

PORTFOLIO

Hyun-Woo Kim

**Department of Artificial Intelligence,
Korea University**

[Curriculum vitae](#)

[Portfolio page](#)

정보



김현우

고려대학교 인공지능학 석사 졸업 - 23.02
GPA: 4.13/4.5 (96.3/100) (지도교수: 이성환)
한성대학교 IT융합공학 학사 졸업 - 21.02
GPA: 3.57/4.5 (90.7/100) (지도교수: 오희석)

연구분야

- Computer Vision
- Multi-view geometry
- 2D/3D Human Pose Estimation
- Temporal Action Localization

어학

Opic: IH

기타

고려대학교 산학협력단 계약직
부서: 인공지능연구센터
기간: 2023.06.16~2023.12.29

논문



SCIE

First author

1. Cross-View Self-Fusion for Self-Supervised 3D Human Pose Estimation in the Wild, *16th Asian Conference on Computer Vision*, **Published**, 2022.12.04, (I.S. 5.7)
2. MHCanonNet: Multi-Hypothesis Canonical Lifting Network for Self-supervised 3D Human Pose Estimation in the wild Video, *Pattern Recognition*, **Published**, 2024.01.01, (I.S. 19.7)

Second/third author

1. EGPose: Explicit and Geometric Self-Supervision for 3D human Pose Estimation in the wild Video, *Procedia Computer Science*, **Published**, 2023.06.18
2. Masked Kinematic Continuity-aware Hierarchical Attention Network for Pose Estimation in Videos, *Neural Networks*, **Published**, 2024.01.01 (I.S. 11.1)
3. Calibrated Attention Masking Network for Temporal Action Localization, *Pattern Recognition*, **Under review**

Projects



기 간: 2021.05~2021.10
발주처: (주)브이씨
근무처: 고려대학교 패턴인식 및 머신러닝 연구실
프로젝트명: 골프 트레이닝을 위한 인공지능 기반 골프 스윙 분석 알고리즘 개발

기 간: 2020.07~2020.11
발주처: 한국전자통신연구원(ETRI)
근무처: 한성대학교 Visual Intelligence 연구실
프로젝트명: 중대형 공간용 초고해상도 비정형 플렌옵틱 동영상 저작/재생 플랫폼 기술 개발

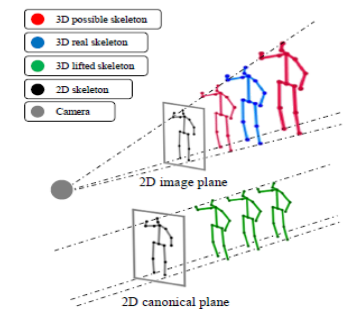
Awards



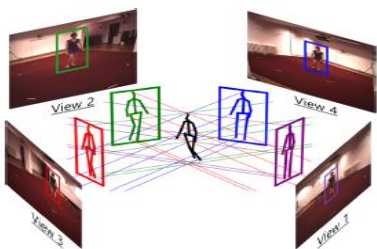
수상일: 2021.12.17
대회명: 미소 인공지능 모델 개발 챌린지, 대상
주관/주최: 과학기술정보통신부, 한국지능정보사회진흥원, (주)미소정보기술

수상일: 2020.11.30
대회명: 2020 공개SW 개발자대회, 동상
주관/주최: 과학기술정보통신부, 정보통신산업진흥원, (주)엘에스웨어

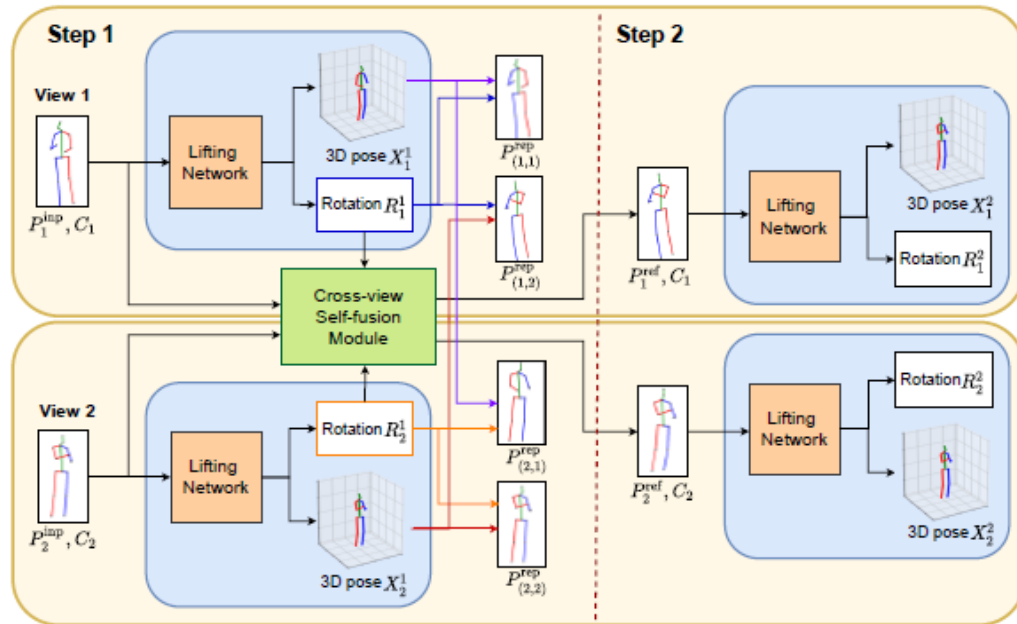
Cross-View Self-Fusion for Self-Supervised 3D Human Pose Estimation in the Wild



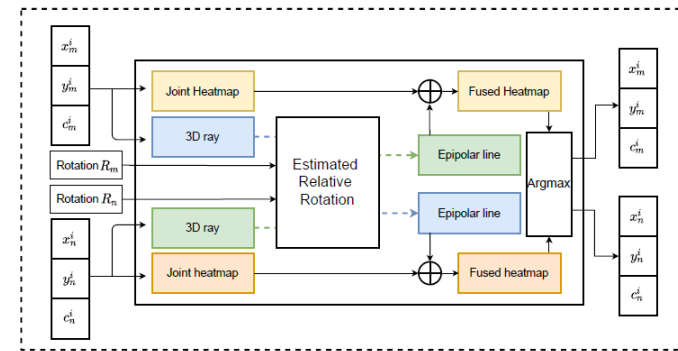
Compare of general and canonical camera



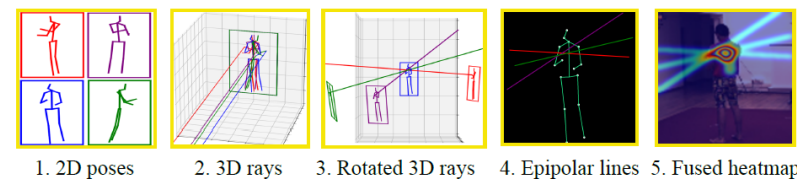
Visualization of Canonical space



Overview of framework



Cross-view self-fusion module



Process of cross-view self-fusion

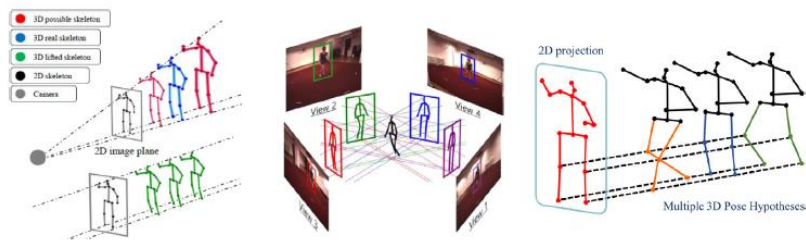
Introduction

- 기존 3D human pose estimation에서 지배적인 Fully-supervision 접근법은 실험 환경에서 모션캡처를 통해 구축된 많은 양의 3D annotation을 필요로 한다.
- 실제 다양한 인간 활동은 실험 환경 밖에서 일어나며 모션캡처 슈트를 입지 않는다.
- 3D annotation을 요구하지 않는 방법들은 Multi-view 정보를 사용하며 이를 위해 정확한 2D keypoint와 Camera calibration을 요구한다
- 우리는 이런 문제들을 해결하기 위해 어떠한 Annotation도 요구하지 않으며 multi-view를 사용하기 위해 Camera calibration을 필요로 하지 않는다.

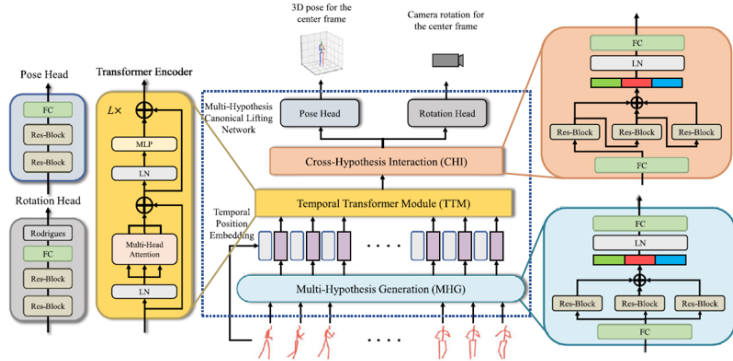
Contribution

- 우리는 root joint zero centering과 Frobenius norm을 통해 원근법 문제를 해결하며 canonical form을 이용하여 상대적 카메라 회전을 추정할 수 있다.
- 우리는 어떠한 카메라 파라메타 없이 multi-view 정보를 이용하여 부정확하게 측정된 입력 2D pose를 개선하는 Cross-view self-fusion 모듈을 제안한다
- 개선된 입력 정보를 통해 멀티뷰 일관성을 더욱 강화하고 이는 다시 정확한 추정을 이끌며 정확한 추정은 다시 더욱 정확한 입력 값 개선을 이끄는 선순환을 이루어 낸다

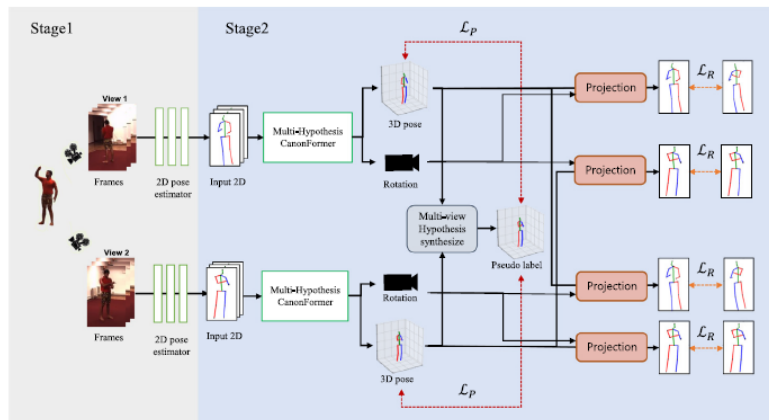
MHCanonNet: Multi-Hypothesis Canonical Lifting Network for Self-supervised 3D Human Pose Estimation in the wild Video



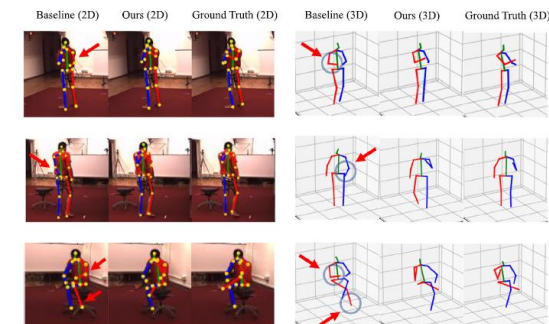
Visualization of Canonical space and Hypotheses



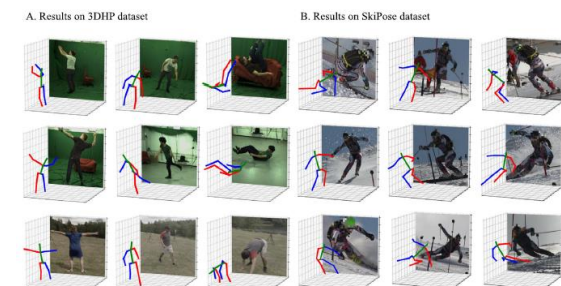
Architecture of our proposed network



Overview of our self-supervised training framework



Comparison results



Qualitative results of our approach

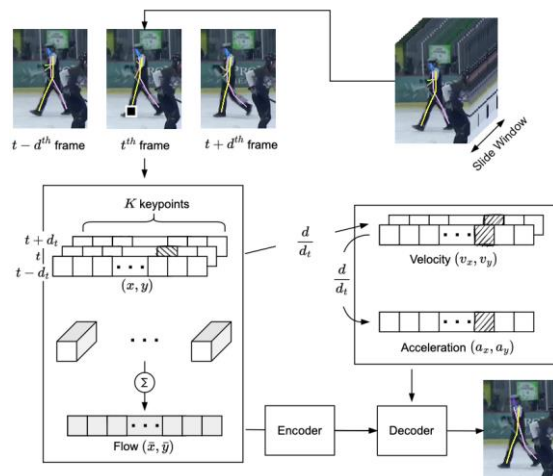
Introduction

- 영상에서 Temporal 정보를 이용하는 Transformer 기법이 3D human pose estimation 분야에 적용됨으로써 뛰어난 성능을 보여주었다.
- 기존 3D human pose estimation에서 지배적인 Fully-supervision 접근법은 실험 환경에서 모션캡처를 통해 구축된 많은 양의 3D annotation을 필요로 한다.
- 3D annotation을 요구하지 않는 방법들은 Multi-view 정보를 사용하며 이를 위해 정확한 2D keypoint와 Camera calibration을 요구한다
- 우리는 이런 문제들을 해결하기 위해 어떠한 Annotation도 요구하지 않으며 multi-view를 사용하기 위해 Camera calibration을 필요로 하지 않는다.

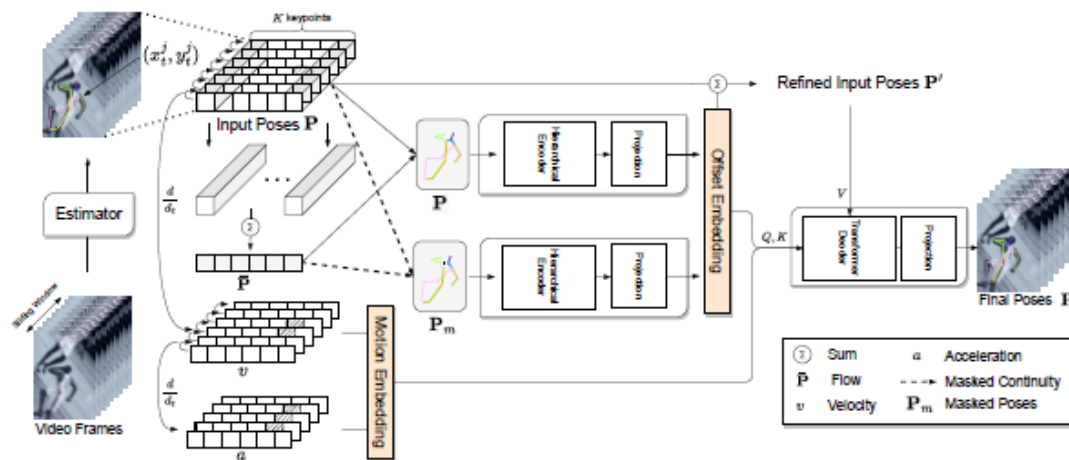
Contribution

- 우리는 어떠한 카메라 파라메타 없이 multi-view 정보를 이용하여 multi-view 일관성 제약을 만족시킴으로써 3D annotation 없이 self-supervised training을 이루어 낸다.
- 대량의 3D annotation을 요구하는 Transformer를 이용하기 위해 네트워크가 추정된 값을 Pseudo label로 취급하여 문제를 해결한다.
- 부정확하게 측정된 입력 2D pose를 고려하여 추정값에 대한 multiple hypotheses의 temporal representations를 학습하는 MHCanonNet을 제안한다.

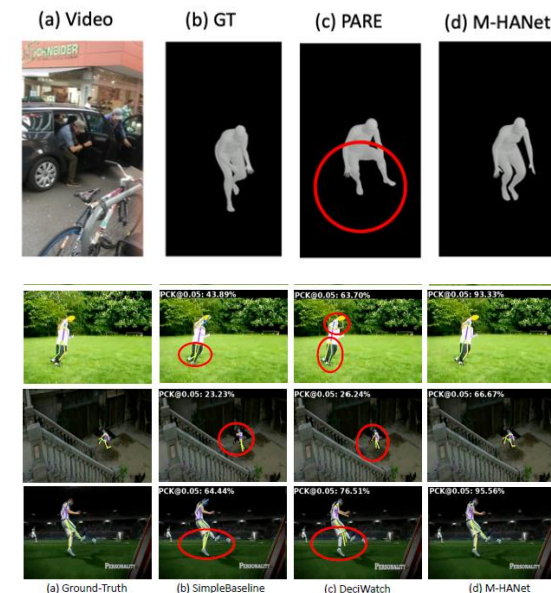
Masked Kinematic Continuity Hierarchical Attention Network for Pose Estimation in Videos



Overview of M-HANet with masked kinematic keypoint features



The overall architecture of M-HANet



Comparative analysis and Qualitative results

Introduction

- Occlusion은 연속된 프레임 간의 불확실성을 초래하여 부드럽지 않은 결과를 만들어 낸다
- 기존 방법들은 Occlusion과 Jitter를 분리하여 다루지만 이 두 가지 문제는 서로 관련이 있다.
- 기존 Transformer를 기반으로 한 refinement 방법들은 시간 정보를 활용하는데 비해 공간적 이미지 특성을 활용하지 못하는 단점이 있다.

Contribution

- Masked kinematic correlation의 kinematic keypoint feature를 사용함으로써 occlusion에 강인하고 jitter를 완화하였다.
- 우리는 masked multi-scale spatio-temporal attention을 결합한 hierarchical transformer encoder를 제시하는데 이는 크고 작은 사람의 body motion들을 represent하기 위해 모든 layer들의 attention map들을 사용하는 hierarchical approach를 사용한다.
- training losses에 따라 online objectives를 선택하여 refined input poses와 final poses간의 position 및 acceleration errors의 cross-optimization을 가능하게 하는 online mutual learning을 제안한다.

Calibrated Attention Masking Network for Temporal Action Localization

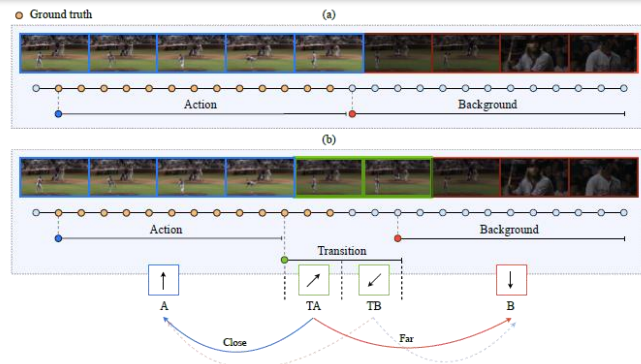
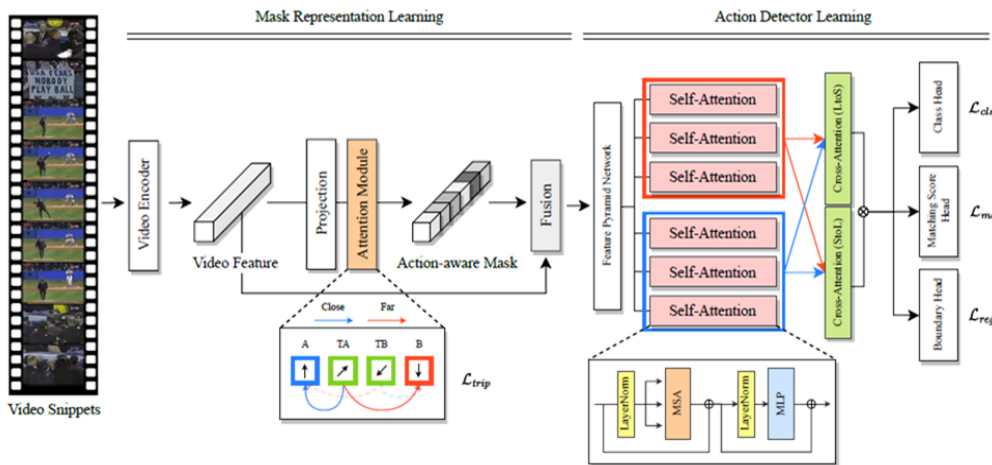
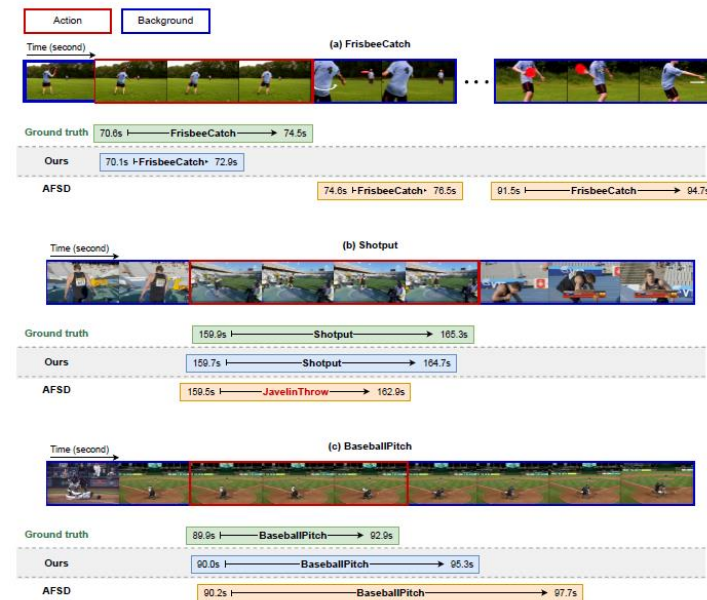


Illustration of mask representation learning



The overall architecture of CAMNet

Visualization of the action-aware attention processes



Qualitative results of our method on THUMOS14

Introduction

- Temporal Action Localization (TAL)의 목표는 편집되지 않은 긴 비디오 내에서 action의 시작, 끝 그리고 클래스를 예측하는 것이다.
- 기존 방법에서 snippet-level features는 제한되고 부족한 시간 정보로 인해 연속된 프레임 간의 불확실성을 야기하며 action 발생의 정확한 예측을 방해할 수 있다.
- 우리의 접근법은 비디오 features를 강화하기 위해 action-aware attention을 동시에 사용하며, 기존 시간 관계를 고려하기 위해 self-, cross-attention 메커니즘을 동시에 사용한다.

Contribution

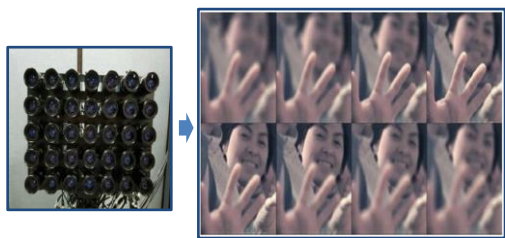
- 우리는 scene ambiguity를 해결하여 향상된 성능을 성취하기 위해 Calibrated Attention Masking Network (CamNet)을 제안한다.
- Action-aware attention module (AAM)은 scene ambiguity에 action-aware attention을 집중함으로써 snippet-level video features를 개선하는 action-aware mask를 생성한다.
- Multi-scale features들을 두 그룹으로 병합하고 cross-attention을 사용함으로써 내재적 시간 관계를 represent하는 Group attention module (GAM)을 설계한다

수행 기업과제 요약

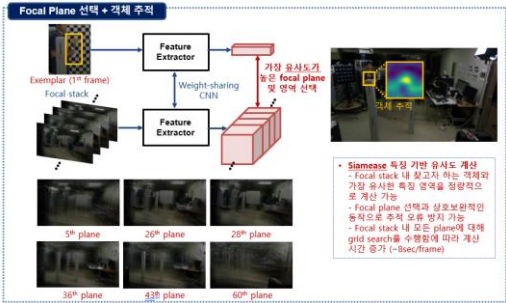
중대형 공간용 초고해상도 비정형 플렌옵틱 동영상
저작/재생 플랫폼 기술 개발



기간: 2020.07~2020.11
발주처: 한국전자통신연구원(ETRI)
근무처: 한성대학교 Visual Intelligence 연구실
목표: 컬러 영상에 대한 2차원 혹은 3차원 객체 추적 기술의 동향을 파악하고 최적의 알고리즘을 선정하여 구현함으로써 플렌옵틱 영상 입력에 대한 성능 검증 및 특성을 비교 분석하고 개선방안을 연구
방법:
• NIQE, Sharpness 1/2를 통해 최적의 선명도를 가진 프레임 선정
• $(t - 1)^{th}$ 프레임의 포커스 정보 기반 탐색 영역 제한
• 포컬 영역 별 최대 유사도 영역 추적
• 후보 영역 스케일 별 최대 유사도 영역 추적
결과:
• Ground-Truth 제작
• 새로운 모델 개발



Plenoptic images



Proposed visual object tracking method

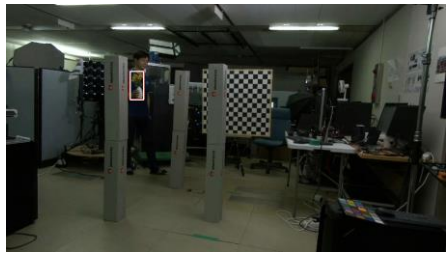
결과



정성적 평가

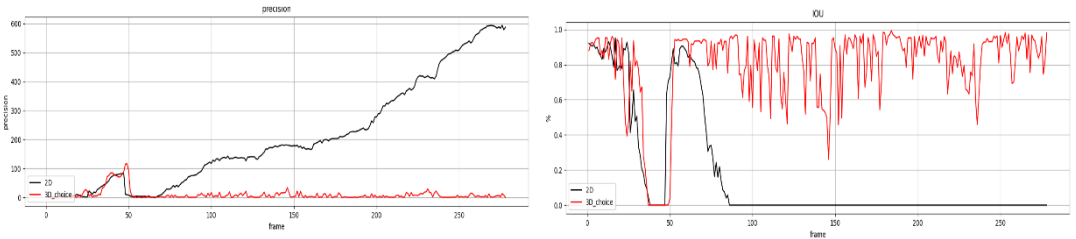


2D 객체 추적 결과



제안 기술 결과

정량적 평가



플렌옵틱 영상 이름	성능 지표	2D 영상 사용	플렌옵틱 영상 사용 (제안된 방법)
NonVideo4_0	Precision (거리)	71.85	3.37
	IoU (%)	20.13	83.04
NonVideo4_1	Precision (거리)	27.39	7.03
	IoU (%)	46.03	67.53
Video3	Precision (거리)	81.08	3.08
	IoU (%)	30.98	91.66

수행 기업과제 요약

골프 트레이닝을 위한 인공지능 기반 골프 스윙 분석 알고리즘 개발



기간: 2021.05~2021.10

발주처: (주)브이씨

근무처: 고려대학교 패턴인식 및 머신러닝 연구실

목표:

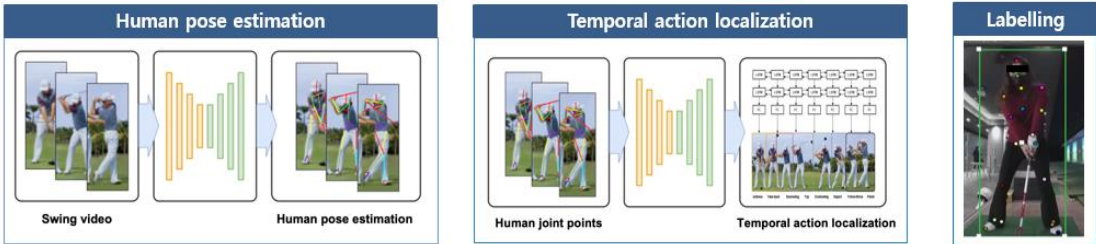
- 골프 스윙 영상에서 골프채를 포함한 관절점 추정 알고리즘 개발
- 골프 스윙 영상에서 주요 스윙 동작 프레임 탐지 알고리즘 개발
- 골프 스윙 영상에서 관절점과 주요 스윙 동작 프레임

방법:

- HRNet기반의 Scalable Pose Network 제안
- Blur augmentation과 Refinement Network 통한 Jitter 문제 개선
- SwingNet기반의 Pose-Guided SwingNet (PGSwingNet) 제안
- Continuous learning과 Auto-labeling을 통한 데이터셋 라벨링

결과:

- 2D HPE: [PCKh@0.5](#) 기준 목표성능(85%) 성취(94.24%)
- TAL: PCE 기준 목표성능(75%) 성취(91.09%)
- Labeling: 3,096개의 비디오중 2,000개의 비디오 선별 후 데이터셋 구축



결과



정성적 평가

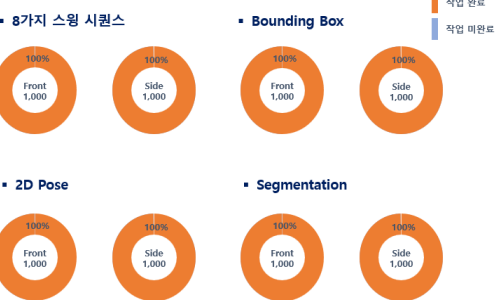


정량적 평가

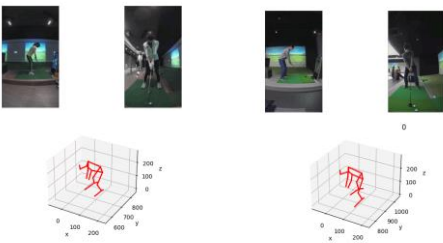
Result of proposed method (%)
Percentage of Correct Events (PCE)

Address	Take-back	Backswing	Top	Downswing	Impact	Follow-through	Finish	Average PCE	Avg w/o AD & F
50.45	83.69	89.12	80.97	96.98	99.70	96.07	40.48	79.68	91.09

Annotation 작업 진행도



후속과제: 3차원 인체 관절점 위치 추정



The Others

Awards



대회명: 미소 인공지능 모델 개발 챌린지, 대상
주관/주최: 과학기술정보통신부, 한국지능정보사회
진흥원, ㈜미소정보기술
주 제: 영유아 발달 장애 예측 모델 개발

수상일: 2021.12.17

대회명: 2020 공개SW 개발자대회, 동상
주관/주최: 과학기술정보통신부, 정보통신산업진흥원,
(주)엘에스웨어
주 제: 안티드론로봇 개발

수상일: 2020.11.30

대회명: 제16회 한성공학경진대회, 은상
주관/주최: 한성대학교
주 제: 안티드론로봇 개발

수상일: 2020.09.25

대회명: 제16회 한성공학경진대회, 동상
주관/주최: 한성대학교
주 제: AI 버섯 도감 어플 개발

수상일: 2020.09.25

대회명: 창의융합 성과 경진대회, 동상
주관/주최: 한성대학교
주 제: 안티드론로봇 개발

수상일: 2020.07.22

대회명: 제15회 한성공학경진대회, 후원업체상, 동상
주관/주최: 한성대학교
주 제: 안전한 자율주행 체험을 위한 FPV 자율주행
RC카 개발

수상일: 2019.09.27

Summary



끊임없는 도전과 다양한 경험을 통해 다음과 같은 역량을 쌓을 수 있었습니다.

- 인공지능과 기계학습에 대한 이해
- 요구사항에 맞는 인공지능 개발 능력과 문제 해결에 결정적인 역할을 하는 창의력
- 맡은 업무를 끝까지 완수하는 책임감과 정신력
- 연구주제를 선정하고 구체화 시키는 능력
- 포기하지 않고 끝까지 연구 결과를 도출할 수 있는 연구정신

지금까지 쌓아 온 연구경험과 도전정신을 바탕으로, 사업 성장에 기여하는
인공지능 연구원이 되고 싶습니다. 인공지능 기술을 연구/개발하고, 이를 창의
적으로 융합하여 새로운 가치를 만들어 내는 연구/개발자가 되겠습니다.

아직 기업의 요구사항에 정확히 맞지 않더라도,
아직 기업이 바라는 기술이 없더라도,
빠르게 필요 기술을 습득하고 요구사항을 충족하는
신입사원, 열심히 일하는 일꾼이 되겠습니다. 한번만 기회를 주세요!

Thank you