Let Your Fingers do the Walking: A Unified Approach for Efficient Short-, Medium-, and Long-Distance Travel in VR

Zhixin Yan*
Worcester Polytechnic Institute

Robert W. Lindeman†
HIT Lab NZ
Worcester Polytechnic Institute

Arindam Dey‡
HIT Lab Australia

Abstract

The tradeoff between speed and precision is one of the challenging problems of travel interfaces. Sometimes users want to travel long distances (e.g., fly) and care less about precise movement, while other times they want to approach nearby objects in a more-precise way (e.g., walk), and care less about how quickly they move. Between these two extremes there are scenarios when both speed and precision become equally important. In real life, we often seamlessly combine these modes. However, most VR systems support a single travel metaphor, which may only be good for one range of travel, but not others.

We present a new VR travel framework which supports three separate multi-touch travel techniques, one for each distance range, but that all use the same device. We use a unifying metaphor of the user's fingers becoming their legs for each of the techniques. We are investigating the usability and user acceptance of the fingers-as-legs metaphor, as well as the efficiency and naturalness of switching between the different travel modes. We conducted an experiment focusing on user performance using the three travel modes, and compared our multi-touch, gesture-based approach with a traditional Gamepad travel interface. The results suggest that participants using a Gamepad interface are more time efficient. However, the quality of completing the tasks with the two input devices was similar, while ForcePad user response was faster for switching between travel modes.

Keywords: 3D travel interface, multi-touch gestures.

Index Terms: I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction techniques; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual reality

1 Introduction

Travel is one of the most basic and common Virtual Reality (VR) tasks. Designing a good travel interface to change the position and orientation of one's virtual representation from point A to B in a Virtual Environment (VE) is still a challenging problem. One of the problems is to effectively, efficiently, and realistically map the user's operation in the finite real-world space to locomotion in an infinite VE. As suggested in [2], more natural locomotion can enhance the sense of presence. Walking-based interfaces (e.g., [3][4]) and leaning-based interfaces (e.g., [5][6][7]) are two common types of natural locomotion interfaces. However, in many applications, travel is not the goal, but a way to reach a location in order to perform other tasks like selection and manipulation [1]. A good travel interface should also produce low fatigue for long-term use. VR researchers and arcade game developers have implemented several low-fatigue travel solutions, such as finger walking [9], and touch-based travel like [13]. The most popular one is WASD+Mouse which is the basic set in every FPS game. The drawbacks of this technique are also obvious, in that it can only provide discrete speed control, both hands are occupied during travel, and the user can only travel where they are looking.

Another challenge to designing travel interfaces is scalability. In our real life, we have many different modes of transportation for different travel purposes. We would like to walk to a place nearby, drive to somewhere miles away, and take a flight to a different city. Similarly, in VEs, we also have different types of travel needs, such as short-distance, long-distance, and flying.

In this paper, we propose a low-fatigue VR travel framework which provides smooth transitions between three types of travel, walking, Segway, and surfboard, with one multi-touch device. The main idea is that by mapping one-handed gestures to lower body motion, users can travel in a low-fatiguing and intuitive way while working on other VR tasks, like picking up objects or moving virtual widgets, with the other hand.

To evaluate our travel framework, we designed and conducted a user study focusing on user performance and behaviors under three travel techniques, and compared our multi-touch gesture-based approach with a Gamepad based travel interface. Our main objective was to investigate quantitative and qualitative usability of multi-touch gesture- and pressure-based travel interfaces.

2 Our Multi-touch, Gesture-based Travel Interface

We designed a multi-touch, gesture-based travel interface with a multi-touch, force-sensing device, the Synaptics ForcePad, which can detect both the position and pressure of up to five fingers individually, and provide 6-bit resolution and up to 1,000g of force sensing. With the 2D-touch+pressure information, we are able to map the user's two-finger gestures to virtual foot gestures and locomotion of the virtual character for the three travel modes.

2.1 Walking Gesture

In our framework, walking is designed as a low-speed but high-precision surface travel interface. Based on this design goal, there are 3-DOFs of movement including forward/backward, strafing, and turning (yaw). The first two DOFs, 2D translations on the ground surface, are controlled by a two-finger gesture mimicking a bipedal walking motion (Figure 1a). The trails of each finger are translated into virtual-world locomotion. This allows users to control the speed by the frequency and distance of each "step." The third DOF (yaw) is a rate-controlled mapping, implemented by pressing on either the left or right side of the pad (Figure 1b). The pressure on each side will determine angular speed of turning left or right. The translation and rotation gestures are independent, and users are able to do both translation and rotation simultaneously like our natural walking experience. To help the user distinguish the translation area and rotation area, we added

*zyan@wpi.edu
†gogo@wpi.edu
‡aridey@gmail.com

IEEE Symposium on 3D User Interfaces 2016
19–23 March, Greenville, SC, USA
978-1-5090-0842-1/16/$31.00 ©2016 IEEE

tape to the rotation areas on the left and right sides to provide passive haptic feedback.

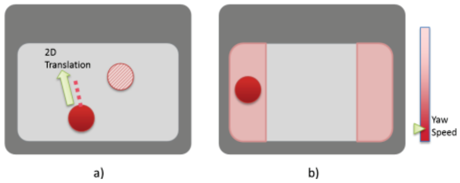

Figure 1: Illustration of walking metaphor gesture mapping.

2.2 Segway Gesture

To achieve faster speed on the ground, the user can put both her index finger and middle finger horizontally aligned on the ForcePad to switch to Segway mode. As a vehicle, a Segway only has two DOFs. One is moving forward or backward, with the speed determined by how much weight the user puts on her toes or heals. The other, steering, is controlled by the difference between left-foot and right-foot pressure.

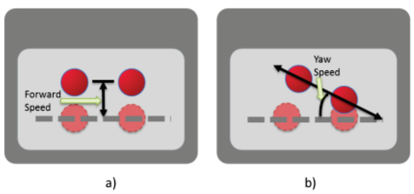

Figure 2: Illustration of Segway metaphor gesture mapping.

We define a baseline once the user triggers Segway mode (Figure 2a). The distance between the middle point of the two fingers and the baseline is mapped to the speed of moving forward or backward. Rotation mapping in the second solution is position controlled done by mapping the Y-axis difference between the two fingers on the ForcePad to the angular speed, moving towards the desired Segway pose from the current perspective (Figure 2b). The baseline is reset each time the fingers touch the ForcePad.

2.3 Surfing Gesture

If the user wants to travel at a much faster speed than a Segway, say to travel to the other side of a map, or fly to somewhere across a very deep valley or river, he can switch to flying surfboard mode by placing his index finger and thumb in a vertical line on the touchpad (see Figure 3). Similar to the mapping proposed by Wang & Lindeman [8], the pressure difference between the front (index) finger and the back (thumb) finger is the pitch angle of the board. The X-axis difference between the two fingers is mapped to the angular (yaw) speed for steering. The third DOF, the speed of moving forward, is controlled by the Y-axis difference between the two fingers on the ForcePad, similar to the idea of two-handed flying introduced by Mine et al. [10].

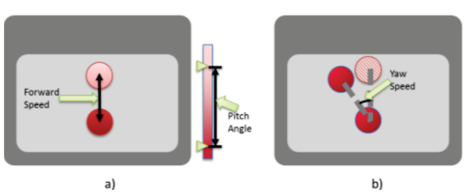

Figure 3: Illustration of surfing metaphor gesture mapping.

Switching between walking, Segway, and surfing is accomplished by removing the fingers from the pad, and placing them down using the preferred metaphor (alternating down-swipes for walking, simultaneous horizontal touch for Segway, and simultaneous vertical touch for surfing).

3 Experiment

3.1 The Virtual Environment

Our virtual world was developed using the Unity3D Game Engine. In the virtual world, there was a large maze with four platforms on the four corners (Figure 4a). Trees, grass, and street lanes were included to increase the realism, provide motion cues, and to indicate valid paths in the maze. To aid in wayfinding during the study, we put an arrow right in front of user's view to show the directions in walking and Segway modes (Figure 4b). The arrow always pointed to the current item to collect (walking) or intersection to cross (Segway). For ForcePad subjects, we displayed a widget in the top-right corner to show the location and pressure of their fingers on the ForcePad. Specifically for Segway trials, we drew the baseline on the widget to help subjects control the speed of movement and rotation.

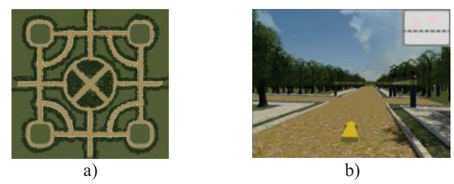

Figure 4: the virtual environment.

3.2 Gamepad Travel Interfaces

Figure 5 shows the mappings for the three Gamepad travel interfaces. Each used the same travel modes and same rotation and translation DOF control as the multi-touch, gesture-based mechanics. The main idea was to map the two joysticks on the Gamepad to locomotion and to use the buttons on the right for switching between travel modes (Figure 5a).

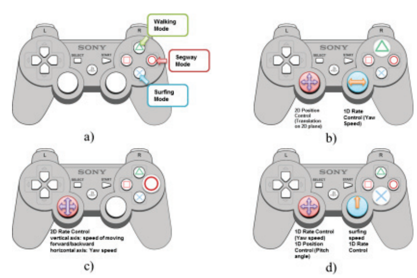

Figure 5: Gamepad control mappings.

For the Gamepad walking interface (Figure 5b), the left joystick was mapped to 2D translation, and the horizontal axis of the right joystick was mapped to yaw rotation. The Gamepad Segway interface (Figure 5c) only used the left joystick, similar to racing games. The vertical axis of the left joystick controlled speed for moving forward and backward, while the horizontal axis of the left joystick was used for yaw rotation. The Gamepad surfing interface (Figure 5d) treated the left joystick as a surfboard. The pitch and yaw of the joystick were mapped to pitch angle and yaw speed. The vertical axis of the right joystick was used to control the speed of movement.

3.3 Tasks

For each travel mode, we designed different tasks to compare the two control devices, including an object-collection task for the walking interface, a path-following task for the Segway, and a breadcrumb-following flying task for the surfing interface. In the walking task, participants were asked to collect targets (spinning pumpkins). A new pumpkin would appear only after the participant collected the previous one. The distance between each target was the same, while the angles between edges were chosen from 36º, 72º, and 108º, depending on the difficulty level. In the Segway task, participants had to follow a certain path in the maze to reach a goal as fast as possible while minimizing collisions with blockers and trees along the road. We also designed three difficulty levels for the Segway task based on the length of the path and number of sharp turns. The paths used for Segway are illustrated in Figure 6.

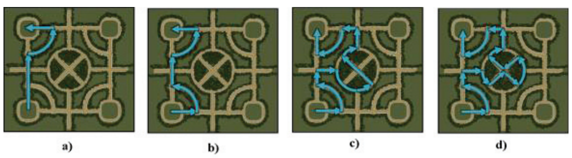

Figure 6: Illustration of sample paths in the Segway tasks. From a) to d) the difficulty levels are training, easy, medium, and hard. Higher difficulty level had longer paths and sharper turns

For the surfing task, participants were asked to follow a path in the sky. Although both pitch and yaw controls were active, the flying path was designed with only pitch variation as we found it was already very challenging during our pilot study. The breadcrumb paths were generated using a bell-shaped function with different height and climbing rates. The harder the task, the higher and faster participants had to climb during the study. The *y* value could vary from Equation 1,

$$y = h * \frac{1}{(1 + \left(\frac{p - 0.5}{a}\right)^{2b})} \qquad (1)$$

where *h* was the maximum height of the path, *a* and *b* were two variables to control the shape of the bell function. Variable *p* was calculated by dividing the distance between the start and end points by the distance already traveled.

3.4 User Study Procedure

When a participant arrived, he was first asked to read and sign the IRB-approved consent form, then asked to fill out a general information form which included demographic questions and gaming experience. After that, he sat on a stationary chair and wore an eMagin Z800 Head-Mounted Display. Participants in the Gamepad group used a Sony PS3 Controller held in two hands, while those in the ForcePad group had a wooden board laid across their lap to provide a stable support platform for the ForcePad device. Then they were asked to face front to calibrate the inertial (SpaceFusion) tracker mounted on the HMD. The setups for both groups are shown in Figure 7.

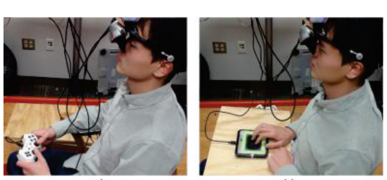

Figure 7: Picture of system setup for both the Gamepad group (a) and the ForcePad group (b)

A training session was designed to ensure at least a minimum level of proficiency for participants in each group. To complete the training session, participants had to finish the task within 80 seconds for walking training, 120 seconds for Segway training and 100 seconds for Surfing training. Additionally, the Segway training tasks required participants to complete the task with fewer than five collisions. For surfing training, they had to be within a tolerance of the breadcrumb path. If a participant failed to pass one type of task, he had to do the particular task again until his performance met the minimum requirement.

In the study trials, participants were asked to complete two sessions, each with nine trials with three levels of difficulty (easy, medium, and hard) using the three travel modes with one control device. We used a 9x9 Latin square to counterbalance the trials and minimize learning effects.

After finishing, participants were asked to fill out a NASA TLX questionnaire [11] and a post-test questionnaire about realism, sense of presence, usability, and fun using six-point Likert scales for their general travel experience with the control device, as well as the three travel modes. We also interviewed every participant to collect comments about their travel experience. The whole study took about 45 minutes for a participant to complete.

Of 32 subjects, 28 successfully finished the study. There were 14 males overall, with six in the ForcePad group, and eight in the Gamepad group. Subject ages ranged from 18 to 35 years (M = 20.9, SD = 3.8), and gaming experience from 1 to 6 points (M = 3.3, SD = 1.5). There was no significant gaming experience difference between the groups.

3.5 Hypothesis

We formulated three hypotheses before conducting the experiment.

H1: Participants using the ForcePad will have a deeper sense of presence than Gamepad participants.

H2: The Gamepad is a more abstract device, and so will be rated higher regarding ease of use.

H3: Participants in the ForcePad group will remember how to transition between the multi-touch gestures better than those in the Gamepad group.

For H1, since finger walking mimics the movement of the legs, we believe it will lead to a deeper sense of presence. In terms of H2, McMahan et al. [12] found that non-natural interfaces may be more efficient than more natural ones. In terms of H3, as the finger gestures for the different travel modes are vastly different from each other, and the mode switch for the Gamepad is pressing three buttons located close to each other, we believe participants in the ForcePad group will remember the transitions better and be more aware of their current travel mode.

4 Results & Discussion

4.1 Subjective Results

We ran Mann-Whitney U tests to analyze the differences between the two travel devices as rated by the participants through the post-questionnaire. Surprisingly, we did not find any significant differences between the two travel devices. Hence, our first three hypotheses were refuted.

The ratings from the NASA TLX were analyzed using a one-way ANOVA for all questions. In the analysis, we noticed that, while other questions had no significant differences, participants using the Gamepad had significantly less frustration than ForcePad users: $F(1, 26) = 128.571$, $p = .013$, $\eta^2_p = .215$.

Surprisingly the results refute both **H1** and **H2**. There is no significant difference between the two travel interfaces in participants' mind regarding the presence, ease of use, or fun. However, the data from the NASA TLX questionnaire showed that the multi-touch, gesture-based interface was harder to interact with. Combining the results from informal interviews after the study, four out of 14 participants from the ForcePad group complained about the limited workspace. More participants in the ForcePad group (seven vs four in the Gamepad group) complained about surfing mode being too hard. As one of the participants explained, when increasing the pressure of one finger, the board would start to turn. Because the human hand is neither symmetric nor rigid, whenever the pressure changed, the touch point shifted, which led to frustration for ForcePad participants..

4.2 User Performance Results

During our user study, we collected user performance data including time to complete each trial, 3D position and orientation at each time frame, events such as collisions with trees and blockers in the scene for Segway tasks, and time for entering every target/intersection for both walking and Segway tasks, for later quantitative analysis. The differences that were significant or showed a strong trend are shown in Table 1.

Table 1: User performance results. W, S, and F are Walking, Segway and Surfing (Flying). When users tried to trigger a travel mode, this was recorded as response time. When they successfully finished a mode switch, this was recorded as correct response time.

	Gamepad		ForcePad		F	df	η^2_p
	M	SD	M	SD			
Completion time (W)	12.1	3.2	20.9	2.5	36.1***	1/26	.581
Completion time (S)	44.8	4.8	68.6	21	17.8***	1/26	.406
Completion time (F)	15.1	6.4	25.6	5.8	20.8***	1/26	.444
Path deviation (W)	6.67	.63	7.55	1.1	6.74*	1/25	.212
Overshoot (S)	.742	.36	.196	.24	21.4***	1/24	.472
Response time (W)	1.61	.43	1.07	.46	9.97**	1/26	.277
Response time (S)	1.61	.43	1.00	.46	13.2***	1/26	.337
Correct response time (W)	1.61	.43	1.07	.46	9.97**	1/26	.277
Correct response time (S)	1.96	.46	2.2	1.1			

*p < 0.05, **p < 0.01, ***p < 0.001

This result supports **H2,** meaning that the Gamepad device is more efficient in time. The result is similar to the work of McMahan et al. [12], who hypothesized that natural interfaces might have more muscle groups involved and greater system latency than non-natural interfaces. In our case, the multi-touch, gesture-based interface has more hand and finger motion involved than just moving both thumbs. At the same time, the gesture classification brings additional latency to the system which might be another reason for the poorer performance compared to the Gamepad interface.

Although Gamepad interfaces are more efficient in time, we found that multi-touch, gesture-based interfaces performed better in some of the measurements. In the Segway trials, we found participants in the Gamepad group often overshot intersections and also kept zigzagging during travel because the joystick on the Gamepad was either too unresponsive or too sensitive. Three participants from the ForcePad group said that operation was very smooth in Segway mode, while no one in the Gamepad group especially liked Segway mode. The result indicates that multi-touch, gesture-based interfaces can be used for tasks needing subtle and precise operation.

The measurement of mode-switch time partially supports our hypothesis **H3**. We could see that participants using the ForcePad interface responded to the mode switch instruction significantly faster in walking and Segway trials. However, it was not the case for Segway trials regarding correct response time. One of the reasons could be the extra latency the multi-touch, gesture-based interface caused due to gesture classification. Another possible explanation is that there were not enough cues about the location of the ForcePad in the real world. The lack of proprioception might have led to unexpected touch points. For Segway and surfing modes, which required fairly precise two-finger touches to trigger them (±20º), once the fingers left the ForcePad, participants had to test the relative position of the fingers and ForcePad, and then adjust their fingers to the right position. Additional passive haptic feedback could be helpful for locating the control devices.

We would like to improve the finger-based interactions in the future, and with better touch and pressure sensitivity of the hardware, we believe we can significantly improve the usability of this type of travel metaphor. It will be interesting to integrate other interactions, such as shooting, together with travel in a use case and evaluate how users can manage multiple tasks through this interaction method. We believe this work provides innovative design ideas for interactions with virtual environments.

References

[1] D. A. Bowman, E. Kruijff, J. J. LaViola, and I. Poupyrev. *3D User Interfaces: Theory and Practice*. Addison, 2004.

[2] M. Usoh, K. Arthur, M. C. Whitton, R. Bastos, A. Steed, M. Slater, and F. P. Brooks, Jr. Walking > walking-in-place > flying, in virtual environments. *Proc. of the 26th Annual Conf. on Computer Graphics and Interactive Techniques, SIGGRAPH '99*, pp. 359–364..

[3] H. Iwata, H. Yano, H. Fukushima, and H. Noma. CirculaFloor: A Locomotion Interface Using Circulation of Movable Tiles. *Proc. of 2005 IEEE Virtual Reality Conf.*, pp. 223–230.

[4] S. Razzaque, Z. Kohn, and M. C. Whitton. Redirected walking. *Proc. of 2001 Eurographics*, pp. 289–294.

[5] S. Beckhaus, K. J. Blom, and M. Haringer. ChairIO - the Chair-Based Interface. Magerkurth and Rötzler (eds.), *Concepts and Technologies for Pervasive Games: A Reader for Pervasive Gaming Research,* vol. 1, (Shaker Verlag), pp. 231–264, 2007.

[6] J. Wang and R. W. Lindeman. Leaning-based Travel Interfaces Revisited: Frontal Versus Sidewise Stances for Flying in 3D Virtual Spaces. *Proc. of ACM Symp. on Virtual Reality Software and Technology (VRST) 2012*, pp. 121-128.

[7] D. Valkov, F. Steinicke, G. Bruder, and K. H. Hinrichs. Traveling in 3D Virtual Environments with Foot Gestures and a Multi-Touch enabled WIM. *Proc. of 2010 IEEE Virtual Reality Conf.*, pp. 171-180.

[8] J. Wang and R. W. Lindeman. Comparing isometric and elastic surf-board interfaces for leaning-based travel in 3d virtual environments. *Proc. of 2012 IEEE Symp. on 3D User Interfaces (3DUI), pp.* 31-38.

[9] J. Kim, D. Gračanin, K. Matković, and F. Quek. Finger walking in place (FWIP): A traveling technique in virtual environments. *Smart Graphics*, pp. 58–69. Springer, 2008.

[10] M. R. Mine, F. P. Brooks Jr, and C. H. Sequin. Moving objects in space: exploiting proprioception in virtual-environment interaction. *Proc. of the 24th annual conf. on Computer graphics and interactive techniques*, pp. 19–26. ACM Press/Add. -Wesley Publishing, 1997

[11] S. G. Hart and L. E. Staveland. Development of NASA-TLX (task load index): Results of empirical and theoretical research. *Advances in psychology*, 52:139–183, 1988.

[12] R. P. McMahan, A. J. D. Alon, S. Lazem, R. J. Beaton, D. Machaj, M. Schaefer, M. G. Silva, A. Leal, R. Hagan, and D. Bowman. Evaluating natural interaction techniques in video games. *Proc. of the 2010 3D User Interfaces (3DUI) Symp., pp.* 11-14.

[13] A. Benzina, A. Dey, M. Tonnis, and G. Klinker, Empirical evaluation of mapping functions for navigation in virtual reality using phones with integrated sensors. *Proc. of the 10th Asia Pacific Conf. on Computer Human Interaction*, pp. 149-158. 2012.