Locomotion by Natural Gestures for Immersive Virtual Environments

Andrea Ferracani, Daniele Pezzatini, Jacopo Bianchini, Gianmarco Biscini, Alberto Del Bimbo
MICC - University of Florence
Firenze, Italy
[name.surname]@unifi.it

ABSTRACT

In this paper we evaluate methods to move 'naturally' in an Immersive Virtual Environment (IVE) visualised through an Head Mounted Display (HMD). Natural interaction is provided through gesture recognition on depth sensors' data. Gestural input solutions in the literature to provide locomotion are discussed. Two new methods for locomotion are proposed, implemented in a framework used for comparative evaluation. Perceived naturalness and effectiveness of locomotion methods are assessed through qualitative and quantitative measures. Extensive tests are conducted on the locomotion considering also: 1) obstacles in navigation; 2) interaction with virtual objects during locomotion. This is done with the aim to identify methods capable to provide a full body experience in an IVE. Results show that one of the methods for locomotion we propose has a performance comparable to established techniques in literature. Outcomes may be exploited to improve the naturalness of users' movements in IVEs and help to unlock new strategies in providing IVEs for learning, training, collaboration and entertainment, also with respect to users with disabilities.

Keywords

Locomotion; Immersive Virtual Reality; Natural Interaction; Head Mounted Display; Motion Tracking

1. INTRODUCTION AND RELATED WORK

Effective IVEs require intuitive interfaces controlled in a way that resembles real world experiences [16]. 'Incompatible spaces' is a common issue that researchers in HCI have to face when providing natural interaction in IVEs. In fact, IVEs allow free movement and infinite walking but the physical environment where the simulation is taking place presents spatial constraints. There are several solutions in the literature to allow infinite walking in IVEs, still maintaining in users a realistic sensation of walking. These solutions can be classified in four groups which exploit:

AltMM'16, October 15-19, 2016, Amsterdam, Netherlands
© 2016 ACM. ISBN 978-1-4503-4521-7/16/10...$15.00
DOI: http://dx.doi.org/10.1145/2983298.2983307

a) *Additional Hardware*: unidirectional and omnidirectional treadmills, footpads and rotating spheres have been used to simulate natural walking maintaining fixed the user position in the environment [5, 12]. These approaches are not easy to set up, require to secure users, are cumbersome and costly. Furthermore Natural User Interfaces (NUIs) do not contemplate the mediation of physical devices as controllers;
b) *Redirected Walking*: a set of reorientation and repositioning techniques which exploit virtual *stimuli* [13], e.g. giving the impression of walking straight to users moving in a circle [11] or using procedural layout generation [17]. Although these methods provide a good sense of presence, obstacles and physical constraints of the environment are still an issue;
c) *Software-based navigation*: interfaces featuring positional tracking supported by navigation tools. In [1] the tracking area, visualised as a 'magic carpet', can be repositioned using an appropriate tool for long-distance navigation. In [3] positional tracking is used in a restricted walking space whose physical boundaries, displayed in the IVE as a barrier tape, can be moved with a joystick. These solutions are not fully natural and require from users additional cognitive efforts while moving;
d) *Gesture Recognition using cameras*: vision-based methods for locomotion recognition have the advantage of not requiring additional hardware as interface controller. They solve several issues with respect to the solutions in a) and b), i.e. infinite walking and space constraints. However, it is difficult to design and agree on a natural gesture to move. Furthermore, fatigue can affect the use of gesture-controlled interfaces, especially when reproducing a continuous action such as walking.

Methods of locomotion proposed in this paper fall within solutions in d). These are the most appropriate for NUIs mimicking real world interactions without the need of specific controllers. Several gestures have been defined in the literature which allow infinite walking in IVEs. Walking-In-Place (WIP) is the most common interaction paradigm: users can move in the IVE while remaining stationary [4, 11, 18]. Although WIP is usually referred as a form of compensating locomotion, the gesture is less frustrating for users than natural locomotion. Users moving naturally should repeatedly go forward and backward due to physical space constraints [8]. The Shake-Your-Head gesture in [15] allows the user to interact with the interface through head oscillations (i.e. as a transposition of the head movements observable in natural walking). Unlike the WIP technique, the user can both stand or sit in front of the interface. This solves the fatigue problem caused by both standing and walking. Arm-

Swing is a gesture performed oscillating the arms alternatively along the hips by a person as it is observed in natural walking. There's no implementation in the literature of a specific recogniser for Arm-Swing but the gesture is ranked second in the user study conducted in [8] where participants were given complete freedom in choosing gestures to complete tasks in a videogame. Free hand interactions have also been proposed and evaluated in literature to support locomotion in IVEs [2] as a mean to determine the direction of the movement.
The paper is organised as follows: in Sec. 2 we discuss the locomotion methods proposed and used in the evaluation; in Sec. 3 the framework and the input/output devices are presented; results, assessed through qualitative and quantitative experiments, are shown in Sec. 4.

2. NATURAL INTERACTIONS

Defining gestures in 3D IVEs exploiting natural interaction is easier than in 2D interfaces for the higher expressiveness that can be obtained by users simply acting like they do in the real world. 'Guessability' studies exploiting user-centered design show that in this scenario users' gestures are dominantly physical (e.g. walking moving knees) and metaphorical (e.g. selecting objects through pointing) [8, 10]. Building upon these studies, we evaluate four gestures for locomotion in IVEs (see Fig.1). Among these gestures, two are derived from the literature whilst the two others are novel. Locomotion methods have been chosen considering: 1) if gestures have been validated in similar studies; 2) the naturalness of the gestures with respect to the real world.
WIP (Walk-In-Place) The user walks in a stationary position. It is the most used in the literature, validated through qualitative and quantitative studies [4, 6, 11, 18];
Swing (Arms Swing) The idea is to replicate the natural oscillations of the arms during locomotion. It is a gesture demonstrated being actually performed by users freely interacting with a IVE [8];
Tap We propose a metaphorical gesture [10] for locomotion consisting in a tap with the index finger in the direction the user wants to start walking. It is a gesture not so far from the real world: people commonly use the index finger to show a walking direction;
Push We propose a metaphorical gesture consisting in closing and opening the hand while translating the hand itself forward with respect to the user elbow. In the real world it is the typical gesture to control locomotion machines moving a lever.
Shake-Your-Head gesture was not included in the study for two main reasons: 1) it can neither be classified as a natural gesture nor as a metaphorical one because the gesture has never been proposed by users in guessability studies; 2) it may cause motion sickness if used repeatedly in a HMD setup.
As regard to locomotion we must point out that the framework provides discrete and not continuous gestures in time. The reason is that users are aware of the fact that they are using methods of compensating locomotion and not natural locomotion. This is an essential feature for the usability of the IVE that otherwise: 1) it would strain too much the user with continuous activity (i.e. using WIP and Swing); 2) it would force the user to hold at least one of the hands always busy making it difficult to interact with virtual objects (i.e. using Tap). Once activated locomotion can be stopped with

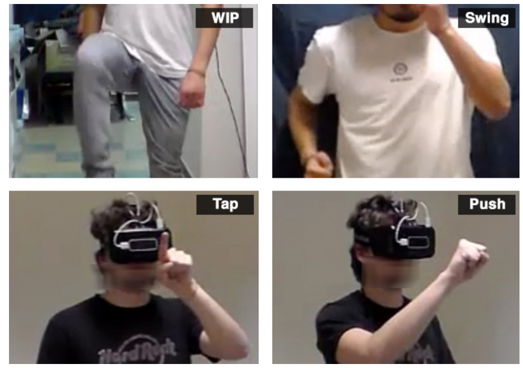

Figure 1: The four gestures for locomotion evaluated in this paper.

a 'Stop gesture' that the user can perform opening his hand in his field of view. This gesture is motivated in [10] where it is demonstrated to be the preferred one by users performing a generic 'stop' action.

3. THE FRAMEWORK

The framework[1] consists in a library we developed that enables a first person controller to navigate and interact in IVEs created for the Unity3D engine[2] moving through the natural gestures described in Sec. 2. Basic interaction with virtual objects is also made available. The library allows to easily connect the interactive IVE with output and input devices, namely with an Head Mounted Display which visualises the 3D environment, and two tracking devices which provide the motion data gestures' detection relies on:
• *A Kinect v2*[3]. It tracks 25 body joint with millimetre accuracy and provides frame by frame data by which the WIP and the Swing gestures for locomotion are detected;
• *A Leap Motion*[4]. It tracks positions and rotations of each finger bone (24 per hand); mounted on the HMD facing in the user's field of view it is used to track hand movements and detect Tap and Push gestures for locomotion, Stop for interrupting locomotion, and gestures for interaction with virtual objects (i.e. pointing and grabbing).
For WIP and Swing gestures recognition we exploited the Microsoft Visual Gesture Builder NUI tool that generates gesture databases used to perform run-time detection through machine learning techniques (e.g. AdaBoost) applied to skeleton data. Leap Motion SDK instead provides Tap and Grab gestures recognition natively. For Push and Point gestures we have trained *ad hoc* classifiers. Looking direction equals walking direction in HMD for all the different gestures.
The library also includes UI components helpful to the user while exploring the IVE. Indicators of current direction and state of gesture recognition are superimposed on the 3D environment in order to give users proper awareness due to the absence of proprioceptive feedback. Furthermore, a virtual representation of user's hands is provided in the 3D environment to enhance sense of presence and ease virtual interactions.

[1]Demo video available at https://vimeo.com/172710194/
[2]https://unity3d.com/
[3]https://developer.microsoft.com/en-us/windows/kinect
[4]https://www.leapmotion.com/

4. EXPERIMENTAL RESULTS

An evaluation was conducted to determine how the proposed methods for locomotion in IVEs perform in terms of effectiveness and perceived naturalness. The four locomotion methods presented in Sec. 2 (i.e. WIP, Swing, Tap, Push) are evaluated comparatively, asking users to complete tasks of increasing difficulty.

Participants and procedure.

Evaluation was conducted with 19 participants (11 males and 8 females) aged between 21 and 39 years old (average 26.4, $\sigma = 5.8$). None of the participants had previous experience with IVEs or HMDs, but they reported a medium to high familiarity with technology (average of 4.4 on a 1 to 5 rating scale) and previous experience with first-person video games (average of 3.8 on a 1 to 5 rating scale). Locomotion methods and gestures for interactions were explained to all participants before the test. At the end of the session, participants were asked to fill a questionnaire.

Tasks and setting.

For the tests we created an IVE representing a forest. Two position in the virtual environment were defined by visual markers: a starting position A and a destination position B (see Fig.2).

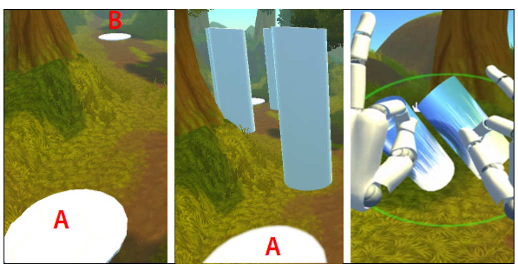

Figure 2: Task scenarios (i.e. T1, T2, T3).

Participants were asked to perform six tasks using all the four locomotion methods. In the easiest task users were asked to move from A to B. Other tasks were defined combining further difficulties such as going back to position A, avoiding obstacles placed along the locomotion path and, at the same time, bringing an object from a position to another. Most of the cited works in the literature evaluate the naturalness of gestures for locomotion and interaction with virtual objects as separate topics [2, 9, 19] and, to the best of our knowledge, solving both problems together it is still an open issue which needs to be addressed in IVEs applications. For this reason, we introduced some tasks that contemplate the use of the Grab gesture to relocate a virtual object in the environment. The following tasks were defined:
T1 Move from position A to position B.
T2 Move from position A to position B and back to A.
T3 Move from position A to position B, avoiding obstacles on the path.
T4 Move from position A to position B and then back to A, avoiding obstacles on the path.
T5 Move from position A to position B, grab an object and then bring it back to A.
T6 Move from position A to position B, grab an object and then bring it back to A, avoiding obstacles on the path.
The order of the used locomotion methods was randomised so to eliminate potential order-related bias. Since the Swing and Grab gestures are incompatible (i.e. Swing assumes that both arms are occupied), results of T5 and T6 are n.a.

Measures.

Locomotion techniques in the framework were evaluated using both qualitative and quantitative methods. Naturalness and effectiveness of locomotion were assessed using the following measures:
Perceived Naturalness. Following the heuristic evaluation method for natural engagement in IVEs proposed in [14], we provided a questionnaire to collect subjective measures of naturalness of locomotion gestures from the participants, expressed on a 1 to 7 scale.
Overall preference. At the end of each session, we asked users to indicate which method they preferred.
Time Completion. A quantitative measure of the time required to complete each task. We did not define a maximum execution time and all users were able to complete all the tasks.
Collision Avoidance. This measure was proposed in [7] as a meaningful way to evaluate locomotion in IVEs. In two of the tasks including obstacles (i.e. T4 and T6) we counted the number of collisions occurred.

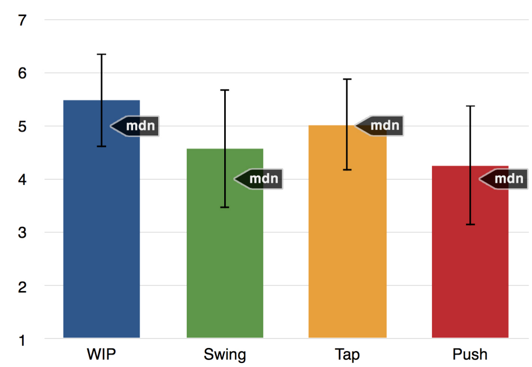

Figure 3: Perceived Naturalness of locomotion methods. The higher the better. The black bars stand for the standard deviation.

Results.

Qualitative and quantitative results were statistically analysed to obtain a comparative evaluation of the four locomotion methods. Qualitative comparison in terms of *Perceived Naturalness* is shown in Fig. 3. All methods have a good rating, but highest scores were obtained by WIP (avg 5.47, mdn 5) and Tap (avg 5.02, mdn 5). Results of the *Time Completion* (see Table 1) and *Collision Avoidance* (see Table 2) tests reveal sensible differences between methods in terms of effectiveness. WIP and Tap methods result to be the fastest and less prone to collisions in almost every task. Tap in particular performs better than other methods in T5 (Table 1) and in T4 and T6 (Table 2). The gesture seems to overcome WIP in tasks that contemplate hand-based interaction (i.e. Grab) and *Collision Avoidance*. An explanation could be given by the verbal considerations of some testers that reported WIP to require a sort of bilateral integration between hands and legs. Results from the *Overall preference* questionnaire indicate that more than half of testers (10 out of 19) would choose Tap as locomotion method, while the remaining preferences were for Push (6 out of 19) and WIP (3 out of 19). The outcomes of the evaluation suggest that even though WIP is by far the most used locomotion technique in IVEs, novel gestures such as Tap could be adopted with comparable results in terms of effectiveness of user experience. Results in Tables 1 and 2 are preliminary: analysis of variance for statistical significance of means between groups of testers are needed and will be the subject of future work.

Table 1: Time Completion in seconds. The lower the better.

	WIP		Swing		Tap		Push	
	Avg	σ	Avg	σ	Avg	σ	Avg	σ
T1	**15**	2	**15**	2	16	3	21	6
T2	40	7	39	6	**30**	2	35	5
T3	**17**	2	19	7	20	5	26	9
T4	43	10	42	10	**37**	6	52	16
T5	56	13	n.a.	n.a.	**51**	9	64	21
T6	**53**	21	n.a.	n.a.	57	23	75	34

Table 2: Collision Avoidance results showing number of collisions. The lower the better.

	WIP		Swing		Tap		Push	
	Avg	σ	Avg	σ	Avg	σ	Avg	σ
T4	0.46	0.78	0.54	0.77	**0.38**	0.61	0.84	0.98
T6	0.92	1.11	n.a	n.a	**0.61**	0.96	1.00	0.91

5. CONCLUSIONS

Providing IVEs' users with the best natural experience is a challenging task. Commonly IVEs are mediated by displays mounted on the head and there's a physical gap between real and virtual space. Infinite locomotion in virtual environments collides with the constraints of their fruition in spaces closed by walls or obstructed by obstacles. Natural interaction provides a solution to these issues through gesture recognition. In this paper we identify and comparatively evaluate four methods of locomotion (i.e. WIP, Tap, Swing, Push). Qualitative and quantitative experiments are conducted through user testing. Results show that two of the four methods perform better than the others (i.e. WIP and Tap) and that the Tap gesture we propose has similar and in some tasks better performance than the well established WIP locomotion technique. This evidence may be useful to researchers and interaction experts for designing IVEs and for providing whole body natural experiences. Furthermore, performance of Tap suggests that hand-based gestures for locomotion deserve further investigation. Although being metaphorical the Tap gesture was perceived as natural by testers. Its adoption could provide some advantages in certain scenarios: for example, it could be used in configurations with a seated user, resulting in a reduction of physical fatigue, and improve accessibility to IVEs even for users with reduced mobility.

References

[1] J. Butterworth, A. Davidson, S. Hench, and M. T. Olano. 3dm: A three dimensional modeler using a head-mounted display. In *Proceedings of the 1992 Symposium on Interactive 3D Graphics*, I3D '92, pages 135–138, New York, NY, USA, 1992. ACM.

[2] G. Caggianese, L. Gallo, and P. Neroni. Design and preliminary evaluation of free-hand travel techniques for wearable immersive virtual reality systems with egocentric sensing. In *Augmented and Virtual Reality*, pages 399–408. Springer, 2015.

[3] G. Cirio, M. Marchal, T. Regia-Corte, and A. Lécuyer. The magic barrier tape: A novel metaphor for infinite navigation in virtual worlds with a restricted walking workspace. In *Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology*, VRST '09, pages 155–162, New York, NY, USA, 2009. ACM.

[4] J. Feasel, M. C. Whitton, and J. D. Wendt. Llcm-wip: Low-latency, continuous-motion walking-in-place. In *3D User Interfaces, 2008. 3DUI 2008. IEEE Symposium on*, pages 97–104, March 2008.

[5] H. Iwata and Y. Yoshida. Path reproduction tests using a torus treadmill. *Presence: Teleoperators and Virtual Environments*, 8(6):587–597, 1999.

[6] J.-S. Kim, D. Gračanin, K. Matković, and F. Quek. Sensor-fusion walking-in-place interaction technique using mobile devices. In *2012 IEEE Virtual Reality Workshops (VRW)*, pages 39–42, March 2012.

[7] S. Livatino and C. Köffel. Handbook for evaluation studies in virtual reality. In *Virtual Environments, Human-Computer Interfaces and Measurement Systems, 2007. VECIMS 2007. IEEE Symposium on*, pages 1–6. IEEE, 2007.

[8] J. Norton, C. A. Wingrave, and J. J. LaViola Jr. Exploring strategies and guidelines for developing full body video game interfaces. In *Proceedings of the Fifth International Conference on the Foundations of Digital Games*, pages 155–162. ACM, 2010.

[9] K.-B. Park and J. Y. Lee. Comparative study on the interface and interaction for manipulating 3d virtual objects in a virtual reality environment. *Transactions of the Society of CAD/CAM Engineers*, 21(1):20–30, 2016.

[10] T. Piumsomboon, A. Clark, M. Billinghurst, and A. Cockburn. User-defined gestures for augmented reality. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '13, pages 955–960, New York, NY, USA, 2013. ACM.

[11] S. Razzaque, D. Swapp, M. Slater, M. C. Whitton, and A. Steed. Redirected walking in place. In *Proceedings of the Workshop on Virtual Environments 2002*, EGVE '02, pages 123–130, Aire-la-Ville, Switzerland, Switzerland, 2002. Eurographics Association.

[12] K. M. Stanney and M. Zyda. Virtual environments in the 21st century. *Handbook of virtual environments: Design, implementation, and applications*, pages 1–14, 2002.

[13] E. A. Suma, D. M. Krum, and M. Bolas. *Human Walking in Virtual Environments: Perception, Technology, and Applications*, chapter Redirected Walking in Mixed Reality Training Applications, pages 319–331. Springer New York, New York, NY, 2013.

[14] A. Sutcliffe and B. Gault. Heuristic evaluation of virtual reality applications. *Interacting with computers*, 16(4):831–849, 2004.

[15] L. Terziman, M. Marchal, M. Emily, F. Multon, B. Arnaldi, and A. Lécuyer. Shake-your-head: Revisiting walking-in-place for desktop virtual reality. In *Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology*, pages 27–34. ACM, 2010.

[16] L. Turchet. Designing presence for real locomotion in immersive virtual environments: an affordance-based experiential approach. *Virtual Reality*, 19(3-4):277–290, 2015.

[17] K. Vasylevska, H. Kaufmann, M. Bolas, and E. A. Suma. Flexible spaces: Dynamic layout generation for infinite walking in virtual environments. In *3D User Interfaces (3DUI), 2013 IEEE Symposium on*, pages 39–42, March 2013.

[18] J. D. Wendt, M. C. Whitton, and F. P. Brooks. Gud wip: Gait-understanding-driven walking-in-place. In *2010 IEEE Virtual Reality Conference (VR)*, pages 51–58, March 2010.

[19] Y. Zhang, S. Stellmach, A. Sellen, and A. Blake. The costs and benefits of combining gaze and hand gestures for remote interaction. In *Human-Computer Interaction*, pages 570–577. Springer, 2015.