Using Head Tracking Data for Robust Short Term Path Prediction of Human Locomotion

Thomas Nescher and Andreas Kunz

Innovation Center Virtual Reality (ICVR)
Institute of Machine Tools and Manufacturing
ETH Zurich, Switzerland
{nescher,kunz}@iwf.mavt.ethz.ch

Abstract. Modern interactive environments like virtual reality simulators or augmented reality systems often require reliable information about a user's future intention in order to increase their immersion and usefulness. For many of such systems, where human locomotion is an essential way of interaction, knowing a user's future walking direction provides relevant information.

This paper explains how head tracking data can be used to retrieve a person's intended direction of walking. The goal is to provide a reliable and stable path prediction of human locomotion that holds for a few seconds. Using 6 degrees of freedom head tracking data, the head orientation and the head's movement direction can be derived. Within a user study it is shown that such raw tracking data provides poor prediction results mainly due to noise from gait oscillations. Hence, smoothing filters have to be applied to the data to increase the reliability and robustness of a predictor.

Results of the user study show that double exponential smoothing of a person's walking direction data in combination with an initialization using the head orientation provides a reliable short term path predictor with high robustness.

Keywords: prediction, human path prediction, head tracking, walking direction, facing direction, virtual reality, augmented reality, exponential smoothing.

1 Introduction

Many research fields require a forecast about future actions, which can be achieved by prediction. For this, mathematical models exist for a projection into the future, which are fed by observations of past and current actions. Also for the fields of virtual and augmented reality, it is beneficial to apply such prediction for instance to human locomotion. This would allow calculating the reaction of the computer generated environment not only based on the current position and orientation of the user, but also his intended walking trajectory - predicted from past observations - could be taken into account. By this, more

M.L. Gavrilova et al. (Eds.): Trans. on Comput. Sci. XVIII, LNCS 7848, pp. 172–191, 2013.

time exists to prepare the system's adequate reaction on the most probable next action of the user.

Systems that rely on human locomotion could be significantly improved by prediction, not only regarding responsiveness, but also regarding the quality of the displayed environment. As an example, the system could output situation-specific information (e.g. a warning) before the user is in this situation. Moreover, location-based information systems could be improved by displaying prediction-based information. Thus, the system could e.g. inform a user that the subway service is not running prior he goes down the stairs. Another example could be a virtual environment, which could pre-calculate objects that the user will most probably experience in the next moment.

For many years, navigating in virtual environment was only possible using a joystick or a spacemouse, which did not address the most intuitive way of human locomotion. Thus, more recent research projects [1] also address this human perception entity to increase immersion and to provide more natural information e.g. about objects' sizes and distances to each other, which also facilitates a better and more intuitive orientation. However, today's step-in-place solutions and treadmills cannot completely fulfill this requirement, and thus new systems should go far beyond this current experience. New systems should be even more immersive, non-obtrusive and should not hinder the user to perform his task in a virtual environment. In order to realize such systems, new tracking systems will be required that can measure the user's position more reliably and more precisely. Moreover, the measured position and orientation values will be used as observation to predict the user's future actions.

In order to predict a user's future plans of action when walking in a real or virtual environment, a predictor needs to be fed by tracking data. For detecting angular rate and acceleration, sensors with a very small form factor exist (e.g. [2, 3]), which can be easily attached to the user, e.g. in the head-up display or attached to AR-glasses. If the sensor signals are combined by sensor fusion, it is possible to realize a continuous position and orientation tracking [4].

In order to realize real walking in a virtual environment, it is crucial to attach the sensors to the visual output device, e.g. to the head-mounted display (HMD), since this is the only position to track a user's head orientation (i.e. gaze direction).

If real walking in a virtual environment should be realized, the problem occurs that usually the virtual space is larger than the physical space that is equipped with a tracking system. Thus, it is crucial to realize a so-called redirected walking, which compresses an ideally unlimited virtual space into a physically limited real space. Such a compression can be achieved by guiding the user on a curved or scaled path [5–7]. Such a compression takes benefit from the human rules of cognition, which define that the visual perception overrides the haptic sensation of walking. Thus, the user will follow a visual goal, while not taking into account the placement and orientation of his feet anymore, resulting in the fact that he does not note any compression. However, the compression and redirection needs to be carefully adapted to the available physical room, which requires planning ahead any compression procedure. For doing so, a path prediction is inevitable.

The tracking data of a person, i.e. his position and orientation, are the prerequisite for realizing a feasible prediction about the user's future path. This is in particular challenging since the user could make decisions in the virtual environment regarding the path he wants to go. While many publications (e.g. [8]) predict pedestrians' movements from an exocentric point of view, this paper approaches the problem from egocentric perspective.

There are three different time intervals for which a path prediction could be done. The shortest prediction interval for human locomotion is in the range of milliseconds and is physically limited by the human's abilities of movement like e.g. maximum acceleration. Short term prediction as the next time interval is in the range of some seconds and is determined by human way finding abilities and the decided direction in which a human wants to walk. This short term prediction is not influenced by any environmental constraints like obstacles, but just depends on the intended walk direction of the user. The third time interval is the long term path prediction and ranges from seconds to minutes. Obviously, the long term path prediction has to take into account the environment, while it is also influenced by a person's cognitive map of the environment and the planned destination he wants to reach.

The focus of this paper is on the short term prediction of human locomotion and how navigation or direction decisions are related to head tracking data that can be acquired by an inside-out tracking system. It shows a user study, in which participants have to walk in a maze-like environment and make decisions for different walking directions. The walking trajectories are recorded and analyzed, and show the challenges when trying to make a robust short term walking prediction. For this prediction, different approaches are introduced and discussed, while the remainder of this paper gives an evaluation of these different approaches.

2 Background

If locomotion interfaces such as treadmills should be replaced by real walking, prediction becomes especially important. Such a real locomotion interface was e.g. introduced by Peck et al., by which users could experience a virtual environment that is larger as the available and tracked physical space [9]. The system benefits from the psychological effect that vision dominates the proprioceptive sensation of walking. Thus, it was possible to slowly rotate the virtual environment around the user without noticing this [10]. This results in an imperceptible redirection of user, which keeps him the physically limited real space. However, in order to plan this redirection properly, it is crucial to have knowledge about the user's future path.

A very similar locomotion interface was proposed by Nitzsche et al., but with two different path prediction approaches [6]. One approach does not base on any targets and derives the user's future walking direction simply from the head orientation, assuming that the user always walks in the direction he looks. However, if specific targets can be identified in a virtual environment, the approach from the above can be further detailed. Every target or potential goal in the virtual environment is provided with a so-called weight coefficient. If an object or a target is in the user's field of view, the weight factor will be increased as long as the user looks at it, otherwise the value will be decreased. Now, the predicted walking direction of the user is simply towards the target with the highest weighting coefficient.

Another approach for the user's future path was introduced by [11]. Here, the future walking direction is simply determined by a linear extrapolation of the user's previous path.

Another hybrid approach for predicting the user's future path was proposed by Interrante et al. [12]. They used the past n seconds of a person's locomotion and calculate the average direction in this time interval for a prediction. If person is not moving, the facing direction is used instead. The influence of the facing direction on the predicted direction is decreased when a user starts walking, while on the other hand the influence of the averaged walking direction becomes more significant for the prediction.

Walking and facing direction were also used in a hybrid approach by Steinicke et al. [7]. Here, the prediction bases on the walking direction, while the facing direction is used for verification. Steinicke et al. consider that walking and facing direction are not consistent all the time. The walking direction is only used as prediction, if the angle between walking and facing direction is smaller than 45 degrees. In all other cases, no reliable prediction is possible.

Although many different approaches exist for a short-term path prediction in virtual environments, so far no evaluation and comparison was done of the proposed methods.

Research results from psychology or neuroscience provide insights how gaze behavior is related to the walking direction. Grasso et al. studied how individuals make eye and head movements to align with their future walking trajectory during walking around corners [13]. Furthermore, Hollands et al. analyzed how human gaze behavior is associated with maintaining and changing the direction of locomotion [14].

Knowing a person's intended action by using motion prediction is of particular interest in the research field of human-robot interaction. In order to derive a human's intentional actions from the relative movements between human and robot, Koo et al. used machine learning algorithms [15]. However, the prediction is regarded from an exocentric perspective, as interactions between the human and the robot are analyzed by the robot's artificial intelligence software.

Urban planning and transportation is another research field, in which the movement patterns of pedestrians are analyzed, see e.g. [16]. These complex models also take into account the interactions among pedestrians e.g. for collision avoidance, and are typically used for large scale simulations of pedestrian flow. It is obvious that these models are more suited for a long-term path prediction.

For tracking users, various systems were proposed, as for example shoe-mounted inertial sensors [17] or ultrasound-aided systems [18]. The focus of these systems is on providing a stable and accurate determination of the user's position. For the estimation of a user's position, typically a Kalman Filter (KF) is used, which can also be seen as some kind of predictive tracking. Such a KF was e.g. used by Kiruluta et al. [19], who proposed a method for predictive head movement tracking. Double exponential smoothing was used by LaViola [20], while Rhijn [21] gives a comparison of different filtering and prediction methods.

Predictive tracking algorithms like those mentioned in the above typically provide prediction times of several milliseconds up to one second and are thus often used to compensate for system latencies. These predictive tracking algorithms typically use a movement and noise model that fits such very short prediction times, e.g. 100 ms in [20].

Unlike the approaches in the above, the goal of this paper is not to make a robust estimation of the head position, but reliably predict the *intended* walking direction. The goal is to realize a prediction for several seconds rather than the very short-term prediction in the range of milliseconds.

3 Formal Derivation of Short Term Path Prediction from Tracking Data

A portable tracking system that is carried by a user typically provides 6 DOF data, i.e. position and orientation, at discrete time steps. Given a known and constant update rate, an approximation of the derivative from the discrete time samples gives information about the current speed and acceleration.

In our test setup, we attached the camera of the inside-out tracking system to the user's head and applied a static coordinate transformation to achieve the center of the user's head. Having this, the facing direction can be determined as the axis from this head center to the nose tip. For predicting a user's future direction of movement, tracking data can be interpreted and extrapolated in different ways as presented in the following.

For normal walking in a virtual environment, the problem can be reduced to 3 DOF in the walking plane: 2 dimensions for position, while the orientation is a rotation around the normal of the walking plane. The position data at time t is referred to as $\boldsymbol{x}_t$ where $\boldsymbol{x}$ denotes the 2-dimensional position vector and t is the discrete time index starting at time 0.

3.1 Facing Direction

The facing direction can be expressed as a 2-dimensional direction vector (in the reference frame of the user's head), which can be calculated from the orientation data. The normalized facing vector is denoted as $\boldsymbol{f}$.

It is assumed that a user's facing direction indicates his intended direction of movement and thus can be used for prediction. This is in particular relevant if the user stands still, i.e. the recorded position data is constant. In this case, the orientation data from the head tracker is the only relevant data that can be used for prediction.

If the virtual environment has known targets, $\boldsymbol{f}$ can be used for prediction. For doing so, this target at position $\boldsymbol{p}$ is chosen that has the smallest angular deviation from the current facing direction $\boldsymbol{f}_t$. Using the scalar product, the angle between a target and the facing direction can be calculated as:

$$\theta_t = arccos\left(\frac{\boldsymbol{f}_t \cdot (\boldsymbol{p} - \boldsymbol{x}_t)}{|\boldsymbol{f}_t||(\boldsymbol{p} - \boldsymbol{x}_t)|}\right) \tag{1}$$

However, the facing direction of a user does not necessarily correspond to his gaze direction, which could result in major problems when using the facing direction for the prediction. Humans often only move their eyes instead of their complete head, in particular if two targets are close to each other. This problem could only be solved, if an additional eye tracker would also be employed for the tracking system.

3.2 Walking Direction and Speed

A person's current and past movement can be derived from the change of the position $\boldsymbol{x}$ over time. This results in a displacement vector $\boldsymbol{w}_t$ that is defined as follows:

$$\boldsymbol{w}_t = \boldsymbol{x}_t - \boldsymbol{x}_{t-1} \tag{2}$$

The displacement vector gives the direction of movement between two time steps t-1 to t. If the data is acquired with a constant sampling interval τ, the user's current speed is given by $|\boldsymbol{w}_t|/\tau$.

Following the above definition of short term path prediction, the goal is to only predict the user's path for a short time horizon, and thus we assume to have no further information about the virtual environment. Hence, we can further assume that due to energy minimization the user will approach the target on a straight line. A user's future path can therefore be intuitively predicted from the current walking direction. This means in other words that we will use linear extrapolation of the current walking direction to predict a user's future path.

Like for the facing direction above, $\boldsymbol{w}$ can be used to determine the chosen target in a virtual environment using the angular deviation.

If a person is standing still, $\boldsymbol{w}$ equals zero and no prediction can be made. This means that the displacement vector could give significantly wrong results for the predictor in case of slow movements. This is in particular the case if $|\boldsymbol{w}|$ is of the same magnitude or even smaller than tracking system's noise. Thus, the lower limit for the displacement vector must be carefully chosen depending on the characteristics of the tracking system (update rate, noise, etc.).

3.3 Smoothing and Robustness

Since tracking data is typically noisy, some data smoothing is required to reduce the effect of noise on the path prediction. Too much noise will make the path prediction unstable and thus worthless. A further disturbing effect for prediction is that the movement of the body during walking is not necessarily aligned with the intended walking direction. A tracking system that is mounted on the user's head, will also measure movements to the side as well as up and down due to the mechanisms of the human gait [22, 23]. Thus, the effects of noise and gait oscillations have to be filtered out in order to achieve a reliable prediction of the intended walking direction.

In the following section, various approaches for data smoothing are presented. These smoothers could be applied to the $\boldsymbol{f}_t$ or the $\boldsymbol{w}_t$ vectors. However, it was shown in Sect. 5 that smoothing is in particular important for predicting the walking direction. Hence, the equations are presented for the $\boldsymbol{w}_t$ vectors. $\boldsymbol{s}_t$ denotes the smoothed path prediction at time t.

Unweighted Moving Average. The moving average method is one of the simplest methods to smooth data. It is given by:

$$\boldsymbol{s}_t = \frac{1}{k} \sum_{i=0}^{k-1} \boldsymbol{w}_{t-i} \tag{3}$$

Here, k is the time horizon over which the arithmetic mean is calculated. Thus, $\boldsymbol{s}_t$ is the average displacement of the past k time steps.

Latency is one of the major problems of the unweighted moving average, meaning that at least k samples must be recorded before a prediction can be made. Another problem is given by the fact that the $\boldsymbol{w}_t$ are displacement vectors. Thus, the moving average rewritten using (2), reduces to

$$\boldsymbol{s}_t = \frac{1}{k} \sum_{i=0}^{k-1} (\boldsymbol{x}_{t-i} - \boldsymbol{x}_{t-i-1}) = \frac{1}{k}(\boldsymbol{x}_t - \boldsymbol{x}_{t-k}) \tag{4}$$

As a result, all position samples in the time interval between t-1 and t-k+1 are actually ignored by the smoother. Thus, the prediction will become unstable e.g. due to noise, if the time horizon is not chosen properly.

Exponential Smoothing. When using a simple moving average smoother, all k past measurements will have the same weight, which might not be feasible since older values are of less importance than newer ones. Thus, an exponential smoothing can be used which weighs past measurements with an exponentially decaying factor.

$$\boldsymbol{s}_0 = \boldsymbol{w}_0 \tag{5}$$

$$\boldsymbol{s}_t = \alpha \boldsymbol{w}_t + (1-\alpha)\boldsymbol{s}_{t-1} \tag{6}$$

$\boldsymbol{s}_t$ can be rewritten as

$$\boldsymbol{s}_t = \alpha \boldsymbol{w}_t + \alpha \sum_{i=1}^{t-1} (1-\alpha)^i \boldsymbol{w}_{t-i} + (1-\alpha)^t \boldsymbol{w}_0 \tag{7}$$

With exponential smoothing, all past measurements are included into the current prediction (infinite impulse response filter). The efficiency of the smoothing can be controlled with the factor $\alpha \in (0, 1)$. If α is chosen close to 1, there is only a little smoothing since new measurements are weighted higher. On the other hand, if α is chosen close to 0, there is a high level of smoothing.

Choosing the correct value of α for exponential smoothing is a delicate task. If α is chosen too high, the smoothing is low, and noise will influence the correct prediction. If α is chosen too low, the latency of the smoother increases and real changes in the walk direction might be detected too late. In order to set a mathematical basis for the correct adjustment of α, the following limit case can be regarded. It is assumed that the measurements were constant with the value v_o, while the smoothed value was stable at $s_o = v_o$. At the time t-k, the measurement abruptly changes to a new constant value v_n. Now, the question is how many time steps k it will take until s_t reaches the factor q of the new measurement value v_n. Hence s_t becomes

$$\boldsymbol{s}_t = \alpha v_n + \alpha v_n \sum_{i=1}^{k} (1-\alpha)^i + \alpha v_o \sum_{i=k+1}^{t-1} (1-\alpha)^i + (1-\alpha)^t v_o \tag{8}$$

Now the factor of change from the old to the new measurement is given as

$$q = \frac{s_t - v_0}{v_n - v_o} \tag{9}$$

$$= 1 - (1-\alpha)^{k+1} \tag{10}$$

$$\alpha = 1 - (1-q)^{1/k+1} \tag{11}$$

Using the equations for the summation of geometric series on (8) and inserting in (9) gives an equation for determining α. For instance, if the smoother is to follow a step function to 80% ($q = 0.8$) within the next $k = 180$ measuring time steps, α should be about 0.009.

Double Exponential Smoothing. Exponential smoothing can be improved if there is a trend in the data like the change of the walking direction. This so-called double exponential smoothing is given by

$$\boldsymbol{s}_0 = \boldsymbol{w}_0 \tag{12}$$

$$\boldsymbol{s}_t = \alpha \boldsymbol{w}_t + (1-\alpha)(\boldsymbol{s}_{t-1} + \boldsymbol{b}_{t-1}) \tag{13}$$

$$\boldsymbol{b}_t = \beta(\boldsymbol{s}_t - \boldsymbol{s}_{t-1}) + (1-\beta)\boldsymbol{b}_{t-1} \tag{14}$$

The $\boldsymbol{b}_t$ vectors represent the current trend in the data. Like in the normal exponential smoothing, α is the data smoothing factor. $\beta \in (0, 1)$ is the so-called trend smoothing factor, which controls how much the current trend is influenced by the change in the smoothed prediction output over time. $\boldsymbol{b}_0$ defines the initial trend in the data.

Following the definition of $\boldsymbol{w}_t$ in (2), it becomes obvious that $\boldsymbol{w}_0$ is not defined by the data. In order to apply the smoothing methods, prediction thus either has to start at time t=1 or some initial value for $\boldsymbol{w}_0$ is required. The following method is proposed for a path prediction with a double exponential smoothing. Under the assumption that a person starts moving at the time t=0, the facing direction at time 0 is used to set $\boldsymbol{w}_0 = c\boldsymbol{f}_0$. This approach is similar to the one proposed by [12], but is more explicitly formalized here. It is now further assumed that the initial facing direction is the most likely direction of movement, which allows setting the initial trend to $\boldsymbol{b}_0 = (0, 0)$. The magnitude of $\boldsymbol{w}_0$ influences how much impact it has on the smoothed output. Therefore the constant c should be chosen so that it reflects a reasonable speed. As an example, c could be defined by taking into account the averaged speed of human walk $\bar{v}$ together with the given update rate r as $c = \bar{v}/r$.

4 User Study

Within the user study we conducted, participants had to walk inside a simple maze-like environment. There were different possible paths through this maze, all of them with the same complexity. Thus, the participants were forced to make a choice which way to go. During the study, users wore a head-mounted tracking system (inside-out system) in order to record to walking trajectory. The goal of the study was to measure head position and orientation while users had to make a choice, and how this choice of a direction relates to the recorded data. Using the set of recorded user paths, a comparison of different path prediction approaches and the adjustment of smoothing parameters was possible.

4.1 Experimental Setup

Movable walls were installed in 5m x 7m room to realize a T-shaped maze. The design and the dimensions of this shape are shown in Fig. 1. The maze had an obstacle in its center in order to force subjects to decide whether they want to go left or right, see Fig. 2. At both ends of the maze, for different calendar pictures were attached (see Fig. 1). In order to avoid any biasing, the maze was perfectly symmetric (left, right) and no distractions were present except for the calendar pictures. However, these pictures could not influence the user during his decision phase whether to move left or right, since they were not visible at the beginning. The participants started the user study in the center of the maze's lower end.

Paper markers for the the Intersense IS-1200 tracking system [24] were attached to the ceiling of the room, in which the T-maze was installed. The sensor of the IS-1200 tracking system (inside-out tracking system) can be easily attached to an HMD. It allows tracking 6 degrees of freedom, i.e. the user's position and orientation, at an update rate of 180 Hz. Since the user should not be irritated by any wire connections, a notebook for recording data was mounted in a backpack the user had to wear during the study. The tracking system was attached to a Triviso Scout HMD. Since we built a real maze, the HMD was only

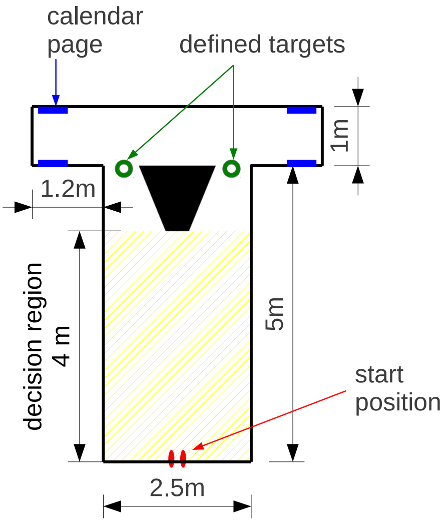


Fig. 1. Layout of the symmetric T-maze. Walls are in black and the four blue bars mark the position of the calendar pictures. The decision region and the targets (left, right) are only used for the analysis and are not visible for participants.

used to mount the tracking system properly on a subject's head, while we did not use the HMD's visualization capabilities. The maze environment and user wearing the notebook and the tracking system are shown in Fig. 2. As mentioned before, a static coordinate transformation was applied to the achieved tracking data in order to actually determine the approximate center of a participant's head.

In order to avoid any deficiencies of a virtual environment, a physical maze environment was constructed deliberately instead of a virtual environment. Thus, the HMD was not used even though the hardware would allow visiting a virtual environment. By this, we avoid any biasing due to the HMD's limited field of view (FOV) or due to simulator latencies.

4.2 Participants

In total, 11 participants took part in the experiment (7 male and 4 female, median age 31). The participants came from the institute and included students, senior researchers, as well as administrative staff. None of the participants was aware of the purpose of the user study.

4.3 Tasks and Conditions

In order to analyze if there is a difference regarding the walking direction, the facing direction or the whole user trajectory between a known and unknown environment, both cases were investigated. It might happen for instance that in

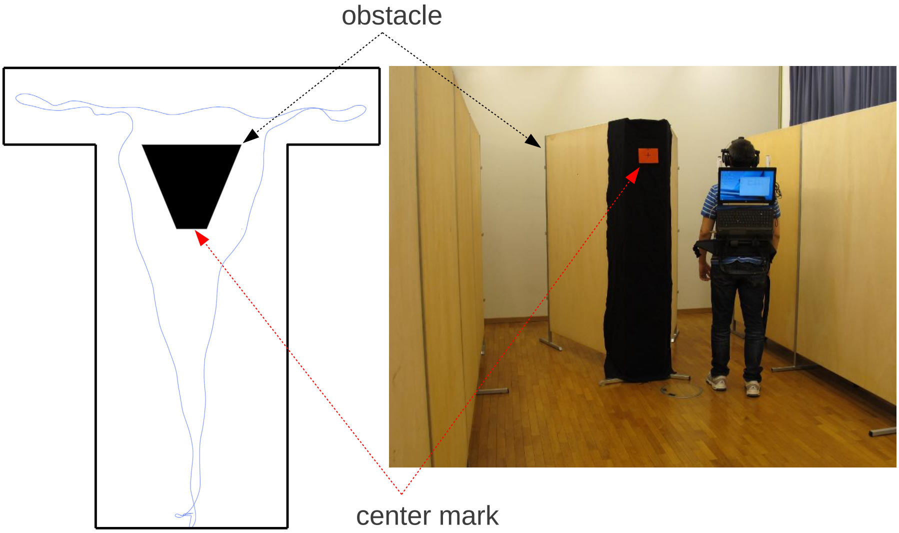


Fig. 2. Left: a typical walk path through the T-maze. Right: a subject during the user study. The subject wears the backpack with the notebook and the HMD with the tracking system. The obstacle forces participants to either walk left or right. The center mark is used for the initial calibration of the facing direction.

an unknown environment users will decide later whether they want to go left or right, or their facing direction might change more often due to orientation behavior.

Explore Condition (EX). Within the first task, the subjects were new to the environment and had no information about the maze. Their task was to walk through all corridors in the maze and to return to the starting point afterwards. The calendar pictures in the maze were of no relevance for this task.

Count Task Condition (CO). For the second task, the calendar pictures at the ends of the T-maze became relevant (see Fig. 1). Within this task, participants had to count the total number of Sundays on all 4 calendar pictures, and then had to return to their starting position. For this second condition, all participants were already familiar with the maze and knew where to find the targets.

4.4 Procedure

Prior to the study, all participants got oral explanations about the task and procedure. The participants were prepared in a separate room, where the HMD and the backpack were mounted. Then they were guided blindfolded to the starting position, so that they had no prior knowledge about the T-maze. After placing themselves properly at the start position, they were instructed to look at the center mark for 4 seconds for calibration before starting to walk (see Fig. 2). This initial calibration of the facing direction was performed to align the orientation of the tracking system with a participant's facing direction. Once a participant finished the first task (EX condition), the second task (CO condition) was explained to him.

5 Results and Discussion

The tasks of the user study were finished by 10 out of 11 participants. Although one participant did not properly finish the task in the EX condition, he could participate again in the CO condition. This results in 10 correct paths for the explore task (EX), and in 11 for the count task (CO).

In the following plots, the analysis of the tracking data is limited to the so-called decision region, which is within the first 4 m from the start position towards the obstacle. This direction is referred to as y-axis and the value y indicates the position in meters from 0 at the start position. Within this region, all participants had to decide whether they want to go left or right. Within the T-maze, two targets - i.e. the passage ways - were defined at a distance of $y = 5$ from the starting point. These targets were named 'left' and 'right' and are thus located in the centers of the possible paths around the obstacle, see Fig. 1.

As suggested in Sect. 3.1, the angular deviation of the prediction from the targets is used for the evaluation of the path prediction. Next, the hit rate $r \in [0, 1]$ is calculated, which is the relative amount of time that a predictor predicts the correct target that a subject finally walked to. As shown in Fig. 3, all participants had already decided for a specific target until $y = 3$. Based on these results, the calculation of the hit rate is limited to the first 3 m ($y \in [0, 3]$), which is referred to as the evaluation area.

In Fig. 3, all paths that were recorded during the user study are shown. For most paths, the gait oscillations of the user's head are clearly visible. The paths also show that most of subjects decided very early, i.e. $y \leq 1$, for left or right. This is in particular the case for the CO condition, in which only one of 11 subjects changed the initially chosen direction after $y \approx 1.4$. During the EX condition, three to four of 10 subjects decided later at $y \approx 2.5$. This indicates that an early path prediction is more difficult if an environment is unknown to the user.

The average walk speed in the decision area for both conditions was 0.95 m/sec, with a standard deviation of 0.22 m/sec.

5.1 Facing Direction

Results. The angular deviation of the facing direction from the targets is shown in Fig. 4. Three different subjects are represented by three plots while they moved from $y = 0$ to $y = 4$.

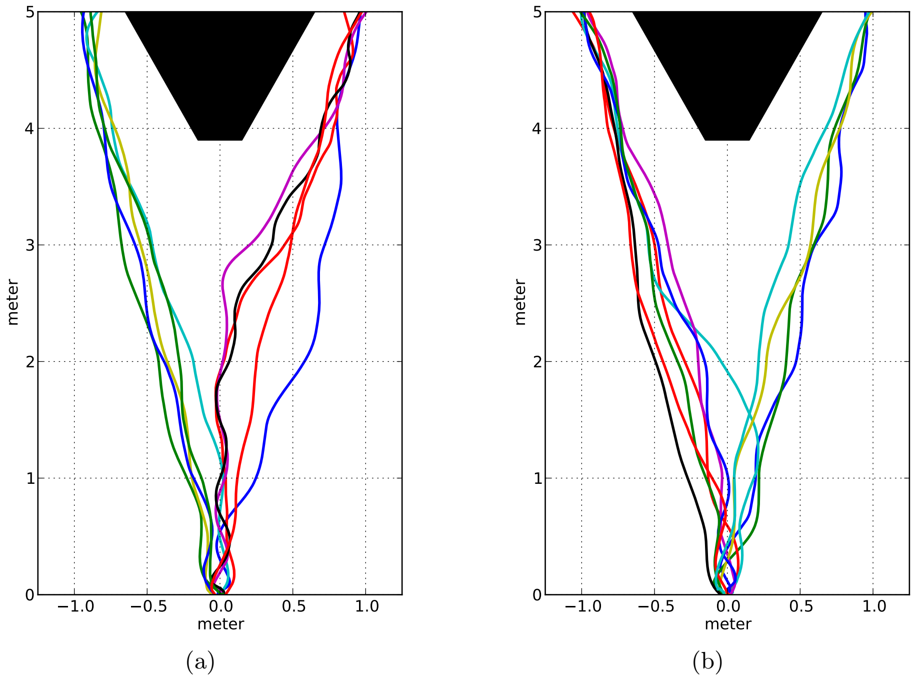


Fig. 3. All user paths from the study. Only the lower part of the T-maze is shown. (a) EX condition paths and (b) CO condition paths.

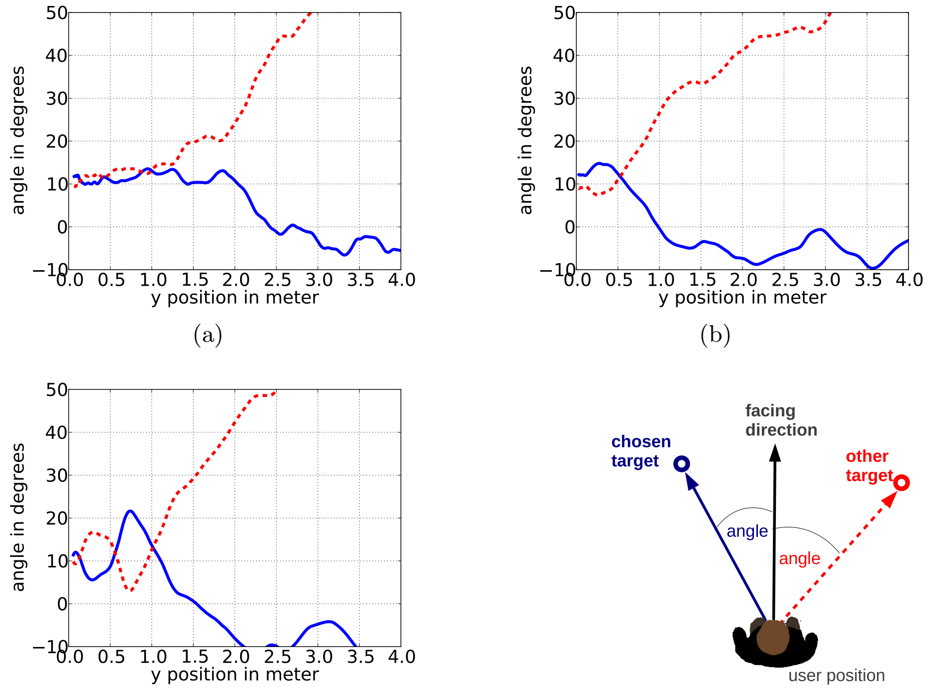
Based on the facing direction, a predictor would thus predict the target which is closer 0 degree. In Fig. 4(c) for instance, the predictor would choose the wrong target for $y \in [0.6, 1]$, while its prediction would be correct for $y > 1$.

Table 1 shows the relative hit rate r within the evaluation area.

Table 1. Relative hit rate of the facing direction predictor averaged over all participants within the evaluation area $y \in [0, 3]$

	EX	**CO**	**both**
mean	0.91	0.95	0.93
standard deviation	0.09	0.06	0.08

Discussion. The data shown in Fig. 4 is the raw data from the tracking system. It only contains little noise and thus the plots represent quite well a subject's head orientation. For this case, smoothing cannot further improve the prediction. In Fig. 4(c) at position $y = 0.7$ for example, the participant turned his head to look towards another target.



(a) (b) (c) (d) Explanation of deviation angles

Fig. 4. Angular deviation of the facing direction from the two targets while a subject is moving from y=0 to y=4. (a), (b) were recorded during the EX condition and (c) was recorded during the CO condition. A solid blue line shows the deviation from the chosen target direction and the dashed red line the deviation from the other target, see (d). The angular deviation is positive for rotations from the targets towards the center of the obstacle.

Generally spoken, the facing direction should be used with care for a prediction. This is mainly because of two reasons. First, the head orientation does not necessarily correspond to a person's actual gaze, since a person might just turn his eyes instead of the whole head. Second, the facing direction is not always aligned with the walking path, but can be highly influenced by visual distractions in the environment. Although the performed user study was designed to have no visual distractions, any picture on the wall would have, however, made a subject looking at it while walking towards the target. Thus, the average facing direction would have been biased by the picture and thus would not point to the target.

For the user study mentioned in the above, it can be assumed that the high hit rate is mainly due to the experimental design. Nevertheless, [14] shows that most of the time the gaze direction is aligned with the walking path. This could also explain the high hit rate, but does not imply that the facing direction is a stable predictor in a general environment.

Because of the user study's design (see Sect. 4.4), an analysis of the initial facing direction when a person just starts walking, is not possible and is omitted.

5.2 Walking Direction and Smoothing

Results. The angular deviation of the walking direction from the targets for the EX and CO condition are shown in Fig. 6. The four plots represent four different subjects while they moved from $y = 0$ to $y = 4$. The walking direction is calculated as given in Sect. 3.2. The angular deviation from the chosen target (out of two) is represented by the solid line, while the dashed line shows the deviation from the other target. Gray lines show the raw measurement data without any smoothing. See Fig. 5 for an explanation.

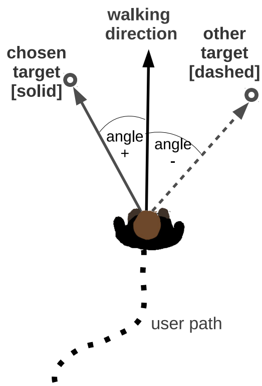


Fig. 5. Definition of the deviation angles for the walking direction analysis (see Fig. 6)

As for the facing direction above, the relative hit rate r is calculated within the evaluation area, see Table 2. Further, an additional robustness factor is introduced in order to demonstrate how smoothing could improve the robustness of the prediction.

As it could be seen from the above analysis, the predictor could switch between the two targets depending on smallest angle between the facing direction and one of the given targets. Any switch, however, is counterproductive to the robustness of a predictor and should be avoided. Thus, the relative robustness measures the average reduction in the number of prediction switches between the two predefined targets for a smoothed prediction in relation to the raw walking direction data. A relative robustness of 2 for the exponential smoother for example means that the exponentially smoothed predictor makes only half as many switches between left and right target as the raw walk direction predictor.

Discussion. In all plots, gait oscillations are clearly visible. These gait oscillations are the largest and most challenging disturbance when using walking direction data for prediction. Compared to the magnitude of gait oscillations,

Table 2. Relative hit rate of the walking direction predictor averaged over all participants within the evaluation area $y \in [0, 3]$. The standard deviation is denoted by sd. The relative robustness shows the improvement of the smoothing on the prediction relative to the raw walk direction prediction.

		EX	CO	both
	mean	**0.81**	**0.87**	**0.84**
raw walk direction	sd	0.13	0.10	0.12
	rel. robustness	1	1	1
	mean	**0.85**	**0.85**	**0.85**
exp. smoothed	sd	0.16	0.16	0.16
	rel. robustness	**2.0**	**1.8**	**1.9**
	mean	**0.86**	**0.86**	**0.86**
double exp. smoothed	sd	0.16	0.16	0.16
	rel. robustness	**2.0**	**2.1**	**2.0**

tracking noise is negligible. In order to overcome disturbances from gait oscillations, it is essential to apply smoothing. In Fig. 6, the colored lines represent smoothed values using either exponential or double exponential smoothing. The initial value of $\boldsymbol{w}_0$ is chosen as suggested in Sect. 3.3 ($\boldsymbol{f}_0$ points to the center of the obstacle). Compared to this, smoothing with a moving average performs worst due to the problems mentioned in the above. Hence, the results are not presented here.

Like for the facing direction, a predictor based on the walking direction also predicts that target which is closer to 0 degree in the plot. This means in other words that a wrong target is predicted if the dashed line is closer to 0 than its solid counterpart.

As shown in the above, the efficiency of the smoothers strongly depends on the correct setting of the smoothing factors α and β. For the given experiment, a smoothing factor of $\alpha = 0.004$ (for normal and double exponential smoothing) and a trend smoothing factor of $\beta = 0.004$ turned out to work best. These factors provide a stable prediction, while the smoothers still react in a reasonable time to changes in the intended direction of movement. Using (9) and the given update rate of 180 Hz, the percentage of change can be estimated. One second after a change in the walking direction, roughly 50% of that change will be included into the smoothed walking direction and roughly 75% after two seconds.

As it can be seen from Table 2, the relative hit rate is not significantly higher for a smoothed predictor in relation to an unsmoothed predictor. However, there is a significant improvement in the robustness of the prediction. Regarding robustness, the double exponentially smoothed predictor performs best.

Figure 6(a) and (b) both show data from subjects who did not decide right away for a target and first walked roughly straight forward. As it can be seen in Fig. 6(b) for instance, the double exponential smoother is slightly faster in detecting the decision and outruns the exponential smoother by $y \approx 2.2$. Figure 6(d) shows the extreme case, in which a participant changed his decision and turned around (see also Fig. 3). Similarly, the double exponential smoother is

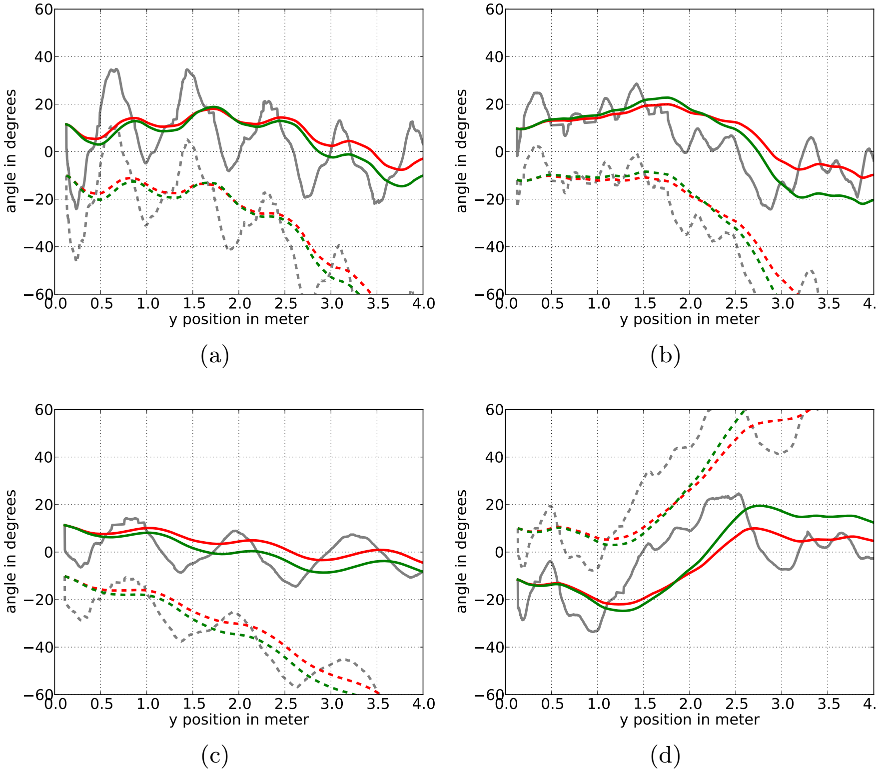


Fig. 6. Angular deviation of the walking direction from the two targets while a subject is moving from y=0 to y=4. (a) and (b) are from the EX condition, (c) and (d) from the CO condition. A solid line represents the deviation from the chosen target and a dashed line the deviation from the other target. The gray line is the raw data, red is the exponentially smoothed data and green is the double exponentially smoothed data.

faster in detecting this change. Figure 6(c) shows a typical plot of a subject who walked straight to the chosen target. In such cases, both smoothers provide reliable and robust predictions.

6 Conclusion

In this paper it is analyzed how head tracking data can be used to retrieve a person's intended direction of travel. A tracking system mounted on a person's head provides position and orientation data. As it is shown in this paper, this data suffices to make a robust and reliable short term path prediction that holds for a few seconds.

Using head position data, a displacement vector can be calculated to retrieve the head's movement direction. Intuitively, one would use this direction to extrapolate the person's intended direction of travel. However, as it is shown in the conducted user study, raw position data from the tracking data gives quite poor prediction results. In fact due to the mechanics of human gait, a person's head movement is not nicely aligned with the intended direction of travel.

Hence, in order to increase the reliability and the robustness of the prediction, a smoothing of the tracking data is required. Different smoothing methods are introduced that are suitable for path prediction. Especially double exponential smoothing turns out to work best and doubles the robustness of the prediction compared to using raw tracking data.

Nevertheless, using smoothing increases the latency of the path prediction. Hence, a suitable trade-off has to be found between robustness and latency. For this purpose, an equation is derived that allows choosing the proper smoothing factor depending on the requirements.

Head orientation data actually is a person's facing direction, which provides another approach for path prediction. However, it is problematic to use only the facing direction for prediction, since a person might move the eyes instead of the head. This could happen in particular if some targets or close to each other and thus making it obsolete to turn the head. Nevertheless, if a person is not moving or just starts to walk, the facing direction can be used as an initialization value for the walking direction smoother.

7 Future Work

To overcome the problem with the facing direction, eye trackers could be used together with a regular tracking system in order to correctly identify the user's gaze direction. Even though Hollands et al. [14] thoroughly studied how gaze behavior and walking direction are related to each other, it requires further research to answer the question how gaze behavior could be used for prediction.

The conducted user study has shown that the major problem for path prediction from head tracking data is "gait noise". So far, this problem has been solved with generic smoothing approaches while in fact this noise is partially systematic. It depends on the mechanics of human gait, the step length, person specific walking patterns, and so forth. Hence, these patterns instead could be learned from the data to better filter the data. E.g. a physical gait model together with a Kalman filter could provide far better smoothing with low latency than exponential smoothing. Similar approaches used for so-called walk-in-place interaction devices might be adapted for this purpose [25].

Additional sensors could also be used to get more data for better prediction. For instance, accelerometers might be used to detect steps and thus learn the step length from the user data. Knowing the step length for example would allow an automatic adaptation to a person in order to improve the prediction. Other approaches could be more suitable for pocket worn mobile phones. In such a scenario, the displacement of the pelvis while walking might give additional information for path prediction [26].

Acknowledgment. The authors would like to thank the Swiss National Science Foundation (project number 127298) for funding this work.

References

1. Souman, J.L., Giordano, P.R., Schwaiger, M., Frissen, I., Thümmel, T., Ulbrich, H., Luca, A.D., Bülthoff, H.H., Ernst, M.O.: Cyberwalk: Enabling unconstrained omnidirectional walking through virtual environments. TAP 2008: Transactions on Applied Perception 8(4), 25:1–25:22 (2008)
2. Harms, H., Amft, O., Winkler, R., Schumm, J., Kusserow, M., Troester, G.: Ethos: Miniature orientation sensor for wearable human motion analysis. IEEE Sensors, 1037–1042 (2010)
3. Leuenberger, K., Gassert, R.: Low-power sensor module for long-term activity monitoring. In: 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC, pp. 2237–2241 (2011)
4. Welch, G., Bishop, G.: Scaat: incremental tracking with incomplete information. In: Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 1997, pp. 333–344. ACM Press/Addison-Wesley Publishing Co. (1997)
5. Razzaque, S., Swapp, D., Slater, M., Whitton, M.C., Steed, A.: Redirected walking in place. In: Proceedings of the Workshop on Virtual Environments, EGVE 2002, pp. 123–130. Eurographics Association (2002)
6. Nitzsche, N., Hanebeck, U.D., Schmidt, G.: Motion compression for telepresent walking in large target environments. Presence: Teleoperators and Virtual Environments 13(1), 44–60 (2004)
7. Steinicke, F., Bruder, G., Kohli, L., Jerald, J., Hinrichs, K.: Taxonomy and implementation of redirection techniques for ubiquitous passive haptic feedback. In: Proceedings of the 2008 International Conference on Cyberworlds, CW 2008, pp. 217–223. IEEE (2008)
8. Gandhi, T., Trivedi, M.: Image based estimation of pedestrian orientation for improving path prediction. In: 2008 IEEE Intelligent Vehicles Symposium, pp. 506–511 (2008)
9. Peck, T., Fuchs, H., Whitton, M.: The design and evaluation of a large-scale real-walking locomotion interface. TVCG 2012: IEEE Transactions on Visualization and Computer Graphics 18(7), 1053–1067 (2012)
10. Razzaque, S., Kohn, Z., Whitton, M.C.: Redirected walking. In: Proceedings of Eurographics, pp. 289–294 (2001)
11. Su, J.: Motion compression for telepresence locomotion. Presence: Teleoperators and Virtual Environments 16(4), 385–398 (2007)
12. Interrante, V., Ries, B., Anderson, L.: Seven league boots: A new metaphor for augmented locomotion through moderately large scale immersive virtual environments. In: IEEE Symposium on 3D User Interfaces, 3DUI 2007, pp. 167–170. IEEE (2007)
13. Grasso, R., Prvost, P., Ivanenko, Y.P., Berthoz, A.: Eye-head coordination for the steering of locomotion in humans: an anticipatory synergy. Neuroscience Letters 253(2), 115–118 (1998)
14. Hollands, M., Patla, A., Vickers, J.: "look where you're going!": gaze behaviour associated with maintaining and changing the direction of locomotion. Experimental Brain Research 143, 221–230 (2002)
15. Koo, S., Kwon, D.S.: Recognizing human intentional actions from the relative movements between human and robot. In: The 18th IEEE International Symposium on Robot and Human Interactive Communication, pp. 939–944 (2009)

16. Hoogendoorn, S.P.: Pedestrian flow modeling by adaptive control. Transportation Research Record: Journal of the Transportation Research Board 1878(1), 95–103 (2004)
17. Foxlin, E.: Pedestrian tracking with shoe-mounted inertial sensors. CGA 2005: Computer Graphics and Applications 25(6), 38–46 (2005)
18. Fischer, C., Muthukrishnan, K., Hazas, M., Gellersen, H.: Ultrasound-aided pedestrian dead reckoning for indoor navigation. In: Proceedings of the First ACM International Workshop on Mobile Entity Localization and Tracking in GPS-Less Environments, MELT 2008, pp. 31–36. ACM (2008)
19. Kiruluta, A., Eizenman, M., Pasupathy, S.: Predictive head movement tracking using a kalman filter. IEEE Transactions on Systems, Man, and Cybernetics 27(2), 326–331 (1997)
20. LaViola, J.J.: Double exponential smoothing: an alternative to kalman filter-based predictive tracking. In: Proceedings of the Workshop on Virtual Environments 2003, EGVE 2003, pp. 199–206. ACM (2003)
21. van Rhijn, A., van Liere, R., Mulder, J.: An analysis of orientation prediction and filtering methods for vr/ar. In: Proceedings of the IEEE Virtual Reality, VR 2005, pp. 67–74 (March 2005)
22. Hirasaki, E., Moore, S.T., Raphan, T., Cohen, B.: Effects of walking velocity on vertical head and body movements during locomotion. Experimental Brain Research 127, 117–130 (1999)
23. Terrier, P., Schutz, Y.: How useful is satellite positioning system (gps) to track gait parameters? a review. Journal of NeuroEngineering and Rehabilitation 2(1), 28 (2005)
24. Foxlin, E., Naimark, L.: Vis-tracker: A wearable vision-inertial self-tracker. In: Proceedings of the IEEE Conference on Virtual Reality, VR 2003, p. 199. IEEE, Washington, DC (2003)
25. Wendt, J., Whitton, M., Brooks, F.: Gud wip: Gait-understanding-driven walking-in-place. In: Proceedings of the IEEE Conference on Virtual Reality, VR 2010, pp. 51–58 (March 2010)
26. Zijlstra, W., Hof, A.: Displacement of the pelvis during human walking: experimental data and model predictions. Gait and Posture 6(3), 249–262 (1997)