

고해상도 단일 카메라 기반의 골프 Impact 검출 방법

유 주 한*

A Golf Impact Detection Method based on High-resolution Mono Camera

Ju Han Yoo*

이 논문은 2022학년도 세명대학교 교내학술연구비 지원에 의해 수행된 연구임

요 약

본 논문에서는 골프공과 플레이어의 움직임을 이용한 고해상도 단일 카메라 기반의 골프 Impact 검출 방법을 제안한다. 먼저, 모양과 크기를 기반으로 추출된 골프공 후보 중에서 확률론적 투표 방법을 이용해서 골프공을 검출하고, 검출된 골프공 위치에서 프레임 사이의 픽셀값들의 차이를 계산함으로써 골프공의 움직임을 추정한다. 다음으로, 스윙 움직임이 많은 부분을 기준으로 플레이어의 움직임 정보를 추출하고, 마지막으로 이 정보와 골프공의 움직임을 함께 고려해서 Impact를 검출한다. 제안하는 방법의 골프공과 Impact 검출 정확도를 계산하기 위해서 골프컨트리클럽에서 스윙 데이터셋을 구축하였으며, 실험을 통해서, 제안하는 방법이 고해상도 이미지에서도 빠르고 정확하게 골프공과 Impact를 검출할 수 있었음을 확인하였다.

Abstract

In this paper, we propose a high-resolution mono camera-based golf impact detection method using the golf ball and player motion. First, the golf ball is detected by using a probabilistic voting method among candidates of the golf ball are extracted based on the shape and size, and the motion of the golf ball is estimated by calculating the difference in pixel values between frames at the detected golf ball position. Finally, the player motion information is extracted based on the part where the swing movement is high, and the impact is detected by considering this information and the motion of the golf ball together. A swing dataset was constructed from a golf country club to calculate the accuracy of the golf ball and the impact detection, and our experiments show that the proposed method can quickly and accurately detect the golf ball and the impact even in high-resolution images.

Keywords

golf swing analysis, golf impact detection, golf ball detection, golf swing, golf pose, pose segmentation

* 세명대학교 컴퓨터학부 교수
- ORCID: <http://orcid.org/0000-0002-2329-7375>

· Received: Oct. 12, 2022, Revised: Dec. 07, 2022, Accepted: Dec. 10, 2022
· Corresponding Author: Ju Han Yoo
School of Computer Science, Semyung University, Jecheon-si,
Chungcheongbuk-do, Korea
Tel.: +82-43-649-1280, Email: unchinto@semyung.ac.kr

1. 서 론

골프는 코로나 19 팬데믹 이후 가장 호황을 맞은 스포츠 중 하나가 되었다. 이것은 골프가 타인과의 접촉률이 비교적 낮은 ‘언택트 레저(Untact leisure)’라는 인식과 소득 구조 변화에 따른 젊은 세대 유입 등으로 인한 골프 이용자가 증가하였기 때문이다[1]. 골프 인구 증가는 상대적으로 강사 공급 부족으로 이어졌고, 이것은 언택트 인식과 더불어 혼자서 골프를 배울 수 있는 서비스의 수요로 이어졌다. 이러한 서비스 중에서, 영상으로 자신의 골프 스윙 자세를 보여주고, 그것에 대한 분석 결과를 제공해 주는 서비스가 대표적이다. 이러한 서비스는 사용자에게 자신의 자세를 보고 싶어 하는 욕구를 충족시켜 줄 수 있는 동시에 자세 교정을 통해서 실력향상을 기대할 수 있다는 장점이 있다.

골프스윙 자세를 분석하기 위해서 Sway, Head-up, X-factor, V-zone 등의 여러 가지 요인들을 고려할 수 있는데, 이러한 요인들은 골프스윙을 특정 동작들로 나눈 후 각각의 동작들 사이에서의 팔과 클럽의 움직임을 통해서 분석한다. 일반적으로, 골프에서 스윙은 Address, Back Swing, TB(Top of Back), Downswing, Impact, FT(Follow-through), Finish의 7가지 동작으로 분류될 수 있고[2]-[5], 이러한 스윙 동작들은 앞서 언급한 요인들과 다음과 같은 관계가 있다.

스웨이(Sway) : Address에서 Tb까지 혹은 Tb에서 Impact까지의 스윙 동작 중에서 상체가 횡 방향으로 움직이는 현상이다. 이 현상이 나타나면 스윙 중심이 무너져 미스샷을 유발할 수 있다[6]. 브이존(V-zone) : 이것은 허리, 어깨, 골프공을 연결하는 V 모양의 존을 의미한다. Tb에서 Impact까지 스윙 동작을 할 때, 이 존을 벗어나게 되면 훅이나 슬라이스가 발생할 수 있다. 헤드업(Head-up) : Impact 전에 플레이어의 중심축 높이가 변해서 플레이어의 헤드가 위로 올라가는 현상이다. 이것은 스윙의 밸런스를 무너뜨리는 원인이 되고 미스샷을 유발할 수 있다[7].

이러한 요인들은 공과 클럽의 면이 만나는 순간인 Impact를 기반으로 분석할 수 있다는 공통점을 가지고 있다. 이 Impact는 클럽 페이스의 각도, 위치, 속도 등 여러 가지 요소가 공이 날아가는 방향

을 결정하기 때문에 골프스윙 동작 중에서 가장 중요한 동작이다. 결국, 정확한 Impact를 위해서 Back Swing과 Downswing 등의 동작들을 연습한다고 볼 수 있다. 따라서, 본 논문에서는 이러한 Impact를 자동으로 검출하는 방법을 제안한다. 또한, 제안하는 방법은 높은 해상도의 이미지들을 빠르게 분석할 수 있어서, 앞서 언급한 자신의 자세를 보고 싶어 하는 플레이어에게 자동으로 양질의 Impact 이미지를 제공할 수 있는 장점이 있다.

일반적으로, 플레이어들은 근육-관절의 워업, 리듬, 템포, 스윙궤도 등을 확인하기 위해서 실제 스윙을 하기 전에 연습스윙을 한다. 기존의 스윙 동작을 검출하는 방법[8]-[11]들은 플레이어의 움직임 기준으로 스윙 동작을 검출하기 때문에, Impact가 없는 연습스윙에서 Impact를 검출하는 단점을 가지고 있다. 또한, 캐디와 다른 플레이어들로 인해서 플레이어 영상의 가려짐이 유발될 수 있고, 이 경우에는 플레이어의 스윙 동작을 구분하는 것이 어려울 수 있다. 이러한 문제점을 해결하기 위해서, 본 논문에서는 플레이어의 가려짐 유무를 확인하고 플레이어의 움직임뿐만 아니라 골프공을 함께 이용해서 Impact를 검출하는 방법을 제안한다. 이것은 앞서 언급한 가려짐으로 인한 Impact 동작의 오검출을 줄일 수 있고 연습스윙에서 검출될 수 있는 Impact를 제거함으로써 정확도를 높일 수 있다.

제안하는 방법은 먼저 취득된 스윙 영상에서 실제 티샷을 하는 플레이어를 검출하고 검출된 플레이어 영역에서 가려진 상태를 확인한다. 이러한 전처리 과정을 통해서, 가려짐이 없는 시퀀스에 대해서만 동작 분석을 수행한다. 그리고 플레이어의 움직임만을 기준으로 Impact를 판단하는 것이 아니라, 골프공의 움직임도 추가로 고려해서 Impact를 판단한다. 이것은 실제 Impact가 일어나면 Tee 위에 있는 골프공이 없어진다는 강력한 단서 때문에 가능하다. 이를 위해서, 골프공을 검출한 다음 그 위치 영역에서의 골프공 존재의 유무를 판단한다. 그러면 골프공이 없어진 시점이 Impact가 발생한 시점이라고 판단할 수 있다. 또한, 이것은 연습스윙과 Impact가 포함된 실제 스윙을 구분하는데 중요한 단서로 이용될 수 있다.

성능 평가를 위해서 골프컨트리클럽에서 취득한 고해상도 스윙 이미지들로 구성된 데이터셋을 구축하였고, 이것은 다양한 색을 가진 골프공, 조도 변화, 연습스윙 등을 포함한다.

본 논문의 구성은 다음과 같다. 1장의 서론에 이어 2장에서는 본 논문에서 제안하는 방법과 관련된 연구를 설명하고, 3장에서는 제안하는 방법에 대하여 설명한다. 4장에서는 실험 및 결과에 대해서 분석하고, 마지막으로 5장에서 결론을 도출한다.

II. 관련 연구

골프스윙 동작을 자동으로 검출하는 연구들은 카메라를 이용하는 방법뿐만 아니라 거리 센서를 추가로 이용하는 방법, 관성 및 가속도계와 같은 센서들을 이용하는 방법[12] 등으로 나눌 수 있다.

[8][13]은 Gyroscope, Strain Gage Sensors, Accelerometer 등의 멀티 센싱 데이터들을 CNN(Convolutional Neural Network)기반 분류기를 이용하여 골프 동작을 검출하였다. [9]는 각 신체 부위에 부착된 단일 관성 측정장치로부터 골프스윙 동작들을 구분하는 방법을 제안하였다. 이를 위해서 그것은 시계열 데이터의 특성을 자동으로 학습할 수 있는 BLSTM(Bidirectional LSTM)[14] 및 CNN 모델을 이용하였다. 이런 방법들은 정확도가 높지만, 클럽이나 신체에 별도의 센서를 부착하여 사용하기 때문에 불편할 뿐만 아니라 스윙 동작을 확인하기 위한 별도의 카메라가 필요하다는 단점이 있다.

[10]은 키넥트와 같은 RGB-D 카메라 이용하여 깊이 정보를 기준으로 각각의 관절들을 추정하고 그것을 기반으로 스윙 동작을 검출하는 방법을 제안하였다. 이 방법은 단순히 자세만을 이용하여 스윙 동작을 검출하기 때문에, 연습스윙뿐만 아니라 부분적인 스윙 동작도 많이 발생하는 실외 골프 컨트리클럽에서는 실제 Impact가 포함된 스윙 동작을 검출할 수 없다. 또한, 깊이 정보를 위한 IR 센서가 추가로 필요하므로 센서의 크기가 크고 실외 환경에서는 깊이 정보를 취득하기 어렵다는 단점이 있다.

[11]은 RGB 이미지만을 이용한 DNN(Deep Neural Network) 기반의 스윙 동작을 검출하는 방법을 제

안하였다. 그것을 위해서 그들은 유튜브에서 취득한 프로 선수들의 골프스윙 시퀀스로 구성된 저해상도 데이터베이스인 GolfDB를 구축하였고, 이 데이터베이스에서 골프스윙 동작을 구분하는 SwingNet이라는 LSTM기반의 경량의 네트워크 모델을 제안하였다. 하지만, 이것은 학습을 위해서 많은 양의 가공된 데이터가 필요하고 연산량도 많아서 다양한 실외 환경을 가지고 있는 골프장에 적용하는 것은 어렵다. 그리고 이 방법 역시 자세만을 이용하여 스윙 동작을 검출하기 때문에 연습스윙에서 Impact를 검출할 수 있는 오인식이 발생할 수 있다.

최근에는, 사용자가 자신의 스윙 동작과 전문가의 스윙 동작의 차이점을 직관적으로 이해할 수 있도록 신경망을 활용한 골프스윙 분석 방법을 제안하였다[15]. 그것은 먼저 신경망을 사용하여 다양한 스윙 단계의 동작을 정렬하고, 동작 불일치 검출기를 이용해서 네트워크에서 학습한 잠재 공간의 동작 간의 미세한 차이[16]를 검출한다. 이 방법은 2D 관절 정보를 추출하고 그것을 이용해서 다시 3D 위치 정보를 추정하기 때문에 오차가 누적될 수 있는 단점이 있다. 또한, 딥러닝 기반의 모델과 그것을 학습하기 위해서 앞서 언급한 GlofDB를 이용하였기 때문에 고해상도 이미지 적용이 어렵고 많은 계산량이 필요하다.

이미지 기반의 방법들은 스윙 동작을 구분하기 위해서 플레이어의 움직임만을 고려하거나 딥러닝 기반의 특징 추출 및 분류 방법을 이용하였다. 따라서, 이 방법들은 앞서 언급한 연습스윙과 실제 스윙을 구분해 낼 수 없고, 실외 환경에 적용하기 위해서는 많은 학습 데이터와 계산량이 필요하다는 단점이 있다. 본 논문에서는, 골프스윙 동작 중에서 중요성이 높은 Impact를 고해상도의 RGB 단일 이미지만을 이용하여 실내뿐만 아니라 실외 환경에서도 강인하게 검출하는 방법을 제안한다. 이것은 플레이어의 움직임뿐만 아니라 골프공의 움직임도 함께 고려하기 때문에, 연습스윙이 아닌 실제 스윙에서 Impact를 검출할 수 있다. 또한, 딥러닝 기반의 특징 추출 및 분류를 위한 학습을 하지 않기 때문에 골프스윙에 대한 많은 학습 데이터와 시간, 계산량이 필요 없다는 장점이 있다.

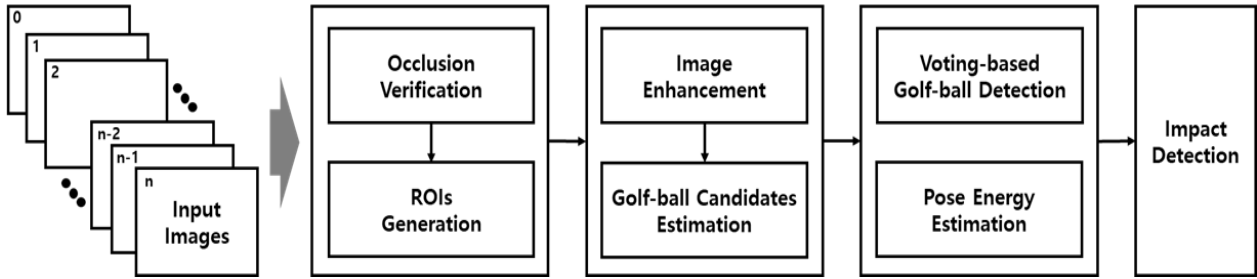


그림 1. 제안하는 방법의 구성도
Fig. 1. Block diagram of the proposed method

III. 제안하는 방법

그림 1에서 보는 것과 같이, 제안하는 방법은 먼저 연속된 스윙 이미지 시퀀스를 입력으로 받고 ROI(Region of Interest)를 구성한다. 다음으로 실외 환경에서 취득한 입력 이미지를 조명 변화와 노이즈에 강인한 필터링 과정을 수행하고, 골프공의 모양과 크기를 기반으로 골프공 후보들을 스윙 시퀀스에서 추출한다. 그 후보 중에서, 확률론적 투표 방법[17]을 이용해서 최종적으로 골프공을 검출하고, 마지막으로, 플레이어와 골프공의 움직임에 의하여 Impact를 검출한다.

3.1 Occlusion Verification

본 논문에서 이용되는 입력 이미지는, 그림 2에서 보는 것과 같이, 플레이어의 측면에 설치된 카메라에 의해서 취득된다. 일반적으로 이러한 환경에서는, 티샷에 방해가 되지 않도록 공이 나아가는 방향의 반대편에서 캐디나 다른 플레이어들은 대기한다. 그러면, 그들은 이미지에서 플레이어보다 크게 나타날 수 있고 상대적으로 작게 나타나는 플레이어를 가릴 수 있다. 따라서, 입력 이미지에서 복수의 사람이 검출될 때는 상대적으로 크기가 작은 플레이어를 기준으로 다른 사람이 겹치지 않으면 가려짐이 없다고 판단하고 Impact 검출을 진행한다. 이것을 위해서 제안하는 방법에서는 플레이어를 검출하기 위해서 사람에 대해 사전 학습된 웨이트를 가진 YOLO[18][19] 모델이 이용된다. 따라서 제안하는 방법에서는 추가적인 학습이 필요 없다.

3.2 ROIs Construction

플레이어가 티샷하는 곳을 Teeing Area라고 한다. 이 영역은 일반적으로 2개의 티마크로 표시되며, 플레이어는 이 영역의 임의의 위치에 티를 꽂고 그 위에 골프공을 놓고 친다. 결국, 넓은 Teeing Area에서 플레이어가 어디에서 티샷하는지 알 수 없으므로, 이미지 전체 영역에서 Impact를 분석해야 한다. 하지만 이것은 많은 계산량이 필요하므로, 스윙이 이루어지는 영역에서만 동작을 분석하면 그 계산량을 줄일 수 있다. 그것을 위해서 먼저 앞서 Occlusion Verification에서 검출한 플레이어의 영역을 기준으로 플레이어 몸통에 대한 ROI(P_{roi})를 설정한다. 이 영역은 스윙하는 동안 팔의 움직임을 포함한 플레이어의 움직임이 가장 큰 영역이기 때문에 플레이어의 움직임을 계산하기 위해서 이용된다. 그리고 골프공의 위치는 플레이어 발 앞쪽 아래 영역으로 정해지고 그 영역을 골프공 ROI(G_{roi})로 설정한다.



그림 2. 입력 이미지와 관심 영역 예
Fig. 2. Example input image and region of interest

그림 2는 입력 이미지의 뷰포인트와 Teeing Area (흰색 티마크 사이 영역)에서의 플레이어(빨간색 사각형), 플레이어의 몸통(파란색 사각형), 골프공(노란색 사각형)의 ROI를 나타낸다.

3.3 Image Enhancement

실내 골프장은 조명의 밝기가 일정하게 유지되는 반면에, 제안하는 방법이 적용되는 실외 골프장은 날씨의 영향을 많이 받기 때문에 입력 이미지에서의 밝기 변화가 크다. 따라서, 밝기 변화에 강인한 이미지 분석을 위해서 P_{roi} 와 G_{roi} 에 다음과 같은 이미지 Enhancement 방법들을 수행한다.

먼저, 골프공을 검출하기 위해 G_{roi} 에 대해서 Green 채널을 제외한 Red와 Blue 채널을 추출한다. 이것은, 그림 3에서 보는 것과 같이, Green 채널에서의 골프공보다 Red와 Blue 채널에서의 골프공이 녹색이 많은 배경과 구분이 더 잘 되기 때문이다. 그리고, 이미지에서 상대적으로 작은 크기의 골프공을 검출하기 위해서 골프공의 형태 정보뿐만 아니라 주변보다 더 밝거나 어두운 영역인 Blob 정보를 함께 이용한다.

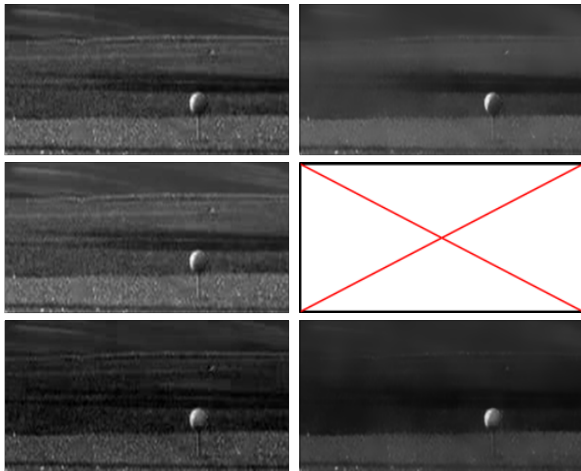


그림 3. 컬러 분할된 골프공 ROI(왼쪽 줄)와 그것의 Bilateral Filter 결과(오른쪽 줄) 예. Red 컬러공간 이미지(위), Green 컬러공간 이미지(중간), Blue 컬러공간 이미지(아래)

Fig. 3. Example of color segmented golf ball ROI(left row) and its Bilateral Filter result(right row) example. red color space image(top), green space image(middle), blue color space image(bottom)

이를 위해서, 앞서 원본 이미지에서 분리된 Red와 Blue 채널에 Blob의 원형 모양을 유지하면서 노이즈를 줄일 수 있는 Bilateral Filter[20][21]를 적용한다. 그림 3의 왼쪽 줄에서는 컬러 분할된 G_{roi} 를 보여주고, 오른쪽 줄에서는 각 채널에 Bilateral Filter 결과를 보여준다. 이것을 통해서, 그 필터가 골프공의 모양을 유지하면서 배경의 노이즈는 효과적으로 줄이는 것을 알 수 있다. 제안하는 방법에서는 Green 채널을 이용하지 않기 때문에, 그것의 Bilateral Filter 결과는 나타내지 않았다.

3.4 Golf-ball Candidates Estimation

골프공은 입력 이미지에서 나타날 수 있는 다른 물체와는 다르게 둥근 모양의 균일한 Intensity를 가진다. 이것은 이미지에서 비슷한 성질을 가지는 픽셀들로 구성된 집합인 Blob 형태로 나타난다. Blob 검출[22]은 비슷한 Intensity를 가지는 픽셀들을 라벨링 한 다음, 근접한 픽셀들까지 군집화하고, 이렇게 군집화된 모양의 종류에 따라서 분류를 추가로 수행할 수 있다. 따라서, 이미지에서 원형 모양의 골프공을 노이즈에 강인하게 검출하는 데 효과적으로 이용될 수 있다. 제안하는 방법에서는, 앞서 추출된 Red와 Blue 채널 이미지에서 다음과 같은 3가지 제약사항들을 이용하여 Blob을 추출한다.

- Circularity : Blob의 둥근 모양의 정도, 이미지에 나타나는 골프공의 둥근 모양 정도
- Area : Blob 크기, 이미지에 나타나는 골프공의 크기
- Distance : 추출된 Blob 사이의 최소거리, 이미지에 나타나는 Blob 사이의 거리가 이 값보다 가까우면 하나의 Blob으로 고려함

추출된 원형 모양의 Blob을 b^{ik} 로 나타내고 이것은 골프공 후보로 선택된다. 여기에서, i 는 프레임 Index를 나타내고, k 는 i 번째 프레임에서 추출된 Blob의 Index를 나타낸다. 그리고, 각각의 b^{ik} 는 그것의 중심점 좌표 b_c^{ik} 와 지름 b_s^{ik} 를 가진다. 본 논문에서는 이러한 Blob들을 전체 스윙 시퀀스에서 추출하고 이것들을 골프공 후보로 고려한다. 그림 4는 Red와 Blue 채널에서 추출된 모든 Blob 결과 예를 보여준다.

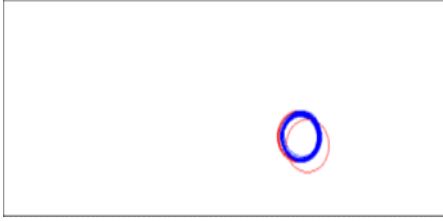


그림 4. Blob 추출 결과 예. Red 채널에서 추출된 Blobs(빨간색 원), Blue 채널에서 추출된 Blobs(파란색 원)
Fig. 4. Example blob extraction result. Blobs extracted from the Red channel(red circles), blobs extracted from the Blue channel(blue circles)

이 그림에서 보는 것과 같이, 실제 골프공이 있는 위치에서 많은 Blob이 추출되었지만, 그렇지 않은 위치에서도 노이즈에 의해 Blob들이 추출되었다. 따라서, 추가적인 방법을 이용해서 이러한 후보 중에서 최종적으로 골프공의 위치를 검출한다.

3.5 Voting-based Golf-ball Detection

G_{roi} 에서, 추출된 b^{ik} 는 Address에서 Impact까지 나타나고 Impact부터 Ft까지는 나타나지 않는다. 결국, 전체 스윙 시퀀스에서 고려한다면, 실제 골프공은 일정한 위치에서 가장 오랫동안 머물러 있게 되고, 그 위치에서 Blob도 가장 많이 추출된다. 이것은 노이즈에 의해서 실제 골프공이 없는 위치에서도 추출될 수 있는 Blob들을 효과적으로 제거하는데 도움이 된다. 이를 위해서, 제안하는 방법에서는, 차선의 소실점을 노이즈에 강인하게 추정하기 위해서 제안된 확률론적인 투표 방법[17]을 이용하여 골프공의 위치를 검출한다. 이미지에서 b^{ik} 는 다음과 같이 Gaussian 분포로 고려한다.

$$P_b^{ik}(x,y;u,v,\sigma_{s;x},\sigma_{s;y}) = \frac{1}{2\pi\sigma_{s;x}\sigma_{s;y}} e^{-\left[\frac{(x-u)^2+(y-v)^2}{2\pi\sigma_{s;x}\sigma_{s;y}}\right]} \quad (1)$$

여기에서 $(x,y) \in G_{roi}^i$, (u,v) 는 b_c^{ik} 의 좌표를 나타낸다. 그리고 $\sigma_{s;x}$ 와 $\sigma_{s;y}$ 는 x 축 y 축으로의 표준편차이다. 이 논문에서는, 골프공의 모양이 원형이기 때문에 다음과 같이 Gaussian 분포는 Isotropic이라고 가정한다.

$$\sigma_s = \sigma_{s;x} = \sigma_{s;y} = \sqrt{\sigma_s^2 + \sigma_s^2} \quad (2)$$

$\sigma_s = a \cdot (1/b_s^{ik})$ 이고, a 는 Scale Factor이다. 그러면, Voting 함수 P_w 는 다음과 같다.

$$P_w(x,y) = \sum_{i,k}^N P_b^{ik}(x,y;u,v,\sigma_s) \quad (3)$$

여기에서, $N = i \times k$ 이다. 마지막으로, 골프공의 위치는 다음과 같이 추정할 수 있다.

$$V(x,y) = \underset{x,y}{\operatorname{argmax}} [P_w(x,y)] \quad (4)$$

3.6 Pose Energy Estimation

골프스윙에서, Address, Tb, Impact, Ft의 주요 동작들은 그 움직임의 크기가 일정한 연속된 형태를 가지면서 나타난다. 예를 들어, 움직임이 크면 강, 작으면 약이라고 한다면, 스윙 시퀀스에서의 플레이어 몸통의 움직임 정도는 “Address(약)→(중)→Tb(약)→(중강)→Impact(강)→(중강)→Ft(약)”으로 나타난다. 이러한 성질을 이용하기 위해서, P_{roi} 에서 움직임 크기와 비례하는 에너지 M 을 Optical Flow[23][24]를 이용하여 다음과 같이 정의한다.

$$M^i = \sum_{(x,y) \in P_{roi}^i} O^i(x,y) \quad (5)$$

여기에서, $O^i(x,y)$ 는 i 번째 이미지의 픽셀 (x,y) 에서의 Optical Flow 값을 나타낸다. 최종적으로, 제안하는 방법에서는 이미지마다 P_{roi}^i 의 영역의 크기가 다르게 나타나기 때문에 M^i 를 P_{roi}^i 에 포함된 픽셀의 개수로 나눈 정규화된 M^{i*} 을 이용한다.

3.7 Impact Detection

Impact 직후에는 골프공이 없어지기 때문에 골프공의 존재 유무를 이용해서 우리는 Impact가 발생했는지 판단할 수 있다. 이를 위해서, 골프공 영역에서

이웃한 프레임 사이의 픽셀값의 차이가 있으면 전경, 그렇지 않으면 배경으로 구분한다. 이 전경은 Impact가 발생할 때 생성되기 때문에 이것은 Impact를 검출하는데 중요한 단서가 될 수 있다. 이 전경을 추출하기 위해서, 제안하는 방법에서는 BGS (Background Subtraction) 방법을 이용한다[25][26].

G_{roi} 에서 BGS의 전경 영역의 Pixel 개수를 score S 라고 정의한다. 그러면, 일반적으로, 전체 스윙 시퀀스에서 Impact와 Address 동작을 할 때 S 값이 크게 나타난다. 이것은 Impact 동작일 때 클럽헤드가 빠르게 골프공을 타격함으로써 프레임 사이의 픽셀값 차이가 크기 발생하기 때문이다.

또한, Address 동작일 때는 클럽헤드를 골프공 뒤에 위치시키고 천천히 Tb 동작으로 이어지는 움직임에 의해서 프레임 사이의 픽셀값 차이가 발생하기 때문에 S 값이 크게 나타난다. 본 논문에서는 이 두 가지 동작에서 Impact를 구분하기 위해 다음과 같은 제약조건을 고려한다.

- Impact의 골프 클럽 속도가 Address의 클럽보다 빠르므로, Address보다 Impact에서의 S 값이 나타나는 프레임의 개수가 적다.
- 골프 클럽의 빠른 속도로 인해서 Impact보다 Address에서의 S 값이 작다.

그림 5는 실제 스윙 동작을 하는 동안에 S 값의 변화를 그래프로 나타낸 것이고 Address와 Impact 동작을 표시하였다. 이 그래프를 보면, S 값이 나타난 Impact의 폭이 Address의 폭보다 좁게 나타나는데, 이것은 앞서 언급한 Impact 동작 속도가 Address보다 빠르기 때문이다. 또한, 이것은 빠른 Impact 속도 때문에 이웃한 프레임 사이의 움직임 차이가 명확하지 않을 수 있으므로, Impact의 S 값이 Address의 그것보다 작게 나타나는 것을 보여준다. 스윙 전체 구간에서 보면, Address나 Impact가 아닌 프레임에서도 상대적으로 작은 S 값이 노이즈처럼 나타난다. 이것은 스윙하는 동안 발생하는 그림자, 나뭇잎, Tee와 연결된 악세서리 등의 움직임 때문이다.

따라서, 본 논문에서는 최종적으로 Impact를 판단하기 위해서 S 값이 큰 2개의 프레임을 검출하고 그 중에서 앞서 언급한 M^{i*} 가 큰 프레임에서 Impact가 발생했다고 판단한다. 왜냐하면, Address보다 Impact의 움직임이 더 크기 때문이다.

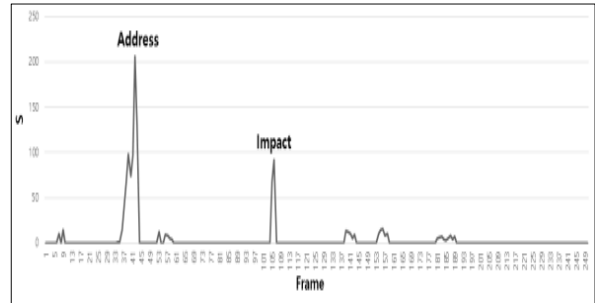


그림 5. 스윙 동작을 하는 동안의 S 값 변화
Fig. 5. S value change during swing motion

IV. 실험 결과

4.1 데이터셋

본 연구의 실험을 위해서 스윙 동작 이미지를 취득하기 위한 시스템을 골프컨트리클럽에 설치하였다. 해당 시스템에서는 60fps의 컬러 RGB 카메라가 플레이어의 측면에(그림 2에서 보는 것과 같이) 설치되고, FHD(1920x1080)의 고해상도로 취득된 이미지들이 서버로 전송된다.

본 시스템을 통해서 조사된 100명의 일반인의 평균 스윙 시간은 S_{TT} (Tb에서 Impact까지 걸린 시간)에서는 약 390ms, S_{IF} (Impact에서 Ft까지 걸린 시간)에서는 약 410ms이었다. 이것은 Address를 하기 전에 하는 준비 동작 시간이 제외된 것이다. 따라서 본 연구에서는 안정적인 분석을 위해서 실제 플레이어가 Address 동작을 취하고 Ft까지의 시간을 약 4초(총 240프레임)로 설정하고 연속된 이미지들을 하나의 시퀀스로 취득한다. 그러나 이 시스템에서는 취득 모듈과 전송 모듈이 비동기로 동작하기 때문에, 전송 지연이나 오류에 의해서 정해진 시간 동안 서버에 저장되는 이미지 개수는 일정하지 않을 수 있다. 따라서 실제로 4초 동안 저장된 한 개의 시퀀스 이미지 개수는 220~240장으로 변동이 있었다.

본 연구에서는, 제안하는 방법에 대한 골프공과 Impact 검출의 정확도를 측정하기 위해서 골프컨트리클럽에서 취득한 82개의 시퀀스 (총 20,763프레임, 1080p 해상도)를 가지는 데이터셋을 구축했다. 이것은 측정하고자 하는 대상에 따라서 Positive와 Negative 이미지의 구성이 다르다.

먼저, 골프공 검출 정확도를 측정하기 위한 데이터셋은 골프공의 유무를 기준으로 Positive와 Negative로 나누고, Positive 시퀀스가 78개, Negative 시퀀스가 4개로 구성된다. 그림 5의 왼쪽 열에서 보는 것과 같이, 이것은 흰색의 골프공뿐만 아니라 빨간색, 노란색, 녹색등 다양한 색을 가진 골프공들을 포함한다. Impact 검출 정확도를 측정하기 위한 데이터셋은 72개 Positive, 10개 Negative 시퀀스로 구성된다. 여기에서 Positive 시퀀스는 실제로 Impact가 발생한 경우이고, Negative 시퀀스는 연습스윙, 카메라 가리어짐, 취득시스템 오류 등으로 인해서 Impact가 실제로 발생하지 않았거나 스윙 동작을 확인할 수 없는 경우이다.

4.2 Golf-ball Detection Evaluation

골프공을 검출의 정확도를 판단하기 위해서, 표 1과 같이 Precision, Recall, F1 score가 평가지표로 이용된다. 제안하는 방법에서는 각각 0.97로 높은 정확도를 보여준다.

표 1. 제안하는 방법의 정확도

Table 1. Accuracy of the proposed method

	Golf ball	Impact
Precision	0.97	0.93
Recall	0.97	1
F1	0.97	0.96

그림 6의 오른쪽 열은 다양한 색의 골프공을 검출한 결과를 나타낸다. 그것은 골프공이 넓은 Teeing Area 내의 오른쪽이나(두 번째 행) 왼쪽에(세 번째 행) 위치하거나, 랜덤한 위치에 있더라도 제안하는 방법이 정확하게 그 위치를 찾아낼 수 있다는 것을 보여준다. 그리고, 그림 7에서 보는 것과 같이, 제안하는 방법은 주변 환경이 다른 플레이어에 의해 가려지고 그의 다리 사이에 있는 골프공도 정확하게 검출할 수 있다. 하지만, 그림 8에서 보는 것과 같이, 골프공의 모양과 크기가 비슷한 Tee 악세서리가 있는 경우에는 제안하는 방법이 골프공을 정확하게 검출하지 못한다. 이것은 골프공의 모양과 크기가 비슷한 Tee 악세서리가 골프공과 함께 Blob

으로 추출되었기 때문이다. 그리고, Impact가 발생하면 골프공뿐만 아니라 Tee 악세서리도 함께 움직일 수 있어서 Impact를 검출할 때 이용한 BGS 정보를 추가로 이용하더라도 골프공과 이 악세서리를 구분할 수 없었다.



그림 6. Teeing area 내에 있는 다양한 위치와 색을 가진 골프공(왼쪽 열)과 골프공 검출 결과(오른쪽 열)의 예
Fig. 6. Examples of golf balls of various positions and colors in the teeing area(left column), and examples of golf ball detection results(right column)



그림 7. 복잡한 환경에서 검출된 골프공 예(빨간색 사각형)
Fig. 7. Example of a golf ball detected in a complex environment(red rectangle)



그림 8. 골프공 오검출 예
Fig. 8. Example of misdetection of a golf ball

4.3 Impact Detection Evaluation

Impact 할 때 클럽의 속도가 매우 빠르므로, 일반적인 카메라를 이용하는 시스템에서는 그 순간의 프레임을 취득하기가 어렵다. 그림 9에서 보는 것과 같이, 본 연구에서 이용한 시스템에서도 Impact 순간의 프레임이 취득된 것이 아니라 그 순간의 전과 후의 프레임들만 취득되었다(실제 Impact는 $[t]$ 와 $[t+1]$ 프레임 사이에서 발생). 따라서, Impact는 그것이 발생한 하나의 프레임으로 결정될 수 있는 것이 아니라 일정한 간격을 가지는 구간으로 나타내는 것이 합리적이다.

본 논문에서는, Impact 검출 정확도를 판단하기 위해서 그것의 GT(Ground Truth)를 하나의 프레임이 아니라 일정한 구간으로 고려한다. 이것을 위해서 [13]에서 이용된 PCE(Percentage of Correct Events)를 이용한다. 이것은 원하는 이벤트가 포함된 프레임 개수를 결정하기 위해서 tolerance δ 를 이용하는데, 이것은 사람마다 다를 수 있는 GT 판단의 변동성을 최소화하기 위해서 Sampling Frequency와 프레임 개수를 고려한다. 이것으로부터 본 논문에서는 δ 를 2로 정한다. 그러면 그림 9의 시퀀스 예에서, 빨간색 사각형으로 표시한 $[t]$ 와 $[t+1]$ 이 Impact의 GT 프레임으로 고려될 수 있다.

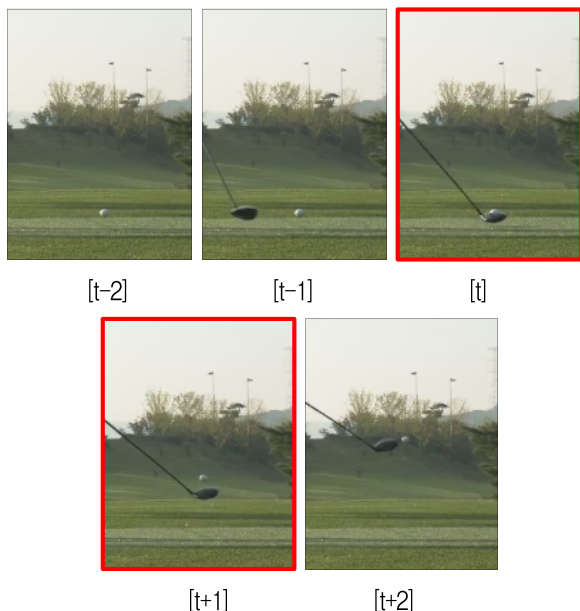


그림 9. Impact 전후의 프레임 예

Fig. 9. Examples of frames before and after the Impact

표 1에서 보는 것과 같이, Impact 검출에 대한 Precision, Recall, F1 score는 각각 0.93, 1, 0.96이다. 이것은 제안하는 방법이 정확한 Impact 순간의 프레임이 없어도 Impact를 검출할 수 있다는 것을 보여준다. 그림 10은 조도 변화가 다양한 환경의 시퀀스에서 Impact 검출 결과를 나타낸다. 빨간색 사각형이 Impact라고 검출한 프레임이고, 그것은 제안하는 방법이 조도 변화가 크고 그림자가 많이 발생하는 실외 환경에서도 정확하게 Impact를 검출할 수 있다는 것을 보여준다.



그림 10. 다양한 환경(행)에서의 Impact 검출 결과(중간 열)의 예. 왼쪽과 오른쪽 열은 Impact 바로 전·후의 프레임

Fig. 10. Examples of impact detection results(middle column) in various environments(rows), the left and right columns are frames immediately before and after Impact

스윙 동작을 했지만, 골프공을 타격하지 않으면 연습스윙으로 고려할 수 있다. 그림 11은 골프공과 아주 가깝게 그것의 위나 옆으로 클럽이 지나가는 연습스윙의 예를 보여준다. 클럽이나 신체의 움직임만을 이용한다면 이러한 연습스윙에서는 [t] 프레임을 Impact로 검출하겠지만, 제안하는 방법은 골프공의 움직임을 함께 고려하기 때문에 이러한 연습스윙에서는 Impact가 없다고 판단한다(True negative).



그림 11. 연습 스윙에 대한 시퀀스 이미지(행)의 예. 왼쪽 열부터 [t-1], [t], [t+1] 프레임

Fig. 11. Example sequence images(rows) for practice swings. [t-1], [t], [t+1] frames from the left column

그림 12는 제안하는 방법이 골프공은 정확하게 검출하였지만, Impact를 잘못 검출한 예를 보여준다. 이 시퀀스에서 GT는 [t]와 [t+1] 프레임이지만, 제안하는 방법은 [t-1] 프레임을 Impact라고 판단하였다. 이것은 첫 번째 그림에서 보는 것과 같이, 골프공 영역 검출 결과가 노이즈로 인해서 실제 골프공보다 크게 추정되었고(빨간색 사각형) 이 영역에서의 클럽 움직임이 [t-1] 프레임에서 크게 나타났기 때문이다.

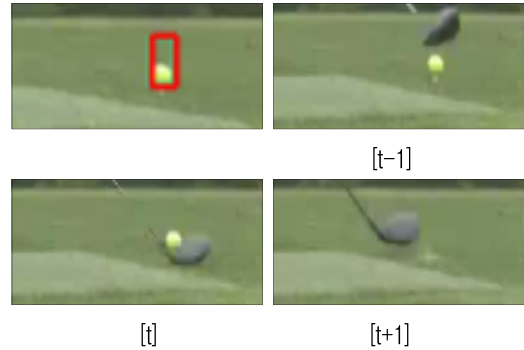


그림 12. Impact 검출 실패의 예. 골프공 검출 결과(첫 번째). Impact 순간[t]의 전-후 프레임

Fig. 12. Example of impact detection failure. golf ball detection result(first). frames before and after the impact[t] moment

4.4 Computational Time

제안하는 방법의 평균 연산 속도는 데스크탑 컴퓨터(intel i7-10700, only CUP) 환경에서 약 4초이다. 이것은 240장의 FHD 이미지로 구성된 하나의 시퀀스를 처리하는데 걸린 평균 시간으로 모든 이미지를 로드 하는 시간을 포함한 것이다. 여기에서, 이미지 로드 시간을 제외한 오직 알고리즘만의 평균 연산 시간은 약 1.4초이다. 이것은 제안하는 방법이 GPU 없이 CPU만으로 빠르게 고해상도 이미지들을 처리할 수 있는 것을 보여준다.

V. 결론 및 향후 과제

본 논문에서는 골프 플레이어의 동작 정보와 골프공의 움직임 정보를 이용해서 고해상도 단일 카메라 기반의 Impact를 검출하는 방법을 제안하였다. 제안하는 방법은 먼저 플레이어의 가려짐이 없는 시퀀스를 대상으로 골프공과 플레이어를 포함하는 ROI를 설정한다. 골프공 ROI에서 Blob들을 추출함으로써 골프공 후보들을 생성하고, 이 후보들로부터 확률적인 투표 방법을 통해서 최종적으로 골프공을 검출한다. 마지막으로 Impact는 골프공과 플레이어의 움직임을 함께 고려해서 검출한다. 실험을 통해서 우리는 제안하는 방법이 실외 환경에서 취득한 고해상도 이미지에서 노이즈에 강인하게 골프공과 Impact를 정확하게 검출할 수 있다는 것을 알 수 있다.

하지만, 제안하는 방법은 골프공과 비슷한 모양과 크기를 가지는 Tee 악세서리가 있는 환경에서는 정확도가 떨어지는 한계를 보여주었다. 앞으로는 골프공의 모양과 크기뿐만 아니라 추가적인 특징에 관한 연구와 플레이어의 움직임의 세부적으로 분석할 수 있는 연구를 진행할 계획이다.

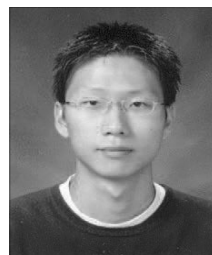
References

- [1] H. Kim, "The Golf Market Changing by COVID-19 and Propose Post-Pandemic Policy for Golf Industry : In Focus on Golf Club", *Journal of Golf Studies*, Vol. 16, No. 2, pp. 99-110, Jun. 2022. <https://doi.org/10.34283/ksgs.2022.16.2.9>.
- [2] J. Park and S. Kwak, "Golf Swing Classification Using Fuzzy System", *Journal of Broadcast Engineering*, Vol. 18, No. 3, pp. 380-392, May 2013. <https://doi.org/10.5909/jbe.2013.18.3.380>.
- [3] M. Ueda, H. Negoro, Y. Kurihara, and K. Watanabe, "Measurement of Angular Motion in Golf Swing by a Local Sensor at the Grip End of a Golf Club", in *IEEE Transactions on Human-Machine Systems*, Vol. 43, No. 4, pp. 398-404, Jun. 2013. <https://doi.org/10.1109/tsmc.2013.2266896>.
- [4] Y. L. Hsu, Y. T. Chen, P. H. Chou, Y. C. Kou, Y. C. Chen, and H. Y. Su, "Golf swing motion detection using an inertial-sensor-based portable instrument", *IEEE International Conference on Consumer Electronics-Taiwan*, Nantou, Taiwan, pp. 1-2, May 2016. <https://doi.org/10.1109/icce-tw.2016.7521016>.
- [5] M. Bourgain, P. Rouch, O. Rouillon, P. Thoreux, and C. Sauret, "Golf Swing Biomechanics: A Systematic Review and Methodological Recommendations for Kinematics", *Sports*, Vol. 10, No. 6, Jun. 2022. <https://doi.org/10.3390/sports10060091>.
- [6] B. Najafi, J. Lee-Eng, J. Wrobel, and R. Goebel, "Estimation of Center of Mass Trajectory using Wearable Sensors during Golf Swing", *Journal Sports Science and Medicine*, Vol. 14, No. 2, pp. 354-363, May 2015.
- [7] C. S. Ha, "The kinematic and kinetic variables of head up putting involved in golf", *Korea Journal Sports Science*, Vol. 17, No. 1, pp. 675-684, Feb. 2008.
- [8] L. Jiao, R. Bie, H. Wu, Y. Wei, J. Ma, A. Umek, and A. Kos, "Golf swing classification with multiple deep convolutional neural networks", *International Journal of Distributed Sensor Networks*, Vol. 14, No. 10, Oct. 2018. <https://doi.org/10.1177/1550147718802186>.
- [9] M. Kim and S. Park, "Golf Swing Segmentation from a Single IMU Using Machine Learning", *Sensors*, Vol. 20, No. 16, Aug. 2020. <https://doi.org/10.3390/s20164466>.
- [10] S. Park, J. Y. Chang, H. Jeong, J. Lee, and J. Park, "Accurate and efficient 3d human pose estimation algorithm using single depth images for pose analysis in golf", In *Proc. of the IEEE conference on computer vision and pattern recognition workshops*, pp. 49-57, 2017. <https://doi.org/10.1109/cvprw.2017.19>.
- [11] W. McNally, K. Vats, T. Pinto, C. Dulhanty, J. McPhee, and A. Wong, "GolfdB: A video database for golf swing sequencing", In *Proce. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2019. <https://doi.org/10.1109/cvprw.2019.00311>.
- [12] S. H. Hwang, K.-R. Ko, and S. B. Pan, "Golf Swing Analysis System Using Camera and Inertial Sensor", *The Journal of Korean Institute of Information Technology*, Vol. 15, No. 4, pp. 139-147, Apr. 2017. <https://doi.org/10.14801/jkiit.2017.15.4.139>.
- [13] L. Jiao, H. Wu, R. Bie, A. Umek, and A. Kos, "Multi-sensor Golf Swing Classification Using Deep CNN", *Procedia Computer Science*, Vol. 129, pp. 59-65, 2018. <https://doi.org/10.1016/j.procs.2018.03.046>.

- [14] A. Graves and J. Schmidhuber, "Framewise phoneme classification with bidirectional LSTM and other neural network architectures", *Neural Networks*, Vol. 18, No. 5-6, pp. 602-610, Aug. 2005. <https://doi.org/10.1016/j.neunet.2005.06.042>.
- [15] C. C. Liao, D. H. Hwang, and H. Koike, "AI Golf: Golf Swing Analysis Tool for Self-Training", in *IEEE Access*, Vol. 10, pp. 106286-106295, Oct. 2022. <https://doi.org/10.1109/access.2022.3210261>.
- [16] Y. Tang, L. Zhao, Z. Yao, C. Gong, and J. Yang, "Graph-based motion prediction for abnormal action detection", *Proc. 2nd ACM Int. Conf. Multimedia Asia*, No. 63, pp. 1-7, Mar. 2021. <https://doi.org/10.1145/3444685.3446316>.
- [17] J. H. Yoo, S. W. Lee, S. K. Park, and D. H. Kim, "A robust lane detection method based on vanishing point estimation using the relevance of line segments", *IEEE Transactions on Intelligent Transportation Systems*, Vol. 18, No. 12, pp. 3254-3266, Dec. 2017. <https://doi.org/10.1109/tits.2017.2679222>.
- [18] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection", In *Proc. of the IEEE conference on computer vision and pattern recognition*, pp. 779-788, 2016. <https://doi.org/10.1109/cvpr.2016.91>.
- [19] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement", *arXiv preprint arXiv:1804.02767*, Apr. 2018. <https://doi.org/10.48550/arXiv.1804.02767>.
- [20] F. Banterle, M. orsini, P. Cignoni, and R. Scopigno, "A Low-Memory, Straightforward and Fast Bilateral Filter Through Subsampling in Spatial Domain", *Computer Graphics Forum*, Vol. 31, No. 1, pp. 19-32, Dec. 2011. <https://doi.org/10.1111/j.1467-8659.2011.02078.x>.
- [21] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images", *Sixth International Conference on Computer Vision, Bombay, India*, pp. 839-846, Jan. 1998. <https://doi.org/10.1109/iccv.1998.710815>.
- [22] S. Mallick, "Blob Detection Using OpenCV (Python, C++)", <https://learnopencv.com/blob-detection-using-opencv-python-c> [accessed: Oct. 12, 2022]
- [23] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision", *Proc. of Imaging Understanding Workshop*, pp. 121-130, Apr. 1981.
- [24] D. L. Bruce, "Generalized Image Matching by the Method of Differences", *Carnegie Mellon University*, 1985.
- [25] Z. Zivkovic, "Improved adaptive Gaussian mixture model for background subtraction", *Proc. of the 17th International Conference on Pattern Recognition, Cambridge, UK*, Vol. 2, pp. 28-31, Aug. 2004. <https://doi.org/10.1109/icpr.2004.1333992>.
- [26] Z. Zivkovic and F. V. D. Heijden, "Efficient adaptive density estimation per image pixel for the task of background subtraction", *Pattern recognition letters*, Vol. 27, No. 7, pp. 773-780, May 2006. <https://doi.org/10.1016/j.patrec.2005.11.005>.

저자소개

유 주 한 (Ju Han Yoo)



2005년 : 경희대학교
전자정보학부(공학사)
2019년 : 고려대학교
컴퓨터전파통신공학과(공학박사)
2022년 ~ 현재 : 세명대학교
컴퓨터학부 조교수
관심분야 : 영상처리, 패턴인식,
머신러닝, 로봇 비전, 컴퓨터 그래픽스