ORIGINAL ARTICLE

Evaluation of the Oculus Rift S tracking system in room scale virtual reality

Riccardo Monica[1] · Jacopo Aleotti[1]

Abstract

In specific virtual reality applications that require high accuracy it may be advisable to replace the built-in tracking system of the HMD with a third party solution. The purpose of this research work is to evaluate the accuracy of the built-in tracking system of the Oculus Rift S Head Mounted Display (HMD) in room scale environments against a motion capture system. In particular, an experimental evaluation of the Oculus Rift S inside-out tracking technology was carried out, compared to the performance of an outside-in tracking method based on the OptiTrack motion capture system. In order to track the pose of the HMD using the motion capture system the Oculus Rift S was instrumented with passive retro-reflective markers and calibrated. Experiments have been performed on a dataset of multiple paths including simple motions as well as more complex paths. Each recorded path contained simultaneous changes in both position and orientation of the HMD. Our results indicate that in room-scale environments the average translation error for the Oculus Rift S tracking system is about 1.83 cm, and the average rotation error is about 0.77°, which is 2 orders of magnitude higher than the performance that can be achieved using a motion capture system.

Keywords Head mounted displays · Tracking technologies · Room-scale virtual reality

1 Introduction

In many virtual reality applications that require high accuracy of Head Mounted Display (HMD) tracking, it may be advisable to replace the built-in tracking system of the HMD with a third party solution (Debarba et al. 2018). For example, interaction with physical objects in industrial or clinical medicine tasks requires a highly accurate correspondence between the virtual environment and the real world. The goal of this work is to provide a quantitative comparison between the built-in tracking system of the Oculus Rift S HMD with respect to the accuracy that can be attained by exploiting a motion capture system, which acts as ground truth. Indeed, motion capture systems work at high speed and achieve a sub-millimeter accuracy (Merriaux et al. 2017).

✉ Jacopo Aleotti
jacopo.aleotti@unipr.it

Riccardo Monica
riccardo.monica@unipr.it

1   Department of Engineering and Architecture, University of Parma, Parco Area delle Scienze, 181A, 43124 Parma, Italy

The Oculus Rift S belongs to the second generation of consumer VR HMDs (since 2016). It is a tethered device that exploits the hardware of an external computer (CPU, graphics card, and RAM) to deliver high quality virtual reality experiences. The Oculus Rift S does not require any external device for positional tracking. Instead, it features five cameras that enable inside-out tracking.

In general, the two common approaches for HMD tracking are called outside-in and inside-out (Rolland et al. 1999). In outside-in systems multiple fixed external cameras are used to track the pose (3D position and 3D orientation) of the HMD. In particular, the external cameras track a set of reference points located on the headset and on the controllers (if any). Usually, the set of reference points is a pattern (constellation) of IR LEDs or passive (retro-reflective) markers. The pose of the HMD can be obtained in an absolute reference frame defined in a calibration step. Outside-in tracking systems are generally faster and more accurate than inside-out systems. Moreover, the localization accuracy of outside-in systems can be improved by adding more cameras. Other advantages of the outside-in technologies are that they work even in the dark, they can be used to track the HMD and the body of the user simultaneously (also including external

rigid objects), and that hand controllers can be tracked even if the user has them behind his/her back. The disadvantages of outside-in tracking systems are that the HMD must be instrumented with reference points, and that these systems are much more expensive.

Inside-out tracking systems use cameras placed on the HMD looking outward. An algorithm based on visual-inertial odometry determines in real-time the position and the orientation of the HMD by observing low-level features of the surrounding environment. The pose of the HMD can be determined only relative to the initial headset configuration. Inside-out HMD tracking systems are easier to set up and offer reduced costs. In particular, calibration is straightforward as there is no need to install fixed cameras with mounts or to instrument the environment with markers. The main disadvantage of inside-out technologies is that tracking is less accurate.

The main contribution of this paper, which was not considered in previous works, is the evaluation of the Oculus Rift S inside-out tracking technology in a room scale virtual reality setup, against an outside-in tracking system based on the OptiTrack motion capture. To this purpose, the Oculus Rift S was instrumented with passive markers and calibrated. A dataset of HMD movements of a user walking around the environment has been recorded. Each recorded path contains simultaneous changes in both position and orientation of the HMD. The dataset includes paths that vary from simple straight motions to more complex and longer random walks. Our results indicate that in room-scale environments the average translation error for the Oculus Rift S tracking system is about 1.83 cm, and the average rotation error is about 0.77°, which is 2 orders of magnitude higher than the performance that can be achieved using a motion capture system.

The paper is organized as follows. Section 2 reviews the state-of-the-art research on the evaluation of HMDs tracking accuracy. Section 3 describes the method used in this study, including the experimental setup, the calibration and data acquisition techniques, the acquired dataset, and the evaluation metrics. Section 4 illustrates the experimental results, while Sect. 5 draws conclusions.

2 Related work

The closest work to ours is by Jost et al. (2021), where a quantitative evaluation of the Oculus Rift S was carried out in a controlled and small-scale environment using an industrial robot to move the HMD. Translation and rotation were tested separately. The results indicated a high accuracy for both translation (1.66 ± 0.74 mm) and rotation ($0.34 \pm 0.38°$). The main differences to our work are that we consider more ample movements performed in a room-scale

environment, and that the movements are more complex, i.e., they contain changes in both rotation and translation.

Most previous works on the evaluation of HMDs tracking accuracy focused on devices that belong to the first generation of consumer VR (since 2016), like the Oculus Rift (DK1, DK2 and CV1) and the HTC Vive. The rotation accuracy of the Oculus Rift DK1 was evaluated by Xu et al. (2015) showing a good estimate of full range motions in cervical spine mobility measurements. The validity of the Oculus Rift DK2 to assess postural changes during balance tasks was investigated by Marchetto and Wright (2019). It was shown that the HMD may be successfully used for assessing postural control without external posturography equipment. A user study was conducted by Chessa et al. (2019) to evaluate the perceptual quality of the Oculus Rift DK2 for immersive virtual reality. The device enabled a strong sensation of presence and did not provoke undesired effects such as cybersickness or fatigue in short tasks. A computer vision approach was presented by Chang et al. (2016), using a high-speed camera, to evaluate timing and accuracy of the Oculus Rift DK2.

An evaluation of the HTC Vive HMD was performed by Niehorster et al. (2017) at static poses along a grid of lines drawn on the floor. An analysis of the spatial tracking performance of the HTC Vive HMD was conducted in small scale environments by Jost et al. (2019) using a motion capture system as ground truth, showing high accuracy. A similar analysis was carried out, in larger environments, by Ikbal et al. (2021) using an industrial robot as ground truth source. The results indicated an average error of about 3 mm and 0.5°. The HTC Vive lighthouse positioning system was evaluated by Greiff et al. (2019) for tracking micro unmanned aerial vehicles, showing sub-centimeter position accuracy. A simplified error model for HTC Vive tracking system was proposed by Wu et al. (2020). The method can be adopted to predict in advance the magnitude of tracking errors in a given configuration of multiple lighthouses (transmitters) and receivers.

A comparison between Oculus Rift HMDs and the HTC Vive was presented in different works. In Suznjevic et al. (2017) the HTC Vive and the Oculus Rift CV1 were compared in terms of ease of use, intuitiveness and quality of experience when performing pick and place tasks in virtual reality. In general, the HTC Vive was marginally better. In Borrego et al. (2018) the Oculus Rift CV1 and the HTC Vive were evaluated in terms of accuracy and jitter. Both devices showed good and similar performance at sitting, while the HTC Vive presented worse accuracy and jitter at standing height, even though it must be recalled that the HTC Vive provides a working area twice as large as that of the Oculus Rift CV1. In Lubetzky et al. (2019) head tracking performance of the Oculus Rift CV1 was compared against the HTC Vive HMD during static and dynamic standing tasks in virtual environments. The

results indicated excellent agreement between the two HMDs with respect to a motion capture system. A weaker agreement was observed for vertical displacement in a static task and moderate agreement was observed for pitch and yaw displacement in a dynamic task.

In Bauer et al. (2021) the performance of the HTC Vive Pro HMD was evaluated, showing a high reproducibility of a few millimeters. However, the HTC Vive Pro tracking system has issues when several lighthouses are used, and it has systematic effects like a tilted reference plane. Other studies involved the HTC Vive tracker (a small device that includes the same tracking technology of the Vive HMD) and its motion controllers. A hybrid tracking system was developed by Groves et al. (2019) using the HTC Vive Pro controller, which enabled optical tracking of a surgical instrument with respect to the HMD, achieving sub-millimeter accuracy. The accuracy of the HTC Vive tracker was investigated by Borge et al. (2018) and by Ameler et al. (2019). In Borge et al. (2018) a robot was used as ground truth, while in Ameler et al. (2019) the OptiTrack motion capture system served as reference. An accuracy ranging from sub-millimeter to millimeter was obtained. The accuracy of the Vive trackers for rehabilitation and medical tracking tasks was investigated by van der Veen et al. (2019), suggesting that the HTC Vive sensors can be used successfully for clinical analysis of human motions. The static accuracy of HTC Vive tracker and motion controller was evaluated by Spitzley and Karduna (2019). The measured errors of both VIVE sensors were below 0.4° and 3 mm. In Flueratoru et al. (2020) the HTC Vive tracker was adopted as ground truth system for UWB indoor localization, while in Lwowski et al. (2020) the HTC Vive Tracker was employed for robot localization. An investigation of the HTC Vive tracking system for gait analysis was carried out by Guaitolini et al. (2021) indicating that the device can accurately monitor gait parameters. In Palma et al. (2021) an augmented reality system was proposed that allows users to interact with a 3D-printed copy of an artefact in a virtual environment using a physical replica (tracked by the HTC Vive tracker) as a tangible user interface.

Approaches for six degrees of freedom human body pose estimation based on the HTC Vive lighthouse transmitters were presented in Caserman et al. (2019), and in Jansen et al. (2019) for automatic calibration. In Vox et al. (2021) a method for human body tracking was developed, based on the HTC Vive tracker and on an inverse kinematic model of the human body, and it was compared against a marker-based optical motion capture system showing some inaccuracies.
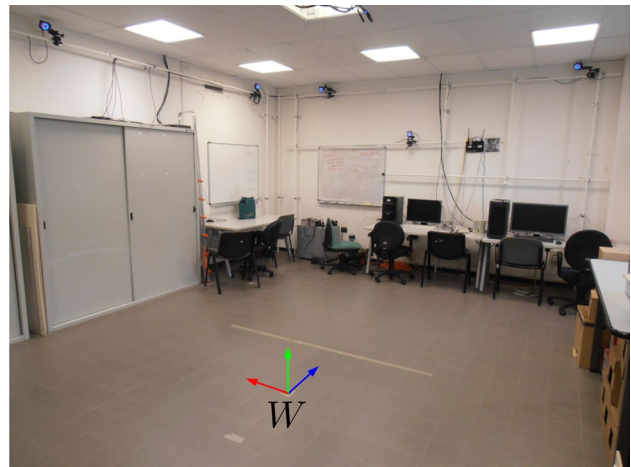

Fig. 1 The room scale environment for the experimental evaluation. The OptiTrack world reference frame *W* is located on the floor in the center of the room, with the *y*-axis pointing upward

Fig. 2 OptiTrack Prime 13 camera

3 Methods

3.1 Experimental setup

The experimental setup consists of a room of size $8.2 \times 5.5 \times 2.9$ m, shown in Fig. 1. In order to perform the outside-in tracking of the HMD an OptiTrack motion capture system was adopted with twelve Prime 13 cameras. This configuration allows an effective capture volume of about $5 \times 3 \times 2.5$ m, with a precision of about 0.2 mm. The Prime 13 camera (shown in Fig. 2) is a high speed IR

Fig. 3 Oculus Rift S instrumented with reflective markers

Table 1 Desktop computer and software specifications

CPU	Intel Core i7-7700 @ 3.6 GHz
Memory	16 GB RAM
Video Card	NVidia GeForce GTX 1070
Video Memory	8 GB RAM
Operating system	Windows 10
Mocap software	Motive version 2.0.1
3D Engine	Unity version 2019.4.21

Fig. 4 Main reference frames used for data acquisition, calibration and evaluation. Axes *x*, *y* and *z* are displayed using red, green and blue arrows, respectively

Fig. 5 Data acquisition (top) and processing (bottom) workflow

sensor (Gigabit Ethernet, 240 maximum frame rate) that provides sub-millimeter accuracy, and that has a range of about 12 m. The camera resolution is 1280×1024 (1.3 MP). The OptiTrack system provides on-camera image analysis for detection of marker location, size and roundness, that relieves the CPU from computation of low-level information.

The experimental setup also comprises an Oculus Rift S HMD, instrumented with six passive retro-reflective markers as shown in Fig. 3. The six markers define a single rigid body and are tracked with six degrees of freedom by the OptiTrack system. The Oculus Rift S is a tethered HMD, with a 5-meter cable (with DisplayPort and USB 3.0 connections). A desktop computer running Unity 3D and Motive (the optical motion capture software by OptiTrack) was adopted for data recording and to generate the virtual reality environment. Hardware and software specifications are provided in Table 1.
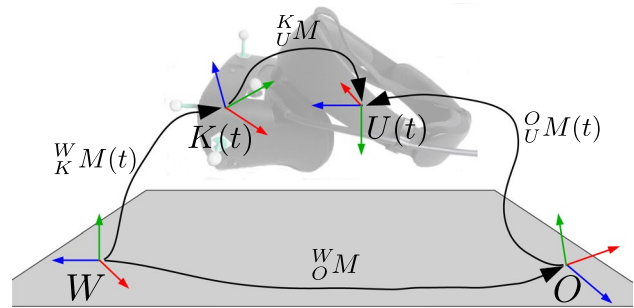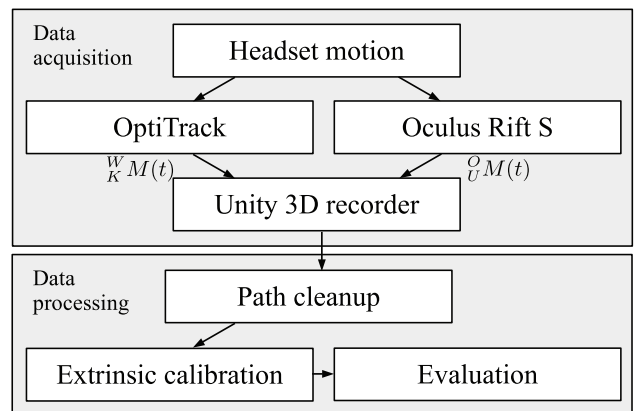
3.2 Data acquisition and processing

Multiple reference frames are defined in the proposed setup as illustrated in Fig. 4. The fixed world reference frame *W* of the OptiTrack motion capture system (also shown in Fig. 1) is located on the floor of the room. Reference frame *W* is known after a one-time calibration phase of the OptiTrack system. Reference frame $K(t)$ is attached to the HMD rigid body and it is tracked by the OptiTrack software (Motive). The position and the orientation of reference frame $K(t)$ with respect to the HMD rigid body are constant over time, and they depend on the configuration of the markers on the headset. Reference frame *O* is the world reference frame of the Oculus Rift S inside-out tracking system. In general, reference frames *W* and *O* are different, moreover, the origin of reference frame *O* may change for each recorded path as it depends on the initial configuration of the HMD. Reference frame $U(t)$ is attached to the HMD rigid body and it is

tracked by the Oculus Rift S tracking system. In particular, reference frame U is located at the midpoint of the user's eyes, with forward (Z-axis) and down (Y-axis) vectors.

Data acquisition and processing was carried out by using a custom Unity 3D script, according to the workflow displayed in Fig. 5. A dataset of HMD paths was recorded by a single user walking around in the room scale environment while wearing the headset. The Unity script, which operates at 60 frames per second, records at each frame t (Unity 3D recorder block in Fig. 5) the 4×4 transformation matrix ${}_{U}^{O}M(t)$ of reference frame $U(t)$ with respect to O as measured by the Oculus Rift S, and the 4×4 transformation matrix ${}_{K}^{W}M(t)$ of reference frame $K(t)$ with respect to W as measured by OptiTrack. Data acquisition in Unity was performed by the Oculus Platform SDK and by the OptiTrack Unity Plugin, respectively.

In order to evaluate the tracking performance of the HMD, recorded paths are processed offline. First, an automated cleanup procedure removes invalid data caused by tracking failures. Then, an extrinsic calibration approach is applied to each path as ${}_{U}^{O}M(t)$ and ${}_{K}^{W}M(t)$ refer to different reference frames. Details about the extrinsic calibration procedure are provided in Sect. 3.3. A description of the acquired dataset is reported in Sect. 3.4. Finally, evaluation was carried out as described in Sect. 3.5.

3.3 Extrinsic calibration

This section describes the extrinsic calibration procedures that are required to evaluate the tracking accuracy of the Oculus Rift S HMD. Since transformations ${}_{U}^{O}M(t)$ and ${}_{K}^{W}M(t)$ track two different reference frames on the HMD a one-time calibration procedure is required to obtain ${}_{U}^{K}M$, i.e., the fixed 4×4 transformation matrix of reference frame $U(t)$ with respect to $K(t)$, as described in Sect. 3.3.1.

3.3.1 Extrinsic calibration between reference frames K and U

As frames $K(t)$ and $U(t)$ are related by a constant transformation ${}_{U}^{K}M$, ${}_{U}^{K}M$ can be estimated by applying an extrinsic calibration algorithm given multiple synchronized samples of ${}_{U}^{O}M(t)$ and ${}_{K}^{W}M(t)$ taken at different poses of the headset. To this purpose a specific calibration path of the HMD was recorded that consists mainly of (in place) rotational movements around multiple axes, as these movements are known to be the most effective for this type of calibration. A set of sampled data $\left\{ {}_{U}^{O}M_c(t), {}_{K}^{W}M_c(t) \right\}$ was then extracted from the calibration path, where subscript c stands for "calibration".

As shown in Fig. 4, the reference frames are related as follows:

$$ {}_{K}^{W}M_c(t) \, {}_{U}^{K}M \, {}_{U}^{O}M_c(t)^{-1} = {}_{O}^{W}M \tag{1} $$

By using (1) for two frames, t and $(t-1)$, an equation in the form of $AX = XB$ is obtained, where:

$$
\begin{aligned}
A &= {}_{K}^{W}M_c(t-1)^{-1} \, {}_{K}^{W}M_c(t) \\
B &= {}_{U}^{O}M_c(t-1)^{-1} \, {}_{U}^{O}M_c(t) \\
X &= {}_{U}^{K}M
\end{aligned} \tag{2}
$$

Equations $AX = XB$ where $A = A_{i-1}^{-1} A_i$ and $B = B_{i-1}^{-1} B_i$ are solved for X given multiple pairs (A_i, B_i) by using the standard formulation by Horaud and Dornaika (1995). To ensure a sufficiently large change in rotation between two consecutive samples, data $A_i = {}_{K}^{W}M_c(t_i)$ and $B_i = {}_{U}^{O}M_c(t_i)$ are sampled from the calibration path whenever the rotation becomes larger than $5°$. That is, t_i is the lowest t so that:

$$
\begin{cases}
\left| \angle \left({}_{K}^{W}M_c(t) \, A_{i-1}^{-1} \right) \right| \geq 5° \\
\left| \angle \left({}_{U}^{O}M_c(t) \, B_{i-1}^{-1} \right) \right| \geq 5°
\end{cases} \tag{3}
$$

where, given a transformation matrix T, operator $\angle(T)$ denotes the rotation angle of the axis-angle representation of the rotation matrix of T.

3.3.2 Extrinsic calibration between reference frames O and W

The transformation matrix ${}_{O}^{W}M$ of reference frame O with respect to W can not be determined in advance for all recorded paths used for the experimental evaluation, as the initial configuration of reference frame O may potentially change for each recorded path. In this work two different approaches are compared to calibrate the transformation between reference frames W and O for each single path. The two calibration methods are based on the alignment of paths ${}_{U}^{W}M(t)$ and ${}_{U}^{O}M(t)$. The first approach is named Single State (SS) alignment, while the second approach is named Multiple States (MS) alignment, as in Zhang and Scaramuzza (2018).

The Single State alignment method exploits only the configuration of the HMD reference frame at the beginning of the path, i.e., when the tracking drift is not present. Given initial transformation ${}_{U}^{W}M(0) = {}_{U_0}^{W}M$ of the headset as measured by the motion capture system, and the initial transformation ${}_{U}^{O}M(0) = {}_{U_0}^{O}M$ as measured by the Oculus Rift S, the Single State alignment transformation ${}_{O}^{W}M_{ss}$ is computed as:

$$ {}_{O}^{W}M_{ss} = {}_{U_0}^{W}M \, {}_{U_0}^{O}M^{-1} \tag{4} $$

In the Multiple States alignment method the translation component ${}_{U}^{W}\mathcal{T}(t)$ of the entire path ${}_{U}^{W}M(t)$ is used for the alignment, as well as the translation component ${}_{U}^{O}\mathcal{T}(t)$ of of the entire path ${}_{U}^{O}M(t)$. In particular, the Multiple States alignment transformation ${}_{O}^{W}M_{ms}$ is obtained by minimizing the mean square distance between ${}_{U}^{W}\mathcal{T}(t)$ and ${}_{U}^{O}\mathcal{T}(t)$:
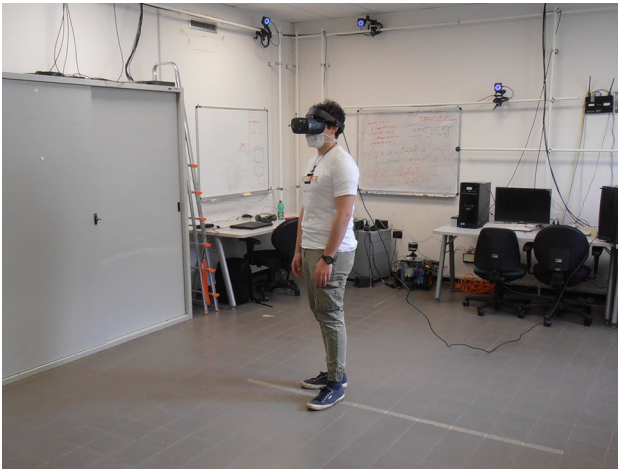
Fig. 6 An image of the user wearing the HMD while recording the dataset

Fig. 7 The 3D virtual reconstruction of the room

$$
{}_{O}^{W}M_{\mathrm{ms}} = \underset{M}{\arg\min} \sum_{t} \left\| {}_{U}^{W}\mathcal{T}(t) - M \, {}_{U}^{O}\mathcal{T}(t) \right\|^{2} \tag{5}
$$

3.4 Dataset

The experimental evaluation was conducted on a custom dataset containing a set of recorded HMD paths of a user walking around the environment (Fig. 6). For the dataset acquisition the user wore the HMD that displayed a 3D virtual reconstruction of the room (Fig. 7). The user was free to rotate his head around during the experiments. Therefore, each recorded path of the dataset contains simultaneous changes in both position and orientation of the HMD. The dataset contains a total of 85 paths, organized in five subsets of paths as follows:

– *Line*: 20 straight line segments with different directions, traveled back and forth, with average length of about 3 m and average duration of 96 s.
– *Circle*: 20 closed paths with circular shape, with average radius of about 1.5 m and average duration of 116 s.
– *Eight*: 20 "8-shaped" paths, covering a surface of about 4 × 2 m, with average duration of 109 s.
– *Dynamic*: 20 closed paths with circular shape, similar to the *Circle* subset, in a dynamic environment with two people walking closely beside the user. Each person carried a vertical square panel (side 85 cm) at head level to increase clutter.
– *Random*: 5 random walking paths, with average duration of 89 s.

It must be noticed that the OptiTrack system may lose tracking of the HMD for a few frames in certain conditions. For example when the user walks close to the corners of the room or when the HMD is occluded. In these cases invalid measurements were discarded and excluded from the evaluation (Path cleanup block in Fig. 5).

3.5 Evaluation

This section describes the evaluation metrics that have been used to assess the tracking accuracy. Data analysis was performed by computing both translation and orientation errors. Transformation ${}_{U}^{W}M(t) = {}_{O}^{W}M \, {}_{U}^{O}M(t)$ which represents the pose of the HMD U with respect to the world reference frame W as measured by the Oculus Rift S built-in tracking system, is compared with the ground truth pose ${}_{U}^{W}M'(t) = {}_{K}^{W}M(t) \, {}_{U}^{K}M$ from the OptiTrack system. The absolute translation error $dT(t)$ for each sample at time t was computed as the Euclidean distance between the translation vectors ${}_{U}^{W}\mathcal{T}(t)$ and ${}_{U}^{W}\mathcal{T}'(t)$ of ${}_{U}^{W}M(t)$ and ${}_{U}^{W}M'(t)$, respectively:

$$
dT(t) = \left\| {}_{U}^{W}\mathcal{T}(t) - {}_{U}^{W}\mathcal{T}'(t) \right\| \tag{6}
$$

The absolute rotation error $dR(t)$ for each sample at time t was computed as the rotation angle of the axis-angle representation of rotation matrix ${}_{U}^{W}\mathcal{R}(t)^{\top} \, {}_{U}^{W}\mathcal{R}'(t)$, i.e.,

$$
dR(t) = \left| \angle\left({}_{U}^{W}\mathcal{R}(t)^{\top} \, {}_{U}^{W}\mathcal{R}'(t) \right) \right| \tag{7}
$$

4 Experimental results

The translation error $dT(t)$ and the rotation error $dR(t)$, averaged over each subset of paths and over the complete dataset, are reported in Tables 2 and 3, respectively. Table 2 and

Table 2 Mean, standard deviation (SD) and maximum (Max) translation error (cm), for each subset of paths and over all paths

	SS alignment			MS alignment		
	Mean	SD	Max	Mean	SD	Max
Line	0.65	0.42	2.25	0.26	0.15	2.20
Circle	1.97	1.11	4.92	1.21	0.31	3.40
Eight	1.46	0.77	5.35	0.88	0.48	3.62
Dynamic	2.89	1.74	8.10	1.62	0.61	3.52
Random	3.31	1.36	8.26	2.58	1.14	7.85
All	1.83	2.30	8.26	1.12	1.39	7.85

Table 3 Mean, standard deviation (SD) and maximum (Max) rotation error (degrees), for each subset of paths and over all paths

	SS alignment			MS alignment		
	Mean	SD	Max	Mean	SD	Max
Line	0.38	0.27	6.51	0.37	0.26	6.36
Circle	0.78	0.51	6.90	0.57	0.50	6.76
Eight	0.64	0.43	6.01	0.50	0.42	5.89
Dynamic	0.78	0.54	6.17	0.68	0.48	5.62
Random	2.03	1.57	15.38	2.05	1.56	15.35
All	0.77	1.08	15.38	0.66	1.01	15.35
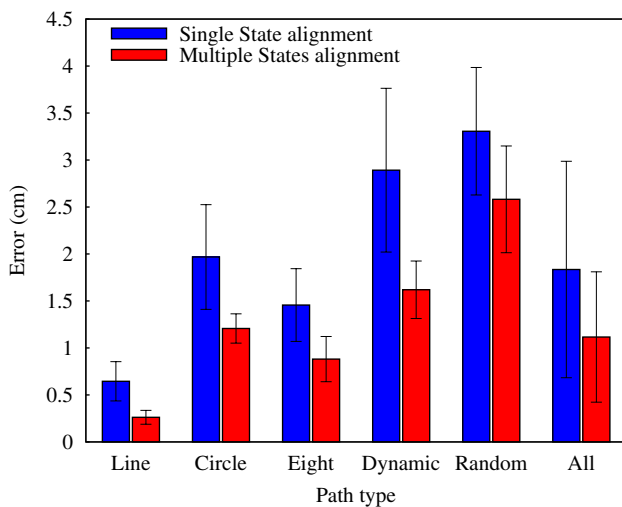
Fig. 8 Average translation error $dT(t)$ for the five subsets of paths (*Line*, *Circle*, *Eight*, *Dynamic*, *Random*) and for the complete dataset (All). Results are shown for the two alignment methods: Single State alignment and Multiple States alignment. The vertical error bars show the standard deviation
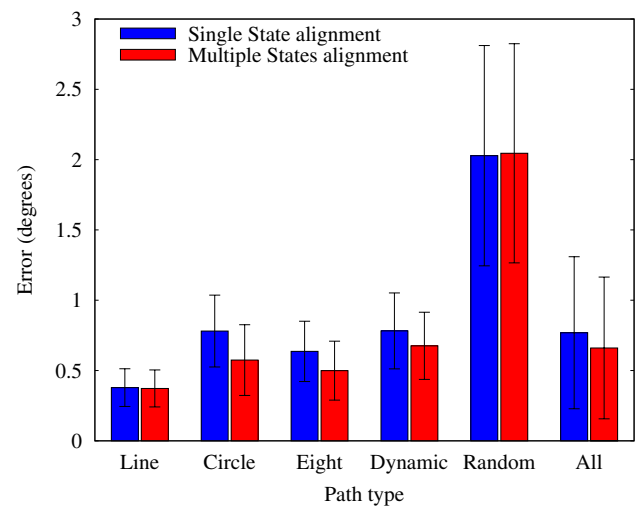
Fig. 9 Average rotation error $dR(t)$ for the five subsets of paths (*Line*, *Circle*, *Eight*, *Dynamic*, *Random*) and for the complete dataset (All). Results are shown for two alignment methods: Single State alignment and Multiple States alignment. The vertical error bars show the standard deviation

Table 3 also report the standard deviation and the maximum error. Data are also illustrated in Fig. 8 and in Fig. 9. The average error computed on the whole dataset is about 1.83 cm and 0.77° (SS alignment method), and 1.12 cm and 0.66° (MS alignment method). The lowest error was obtained for the *Line* paths, due to their simple shape. Conversely, the more complex paths in the *Random* subset have an average error which is significantly higher than all other path types. The average error of *Circle* and *Eight* paths, that have

an intermediate complexity, is contained between the average error of *Line* paths and *Random* paths. The *Eight* paths have a slightly lower error than *Circle* paths, possibly due to the longer average duration of *Circle* paths compared to *Eight* paths (116 s and 109 s as reported in Sect. 3.4). The error of *Dynamic* paths is slightly higher than the error for *Circle* paths. Therefore, it can be observed that the Oculus Rift S native tracking system is rather robust to dynamic environments.
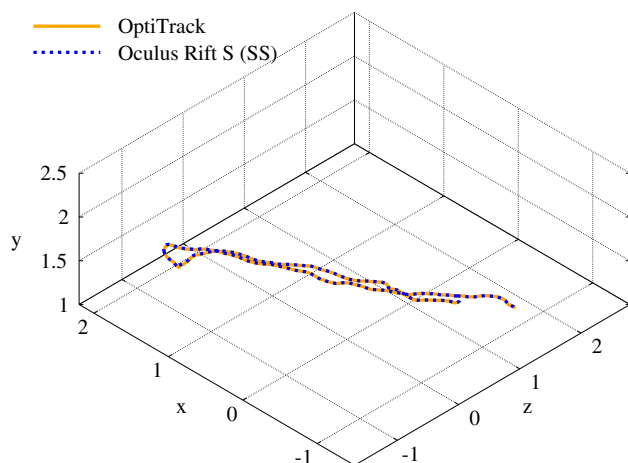
Fig. 10 Example *Line* path tracked by the motion capture system and the corresponding SS-aligned path tracked by the Oculus Rift S
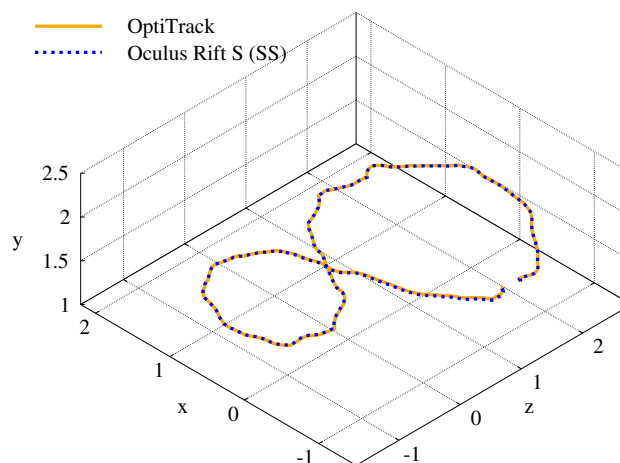

Fig. 12 Example *Eight* path tracked by the motion capture system and the corresponding SS-aligned path tracked by the Oculus Rift S
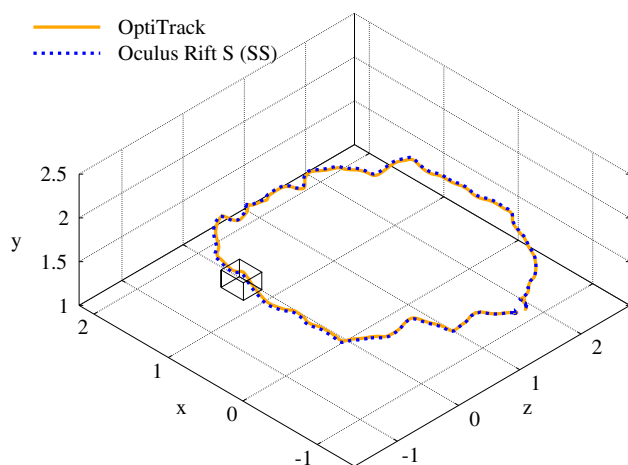

Fig. 11 Example *Circle* path tracked by the motion capture system and the corresponding SS-aligned path tracked by the Oculus Rift S. The region in the black cube is shown enlarged in Fig. 14
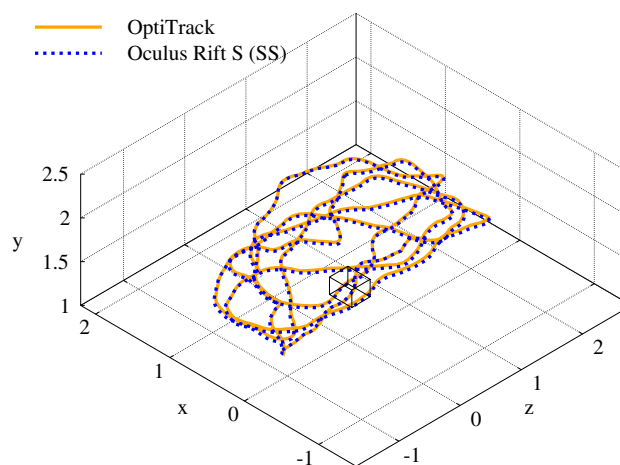

Fig. 13 Example *Random* path tracked by the motion capture system and the corresponding SS-aligned path tracked by the Oculus Rift S. The region in the black cube is shown enlarged in Figure 15

In general, the average translation error is lower for the Multiple States (MS) calibration method, compared to the Single State (SS) method. This result is expected, since calibration using the MS alignment method minimizes the distance between all corresponding points in ${}^{W}_{U}\mathcal{T}(t)$ and ${}^{W}_{U}\mathcal{T}'(t)$. The average rotation error of Single State and Multiple States alignment methods is comparable, since the MS method considers only the position distance between corresponding points. It must be observed that the Multiple States alignment approach requires the path to be known in advance, while the Single State alignment approach only requires the first frame. Hence, the use of the Single State alignment method is more suitable for online virtual reality applications.

Example paths from the dataset, tracked by the Oculus Rift S and by the motion capture system, are shown in Figs. 10, 11, 12 and 13. Enlarged views of some example paths are displayed in Fig. 14 and in Fig. 15. As expected, the Oculus Rift S path obtained through MS alignment is closer to the ground truth OptiTrack path than the SS-aligned path.

The translation and rotation errors over time for the *Circle* path in Fig. 11 and the *Random* path in Fig. 13 are shown in Figs. 16 and 17, respectively. In the *Circle* path, the average translation error is 1.55 cm for the Single State alignment method. The translation error obtained by the Single State alignment approach increases at the beginning of the path when the user moves away from the starting position, and it decreases near the end of the
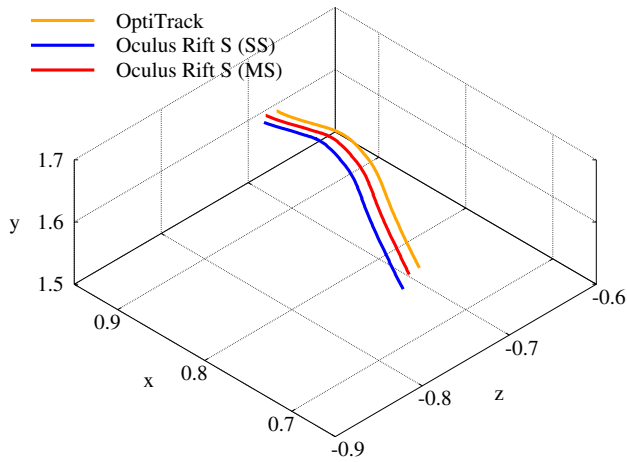
Fig. 14 Enlarged view of a tracked path from Figure 11 (including also the tracked path obtained from the Multiple States calibration)
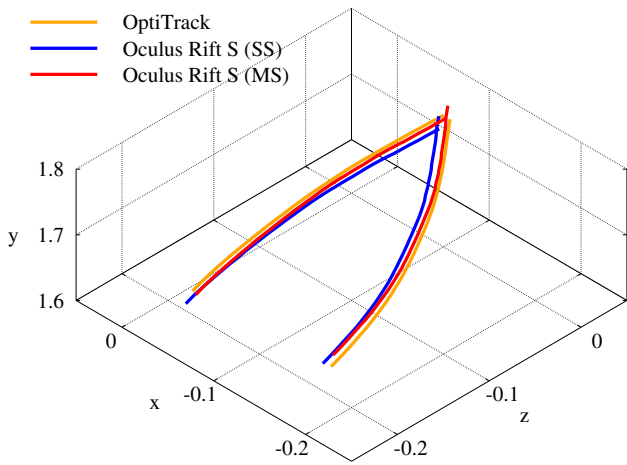

Fig. 15 Enlarged view of a tracked path from Figure 13 (including also the tracked path obtained from the Multiple States calibration)

Fig. 16 Tracking error for the *Circle* path in Fig. 11
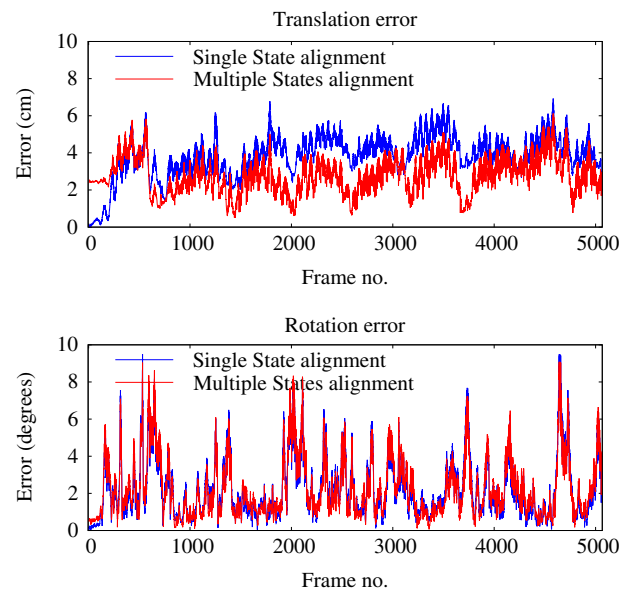
Fig. 17 Tracking error for the *Random* path in Fig. 13

path, when the user comes back to the initial position, thus suggesting a not negligible error in the estimated rotation component of ${}^{O}_{W}M_{ss}$. Conversely, the translation error obtained by the Multiple States alignment method is rather constant, about 1.08 cm on average, thus suggesting that the Multiple States alignment method provides a better calibration of the reference frames. In the *Circle* path, the average rotation error is about 0.52° for the Single State alignment, and 0.37° for the Multiple States alignment. In the *Random* path, the average translation error is about 2.90 cm, and the average rotation error is 2.3° (with MS alignment), which are significantly larger than in the *Circle* path. The larger errors in the *Random* path are due to the more complex shape of the path that includes frequent changes in motion direction and speed.
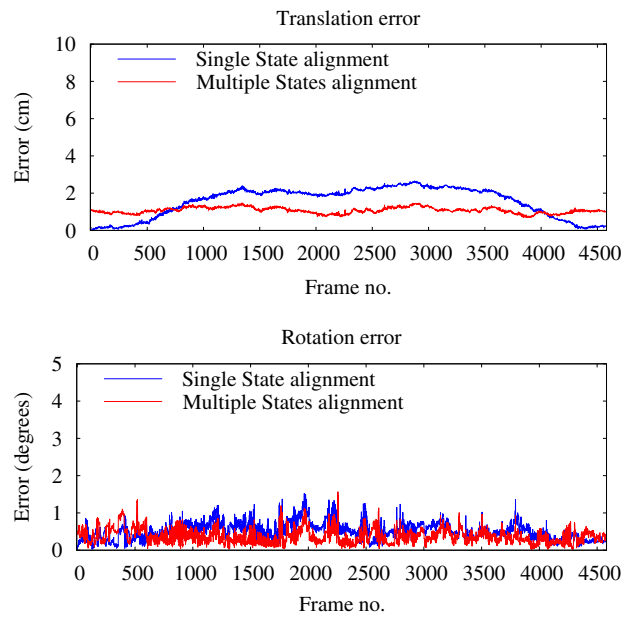
Repeatability in calibration between reference frames *K* and *U* (Sect. 3.3.1) has been assessed by rerunning the calibration procedure on 20 different calibration paths of the headset. The results indicate that the standard deviation of the translation is about 0.24 cm, whereas the standard deviation of the rotation angle in the axis-angle representation is about 0.44 degrees.

5 Conclusions

This work investigated the tracking accuracy of the Oculus Rift S HMD in room scale environments. The built-in tracking algorithm of the Oculus Rift S was compared to the performance that can be achieved by using an Opti-Track motion capture system. The results show that, in room-scale environments, the translation and rotation accuracy of the built-in HMD tracking system is about 1.83 cm and 0.77° on average. Therefore, it may be concluded that in most virtual reality applications the inside-out tracking system of the Oculus Rift S is more than adequate, however, for specific virtual reality tasks requiring high quality tracking it may be advisable to replace the built-in tracking system of the Oculus Rift S with a third party solution. Moreover, it can be observed that the proposed method to evaluate the accuracy of the Oculus Rift S tracking system is general and it can be applied to other HMDs. Future work will investigate more robust tracking algorithms by combining data from the motion capture equipment and from the HMD built-in tracking system.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

References

Ameler T, Blohme K, Brandt L, Brüngel R, Hensel A, Huber L, Kuper F, Swoboda J, Warnecke M, Warzecha M, Heß D, Frömke J, Schmitz-Stolbrink A, Friedrich CM (2019) A comparative evaluation of steamvr tracking and the optitrack system for medical device tracking. In: Proceedings of the 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC), pp 1465–1470

Bauer P, Lienhart W, Jost S (2021) Accuracy investigation of the pose determination of a VR system. Sensors 21(5):1622

Borges M, Symington A, Coltin B, Smith T, Ventura R (2018) HTC Vive: analysis and accuracy improvement. In: Proceedings of the IEEE/RSJ international conference on intelligent robots and systems (IROS), pp 2610–2615

Borrego A, Latorre J, Alcañiz M, Llorens R (2018) Comparison of oculus rift and HTC Vive: feasibility for virtual reality-based exploration, navigation, exergaming, and rehabilitation. Games Health J 7(3):151–156

Caserman P, Garcia-Agundez A, Konrad R, Göbel S, Steinmetz R (2019) Real-time body tracking in virtual reality using a Vive tracker. Virtual Real 23(2):155–168

Chang CM, Hsu CH, Hsu CF, Chen KT (2016) Performance measurements of virtual reality systems: quantifying the timing and positioning accuracy. In: Proceedings of the 24th ACM international conference on multimedia, pp 655–659

Chessa M, Maiello G, Borsari A, Bex PJ (2019) The perceptual quality of the oculus rift for immersive virtual reality. Hum-Comput Interact 34(1):51–82

Debarba HG, de Oliveira ME, Lëdermann A, Chaqué S, Charbonnier C (2018) Tracking a consumer hmd with a third party motion capture system. In: Proceedings of the IEEE conference on virtual reality and 3D user interfaces (VR), pp 539–540

Flueratoru L, Simona Lohan E, Nurmi J, Niculescu D (2020) HTC Vive as a ground-truth system for anchor-based indoor localization. In: Proceedings of the 12th international congress on ultra modern telecommunications and control systems and workshops (ICUMT), pp 214–221

Greiff M, Robertsson A, Berntorp K (2019) Performance bounds in positioning with the VIVE lighthouse system. In: Proceedings of the 22th international conference on information fusion (FUSION), pp 1–8

Groves LA, Carnahan P, Allen DR, Adam R, Peters TM, Chen ECS (2019) Accuracy assessment for the co-registration between optical and VIVE head-mounted display tracking. Int J Comput Assist Radiol Surg 14(7):1207–1215

Guaitolini M, Petros FE, Prado A, Sabatini AM, Agrawal SK (2021) Evaluating the accuracy of virtual reality trackers for computing spatiotemporal gait parameters. Sensors 21(10):3325

Horaud R, Dornaika F (1995) Hand-eye calibration. Int J Robot Res 14(3):195–210

Ikbal MS, Ramadoss V, Zoppi M (2021) Dynamic pose tracking performance evaluation of HTC Vive virtual reality system. IEEE Access 9:3798–3815

Jansen W, Laurijssen D, Daems W, Steckel J (2019) Automatic calibration of a six-degrees-of-freedom pose estimation system. IEEE Sens J 19(19):8824–8831

Jost TA, Drewelow G, Koziol S, Rylander J (2019) A quantitative method for evaluation of 6 degree of freedom virtual reality systems. J Biomech 97:109379

Jost TA, Nelson B, Rylander J (2021) Quantitative analysis of the oculus rift S in controlled movement. Disabil Rehabil Assist Technol 16(6):632–636

Lubetzky AV, Wang Z, Krasovsky T (2019) Head mounted displays for capturing head kinematics in postural tasks. J Biomech 86:175–182

Lwowski J, Majumdat A, Benavidez P, Prevost JJ, Jamshidi M (2020) HTC Vive tracker: accuracy for indoor localization. IEEE Syst, Man, Cybern Mag 6(4):15–22

Marchetto J, Wright WG (2019) The validity of an oculus rift to assess postural changes during balance tasks. Hum Factors 61(8):1340–1352

Merriaux P, Dupuis Y, Boutteau R, Vasseur P, Savatier X (2017) A study of vicon system positioning performance. Sensors 17(7):1591

Niehorster DC, Li L, Lappe M (2017) The accuracy and precision of position and orientation tracking in the HTC Vive virtual reality system for scientific research. i-Perception 8(3):1–23

Palma G, Perry S, Cignoni P (2021) Augmented virtuality using touch-sensitive 3D-printed objects. Remote Sens 13(11):2186

Rolland J, Baillot Y, Goon A (1999) A survey of tracking technology for virtual environments. In: Fundamentals of Wearable Computers and Augmented Reality (chap. 3), CRC Press, pp 110–127

Spitzley KA, Karduna AR (2019) Feasibility of using a fully immersive virtual reality system for kinematic data collection. J Biomech 87:172–176

Suznjevic M, Mandurov M, Matijasevic M (2017) Performance and QoE assessment of HTC Vive and Oculus Rift for pick-and-place tasks in VR. In: Proceedings of the 9th international conference on quality of multimedia experience (QoMEX), pp 1–3

van der Veen SM, Bordeleau M, Pidcoe PE, France CR, Thomas JS (2019) Agreement analysis between vive and vicon systems to monitor lumbar postural changes. Sensors 19(17):3632

Vox JP, Weber A, Wolf KI, Izdebski K, Schüler T, König P, Wallhoff F, Friemert D (2021) An evaluation of motion trackers with virtual reality sensor technology in comparison to a marker-based motion capture system based on joint angles for ergonomic risk assessment. Sensors 21(9):3145

Wu R, Pandurangaiah J, Blankenship GM, Castro CX, Guan S, Ju A, Zhu Z (2020) Evaluation of virtual reality tracking performance for indoor navigation. In: Proceedings of the IEEE/ION position, location and navigation symposium (PLANS), pp 1311–1316

Xu X, Chen KB, Lin JH, Radwin RG (2015) The accuracy of the oculus rift virtual reality head-mounted display during cervical spine mobility measurement. J Biomech 48(4):721–724

Zhang Z, Scaramuzza D (2018) A tutorial on quantitative trajectory evaluation for visual(-inertial) odometry. In: Proceedings of the IEEE/RSJ international conference on intelligent robots and systems (IROS), pp 7244–7251