信学技報 MVE2009-55, Vol.109, No215, pp.119-124 (2009)

社団法人 電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
IEICE Technical Report

ハンズフリー全方位ウォークスルーシミュレータ

朴 鴈振　大隈 隆史　興梠 正克　石川 智也　蔵田 武志

†産業技術総合研究所サービス工学研究センター
〒305-8568 つくば市梅園 1-1-1 中央第二
E-mail: {anjin.park,takashi-okuma, m.kourogi,tomoya-ishikawa,t.kurata}@aist.go.jp

あらまし 本稿では、ユーザの方向感覚を保持するための全方位表示ウォークスルーシミュレータ（WTS）について述べる。本シミュレータでは、歩行者デッドレコニング(PDR)に用いられる小型無線センサモジュールによってユーザの足踏み（WIP）や向きの計測することで仮想環境内を移動する。これによりハンズフリーな操作が実現されるため、環境側の設計とハンドヘルドデバイスや手持ちの紙媒体資料の組み合わせの評価を行うことができる。さらに WTS ユーザと指示者などの遠隔ユーザとの対面コミュニケーションを写実的アバタを通じて行うこともできる。

キーワード ウォークスルーシミュレータ，全方位ディスプレイ，ハンズフリー足踏みインタフェース，写実的アバタ

Hands-free Omni-directional Walkthrough Simulator

Anjin PARK, Takashi OKUMA, Masakatsu Kourogi, Tomoya Ishikawa, and Takeshi KURATA

Center for Service Research, AIST

E-mail: {anjin.park,takashi-okuma,m.kourogi,tomoya-ishikawa,t.kurata}@aist.go.jp

Abstract We present a walkthrough simulator (WTS) in which the display is omni-directional so as to preserve the viewer's sense of direction. A small-sized wireless sensor module for locomotion interface based on walking-in-place (WIP), detecting the viewer's step and orientation, is also presented to navigate the virtual constructions. The sensor module attached to the viewer's belt realizes the hands-free operation, which makes it possible to evaluate the usability of handheld devices or even just a paper map while walking around in virtual environments. In addition, the proposed WTS supports face-to-face communication between the simulator and a remote computer to help the subject reach their destination. The WTS is then applied to navigation in photorealistic environment for viewers to tour the remote indoor place without visiting in person, as an application of the proposed simulator. Based on the hands-free operation and omni-directional display, the proposed WTS can provide the viewers with a feeling that they experience the virtual environment as similar as possible with the real world.

Keyword Walkthrough Simulator, Omni-directional Display, Hands-free Locomotion Interface, Photorealistic Avatar

1. Introduction

Before constructing or renovating houses or buildings, details such as suitability of the floor layout and navigation signs and whether customers will feel comfortable in the buildings should be considered. Virtual reality techniques are used to investigate virtual structures in details. However, since information is generally displayed through monitors and a keyboard or mouse is used to navigate the structure, it is difficult to evaluate the relationship between the details of the structure and the sense of absolute direction of the viewer [1]. Therefore, we have developed the walkthrough simulator (WTS) in order to enable subjects to navigate virtual constructions from the perspective of the customer. The final goal is to create a 360-degree virtual world displayed around the subjects as they physically walk through the virtual world.

In the beginning of the 1990's, a head-mounted display (HMD) with head-tracking devices that detect the location and orientation of the viewer to simulate the correct view was presented to provide 360-degree virtual scene [2]. However, wearing the HMD for more than ten minutes is visually stressful, because the field of view of HMD is limited compared with human eye. Therefore, CAVE™ [3] was presented to provide wide field of view as an alternative of HMD, which is a cubic with display screens that surrounds a viewer. The goal of the original version of the CAVE is to provide viewers with a 10′×10′×9′-sized theater, using rear and down-projection screens. However, since the rear- and down-projection screen requires a large backward, especially high ceiling, it is usually installed in two-stories-height rooms. For optimum configuration of display screen in limited space, Iwata [4] proposed Garnet Vision that is a closed screen in which image covers full angle around the viewer, optimized by determining two criteria, space efficiency and pixel efficiency, and Ensphered Vision that is an image display system for full-surround spherical screen. However, since the Garnet Vision was built with 12 projects and dodecahedron -shaped screens and Ensphered Vision was built with a spherical screen and convex and plain mirrors, it is fairly difficult to develop and lead to cost. Moreover, since the purpose of these systems [3,4] was to just display the virtual scenes around viewers, they did not enable viewers to navigate virtual construction as the viewers physically walk.

To allow the viewer to navigate the virtual construction in limited space, locomotion interface based on walking-in-place (WIP), which results in changing the location of the viewer's eye in the virtual scene, has become a popular topic [5]. The locomotion interface can be classified into three majors, i.e. pressure sensor-based, treadmill-based, and accelerometer-based. The pressure sensor-based approaches use pressure sensors under the feet to detect when a foot is on the sensor [6]. The treadmill-based approaches use treadmill to synchronize the virtual world with leg motion [7]. Lastly, accelerometer sensors [8] are used to detect users' movement by measuring acceleration in three directions x, y and z, sensing the detectable ankle motions for the foot movement.

However, the pressure sensor and treadmill-based approaches have their own drawbacks when used in the limited spaces. In the pressure sensor-based approaches, the viewers should be careful not to lose their balance and tumbled over, because they stand on the pressure sensor to move their feet, thereby it is not easy for the viewers to be immersed in navigating the virtual construction. The treadmill-based approaches make too much noise when users are walking on the treadmill, thereby communication, which is one of important factors in the virtual reality, is not easy. Moreover, since the treadmill is not safe, compared with the other approach, users hold a safety device to walk on the treadmill. With the accelerometer-based interface, the viewers have little trouble, but additional sensors or methods are required to detect more precise movements needed to navigate the virtual environments.

This paper proposes a novel hands-free omni-directional walkthrough simulator. The display is omni-directional so as to preserve the viewer's sense of direction, and a hexahedron simulator in which the viewer can stand on the center is built in simple and compact form to be suitable for our goal. In addition, a novel locomotion interface based on WIP to navigate virtual constructions displayed on the omni-directional screen is presented. The interface detects five movements defined based on actual people's movements when they are walking to their destination in a building, with a small-sized wireless sensor module that detects WIP and orientation. Hands-free operations can be realized with the proposed interface, because the small-size sensor module is attached to the viewer's belt. That makes it possible to evaluate the usability of handheld devices or even just a paper map while walking around in virtual constructions.

The proposed simulator also provides communication between the viewer in the simulator and a guide in a remote computer. The purpose of this communication is to provide instructions to viewers to help them reach their destination, like guides in some buildings, such as public institutions. The guide will be displayed as *Photorealistic Avatar*, in which the appearance of an actual person is used as CG texture in the virtual world, and the communication is performed with the photorealistic avatar and 3D voices.

Consequently, the proposed WTS can help to experience the virtual world with several factors that are experienced in the real world, which are omni-directional sense of direction of users, hands-free operations, and communication with a guide displayed as a photorealistic avatar, even if the world is virtually constructed and the user is walking in place.

The remainder of this paper is organized as follows. The overall structure of the proposed WTS is introduced in section 2, and section 3 describes how to detect users' movement and rotation and how to navigate the virtual world using sensor modules. Then, section 4 describes a method to generate photorealistic avatars. Some experimental results are presented in section 5, and the final conclusions are given section 6.

2. Walkthrough Simulator

The goal of the proposed WTS is to create a 360-degree virtual construction around a viewer who wants to check the suitability of the floor layout or tour the indoor environments as they physically walk through the virtual construction. In other words, the goal is to make a simulator that the viewer experiences the virtual environment as similar as possible with the real world. Figure 1 shows the schematic diagram of the proposed WTS. The hexahedral-shaped device as shown in the left-top side of Figure 1 is the WTS. The virtual construction is displayed around the viewer inside the WTS using four projectors as shown in the right-hand side of Figure 1.The viewers attach the sensor to their belt as shown in Figure 1(C), thus the viewer's rotation and step rates are estimated near the center of gravity of human body.

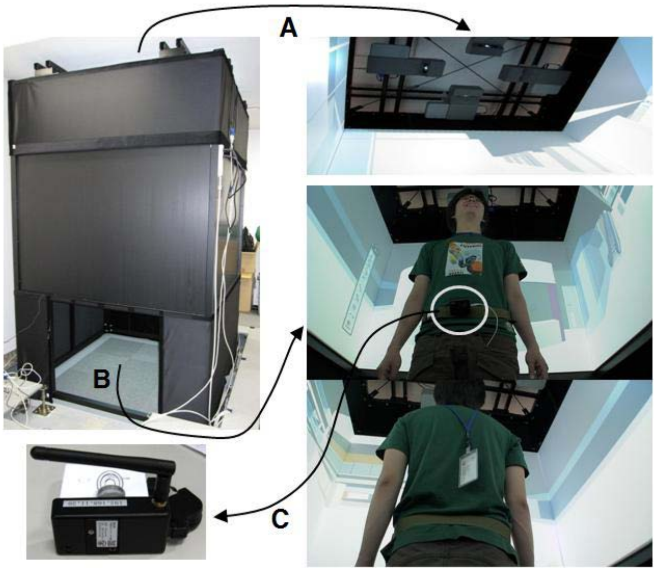

Figure 1. Schematic diagram of WTS: (A) multi-projector, (B) virtual scene displayed around viewer in WTS, and (C) sensor module.

A configuration of an omni-directional display screen to display a 360-degree virtual construction is described in Section 2.1, and a locomotion interface recognizing five movements defined based on actual people is described in Section 2.2.

2.1. Omni-directional Display Screen

The proposed WTS displays virtual constructions on an omni-directional display screen in order to provide the viewers with a feeling that they are standing inside real environment. That makes it possible to enable the viewers to navigate the virtual constructions while preserving the viewer's sense of absolute direction.

The shape of the WTS is a hexahedron to cover 360-degree horizontal view in which the viewer stands on the center, and each wall is used as a display screen. To determine the size of the screen and configuration of the WTS, it is assumed that the viewer's height is 170 centimeter (cm) and the height of eyes to see the virtual construction from the floor is 160 cm. Then, the center of the display screens is perpendicular to the eyes of viewer to make the viewer immersed in the virtual constructions, and the size of each screen is 1600×1200, as aspect ratio of projected image is 4:3. The vertical angle of view of the proposed WTS is about 74 degree, $\tan^{-1}\left(\frac{600}{800}\right) \times 2$, whereas pixel efficiency used to determine optimal configuration of display system in [4] is 100%. The pixel efficiency means how many pixels from a projector are displayed on each screen, thus the proposed WTS can provide undistorted high-resolution images to the viewer.

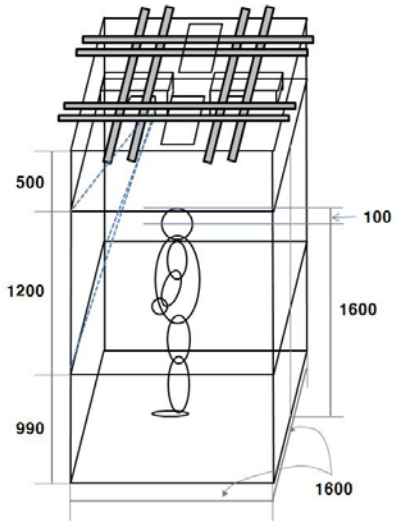

Figure 2. Design drawing of configuration of the proposed WTS.

The projectors are set above the viewer, and each project displays image corresponding to each viewport to a screen. The location of projections is designed to maximize space efficiency used in [4], and is designed to overcome a limitation of rear-projection-based screen that the viewer obstructs light from the projector and large shadow is projected on the screen. The space efficiency is determined by the ratio of displayed volume to overall dead volume occurred by projection installation and the space efficiency of proposed simulator is 3:5 (1200:500+1000). Figure 2 shows a sketch of the configuration of the proposed WTS. The WTS can be installed in a normal-sized room, as the overall size of the proposed WTS is 1600(W)x2600(H)x1600(D), and can be easily built, as only four flat screens and four projects are used for the proposed WTS.

2.2. User Interface

To navigate the virtual construction displayed around in the viewer, a small-sized wireless sensor module, detecting the viewer's WIP and orientation, is used, and Section 3 describes about the sensor module. In this section, we explain how to design user input interfaces for the viewers to navigate virtual constructions, based on an actual person's movement.

In the most cases, an actual person stands (1), turns (2), walks forward (3), walks forward while turning left (4) or right (5) to go his/her destination in a building. Therefore, the user input interface is designed based on these five movements. The five movements can be classified into three cases: 1) composed of single action, 2) composed of more than actions, and 3) composed of more than one movement. The first and second movements, standing and turning, are included in the first case. The standing is recognized when the viewer does not have any other movements. The turning is recognized when the viewer rotates his/her body from a principal axis, and the orientation angle is estimated by the sensor module.

The second case includes walking forward, which is more complicated movements, and user interface for this movement is designed as similar as possible with actual people's movement in limited spaces. An actual person walks forward by moving firstly the left or right foot to forward and then moving the other foot to much more forward. In the proposed interface, the viewer moves his/her feet in an up-down fashion similar to really walking. In this case, the sensor detecting gravity is used to check his/ her foot is moved in an up-down fashion. The most natural movements in navigating buildings, walking forward and rotation, are performed on the center in the simulator.

The third case includes walking forward while turning left or right. These movements are detected by integrating walking forward or backward with rotating left or right. Based on these five movements detected by a sensor module in the WTS is used to navigate the virtual construction.

3. Navigation using Sensor Modules

To recognize the viewer's movement for a locomotion interface based on WIP, described in Section 2.2, we use a sensor module composed of an accelerometer, a gyrosensor, a magnetometer, and thermometers for each axis of the gyrosensor.

Based on analyzing patterns for vertical vector caused by human walking locomotion, we have observed that there is a very simple pattern similar to sine waves with different phases but at the same frequency. Therefore, peak detection and a gradient test algorithm for vertical acceleration are used to detect a cycle of walking in place. However, it is not possible to detect vertical acceleration using only the accelerometer, because attitudes on the sensor module are always changed while walking. To solve this problem, the present paper estimates the direction of gravity, referred to as gravitational vector, to adjust the effect of attitude changes on the attached sensors while walking using a Kalman filter framework proposed in [11]. Here, an initial gravitational vector for the Kalman filter framework is set an output value when the accelerometer is stationary, and the output values of the accelerometer after initialization of the Kalman filter framework are projected to the gravitational vector to estimate vertical acceleration.

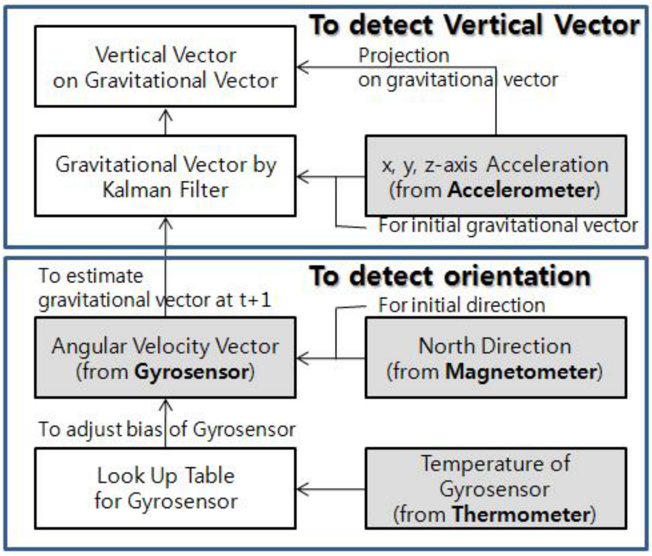

Figure 3. Flowchart to detect WIP and orientation using sensor module.

The orientation of viewers in the WTS is estimated using three sensors, the magnetometer, gyrosensor, and thermometer. The magnetometer estimates actual north direction, which is used as an initial direction for angular velocity vector, and the gyrosensor estimates angular velocity vector. The angular velocity vector is also needed to estimate the gravitational vector generated by the Kalman filter framework. The gyrosensor has sometimes errors, referred to as bias, between actual output of the gyrosensor and the estimated value. In the experiments, there are some rules between the bias and temperatures of each axis of the gyrosensor. The temperatures are estimated by the thermometer. Therefore, we make a look-up table to estimate the bias, and the bias is used to compensate errors between the actual output and estimated value. For more information to detect human's locomotion interface, please refer to the paper of Kourogi and Kurata [11].

Figure 3 shows a flowchart to detect the WIP and orientation using the sensor module, where the gray color indicates values from the each sensor.

4. Photorealistic Avatar for Face-to-Face Communication

In some buildings, such as public institutions, guides provide instructions to visitors or customers to help them reach their destination. In the virtual building, guides are displayed as photorealistic avatars. The image of the person of the person that is used to create the photorealistic avatar is extracted by the camera in front of the camera attached to the remote computer, and the photorealistic avatar is displayed in the virtual construction inside the simulator, as shown in Figure 4.

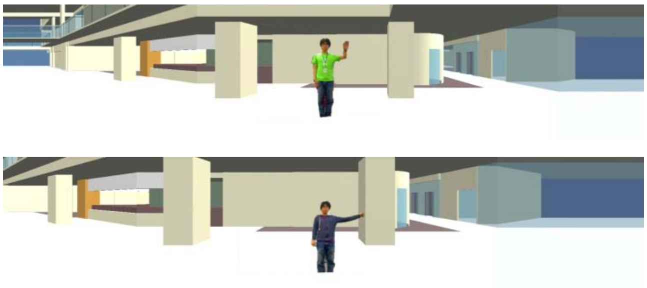

Figure 4. Photorealistic avatar integrated into a virtual building.

In the present paper, it is assumed that the moving foreground in front of the fixed camera outside the simulator is an individual whose image will be used to generate photorealistic avatar. Therefore, background subtraction is used to extract the appearance of the guide from images capture by a camera. There are four requirements for the background subtraction: 1) extraction should be performed in real-time, 2) memory consumption should be limited, 3) image should be extracted with little noise, and 4) the boundaries of the avatar should be clear. Accordingly, the present paper uses a codebook-based method using a graph-cut for background subtraction that satisfies these requirements. In this method, a codebook-based approach [9] is used for real-time implementation and memory minimization, and a graph-cut [10] is used to eliminate noise and clarify boundaries. Please refer to the paper [12] for more information about the background subtraction.

5. Experimental Results

The sensor module we used in this paper is composed of four components, as shown in Table 1, and each sensor data was sampled at 100Hz. All the components are included in one sensor module, as shown in Fig. 1(C), and the size of the sensor module is 7.5cm×3.5cm×1.5cm. Therefore, it was easy to attach the sensor module to the viewers' belt, thereby realizing the hands-free operation, which makes it possible to evaluate the usability of handheld devices while walking around in virtual environment.

Table 1. Lists of name of sensors and manufactures

	Name	Manufacture
Accelerometer	LIS3LV02DQ	ST micro
Gyrosensor	XV-3500CB	Epson Toycom
Magnetometer	AM1304	Aichi Steel
Thermometers	TMP35	Analog Devices

(a)

(b)

Figure 5. Example of photorealistic construction: (a) actual office of our center and (b) 3D reconstruction based on several photos.

As an application, we apply the proposed WTS to navigation in photorealistic constructions modeled by photos captured from real indoor environments. Virtualized real scenes reconstructed from photos or sequential images can help to enable virtual constructions to enhance reality, because this can help to reduce the gap between the real and virtual world for the proposed WTS. The reconstructed virtualized real scenes are referred to as Photorealistic Construction in this paper. We use an interactive indoor 3D constructor, which enables producers to create 3D photorealistic constructions efficiently and intuitively. The purpose of the application is that the viewers experience a 360-degree remote indoor place displayed around them, without visiting the place in person. The photorealistic construction was constructed based on a method proposed by Ishikawa et al. [13], and Figure 5 shows a result of photorealistic construction and an actual environment of our office.

6. Conclusion

In the present paper, we presented a walkthrough simulator in which the display is omni-directional so as to preserve the viewer's sense of direction. In addition, the input interface realizes hands-free operation. That made it possible to evaluate the usability of handheld devices or even just a paper map while walking around in virtual environments. To improve the virtual environments, the proposed WTS supported face-to-face communication between the WTS and a remote desktop PC, and was also applied to navigate photorealistic constructions.

However, the vertical angle of view of the proposed WTS was not enough to be satisfied. Therefore, in the future studies, we will investigate how to improve the vertical angle with high-resolution images without any distortions.

Acknowledgement: This research was entrusted by the Ministry of Economy, Trade and Industry (METI).

References

[1] M. Usoh, K. Arthur, M. Whitton, R. Bastos, A. Steed, M. Slater, and P.J. Frederick. Walking > walking-in-place > flaying, in virtual environments, Proceedings of SIGGRAPH, pp. 359-364, 1999.

[2] M.A. Teitel. The Eyephone: a head-mounted stereo display, Proceedings of SPIE, vol. 1256, pp. 168-171, 1990.

[3] C. Cruz-Neira, D.J. Sandin, T.A. Defanti, R.V. Kenyon, and J.C. Hart. The Cave: audio visual experience automatic virtual environment, Communication of ACM, vol. 35, no. 6, pp. 64-72, 1992.

[4] H. Iwata. Full-surround image display technologies, International Journal of Computer Vision, vol. 58, no. 3, pp. 227-235, 2004.

[5] J.M. Hollerbach. Locomotion interface, in Handbook of Virtual Environments Technology, K.M. Stanney, ed., Lawrence Erlbaum Associates, pp. 239-254, 2006.

[6] L. Bouguila, F. Evequoz, M. Courant, and B. Hirsbrunner. Walking-pad: a step-in-place locomotion interface for virtual environment, Proceedings of International Conference on Multimodal Interfaces, pp. 77-81, 2004.

[7] L. Lichtenstein, J. Barabas, R.L. Woods, and E. Peli. A feedback-controlled interface for treadmill locomotion in virtual environments, ACM Transactions on Applied Perception, vol. 4, issue 1, 2007.

[8] S. Barrera, H. Takahashi, and M. Nakajima. Hands-free navigation methods for moving through a virtual landscape walking interface virtual reality input devices, Proceedings of the Computer Graphic International, pp. 388-394, 2004.

[9] K. Kim, T.H. Chalidabhongse, D. Harwood, and L. Davis. Real-Time Foreground-Background Segmentation using Codebook Model, Real-Time Imaging, vol. 11, no. 3, pp. 172-185, 205.

[10] Y. Boykov, O. Veksler, and R. Zabih. Fast Approximation Energy Minimization via Graph Cuts, IEEE Trans. Pattern Anal. Mach. Intell., vol. 23, no. 11, pp. 1222-1239, 2001.

[11] M. Kourogi and T. Kurata. Personal Positioning based on Walking Locomotion Analysis with Self-Contained Sensors and a Wearable Camera, Proceedings of International Symposium on Mixed and Augmented Reality, pp. 103-112, 2003.

[12] A. Park, T. Okuma, and T. Kurata. Codebook-based Background Subtraction to Generate Photorealistic Avatars in a Walkthrough Simulator, Proceedings of International Symposium on Visual Computing, to be accepted, 2009.

[13] T. Ishikawa, K. Thangamani, M. Kourogi, A.P. Gee, W.W. Mayol-Cuevas, K. Jung, and T. Kurata. In-Situ 3D Indoor Modeler with a Camera and Self-contained Sensors, Proceedings of International Conference on Human-Computer Interaction, pp. 454-464, 2009.