Metrics for 3D Object Pointing and Manipulation in Virtual Reality

The Introduction and Validation of a Novel Approach in Measuring Human Performance

Eleftherios Triantafyllidis

The School of Informatics, The University of Edinburgh, Scotland, United Kingdom

Abstract

With the rapid growth in virtual reality technologies, multimodal/multisensory integration and robotics, teleoperation is becoming increasingly more immersive, elucidating human perception. Consequently, with the hopes of understanding how these new technologies truly affect human perception, the need to assess human motor performance has been a longstanding research goal. More specifically, assessing the performance of human motor skills, particularly during teleoperation, is challenging due to the complex spatial settings associated with 3D space. A promising human performance model is Fitts' law; one of the most widely used models in HCI history. While a promising basis for measuring performance, the law suffers in terms of simplicity when full 3D space is considered. Even though there has been a collective, multi-disciplinary effort in extending Fitts' Law to 3D space, a compelling standardized metric in 3D is missing, aggravating inter-study comparability.

Introduction

With the recent spike in 3D displays and mixed reality, current technologies have opened new pathways in 3D manipulation and pointing of virtual objects. This has led to an increase in teleoperation research – allowing humans the ability to remotely inhabit a foreign body, e.g. a robot as an avatar.

However, an important question comes to mind, how do we effectively measure and quantify such increased performance efficiently and as clearly as possible?

To propose a higher dimensional metric for assessing human performance, we investigated Paul Fitts' original predictive model, short for Fitts' law. The law measures the time required to point a generic mouse cursor, to a target on a screen. This is formulated as:

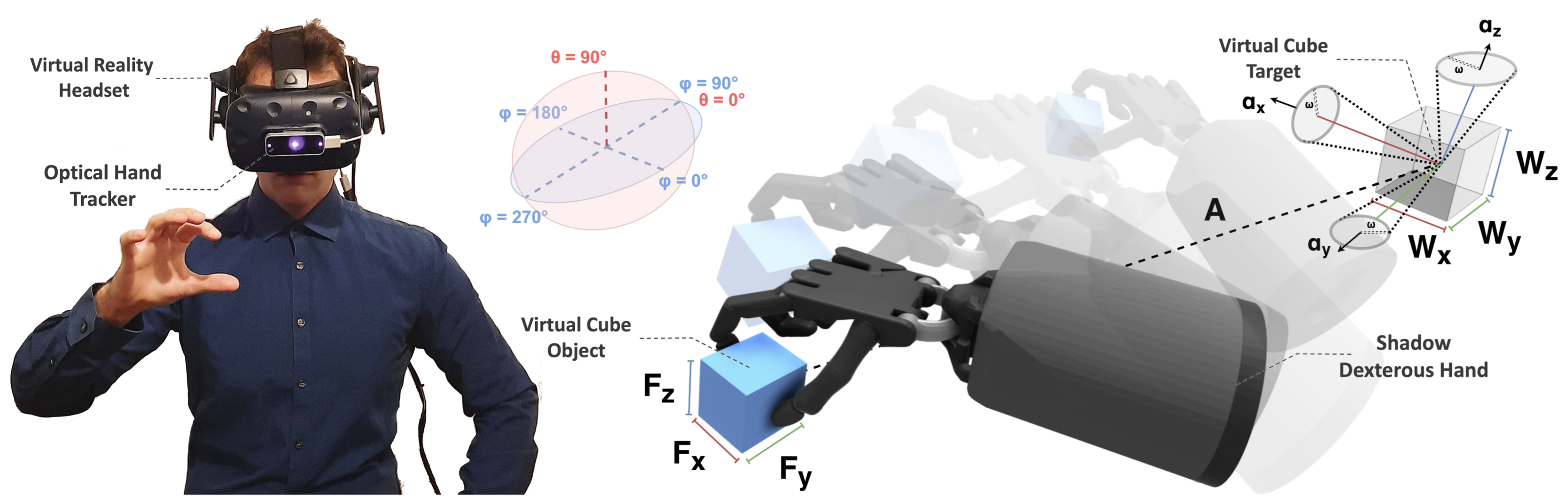
$$MT = a + b \cdot ID, \text{where } ID = log_2\left(\frac{2A}{W}\right). \quad (1)$$

Where, A represents the distance between a cursor and a target. W represents the width of the target area. The logarithmic term ID, represents the *Index of Difficulty* of the task, measured in bits per second. The resultant MT is measured in seconds. The constants a and b represent the y-intercept and slope respectively and are derived via regression analysis.

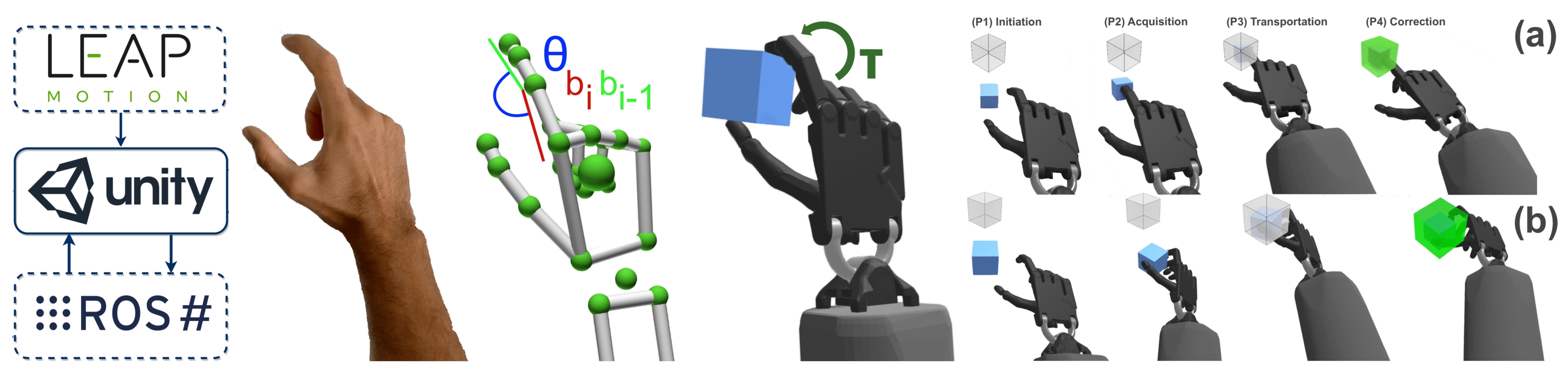
While the law suffers in terms of simplicity as it is was originally designed for 1D and 2D translational movement tasks, its simplicity and subsequent popularity render its pursuit of extending it to 3D space of significant importance. Finally, the derivation of such a metric may increase standardization as it encapsulates time and spatial based metrics under a single model.

Methodology and Approach

Four experiments (E1 to E4) were designed and conducted with progressively higher spatial complexity to study and compare existing metrics thoroughly. Our research goal was to quantify the difficulty of these 3D tasks and model human performance. Consequently, we evaluated the most popular 2D and "3D" extensions of Fitts' original 1D law in full 3D as seen in teleoperation. To achieve this, we conducted a user study (N=20) and evaluated each model on the experimental results, including ours, via linear regression analysis (R^2), to determine the model fitting.



An operator interacts with virtual objects in full 3D virtual reality with all task-related spatial variables.



From left to right. The simulation environment setup, our implemented hand-retargeting approach and the two types of object interaction (a) pointing tasks and (b) manipulation tasks.

The Derivation of our 3D Model

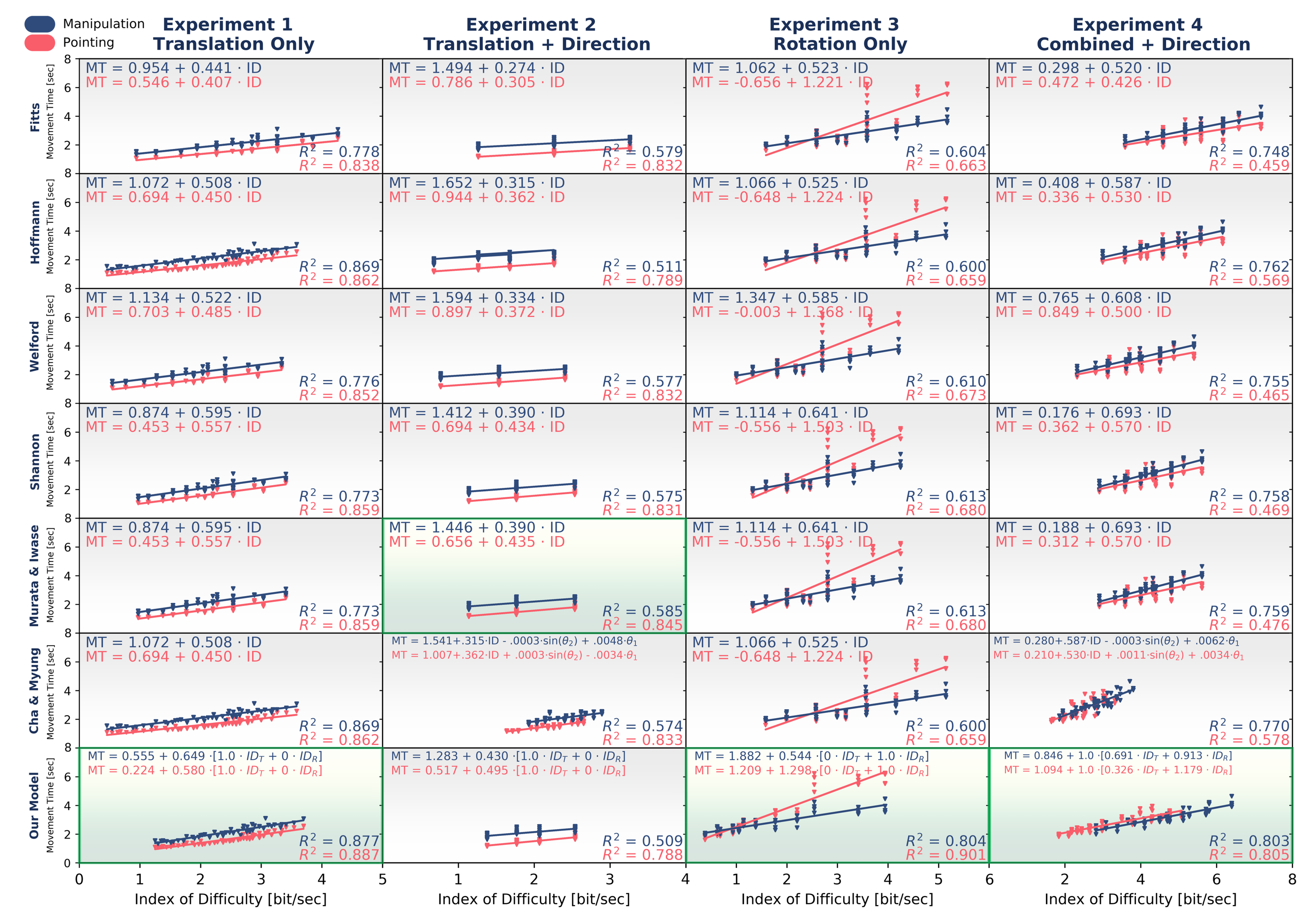
From our observations and from the experimental results, we formulated an improved and more suitable 3D formulation as:

$$\text{Final Model: } \begin{cases} MT &= a + b[c \cdot ID_t + d \cdot ID_r], \\ ID_t &= log_2\left(\frac{2A}{F+W} + 1\right), \\ ID_r &= log_2\left(\frac{2\alpha}{\omega^2} + 1\right), \end{cases} \quad (2)$$

This model encapsulates both translational and rotational movements under one model in full 3D based on Fitts' law and shows overall better and superior model fitting than existing formulations. For more information: *10.1109/MRA.2021.3090070.*

Results Highlights

The graph below illustrates all linear regression plots with the respective R^2 stemming from the most widely used model extensions of Fitts' Law, including the proposal of ours, on the experimental results (E1-E4) including pointing and manipulation tasks.



Regression plots of all models across E1 to E4, depicting line equations and r^2 fitting. Green boxes represent the best model in the column group.

Conclusion

In this work we conducted four experiments (E1 to E4), each adding progressively higher spatial complexity to study human movement and performance in full 3D space entailing translational and rotational movements in both object pointing and manipulation. In the most basic form of 3D movement (E1), we observed that existing approaches can be used to adequately model human movement, but were insufficient when spatial arrangements were introduced (E2). Moreover, the majority of these approaches did not model the rotational movements in E3 well and were insufficient when combining translation and rotation with spatial arrangements as studied in E4. Consequently, this led us to the proposal of a new performance model that can model human performance in full 3D space better than existing formulations.

Acknowledgements

This research is supported by the Engineering and Physical Sciences Research Council (EPSRC), the EPSRC FairSpace project (EP/R026092/1), EPSRC CDT in RAS (EP/L016834/1), and H2020 project Harmony (101017008). Ethics Approval 2019/22258.

Contact Details

Email: eleftherios.triantafyllidis@ed.ac.uk

Web: https://www.edinburgh-robotics.org/students/eleftherios-triantafyllidis

Youtube : https://www.youtube.com/channel/UC39bFeeUDUg2GxqqKrdkOpw