

# Movement tracking and action classification for human behaviour under threat in virtual reality

Ulises Daniel Serratos Hernandez <sup>1</sup>, Jack Brookes <sup>1</sup>, Samson Hall <sup>1</sup>, Juliana K. Sporrer <sup>1</sup>, Sajjad Zabbah <sup>1</sup>, Dominik R. Bach <sup>1 2</sup>

<sup>1</sup> University College London, Max Planck UCL Centre for Computational Psychiatry and Ageing Research and Wellcome Centre for Human Neuroimaging- UCL Queen Square Institute of Neurology, London, United Kingdom

<sup>2</sup> University of Bonn, Hertz Chair for Artificial Intelligence and Neuroscience, Bonn, Germany

## Introduction

Understanding and characterising human movements is complex due to the diversity of human actions and their inherent inter, intra, and secular variability. Traditional marker-based, and more recently, some marker-less motion capture (MoCap) systems have demonstrated to be reliable tools for movement analysis. However, in complex experimental set ups involving virtual reality (VR) and free movements (as in [1]), accuracy and reliability tend to decrease due to occlusion, sensor blind spots, marker detachment, and other artifacts. Furthermore, when actions are less distinct, e.g., fast walk and slow run, current classification methods tend to fail when actions overlap, which is expected as even researchers struggle to manually label such actions.

## Research Question

Can current marker-less MoCap systems, pose estimation (PE) algorithms, and advanced action classification (AC) methods: (1) accurately track participant movements in VR; (2) cluster participant actions.

## Methods

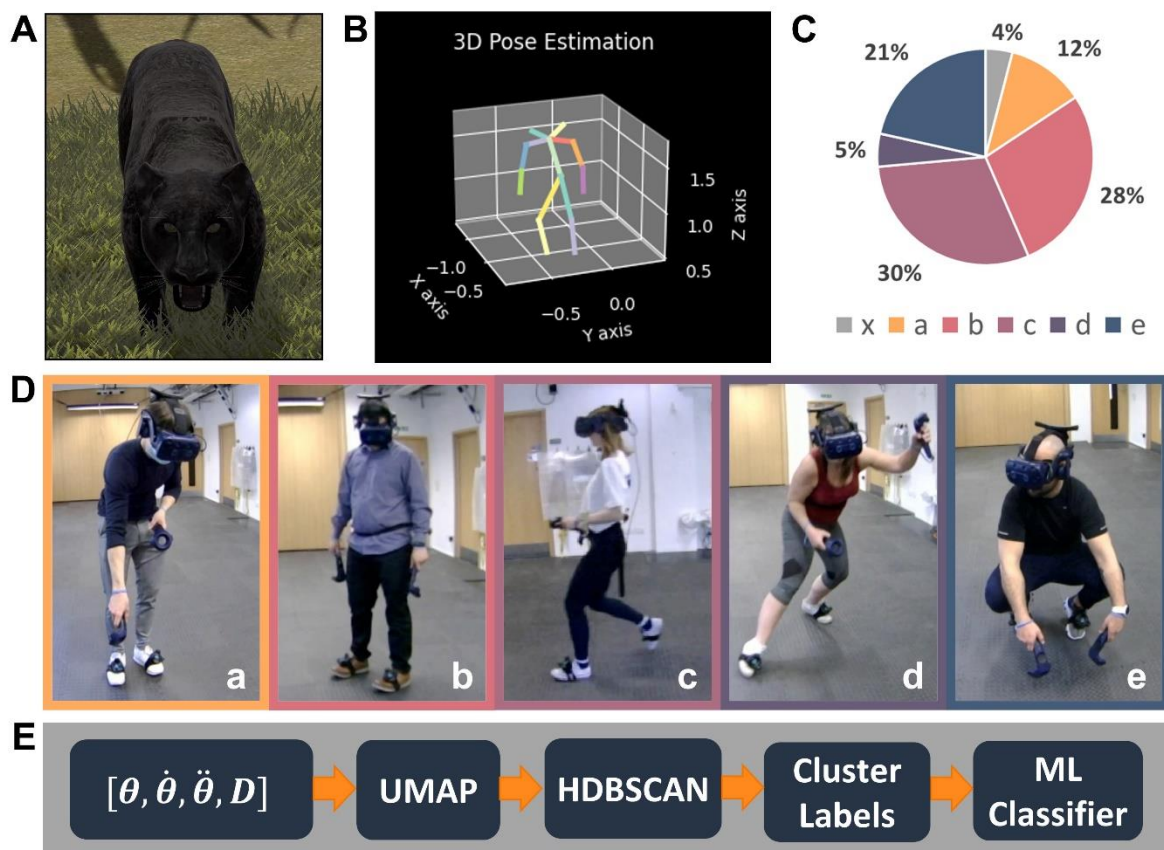
The experiment consisted of avoiding threats (Fig. 1A) whilst collecting fruit in VR environments (n=29 participants, 5x10m area), see [1]. The Unity<sup>®</sup> software [2], based on the Unity Experiment Framework [3], was used to create the VR experiment, which was streamed through an HTC vive pro (HTC Corporation) VR headset. Movements were recorded using 5 ELP cameras (1280x720 @120Hz) synchronised with the Open Broadcaster Software<sup>®</sup> (OBS) [4]. Openpose [5] was employed for PE (Fig. 1B). Euclidean distances, and angular positions, velocities, and accelerations were derived from cartesian positions. Finally, Uniform Manifold Approximation and Projection (UMAP) was used to embed high-dimensional features into a low-dimensional space, and Hierarchical Density Based Spatial Clustering of Applications (HDBSCAN) was used for classification (see Fig. 1E), similar to B-SOiD [6].

## Results

Participants were virtually killed by the threat in 223 episodes, for which the participants' last poses were estimated. After applying UMAP and HDBSCAN, 5 pose clusters were found (see Fig. 1C-D), which depict: (a) stand up, picking fruit with slow escape; (b) stand up, arms extended and slow escape; (c) long retreat at fast speed; (d) short retreat at medium speed; (e) crouching and picking fruit; (x) 4% unlabelled.

## Discussion

Marker-less MoCap and PE methods were mostly successful for participants' last poses. However, in some cases, and during exploration, tracking was lost due to occlusion and sensor blind spots. The results from the AC methods are an indication of the potential use of unsupervised methods to find participant actions under threat in VR. Nevertheless, such clustering is rather general, and had some AC errors, which could not be quantified as further work is needed to understand and define where the threshold of overlapping actions occurs. The results are exciting and promising; however, further investigation is needed to validate the findings, and to improve the AC methods.



**Figure 1.** (A) Panther threat experienced by participants in VR episodes; (B) Participant estimated 3D pose; (C) Pose clusters: [a] stand up, picking fruit with slow escape; [b] stand up, arms extended and slow escape; [c] long retreat at fast speed; [d] short retreat at medium speed; [e] crouching and picking fruit; [x] 4% unlabelled; (D) Pose cluster examples; (E) Methodology used in this investigation.

## References

- [1] J. K. Sporrer, J. Brooks, S. Hall, S. Zabbah, U. D. Serratos Hernandez, and D. R. Bach, "Computational Characteristics of Human Escape Decisions," *PsyArXiv*, 2020.
- [2] "Unity Technologies. (2023), Unity." <https://unity.com> (accessed Apr. 23, 2023).

- [3] J. Brookes, M. Warburton, M. Alghadier, M. Mon-williams, and F. Mushtaq, "Studying human behavior with virtual reality : The Unity Experiment Framework," pp. 455–463, 2020.
- [4] "Open Broadcaster Software. (2023), OBS." <https://obsproject.com> (accessed Apr. 23, 2023).
- [5] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, "OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields," 2018, doi: 10.48550/arxiv.1812.08008.
- [6] A. I. Hsu and E. A. Yttri, "B-SOiD, an open-source unsupervised algorithm for identification and fast prediction of behaviors", doi: 10.1038/s41467-021-25420-x.