재 시점의 RGB 프레임을 생성하는 외삽 방법으로 수행하는 것 대비 객관적으로는 PSNR 기준 0.72dB 향상된 성능을 보였으며, 프레임 레이트 상향 변환을 과거 시점의 RGB 프레임과 미래 시점의 RGB 프레임을 이용하여 현재 시점의 RGB 프레임을 생성하는 내삽 방법으로 수행하는 것 대비 객관적으로는 PSNR 기준 0.28dB 낮은 성능을 보였으나, 미래 시점의 프레임 없이 수행이 가능했다. 프레임 레이트 상향 변환을 단일 밝기 불변 가정을 사용하여 과거 시점의 이벤트 프레임과 현재 시점의 이벤트 프레임을 이용하여 현재 시점의 RGB 프레임을 생성하는 것 대비 객관적으로는 PSNR 기준 0.03dB 차이로 거의 동일한 성능을 보였고, 특히 급격한 밝기 변화가 발생하는 장면에 대해서는 객관적, 주관적 화질 모두 개선되었음을 보였다. 이를 통해 RGB 프레임의 시간 해상도 향상에 있어 이벤트 카메라의 시간적 특성을 효과적으로 활용할 수 있음을 확인하였다. 뿐만 아니라, 제안하는 방법은 이벤트로부터 복셀 그리드를 생성한 후 U-Net 딥러닝 모델을 사용하여 광류 추정을 하는 방법인 MotionPriorCM^[19] 대비 낮은 광류 추정 성능을 보였지만, 컴퓨팅 파워가 제한적인 환경에서 훨씬 더 적은 실행 시간을 가짐을 확인하였다.

향후 연구에서는 RGB 영상 프레임 레이트 상향 변환에서 더 높은 정밀도와 강건성을 위하여, 광류 추정에서 제안하는 데이터 항에 최적화된 정규화 항을 결합하는 것을 계

획하고 있다. 또한, 시공간 정보의 보존을 위하여, 이벤트로부터 복셀 그리드를 생성하여 처리하는 것 또한 계획하고 있다. 이러한 계획을 RGB 이벤트 다중 카메라 영상 처리 시스템으로 구현함에 있어서, 실제 응용에서의 실행 시간을 줄이기 위한 SIMD(Single Instruction Multiple Data) 최적화, FPGA(Field-Programmable Gate Array) 포팅 등을 고려할 예정이다.

참 고 문 현(References)

- [1] Scicluna, Adam, et al. "Towards Robust Perception for Assistive Robotics: An RGB-Event-LiDAR Dataset and Multi-Modal Detection Pipeline." 2024 IEEE 20th International Conference on Automation Science and Engineering (CASE). IEEE, 2024.
doi: <https://doi.org/10.1109/CASE59546.2024.10711832>
- [2] Dae Yeon Ju, et al. "High-speed Event-based Region Proposal Method." Proceedings of the Korean Institute of Broadcast and Media Engineers Conference (2024): 208-210. <https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE11995456>
- [3] Dae Yeon Ju, et al. "ROI Event Frame Generation Method for Enhancing Object Detection Performance using DVS." Proceedings of the Korean Institute of Communication Sciences Conference (2024): 705-706. <https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE11737236>
- [4] Gehrig, Daniel, and Davide Scaramuzza. "Low-latency automotive vision with event cameras." Nature 629.8014 (2024): 1034-1040.
doi: <https://doi.org/10.1038/s41586-024-07409-w>
- [5] Gallego, Guillermo, et al. "Event-based vision: A survey." IEEE transactions on pattern analysis and machine intelligence 44.1 (2020): 154-180.
doi: <https://doi.org/10.1109/TPAMI.2020.3008413>
- [6] Javier Sánchez Pérez, Enric Meinhardt-Llopis, and Gabriele Facciolo, TV-L1 Optical Flow Estimation, Image Processing On Line, 3 (2013), pp. 137 - 150.
doi: <https://doi.org/10.5201/ipol.2013.26>
- [7] Luo, Yinhui, et al. "A review of homography estimation: advances and challenges." Electronics 12.24 (2023): 4977.
doi: <https://doi.org/10.3390/electronics12244977>
- [8] Rosten, Edward, and Tom Drummond. "Machine learning for high-speed corner detection." Computer Vision - ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006. Proceedings, Part I 9. Springer Berlin Heidelberg, 2006.
doi: https://doi.org/10.1007/11744023_34
- [9] Calonder, Michael, et al. "Brief: Binary robust independent elementary features." Computer Vision - ECCV 2010: 11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5-11, 2010, Proceedings, Part IV 11. Springer Berlin Heidelberg, 2010.
doi: https://doi.org/10.1007/978-3-642-15561-1_56
- [10] Ou, Yangming, et al. "DRAMMS: Deformable registration via attribute matching and mutual-saliency weighting." Medical image analysis 15.4 (2011): 622-639.
doi: <https://doi.org/10.1016/j.media.2010.07.002>
- [11] Wachinger, Christian, et al. "Gaussian process interpolation for uncertainty estimation in image registration." International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer International Publishing, 2014.
doi: https://doi.org/10.1007/978-3-319-10404-1_34
- [12] Zheng, Xu, et al. "Deep learning for event-based vision: A comprehensive survey and benchmarks." arXiv preprint arXiv:2302.08890 (2023).
doi: <https://doi.org/10.48550/arXiv.2302.08890>
- [13] Muglikar, Manasi, et al. "How to calibrate your event camera." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021.
doi: <https://doi.org/10.48550/arXiv.2105.12362>
- [14] DSEC-flow: Optical Flow Benchmark - DSEC, <https://dsec.ifl.uzh.ch/uzh/dsec-flow-optical-flow-benchmark/> (accessed August. 26,2025)
- [15] Brebion, Vincent, Julien Moreau, and Franck Davoine. "Real-time optical flow for vehicular perception with low-and high-resolution event cameras." IEEE Transactions on Intelligent Transportation Systems 23.9 (2021): 15066-15078.
doi: <https://doi.org/10.1109/TITS.2021.3136358>
- [16] Shiba, Shintaro, Yoshimitsu Aoki, and Guillermo Gallego. "Secrets of event-based optical flow." European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2022.
doi: https://doi.org/10.1007/978-3-031-19797-0_36
- [17] Adarve, Juan David, and Robert Mahony. "A filter formulation for computing real time optical flow." IEEE Robotics and Automation Letters 1.2 (2016): 1192-1199.
doi: <https://doi.org/10.1109/LRA.2016.2532928>
- [18] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical image computing and computer-assisted intervention. Cham: Springer international publishing, 2015.
doi: https://doi.org/10.1007/978-3-319-24574-4_28
- [19] Hamann, Friedhelm, et al. "Motion-prior contrast maximization for dense continuous-time motion estimation." European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2024.
doi: https://doi.org/10.1007/978-3-031-72646-0_2
- [20] Paredes-Vallés, Federico, et al. "Taming contrast maximization for learning sequential, low-latency, event-based optical flow." Proceedings of the IEEE/CVF international conference on computer vision. 2023.
doi: <https://doi.org/10.48550/arXiv.2303.05214>
- [21] Cuadrado, Javier, et al. "Optical flow estimation from event-based cameras and spiking neural networks." Frontiers in Neuroscience 17 (2023): 1160034.
doi: <https://doi.org/10.3389/fnins.2023.1160034>
- [22] Tian, Yi, and Juan Andrade-Cetto. "SDformerFlow: Spiking neural

- network transformer for event-based optical flow.” International Conference on Pattern Recognition. Cham: Springer Nature Switzerland, 2024.
doi: https://doi.org/10.1007/978-3-031-78354-8_30
- [23] Tulyakov, Stepan, et al. “Time lens: Event-based video frame interpolation.” Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021.
doi: <https://doi.org/10.48550/arXiv.2106.07286>
- [24] Palwal, Avinash, and Nima Khademi Kalantari. “Deep slow motion video reconstruction with hybrid imaging system.” IEEE transactions on pattern analysis and machine intelligence 42.7 (2020): 1557-1569.
doi: <https://doi.org/10.1109/TPAMI.2020.2987316>
- [25] Hu, Yuhuang, Shih-Chii Liu, and Tobi Delbruck. “v2e: From video frames to realistic DVS events.” Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021.
doi: <https://doi.org/10.48550/arXiv.2006.07722>
- [26] Jiang, Huaizu, et al. “Super slomo: High quality estimation of multiple intermediate frames for video interpolation.” Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.
doi: <https://doi.org/10.48550/arXiv.1712.00080>
- [27] Gehrig, Mathias, et al. “Dsec: A stereo event camera dataset for driving scenarios.” IEEE Robotics and Automation Letters 6.3 (2021): 4947-4954.
doi: <https://doi.org/10.1109/LRA.2021.3068942>
- [28] normalized cross correlation, <https://www.sciencedirect.com/topics/computer-science/normalized-cross-correlation/> (accessed August. 26, 2025)
- [29] Baker, Simon, et al. “A database and evaluation methodology for optical flow.” International journal of computer vision 92.1 (2011): 1-31.
doi: <https://doi.org/10.1007/s11263-010-0390-2>

저자 소개



김 영 육

- 2024년 2월 : 광운대학교 컴퓨터공학과 학사
- 2024년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 석사과정
- ORCID : <https://orcid.org/0009-0002-1120-110X>
- 주관심분야 : RGB-이벤트 융합 영상처리, 영상 압축, 컴퓨터 비전



주 대 연

- 2023년 2월 : 광운대학교 전자공학과 학사
- 2025년 2월 : 광운대학교 컴퓨터공학과 석사
- 2025년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 박사과정
- ORCID : <https://orcid.org/0009-0006-7956-3077>
- 주관심분야 : 영상 압축, 컴퓨터 비전



박 장 수

- 2016년 8월 : 광운대학교 컴퓨터공학과 학사
- 2019년 2월 : 광운대학교 컴퓨터공학과 석사
- 2019년 8월 ~ 2023년 1월 : 아이브스 주식회사 선임연구원
- 2023년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 박사과정
- ORCID : <https://orcid.org/0009-0005-9312-6781>
- 주관심분야 : 영상신호처리, 고성능컴퓨팅

저자소개

박주현



- 1993년 2월 : 전남대학교 전자공학과 공학사
- 1995년 2월 : 전남대학교 전자공학과 공학석사
- 1999년 2월 : 전남대학교 전자공학과 공학박사
- 1999년 3월 ~ 2002년 4월 : 한국전자통신연구원 선임연구원
- 2002년 9월 ~ 2009년 3월 : (주)마무리안디자인 연구소장
- 2012년 3월 ~ 2018년 2월 : (주)팹셀플러스 재직
- 2021년 2월 ~ 현재 : (주)아르고 재직
- ORCID : <https://orcid.org/0009-0009-5954-7418>
- 주관심분야 : SoCaaS, 스마트에너지 및 보안 SoC 반도체

정해룡



- 1993년 2월 : 서강대학교 전자공학과 공학사
- 1995년 2월 : 서강대학교 전자공학과 공학석사
- 1995년 3월 ~ 2000년 8월 : 삼성전자 선임연구원
- 2000년 9월 ~ 2006년 4월 : 이오네스 책임연구원
- 2006년 5월 ~ 2009년 12월 : 포스코ICT 수석연구원
- 2010년 7월 ~ 2012년 7월 : TSMC Korea
- 2012년 8월 ~ 2018년 8월 : 동부하이텍 수석연구원
- 2020년 3월 ~ 2023년 2월 : 에이디테크놀로지 전무
- 2023년 3월 ~ 현재 : 아르고 연구소장
- ORCID : <https://orcid.org/0009-0006-5941-5472>
- 주관심분야 : SoC, 보안, 스마트그리드

심동규



- 1993년 2월 : 서강대학교 전자공학과 공학사
- 1995년 2월 : 서강대학교 전자공학과 공학석사
- 1999년 2월 : 서강대학교 전자공학과 공학박사
- 1999년 3월 ~ 2000년 8월 : 현대전자 선임연구원
- 2000년 9월 ~ 2002년 3월 : 바로비전 선임연구원
- 2002년 4월 ~ 2005년 2월 : University of Washington Senior research engineer
- 2005년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 교수
- ORCID : <https://orcid.org/0000-0002-2794-9932>
- 주관심분야 : 영상신호처리, 영상압축, 컴퓨터비전