Natural Walking in Virtual Reality

A Review

Nilsson, Niels Chr.; Serafin, Stefania; Steinicke, Franke; Nordahl, Rolf

Published in:
Computers in Entertainment

DOI (link to publication from Publisher):
10.1145/3180658

Publication date:
2018

Document Version
Accepted author manuscript, peer reviewed version

Link to publication from Aalborg University

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy
If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.
Natural Walking in Virtual Reality: A Review

NIELS CHRISTIAN NILSSON and STEFANIA SERAFIN, Aalborg University Copenhagen
FRANK STEINICKE, University of Hamburg
ROLF NORDAHL, Aalborg University Copenhagen

Recent technological developments have finally brought virtual reality (VR) out of the laboratory and into the hands of developers and consumers. However, a number of challenges remain. Virtual travel is one of the most common and universal tasks performed inside virtual environments, yet enabling users to navigate virtual environments is not a trivial challenge—especially if the user is walking. In this article, we initially provide an overview of the numerous virtual travel techniques that have been proposed prior to the commercialization of VR. Then we turn to the mode of travel that is the most difficult to facilitate, that is, walking. The challenge of providing users with natural walking experiences in VR can be divided into two separate, albeit related, challenges: (1) enabling unconstrained walking in virtual worlds that are larger than the tracked physical space and (2) providing users with appropriate multisensory stimuli in response to their interaction with the virtual environment. In regard to the first challenge, we present walking techniques falling into three general categories: repositioning systems, locomotion based on proxy gestures, and redirected walking. With respect to multimodal stimuli, we focus on how to provide three types of information: external sensory information (visual, auditory, and cutaneous), internal sensory information (vestibular and kinesthetic/properciptive), and efferent information. Finally, we discuss how the different categories of walking techniques compare and discuss the challenges still facing the research community.

CCS Concepts: • Human-centered computing → Virtual reality;

Additional Key Words and Phrases: Virtual reality, virtual travel, walking, naturalness

ACM Reference format:
https://doi.org/10.1145/3180658

1 INTRODUCTION

Virtual reality (VR) is no longer confined to the laboratories of larger public and private institutions. In 2016, VR entered the homes of consumers for the first time. We use the designation VR to denote systems that, through high-fidelity tracking and displays, allow users to interact naturally within computer-generated environments; that is, VR supports a sensorimotor loop similar to that of the real world and thereby enables users to perceive and act as they would in reality. The popularization of VR was partially instigated by a generation of young entrepreneurs and...
crowdfunding campaigns (Morie 2014) and has since been intensified through the involvement of
large technology corporations—most prominently Facebook, Samsung, Google, and Sony. Despite
its recent popularity, VR is by no means a novelty. In fact, it has been more than 50 years since
Sutherland (1965) presented his vision of the ultimate display. Specifically, Sutherland described
the ultimate display as “a room within which the computer can control the existence of matter.
A chair displayed in such a room would be good enough to sit in. Handcuffs displayed in such
a room would be confining, and a bullet displayed in such a room would be fatal. With appro-
priate programming such a display could literally be the Wonderland into which Alice walked”
(Sutherland 1965, p. 508). However, natural walking remains one of the biggest challenges facing
researchers and developers aspiring to provide users with access to digital wonderlands such as
the ones envisioned by Sutherland.

In this article, we present a review of the various approaches to facilitating natural walking in
virtual environments. Unlike critical reviews, the selection of sources did not involve a search of
all potentially relevant work based on reproducible criteria (Cook et al. 1997). Instead, the litera-
ture forming the basis for the narrative review was identified based on referral sampling of both
well-known and recent work detailing surveys of topics relevant to natural walking in virtual en-
vironments (Bowman et al. 2004; Fontana and Visell 2012; Nilsson et al. 2016b; Steinicke et al. 2013;

The general aim of the review is to provide an overview of the large body of work on walk-
ing within virtual environments. The review seeks to position virtual walking within the wider
category of virtual travel techniques (Section 2) and highlight the challenges that are particularly
pertinent in relation to virtual walking (Section 3). The two primary challenges associated with
natural walking in VR are (1) enabling unconstrained walking in virtual environments that are
larger than the physical tracking space and (2) providing users with appropriate multisensory
stimuli in response to their interaction with the virtual environment. Consequently, the review
aims to survey and categorize various attempts at meeting these two challenges (Section 4 and 5).
Finally, based on the literature surveyed throughout the article, we present current challenges and
potential directions for future work (Section 6).

2 VIRTUAL TRAVEL TECHNIQUES

To most people, the act of moving from one place to another is a common everyday activity.
We cover shorter distances on foot, and longer distances are traversed with the aid of humanly
propelled or motorized vehicles, such as bicycles, cars, and planes. In regard to interaction with
three-dimensional interfaces, Bowman et al. (2004) similarly describe virtual travel as one of the
most common and universal forms of interaction. Moreover, virtual travel is usually secondary to
other tasks, such as exploration, searching, and maneuvering. Thus, it is necessary for users to be
able to virtually travel with relative ease and without needing to assign much explicit attention to
the act of traveling itself. However, enabling users to do so is not trivial, especially if the person
traveling through a virtual environment is doing so using a commercially available VR system set
up inside an average living room.

Generally speaking, the task of moving from one place to another in real and virtual environ-
ments (i.e., navigation) can be broken down into wayfinding (the cognitive component involving
path planning and decision making) and travel (the motor component) (Bowman et al. 2004). Bow-
man et al. (1997) describe that travel can be further decomposed into three subtasks: direction
or target selection (specification of where to move), velocity/acceleration selection (specification of
movement speed), and conditions for input (specification of how travel is instigated, continued, and
terminated). For example, when traveling by car, the driver uses the steering wheel to select the
direction of heading, and the velocity is selected using a combination of the gas pedal, shift lever,
and breaks, which also define the conditions for input. While this decomposition provides a useful lens through which to view individual travel techniques, it does not provide a broader picture of how virtual travel may be accomplished. Inspired by existing categorizations of travel techniques (Bowman et al. 2004; Slater and Usoh 1994; Suma et al. 2012; Wendt 2010), Nilsson et al. (2016b) divide existing travel techniques into dichotomous categories. Specifically, they distinguish between travel techniques based on whether the user is stationary or moving, whether the techniques involve virtual vehicles or not, and whether the techniques qualify as mundane or magical. Figure 1 visualizes the taxonomy that is described in more detail throughout the following.

### 2.1 Mobile and Stationary Travel Techniques

The distinction between mobile and stationary travel techniques is inspired by previous work, where travel techniques are categorized based on the presence or absence of physical movement (Bowman et al. 1999; Wendt 2010). This distinction is important because some commercially available VR systems do not include positional tracking. Specifically, this is the case in regard to VR systems powered by mobile phones (e.g., Samsung Gear VR, Google Daydream, and Google Cardboard), which only include tracking of the user’s head orientation. Such systems make it impossible to rely on mobile travel techniques. Instead, direction selection is often accomplished using gaze-directed steering, where the direction is derived from the head orientation, and movement is instigated through button clicks or by sustained fixation on a specific point in the environment. Gaze-directed steering may be awkward if vertical movement is possible (e.g., when flying), and the coupling of travel direction and gaze direction prevents the user from orienting himself or herself while moving (Bowman et al. 2004). The question of whether a travel technique demands physical movement or not is also important in relation to higher-fidelity systems like the HTC Vive and the Oculus Rift. Such systems include positional tracking, and the user is therefore able to change the virtual viewpoint through physical movement. However, these systems only afford

![Fig. 1. Nilsson, Serafin, and Nordahl’s (2016b) taxonomy of virtual travel techniques: The vertical axis subdivides the techniques based on whether the travel technique represents body-centric or vehicular travel. The horizontal subdivides the techniques based on whether the interaction metaphor is mundane or magical. The division of each cell represents the degree of user movement relative to the physical environment.](image-url)
movement within a limited physical space. This is problematic because the user’s virtual movement only should be constrained by the virtual architecture and topography. At best, leaving the tracked area may hamper the user’s sensation of being in the virtual environment. At worst, it may be dangerous because the user is oblivious to physical obstacles while wearing the head-mounted display (HMD).

2.2 **Mundane and Magical Travel Techniques**

The distinction between mundane and magical travel techniques has been adopted from Slater and Usoh (1994) and Whitton and Razzaque (2008). A travel technique is classified as mundane or magical based on whether movement in the virtual environment is limited by physical constraints, such as the laws of physics, biological evolution, or the current state of technological development. Thus, any technique that allows the user to travel in a manner that cannot be accomplished in the real world qualifies as magical, and techniques that mimic real travel are considered mundane. There are some important differences between magical and mundane travel techniques. First, magical travel techniques may allow users to travel great distances within the virtual environment without requiring any physical movement. Moreover, Bowman et al. (2012) describe that magic interaction techniques may be designed purposely to favor task performance and usability over the familiarity accompanying techniques that mimic real-world interactions. For example, if a user is supposed to traverse great distances within the virtual environment, teleportation will be much faster and less tiring compared to walking or piloting a virtual vehicle. With that being said, sometimes mundane travel techniques are easier to use because they are based on a familiar type of interaction, and the scenario itself may demand a technique that is possible in real life (e.g., during training scenarios or narrative experiences unfolding in a world that adheres to real-world constraints).

2.3 **Vehicular and Body-Centric Travel Techniques**

Finally, Nilsson et al. (2016b) distinguish between vehicular and body-centric travel techniques, that is, techniques that simulate travel by means of a virtual vehicle and techniques that simulate movement generated by using the body to exert forces to the environment (e.g., walking, running, or swimming). When traveling by means of a virtual vehicle, the user indirectly produces movement by manipulating the controls while remaining stationary relative to the vehicle. Thus, vehicular travel techniques do not require a large tracking space. Even when the user is physically stationary, movement of the virtual viewpoint can produce compelling self-motion illusions (e.g., Hettinger et al. (2002) and Warren and Wertheim (1990)). However, compelling illusions of self-motion may come at a cost since they are believed to elicit cybersickness (Davis et al. 2014) or VR sickness (Fernandes and Feiner 2016). Particularly, the dominant view holds that cybersickness results from a conflict between external sensory information and vestibular sensations (e.g., the user visually perceives motion but the vestibular system indicates that he or she is stationary) (Davis et al. 2014).

3 **NATURAL WALKING IN VIRTUAL ENVIRONMENTS**

Humans do, as suggested, routinely navigate their surroundings on foot, and they generally do so with relative ease and without assigning much explicit attention to the performed movements or the sensory stimuli produced as a result of these movements. However, the task of enabling users to walk through virtual worlds is anything but trivial. According to Nordahl et al. (2012b), this task can be broken down into at least two separate, yet interrelated, challenges: (1) creation of travel techniques that mimic the experience of real walking without requiring a physical space of the same size as the virtual environment and (2) provision of appropriate multisensory stimuli.
Fig. 2. Three general categories of walking techniques that provide users with relatively natural walking experiences when the virtual environment is smaller than the physical space: repositioning systems, proxy gestures, and redirected walking. It is possible to further distinguish between repositioning systems that are either active or passive, proxy gestures that are based on movement of the legs or upper body, and redirected walking that is based on application of gains or manipulation of the virtual architecture.

resulting from the user’s interaction with the virtual environment (e.g., the sight, sound, and touch accompanying each step). Throughout the following, we review work addressing both of these challenges.

4 WALKING TECHNIQUES

Within the academic community, numerous solutions to the problem of allowing users to naturally walk through large virtual environments have been proposed. Generally, these techniques fall into one of three categories: repositioning systems, proxy gestures, and redirected walking. Figure 2 visualizes the three categories and the potential subdivisions that will be described in the following.

4.1 Repositioning Systems

Repositioning systems essentially counteract the forward movement of the user and thereby ensure that he or she remains in a relatively fixed position. Thus, following the taxonomy presented in Section 2, repositioning systems offer stationary virtual travel. It is possible to distinguish between systems that reposition the user actively or passively.

Active repositioning often relies on elaborate mechanical setups in order to cancel the user’s forward movement. One of the simplest examples of an active repositioning system is the traditional, linear treadmill (Fease et al. 2011; Kassler et al. 2010; Powell et al. 2011). An inherent disadvantage of such treadmills is that the user can only walk forward, and if the application requires turning, this will have to be done in an indirect manner (e.g., based on the user’s head orientation or using a joystick) (Bowman et al. 2004). Notably, efforts have also been made to facilitate natural walking using omnidirectional treadmills that allow the user to freely walk in any direction (Darken et al. 1997; Iwata 1999; Noma 1998; Souman et al. 2011). A potential limitation of such techniques is that motion of the treadmill may cause the user to lose his or her balance during turns and sidesteps (Bowman et al. 2004). Other examples of repositioning systems include motorized floor tiles that move in the opposite direction of the walker’s direction (Iwata et al. 2005), cancellation of the walker’s steps through strings attached to his or her shoes (Iwata et al. 2007), and a human-sized hamster ball (Medina et al. 2008). Three examples of active repositioning systems can be seen in Figures 3(a) to 3(c).

Passive repositioning offers a simpler and less expensive alternative to active repositioning. Generally, passive repositioning systems rely on friction-free platforms that prevent the forces generated during each step from moving the user forward (Avila and Bailey 2014; Cakmak and Hager 2014; Huang 2003; Iwata and Fuji 1996; Swapp et al. 2010; Walther-Frank et al. 2013). The Virtuix
4.2 Proxy Gestures

Locomotion based on proxy gestures requires the user to perform gestures that serve as a proxy for actual steps. It is possible to distinguish between different subcategories depending on what part of the body is used to perform the proxy gesture. In this article, gestures will be classified as upper and lower body gestures for simplicity.

Because the aim is to produce a walking experience that resembles real walking, proxy gestures often rely on lower-body movement. The most common approach is so-called walking-in-place (WIP) techniques. When traveling through virtual worlds using such techniques, the user performs stepping-like movements on the spot. The steps in place may be registered based on a physical interface detecting discrete gait events (e.g., Bouguila et al. (2005, 2003) and Richard et al. (2007)) or using motion tracking systems enabling continuous detection of the position and velocity of limbs (e.g., Bruno et al. (2013), Feasel et al. (2008), Slater et al. (1993), and Wendt et al. (2010)). Even though self-reported measures have revealed that users find real walking more simple, straightforward, and natural (Usah et al. 1999), WIP techniques do come with a number of advantages: (1) WIP techniques are convenient and inexpensive (Feasel et al. 2008), (2) WIP techniques can provide some of the proprioceptive feedback inherent to real walking (Slater et al. 1994), (3) WIP techniques may provide an increased sensation of “being there” in the virtual environment compared to techniques where movement is produced by pressing a button (Slater et al. 1995), and (4) WIP techniques may be comparable to real walking in terms of users’ performance on simple spatial orientation tasks (i.e., tasks requiring users to locate objects and point to a specific location once the navigation is completed) (Williams et al. 2011). WIP techniques have been implemented using commercially available hardware such as Microsoft’s Kinect (Suma et al. 2011), Nintendo’s Wii Balance Board (Filho et al. 2012; Williams et al. 2011), and the inertial data obtained from the sensors of a mobile HMD (Pfeiffer et al. 2016; Tregillus and Folmer 2016). Most WIP techniques involve a gesture reminiscent of marching on the spot or walking up a flight of stairs (Figure 4(a)). However, recent work suggests that the experience of walking may be perceived as more natural if the user performs a gesture that better matches real walking in terms of perceived energy expenditure, such as alternately tapping each heel against the ground (Nilsson et al. 2013) (Figure 4(b)). Notably, tapping in place produces knee movements that have been described as gestural input for virtual locomotion elsewhere (Guy et al. 2015; Punpongsan et al. 2016; Templeman et al. 1999).
While steps in place appear to be the most common lower-body gesture, alternatives have been proposed. For example, Zielinski, McMahan, and Brady (2011) combined common WIP locomotion for forward movement with a leg-based pinch gesture for sidestepping (i.e., the user would step to the side with one foot and then slide this foot inward along the floor toward the other foot). Moreover, as part of their efforts to identify non-tiring and easily accessible gestures for virtual travel, Guy et al. (2015) and Punpongsanon et al. (2016) explored the use of a number of gestures including placing one foot in front of or behind the center of gravity, hip rotations around the body’s longitudinal axis, and sideways hip swings.

At first glance, upper-body gestures seem less suited if one wishes to facilitate a natural walking experience. Nevertheless, it has been proposed that relatively natural experiences can be achieved based on gestures devoid of explicit leg movement, such as swinging one’s arms back and forth (McCullough et al. 2015; Nilsson et al. 2013; Wilson et al. 2016) (Figure 4(c)). Moreover, the results of a study performed by Nilsson et al. (2013) suggest that users experience this gesture equally as natural as WIP locomotion and less fatiguing. The authors describe that a possible reason this gesture was perceived as relatively natural is that it involves a rhythmic swinging of the arms similar to the ones sometimes occurring during real walking (Zehr and Haridas 2003). An inherent disadvantage of this approach is that it leaves the user unable to interact with his or her arms while walking. Nevertheless, arm swinging may prove to be a meaningful solution for certain applications because systems such as the Oculus Rift currently only support tracking of the head and hands. Notably, an implementation of this locomotion technique, aptly dubbed ArmSwinger, is currently available to Unity developers (electricnightowl.com). Other forms of upper-body gestures include head swaying (Terziman et al. 2010), shoulder rotation (Guy et al. 2015), sideways and forward leaning of the torso (Guy et al. 2015; Kitson et al. 2017; Langbehn et al. 2015), and even finger-based gestures (i.e., finger tapping on a touchpad (Yan et al. 2016) and alternating clicks on the trigger buttons of two HTC Vive controllers (Sarupuri et al. 2017)).

### 4.3 Redirected Walking

Unlike repositioning systems and proxy gestures, redirected walking does involve physical walking and therefore qualifies as mobile travel techniques following the taxonomy outlined in Section 2. Generally speaking, redirected walking refers to a collection of approaches that make it possible to control the user’s path through the physical environment by manipulating the stimuli used to represent the virtual environment (Suma et al. 2012). It is possible to distinguish between techniques that accomplish redirection based on either perspective manipulation or environmental manipulation. The two types of techniques are discussed in turn below.

![Fig. 4. Three examples of proxy gestures: (a) the traditional WIP gesture, (b) arm-swinging, and (c) tapping in place. The purple arrows illustrate the movement of the body parts used to perform the gesture.](image-url)
Redirection techniques relying on perspective manipulation apply changes to the user’s virtual point of view (or point of audition). Particularly, the manipulation is accomplished by applying gains that affect the mapping between the user’s real virtual movement (e.g., if a gain of 2.0 is applied to the user’s forward translation, then he or she will travel twice as fast in the virtual environment). Common ways of redirecting users using this approach include application of imperceptible translation, curvature, rotation, and bending gains to the movement of the virtual camera. The gains scale and bend the walker’s path, as well as increase or decrease the virtual rotations resulting from physical rotations (Interrante et al. 2007; Langbehn et al. 2017; Razzaque et al. 2001) (Figure 5)). To exemplify, if the user is asked to walk across a virtual soccer field, it is possible to slowly and imperceptibly rotate the field around the user. This will cause him or her to walk in circles even though he or she thinks he or she is walking along a straight path. It is preferable for redirection techniques relying on perspective manipulation to be subtle because overt manipulation would disrupt the natural experience of walking through the virtual environment (Suma et al. 2012). The maximum and minimum gains that can be applied without the user noticing the manipulation (i.e., the perceptual detection thresholds) have been established through empirical evaluations relying on psychophysical methods (Grechkin et al. 2016; Steinicke et al. 2010). Even though subtlety is preferable, redirection techniques relying on overt perspective manipulation exist. Particularly, overt perspective manipulation may become necessary if the user is dangerously close to the boundary of the walking space. Under such circumstances, the system intervenes, the user is instructed to turn around, and during the turn the visual image may be frozen or a gain applied (Williams et al. 2007). An increasingly growing body of work has explored different approaches to decreasing the likelihood of the user detecting the manipulation by using visual distractors (Peck et al. 2011), by manipulating the user’s viewpoint during saccades and eye blinks (Bolte and Lappe 2015; Langbehn et al. 2016), or by using narrative events as opportunities to imperceptively manipulate the user’s path (Grechkin et al. 2015; Neth et al. 2012).

Work on perspective manipulation has primarily relied on manipulation of the visuals presented to the user. Nevertheless, it has been demonstrated that redirection can be accomplished using sound when the user is deprived of visual stimuli (Nogalski and Fohl 2016; Serafin et al. 2013). However, it remains to be seen if the addition of sound can decrease users’ ability to detect visual manipulations under certain circumstances (e.g., in a dimly lit or foggy environment) (Meyer...
Fig. 6. A visual overview of Waller and Hodgson’s (2013) three types of information received by walkers about their environment and their place and movement through it.

et al. 2016; Nilsson et al. 2016). Moreover, the addition of haptic cues in the form of a convex surface wall have been shown to positively influence the ability to redirect users (Matsumoto et al. 2016). Furthermore, several attempts have been made at producing steering algorithms that deploy different types of perspective manipulation to ensure that the user remains within the tracked space (Hodgson and Bachmann 2013; Nescher et al. 2014; Razzaque 2005). However, according to Azmandian, Grechkin, and Rosenberg (2017), the most commonly used algorithm is the Steer-to-Center algorithm that continuously tries to steer the user toward the center of the tracking area. Notably, this approach has been combined with overt interventions in order to ensure the safety of the user (Peck et al. 2011). While most research on subtle perspective manipulation has focused on single user scenarios, recent work has begun to explore the problem of simultaneously redirecting multiple users in a shared space (Azmandian et al. 2017; Bachmann et al. 2013; Holm 2012).

Redirection techniques relying on environmental manipulation do, as the name implies, accomplish the redirection by changing the properties of the virtual environment. In its simplest form, this type of redirection involves creation of virtual environments that match the physical space in size and in terms of potential obstacles (Simeone et al. 2015; Sra et al. 2016). Presenting virtual objects at the boundaries of the tracking space should be able to contain the movement of most walkers. Suma et al. (2011a) devised an approach redirecting users through subtle manipulation of the virtual architecture inspired by change blindness (i.e., the inability of an individual to detect changes in the environment (Matlin 2009)). Specifically, Suma et al. (2011a) were able to manipulate the orientation of doorways behind users’ backs and thereby influence their walking paths. Moreover, Suma et al. (2012) proposed a technique dubbed Impossible Spaces that makes it possible to compress virtual interior environments into comparatively smaller physical spaces by means of self-overlapping architecture (e.g., two adjacent virtual rooms may occupy the same physical space). Finally, if the aim is not to replicate the spatial layout of a real space, then the technique Flexible Spaces can provide unrestricted walking within a dynamically generated interior virtual environment (Vasylevska and Kaufmann 2017b; Vasylevska et al. 2013).

The fact that redirected walking involves actual walking means that a larger physical space is required. However, it also means that the user receives vestibular self-motion information, which may aid the walker’s understanding of the size of the environment and improves spatial understanding (Bowman et al. 2004). Notably, using a specific scenario, researchers have been able to successfully redirect walkers in areas as small as 6m × 6m (Suma et al. 2015) and 4m × 4m (Langbehn et al. 2017). While these results are encouraging, it should be stressed that redirection techniques relying on environmental manipulation generally are limited to interior environments (Suma et al. 2011a). Contrarily, perspective manipulation using gains can be applied in open environments, such as outdoor scenes. However, this form of redirection requires the user to subconsciously compensate for the introduced manipulation, which may impose additional cognitive load, as demonstrated using verbal and spatial working memory tasks (Bruder et al. 2015).
5 MULTIMODAL FEEDBACK DURING WALKING

Walking is an inherently multisensory activity, and several sources provide the walker with information about the surrounding environment as well as the act of walking itself. Waller and Hodgson (2013) present a discussion of the systems that provide individuals with sensory information about the environment and their movement through it. Inspired by this work, we distinguish between three categories of information: (1) \emph{External sensory information} includes information derived from the visual, auditory, and cutaneous senses. The cutaneous senses provide information about interactions at the level of the skin (Robles-De-La-Torre 2006). (2) \emph{Internal sensory information} is produced by the vestibular and the kinesthetic/proprionceptive system. The vestibular system is located in the inner ear and registers angular and linear acceleration of the head. The kinesthetic/proprionceptive system is responsible for detecting positions, orientations, and movements of the musculature and joints (Waller and Hodgson 2013). (3) Finally, \emph{efferent information} relates to attention allocation and efference copy (the neural representation of motor commands from the central nervous system to the musculature) (Waller and Hodgson 2013). Besides vision, hearing, and the cutaneous senses, other external sources can conceivably also provide the walker with spatial information (e.g., the olfactory or gustatory senses). However, during everyday interactions, their contributions are likely to be negligible (Waller and Hodgson 2013). In what follows, we summarize work aimed at providing users with appropriate multisensory stimuli during virtual walking. First we present work related to external sources of sensory information (i.e., visual, auditory, and haptic feedback), and then the role of internal and efferent information is discussed.

5.1 Visual Feedback

Vision serves as a direct, rich, and precise source of spatial information (Waller and Hodgson 2013), and it is central to how walkers perceive their movement through an environment. Specifically, \emph{optic flow} (i.e., “the pattern of visual motion at the moving eye” (Warren et al. 2001, p. 213)) provides the walker with information about translational and rotational movement. That is, expanding and contracting flow fields are indicative of forward and backward movement, respectively. Laminar flow patterns may indicate either rotational or sideways movement. The interpretation of ambiguous external sensory information, such as laminar flow patterns, may vary depending on the nature of the simultaneous internal and efferent information (see Section 5.4).

Optic flow can become disambiguated by information arriving from other sensory modalities and efferent sources. Nevertheless, research on sensory psychology suggests that in the event of a sensory conflict, vision tends to dominate proprioceptive and vestibular sensations (Dichgans and Brandt 1978). It is arguably this visual dominance that makes it possible to subtly redirect walkers by applying translation, curvature, rotation, and bending gains.

Interestingly, self-motion perception during treadmill and WIP walking is prone to distortion. If a user is walking on a treadmill while viewing a virtual environment using an HMD, then one would expect users to find a match between the visually presented speed and the speed of the treadmill belt to be the most realistic. However, contrary to intuition, it has been demonstrated that individuals tend to underestimate visually presented walking speeds when using linear treadmills for virtual locomotion. In other words, walkers are likely to find visually presented speeds too slow if they correspond to the speed of the treadmill belt (Banton et al. 2005; Kassler et al. 2010; Nilsson et al. 2016a; Powell et al. 2011). All the factors influencing this perceptual distortion remain unknown, but research on self-motion during treadmill walking and WIP locomotion have yielded the following findings: (1) if walkers direct their gaze downward or to the side, the underestimation is eliminated (Banton et al. 2005); (2) the underestimation does not appear to be caused by image jitter (Banton et al. 2005); (3) no effect of increased HMD weight or varying peripheral
occlusion has been found (Nilsson et al. 2015a, 2015b); (4) the amount of underestimation appears to be inversely proportional to the size of the display field of view (Nilsson et al. 2014a); (5) similarly, the degree of underestimation seems to be inversely proportional to the size of the geometric field of view (Nilsson et al. 2015b); (6) the amount of identified underestimation may vary depending on study methods (Nilsson et al. 2015b); (7) high step frequencies may lead to a larger degree of underestimation, but the evidence is somewhat equivocal with respect to this effect (Durgin et al. 2007; Kassler et al. 2010; Nilsson et al. 2014b); (8) finally, the degree of underestimation may vary slightly depending on whether the user is walking on a treadmill or walking in place (Nilsson et al. 2016a).

5.2 Auditory Feedback

Even though vision tends to dominate spatial perception, audition does provide stationary and moving observers with information about the surrounding environment. Particularly, Waller and Hodgson (2013) describe that audition can provide the individual with information about the size of the environment and the position of objects and events within that environment—assuming that they emit sounds, that is. As evident from the literature on self-motion illusions (i.e., vection), sound sources moving relative to the listener can also influence motion perception. According to Väljamäe (2009), three cues are central for discrimination of auditory motion: binaural cues (e.g., interaural time and level differences), the Doppler effect (i.e., frequency shifts during relative movement between a sound source and listener), and sound intensity (e.g., intensity changes providing an estimate of time to arrival). Larsson et al. (2010) describe that approaches to sound spatialization fall into two general categories: sound-field-related methods (multichannel loudspeaker systems) and head-related methods (spatial rendering of sound delivered via headphones). Rendering and perception of distant auditory objects have been studied extensively. However, comparatively little work has explored how to make the sensation of walking more natural by providing auditory cues related to the interaction between users’ feet and the virtual ground.

Footstep sounds are frequently used in movies to provide information about the presence of characters on and off screen, and in case of computer games, such sounds may also be used to inform the player about the surface being traversed and to produce a sense of weight and embodiment. In these cases, footstep sounds are often based on recordings retrieved from sound libraries or produced by Foley artists (Nordahl et al. 2011). An alternative to recordings is physics-based sound synthesis algorithms. Among the pioneers of this approach is Cook (1997), who proposed a number of physically informed stochastic models (PhiSMs) simulating everyday sonic events. Cook’s (2002) work includes algorithms for simulating the sound generated when walking on varying surfaces. Similarly, Fontana and Bresin (2003) devised physically informed models with the ability to reproduce the sound of footsteps on several stochastic surfaces.

Visell et al. (2009) distinguish between two broader categories of interfaces for controlling the feedback generated from footsteps: instrumented floors (rigid surfaces augmented with sensors and actuators) and instrumented shoes (footwear augmented with sensors and actuators). An advantage of instrumented floors is that they do not require users to wear additional equipment. However, unlike wearable solutions, they are likely to limit the user’s movement to a relatively small walking area, and current solutions are relatively impractical and expensive. Nordahl et al. (2011) developed a real-time, physics-based sound synthesis engine and integrated it in a VR setup using an instrumented floor. Condenser boundary microphones embedded in the floor detected the user’s footsteps, which were used to drive the sound synthesis engine. This system was able to synthesize the sound of walking on different solid and aggregate surfaces, including wood, snow, and grass. The evaluation of the system generally yielded promising results with respect to surface recognition, and when the participants made incorrect judgments, they frequently mistook aggregate and solid surfaces for other surfaces belonging to the same general category. Moreover, an additional
study revealed that participants found it easier to recognize simulated surface materials when these materials were presented alongside semantically congruent environmental sounds (Nordahl et al. 2011). Notably, it has been proposed that ground reaction forces can be derived from microphones placed both on the floor and on shoes (Serafin et al. 2009). Earlier work by Nordahl (2005, 2006), relied on instrumented shoes for detecting foot-ground interactions and demonstrated that simulated footstep sounds can significantly increase participants’ sensation of presence in the virtual environment and the amount of movement performed by the user. A similar system for producing footstep sounds was developed by Papetti et al. (2009).

5.3 Haptic Feedback

The third source of external sensory information is the cutaneous senses. Particularly, the somatosensory pressure receptors inform the walker about physical contact with objects on the path (Waller and Hodgson 2013). Haptic feedback is intended to provide such information by reproducing forces, movements, and other cutaneous sensations felt via the sense of touch (Marchal et al. 2013). In a manner similar to Lindeman et al. (1999), we distinguish between two approaches for supplying individuals with cutaneous information during virtual walking. Passive-haptic feedback is generated simply by virtue of the objects’ physical properties (e.g., its shape and texture). Contrarily, active-haptic feedback is controlled using a computer and delivered using a haptic display (e.g., vibrotactile actuators).

Meehan et al. (2002) describe a study combining passive-haptic feedback with a stressful virtual environment. The stressful environment comprised a virtual pit and the participants were required to look over a virtual precipice. The passive-haptic feedback comprised a 1.5” wooden ledge collocated with the virtual precipice. Among other things, the study revealed that the participants exhibited stronger fear responses, as assessed by means of physiological measures, when exposed to the passive-haptic feedback ledge compared to when they were standing on the floor. More recently, Suma et al. (2011b, 2013) developed a virtual environment wherein users are redirected over real-world concrete and gravel using a redirection technique inspired by change blindness (see Section 4.3). Particularly, the redirection ensured that the users would physically step on the correct surface material whenever this material was present in the virtual environment.

With respect to active-haptic feedback, Marchal et al. (2013) stated that the addition of even low-fidelity tactile feedback may increase the sensation of presence in an audiovisual virtual environment. Along similar lines, Srinivasan and Basdogan (1997) suggested that the potential gains of adding such tactile feedback may be larger than improving the quality of the feedback delivered to an existing modality (e.g., the visual display). Passive-haptic feedback will necessarily be floor based. However, like auditory feedback, active-haptic feedback can be delivered based on interaction with either instrumented floors or shoes. In fact, much recent research on auditory and haptic feedback for virtual walking has been performed in parallel using multimodal interfaces. Marchal et al. (2013) describe that because the haptic and auditory stimuli share the same origin (i.e., physical contact between the foot and the ground), the high-frequency information in the mechanical signals of the two are often closely related. For that reason, the same physics-based algorithms have been used to produce the signals fed to both auditory and haptic displays (Nordahl et al. 2012b). Much of the recent work on audiohaptic feedback for virtual walking was performed as part of the Natural Interactive Walking (NIW), FET–Open EU project (FP7–ICT–222107), which explored the use of both instrumented floors and shoes.

Law et al. (2008, 2009) produced a haptic display based on an instrumented floor. Particularly, they developed a system that is able to provide users with active-haptic feedback through actuated floor tiles. The system included 36 square tiles (30.5cm × 30.5cm) arranged in a 6 × 6 matrix...
in the center of a CAVE-like environment, including floor projections of the virtual ground surface. Recent work by Kruijff et al. (2016) demonstrated that the addition of, among other things, vibrotactile feedback could improve stationary users’ sensation of self-motion.

Much of the work using instrumented shoes for delivering active-haptic feedback has focused on how accurately participants could recognize virtual surfaces rendered using the augmented footwear. Work by Nordahl et al. (2010) suggests that auditory feedback may yield superior recognition performance compared to haptic feedback, and the combination of auditory and haptic feedback did not produce a significant improvement. Serafin et al. (2010) similarly found that auditory feedback was superior to haptic feedback with respect to recognition. However, they did find that audiohaptic feedback improved recognition in some cases. Nordahl et al. (2012a) performed a study exploring whether the addition of audiohaptic simulation of foot-ground interaction influences perceived realism and presence. They used a stressful virtual environment inspired by the pit room used by Meehan et al. (2002). The results did not reveal any significant differences in terms of presence. However, the participants did find that the addition of audiohaptic feedback made the experience seem more realistic (Nordahl et al. 2012a). For a comprehensive overview of audiohaptic feedback for walking and the outcomes of the NIW project, refer to Fontana and Visell (2012).

5.4 Internal and Efferent Information

As suggested, it is possible to distinguish between at least three types of internal sensory information: vestibular, proprioceptive, and kinesthetic information. Aside from contributing to the sensation of self-motion (Riecke et al. 2005), information originating from the vestibular system supports various oculomotor and postural reflexes, and it is thought to play a central role for spatial updating and dead reckoning (Waller and Hodgson 2013). While the terms “kinesthetic” and “proprioceptive” often are used interchangeably, Waller and Hodgson (2013) describe that it is possible to distinguish between kinesthetic and proprioceptive information. Particularly, kinesthetic information relates to “information about the movement of one’s limbs or effectors” [p. 8], whereas proprioceptive information relates to the “relatively static position or attitude of the musculature” [p. 8]. In regard to the act of walking, this implies that kinesthetic information enables the walker to take steps without visually confirming that the action is being performed as intended, and the proprioceptive information makes an individual aware of the position of the lower limbs even in the absence of motion. Especially proprioception is believed to positively influence performance on heading, turn, and distance estimation, as well as spatial updating. The final category of information discussed by Waller and Hodgson (2013) is efferent information. Particularly, efference copy is pertinent to the current discussion. Efference copy refers to the neural representation of motor commands from the central nervous system to the musculature (Waller and Hodgson 2013). This simultaneous record of current motor commands is used to predict sensory stimuli and modulate the response of the associated sensory modality (Pynn and DeSouza 2013). Moreover, it is believed that the information provided by efference copy enables predictions about the consequences of performed actions before they have occurred (Harris et al. 2002). Thus, efference copies, among other things, enable individuals to discern stimuli generated by external events in the environments from similar stimuli produced by their own actions (Waller and Hodgson 2013). For example, it is efferent information about one’s own motor commands that makes it possible to disambiguate laminar optic flow produced during head turns from the similar pattern on the retina resulting from circular environmental movement.

Internal and efferent information is particularly relevant when considering how the different walking techniques outlined in Section 4 are experienced. With respect to vestibular stimulation, subtle redirection techniques are generally superior to repositioning systems and WIP techniques.
Because repositioning systems and WIP techniques are designed to limit translational movement, users are guaranteed conflicting visual and vestibular information. Contrarily, redirected walking involves vestibular stimulation, and as long as the applied gains remain low enough, the conflict between vision and the vestibular sense will remain unnoticed. Notably, it has been suggested that forward leaning, which produces vestibular stimulation, may elicit stronger illusions of self-motion on behalf of stationary users (Kruijff et al. 2015). Thus, WIP locomotion involving leaning, such as the technique proposed by Langbehn et al. (2015), could help compensate for the missing vestibular information.

Because subtle redirection involves actual steps, approaches belonging to this category should provide more natural kinesthetic and proprioceptive information than WIP techniques. In this regard, some repositioning systems are also likely to exceed WIP techniques. Moreover, alternative gestural input, such as tapping in place and arm swinging, may be perceived to be at least as natural as the traditional WIP gesture (see Section 4.2). A possible explanation is that tapping in place, like steps in place, results in kinesthetic information reminiscent of that generated during real walking. However, tapping in place better matches real walking in terms of perceived effort (Nilsson et al. 2013). Similarly, arm swinging provides a better match with respect to perceived effort, and relevant kinesthetic information is also generated from rhythmic swinging of the arms, which is known to occur during walking (Zehr and Haridas 2003).

6 CURRENT CHALLENGES AND FUTURE WORK

Recent advances in display and tracking technology have made VR accessible to consumers in an unprecedented manner. The popularization of VR will (hopefully) mean that an increasingly large number of people will be navigating through familiar, foreign, and fantastic virtual environments on a regular basis. Indeed, virtual travel is one of the most common and universal tasks performed inside virtual environments. As pointed out in Section 2, numerous different virtual travel techniques have been proposed. We described that it is possible to distinguish between travel techniques based on whether the user is physically moving or not, whether the techniques involve a virtual vehicle or not, and whether they mimic real-world travel or employ a magical interaction metaphor (Section 2). Techniques involving physical movement are particularly problematic because the user’s ability to travel virtually will be constrained by the size of the tracked space. Both vehicular and magical travel techniques largely circumvent this issue. However, many scenarios will demand that the user is able to navigate large virtual environments on foot.

The research community has yet to produce a system that facilitates unconstrained and natural walking within large virtual environments. Specifically, we argued that this task involves two challenges: (1) creation of travel techniques that sufficiently mimic the experience of real walking without requiring a physical space of the same size as the virtual environment and (2) provision of appropriate multisensory stimuli resulting from the user’s interaction with the virtual environment. As discussed in Section 4, repositioning systems, proxy gestures, and redirected walking all offer potential solutions to the first of the two challenges.

An advantage of active repositioning systems is that they enable users to take actual steps, thus ensuring correct proprioceptive/kinesthetic feedback. Moreover, such systems can successfully confine the user’s movement to an area of limited size. However, most current implementations are relatively cumbersome and expensive, and a limitation of large mechanical setups, such as omnidirectional treadmills, is that they may cause users to lose balance during turns and sidesteps (Bowman et al. 2004). Passive repositioning systems relying on friction-free platforms offer an inexpensive alternative. However, the community has yet to empirically establish how well these systems perform with respect to factors such as perceived naturalness, spatial performance, task performance, and simulator sickness. This is an area ripe for future work.
There is obviously a limit to how well proxy gestures can mimic the experience of real walking. Nevertheless, a considerable advantage of such approaches is that they are relatively inexpensive and require very little physical space. Historically, research has focused on lower-body gestures and particularly WIP techniques. Moreover, much of this work has focused on optimization of algorithms for step detection and velocity estimation and exploration of different hardware for detecting the user’s movements. Future work should continue to improve techniques with respect to the virtual locomotion speed control goals introduced by Feasel et al. (2008): smooth between-step locomotion speed, continuous within-step speed control, real-world turning and maneuvering, and low starting and stopping latency. However, previous work has almost exclusively focused on gestural input for forward movement. Future work should establish what upper- and lower-body gestures provide the most natural experience of walking forward, backward, and laterally. Moreover, because many systems do not offer full-body tracking, it is necessary to determine what steering methods users will find the most natural (e.g., gaze-directed steering, torso-directed steering, or hip-directed steering) (Nilsson et al. 2016b). Finally, it remains to be seen whether a sense of ownership of a virtual body exhibiting normal gait behavior can be sustained during locomotion based on proxy gestures.

Because repositioning systems and proxy gestures are designed to limit translational movement, these approaches involve limited vestibular self-motion information. Redirected walking is arguably the most natural of the three general approaches because the user is physically moving. However, this generally means that redirected walking will require a much larger physical space than repositioning systems and proxy gestures. It is well documented that users can be subtly redirected through perspective manipulation (i.e., application of gains) and environmental manipulation (i.e., self-overlapping virtual architecture). However, previous work establishing detection thresholds for perspective manipulation relied on displays that do not compare to current-generation HMDs. Moreover, threshold estimates vary significantly depending on estimation methods, and users’ sensitivity to the redirection is likely to vary depending on individual differences and the attentional demands of the virtual task. Thus, estimation of detection thresholds for visual (and acoustic) gains remains an important area of research. Environmental manipulation offers a relatively safe approach to redirected walking, insofar as the interior environment does not exceed the bounds of the tracking space. However, the same cannot be said of perspective manipulation using gains. Thus, it is important for future work to explore nonintrusive ways of intervening when the user walks too close to bounds of the tracked space. Creating opportunities to imperceptibly manipulate the user’s path using narrative events is a promising direction for future work. A major challenge facing work on redirected walking is generalizability. Ideally, the same redirection algorithm should be applicable across different virtual environments and scenarios. In this regard, automatic calculation of navigable paths and decision points from a given virtual environment and prediction of users’ future paths remain major challenges. Finally, we have yet to learn how exactly subtle redirection influences factors such as simulator sickness and cognitive load.

With respect to the second challenge of providing appropriate multisensory stimuli, three issues were emphasized in Section 5: facilitation of natural motion perception, delivery of natural feedback representing foot-ground interaction, and ensuring correct internal and efferent information. As discussed in Section 5.1, users tend to underestimate visually presented walking speeds during treadmill and WIP locomotion. However, we still do not know exactly what causes this perceptual distortion, or if it is equally prevalent when using current-generation HMDs. As a consequence, it may be necessary to establish HMD-specific guidelines describing what gains to apply in order to produce perceptually natural motion perception. It is worth stressing that underestimations of visual walking speeds generally have been observed when the user is exposed to relatively
artificial scenarios (e.g., walking in a straight line down long corridors while the gaze is fixated on the end of the corridor). Thus, it seems likely that the issue will be far less prevalent, and possibly nonexistent, in relation to more complex scenarios.

Sections 5.2 and 5.3 discussed the use of instrumented floors and footwear for delivering auditory and haptic feedback accompanying foot-ground interactions. Instrumented floors are a promising approach to delivering audio-haptic feedback to users relying on WIP locomotion. However, this approach is not practical in relation to redirected walking, which requires a large walkable area. Instead, it seems possible to combine redirected walking with instrumented footwear. In addition to providing more natural feedback to the user, it also seems possible that instrumented shoes could be used to decrease users’ ability to detect perspective manipulations. For example, it has been demonstrated that users walking along a curved path are more likely to believe that they are walking straight when touching a wall that is physically curved but appears straight in the virtual environment (Matsumoto et al. 2016). Audio-haptic feedback delivered at the feet could similarly be used to support visually presented perspective manipulations.

Finally, the internal and efferent information accompanying repositioning systems, proxy gestures, and redirected walking was discussed in Section 5.4. As argued above, the three approaches to virtual walking vary greatly in this regard. Repositioning systems provide relatively accurate proprioceptive/kinesthetic feedback. However, the fact that the user remains stationary comes at a cost since vestibular information is limited. Approaches relying on proxy gestures are for the same reason limited with respect to vestibular feedback. However, these approaches are considerably less expensive and can be implemented using off-the-shelf hardware. With respect to internal and efferent information, redirected walking arguably mimics real walking the best. Nevertheless, this approach requires a considerable amount of space and the community has yet to produce a redirection algorithm that is truly generalizable.

We have yet to see a commercially viable solution that is able to replicate the experience of real walking within a limited physical space while providing high-fidelity multisensory feedback. Nevertheless, great strides have been made since Sutherland (1965) presented his ultimate display that would allow users to step into (and walk around in) a digital wonderland. This article detailed an overview of this work and highlighted some of the challenges that are likely to inform future work on virtual walking.

REFERENCES


Received April 2017; revised December 2017; accepted December 2017