



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2022년02월22일
(11) 등록번호 10-2366187
(24) 등록일자 2022년02월17일

(51) 국제특허분류(Int. Cl.)
G06T 3/40 (2006.01) G06N 3/04 (2006.01)
G06N 3/08 (2006.01)
(52) CPC특허분류
G06T 3/4053 (2013.01)
G06N 3/04 (2013.01)
(21) 출원번호 10-2020-0070044
(22) 출원일자 2020년06월10일
심사청구일자 2020년06월10일
(65) 공개번호 10-2021-0153229
(43) 공개일자 2021년12월17일
(56) 선행기술조사문헌
Wang et al. Event-based High Dynamic Range
Image and Very High Frame Rate Video
Generation using Conditional Generative
Adversarial Networks, CVPR, 2019년, pp.
10073-10082. 1부.*
KR1020180014992 A
KR1020190086334 A
KR1020200015589 A
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
광주과학기술원
광주광역시 북구 첨단과기로 123 (오룡동)
한국과학기술원
대전광역시 유성구 대학로 291(구성동)
(72) 발명자
최종현
광주광역시 북구 첨단과기로 123 광주과학기술원
AI대학원
모하마드 모스타파비
광주광역시 북구 첨단과기로 123 광주과학기술원
전기전자컴퓨터공학부
윤국진
대전광역시 유성구 대학로 291
(74) 대리인
특허법인지원

전체 청구항 수 : 총 10 항

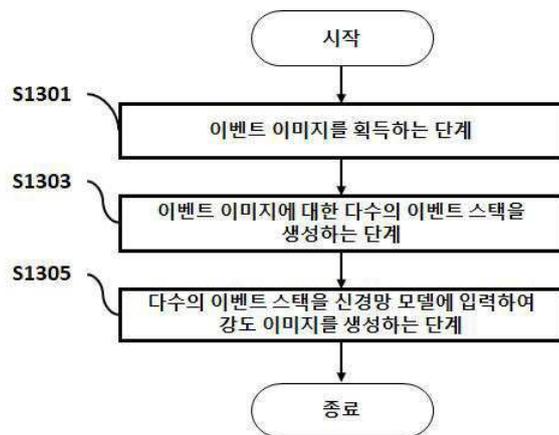
심사관 : 김창원

(54) 발명의 명칭 **고해상도 강도 이미지 생성 방법 및 장치**

(57) 요약

본 발명은 고해상도 강도 이미지 생성 방법 및 장치에 관한 것이다. 본 발명의 일 실시예에 따른 고해상도 강도 이미지 생성 방법은 (a) 이벤트 이미지를 획득하는 단계; (b) 상기 획득된 이벤트 이미지에 대한 다수의 이벤트 스택을 생성하는 단계; 및 (c) 상기 다수의 이벤트 스택을 신경망(neural network) 모델에 입력하여 강도 이미지(intensity image)를 생성하는 단계;를 포함할 수 있다.

대표도 - 도13



(52) CPC특허분류

G06N 3/08 (2013.01)
 G06T 3/4046 (2013.01)
 G06T 2207/20084 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 1711102946
 과제번호 NN25200
 부처명 과학기술정보통신부
 과제관리(전문)기관명 정보통신기획평가원
 연구사업명 선도기술확보 지능형반도체 기술개발(설계)
 연구과제명 활성화/커널데이터의 압축/복원을 통한 초저전력 모바일 딥러닝 반도체 기술 개발
 기여율 1/3
 과제수행기관명 전자부품연구원
 연구기간 2020.01.01 ~ 2020.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호 1711109566
 과제번호 2018R1A2B3008640
 부처명 과학기술정보통신부
 과제관리(전문)기관명 한국연구재단
 연구사업명 이공분야기초연구사업
 연구과제명 이벤트 카메라 기반 시각 지능을 위한 컴퓨터 비전 알고리즘 연구
 기여율 1/2
 과제수행기관명 한국과학기술원
 연구기간 2020.03.01 ~ 2021.02.28

이 발명을 지원한 국가연구개발사업

과제고유번호 1711097556
 과제번호 2017M3C4A7069369
 부처명 과학기술정보통신부
 과제관리(전문)기관명 한국연구재단
 연구사업명 원천기술개발사업 (차세대정보컴퓨팅기술개발사업)
 연구과제명 비전 모델 기반 공간 상황 인지 원천기술 연구
 기여율 1/6
 과제수행기관명 한국과학기술원
 연구기간 2020.04.01 ~ 2020.12.31

명세서

청구범위

청구항 1

각 픽셀의 강도 값(intensity value)이 변경됨에 따라 생성된 이벤트를 포함하는 이벤트 이미지를 획득하는 단계;

상기 획득된 이벤트 이미지에 대한 다수의 이벤트 스택을 생성하는 단계;

상기 다수의 이벤트 스택의 시퀀스(sequence)를 신경망(neural network) 모델에 입력하여 다수의 광학 플로우(optical flow)를 생성하는 단계;

상기 생성된 다수의 광학 플로우를 이용하여 상기 다수의 이벤트 스택을 수정(rectify)하는 단계;

상기 수정된 다수의 이벤트 스택과 상기 다수의 이벤트 스택에 대한 상태(state) 정보를 이용하여 다수의 중간 강도 이미지(intermediate intensity image)를 생성하는 단계; 및

상기 다수의 중간 강도 이미지를 혼합(mix)하여 상기 강도 이미지를 생성하는 단계;

를 포함하는,

고해상도 강도 이미지 생성 방법.

청구항 2

제1항에 있어서,

상기 다수의 이벤트 스택을 생성하는 단계는,

상기 이벤트 이미지의 각 픽셀에 대한 이벤트 스트림(event stream)을 생성하는 단계; 및

상기 생성된 이벤트 스트림을 분류하여 상기 다수의 이벤트 스택을 생성하는 단계;

를 포함하는,

고해상도 강도 이미지 생성 방법.

청구항 3

삭제

청구항 4

제1항에 있어서,

상기 강도 이미지를 생성하는 단계 이후에,

상기 생성된 강도 이미지와 GT(ground truth) 이미지의 차이(difference)에 기반한 오차(error) 정보를 산출하는 단계; 및

상기 산출된 오차 정보를 이용하여 상기 신경망 모델을 학습시키는 단계;

를 더 포함하는,

고해상도 강도 이미지 생성 방법.

청구항 5

제1항에 있어서,
 상기 강도 이미지를 생성하는 단계는,
 상기 다수의 이벤트 스택의 시퀀스 중 중앙 스택(central stack)을 APS(active pixel sensor) 프레임으로 형성하는 단계; 및
 상기 중앙 스택으로 형성된 APS 프레임과 상기 다수의 이벤트 스택의 시퀀스 중 상기 중앙 스택을 제외한 다수의 주변 스택(nearby stack)을 상기 신경망 모델에 입력하여 상기 강도 이미지를 생성하는 단계;
 를 포함하는,
 고해상도 강도 이미지 생성 방법.

청구항 6

제1항에 있어서,
 상기 강도 이미지를 생성하는 단계 이후에,
 상기 다수의 이벤트 스택의 시퀀스 중 중앙 스택을 상기 생성된 강도 이미지로 형성하는 단계; 및
 상기 중앙 스택으로 형성된 강도 이미지와 상기 다수의 이벤트 스택의 시퀀스 중 상기 중앙 스택을 제외한 다수의 주변 스택을 상기 신경망 모델에 입력하여 다른 강도 이미지를 생성하는 단계;
 를 더 포함하는,
 고해상도 강도 이미지 생성 방법.

청구항 7

각 픽셀의 강도 값(intensity value)이 변경됨에 따라 생성된 이벤트를 포함하는 이벤트 이미지를 획득하는 획득부; 및
 상기 획득된 이벤트 이미지에 대한 다수의 이벤트 스택을 생성하고,
 상기 다수의 이벤트 스택의 시퀀스(sequence)를 신경망(neural network) 모델에 입력하여 다수의 광학 플로우(optical flow)를 생성하고,
 상기 생성된 다수의 광학 플로우를 이용하여 상기 다수의 이벤트 스택을 수정(rectify)하고,
 상기 수정된 다수의 이벤트 스택과 상기 다수의 이벤트 스택에 대한 상태(state) 정보를 이용하여 다수의 중간 강도 이미지(intermediate intensity image)를 생성하고,
 상기 다수의 중간 강도 이미지를 혼합(mix)하여 상기 강도 이미지를 생성하는 제어부;
 를 포함하는,
 고해상도 강도 이미지 생성 장치.

청구항 8

제7항에 있어서,
 상기 제어부는,
 상기 이벤트 이미지의 각 픽셀에 대한 이벤트 스트림(event stream)을 생성하고,
 상기 생성된 이벤트 스트림을 분류하여 상기 다수의 이벤트 스택을 생성하는,

고해상도 강도 이미지 생성 장치.

청구항 9

삭제

청구항 10

제7항에 있어서,

상기 제어부는,

상기 생성된 강도 이미지와 GT(ground truth) 이미지의 차이(difference)에 기반한 오차(error) 정보를 산출하고,

상기 산출된 오차 정보를 이용하여 상기 신경망 모델을 학습시키는,

고해상도 강도 이미지 생성 장치.

청구항 11

제7항에 있어서,

상기 제어부는,

상기 다수의 이벤트 스택의 시퀀스 중 중앙 스택(central stack)을 APS(active pixel sensor) 프레임으로 형성하고,

상기 중앙 스택으로 형성된 APS 프레임과 상기 다수의 이벤트 스택의 시퀀스 중 상기 중앙 스택을 제외한 다수의 주변 스택(nearby stack)을 상기 신경망 모델에 입력하여 상기 강도 이미지를 생성하는,

고해상도 강도 이미지 생성 장치.

청구항 12

제7항에 있어서,

상기 제어부는,

상기 다수의 이벤트 스택의 시퀀스 중 중앙 스택을 상기 생성된 강도 이미지로 형성하고,

상기 중앙 스택으로 형성된 강도 이미지와 상기 다수의 이벤트 스택의 시퀀스 중 상기 중앙 스택을 제외한 다수의 주변 스택을 상기 신경망 모델에 입력하여 다른 강도 이미지를 생성하는,

고해상도 강도 이미지 생성 장치.

발명의 설명

기술 분야

[0001] 본 발명은 고해상도 강도 이미지 생성 방법 및 장치에 관한 것으로, 더욱 상세하게는 이벤트 이미지를 이용한 고해상도 강도 이미지 생성 방법 및 장치에 관한 것이다.

배경 기술

[0003] 뉴로모픽(neuromorphic) 카메라로 지칭될 수 있는 이벤트 카메라는 낮은 대기 시간과 낮은 전력 소비로 저렴한 비용과 높은 동적 감지 범위의 장점을 이용하여 컴퓨터 비전 및 로봇 공학 분야의 발전과 함께 발전하였다.

[0004] 이벤트 카메라는 낮은 대기 시간, 낮은 전력 및 낮은 메모리 소비 제약으로 높은 동적 범위 출력을 요구하는 시스템에서 높은 적용 가능성으로 인해 많은 주목을 받았다.

[0005] 대부분의 상용 이벤트 카메라는 효율성을 위해 상대적으로 저해상도 이벤트 스트림을 생성한다. 이에, 많은 응용 프로그램에 대한 수많은 제안이 있지만, 이벤트에서 고해상도 이미지를 추정하는 것에 대한 연구는 미흡한 실정이다.

선행기술문헌

비특허문헌

[0007] (비특허문헌 0001) [비특허문헌 1] Patrick Bardow, Andrew J Davison, and Stefan Leutenegger. Simultaneous optical flow and intensity estimation from an event camera. In IEEE CVPR, pages 884-892, 2016

발명의 내용

해결하려는 과제

[0008] 본 발명은 전술한 문제점을 해결하기 위하여 창출된 것으로, 고해상도 강도 이미지 생성 방법 및 장치를 제공하는 것을 그 목적으로 한다.

[0009] 또한, 본 발명은 다수의 이벤트 스택을 신경망(neural network) 모델에 입력하여 고해상도 강도 이미지(intensity image)를 생성하기 위한 방법 및 장치를 제공하는 것을 그 목적으로 한다.

[0010] 본 발명의 목적들은 이상에서 언급한 목적들로 제한되지 않으며, 언급되지 않은 또 다른 목적들은 아래의 기재로부터 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0012] 상기한 목적들을 달성하기 위하여, 본 발명의 일 실시예에 따른 고해상도 강도 이미지 생성 방법은 (a) 이벤트 이미지를 획득하는 단계; (b) 상기 획득된 이벤트 이미지에 대한 다수의 이벤트 스택을 생성하는 단계; 및 (c) 상기 다수의 이벤트 스택을 신경망(neural network) 모델에 입력하여 강도 이미지(intensity image)를 생성하는 단계;를 포함할 수 있다.

[0013] 실시예에서, 상기 (b) 단계는, 상기 이벤트 이미지의 각 픽셀에 대한 이벤트 스트림(event stream)을 생성하는 단계; 및 상기 생성된 이벤트 스트림을 분류하여 상기 다수의 이벤트 스택을 생성하는 단계;를 포함할 수 있다.

[0014] 실시예에서, 상기 (c) 단계는, 상기 다수의 이벤트 스택의 시퀀스(sequence)를 이용하여 다수의 광학 플로우(optical flow)를 생성하는 단계; 상기 생성된 다수의 광학 플로우를 이용하여 상기 다수의 이벤트 스택을 수정(rectify)하는 단계; 상기 수정된 다수의 이벤트 스택과 상기 다수의 이벤트 스택에 대한 상태(state) 정보를 이용하여 다수의 중간 강도 이미지(intermediate intensity image)를 생성하는 단계; 및 상기 다수의 중간 강도 이미지를 혼합(mix)하여 상기 강도 이미지를 생성하는 단계;를 포함할 수 있다.

[0015] 실시예에서, 상기 (c) 단계는, 상기 생성된 강도 이미지와 GT(ground truth) 이미지의 차이(difference)에 기반한 오차(error) 정보를 산출하는 단계; 및 상기 산출된 오차 정보를 이용하여 상기 신경망 모델을 학습시키는 단계;를 포함할 수 있다.

[0016] 실시예에서, 상기 (c) 단계는, 상기 다수의 이벤트 스택의 시퀀스 중 중앙 스택(central stack)을 APS(active pixel sensor) 프레임으로 형성하는 단계; 및 상기 중앙 스택으로 형성된 APS 프레임과 상기 다수의 이벤트 스택의 시퀀스 중 상기 중앙 스택을 제외한 다수의 주변 스택(nearby stack)을 상기 신경망 모델에 입력하여 상기 강도 이미지를 생성하는 단계;를 포함할 수 있다.

[0017] 실시예에서, 상기 고해상도 강도 이미지 생성 방법은, 상기 (c) 단계 이후에, 상기 다수의 이벤트 스택의 시퀀스 중 중앙 스택을 상기 생성된 강도 이미지로 형성하는 단계; 및 상기 중앙 스택으로 형성된 강도 이미지와 상기 다수의 이벤트 스택의 시퀀스 중 상기 중앙 스택을 제외한 다수의 주변 스택을 상기 신경망 모델에 입력하여

다른 강도 이미지를 생성하는 단계;를 더 포함할 수 있다.

- [0018] 실시예에서, 고해상도 강도 이미지 생성 장치는, 이벤트 이미지를 획득하는 획득부; 및 상기 획득된 이벤트 이미지에 대한 다수의 이벤트 스택을 생성하고, 상기 다수의 이벤트 스택을 신경망(neural network) 모델에 입력하여 강도 이미지(intensity image)를 생성하는 제어부;를 포함할 수 있다.
- [0019] 실시예에서, 상기 제어부는, 상기 이벤트 이미지의 각 픽셀에 대한 이벤트 스트림(event stream)을 생성하고, 상기 생성된 이벤트 스트림을 분류하여 상기 다수의 이벤트 스택을 생성할 수 있다.
- [0020] 실시예에서, 상기 제어부는, 상기 다수의 이벤트 스택의 시퀀스(sequence)를 이용하여 다수의 광학 플로우(optical flow)를 생성하고, 상기 생성된 다수의 광학 플로우를 이용하여 상기 다수의 이벤트 스택을 수정(rectify)하고, 상기 수정된 다수의 이벤트 스택과 상기 다수의 이벤트 스택에 대한 상태(state) 정보를 이용하여 다수의 중간 강도 이미지(intermediate intensity image)를 생성하고, 상기 다수의 중간 강도 이미지를 혼합(mix)하여 상기 강도 이미지를 생성할 수 있다.
- [0021] 실시예에서, 상기 제어부는, 상기 생성된 강도 이미지와 GT(ground truth) 이미지의 차이(difference)에 기반한 오차(error) 정보를 산출하고, 상기 산출된 오차 정보를 이용하여 상기 신경망 모델을 학습시킬 수 있다.
- [0022] 실시예에서, 상기 제어부는, 상기 다수의 이벤트 스택의 시퀀스 중 중앙 스택(central stack)을 APS(active pixel sensor) 프레임으로 형성하고, 상기 중앙 스택으로 형성된 APS 프레임과 상기 다수의 이벤트 스택의 시퀀스 중 상기 중앙 스택을 제외한 다수의 주변 스택(nearby stack)을 상기 신경망 모델에 입력하여 상기 강도 이미지를 생성할 수 있다.
- [0023] 실시예에서, 상기 제어부는, 상기 다수의 이벤트 스택의 시퀀스 중 중앙 스택을 상기 생성된 강도 이미지로 형성하고, 상기 중앙 스택으로 형성된 강도 이미지와 상기 다수의 이벤트 스택의 시퀀스 중 상기 중앙 스택을 제외한 다수의 주변 스택을 상기 신경망 모델에 입력하여 다른 강도 이미지를 생성할 수 있다.
- [0024] 상기한 목적들을 달성하기 위한 구체적인 사항들은 첨부된 도면과 함께 상세하게 후술될 실시예들을 참조하면 명확해질 것이다.
- [0025] 그러나, 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라, 서로 다른 다양한 형태로 구성될 수 있으며, 본 발명의 개시가 완전하도록 하고 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자(이하, "통상의 기술자")에게 발명의 범주를 완전하게 알려주기 위해서 제공되는 것이다.

발명의 효과

- [0027] 본 발명의 일 실시예에 의하면, 고 신뢰도의 고해상도 이미지를 생성하기 위해, 이벤트에서 픽셀 단위의 고해상도 강도 이미지를 엔드 투 엔드 방식으로 직접 추정하는 방법을 학습하고, 상세한 세부 정보(detail)와 아티팩트(artifact)가 적은 이미지를 고해상도화 할 수 있다.
- [0028] 또한, 본 발명의 일 실시예에 의하면, APS 프레임을 입력으로 고려하거나 신경망 모델을 반복적으로 학습하여 초기 이미지에 세부 정보를 추가함으로써 더 자세한 세부 정보를 갖는 강도 이미지를 생성할 수 있다.
- [0029] 본 발명의 효과들은 상술된 효과들로 제한되지 않으며, 본 발명의 기술적 특징들에 의하여 기대되는 잠정적인 효과들은 아래의 기재로부터 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

- [0031] 도 1a는 종래의 카메라 이미지를 도시한 도면이다.
- 도 1b는 본 발명의 일 실시예에 따른 이벤트 카메라의 이벤트 이미지를 도시한 도면이다.
- 도 2a 내지 2c는 본 발명의 일 실시예에 따른 이벤트 수 기반 이벤트 스택을 도시한 도면이다.
- 도 3a 내지 3c는 본 발명의 일 실시예에 따른 신경망 모델을 도시한 도면이다.
- 도 4a 및 4b는 본 발명의 일 실시예에 따른 이벤트 스택의 시퀀스를 도시한 도면이다.
- 도 5a 및 5b는 본 발명의 일 실시예에 따른 SRNet을 도시한 도면이다.
- 도 6a 내지 6c는 본 발명의 일 실시예에 따른 고해상도 강도 이미지 생성의 성능 비교를 도시한 도면이다.

도 7a 및 7b는 본 발명의 일 실시예에 따른 고해상도 강도 이미지 생성의 성능 그래프를 도시한 도면이다.

도 8a 및 8b는 본 발명의 일 실시예에 따른 다운스케일 출력의 성능 비교를 도시한 도면이다.

도 9는 본 발명의 일 실시예에 따른 손실 함수에 대한 성능 비교를 도시한 도면이다.

도 10은 본 발명의 일 실시예에 따른 HDR 시나리오에서의 이미지 재구성 비교를 도시한 도면이다.

도 11a는 본 발명의 다른 실시예에 따른 고해상도 강도 이미지 생성을 도시한 도면이다.

도 11b는 본 발명의 또 다른 실시예에 따른 고해상도 강도 이미지 생성을 도시한 도면이다.

도 12a 및 12b는 본 발명의 다양한 실시예에 따른 고해상도 강도 이미지 생성의 성능 비교를 도시한 도면이다.

도 13은 본 발명의 다양한 실시예에 따른 고해상도 강도 이미지 생성 방법을 도시한 도면이다.

도 14는 본 발명의 다양한 실시예에 따른 고해상도 강도 이미지 생성 장치의 기능적 구성을 도시한 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0032] 본 발명은 다양한 변경을 가할 수 있고, 여러 가지 실시예들을 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 이를 상세히 설명하고자 한다.
- [0033] 청구범위에 개시된 발명의 다양한 특징들은 도면 및 상세한 설명을 고려하여 더 잘 이해될 수 있을 것이다. 명세서에 개시된 장치, 방법, 제법 및 다양한 실시예들은 예시를 위해서 제공되는 것이다. 개시된 구조 및 기능상의 특징들은 통상의 기술자로 하여금 다양한 실시예들을 구체적으로 실시할 수 있도록 하기 위한 것이고, 발명의 범위를 제한하기 위한 것이 아니다. 개시된 용어 및 문장들은 개시된 발명의 다양한 특징들을 이해하기 쉽게 설명하기 위한 것이고, 발명의 범위를 제한하기 위한 것이 아니다.
- [0034] 본 발명을 설명함에 있어서, 관련된 공지기술에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우, 그 상세한 설명을 생략한다.
- [0035] 이하, 본 발명의 일 실시예에 따른 고해상도 강도 이미지 생성 방법 및 장치를 설명한다.
- [0037] 도 1a는 종래의 카메라 이미지를 도시한 도면이다. 도 1b는 본 발명의 일 실시예에 따른 이벤트 카메라의 이벤트 이미지를 도시한 도면이다. 도 2a 내지 2d는 본 발명의 일 실시예에 따른 이벤트 수 기반 이벤트 스택을 도시한 도면이다.
- [0038] 도 1a를 참고하면, 기존의 카메라는 고정된 시간 간격으로 모든 픽셀의 프레임을 캡처하게 된다.
- [0039] 반면, 도 1b를 참고하면, 본 발명에 따른 이벤트 카메라는 최대 마이크로초 해상도(microsecond resolution)까지 이벤트 이미지(110)를 출력할 수 있다. 이 경우, 이벤트 이미지(110)의 각 픽셀이 강도 값(intensity value)을 변경할 때마다 이벤트가 생성될 수 있다.
- [0040] 일 실시예에서, 변경된 강도 값이 미리 정의된 임계값보다 큰 경우, 플러스 이벤트로 결정되고, 변경된 강도 값이 미리 정의된 임계값보다 작은 경우, 마이너스 이벤트로 결정될 수 있다.
- [0041] 일 실시예에서, 이벤트 이미지(110)의 각 픽셀에 대한 이벤트 스트림(event stream)(120)이 생성될 수 있다. 이벤트 스트림(120)은 모든 픽셀에 대한 이벤트 위치와 정확한 시간 및 부호 정보를 포함할 수 있다.
- [0042] 이벤트 스트림(120)은 미리 정의된 임계값으로 강도 값의 변화량을 확인하여 픽셀 위치(x, y)의 강도 변화를 플러스 또는 마이너스 부호 (σ)로 비동기적으로 나타낼 수 있다.
- [0043] 즉, 이벤트 스트림(120)은 정확한 타임스탬프(t)를 통해 μs 의 지연 시간을 달성할 수 있으며, 발생 이벤트를 (x, y, t, σ) 형식으로 표현할 수 있다.
- [0044] 다만, 이러한 이벤트 스트림(120)은 신경망 모델로 전달될 수 없다. 이 경우, 도 2a 내지 2c를 참고하면, 이벤트 스트림(120)은 이벤트 수에 기반한 스택킹(stacking based on the number of events, SBN)을 사용하여 더 작은 부분으로 스택킹되어, 다수의 이벤트 스택(210)으로 생성될 수 있다.
- [0045] SBN을 사용하면, 이벤트 스트림(120)의 타임스탬프에서 시작하여, 미리 정의된 수(Ne)에 도달할 때까지 이벤트 수를 계산하고, 이벤트 스택(210)에 하나의 채널을 형성하기 위해 이벤트를 누적할 수 있다. 이러한 과정은 각

이벤트 스택(210)에 대해 c번 반복될 수 있다.

- [0046] 따라서, 각 이벤트 스택(210)은 총 $M=c \times N_e$ 이벤트를 포함하고, $h \times w \times c$ 의 차원을 가질 수 있다. 여기서, h 및 w는 각각 이미지의 폭 및 높이를 나타낸다.
- [0047] 생성된 이벤트 스택(210)은 입력으로 신경망 모델에 공급될 수 있다. 해당 APS 프레임은 GT(ground truth) 스택의 마지막 이벤트의 타임스탬프에서 샘플링될 수 있다.
- [0048] 일 실시예에서, 각 채널에서, 모든 픽셀 값은 처음에 128로 설정될 수 있다. 이벤트가 위치 (x, y)에서 트리거 되면, 동일한 채널에서 (x, y)의 픽셀 값은 256(긍정 이벤트(positive event)) 또는 0(부정 이벤트(negative event))으로 변경될 수 있다.
- [0049] 새로운 이벤트는 이전 이벤트를 무시할 수 있으므로, 시공간 시각적 정보를 보다 잘 보존하기 위하여 M이 신중하게 선택될 수 있다. 프레임 레이트는 시간에 따른 각 이벤트 스택(210) 사이의 겹치는 이벤트 수와 Ne에 의해 결정될 수 있다.
- [0050] 예를 들어, 3개의 채널이 있는 각 이벤트 스택(210)은 3,000개의 이벤트를 포함할 수 있으나, 이에 제한되지 않는다.
- [0052] 도 3a 내지 3c는 본 발명의 일 실시예에 따른 신경망 모델을 도시한 도면이다. 도 4a 및 4b는 본 발명의 일 실시예에 따른 이벤트 스택의 시퀀스를 도시한 도면이다.
- [0053] 도 3a 내지 3c를 참고하면, 다수의 이벤트 스택(210)의 시퀀스(sequence)를 신경망 모델(300)에 입력할 수 있다.
- [0054] 예를 들어, 도 4a 및 4b를 참고하면, 다수의 이벤트 스택(210)의 시퀀스는 3개의 이벤트 스택(3S) 또는 7개의 이벤트 스택(7S)을 포함할 수 있다. 다만, 시퀀스에 포함되는 이벤트 스택의 개수는 제한되지 않는다.
- [0055] 일 실시예에서, 3개의 이벤트 스택(3S)의 시퀀스의 경우, 시퀀스는 n번째 APS 타임스탬프의 이벤트 스택(SBN_n), n번째 APS 타임스탬프 전의 이벤트 스택(SBN_{n-m}) 및 n번째 APS 타임스탬프 후의 이벤트 스택(SBN_{n+m})을 포함할 수 있다.
- [0056] 이 경우, 각 이벤트 스택은 M(예: 3,000)개의 이벤트를 포함할 수 있으며, 종료 위치 m은 M개의 이벤트를 발생시키는 데 필요한 시간에 기반하여 이벤트 타임 라인에 따라 결정될 수 있다. SBN_n 은 3개의 시퀀스 중 중앙 스택(central stack)을 의미할 수 있다.
- [0057] SBN_{n+m} 및 SBN_{n-m} 은 이벤트 스택 간에 겹침($L = 0$)이 없는 경우, 중앙 스택의 시작 또는 끝에서 각각 M개의 이벤트를 포함할 수 있다.
- [0058] 일 실시예에서, 더 높은 프레임 레이트를 생성하기 위해 이벤트 스택 간 오버랩(overlap)이 있을 수 있다. 이 경우, 다음 이벤트 스택의 끝은 중앙에서 겹친 양(M-L)을 뺀 후 M 개의 이벤트를 포함할 수 있다.
- [0059] 일 실시예에서, 해당 타임스탬프(time stamp)에 대한 다수의 이벤트 스택(210)의 시퀀스를 플로우 네트워크(Flow Network, FNet)(310)에 입력하여 다수의 광학 플로우(optical flow)(321)를 생성할 수 있다.
- [0060] 이벤트 스트림을 스택킹함에 따라 이벤트 스택 간의 시간적 관계를 잃을 수 있는데, 광학 플로우가 장면(scene)에서 트리거된 이벤트가 어떻게 이동하고 변경이 발생한 위치를 나타낼 때, 이벤트 스택(210)의 시퀀스와 각 이벤트 스택(210) 쌍 사이의 광학 플로우(321)를 사용하여 이벤트 스택(210) 간의 손실된 시간 관계를 부분적으로 복구할 수 있다.
- [0061] 예를 들어, SBN_{n+m} 및 SBN_{n-m} 각각은 SBN 과 함께 이벤트 스택 사이의 광학 플로우(F_{n+m} 및 F_{n-m})을 예측하기 위해 FNet(310)에 입력될 수 있다. 이에, 이벤트 스택은 FNet(310)에 의해 획득된 광학 플로우와 쌍(pair)으로 연결될 수 있다.
- [0062] 일 실시예에서, 다수의 광학 플로우(321)를 이벤트 기능 수정(Event Feature Rectification, EFR) 네트워크(311)에 입력하여 다수의 이벤트 스택(210)을 수정(rectify)할 수 있다.
- [0063] 이벤트를 스택킹하는 것의 또 다른 단점은 빠른 트리거링 위치에서 이전 이벤트 정보를 덮어버릴 수 있다(overwritten). 덮어 쓰여진 이벤트(overwritten event)는 블러 스택(blurry stack)을 초래하고, 결국 영상 채

구성 품질이 저하될 수 있다.

- [0064] 이벤트 덮어 쓰기를 방지하기 위해, 두 개의 이벤트 스택(210)을 광학 플로우(321)와 연결하고, EFR 네트워크(311)에 입력될 수 있다. EFR 네트워크(311)는 각 이벤트의 세부 정보를 보존하기 위해 이벤트 스택을 통합할 수 있다.
- [0065] 일 실시예에서, EFR 네트워크(311)는 두 개의 컨볼루션 레이어를 포함할 수 있다.
- [0066] EFR 네트워크(311)는 두 개의 이벤트 스택(310)이 광학 플로우(321)와 관련될 수 없는 하나의 이벤트 스택(210)에만 보이는 위치에 이벤트가 있을 때 이미지를 재구성하는데 도움이 될 수 있다. EFR 네트워크(311)에 의해 세 개의 입력을 모두 사용하기 때문에, 이벤트가 강도 재구성을 위해 유지될 가능성이 높다.
- [0067] 이 경우, 중앙 스택은 광학 플로우가 없기 때문에 추정되는 광학 플로우 없이 EFR 네트워크(311)에 입력될 수 있다.
- [0068] 예를 들어, 쌍(pair)을 이룬 이벤트 스택(SBN_{n+m} 및 SBN_n)과 광학 플로우(F_{n+m})의 조합을 수정하여, 수정된 이벤트 스택(RE_{n+m})을 생성할 수 있다.
- [0069] 일 실시예에서, 수정된 다수의 이벤트 스택(322)과 다수의 이벤트 스택에 대한 상태(state) 정보(323)를 고해상도 네트워크(super-resolution network, SRNet)(312)에 입력하여 다수의 중간 강도 이미지(intermediate intensity image)(324)를 생성할 수 있다.
- [0070] SRNet(312)은 수정된 이벤트 스택(RE_{n+m})과 이전 상태 정보($State_n$)를 사용하여, 순차적 모델의 다음 상태 정보($State_{n+m}$)와 중간 강도 출력(intermediate intensity output)(I_{n+m})과 같은 고해상도 강도 이미지를 생성할 수 있다.
- [0071] 일 실시예에서, 초기 이벤트 스택의 경우, 첫 번째 이벤트 스택만 EFR 하위 네트워크에 공급되어 초기 $State_n$ 을 생성할 수 있다.
- [0072] 일 실시예에서, 다수의 중간 강도 이미지(324)를 혼합(mix) 네트워크(313)에 입력하여 강도 이미지(325)를 생성할 수 있다.
- [0073] 혼합 네트워크(313)는 다양한 시간 위치($i=\{n-m, n, n+m\}$)에서 SRNet(312)의 출력(I_i)을 증가시켜 중앙 스택의 타임스탬프(n)에서 상세도가 높은 강도 이미지(O_n)(325)를 재구성할 수 있다. 일 실시예에서, 혼합 네트워크(313)는 컨볼루션 레이어를 사용하여 세부 이미지로 강도 이미지를 재구성할 수 있다.
- [0074] 이벤트 스택(210)은 연속적인 이벤트 스트림을 개별 입력으로 양자화하기 때문에, 각 이벤트 스택(210)은 이미지 재구성에 필요한 모든 세부 정보를 포함하지 않을 수 있다. 따라서, 모든 이벤트 스택(210)으로부터의 중간 강도 출력(324)은 혼합 네트워크(313)에 의해 혼합되어 상세한 세부 정보로 강도 이미지(325)를 재구성할 수 있다.
- [0075] 일 실시예에서, 유사성 손실(similarity loss, Sim) 함수(314)를 이용하여 강도 이미지(325)와 GT(ground truth) 이미지(326)의 차이(difference)에 기반한 오차(error) 정보를 산출할 수 있다.
- [0076] 일 실시예에서, 재구성된 강도 이미지(O)와 GT 이미지(G)가 주어지면, 하기 <수학식 1>과 같이 두 개의 항으로 유사성 손실 함수 $L_{sim}(O, G)$ 를 결정할 수 있다.

수학식 1

[0077]
$$\mathcal{L}_{sim}(O, G) = \mathcal{L}_{\ell_1}(O, G) + \lambda \mathcal{L}_{LPIPS}(O, G)$$

- [0078] 여기서, λ 는 두 항의 밸런싱 파라미터를 나타낸다. 매개 변수를 학습하기 위해 최소화할 수 있다.
- [0079] 먼저, 첫 번째 항($\mathcal{L}_{\ell_1}(O, G)$)은 l_1 norm과 같은 비구조적 손실을 사용하여, 보다 선명한 이미지를 l_2 가 아닌 하기 <수학식 2>와 같이 재구성하여, 출력 이미지에서 낮은 주파수 텍스처로 가장자리를 부드럽게 만들 수 있다.

수학식 2

$$\mathcal{L}_{\ell_1}(O, G) = \|O - G\|_1$$

[0080]

[0081]

두 번째 항($\mathcal{L}_{LPIPS}(O, G)$)의 경우, l_1 은 장면의 구조적 정보를 잃을 수 있기 때문에, LPIPS(Learned Perceptual Image Patch Similarity) 또는 지각적 유사성(perceptual similarity)에 의해 목적 함수의 두 번째 항인 구조적 부족을 보상할 수 있다.

[0082]

구체적으로, 사전 훈련된 네트워크에 의해 인코딩된 한 쌍의 이미지(O, G)가 주어지면, 1번째 레이어의 니어 엔드피처(near end feature)(\hat{O}_{hw}^l)가 추출되고, 그것의 활성화는 채널 차원(H_1, W_1)에 의해 정규화될 수 있다.

[0083]

그런 다음 각 채널은 벡터 w_l 에 의해 스케일링되고 l_2 거리가 계산될 수 있다. 마지막으로, 공간 평균은 하기 <수학식 3>과 같이 LPIPS 손실에 대해 모든 레이어(l)를 통해 이미지 축(h, w)에 대해 계산될 수 있다.

수학식 3

$$\mathcal{L}_{LPIPS}(O, G) = \sum_l \frac{1}{H_l W_l} \sum_{h,w} \|w_l \odot (\hat{O}_{hw}^l - \hat{G}_{hw}^l)\|_2^2$$

[0084]

[0085]

이후, 오차 정보를 이용하여 신경망 모델(300)을 학습시킬 수 있다. 즉, 혼합 네트워크(313)의 출력은 유사도 네트워크(Sim)(314)에 제공되어 오류 정보(Err)를 기반으로 신경망 모델(300)의 매개 변수를 최적화할 수 있다.

[0087]

도 5a 및 5b는 본 발명의 일 실시예에 따른 SRNet(312)을 도시한 도면이다.

[0088]

도 5a 및 5b를 참고하면, SRNet(312)은 제1 내지 제4 네트워크의 조합으로 구성될 수 있다. 예를 들어, 제1 내지 제4 네트워크는 RNet-A(510), RNet-B(520), RNet-C(530) 및 RNet-D(540) 또는 이와 동등한 기술적 의미를 갖는 용어로 지칭될 수 있다.

[0089]

이 경우, 이전 상태(State_n)는 RNet-A(510)로, 이벤트의 입력 RE_{n+1}은 RNet-C(530)로 전달될 수 있다. RNet-C(530)는 입력 이벤트 스택을 처리해야 하므로 다른 3개의 네트워크보다 깊을 수 있다(deep).

[0090]

RNet-A(510)에서 RNet-C(530)의 출력을 빼면 RNet-B(520)는 상태가 아닌 이벤트에서 주로 새로운 세부 정보(e_n)를 전달받을 수 있다.

[0091]

RNet-B(520)의 출력은 다음 상태를 만들기 위해 RNet-C(530)의 출력에 추가될 수 있다. 추가된 결과값은 상태 State_{n+1}을 생성할 수 있고, RNet-D(540)로 전달되어 중간 강도 출력 I_{n+1}을 생성할 수 있다.

[0092]

일 실시예에서, 스택하는 이벤트 스트림의 각 부분이 출력 이미지의 세부 정보를 캡처하고 원래는 연속적이지만 스택킹 방법으로 양자화되기 때문에, SRNet(312)에 대해 반복적인 신경망이 사용될 수 있다.

[0093]

불연속성을 완화하기 위해, 반복되는 신경망의 내부 메모리 상태를 활용하여, 각 수신 스택에 의해 상태가 내부적으로 업데이트됨에 따라, 상세한 세부 정보를 가진 여러 영역을 연속적으로 재구성할 수 있다. 특히, 단일 이벤트 스택은 스택 범위에는 없지만, 이전 스택에서 캡처한 이전에 실행된 이벤트에서 중요한 세부 정보를 부분적으로 놓칠 수 있다.

[0094]

스태킹된 이벤트는 U-net과 같은 심층 신경망에 의해 강도 이미지를 합성할 수 있는 것으로 나타났습니다. 아키텍처적으로, 더 많은 필터와 더 큰 커널 크기로 15 블록 깊이의 ResNet을 사용하여 아이디어를 확장할 수 있다.

[0095]

특히, MISR에서 잘 설계된 네트워크에 따라, 고해상도 강도를 위해 잔존 학습의 힘을 사용할 수 있다. 수정된 이벤트 특징을 SR 강도 생성기(RNet-C)로 전송한다. RNet-C(530)는 전치된(transposed) 컨볼루션 연산의 조합에 의해 초기 SR 강도 이미지 상태를 생성할 수 있다.

[0096]

SRNet(312)은 강도 정보를 추가하면서 입력 RE를 업스케일링하도록 설계될 수 있다.

[0097]

더하여, 2개의 컨볼루션 레이어를 포함하는 5개의 ResNet 블록으로 구성된 3개의 잔여 네트워크(RNet-{A, B,

D))의 조합이 사용될 수 있다. 이러한 네트워크는 수정된 이벤트가 아닌 이전 상태의 특징과 같은 표현을 인코딩하기 때문에, 이 네트워크는 RNet-C보다 얕다(shallower).

[0098] 업샘플링 인코더로 작동하는 RNet-A(510)의 출력을 RNet-C(530)의 출력에서 빼서 내부 오류(internal error)(e_n)를 생성할 수 있다. 이 오류는 하기 <수학식 4>와 같이 현재 수정된 이벤트 스택 RE_{n+m} 이 이전 상태 $State_n$ 과 비교하여 어느 정도 기여하는지 측정하기 위해 사용될 수 있다.

수학식 4

[0099]
$$e_n = RNet-C(RE_{n+m}) - RNet-A(State_n)$$

[0100] 상기 오류는 일반 인코더로 작동하는 RNet-B(520)에 대한 입력으로 제공될 수 있다.

[0101] 일 실시예에서, 하기 <수학식 5>와 같이 RNet-C(530)의 출력과 합산된 RNet-B의 출력으로 다음 상태($State_{n+m}$)가 결정될 수 있다.

수학식 5

[0102]
$$State_{n+m} = RNet-B(e_n) + RNet-C(RE_{n+m})$$

[0103] 일 실시예에서, $State_{n+m}$ 은 최종 디코더(RNet-D)에 제공되어, 하기 <수학식 6>과 같이 중간 강도 출력(I_{n+m})을 생성한다.

수학식 6

[0104]
$$I_{n+m} = RNet-D(State_{n+m})$$

[0105] 일반적으로, RNet-C(530)는 이전 이벤트 스택에 의해 놓쳐진 장면에 대한 세부 정보를 추가하여, 현재 이벤트 스택의 새로운 정보를 이전 상태로 추가할 수 있다.

[0106] 현재 이벤트 스택에 의해 캡처된 일부 영역에 이벤트가 없지만, 이전 이벤트 스택에 의해 캡처된 영역에 장면 세부 정보가 있는 경우에도 이전 상태($State_n$)는 RNet-A(510)를 통해 해당 정보를 숨김 상태(hidden)로 유지하여 오히려 빠진 지역에서 장면 세부 정보를 재구성할 수 있다.

[0108] 도 6a 내지 6c는 본 발명의 일 실시예에 따른 고해상도 강도 이미지 생성의 성능 비교를 도시한 도면이다.

[0109] 도 6a 내지 6c를 참고하면, 정성 분석을 위해, EV에 의한 강도 재구성, EV+MISR 및 실제 방법 및 시뮬레이션된 시퀀스에 대한 본 발명에 따른 방법(Ours)의 성능 비교를 확인할 수 있다.

[0110] 일 실시예에서, EG는 이벤트의 장면 세부정보를 항상 재구성하지는 않으며, 때로는 불완전한 아티팩트(jittery artifacts)을 야기할 수 있다(hallucinate).

[0111] EV는 EG보다 상대적으로 더 나은 이벤트의 장면 세부 정보를 재구성하는 반면, 그림자 같은 아티팩트를 생성하고 장면의 일부 영역을 어렵게할 수 있다.

[0112] 또한, 데이터에 핫 픽셀(hot pixel)이 있는 경우, EV는 핫 픽셀을 필터링하지 않을 수 있다. EV의 결과에 흰색 또는 검은 색 점이 나타나는 반면, 본 발명에 따른 방법은 대부분 명시적인 조작 없이 흰색 또는 검은 색 점을 제거하여 필터링할 수 있다.

[0114] 도 7a 및 7b는 본 발명의 일 실시예에 따른 고해상도 강도 이미지 생성의 성능 그래프를 도시한 도면이다.

[0115] 도 7a 및 7b를 참고하면, 성능 검증을 위해 이벤트 카메라 시뮬레이터(event camera simulator, ESIM)와 4개의 다양한 실제 공개 데이터 세트를 사용하여 생성된 시퀀스가 사용될 수 있다.

- [0116] 정량 분석의 경우, 0(비유사)에서 1(완전 유사) 사이에서의 간격으로서 구조적 유사성(structural similarity, SSIM), PSNR(dB)(로그 스케일)을, 평균 제곱 오차(mean squared error, MSE) 및 지각 유사성(LPIPS)를 두 이미지에서 높은 수준의 기능의 유사성을 평가하기 위한 지표로 사용할 수 있다.
- [0117] 이 경우, 본 발명에 따른 방법(Ours)이 SSIM 값이 높고, MSE 값이 낮으며, LPIPS 값이 낮음을 확인할 수 있으며, 이에 따라, 본 발명에 따른 방법의 성능이 종래의 다른 기술들보다 우수함을 확인할 수 있다.
- [0118] 일 실시예에서, 이벤트에서 직접 재구성 SR 강도 이미지를 최초로 제안하기 때문에 직접비교할 수 있는 방법은 없기 때문에, 먼저 출력을 다운샘플링하고, 동일한 크기의 강도 재구성 방법과 비교하여, 재구성 품질을 평가할 수 있다. 그런 다음, 본 발명에 따른 방법을 최신의 초 고해상도(SR) 방법과 결합된 종래의 강도 재구성 방법과 비교할 수 있다.
- [0119] 일 실시예에서, 고해상도 없는 이미지 재구성의 경우, 본 발명에 따른 방법의 다운 샘플링된 출력을 종래의 기술과 비교할 수 있다. 표기법 간결성을 위해 고역 통과 필터(high pass filter) 방법을 HF로, 매니 폴드 정규화(manifold regularization)은 MR로, 비디오 생성 이벤트(event to video generation)는 EV로, 조건부 GAN에 의한 이벤트 강도(event to intensity by conditional GANs)는 EG 또는 이와 동등한 기술적 의미를 갖는 용어로 지칭될 수 있다.
- [0120] 많은 실제 이벤트 데이터 세트의 평가 프로토콜에 따라 APS 프레임은 GT로 간주될 수 있다.
- [0121] 일 실시예에서, 하기 <표 1> 내지 <표 3>을 참고하면, 본 발명에 따른 방법(Ours)은 LPIPS의 다른 종래의 방법보다 우수함을 확인할 수 있다. 또한, 본 발명에 따라 재구성된 강도 이미지가 종래의 방법보다 지각적으로 더 우수하고, 여러 시퀀스에서 높은 SSIM 점수와 EG와 비슷한 MSE 오류를 나타냄을 확인할 수 있다.
- [0122] EV와 마찬가지로, 합성 시퀀스(synthetic sequence)로만 모델을 학습하고 실제 시퀀스(real world sequence)에 적용할 수 있다. 미세 조정 없이 이 까다로운 제로 샷 데이터 전송설정에서 실제 이벤트에서 다른 방법보다 성능이 우수함을 확인할 수 있다. LPIPS(EV 및 EG)의 두 가지 실행 방법도 학습 기반 프레임 워크를 사용할 수 있다.

표 1

[0124]

Sequence	SSIM (↑)				
	HF	MR	EV	EG	Ours
dynamic_6dof	0.39	0.52	0.46	0.48	0.44
boxes_6dof	0.49	0.45	0.62	0.45	0.61
poster_6dof	0.49	0.54	0.62	0.61	0.63
shapes_6dof	0.50	0.51	0.80	0.56	0.79
office_6dof	0.38	0.45	0.54	0.67	0.68
slider_6dof	0.50	0.50	0.58	0.54	0.59
calibration	0.48	0.54	0.70	0.67	0.71
Average	0.46	0.50	0.62	0.57	0.64

표 2

[0126]

Sequence	MSE (↓)				
	HF	MR	EV	EG	Ours
dynamic_6dof	0.10	0.05	0.14	0.03	0.05
boxes_6dof	0.08	0.10	0.04	0.03	0.02
poster_6dof	0.07	0.05	0.06	0.01	0.02
shapes_6dof	0.09	0.19	0.04	0.03	0.01
office_6dof	0.09	0.09	0.03	0.01	0.01
slider_6dof	0.06	0.07	0.05	0.02	0.02
calibration	0.09	0.07	0.02	0.01	0.01
Average	0.08	0.09	0.05	0.02	0.02

표 3

[0128]

Sequence	LPIPS (↓)				
	HF	MR	EV	EG	Ours
dynamic_6dof	0.54	0.50	0.46	0.45	0.42
boxes_6dof	0.50	0.53	0.38	0.48	0.32
poster_6dof	0.45	0.52	0.35	0.42	0.29
shapes_6dof	0.61	0.64	0.47	0.51	0.38
office_6dof	0.54	0.50	0.41	0.36	0.29
slider_6dof	0.50	0.55	0.44	0.42	0.34
calibration	0.48	0.47	0.36	0.42	0.24
Average	0.52	0.53	0.41	0.43	0.33

[0130]

고해상도 이미지 재구성의 경우, 강도 재구성 알고리즘에 대한 종래의 이벤트를 종래의 SR 방법과 결합하고, 본 발명에 따른 방법(Ours)과 비교할 수 있다. 강도 알고리즘에 대한 종래의 이벤트에 대하여, EV는 SSIM 및 LPIPS의 EG를 대부분의 시퀀스에서 평균적으로 성능을 증가하여 수행하는 학습 방식이기 때문에, EV가 사용될 수 있다.

[0131]

도 7b를 참고하면, 고해상도 알고리즘의 경우, 하나는 SISR(single image SR, SISR)과 다른 하나는 MISR(multiple image SR)의 두 가지 고해상도 알고리즘이 사용될 수 있다. 이 경우, 하기 <표 4>를 참고하면, 본 발명에 따른 방법(Ours)은 큰 마진에 의한 모든 메트릭(metric)에서 SR 알고리즘과 결합된 종래의 강도 재구성 알고리즘보다 성능이 우수함을 확인할 수 있다. ESIM에서 생성한 데이터 세트에서 30개의 시퀀스가 사용될 수 있다.

표 4

[0132]

Method	PSNR (↑)	SSIM (↑)	MSE (↓)	LPIPS (↓)
EV+SISR 2x	11.292	0.384	0.348	0.394
EV+MISR 2x	11.309	0.358	0.347	0.392
Ours	16.420	0.600	0.108	0.172
EV+SISR 4x	11.168	0.396	0.089	0.543
EV+MISR 4x	11.293	0.384	0.087	0.396
Ours	16.068	0.560	0.028	0.253

[0134]

도 8a 및 8b는 본 발명의 일 실시예에 따른 다운스케일 출력의 성능 비교를 도시한 도면이다.

[0135]

도 8a 및 8b를 참고하면, 또 다른 데이터 세트의 시퀀스에 대한 실험을 수행하고, 본 발명에 따른 방법을 종래의 EG 및 EV와 질적으로 비교할 수 있다. 이 경우, 본 발명에 따른 방법은 손가락 끝이나 질감과 같은 동일한 크기의 이미지를 구성할 때 보이지 않는 세부 정보를 나타낼 수 있다.

[0137]

도 9는 본 발명의 일 실시예에 따른 손실 함수에 대한 성능 비교를 도시한 도면이다.

[0138]

도 9를 참고하면, 손실 항목(Loss Term) 분석(Lsim)의 경우, 손실 함수를 제거(ablate)하여, 하기 <표 5>와 같이 이미지 재구성에 대한 각 항목(term)의 영향을 확인할 수 있다.

표 5

[0139]

Loss	PSNR ^(†)	SSIM ^(†)	MSE ^(‡)	LPIPS ^(‡)
\mathcal{L}_i	15.33	0.517	0.034	0.485
\mathcal{L}_{LPIPS}	10.06	0.388	0.454	0.232
\mathcal{L}_{sim} (full)	15.03	0.528	0.032	0.258

[0140]

모든 분석 및 절제(ablation) 연구는 고품질 GT를 사용한 신뢰할 수 있는 정량 분석을 위해 시뮬레이션된 데이터로 수행될 수 있다. L11 항(term)만 사용하면, PSNR에서 더 나은 성능을 볼 수 있지만, 시각적으로 선명하지

많은 이미지를 만들어, 다른 모든 메트릭에서 성능이 저하됨을 확인할 수 있다.

[0141] LPIPS 향만 사용하면, 이미지가 시각적으로 수용 가능하지만, 이벤트가 적은 영역과 가장자리에 점과 같은 아티팩트가 있는 낮은 PSNR의 단점이 있다. 11과 LPIPS의 결합인 손실 함수 L_{sim} 은 PSNR 및 LPIPS가 약간 감소하여 SSIM 및 MSE에서 최고 성능을 발휘하지만 시각적으로 가장 우수한 이미지를 만들 수 있다.

[0143] 고해상도 파라미터 분석의 경우, 두 가지 SR 매개 변수의 영향을 평가할 수 있다. 예를 들어, 두 가지 SR 매개 변수는 출력 품질에 대한 업스케일 팩터(2x, 4x) 및 스택 시퀀스(3S, 7S)의 크기를 포함할 수 있다. 평가 결과는 하기 <표 6>과 같이 나타낼 수 있다.

표 6

Scale	#Stacks	PSNR ⁽¹⁾	SSIM ⁽²⁾	MSE ⁽⁴⁾	LPIPS ⁽⁴⁾
2x	3S	15.46	0.554	0.323	0.191
	7S	16.42	0.600	0.108	0.172
4x	3S	15.03	0.528	0.032	0.258
	7S	16.06	0.560	0.028	0.253

[0145] 3S와 7S를 비교하면, 7S가 모든 메트릭에서 더 나은 성능을 제공한다는 것을 알 수 있다. 이는 시퀀스에 대한 재귀(recursion)가 깊수록 숨김 상태(hidden state)가 보다 안정적으로 생성되어 더 나은 품질의 결과를 얻을 수 있음을 의미할 수 있다. 또한 더 긴 시퀀스를 사용하는 경우, 더 큰 재귀에 대한 롤링을 해제하면(unrolling), 짧은 이벤트의 정보를 유지하는데 도움이 되므로, 짧은 기간 동안만 발생하는 이벤트를 캡처할 가능성이 높을 수 있다.

[0146] 알고리즘이 이벤트가 없는 넓은 공간 위치를 처리하는 것은 쉽지 않기 때문에, 이벤트를 더 큰 이미지로 고해상도화 하는 것이 더 어려울 수 있다. MSE는 2x에 비해 감소했지만, 이는 분모의 숫자가 이미지의 크기로 인해 더 크고 출력 품질과 관련이 많지 않기 때문일 수 있다.

[0148] 도 10은 본 발명의 일 실시예에 따른 HDR 시나리오에서의 이미지 재구성 비교를 도시한 도면이다.

[0149] 도 10을 참고하면, HDR 시퀀스에 대한 정성 분석의 경우, 이벤트 카메라를 사용하는 한 가지 어려운 시나리오는 극도의 동적 범위에서 이벤트를 캡처하는 것일 수 있다.

[0150] 이러한 극한 조건에서 출력을 정성적으로 분석하고 이를 EV와 비교할 수 있다. APS 프레임을 포함한 일반 카메라는 동적 범위가 훨씬 낮고 검은색 영역(카메라가 상단 열에 표시된 감지 범위에서 강도 세부 정보를 감지하지 못하는 경우) 또는 흰색 영역(카메라에 빛이 쏟아져 카메라가 하단 열에 표시된 감지 범위보다 높을 수 없는 경우)을 생성할 수 있다.

[0151] 본 발명에 따른 방법이 더 높은 범위를 다루고 EV와 APS 프레임이 포착하지 못하는 더 많은 구조적 세부 정보를 나타낼 수 있음을 확인할 수 있다.

[0153] 일 실시예에서, 실패 모드(Failure Mode) 분석의 경우, 실패 사례는 대부분 전경 개체(foreground object)가 빠르게 움직일 때 긴 궤적에 대한 배경 세부정보(background detail)가 누락된 것과 관련이 있다. 이러한 시퀀스에서, 본 발명에 따른 방법은 중앙 스택까지 제한된 시간 거리에 있는 장면의 일부만 복구할 수 있다.

[0155] 도 11a는 본 발명의 다른 실시예에 따른 고해상도 강도 이미지 생성을 도시한 도면이다.

[0156] 도 11a를 참고하면, 다수의 이벤트 스택(210)의 시퀀스 중 중앙 스택(central stack)을 APS(active pixel sensor) 프레임으로 형성할 수 있다.

[0157] 중앙 스택으로 형성된 APS 프레임과 다수의 이벤트 스택(210)의 시퀀스 중 중앙 스택을 제외한 다수의 주변 스택(nearby stack)을 신경망 모델(300)에 입력하여 강도 이미지(325)를 생성할 수 있다.

[0158] 즉, APS 프레임을 보완(complementary)으로 사용하는 경우, APS 프레임을 사용하여 신경망 모델(300)의 초기 상태를 중앙 스택으로 훈련하고, 이벤트 스택을 주변 스택(nearby stack)으로 제공할 수 있다. 신경망 모델은 APS 프레임을 포함하는 입력으로부터 더 높은 해상도의 세부 정보를 추가하는 것을 학습할 수 있다.

[0160] 도 11b는 본 발명의 또 다른 실시예에 따른 고해상도 강도 이미지 생성을 도시한 도면이다.

[0161] 도 11b를 참고하면, 본 발명의 일 실시예에 따른 고해상도 강도 이미지 생성 방법에 따라 강도 이미지(325)를

생성한 후, 다수의 이벤트 스택(210)의 시퀀스 중 중앙 스택을 상기 생성된 강도 이미지(325)로 형성할 수 있다.

- [0162] 중앙 스택으로 형성된 강도 이미지(325)와 다수의 이벤트 스택(210)의 시퀀스 중 중앙 스택을 제외한 다수의 주변 스택을 신경망 모델에 입력하여 다른 강도 이미지를 생성할 수 있다.
- [0163] 즉, APS 프레임 사용하지 않지만, 신경망 모델(300)을 두 번의 반복하는 듀오 패스(Duo-Pass)를 사용하는 경우, 첫 번째 패스에서는 신경망 모델(300)의 기본 체계를 사용하여 이벤트 스택을 이용하여 강도 이미지를 생성할 수 있다.
- [0164] 두 번째 패스에서는, 중앙 스택으로 첫 번째 패스의 합성된 강도 이미지를 사용할 수 있다. 듀오 패스를 통해 APS 프레임의 도움 없이 첫 번째 패스에서 누락된 HR 세부 정보를 추가로 복구할 수 있다.
- [0166] 도 12a 및 12b는 본 발명의 다양한 실시예에 따른 고해상도 강도 이미지 생성의 성능 비교를 도시한 도면이다.
- [0167] 도 12a 및 12b를 참고하면, GT로 간주되는 APS 프레임과 함께 원래의 보완(complementary) 및 듀오 패스(Duo-pass)의 샘플 출력을 확인할 수 있다.
- [0168] 듀오 패스 방식은 원본(original)과 비교하여 더 많은 세부 사항을 추가할 수 있다. 보완 방식은 훨씬 더 많은 세부 사항을 추가하고, APS 프레임 방식보다 훨씬 더 많은 디테일로 이미지를 재구성 할 수 있다.
- [0170] 도 13은 본 발명의 일 실시예에 따른 고해상도 강도 이미지 생성 방법을 도시한 도면이다.
- [0171] 도 13을 참고하면, S1301 단계는, 이벤트 이미지(110)를 획득하는 단계이다.
- [0172] S1303 단계는, 이벤트 이미지(110)에 대한 다수의 이벤트 스택(210)을 생성하는 단계이다.
- [0173] 일 실시예에서, 이벤트 이미지(110)의 각 픽셀에 대한 이벤트 스트림(120)을 생성하고, 이벤트 스트림(120)을 분류하여 다수의 이벤트 스택(210)을 생성할 수 있다.
- [0174] 일 실시예에서, 이벤트 스트림(120)은 이벤트 스트림(120)에 포함된 이벤트의 수(number)에 따라 분류되어, 다수의 이벤트 스택(210)이 생성될 수 있다.
- [0175] S1305 단계는, 다수의 이벤트 스택(210)을 신경망 모델(300)에 입력하여 강도 이미지(325)를 생성하는 단계이다.
- [0176] 일 실시예에서, 다수의 이벤트 스택(210)의 시퀀스(sequence)를 이용하여 다수의 광학 플로우(optical flow)(321)를 생성할 수 있다.
- [0177] 일 실시예에서, 다수의 광학 플로우(321)를 이용하여 다수의 이벤트 스택(210)을 수정(rectify)할 수 있다.
- [0178] 일 실시예에서, 수정된 다수의 이벤트 스택(322)과 다수의 이벤트 스택에 대한 상태(state) 정보(323)를 이용하여 다수의 중간 강도 이미지(intermediate intensity image)(324)를 생성할 수 있다.
- [0179] 일 실시예에서, 다수의 중간 강도 이미지(324)를 혼합(mix)하여 강도 이미지(325)를 생성할 수 있다.
- [0180] 일 실시예에서, 강도 이미지(325)와 GT(ground truth) 이미지(326)의 차이(difference)에 기반한 오차(error) 정보를 산출하고, 오차 정보를 이용하여 신경망 모델(300)을 학습시킬 수 있다.
- [0182] 도 14는 본 발명의 일 실시예에 따른 고해상도 강도 이미지 생성 장치(700)의 기능적 구성을 도시한 도면이다.
- [0183] 도 14를 참고하면, 고해상도 강도 이미지 생성 장치(1400)는 획득부(1410), 제어부(1420), 표시부(1430) 및 저장부(1440)를 포함할 수 있다.
- [0184] 획득부(1410)는 이벤트 이미지(110)를 획득할 수 있다. 일 실시예에서, 획득부(1410)는 이벤트 카메라로 구현될 수 있다. 일 실시예에서, 획득부(1410)는 유선 통신 모듈 및 무선 통신 모듈 중 적어도 하나를 포함할 수 있다. 획득부(1410)의 전부 또는 일부는 '송신부', '수신부' 또는 '송수신부(transceiver)'로 지칭될 수 있다.
- [0185] 제어부(1420)는 이벤트 이미지(110)에 대한 다수의 이벤트 스택(210)을 생성하고, 다수의 이벤트 스택(210)을 신경망 모델(300)에 입력하여 강도 이미지(325)를 생성할 수 있다.
- [0186] 일 실시예에서, 제어부(1420)는 적어도 하나의 프로세서 또는 마이크로(micro) 프로세서를 포함하거나, 또는, 프로세서의 일부일 수 있다. 또한, 제어부(1420)는 CP(communication processor)라 지칭될 수 있다. 제어부

(1420)는 본 발명의 다양한 실시예에 따른 고해상도 강도 이미지 생성 장치(1400)의 동작을 제어할 수 있다.

- [0187] 표시부(1430)는 이벤트 이미지(110)를 디스플레이할 수 있다. 일 실시예에서, 표시부(1430)는 강도 이미지(325)를 디스플레이할 수 있다.
- [0188] 일 실시예에서, 표시부(1430)는 고해상도 강도 이미지 생성 장치(1400)에서 처리되는 정보를 나타낼 수 있다. 예를 들면, 표시부(1430)는 액정 디스플레이(LCD; Liquid Crystal Display), 발광 다이오드(LED; Light Emitting Diode) 디스플레이, 유기 발광 다이오드(OLED; Organic LED) 디스플레이, 마이크로 전자기계 시스템(MEMS; Micro Electro Mechanical Systems) 디스플레이 및 전자 종이(electronic paper) 디스플레이 중 적어도 어느 하나를 포함할 수 있다.
- [0189] 저장부(1440)는 이벤트 이미지(110), 다수의 이벤트 스택(210), 강도 이미지(325) 중 적어도 하나를 저장할 수 있다. 일 실시예에서, 저장부(1440)는 신경망 모델(300)을 저장할 수 있다.
- [0190] 일 실시예에서, 저장부(1440)는 휘발성 메모리, 비휘발성 메모리 또는 휘발성 메모리와 비휘발성 메모리의 조합으로 구성될 수 있다. 그리고, 저장부(1440)는 제어부(1420)의 요청에 따라 저장된 데이터를 제공할 수 있다.
- [0191] 도 14를 참고하면, 고해상도 강도 이미지 생성 장치(1400)는 획득부(1410), 제어부(1420), 표시부(1430) 및 저장부(1440)를 포함할 수 있다. 본 발명의 다양한 실시예들에서 고해상도 강도 이미지 생성 장치(1400)는 도 14에 설명된 구성들이 필수적인 것은 아니어서, 도 14에 설명된 구성들보다 많은 구성들을 가지거나, 또는 그보다 적은 구성들을 가지는 것으로 구현될 수 있다.
- [0193] 일 실시예에서, 비디오 재구성(Video reconstruction)의 경우, 본 발명에 따른 방법을 비디오 재구성으로 확장하기 위해, 블라인드 후 처리 방법(blind post-processing method)을 사용하여 강도 이미지들 사이의 시간적 일관성(temporal consistency)을 인코딩할 수 있다.
- [0194] 시간적 일관성을 정량적으로 평가하기 위해, 우리는 하기 <수학식 7>과 같은 두 개의 연속 합성된 프레임들(F_t , F_{t+1}) 사이의 플로우 왜곡 에러(flow warping error)에 기초한 시간적 안정성 메트릭을 산출할 수 있다.

수학식 7

[0195]
$$E_{warp}(F_t, F_{t+1}) = \frac{1}{\sum_{i=1}^N M_t^{(i)}} \sum_{i=1}^N M_t^{(i)} \|F_t^{(i)} - \hat{F}_{t+1}^{(i)}\|_2^2$$

- [0196] 여기서 \hat{F}_{t+1} 은 F_{t+1} 의 왜곡 프레임(warped frame)이고, $M_t \in \{0, 1\}$ 은 계산이 비 폐쇄 영역(non-occlusion region)에만 적용되도록 하기 위한 비 폐쇄 마스크(non-occlusion mask)를 나타낸다.
- [0197] 비교된 모든 방법과 APS가 GT인 것처럼 APS 프레임의 왜곡 오차를 평가하기 위해 APS 프레임을 기반으로 프레임 및 비 폐쇄 맵을 워핑(warping)하는데 사용되는 광학 플로우를 계산할 수 있다.
- [0198] 일 실시예에서, 하기 <표 7>의 EV와 비교하여 다른 크기의 시퀀스(3S와 7S)로 결과를 확인할 수 있다. 본 발명에 따른 후 처리 방식(3S+와 7S+)은 EV와 후 처리 버전(EV+)을 모두 크게 증가하여 성능이 향상됨을 확인할 수 있다.

표 7

[0199]

Ewarp ⁽⁴⁾	APS	3S	7S	EV	3S+	7S+	EV+
dynamic_6dof	0.61	20.35	16.54	8.78	3.42	3.71	5.56
boxes_6dof	1.81	16.69	17.51	15.69	3.58	3.95	9.36
poster_6dof	1.10	18.80	22.66	17.74	4.41	5.91	5.56
shapes_6dof	0.44	24.00	21.23	16.66	2.80	2.63	8.33
office_6dof	0.08	3.62	2.19	0.72	0.36	0.34	0.44
slider_6dof	0.02	0.57	0.34	0.19	0.06	0.04	0.12
calibration	0.36	15.46	9.72	2.99	1.31	1.24	1.62
Average	0.63	14.21	12.89	8.97	2.28	2.55	5.20

- [0201] 본 발명의 일 실시예에 따르면, 엔드-투-엔드 신경망(end-to-end neural network)에 의해 이벤트의 고해상도 강도 이미지를 직접 재구성할 수 있다.
- [0202] 또한, 본 발명의 일 실시예에 따르면, 동일한 크기의 이미지 재구성과 고해상도에서 종래의 기술과 비교하여 세밀한 세부 정보로 고품질 이미지를 재구성할 수 있다.
- [0203] 또한, 본 발명의 일 실시예에 따르면, 누락된 세부 정보를 추가하기 위해 여분의 패스를 수행하고 이벤트 외에도 APS 프레임을 사용하는 보완을 수행하는 듀오 패스 방식으로 확장할 수 있다.
- [0204] 또한, 본 발명의 일 실시예에 따르면, 후 처리를 통해 본 발명에 따른 방법을 비디오로 재구성하여 시간적 일관성을 보장할 수 있다.
- [0206] 이상의 설명은 본 발명의 기술적 사상을 예시적으로 설명한 것에 불과한 것으로, 통상의 기술자라면 본 발명의 본질적인 특징이 벗어나지 않는 범위에서 다양한 변경 및 수정이 가능할 것이다.
- [0207] 따라서, 본 명세서에 개시된 실시예들은 본 발명의 기술적 사상을 한정하기 위한 것이 아니라, 설명하기 위한 것이고, 이러한 실시예들에 의하여 본 발명의 범위가 한정되는 것은 아니다.
- [0208] 본 발명의 보호범위는 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 발명의 권리범위에 포함되는 것으로 이해되어야 한다.

부호의 설명

- [0210] 110: 이벤트 이미지
- 120: 이벤트 스트림
- 210: 이벤트 스택
- 300: 신경망 모델
- 310: 플로우 네트워크
- 311: 이벤트 기능 수정 네트워크
- 312: 고해상도 네트워크
- 313: 혼합 네트워크
- 314: 유사성 손실 함수
- 321: 광학 플로우
- 322: 수정된 다수의 이벤트 스택
- 323: 상태 정보
- 324: 중간 강도 이미지
- 325: 강도 이미지
- 326: GT 이미지
- 510: RNet-A
- 520: RNet-B
- 530: RNet-C
- 540: RNet-D
- 1410: 획득부
- 1420: 제어부
- 1430: 표시부

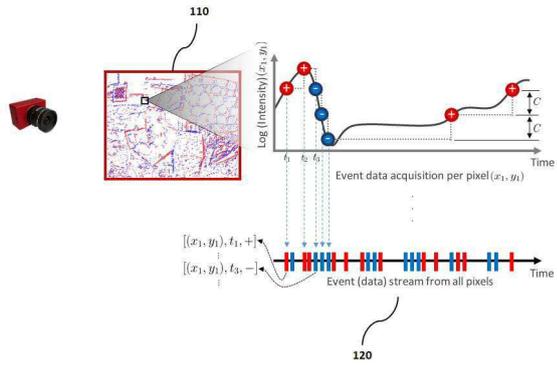
1440: 저장부

도면

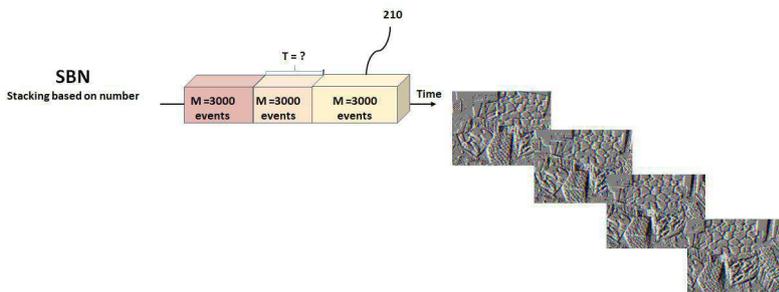
도면1a



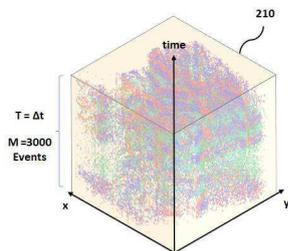
도면1b



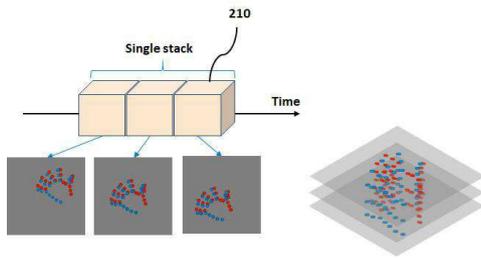
도면2a



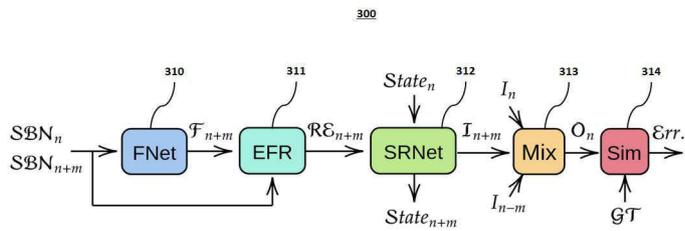
도면2b



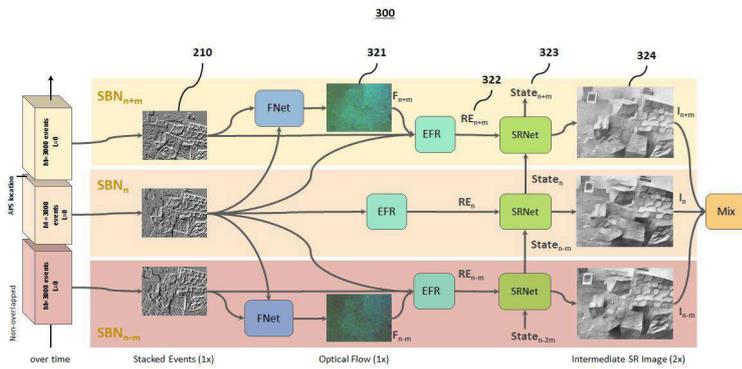
도면2c



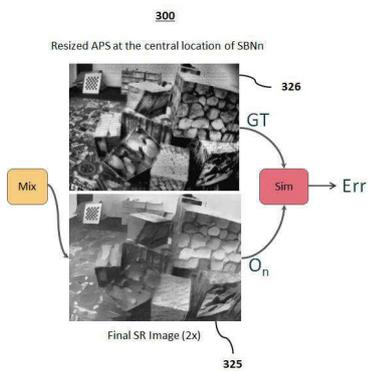
도면3a



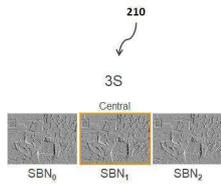
도면3b



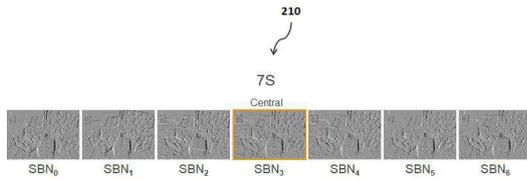
도면3c



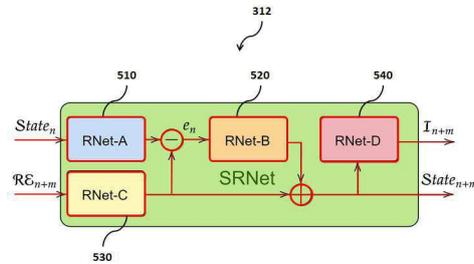
도면4a



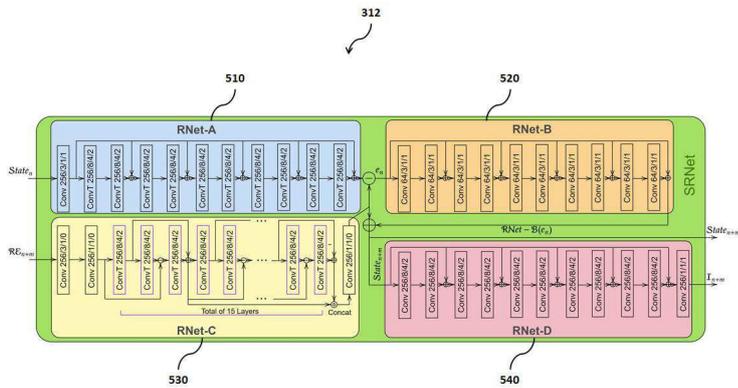
도면4b



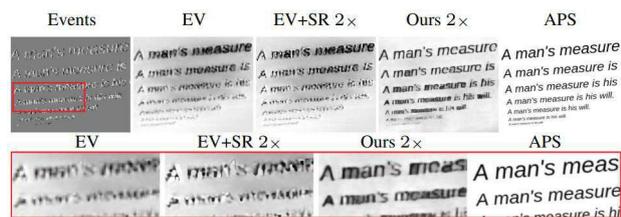
도면5a



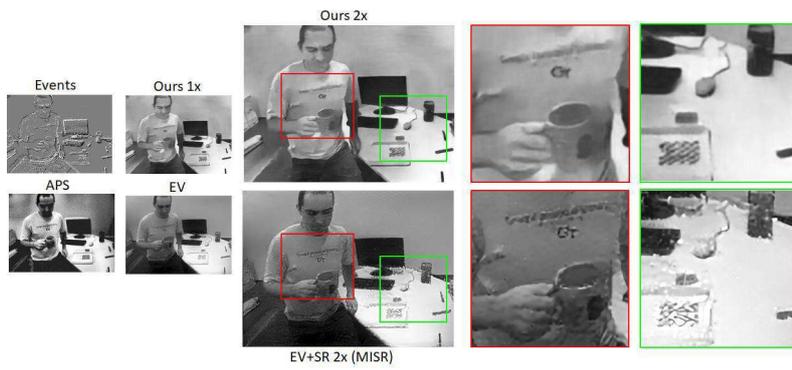
도면5b



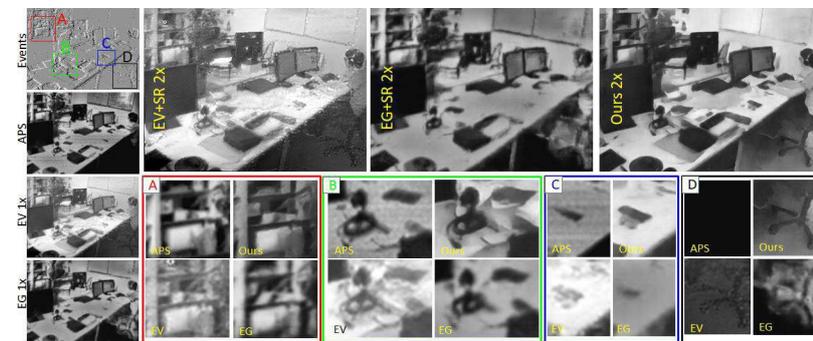
도면6a



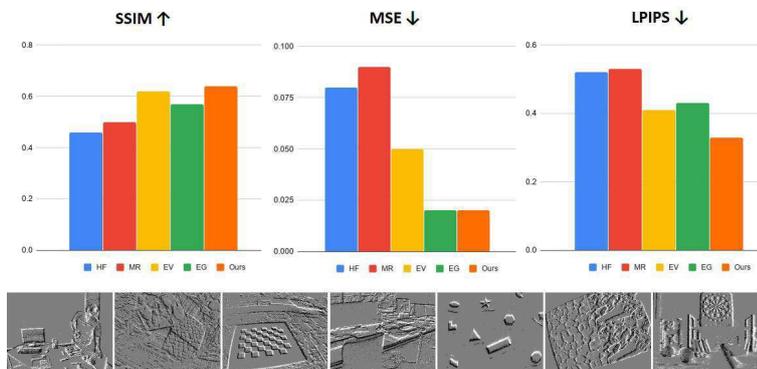
도면6b



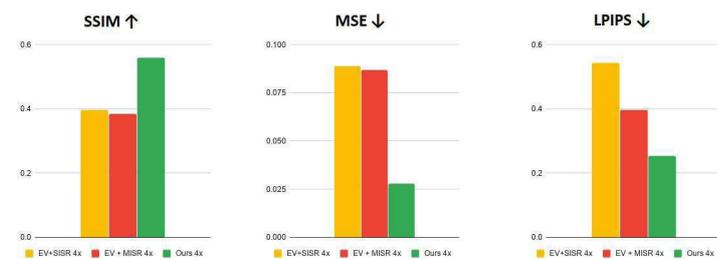
도면6c



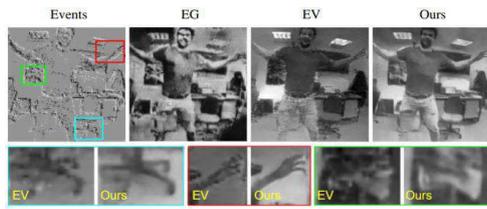
도면7a



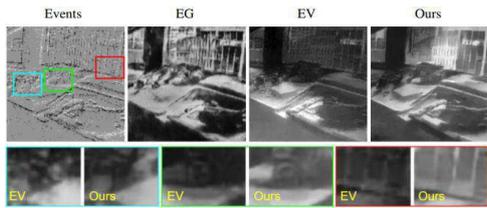
도면7b



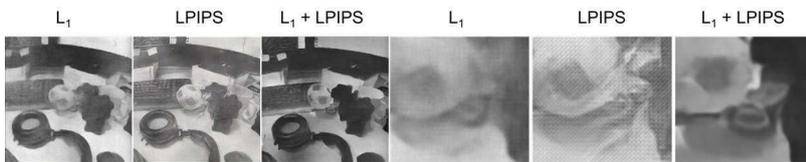
도면8a



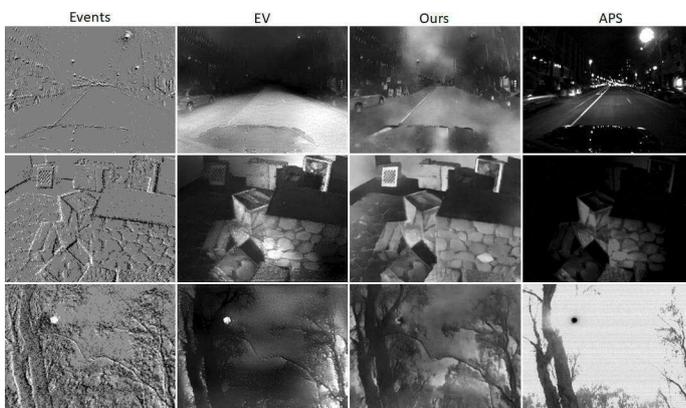
도면8b



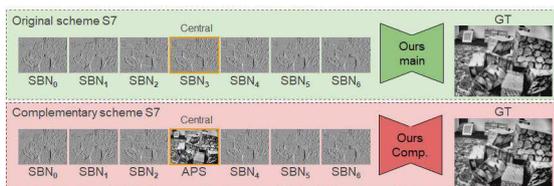
도면9



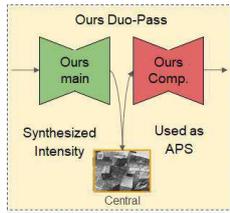
도면10



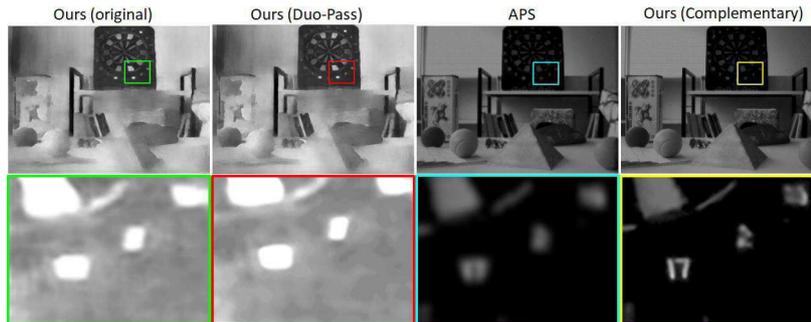
도면11a



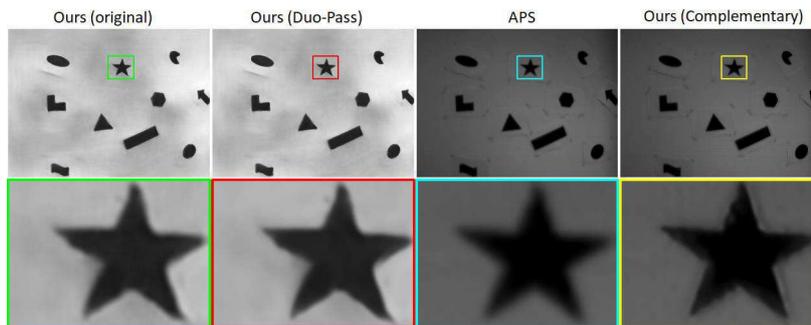
도면11b



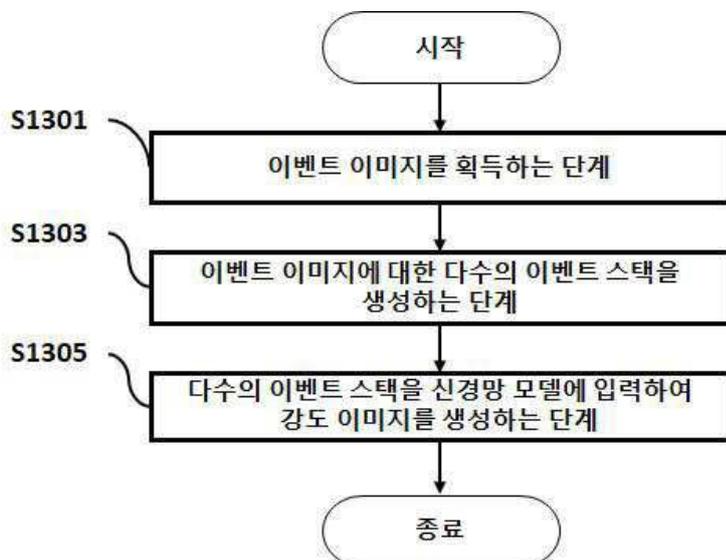
도면12a



도면12b



도면13



도면14

