Shadow Walking: an Unencumbered Locomotion Technique for Systems with Under-floor Projection

David J. Zielinski[1] Ryan P. McMahan[2] Rachael B. Brady[3]

Duke University Virginia Tech Duke University

ABSTRACT

When viewed from below, a user's feet cast shadows onto the floor screen of an under-floor projection system, such as a six-sided CAVE. Tracking those shadows with a camera provides enough information for calculating a user's ground-plane location, foot orientation, and footstep events. We present *Shadow Walking*, an unencumbered locomotion technique that uses *shadow tracking* to sense a user's walking direction and step speed. Shadow Walking affords virtual locomotion by detecting if a user is walking in place. In addition, Shadow Walking supports a sidestep gesture, similar to the iPhone's pinch gesture.

In this paper, we describe how we implemented Shadow Walking and present a preliminary assessment of our new locomotion technique. We have found Shadow Walking provides advantages of being unencumbered, inexpensive, and easy to implement compared to other walking-in-place approaches. It also has potential for extended gestures and multi-user locomotion.

KEYWORDS: Virtual Reality, Navigation, Shadow Walking.

INDEX TERMS: I.3.7 [Three-Dimensional Graphics and Realism]: Virtual Reality; H.5.2 [User Interfaces]: Interaction Styles

1 INTRODUCTION

When viewed from below, a user's feet cast shadows onto the floor screen of an under-floor projection system, such as a six-sided CAVE [1]. These shadows are cast by the user's feet contacting the floor screen, thus blocking overhead illumination from being seen below at the points of contact (see Figure 1). This phenomenon can be utilized to provide a user's ground-plane location, foot orientation, and footstep events by capturing the shadows with a camera and processing differences in luminance. We refer to this as *shadow tracking*.

By using shadow tracking, we have developed an unencumbered locomotion technique called *Shadow Walking*. Shadow Walking is a travel technique that provides virtual locomotion through walking in place (i.e., users move their feet to simulate walking without physically translating their bodies). Unlike most walking-in-place implementations (e.g., [2], [3], [4]), Shadow Walking does not require any sensors or tracking devices to be worn and, hence, is unencumbered. Additionally, Shadow Walking is less prone to step-recognition errors due to tracking the contact between a user's feet and the floor screen as opposed to analyzing a wide range of body motions.

[1]email: djzielin@duke.edu
[2]email: rymcmaha@vt.edu
[3]email: rachael.brady@duke.edu

Another feature of Shadow Walking that we have developed is a sidestep gesture, similar to the iPhone's pinch gesture. By stepping out to a side and sliding the outreached foot along the floor closer to the other foot, Shadow Walking allows for virtual sidestepping (i.e., moving in the direction of the activating outward step). This sidestep gesture is one of many potential gestures that shadow tracking can afford.

In this paper, we explain how we developed shadow tracking for determining ground-plane location and foot orientation in an under-floor projection system. We then describe the implementation of the walking-in-place and sidestep gestures for Shadow Walking. We also present a preliminary assessment of our new unencumbered locomotion technique and discuss the advantages, limitations, and potentials for Shadow Walking as a usable, unencumbered alternative to other walking-in-place implementations.

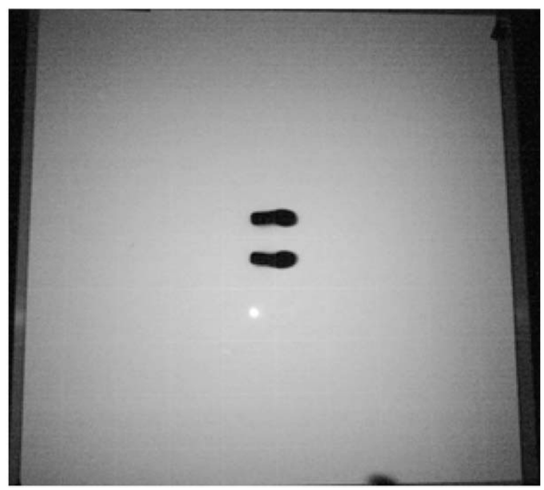

Figure 1. Luminance-only view of user's feet from beneath the floor screen in an under-floor projection system

2 RELATED WORK

Based on our survey of research literature, we have been unable to identify many uses of tracking shadows for interaction capabilities. Echtler, Huber, and Klinker utilized shadow tracking with multi-touch tables to provide the ability to control multiple cursors by hovering over the table surface [5]. This allowed users to separately control cursor movements and "clicks", which were distinguished by touching the table surface. The implementation developed by Echtler, Huber, and Klinker was based on frustrated total internal reflections (FTIR), originally described by Han [6].

Unlike shadow tracking, locomotion techniques have been extensively researched. Bowman et al. have categorized many locomotion techniques, making distinctions between physical locomotion, steering, route planning, targeting, manual manipulation, travel-by-scaling, viewpoint orientation, and velocity specification [7]. One type of physical locomotion that has been investigated as an alternative to real-walking locomotion is walking in place, which overcomes the requirement for a large physical space.

There have been several approaches to implementing walking in place. Slater, Usoh, and Steed used a head-mounted, six-degrees-of-freedom (6-DOF) tracker and a neural network to determine when users are walking in place and based the direction of locomotion on the direction of gaze reported by the head-mounted tracker [2]. The neural network they used required training data from each user to compute the weights for the network using back-propagation. Usoh et al. have since developed a standard network based on his own gait [8].

Templeman et al. implemented walking in place by attaching a 6-DOF tracker to each knee and placing force sensors on the user's shoe insoles [3]. This system, known as Gaiter, used a sophisticated algorithm to recognize natural walking motions, such as stepping forward, backward, and sideways. Feasel, Whitton, and Wendt implemented a similar technique with their low-latency, continuous-motion walking-in-place (LLCM-WIP) technique, which also utilized a 6-DOF tracker attached to the user's chest for direction of locomotion [4].

Contrastingly, Iwata and Fujii implemented walking in place by using omni-direction sliding sandals to allow users to "shuffle" their feet back and forth on a low-friction surface [9]. They used a hoop set around the user's waist to limit the user's movement and provide support for shuffling. Six-DOF tracking was used to determine if the user was walking in place and the direction of walking. Swapp, Williams, and Steed have recently implemented a similar technique using the Wizdish [10].

Despite the amount of research on physical locomotion and walking-in-place techniques, many of the implementation approaches have relied heavily on encumbering attachments, expensive tracking capabilities, and, in some cases, user calibration. These implementations have also been known to suffer from problems of recognition errors [8]. Type I errors occur when the system judges users to be walking in place when they are not. Type II errors occur when the system judges users to not be walking in place when in fact they are. Type I errors are often more costly and disturbing for users since additional locomotion is necessary to correct the undesired locomotion.

3 Shadow Tracking

Shadow tracking provides the capability to track a user's ground-plane location and foot orientation by capturing the shadows of the feet with a camera and processing differences in luminance.

3.1 Image Capturing

In order to capture the shadows of the user's feet, a camera must be mounted beneath the floor screen of an under-floor projection system. In most of these systems, the camera cannot be mounted directly under the floor screen as the floor projector or projection mirror will already be positioned there. Instead, the camera must be mounted to one side and oriented to view the entire floor screen (see Figure 2). This off-center mounting increases the complexity of shadow tracking but is addressed during the processing of the captured images.

For our research, we used a PlayStation Eye, a USB 2.0 video camera designed as a peripheral device for the PlayStation 3. We chose this device due to its low cost (retail $39.99 USD) and video specifications. The PlayStation Eye is capable of providing uncompressed video at 640 x 480 resolution at 60 frames per second (fps) or at 320 x 240 resolution at 120 fps. Unfortunately, due to the available device drivers for our dual-core 2.0 GHz system running Ubuntu 10.4, we were only able to achieve 640 x 480 resolution at 30 fps.

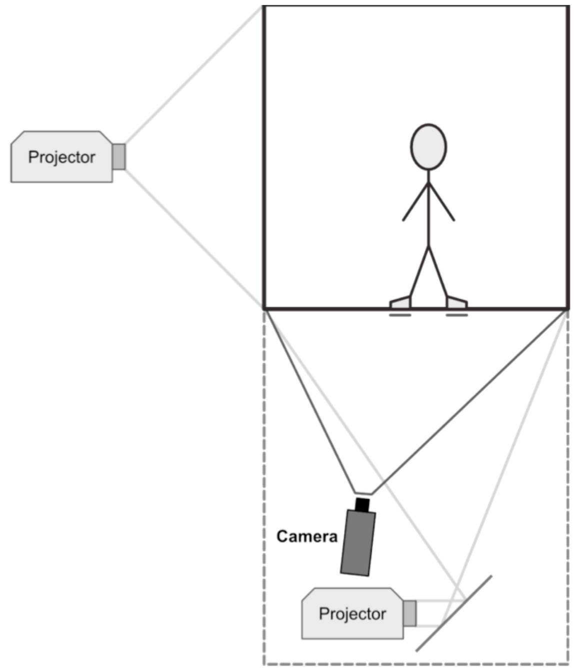

Figure 2. Camera mounting in an under-floor projection system

3.2 Shadow Detection

To process the uncompressed video coming from the PlayStation Eye, we used the Video For Linux Two (V4L2) library. V4L2 is a set of APIs and standards for handling video devices on Linux. In order to begin identifying shadows, we used V4L2's YUYV format, which separates brightness information (Y) from color information (U and V). Specifically, this format represents two pixels, with a Y value for each pixel and shared U and V values for the two combined.

Using the YUYV format, we developed the following algorithm for detecting shadows within each image captured (see Figure 3).

1. *Crop*
   We discard the outer portions of the image that are not part of the floor screen. This relies on the camera being mounted.
2. *Box Blur*
   We perform the standard image filter on the Y values of each pixel to help eliminate single-pixel, luminance noise.
3. *Color Threshold*
   Using an upper threshold (for red and blue) and a lower threshold (for green) on the U and V values, we discard pixels with too much color since shadows lack color.
4. *Brightness Threshold*
   Using an upper threshold on the Y values, we discard pixels with too much brightness.
5. *Connectivity Test*
   We test each remaining pixel to ensure it is connected to other remaining pixels. If not, we discard the pixel as noise.
6. *Cluster Identification*
   While testing for connectivity, we identify clusters of pixels by keeping track of connected pixels in groups sorted by size (i.e., number of pixels).
7. *Position Correction*
   To account for off-center mounting and to convert pixels to units, we multiply the position of each remaining pixel by a calibration matrix. We calculated this calibration matrix by surveying nine points and using the linear least squares fit method to correct for the skewed perspective of the camera.

Figure 3. Shadows detected by algorithm (highlighted green)

The thresholds for the shadow detection algorithm can be adjusted for the expected brightness and colors of specific virtual environments (VEs). It is also important to note that the algorithm must be computed at a rate faster than the frame rate of the camera (30 fps) to avoid latency.

3.3 Tracking

Using the sorted clusters of shadows and their calibrated positions, we are able to determine the user's ground-plane location and foot orientation. We calculate ground-plane location by averaging the center-of-mass positions of the two feet, when both are down (i.e., the second-largest cluster is at least 50 percent the size of the largest cluster). When both feet are not down (e.g., the user is stepping), we do not update the user's ground-plane location.

We calculate the orientation of each foot by determining the furthest point from the center of mass and creating a vector from that point back through the center of mass. This works because the foot is wider in the front than the back, thus placing the center of mass towards the front of the foot. For tracking a user's orientation, we average the orientations of the two feet, when the dot product of the two vectors is greater than zero.

4 Shadow Walking

By using shadow tracking, we developed *Shadow Walking*. Shadow Walking is an unencumbered locomotion technique that affords virtual travel through walking in place and sidestepping.

4.1 Walking In Place

To provide a walking-in-place gesture, we had to detect when a user was stepping. We determined that we could compare the two largest clusters of shadows to make this detection. Essentially, if the second-largest cluster is less than 50 percent the size of the largest cluster, we consider the user to be in mid-step. Once, the second-largest cluster is at least 75 percent the size of the largest cluster, we consider both feet to be down, and we register a full-step event. We use different mid-step and full-step thresholds to avoid registering a flurry of step events due to the size percentage straddling a single threshold.

In addition to detecting when a user was stepping, we had to detect when a user was stepping *in place* to differentiate between the walking-in-place gesture and real walking. To do this, we compare the user's ground-plane locations before and after a full-step event. If the new ground-plane location is closer than 50 percent the longest foot length (i.e., farthest distance between two pixels in the largest cluster) from the original ground-plane location, the user is considered to be walking in place, and we virtually move the user in the direction of the average orientation of the feet. Otherwise if the new ground-plane location is further than 50 percent, the user is considered to be physically walking and no virtual locomotion occurs.

4.2 Sidestepping

In addition to developing the walking-in-place gesture, we also developed a sidestep gesture. Our inspiration for the sidestep gesture was the iPhone's pinch gesture used for panning images. We designed the gesture to be used by stepping out to a side and then sliding the outreached foot along the floor closer to the other foot. We use the vector starting from the user's original location to the outreached location (prior to sliding) for the direction of virtual sidestepping. Figure 4 demonstrates the sidestepping feature of Shadow Walking.

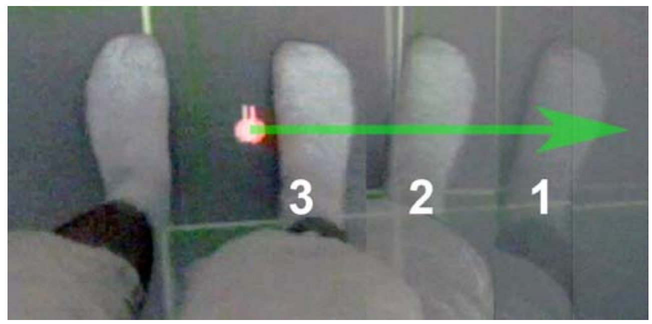

Figure 4. Sidestepping feature of Shadow Walking

5 Preliminary Assessment

Though we have not yet had the opportunity to conduct a formal evaluation of Shadow Walking, we have preliminarily assessed the new locomotion technique based on our own experiences.

Our first subjective finding is that Shadow Walking is easy and natural to use. Without the need for 6-DOF trackers or sensors, we are able to walk into our under-floor projection system and immediately begin using the unencumbered technique. The technique allows for seamless switching between physically walking around and the walking-in-place gesture. We have tried a range of stepping speeds, and Shadow Walking has successfully performed with all of them. We have noticed that the walking-in-place gesture is best performed with a flat-footed step, as opposed to tiptoeing for example (though this is expected due to our foot orientation calculations). We have also assessed the sidestep gesture and found it easy to perform, though it is often not necessary since it is more natural to turn and walk in place.

In addition to our subjective assessments, we have compared the tracking data of our shadow tracking to the data provided by an Intersense IS-900 head tracker. Based on our comparison, shadow tracking has a mean ground-plane location error of 13.4231 cm (std. dev. = 3.1785 cm) and a mean yaw error of 7.3427 degrees (std. dev. = 4.9574 deg.). We believe the large ground-plane location error stems from comparing positional values of a head tracker placed on shutter glasses to a center of mass calculated from the positions of the user's feet. Similarly, we believe the large yaw error (and standard deviation) are due to the fact that the head and feet are disjoint and can be pointed in different directions (especially when walking).

During our preliminary assessment of Shadow Walking, we also discovered that VEs with little overhead illumination make it nearly impossible to track shadows due to the lack of contrast. For instance, a VE with a moonless night sky might not work.

6 Discussion of Shadow Walking

6.1 Advantages

We believe Shadow Walking provides some advantages over traditional implementations of the walking-in-place gesture. First, Shadow Walking is completely unencumbered and requires no attachments to be worn. This avoids problems of users becoming entangled in wires, fatigued from extra weight, or distracted by sensors worn. Though modern 6-DOF trackers may be wireless and lighter-weight, we believe that users still become distracted with how they move these unnatural sensors as opposed to moving their feet with Shadow Walking.

Another advantage of Shadow Walking is its cheap and easy implementation. One concern Feasel, Whitton, and Wendt had with their LLCM-WIP implementation was how to make their system less expensive [4]. For $39.99 USD, we implemented Shadow Walking with a PlayStation Eye camera, the V4L2 library, and our shadow detection algorithm as opposed to expensive tracking systems and neural networks. (Though it is important to note that our technique is limited to under-floor projection systems that tend to be expensive.)

A third advantage of our new locomotion technique over some walking-in-place implementations is the lack of requirement for user calibration. Slater, Usoh, and Steed required users to train to compute the weights for their neural network [2]. Later on, Usoh et al. used a standard network calibrated on his own gait, though this can cause step recognition errors in some cases [8].

6.2 Limitations

Shadow Walking does have one major limitation: it only works in under-floor projection systems, such as a six-sided CAVE. This is an inherent limitation of the enabling shadow tracking capability and cannot be overcome. Unfortunately, most under-floor projection systems are six-sided CAVEs, which are currently sparse considering VE systems in general. Despite this limitation, we hope that further research into shadow tracking and techniques enabled by it (e.g., Shadow Walking) will motivate researchers to investigate under-floor projection systems other than six-sided CAVEs, such as a floor and front wall projection system.

A minor limitation that impairs Shadow Walking is the requirement of sufficient overhead illumination. As assessed, VEs with little to no overhead illumination make it nearly impossible to track shadows due to the lack of contrast. For such VEs, the shadow detection thresholds can be adjusted to afford tracking though there will likely be false positives with such thresholds.

6.3 Potential

We have identified two key potentials for our Shadow Walking technique: extended gestures and multi-user locomotion. As we found in our preliminary assessment, our sidestep gesture was not very useful (because it's easier to turn and walk in place), but it did prove that shadow tracking could provide for several types of extended gestures. A natural progression would be to investigate other "natural" locomotion gestures such as backpedaling or pushing a skateboard even, but we have envisioned developing extended gestures for "magic" functionalities [7]. For instance, the physical motions underlying our sidestep gesture could be used to scale the size of the virtual environment down for quicker locomotion with the walking-in-place gesture. Additionally, we have yet to investigate hands-to-floor gestures, but we expect to find a rich design space for such interactions.

In our current implementation, Shadow Walking utilizes the two largest clusters of shadows based on the assumption that there is one user. This does not mean that the technique cannot be modified for multi-user locomotion though. By examining the locations of each cluster of shadows, pairs of feet should be identifiable. Once clusters of shadows are paired, footsteps within each pair can be tracked. This capability affords the walking-in-place gesture (and possibly extended gestures) to each user in the system, without the need for additional sensors. The research focus then becomes how do multiple users control locomotion within the same VE system. Perhaps only one user should control locomotion at a time or perhaps the system should "fit" the VE to match the average location of all the users.

7 Conclusions and Future Work

We have presented the development of Shadow Walking, an unencumbered locomotion technique for systems with under-floor projection. By utilizing shadow tracking capabilities, Shadow Walking provides walking-in-place and sidestep gestures for virtual travel. Unlike prior walking-in-place implementations, Shadow Walking requires no attachments to be worn, is cheap and easy to implement, and requires no user calibration. But it is limited to under-floor projection systems, like a six-sided CAVE. We have also discussed the potentials of extended gestures (e.g., a world-scaling pinch motion) and multi-user locomotion.

We are currently planning to further investigate the design space of extended gestures and will follow with a formal evaluation of those incorporated into Shadow Walking. We may also investigate the potential of multi-user locomotion.

References

[1] Cruz-Neira, C., Sandin, D.J., and DeFanti, T.A., "Surround-screen Projection-based Virtual Reality: the Design and Implementation of the CAVE", in *Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*, pp. 135-142, 1993.

[2] Slater, M., Usoh, M., and Steed, A., "Taking Steps: The Influence of a Walking Technique on Presence in Virtual Reality", *ACM Transactions on Computer-Human Interaction (TOCHI)*, 2(3), pp. 201-219, 1995.

[3] Templeman, J.N., Denbrook, P.S., and Sibert, L.E., "Virtual Locomotion: Walking in Place through Virtual Environments", *Presence: Teleoperators and Virtual Environments*, 8(6), pp. 598-617, 1999.

[4] Feasel, J., Whitton, M.C., and Wendt, J.D., "LLCM-WIP: Low-Latency, Continuous-Motion Walking-in-Place", in *IEEE Symposium on 3D User Interfaces 2008 (3DUI)*, pp. 97-104, 2008.

[5] Echtler, F., Huber, M., and Klinker, G., "Shadow Tracking on Multi-Touch Tables", in *Working Conference on Advanced Visual Interfaces (AVI)*, pp. 388-391, 2008.

[6] Han, J.Y., "Low-cost Multi-touch Sensing Through Frustrated Total Internal Reflection", in *ACM Symposium on User Interface Software and Technology (UIST)*, pp. 115-118, 2005.

[7] Bowman, D.A., et al., *3D User Interfaces: Theory and Practice*, Addison-Wesley, Boston, 2004.

[8] Usoh, M., et al., "Walking > Walking-in-Place > Flying, in Virtual Environments", in *Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*, pp. 359-364, 1999.

[9] Iwata, H. and Fujii, T., "Virtual Perambulator: A Novel Interface Device for Locomotion in Virtual Environment", in *Virtual Reality Annual International Symposium (VRAIS)*, pp. 60, 1996.

[10] Swapp, D., Williams, J., and Steed, A., "The Implementation of a Novel Walking Interface within an Immersive Display", in *IEEE Symposium on 3D User Interfaces 2010 (3DUI)*, pp. 71-74, 2010.