Enhancing Human Pose Estimation: Precise Feet Alignment Using 6-DoF Detection

Long Hoang, Department of Computer Science
Mentor: Dr. Hao Ji
RSCA Conference 2025

Abstract

Human pose estimation models, such as 4D Humans and GVHMR, have made significant strides in capturing and synthesizing realistic human movements in 3D models. However, these models often overlook the detailed orientation of the feet, resulting in unnatural animations and misaligned foot positioning. This research addresses this limitation by leveraging 6-Degrees-of-Freedom (6-DoF) detection from MediaPipe to correct foot alignment and integrate it with the body posture. Our approach focuses on detecting precise foot orientation (up/down, forward/back, left/right, pitch, roll and yaw) and aligning it with the body to produce more natural and realistic human motion. This ongoing work has successfully implemented foot detection and aims to refine the alignment process for seamless integration into motion models, with potential applications in gaming, virtual reality, robotics, and biomechanics.

Problem Statement

Current human motion models prioritize upper body and limb movements but neglect the detailed orientation of the feet. This oversight leads to several issues:

- Misaligned Feet: Feet often appear disconnected from the body posture, breaking the continuity of motion.
- Unnatural Motion: Poor foot orientation reduces the realism of animations, making movements look robotic or awkward.
- Limited Applications: Fields requiring precise foot positioning, such as virtual reality, gaming, robotics, and biomechanics, are hindered by these inaccuracies.

This research aims to solve these issues by developing a framework for precise foot alignment using 6-DoF detection, ensuring that foot orientation is accurately represented and aligned with the body posture.
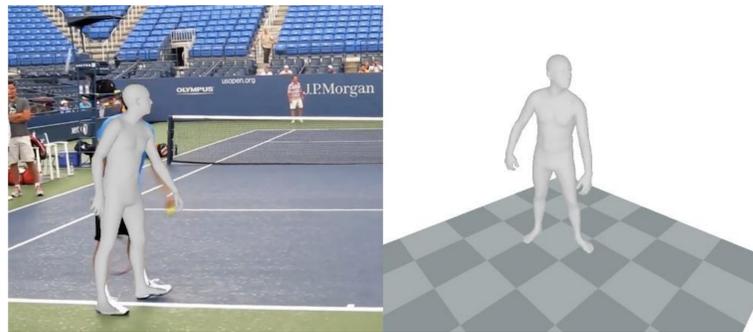

Figure 1: Human pose estimation output from GVHMR model.

Figure 2: Human pose estimation output from 4DHumans model.

Research Objective

The primary objective of this research is to enhance human motion models by correcting foot orientation and aligning it with the body posture. By leveraging 6-DoF detection, we aim to:

- Improve the realism of human motion in animations.
- Ensure accurate foot positioning relative to the body.
- Enable seamless integration of corrected foot data into the rest of the body posture.

Methodology

- 6-DoF Detection with MediaPipe:

+ Use MediaPipe's 6-DoF detection to capture precise foot orientation.
+ Track rotational (roll, pitch, yaw) and translational (x, y, z) axes for each foot.
+ Extract foot position and orientation data relative to the body.

- Feet Alignment Algorithm:

+ Analyze foot orientation data to identify misalignments.
+ Apply corrective transformations to align feet with the body posture.
+ Ensure that the corrected foot orientation matches natural human movement patterns.

- Integration with Human Pose Estimation Models:

+ Combine corrected foot data with human pose estimation models referencing from existing models to create a newly developed model focusing on the precise of the lower section of the human body.
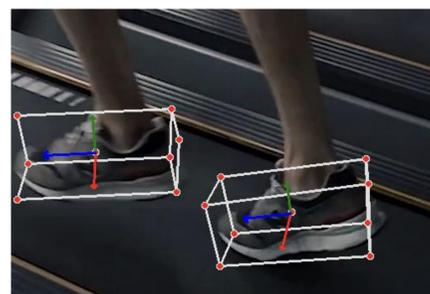+ Refine motion synthesis to ensure seamless and natural human-like motion.
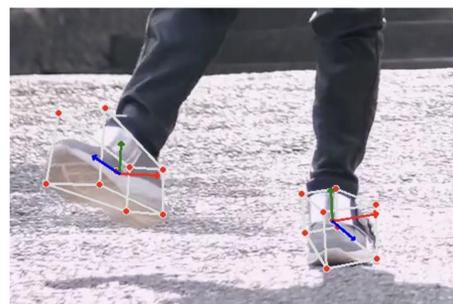

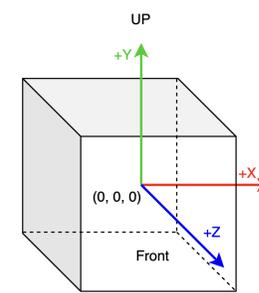Figure 3.1: Shoe detection output using MediaPipe.

Figure 4: Object Coordinate Systems for 6-DoF.
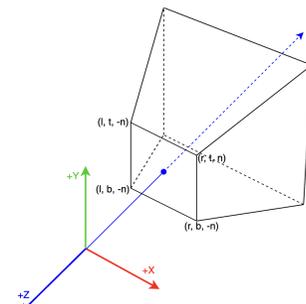

Figure 3.2: Shoe detection output using MediaPipe.

Figure 5: Camera Coordinate Systems for 6-DoF.

Progress

Ongoing Work:

+ Developing algorithms to align feet with body posture.

+ Mapping 2D skeleton-based body posture to 3D volume-based body model.

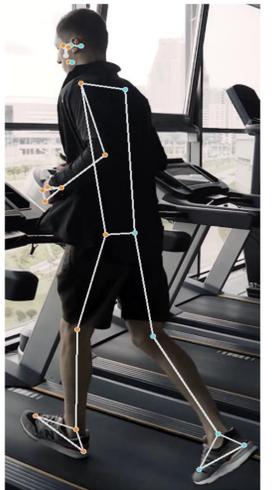+ Refining motion synthesis for realistic integration.

Figure 6: Posture detection.

Conclusion

This research addresses a critical gap in human motion modeling by focusing on precise foot alignment. By leveraging 6-DoF detection, we aim to enhance the realism and accuracy of human motion, paving the way for more natural and lifelike animations. Our work has the potential to significantly improve applications in gaming, virtual reality, robotics, and biomechanics by ensuring accurate and realistic foot positioning.

Future Work

- Validation: Test the model with diverse motion datasets to ensure robustness and generalizability.

- Complex Motions: Extend the framework to handle complex motions, such as running, dancing, and sports activities.

- Real-Time Integration: Integrate the framework with real-time motion capture systems for immediate feedback and corrections.

- User Applications: Develop user-friendly tools for animators and developers to incorporate precise foot alignment into their workflows.

References

- "Mediapipe Solutions Guide | Google Ai Edge | Google AI for Developers." Google, Google, ai.google.dev/edge/mediapipe/solutions/guide. Accessed 12 Feb. 2025.

- Goel, Shubham, et al. "Humans in 4D: Reconstructing and Tracking Humans with Transformers." arXiv.Org, 31 Aug. 2023, arxiv.org/abs/2305.20091.

- Shen, Zehong, et al. "World-Grounded Human Motion Recovery via Gravity-View Coordinates." arXiv.Org, 10 Sept. 2024, arxiv.org/abs/2409.06662.

Acknowledgments