International Conference on Artificial Reality and Telexistence
Eurographics Symposium on Virtual Environments (2014)
T. Nojima, D. Reiners, and O. Staadt (Editors)

Ground Navigation in 3D Scenes using Simple Body Motions

Parinya Punpongsanon[1] Emilie Guy[2] Tamy Boubekeur[2] Daisuke Iwai[1] Kosuke Sato[1]

[1] Osaka University, Japan [2] Telecom ParisTech, CNRS, Institut Mines-Telecom, France

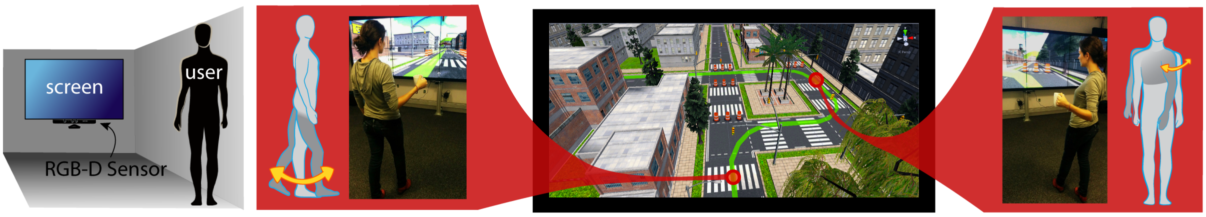

Figure 1: *From left to right: setup and user navigating in a scene while carrying a mug.*

Abstract

With the growing interest in virtual reality, mid-air ground navigation is becoming a fundamental interaction for a large collection of application scenarios. While classical input devices (e.g., mouse/keyboard, gamepad, touchscreen) have their own ground navigation standards, mid-air techniques still lack natural mechanisms for travelling in the scene. In particular, for most applications, the user should navigate in the scene while still being able to interact with its content using her hands, and observe the displayed content moving her eyes and locally rotating her head. Since most ground navigation scenarios require only two degrees of freedom to move forward/backward and rotate the view to the left or right, we propose a mid-air ground navigation control model which lets the user's hands, eyes or local head orientation completely free, making use of the remaining tracked body elements to tailor the navigation. We also study its desired porperties, such as being easy to discover, control, socially acceptable, accurate and not tiring.

Categories and Subject Descriptors (according to ACM CCS): Information Interfaces and Presentation (e.g., HCI) [H.5.2]: User Interfaces—;

1. Introduction

Mid-air techniques are becoming popular for public displays as users do not have to connect, touch or wear any specific device and can instantaneously interact with the system. Thanks to 3D scanners and online 3D databases, the number of available 3D scenes is also growing exponentially. Thus, practionners envision new public applications where the user navigates inside a 3D scene using her own gestures to, for instance, visit a virtual museum or find her way in a mall. In particular, Adhikarla et al. [AWB*14] use gestures similar to the touchscreen ones (rotate, pan and zoom). Whereas Nabiyouni et al. [NLB14] evaluate several travelling metaphors and multiple ways to control the speed. A different approach [SVAG14] is to have the user seated at a desk and use her foot to control the navigation. However, in the context of public environment, using small hand motions or asking the user to seat is not really practical. Ren and colleagues [RLOW13] use a freehand (no hands-on device) gestural technique, with a broom metaphor to navigate: the user hands control the walk and her shoulders control the view rotation. In the work of Roupe et al. [RBSJ14], users lean the bust forward/backward to walk, rotate the shoulders to turn, and raise their arm to stop the motions.

We take a different direction, as we believe that offering a natural mid-air navigation system while preserving the ability to interact with either the virtual or the real world content is a key element for a number of applications. This prevents making use of hands, eyes or local head orientation, and leads us to account for other body components.

2. Ground Navigation using Simple Body Motions

Ground navigation only requires two degrees of freedom: one for walking (backward or forward) and one for rotating the view (left or right). Based on the tracked user body model, we define all motions by measuring the angles made

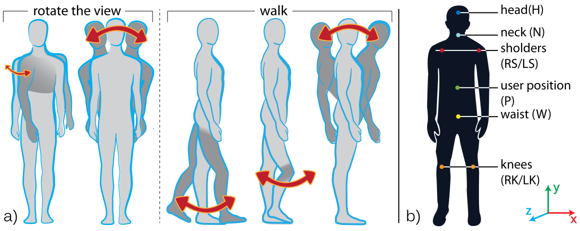

Figure 2: *(a:) Available motions. (b:) Tracked body model.*

by the body components at the current position w.r.t. a reference (rest) pose captured once at the beginning of the session. This makes our system adaptive to the initial pose and robust to the user morphology (e.g., height). Moreover, we couple the motion amplitude and the virtual velocity to ensure a proper speed control. In practice, we use a set of vectors defined over the tracked body key points to compute our angles:

Rotate the view: the user can either *rotate her shoulders* or *lean her bust* (see Fig 2.a). For the former, the vector spans the two shoulders positions **LS** and **RS** (see Fig. 2.b). For the latter, the neck **N** and waist position **W** are used instead.

Walk: the user can either *bend a knee* (right knee to go forward, left to go backward), *perform a step* (forward/backward), or *bend the bust* to walk (see Fig 2.a). For the *bust motion*, we use a vector that goes from the user position **P** to the user head **H**, and we compute the angle in the z axis. For the *bend knee* motion, we use a vector that goes from the left knee **LK** to the right knee **RK**. Finally to compute the *step motion*, we use the same angle than for the *bend knee* motion and we also compare the sum of the knees depth between the rest and current poses, to determine if the user goes forward or backward. We choose knees over feet because they are more likely to be inside the sensor frustum.

3. System Configuration and Implementation

We tracked the user body key points using a RGB-D sensor (e.g., Microsoft Kinect) with the Zigfu SDK, and render the 3D scene using the Unity3D game engine. Our system is divided in three main blocks: *motion measures*, *transfer functions* and *actuators*. The motion measure provides an angle, computed for a given vector from its rest pose, and normalized by a maximum range given as a parameter. The transfer function regulates the motion effect by converting the motion measure into a value between 0 and 1. We use $f(x) = \frac{(x-\alpha)}{1-\alpha}^{\beta}$, if $x \geq \alpha$, 0 otherwise, with α delaying the beginning of the motion to a certain angle value and β controlling the slope. Note that α trades a comfortable rest pose for low amplitude motions to navigate. Finally, the actuator computes the current speed for the walk (resp. view rotation) by modulating the transfer function output with a *walking* (resp. *rotation*) speed factor. Then, it updates the camera position in the virtual scene accordingly.

4. Demonstration

Once the user stands in front of the display (see Fig 1), the system captures her rest pose position (no "T-pose" like position is required). During these few seconds, a displayed image shows the available motions. After the registration step, the user can freely navigate in the scene. User can combine the two motions to move and turn in a synchronous way, and can also control the speed by varying the motion amplitude. Note that other users can stand by, they will not disturb the experience as far as they do not occlude her. The user can also interact with the system while doing something else like grabbing a bag, carrying a PC, or pointing virtual elements to someone. Finally a menu allows to instantaneously change the interactions, so that the user can experience different combinations of motions in the same session.

5. Conclusion and Future Work

We have described a general framework to navigate in a 3D scene using simple user gestures. By restricting the navigation to the ground plane, we can preserve complete freedom for the user hands, eyes and local head rotation, to perform other tasks in the virtual or real world. Although several related approaches [RLOW13, RBSJ14] use the shoulder rotation to rotate the view, we are not aware of a general agreement about the best set of body motions for full ground navigation. So far, we designed different motions for each interaction and let the user try and evaluate several combinations. As future work we will perform user experiments to determine the best motion combination for ground navigation, i.e., one set of motions that would be not tiring, easily understandable, controllable (the two motions have to be as decorellated as possible) and accurate. We will also optimize the different parameters of our system.

References

[AWB*14] Adhikarla V. K., Wozniak P., Barsi A., Singhal D., Kovács P. T., Balogh T.: Freehand interaction with large-scale 3d map data. In *Proc. of IEEE 3DTV-CON* (2014), pp. 1–4. 1

[NLB14] Nabiyouni M., Laha B., Bowman D. A.: Designing effective travel techniques with bare-hand interaction. In *Proc. of IEEE 3DUI* (2014), pp. 139–140. 1

[RBSJ14] Roupé M., Bosch-Sijtsema P., Johansson M.: Interactive navigation interface for virtual reality using the human body. *Computers, Environment and Urban Systems 43* (2014), 42 – 50. 1, 2

[RLOW13] Ren G., Li C., O'Neill E., Willis P.: 3d freehand gestural navigation for interactive public displays. *IEEE CG&A 33*, 2 (2013), 47–55. 1, 2

[SVAG14] Simeone A., Velloso E., Alexander J., Gellersen H.: Feet movement in desktop 3d interaction. In *Proc. of IEEE 3DUI* (2014), pp. 71–74. 1