Effects of Tracking Scale on User Performance in Virtual Reality Games

Ben Burgh, Kyle Johnsen
University of Georgia

ABSTRACT

We explore how scaling a user's tracking data may impact performance in an immersive virtual reality game, which may have implications for fairness and accessibility of many applications. In our study, which used an HTC Vive room-scale VR system, users play the role of a factory worker who must remove deformed bread from a production line. Users were scaled to a reference height, such that taller than average users were rendered shorter and had shorter reach and shorter than average users were rendered taller and had longer reach than normal. Users also performed with unscaled tracking data. Our analysis indicates that there was no systematic advantage of being taller or shorter than normal, and scaling users may have had a detrimental effect. Moreover, scale changes were noticed by many users who had conflicting preferences for various application-specific reasons, indicating that application strategy can be affected by scale. Results suggest that while virtual reality tracking data may be scaled to compensate for user differences in physical height or reach, care must be taken to ensure that performance will benefit.

Keywords: Virtual Reality, Scale, Ergonomics, User Study, Immersion

Index Terms: K.6.1 [Management of Computing and Information Systems]: Project and People Management—Life Cycle; K.7.m [The Computing Profession]: Miscellaneous—Ethics

1 INTRODUCTION

Historically, psychology, training, and productivity applications have dominated the marketable immersive virtual reality landscape [1]. With the release of several mass-market immersive virtual reality systems such as the Oculus Rift and HTC Vive, video games have become a major user case. This shift may force re-examination of some of the assumptions made by virtual reality designers. This paper studies one such assumption, that the user should be rendered at his or her own physical height.

Many video games and other virtual environments provide users with an avatar, a surrogate for their real body in the virtual world. Through careful design, the virtual environment "fits" the avatar perfectly. Everything that should be visible is visible. Everything that should be reachable is reachable. Users know this, and thus have increased confidence. By contrast, in the real world, physical stature and capability differs significantly from person to person, with large mean differences between genders as well as those with disabilities. With the rise of immersive virtual reality for non-training (e.g. gaming) applications, these real world disparities may become prominent. For example, *Is the user going to have to be tall and jump high to dunk in a virtual reality basketball game?*

kjohnsen@uga.edu

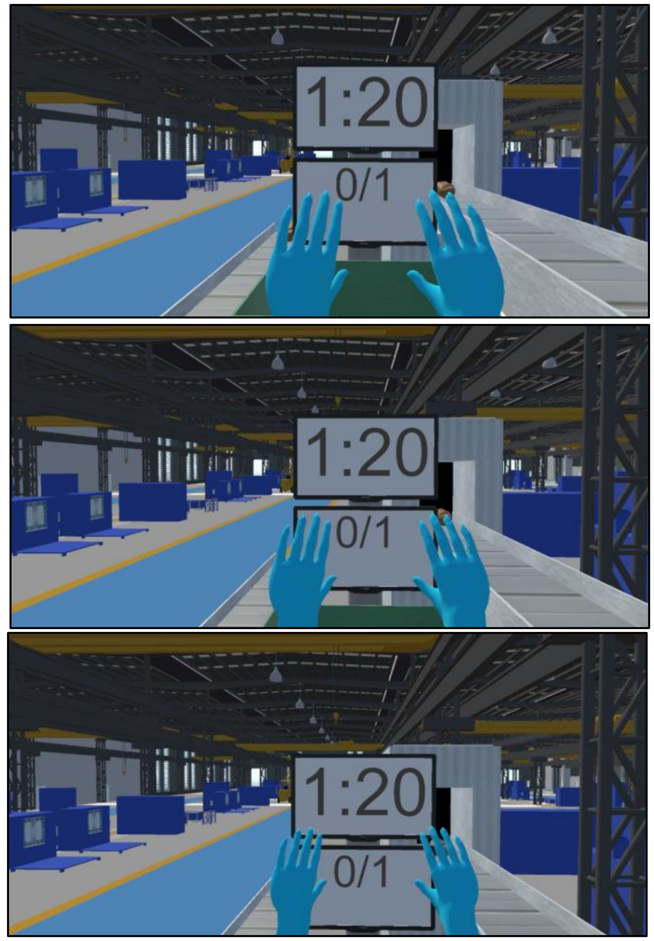

Figure 1: The center view is an average height/reach user. The top view has an 80% scale applied. The bottom has a 120% scale applied.

Facing the prospect of alienating users who may have a disadvantage, game designers may seek to compensate for user differences. Like the real world, virtual reality designers could simply design realistic worlds with tolerances for deviation from the average. However, this is complex, which is a major reason for the existence of the entire ergonomics field [2]. Rather than studying anthropometry, designers may choose to use the malleability of virtual body representation to compensate for user differences. Indeed, this is relatively straightforward to accomplish in immersive virtual reality. By scaling body tracking data, users may be given extended virtual height or reach, or all users may be forced to have the same. This is particularly simple to accomplish when the user's body is not directly tracked or represented, but instead only the tracked objects are visible, e.g. hand-held controllers or surrogates (e.g. virtual hands) for those controllers.

Though users may sense this modification, the extent to which it would affect performance is unclear. Some research would suggest that there would be little effect. It is well known that the visual senses are dominant over kinesthetic ones [3], that users may exhibit extreme "homuncular flexibility" when faced with non-matched avatars [4], and that distance judgements within head-mounted displays often underestimated relative to the real world [5]. This knowledge has already been exploited in the case of, for example, extended reaching [6, 7], redirected touching [8], or providing users with extra limbs or a tail [9]. Moreover, scaling tracking data is not the same as being taller or having longer reach. Tracking performance (e.g. jitter, accuracy) would also be proportionally scaled, as sensor resolution is fixed. So, if scaled up, a shorter user may be better able to reach an object they may also be less able to accurately manipulate it, or if a user were taller they may be better able to see over a barrier, but may also experience increase simulator sickness from head tracking jitter.

These competing aspects motivated us to explore the idea of user scaling in the context of an immersive video game. In this paper, we report on a study (N=38) that had users play a game with both scaled and normal tracking data. We present results and discuss how performance may have been impacted, as well as user detection and preference for the technique, and directions for future research in the area.

2 User Study

The purpose of our study was to determine if by scaling users to all be the same virtual height with proportional reach, if performance variation was similarly scaled, i.e. the game was more "fair". This was considered dependent on whether or not real world height was a predictor of performance. In the absence of this, we predicted that scaling (in either direction) would have a net negative effect on performance and increase variance. In any case, we were interested to see if users noticed scaling and, if so, how they felt it affected their performance.

2.1 Application & System

The context for our study was an immersive game for the HTC Vive virtual reality system that was developed as a demonstration for food-industry tradeshows. In the game, users play the role of a commercial bakery employee who is tasked with inspecting loaves of bread for quality and removing unacceptable bread. As seen in Figure 2, the user is placed in the center of three conveyor belts arranged in a U-shaped configuration. The two side conveyor belts transport bread from ovens for inspection and move in opposite directions. The left conveyor belt is lower than the right. The front conveyor belt is used to transport acceptable bread to packaging, and a bin behind it is used to toss defective bread into. The bread arrives via two conveyor belts, one of which is lower and placed to the left of the user and the other is higher and on the right. The game time remaining, starting at 2 minutes, and the user's score, counted as correctly sorted loaves, is presented on a virtual display in front of the user. The game becomes more challenging as time goes on – what is at first a leisurely chore quickly becomes an overwhelming task as the bread spawns more frequently. This is surprisingly difficult, as users must both be alert and agile, ready to pick up and examine bread on either side.

The HTC Vive system includes a tethered, stereoscopic head-mounted display (1080 x 1200 pixels per eye, 110 degree diagonal field of view, 90Hz refresh) and 2 wireless, wand-shaped controllers. These devices are tracked through HTC's "Lighthouse" system, an optical-inertial system that provides accurate, low-latency results over a room-scale volume similar to [10]. Unity3D was used to develop the application, and the SteamVR Unity Plugin provided access to tracking data. The Vive also includes mechanical interpupillary distance correction, which is sensed and automatically incorporated by Unity and SteamVR to generate correct stereoscopic rendering in applications.

Figure 2: The virtual environment showing the conveyors, various types of bread, GUIs, and "reject" bin,

In the application, the user's body is not represented. Rather, as is commonplace in current generation Vive-based games, only the controllers are represented. Also, at the start of the application, users are required to "put on" virtual gloves, at which point the representation of the controllers are replaced by the gloves.

Of note, this application was not designed for the present study, but rather motivated it. The physical demands placed on the user as well as the requirements of reaching and different height levels raised the potential issue that performance would be moderated by user physical characteristics. We suspected that taller users would have an advantage over shorter users, by increased visibility of upcoming bread and a likelihood of increased reach. We designed a scaling approach that attempts to compensate for physical user height disparities by scaling the world in proportion to height. A reference height, Yv, is measured in the Vive's tracking space (measured HMD vertical position relative to the tracking origin, the floor). The user's virtual height could then be adjusted based on this reference by having the user stand straight, measuring the HMD height, Yu, and computing a scale value, S = Yv/Yu Given this scale, all tracking data is adjusted by a uniform scale by S. Note, this is not the same as scaling the virtual world, as eye separation is not scaled, and thus stereoscopic depth perception would not change. However, it does have the effect of scaling movement speed, potentially also scaling depth cues from head-motion motion parallax. We also considered increasing the user's virtual height by a set amount, akin to standing on a step-stool or only scaling Y-data. However, the former makes lower objects more difficult to reach and neither would help with reachability. Furthermore, the simplicity of a uniform scale was attractive as a general technique for user scaling.

2.2 Study Design

Participants (N=38) each performed two trials, one using scaled tracking data, and one using unscaled tracking data. The order of these were randomized, such that 20 participants used the scaled experience first, and 18 participants used the unscaled experience first. Participants were not told which experience they were using.

Given that average adult heights vary significantly by gender, we chose to use two different reference heights, one for male users and one for female users. We did this so that the average deviation would be similar between genders (rather than female participants or male participants being disproportionately scaled up or down). We obtained the male reference height by measuring the "vive-height" of an average-height US male (177cm) and then subtracted that height by 13.5cm to obtain the reference female height.

Figure 3: A figure showing a user being calibrated in the HTC Vive (and showing the study environment)

It is important to note that the effects of virtual height, with respect to the environment. With a larger height, users would generally be able to reach further. If everyone is scaled to the same height, then there should not be an ideal height to be at.

2.2.1 Measures

The primary measure used in the experience was the performance, the percentage of correctly classified bread, in each experience. We also asked participants to self-report height, age, gender, overall gaming experience, and fitness with single-item survey items. Finally, we asked participants for comments regarding the experiences and if they noticed a difference in their virtual height between the first and second trials and how it affected them.

2.2.2 Population and Environment

We recruited participants from a large undergraduate engineering design laboratory within a large university in the United States. The laboratory conveniently already had an HTC Vive system installed in a cubicle, which was used for the experiment. The cubicle met the minimum requirements for "room-scale" tracking in the SteamVR ecosystem, which is a 1.5mx2m space free of obstacles, and isolated participants during trials (See Figure 3). Participants present in the laboratory were directly recruited to participate in the study. All participants were recruited over a 2-week period. No compensation or other incentive was provided. The overwhelming majority of participants were male (34M, 4F) and of predictable age for undergraduate students (min: 18, max: 21, avg: 19.85).

2.2.3 Procedure

After recruitment and informed consent, participants filled out our demographic survey, and then were asked to put on the Vive HMD and handed the controllers. They were then asked to stand straight upright to measure their vive-height and compute their scale factor (according to gender, as described earlier). Following this, they performed two trials, one of which was at their unscaled (real) height and one at their scaled (average) height in random order. Following the first trial, they were instructed to close their eyes while their scale was adjusted and then they performed the second trial. After completing trials, they were debriefed and asked for comments regarding their experience.

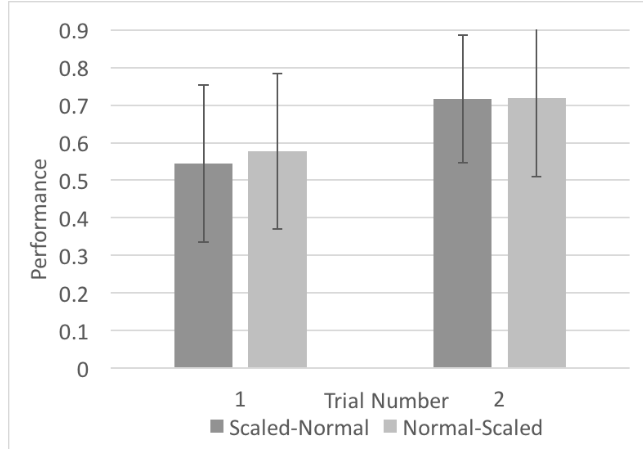

Figure 4: Plot of average performance through both trials according to user scale (larger value is a shorter measured height).

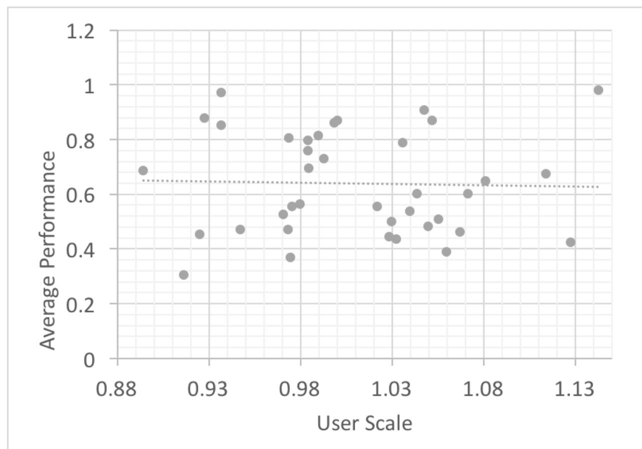

Figure 5: Plot of average performance through both trials according to user scale (larger value is a shorter measured height).

3 RESULTS

All 38 participants completed both trials. 20 participants (18M, 2F) were randomly assigned to the Scaled-Normal order and 18 participants (16M, 2F) were assigned to the Normal-Scaled order. Data was analyzed using IBM SPSS Statistics version 24. Participant heights were normally distributed about the reference height, and the average scale used in the study measured at 101% relative to the reference height. The maximum scale was 114% and the minimum scale was 89%. The average scale for users was +/- 4.8% relative to the reference.

Average performance in trial 1 was 56% (SD=21%) and in trial 2 was 72% (See Figure 4). A Pearson correlation analysis was performed on all measured variables showed a significant positive correlation between average performance and self-reported gaming experience (r=0.49, p=0.002). No other significant correlations were observed in the data, and most importantly, no correlation was observed with performance for physical height, scale, or absolute scale (See Figure 5). Following this, a repeated measures analysis of variance was performed, using the two trials as the repeated measure, order as the independent variable, and gaming as a covariable. The results showed substantial improvement between the two trials (F=39.7, p<.001), with almost all participants

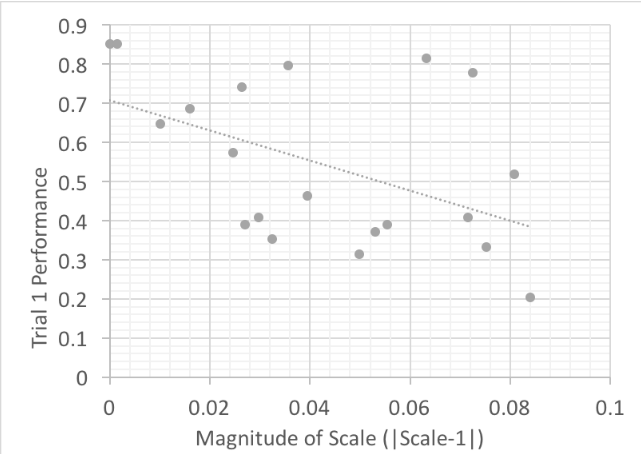

Figure 6: Plot of performance in Trial 1 by users who were scaled

improving greatly between trial 1 (M=56%, SD=21%) and trial 2 (M=72%, SD= 19%).

Given that physical height was not a significant factor, we expected that any scaling would negatively impact performance. To determine this, we performed a correlation analysis between unscaled performance, scaled performance, and absolute scale, looking at each trial independently. For the first trial, there was a significant negative correlation ($r=-0.49$, $p<.05$) within the unscaled group between performance and absolute scale, i.e. larger scaling values (positive or negative) had a net negative predictive relationship with performance (See Figure 6). This relationship was not significant for the second trial.

4 Discussion

The results show that performance within the game was highly dependent on practice with the game and self-reported gaming experience overall. We did not find a significant effect of height on game performance. There are several possible explanations for this. One explanation is that we needed to perform more trials to find the "maximum" performance of a user, i.e. we were still seeing mostly learning effects occurring because of practice and that any effects of height were inconsequential. However, it is more likely that physical height is not predictive of performance for this game. This is partially supported in that scaling height had a negative net impact on the default performance of participants in trial 1.

Users also had opinions about the scale change. 18 out of 38 participants noticed and commented on the scale change (including the most scaled users), although there was not a significant difference in scale between those who noticed and those who did not. Of those who noticed, the change in scale was seen both positively and negatively, with no apparent relationship between the magnitude or direction of scaling. Some users preferred being shorter so that they could better reach the lower conveyor belt. Others preferred being taller so that they could see the higher conveyor height. They wanted to "get used to" the experience rather than have the scale change on them between trials. This suggests that height is a complicated factor within the game, with ways to adapt behaviour based on height to maximize performance.

5 Conclusions and Future Work

Fairness in virtual reality games may be an important issue, given that the input range, unlike traditional controllers, is different user to user. However, unlike in the real world, user ranges can be modified in the virtual world by adjusting tracking data. While our study did not find that height difference significantly impacted performance, virtual height was a factor in user's opinions of the experience. Also, as we discovered, attempting to significantly adjust this virtual height, may have a negative effect on performance.

This preliminary work had many limitations with respect to studying this issue, including the population and application. The study population was 90% male Engineering students. Furthermore, the application was not explicitly designed for the study, nor was it a commercially available video game. These aspects limit the generalizability of the results.

Following this study, we are interested in exploring a game with a more obvious dependence on height and various compensatory mechanisms. There are many possibilities to explore. For example, the game could require users to place objects on shelves, with a scale factor being used to allow shorter users to reach the taller shelf, or a maze with walls that only some users can look over. As designers could not be expected anticipate all such ways that user measurements may affect performance, we believe that in the future, games will need to include some form of this compensation for at least height and reach. This may be a straightforward uniform scale, as we employed in this study, or may be other "adaptive" equipment like we use in the real world (e.g. raised platform, reacher). In any case, these compensatory mechanisms will likely need to be shown to be fair for any user, regardless of physical measurements, to be accepted by gamers, as it will be difficult to regulate their use.

Aknowledgements

References

[1] F. P. Brooks, "Whats Real About Virtual Reality," *IEEE Computer Graphics and Applications,* vol. 19, pp. 16-27, 1999.

[2] J. Dul, R. Bruder, P. Buckle, P. Carayon, P. Falzon, W. S. Marras, *et al.*, "A strategy for human factors/ergonomics: developing the discipline and profession," *Ergonomics,* vol. 55, pp. 377-395, 2012/04/01 2012.

[3] E. Burns, S. Razzaque, A. T. Panter, M. C. Whitton, M. R. McCallus, and F. P. Brooks Jr, "The Hand is Slower than the Eye: A quantitative exploration of visual dominance over proprioception," in *Virtual Reality, 2005. Proceedings. VR 2005. IEEE*, 2005, pp. 3-10.

[4] A. S. Won, J. Bailenson, J. Lee, and J. Lanier, "Homuncular flexibility in virtual reality," *Journal of Computer-Mediated Communication,* vol. 20, pp. 241-259, 2015.

[5] V. Interrante, B. Ries, and L. Anderson, "Distance Perception in Immersive Virtual Environments, Revisited," in *IEEE Virtual Reality Conference (VR 2006)*, 2006, pp. 3-10.

[6] I. Poupyrev, M. Billinghurst, S. Weghorst, and T. Ichikawa, "The go-go interaction technique: non-linear mapping for direct manipulation in VR," in *Proceedings of the 9th annual ACM symposium on User interface software and technology*, 1996, pp. 79-80.

[7] K. Kilteni, J.-M. Normand, M. V. Sanchez-Vives, and M. Slater, "Extending body space in immersive virtual reality: a very long arm illusion," *PloS one,* vol. 7, p. e40867, 2012.

[8] L. Kohli, M. C. Whitton, and F. P. Brooks, "Redirected touching: The effect of warping space on task performance," in *3D User Interfaces (3DUI), 2012 IEEE Symposium on*, 2012, pp. 105-112.

[9] W. Steptoe, A. Steed, and M. Slater, "Human tails: ownership and control of extended humanoid avatars," *IEEE transactions on visualization and computer graphics,* vol. 19, pp. 583-590, 2013.

[10] S. Islam, B. Ionescu, C. Gadea, and D. Ionescu, "Full-body tracking using a sensor array system and laser-based sweeps," in *2016 IEEE Symposium on 3D User Interfaces (3DUI)*, 2016, pp. 71-80.