

Tennis Player Tracking and Distance Measurement Using DeepSORT and Homography

Hyun-Il Kim*, Seung-Bo Park**

*Student, Dept. of Software Convergence Engineering, Inha University, Incheon, Korea

**Professor, Dept. of Software Convergence Engineering, Inha University, Incheon, Korea

[Abstract]

High-cost modern sports analysis systems are predominantly utilized by elite athletes, creating a significant technology gap that limits access for amateur players and sports enthusiasts. To address this issue, this study proposes a practical method for measuring the movement distance of tennis players using only a single, commonly available camera. Our approach combines YOLOv8 and DeepSORT for robust player detection and tracking, followed by a homography transformation based on court keypoints to map 2D image coordinates to real-world metric coordinates. A Kalman filter is then applied to smooth the resulting trajectory and derive a reliable metric of movement distance. When tested on real tennis match footage, the proposed system achieved an average measurement accuracy of 85.26% relative to manual analysis tools, demonstrating its viability as a low-cost alternative.

▶ **Key words:** Tennis Analysis, DeepSORT, Homography, Kalman Filter, Computer Vision

[요 약]

현대 스포츠 분석 시스템은 고비용 문제로 엘리트 선수 중심으로 활용되어, 아마추어 선수나 동호인이 접근하기 어려운 기술 격차 문제가 존재한다. 본 연구는 이러한 문제를 해결하고자 일반적인 단일 카메라 영상만으로 테니스 선수의 이동 거리를 측정하는 방법을 제안한다. YOLOv8과 DeepSORT를 결합하여 선수를 안정적으로 탐지 및 추적하고, 코트 특징점 기반의 호모그래피 변환을 통해 2D 영상 좌표를 실제 코트의 미터(m) 단위 좌표로 변환한다. 이후, 칼만 필터를 통해 궤적을 보정하여 이동 거리에 대한 지표를 산출한다. 실제 테니스 경기 영상에 적용한 결과, 수동 분석 도구 대비 평균 85.26%의 이동 거리 측정 정확도를 달성하였다.

▶ **주제어:** 테니스 분석, DeepSORT, 호모그래피, 칼만 필터, 컴퓨터 비전

-
- First Author: Hyun-Il Kim, Corresponding Author: Seung-Bo Park
 - *Hyun-Il Kim (akadsam@inha.edu), Dept. of Software Convergence Engineering, Inha University
 - **Seung-Bo Park (molaal@inha.ac.kr), Dept. of Software Convergence Engineering, Inha University
 - Received: 2025. 09. 10, Revised: 2025. 10. 01, Accepted: 2025. 10. 22.

I. Introduction

현대 스포츠 분야는 데이터 과학 및 컴퓨터 비전 기술의 발전으로 인하여 스포츠 분석의 지평을 넓히고 있다. 특히, 테니스와 같이 역동적인 움직임이 요구되는 종목에서는 선수의 물리적 성능을 정량적으로 평가하려는 노력이 활발하다. 다수의 카메라를 통해 촬영된 영상을 기반으로 딥러닝 기반의 포즈 추정 및 동작 인식 연구를 통해 서브스타일 구분 및 샷의 위치를 예측하거나[1], UWB(Ultra-Wideband) 추적 시스템을 이용하여 선수들의 이동 거리 및 가속도를 측정하는 연구가 진행되었다[2]. 그러나 이러한 시스템은 고가의 장비, 복잡한 설치 등 높은 진입 장벽을 가지고 있다. 이로 인해 프로 선수가 아닌 아마추어 선수나 생활 체육 동호인들은 자신의 훈련 데이터를 객관적으로 분석하는 데 한계가 존재한다. 즉, 데이터 기반 스포츠 분석의 혜택이 일부 엘리트 계층에 집중되는 '기술 격차' 문제가 존재하는 것이다. 이러한 격차를 메우기 위해 저비용 영상 분석 도구인 Kinovea[3]와 같은 수동 분석 도구가 존재하지만, 분석가가 직접 원근을 보정하고 선수 추적 실패 시 프레임 단위로 보정해야 하는 수동 방식에 의존하고 있어 상당한 시간과 노력이 요구된다.

본 연구는 이러한 문제의식에서 출발하여 단일 카메라 영상만을 활용하여 테니스 선수의 이동 거리를 측정하는 방법을 제안하고자 한다. 단일 카메라 시점에서는 원근 왜곡이 발생하기 때문에 픽셀(pixel) 단위의 이동을 실제 미터(m) 단위의 거리로 변환하는 과정이 필요하다. 따라서 2D 영상 좌표를 실제 코트의 미터 단위 좌표로 변환하기 위해 호모그래피(Homography) 기반 위치 보정을 적용하고 영상 속 선수를 안정적으로 추적하기 위해 DeepSORT 알고리즘을 사용한다. 마지막으로, 추적 과정에서 발생할 수 있는 측정 노이즈를 최소화하고 부드러운 이동 궤적을 얻기 위해 칼만 필터(Kalman Filter)를 활용한다.

II. Related works

스포츠 분야에서 선수의 활동량을 정량적으로 분석하려는 연구는 크게 센서 기반 방식과 비전 기반 방식으로 나뉜다. 본 장에서는 주요 연구 동향을 분석하고, 이를 측정하기 위한 다양한 기법들을 소개한다.

2.1 Sensor Based Analysis

스포츠 분야에서 선수들의 활동량을 측정하는 대표적인 방법으로는 GPS(Global Positioning System)와 GNSS(Global Navigation Satellite System)를 사용하는 방법이 있다. 축구의 경우 운동 동작을 분석하기 위해 GNSS 및 IMU 기반 EPTS 추적 시스템을 사용한다[4]. 하지만 테니스의 경우, 상대적으로 좁은 경기장에서 경기가 이루어지며 주변에 구조물이나 지붕이 있는 실내 코트도 많아 GNSS의 신호 수신에 제한될 수 있다.

UWB를 활용하는 센서 기반 방식은 이동 거리, 속도, 방향 전환, 가속도 등 높은 정밀도로 측정할 수 있는 장점이 있다[2][5]. 하지만 시스템 구축 및 유지보수의 어려움으로 인해 일반 사용자가 사용하기 어렵다는 문제점이 존재한다. 또한, 선수의 움직임을 정확히 분석하기 위해서는 일정 수 이상의 센서 설치와 체계적인 환경 구성이 필요하다는 점에서 실사용에 제약이 따른다.

이러한 한계를 극복하기 위해 컴퓨터 비전 기술을 활용한 분석 연구가 활발히 진행되고 있다. 특히 별도의 장비 착용 없이 영상만으로 분석이 가능하다는 점에서 접근성이 뛰어나다.

2.2 Computer-Vision Based Analysis

테니스 분야의 비전 기반 분석은 주로 다중 카메라 시스템을 통해 이루어졌다. 여러 각도에서 촬영된 영상을 종합하여 선수의 동작을 인식하거나, 샷의 종류 및 위치를 예측하는 연구가 진행되었다[1]. 그러나 다수의 카메라를 사용하기 때문에, 일반적인 환경에 적용하기에는 어려움이 따른다. 이에 단일 카메라를 활용한 연구들이 대안으로 제시되었다. MoveNet을 이용한 포즈 추정을 통해 테니스 포핸드 스트로크를 분석하거나[6] Alpha Pose 및 Pose Flow를 이용하여 프로 선수와 아마추어 선수의 서브를 분석하는 연구가 진행되었다[7]. 하지만 이러한 연구는 '동작'을 분석하는 데 그칠 뿐, 실제 코트 위에서 얼마나 움직였는지를 나타내는 이동 거리와 같은 활동량 데이터를 산출하는 데까지는 이르지 않았다.

테니스 경기에서 선수의 총 이동 거리(Total Distance)는 경기력과 피로 누적, 부상 위험 등에 직결되기 때문에 중요한 지표로 여겨진다. 선수의 에너지 소모량과 체력 소진 정도는 단일 경기에서뿐만 아니라 토너먼트 기간 동안 누적된 거리가 경기 후반부 성과에 큰 영향을 미칠 수 있다. 이처럼 테니스에서는 단순한 동작 분석을 넘어, 실제 경기 중에 발생하는 거리 데이터를 수집하고 분석하는 것

이 선수의 체력 관리, 경기 전략 수립 등에 실질적인 가치를 제공한다. 그럼에도 불구하고, 활동량 데이터를 추출하는 과정은 여전히 어려움이 따른다. 예를 들어, Kinovea와 같은 동작 분석 도구는 영상 내 선수를 추적하여 이동 거리 및 각도 등의 데이터를 산출하는 기능을 제공하지만, 자동 추적 성능의 한계로 인해 정확한 측정을 위해 분석가가 프레임마다 수동으로 위치를 보정해야 한다. 이는 경기 전체를 분석하기에는 매우 비효율적인 방식이며 시간이 많이 소요되는 한계가 존재한다. 따라서 이러한 한계를 극복하고, 중요성이 높은 선수 활동량 데이터를 실용적으로 추출하기 위해서는 선수를 자동으로 감지하고 추적하는 기술이 요구된다.

2.3 Background Technologies

본 연구는 단일 카메라 환경에서 선수의 2D 영상 좌표를 실제 코트의 미터 단위 궤적으로 변환하기 위해, 선수 탐지, 추적, 코트 검출, 시점 변환, 궤적 보정을 순차적으로 수행하는 파이프라인을 제안한다. 본 연구는 각 단계를 테니스 분석 환경에 맞게 통합하여, 선수의 실제 이동 거리를 산출하는 것을 목표로 한다.

2.3.1 YOLO(You Only Look Once)

선수의 위치를 실시간으로 파악하기 위한 첫 단계로, 대표적인 1단계(Single-Stage) 객체 탐지 모델인 YOLOv8n을 사용한다[8]. YOLO는 입력 이미지를 그리드(Grid)로 분할하고, 각 그리드 셀에서 직접 바운딩 박스(Bounding Box)의 좌표와 클래스 확률(Class Probability)을 예측하는 독창적인 구조를 가진다. 이 방식은 후보 영역을 먼저 제안하고 분류를 수행하는 2단계(Two-Stage) 모델보다 연산 속도가 월등히 빨라 비디오 스트림과 같은 실시간 환경에서 선수 탐지를 수행하는 데 매우 효과적이다.

2.3.2 DeepSORT

테니스 경기 영상에서는 선수가 네트나 다른 선수에 의해 가려지는 폐색(Occlusion) 상황이 빈번하게 발생하며, 이로 인해 탐지된 바운딩 박스의 ID가 불안정하게 변경될 수 있다. 이러한 문제를 해결하고 선수를 일관성 있게 추적하기 위해 본 연구는 DeepSORT 알고리즘을 적용한다[9]. DeepSORT는 추적의 정확성과 속도를 모두 고려한 SORT(Simple Online and Realtime Tracking) 알고리즘을 기반으로 한다. SORT는 칼만 필터를 사용하여 객체의 다음 위치를 예측하고, 탐지된 바운딩 박스와의 IoU(Intersection over Union)를 기반으로 데이터를 연결

한다. 여기에 DeepSORT는 심층 신경망(DNN)을 통해 추출된 외형 특징(Apppearance Feature) 정보를 추가로 활용한다. DeepSORT의 추적 절차는 먼저 칼만 필터의 예측과 마할라노비스 거리를 이용해 움직임이 유사한 객체를 연결하고, 여기서 연결되지 않은 객체에 대해서는 외형 특징 간의 코사인 거리를 비교하여 재연결을 시도한다. 즉, 객체가 가려졌다가 다시 나타났을 때, 칼만 필터의 예측만으로 구별이 어려운 경우에도 외형 정보를 비교하여 동일한 ID를 부여함으로써 추적의 연속성을 강건하게 유지한다. 이는 선수의 움직임이 빠르고 복잡한 테니스 환경에서 추적 실패율을 크게 낮출 수 있다. DeepSORT 기반 객체 추적 기술은 스포츠 경기에서 선수의 실시간 움직임 분석[10], 교통 감시에서 차량의 이동 추적[11] 등 다양한 분야에 적용되어, 객체 추적의 핵심 기술로 활용되고 있다.

2.3.3 Homography

호모그래피는 하나의 평면을 다른 관점의 평면으로 투영시키는 변환 행렬을 계산하는 기법이다[12]. 본 연구에서는 이를 이용해 카메라에 비스듬하게 촬영된 코트 평면을 실제 코트 규격에 맞는 조감도(Bird's-eye view) 시점으로 변환한다. 앞서 검출된 영상 속 코트의 특징점(2D 좌표)과 국제 테니스 연맹(ITF) 표준 규격에 따른 실제 코트의 특징점(3D 월드 좌표) 간의 대응 관계를 이용해 3x3 호모그래피 행렬(H)을 계산한다.

이 행렬(H)을 선수의 좌표에 적용하면, 2D 픽셀 위치를 실제 코트 평면상의 미터 단위 좌표로 정밀하게 변환할 수 있다. 이 기법은 평면으로 간주할 수 있는 축구장이나 농구 코트 분석에서 선수 위치를 보정한 연구가 존재한다[13][14]. 본 연구에서는 이를 테니스 코트에 적용하여 거리 측정의 신뢰도를 확보하고자 한다.

2.3.4 Kalman Filter

본 연구에서는 최종 이동 궤적의 안정성을 높이기 위해 칼만 필터를 후처리 과정에 한 번 더 적용한다. 이는 DeepSORT 내부에서 이미지 좌표계(pixel)를 기준으로 객체를 추적하는 칼만 필터와는 별개로, 호모그래피 변환까지 마친 실제 좌표계 상의 궤적을 보정하는 역할을 한다.

객체 탐지 및 추적 과정에서 발생하는 미세한 측정 오류나 프레임 간의 불확실성은 이동 궤적에 노이즈를 유발하여 떨림 현상을 야기한다. 이러한 노이즈를 효과적으로 제거하고 물리적으로 가능한 부드러운 이동 궤적을 생성하기 위해 칼만 필터를 사용한다[15]. 칼만 필터는 선형 동적 시스템에 대한 재귀적 필터로, 예측(Prediction)과 업데이트

트(Update) 두 단계를 반복적으로 수행한다. 예측 단계에서는 이전 상태(위치, 속도)를 바탕으로 현재 시점의 상태를 예측하고 업데이트 단계에서는 탐지 모델(YOLO)과 추적기(DeepSORT)를 통해 새로 측정된 값을 이용해 예측값을 보정한다. 이 과정을 통해 측정값에 포함된 노이즈의 영향을 최소화하고, 일부 프레임에서 선수를 탐지하지 못하더라도 예측을 통해 궤적을 일정 부분 유지할 수 있다. 그 결과, 최종적으로 생성되는 선수의 이동 궤적은 한층 더 안정적이고 신뢰도가 높아진다.

III. Player Distance Estimation Method

3.1 System Architecture Design

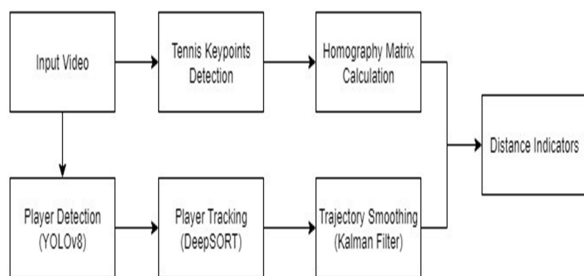


Fig. 1. System Architecture

본 연구에서 제안하는 방법의 전체 흐름도는 Fig. 1과 같다. 선수 검출 및 추적, 코트 특징점 검출, 호모그래피 행렬 계산, 궤적 보정, 이동 거리 산출 다섯 단계로 구성된다.

3.2 Players Detection and Tracking

선수 검출 단계에서는 실시간 객체 검출 성능이 뛰어난 YOLOv8n 모델을 사용한다. 매 비디오 프레임이 입력되면, YOLO 모델은 프레임 내의 모든 객체에 대한 바운딩 박스, 신뢰도 점수(Certainty Score), 클래스 정보를 반환한다. 하지만 YOLO는 두 선수가 잠시 겹쳤다가 떨어지는 경우, 각 선수에게 할당된 ID가 뒤바뀌는 문제가 발생할 수 있다. 이러한 ID 스위칭 문제를 해결하고 각 선수에게 일관된 고유 ID를 부여하기 위해, 추적 성능이 검증된 DeepSORT 알고리즘을 적용하였다. 본 연구에서는 선수가 60프레임(1초) 동안 재관측되지 않으면 해당 트랙을 소실(lost) 처리하도록 하였다. 또한, 추적의 안정성을 위해 DeepSORT가 '확정(Confirmed)' 상태로 판단한 트랙만을 유효한 선수 데이터로 간주하였다. 이 과정을 통해 최종적으로 각 프레임마다 고유 ID가 할당된 선수들의 바운딩 박스 정보를 획득하며,

이 정보는 다음 3.4절에서 선수의 발 위치를 특정하고 좌표를 변환하는 데 핵심적인 입력으로 사용된다.

3.3 Tennis Keyoints Detection

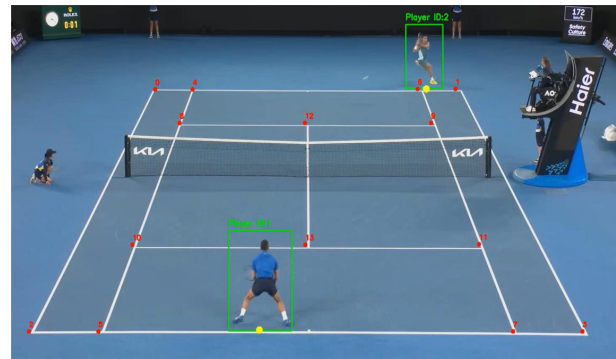


Fig. 2. Tennis Keyoints Detection

영상 내 테니스 코트를 인식하기 위해, 사전 훈련된 TrackNet 아키텍처를 기반으로 하는 테니스 코트 특징점 검출 모델을 사용하였다[16][17]. 해당 모델은 3개의 연속 프레임을 입력으로 사용하는 원본 TrackNet과 달리, 단일 프레임만을 입력받아 코트의 주요 특징점 14개에 대한 좌표를 출력으로 갖도록 재구성한 것이다. 시스템 초기화 단계에서 비디오의 첫 번째 프레임이 입력되면, 모델은 해당 프레임에서 Fig. 2와 같이 코트의 외곽선과 주요 라인이 교차하는 14개의 핵심 포인트를 픽셀 단위로 검출한다. 이 특징점들은 영상 속 코트의 원근과 형태를 정의하는 기준점 역할을 한다.

3.4 Homography Matrix Calculation

호모그래피를 통해 2D 영상 평면에 있는 점을 실제 코트의 2D 평면도(Top-down View)로 맵핑한다. 호모그래피 행렬을 계산하기 위해서는 최소 4쌍의 대응점(Corresponding Points)이 필요하다. 이를 위해 영상 평면의 소스 포인트(Source Points)와 실제 코트 평면의 목표 포인트(Destination Points), 두 점 집합을 사용한다. 소스 포인트로는 3.3절에서 검출된 코트의 네 모서리 픽셀 좌표를, 목표 포인트로는 국제 테니스 연맹(ITF) 규격에 따른 실제 코트의 네 모서리 미터(m) 단위 좌표([0, 0], [10.97, 0], [0, 23.77], [10.97, 23.77])를 각각 사용한다. 이 두 점 집합을 기반으로 소스 포인트를 목표 포인트로 변환하는 호모그래피 행렬을 계산한다. 이렇게 생성된 단일 행렬은 선수의 발 위치 픽셀 좌표를 실제 코트의 미터 단위 좌표로 변환하는 데 사용된다. 이 과정을 통해 카메라의 위치나 각도에 관계없이 일관된 실제 좌표계를 획득할 수 있다.

3.5 Trajectory Smoothing

추적된 바운딩 박스로부터 산출된 선수 발 위치 좌표는 탐지 알고리즘의 한계와 선수의 동적인 자세 변화로 인해 상당한 떨림(jitter) 노이즈를 포함한다. 이러한 불확실성을 효과적으로 처리하기 위해, 본 연구에서는 칼만 필터(Kalman Filter)를 적용하여 좌표를 보정(Smoothing)하였다.

필터 설계를 위해, 추적 대상인 선수의 상태 변수 x_k 는 k 시점에서의 2D 평면상 위치 (p_x, p_y) 와 각 축의 속도 (v_x, v_y) 로 구성된 4차원 벡터로 정의하였다.

$$x_k = \begin{bmatrix} p_x \\ p_y \\ v_x \\ v_y \end{bmatrix}_k \quad (1)$$

테니스 선수의 움직임이 짧은 시간 내에서는 관성을 유지한다는 점에 착안하여 시스템 모델은 계산 효율성과 안정성이 높은 등속도 운동(Constant Velocity Model)으로 가정하였으며, 이에 따른 상태 전이 모델은 식(2)와 같다. 여기서 A 는 상태 전이 행렬, Δt 는 영상의 프레임을 반영한 실제 시간 간격이다.

$$x_k = Ax_{k-1} = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ v_x \\ v_y \end{bmatrix}_{k-1} \quad (2)$$

YOLOv8n과 DeepSORT를 통해 실제로 측정되는 값은 선수의 위치 (p_x, p_y) 뿐이므로, 관측 모델은 식(3)과 같이 정의하였다. 여기서 H 는 관측 행렬이다.

$$z_k = Hx_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ v_x \\ v_y \end{bmatrix}_k \quad (3)$$

칼만 필터의 성능은 잡음 공분산 행렬 설정에 크게 의존한다. 프로세스 잡음 공분산 Q 는 정의된 운동 모델의 불확실성을, 측정 잡음 공분산 R 은 측정값의 노이즈 크기를 나타낸다. 본 연구에서는 바운딩 박스 좌표의 측정 노이즈가 선수의 실제 움직임 변화보다 크다고 판단하여, 필터가 불안정한 측정값보다 내부 모델의 예측을 더 신뢰하도록 R 값을 Q 값보다 상대적으로 크게 설정하였다. 이 과정을 통해 노이즈가 제거된 안정적인 미터 좌표를 획득하였으며, 이를 기반으로 최종 이동 거리를 산출하였다.

3.6 Distance Indicators

이동 거리는 매 프레임 간 위치 변화를 유클리드 거리로 계산하여 누적 합산한다.

$$D_{total} = \sum_{i=2}^N \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2} \quad (4)$$

여기서 i 번째 프레임에서 선수의 실제 좌표를 의미한다. 단, 추적 오류로 인해 선수가 순간적으로 멀리 이동한 것으로 측정되는 경우를 방지하기 위해, 프레임 간 이동 거리가 특정 임계값(1m)을 초과할 경우 누적 계산에서 제외하였다.

IV. Experiment & Discussion

본 장에서는 3장에서 제안한 선수 이동 거리 측정 방법을 검증하고, 실제 테니스 경기 영상에 적용하여 도출한 분석 결과를 제시한다.

4.1 Experimental Environment

본 연구의 성능 평가는 실제 테니스 중계 영상을 대상으로 진행하였다. 분석에는 주요 남자 단식 경기의 일부를 사용하였으며, 상세 내용은 Table 1과 같다. 제안하는 기법의 안정적인 특징점 검출 및 Kinovea 분석을 위해, 카메라 움직임이 없고 코트의 네 꼭짓점을 포함한 전체 영역이 명확하게 보이는 고정된 시점의 영상만을 사용하였다. 이로 인해 코트 일부만 보이는 측면이나 선수를 따라가는 시점의 영상은 분석에서 제외되었다. 3초 이상 랠리가 지속된 인플레이(in-play) 상황으로 한정하였으며, 서브에스나 폴트처럼 랠리가 발생하지 않은 포인트와 포인트 사이의 휴식 시간은 분석에서 제외하였다. 각 분석 영상은 이러한 조건을 만족하는 장면들로 약 5분 분량으로 구성되었다. 본 연구의 실험은 GPU NVIDIA GeForce GTX 1660 SUPER 사양의 Windows 11 환경에서 수행되었으며, 5분 분량의 영상을 분석하는 데 평균 11분이 소요되었다. 수동 분석에는 Kinovea 2023.1.2. 버전을 사용하였다.

4.2 Performance Evaluation

제안 방법의 이동 거리 측정 정확도를 평가하기 위해 수동 비디오 분석 도구인 Kinovea를 사용하여 측정한 값을 Ground Truth, 제안 방법으로 측정된 값을 Proposed System으로 설정하였다. Kinovea는 수동 분석 도구이므로 분석자의 주관적 오차가 발생할 수 있다. 본 연구는 이러한 변동성을 줄이고 데이터의 신뢰도를 확보하고자, 동일 분석가가 각 영상을 3회씩 측정하였다. 최종적으로 세 측정값의 산술 평균을 해당 구간의 실제 이동 거리로 사용하였다. 정확도는 아래 수식(5)을 사용하여 계산하였다.

Table 1. Video Information Used for Analysis

Match	Players	Resolution	FPS
2022 Cincinnati R16	Taylor Fritz vs Andrey Rublev	1920x1080	60
2022 French Open SF	Rafael Nadal vs Alexander Zverev	1920x1080	60
2025 Australian Open QF	Novak Djokovic vs Carlos Alcaraz	1920x1080	60
2025 National Bank Open Final	Ben Shelton vs Karen Khachanov	1920x1080	60

Table 2. Per-Player Results of Measured Distance and Accuracy

Player	Ground Truth (Kinovea)	Proposed System	Accuracy (%)
Taylor Fritz	690.99 m	801.77 m	83.96
Andrey Rublev	676.00 m	733.09 m	91.55
Rafael Nadal	588.35 m	667.75 m	86.50
Alexander Zverev	534.82 m	626.00 m	82.95
Novak Djokovic	613.54 m	734.15 m	80.34
Carlos Alcaraz	637.07 m	768.62 m	79.35
Ben Shelton	624.11 m	686.14 m	90.06
Karen Khachanov	625.52 m	704.51 m	87.37

$$Accuracy (\%) = \left(1 - \frac{|D_{GT} - D_{PS}|}{D_{GT}}\right) \cdot 100 \quad (5)$$

여기서 D_{GT} 는 Kinovea로 측정된 이동 거리를, D_{PS} 는 제안 방법으로 측정된 이동 거리를 의미한다. 총 8명의 선수에 대한 이동 거리 측정 결과 및 정확도는 Table 2와 같다. 실험 결과, 제안 방법은 평균 85.26%의 이동 거리 측정 정확도를 보였다. Andrey Rublev의 경우 91.55%로 가장 높은 정확도를 기록했으며, Carlos Alcaraz의 경우 79.35%로 가장 낮은 정확도를 나타냈다. 이러한 정확도의 편차는 주로 선수 탐지 단계에서 발생하는 바운딩 박스의 불안정성에 기인하는 것으로 분석된다. 프레임 간 발생하는 바운딩 박스의 미세한 떨림이나 선수의 역동적인 자세 변화에 따른 형태 변형이 누적되어 최종 이동 거리 계산에 오차를 유발하는 것이다. 그럼에도 불구하고, 시스템이 별도의 장비 없이 단일 영상만으로 선수의 이동 거리를 고가의 장비나 수작업에 드는 시간 비용 없이 유의미한 수준으로 측정할 수 있음을 확인하였다.

V. Conclusions

본 연구는 고가의 장비나 다중 카메라 시스템 없이, 일반적인 단일 카메라 방송 영상만으로 테니스 선수의 이동 거리를 측정하는 방법을 제안하였다. 이를 위해 YOLOv8n과 DeepSORT를 결합하여 선수를 탐지 및 추적하였고, 코트 특징점 검출과 호모그래피 변환을 통해 2D 영상 좌표를 실제 코트의 미터(m) 단위 좌표로 변환하였다. 마지막으로, 변환된 좌표에 칼만 필터를 적용하여 노이즈를 제거

하고, 이동 거리를 산출하였다. 실험 결과, 제안 방법은 수동 분석 도구인 Kinovea를 기준으로 평균 85.26%의 측정 정확도를 달성하였다. 이는 고가의 장비나 수작업 없이도 신뢰할 만한 수준의 활동량 데이터를 추출할 수 있음을 의미한다. 따라서 본 연구에서 제안한 방법은 기존 수동 분석 방식의 비효율성을 극복하고, 누구나 쉽게 선수 데이터를 분석할 수 있는 실용적인 대안이 될 수 있다는 것을 검증하였다.

본 연구는 8개의 경기 영상으로 실험을 진행하였기에 데이터셋의 다양성과 규모 측면에서 한계가 존재한다. 향후 연구에서는 더 많은 경기와 다양한 촬영 조건의 영상을 확보하여 제안 기법의 일반화 성능을 검증할 필요가 있다. 또한, 성능 평가는 Kinovea를 기준으로 한 총 이동 거리의 상대 정확도라는 단일 지표에 한정되었다. 향후 연구에서는 평균 절대 오차(MAE)나 평균 제곱근 오차(RMSE)와 같은 통계적 오차 지표를 추가하고, 추적 안정성을 평가할 수 있는 MOTA(Multiple Object Tracking Accuracy) 같은 지표를 도입하여 성능을 다각적으로 분석할 필요가 있다. 아울러 본 연구는 선수의 발 위치를 바운딩 박스의 하단 중앙점으로 간주한다. 따라서 선수의 자세 변화에 따라 바운딩 박스의 형태가 변하면서 발의 위치를 추정하는데 오차가 발생한다는 한계점이 존재한다. 향후 연구에서는 인체 포즈 추정(Human Pose Estimation) 기술을 도입하여 바운딩 박스의 하단 중앙점이 아닌, 선수의 발목 키포인트를 직접 검출하고 추적한다면, 선수의 위치 정확도를 향상시킬 수 있을 것이다. 나아가, 이렇게 확보된 좌표를 기준으로 선수의 순간 속도, 가속 및 감속, 방향 전환 패턴 등을 분석하여 선수의 성능과 경기 스타일을 깊이 있게 이해하는 연구로 발전시킬 수 있을 것이다.

ACKNOWLEDGEMENT

This work was supported by the National Research Foundation of Korea (NRF) funded by Korean Government [Ministry of Science and ICT (MSIT)] under Grant RS-2023-00252952.

REFERENCES

- [1] S. V. Mora, "Computer vision and machine learning for in-play tennis analysis: framework, algorithms and implementation," 2018. URL: <https://api.semanticscholar.org/CorpusID:198358283>
- [2] W. Yang, J. Wang, Z. Zhao, and Y. Cui, "Accuracy of an Ultra-Wideband-Based Tracking System for Time-Motion Analysis in Tennis," *Sensors*, Vol. 25, No. 4, p. 1031, 2025. DOI: 10.3390/s25041031
- [3] B. Pueo, A. Penichet-Tomas, and J. M. Jimenez-Olmedo, "Validity, reliability and usefulness of smartphone and kinovea motion analysis software for direct measurement of vertical jump height," *Physiology & Behavior*, Vol. 227, p. 113144, 2020. DOI: 10.1016/j.physbeh.2020.113144
- [4] H. Kim, J. Kim, Y.-S. Kim, M. Kim, and Y. Lee, "Energy-Efficient Wearable EPTS Device Using On-Device DCNN Processing for Football Activity Classification," *Sensors*, Vol. 20, No. 21, p. 6004, 2020. DOI: 10.3390/s20216004
- [5] A. Waqar, I. Ahmad, D. Habibi, and Q. V. Phung, "Analysis of GPS and UWB positioning system for athlete tracking," *Measurement: Sensors*, Vol. 14, p. 100036, 2021. DOI: 10.1016/j.measen.2020.100036
- [6] E. Cheng, R. Dange, and I. Gorelik, "MatchPoint: Your Computer Vision Tennis Coach," Stanford University CS231n Project Report, 2024. <https://cs231n.stanford.edu/2024/papers/matchpoint-your-computer-vision-tennis-coach>
- [7] L. Baily, N. Truong, J. Lai, and P. Nguyen, "Stroke Comparison between Professional Tennis Players and Amateur Players using Advanced Computer Vision," *Proceedings of the 8th International Conference on Sport Sciences Research and Technology Support - icSPORTS*, pp. 44-52, 2020. DOI: 10.5220/0010145800440052
- [8] M. Yaseen, "What is YOLOv8: An In-Depth Exploration of the Internal Features of the Next-Generation Object Detector," *arXiv preprint arXiv:2408.15857*, 2024. <https://arxiv.org/abs/2408.15857>
- [9] N. Wojke, A. Bewley, and D. Paulus, "Simple Online and Realtime Tracking with a Deep Association Metric," *arXiv preprint arXiv:1703.07402*, 2017. <http://arxiv.org/abs/1703.07402>
- [10] Y. Zhang, Z. Chen, and B. Wei, "A Sport Athlete Object Tracking Based on Deep Sort and Yolo V4 in Case of Camera Movement," 2020 IEEE 6th International Conference on Computer and Communications (ICCC), pp. 1312-1316, 2020. DOI: 10.1109/ICCC51575.2020.9345010
- [11] Z. Luo, Y. Bi, X. Yang, Y. Li, S. Yu, M. Wu, and Q. Ye, "Enhanced YOLOv5s + DeepSORT method for highway vehicle speed detection and multi-sensor verification," *Frontiers in Physics*, Vol. 12, 2024. DOI: 10.3389/fphy.2024.1371320
- [12] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 11, pp. 1330-1334, Nov. 2000. DOI: 10.1109/34.888718
- [13] S. Iwase and H. Saito, "Tracking soccer players based on homography among multiple views," *Visual Communications and Image Processing 2003*, pp. 283-292, 2003. DOI: 10.1117/12.502967
- [14] Y. Pandya, K. Nandy, and S. Agarwal, "Homography based Player Identification in Live Sports," 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 5209-5218, 2023. DOI: 10.1109/CVPRW59228.2023.00549
- [15] S. H. Jeevith and S. Lakshmikanth, "Detection and tracking of moving object using modified background subtraction and Kalman filter," *International Journal of Electrical and Computer Engineering (IJECE)*, Vol. 11, No. 1, pp. 217-223, Feb. 2021. DOI: 10.11591/ijece.v11i1.pp217-223
- [16] S. Kosolapov, "Tennis analysis using deep learning and machine learning," Medium. <https://medium.com/@kosolapov.aetp/tennis-analysis-using-deep-learning-and-machine-learning-a5a74db7e2ee>
- [17] Y.-C. Huang, I.-N. Liao, C.-H. Chen, T.-U. Ik, and W.-C. Peng, "TrackNet: A Deep Learning Network for Tracking High-speed and Tiny Objects in Sports Applications," *arXiv preprint arXiv:1907.03698*, 2019. <http://arxiv.org/abs/1907.03698>

Authors



Hyun-Il Kim is an undergraduate student of Dept. of Software Convergence Engineering at Inha University, Korea. He is interested in machine learning, computer vision, and data science.



Seung-Bo Park received the BS, M.S. in Electrical Engineering and Ph.D. degrees in Information Engineering from Inha University, Korea, in 1995, 1997 and 2011, respectively. His research interests include video story

analyzing, semantic contents, video knowledge representation, social network analysis, and A.I. He joined the faculty of the Department of Software Convergence Engineering at Inha University, Incheon, Korea, in 2017.