# Insert Your Own Body in the Oculus Rift to Improve Proprioception

Manuela Chessa, Lorenzo Caroggio, Huayi Huang and Fabio Solari
*Dept. of Informatics, Bioengineering, Robotics, and Systems Engineering,*
*University of Genoa, Viale Causa 13, Genoa, Italy*

Keywords: Head-mounted-displays, Microsoft Kinect, Leap Motion, Virtual Reality, Registration and Calibration, Self-perception.

Abstract: A natural interaction in virtual reality environments, in particular when wearing head-mounted-displays, is often prevented by the lack of a visual feedback about the user's own body. This paper aims to create a virtual environment, in which the user can visually perceive his/her own body, and can interact with the virtual objects, by using a virtual body that replicates his/her movements. To this aim, we have set up an affordable virtual reality system, which combines the Oculus Rift head-mounted-display, a Microsoft Kinect, and a Leap Motion, in order to recreate inside the virtual environments a first-person avatar, who replicates the movements of the user's full-body and the fine movements of his/her fingers and hands. By acting in such an environment, the user is able to perceive him/herself thus improving his/her experience in the virtual reality. Here, we address and propose a solution to the issues related to the integration of the different devices, and to the alignment and registration of their reference systems. Finally, the effectiveness of the proposed system is assessed through an experimental session, in which several users report their feeling by answering to a 5-points Likert scale questionnaire.

## 1 INTRODUCTION AND RELATED WORKS

The main motivation of the present work is to overcome the lack of presence of the user own body in virtual reality (VR) systems, in particular when wearing a head-mounted-display (HMD). The aim of the paper is to create a virtual environment in which the user can visually perceive his/her own body, and can interact with the virtual objects by using a body that reproduces his/her movements. In particular, the focus is on having a realistic and natural interaction through all the body and, in addition, a fine interaction with the hands. VR systems are commonly able to elicit a strong sense of presence, allowing the user to perform several uncommon operations in a safe environment: for example, in medical applications, in (Ahlberg et al., 2002) it was observed that VR simulation was able to predict the surgical outcome; also in (Seymour et al., 2002) it was demonstrated the effectiveness of the VR training in improving the operation room performances.

The use of the hands in a virtual environment, in order to have a better feeling of presence, was discussed in (Tecchia et al., 2014) with the aim of training the user, and in (Beattie et al., 2015) to use CAD packages. The outcomes of such works consist in a more natural way of interacting with the environments, and in the creation of more immersive virtual systems.

However, the impossibility of having the full control and a natural perception of own virtual body often gives a sense of unnaturalness to the user actions and decreases the effectiveness of transferring the skills acquired in the virtual environment to the real operating environment. In fact, one of the most important differences of performing an experience in a virtual context or in a real one, is related to the self-body visual perception. For example, in a CAVE system (Creagh, 2003), users can still see their body: this has, of course, a great impact on the experience, especially in tasks, where the interaction between the user and the virtual environment is required. Avatar representation can be used, but it rarely correspond exactly to the dimensions or current posture of the user.

In the literature, some researchers have used the Kinect in order to animate an avatar inside a VR environment for the Oculus Rift (Lee et al., 2015), but this work has not the aim of recreating the self-perception of the HMD user. Recently, in (Sra and Schmandt, 2015) the authors achieved the tracking of the users body with a Kinect device such that their physical



Chessa, M., Caroggio, L., Huang, H. and Solari, F.
Insert Your Own Body in the Oculus Rift to Improve Proprioception.
DOI: 10.5220/0005851807550762
In *Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2016) - Volume 4: VISAPP*, pages 755-762
ISBN: 978-989-758-175-5

movements are mirrored in the virtual world, in a collaborative environment. Users can see their own avatar and the other person's avatar allowing them to perceive and act intuitively in the virtual environment.

In order to address the proprioception of own body in VR, we propose to fuse the data acquired by an RGB-D camera and by a low-price hand-tracking device, to reconstruct an accurate avatar, which moves in a coherent way with the user. Some authors have recently addressed the problem of fusing information acquired by the RGB-D devices and the Leap Motion, e.g. in (Penelle and Debeir, 2014) the authors create an augmented reality system to be used with amputees patients, and in (Ahmed et al., 2014), where an approach to fuse Kinect range images and Leap Motion data for immersive augmented reality applications is described.

The aim of our work is to set up an affordable VR system, which combines the Oculus Rift HMD, a RGB-D device, i.e. the Microsoft Kinect, and a Leap Motion, in order to allow a user to perceive his/her own body inside the VR environments also by taking into account fine details, such as the fingers' movements, thus bridging the gap between the HMD systems and the CAVE system.

## 2 MATERIAL AND METHODS

The proposed system has been designed using different sensors to track of the entire body of the user. The depth and color images from the Microsoft Kinect [1] provide a quite robust tracking of the users movements and at the same time the Leap Motion [2] provides a more precise tracking of the users hands and fingers position. We have also developed a calibration method that computes the rigid transformation to align the two different frames of references expressed from the two sensors. After the calibration, the data from the Kinect and the Leap Motion (LM) are fused together and used to control a first-person 3D avatar, which is showed inside a virtual environment, by using as HMD the Oculus Rift [3] (OR). The virtual environment has been created by using Unity3D 5 [4], which also makes the assets for acquiring data from Kinect and LM available in its store. The data fusion gives some advantages: more accurate tracking of the users hands and fingers, and extension of the tracking range by providing hand locations when they are not visible to the Kinect.

[1]https://dev.windows.com/en-us/kinect
[2]https://developer.leapmotion.com/
[3]https://developer.oculus.com/
[4]https://unity3d.com/5

### 2.1 Description of the System Setup

Figure 1 shows an overview of the VR system described in the paper. The user stands in front of the Kinect at a distance of about 1.5m wearing the OR. The Leap Motion is attached on the OR, and the Kinect is located in a table. The user is free to move over an area of about 1.5$m^2$ in order to interact with the 3D objects in the virtual environment. The area where the objects appear is set to be in the reaching range for the entire users body.

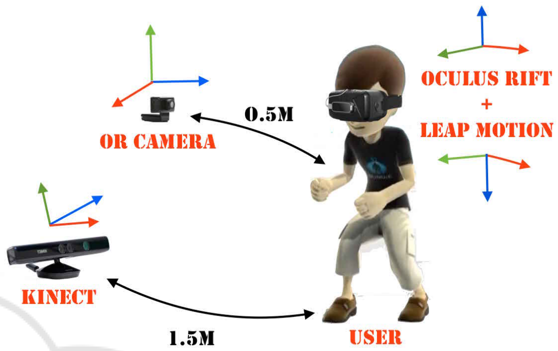


Figure 1: Overview of the proposed system.

The device used for Virtual Reality is the Oculus Rift DK2: its lenses distortion allows the user to have a really immersive experience with an estimated field of view of 100 degrees. The device has a resolution of 2160x1200 pixels and connected via HDMI 1.3 and via USB2.0. The OR is very comfortable with a high frame rate of about 90Hz and contains several sensors such as the accelerometer, the gyroscope and magnetometer, which are used to track user position and movement while wearing it. It is also present an external camera, the OR camera, which is available to track the OR with an update rate of 60Hz and estimate its location in a Cartesian coordinates system centered on the camera.

To make the system works fine, it requires a slightly powerful computer and a large bandwidth to manage the data flow from the two sensors simultaneously. The used machine specification are: a PC with Mother Board Asrock Z77 Extreme4, equipped with graphic card NVIDIA GeForce CTX 960, processor Intel Core i5 2500k 3.30 GHz, 8GB of RAM, operating system Windows 10 Pro 64bit.

### 2.2 Body Tracking

The data stream for tracking the user is acquired by the Kinect, through a synchronization of depth and color images with a resolution of 640x480 pixels at a frame rate of 30Hz. The images are analyzed by the Kinect MS-SDK Assets, available in the Unity asset Store, which uses the Kinect Runtime provided by Microsoft to make the tracking information suitable to move an avatar. The asset gives information on the tracking of twenty joints of the users body which are then aligned with those acquired by the Leap Motion (more details on the alignment procedures are in Section 2.4).

## 2.3 Hand Tracking

The accurate tracking of the users hands and fingers are performed by the Leap Motion, which is a small sensor connected via USB 2.0 and placed on the Oculus Rift support in front the users eyes. It is composed of two wide-angle infrared CCD cameras with an acquisition frequency of about 120 Hz and the detection field is approximately a hemisphere of 0.5m radius above the sensor.

The technology behind the Leap Motion has not be released by the manufacturer yet but the accuracy announced for the position detection of the fingertips is approximately 0.01mm. For more details, in (Weichert et al., 2013) the authors have conducted a study on the accuracy and robustness of the Leap Motion. The acquired data are elaborated by the Leap Motion SDK, which makes available the locations of the center of the palm and of each single finger. The information is expressed in a Cartesian coordinates system centered on the device.



## 2.4 Calibration

Before the calibration step the data acquired by the two sensors are used to produce a virtual avatar as in Figure 2.

To align the data of all sensors, the calibration phase is done in two steps:

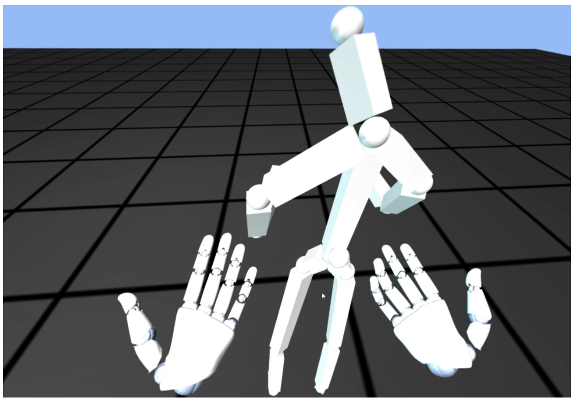

Figure 2: Own avatar inside the VR environment by using uncalibrated data.

- Rigid transformation between common points acquired by the Kinect and the Leap Motion, computed just once.
- Live corrections to overcome the residual offset present between the Kinect and the Leap Motion tracking, and management of the head position over time, computed every frame.

Thanks to the Leap Motion Core Assets plug-in for Unity3D, the Leap Motion and the Oculus Rift are already very well aligned: the two cameras of the Leap Motion allow us to use the Oculus-Leap system as a pass-through head mounted device, and to verify the correctness of the alignment embedded in the plug-in.

The data obtained by the Kinect is properly low pass filtered, in order to remove most of the noise, maintaining realistic movements of the body.

### 2.4.1 Rigid Transformation

To compute the rigid transformation, we use the “least-square rigid motion using SVD” technique (Sorkine, 2009), implemented in C# language, in order to have a fast and reliable implementation of this technique in the Unity3D environment. The method can be formulated in the following way:

$$(\bar{R},\bar{\mathbf{t}}) = \underset{R,\mathbf{t}}{\operatorname{argmin}} \sum_{i=1}^{n} \|(R\mathbf{p_i} + \mathbf{t}) - \mathbf{q_i})\|^2,$$

where:

- $R$ is the rotation matrix between the two sets of points (and $\bar{R}$ the computed estimate) ;
- $\mathbf{t}$ is the translation vector between the centers of mass of the two sets of points (and $\bar{\mathbf{t}}$ the computed estimate);
- $P = \{\mathbf{p_1}, \mathbf{p_2}, ..., \mathbf{p_n}\}$ are the Leap Motion samples;
- $Q = \{\mathbf{q_1}, \mathbf{q_2}, ..., \mathbf{q_n}\}$ are the Kinect samples;
- $n$ is the number of samples.

Such a method requires several correspondences between the two systems to align them; in our case, we have some common joints tracked by both the Kinect and the Leap Motion: the centers of the hands, the wrists, the elbows.

The result of the rigid transformation, done on a single set of samples, often leads to have an alignment visually incorrect; this is due to multiple factors:

- possible coplanar structure of the points: hands, wrists, elbows almost on the same plane;
- noise that affects the Kinect joints: in particular hands and wrists;
- frequent mismatch between the centers of the hands, due to the noise and to the worse accuracy of the Kinect, with respect to the Leap Motion.

To overcome this problem, we have taken in consideration the use of less joints and more samples in time. We decided to remove the centers of the hands from the list of matching points (so we take in consideration 4 points for each frame, the two wrists and the two elbows) and 500 samples over time: during the calibration period, the user have to move his/her arms a little bit, paying attention that both are tracked by the two sensors.

In order to maintain the correct distance between the Oculus Rift point of view (POV) and the hands, as well as its correct rotation, looking at the hands during the phase of rigid transformation, we set the Oculus child-node to the middle point of the Leap Motion matching points in the Unity3D environment: considering that, as explained in (Sorkine, 2009), the rotation is done around the middle point, this ensures to maintain a correct behavior of the POV.

After that, the visual result of the rigid transformation is far better than before, in particular for what concerns the legs (see Figure 3). The Kinect and the Leap Motion hands, however, still do not perfectly overlap over time, because of the different precision and the position of the two sensors.

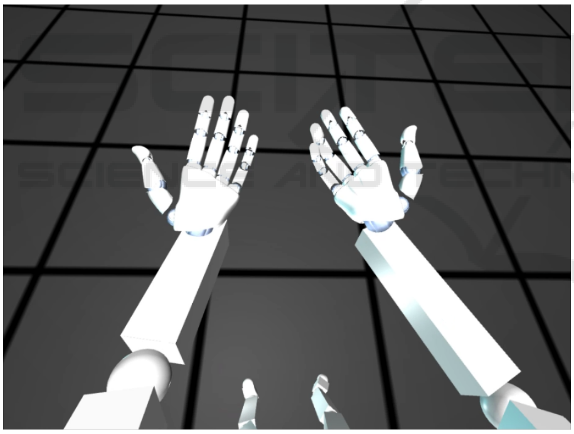

Figure 3: Own avatar inside the VR environment, after the calibration and the rigid trasformation.

#### 2.4.2 Live Corrections

In order to have a unique body structure, which shows a continuity between the arms tracked by the Leap Motion and the rest of the body tracked by the Kinect, we have to better fuse the data from the two sensors. In particular, we choose to use only the positions provided by the Leap Motion, because of its better precision, when the arms are tracked by both sensors, and use the data obtained by the Kinect, when the hands are out of the range of the Leap Motion.

Thanks to the previous rigid transformation, the legs are visually correct, even if they are not so aligned as the hands. In second place, we manage the movements of the point of view of the user in a way that is consistent with the movements of the head, to let the user move around having his/her virtual body, which moves in a coherent way with his/her real movements. There are two procedures to do this:

- Link the Oculus POV to the head joint tracked by the Kinect: this solution allows the user to move in all the area in which the Kinect is able to track the user, but, because of the noise that affects the measurement, the VR camera tends to suffer of small oscillations, that have a bad impact on the general sense of immersion of the user;
- Use the Oculus Positional Tracker (i.e. the OR camera): this is the best solution in term of realism of the movements of the head, but the area is limited to the range and the direction of the Oculus Positional Tracker.

In order to obtain the most realistic solution, we use the second solution; the position computed by the Oculus Positional Tracker is affected by a little drift-offset respect to the position obtained from the Kinect, but it seems not to be relevant in the truthfulness of the body perception (see Figure 4).

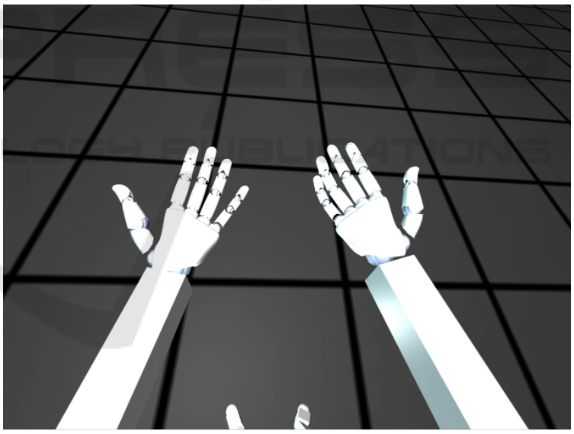

Figure 4: Own avatar inside the VR environment: final result after calibration process.

Finally, Figure 5 shows the system work flow: at each start we have to perform the calibration phase, during which the system collects the data for the rigid transformation, only if they are available from both the Kinect and the Leap Motion. Once the system is calibrated, we apply live corrections to use the best data and animate correctly the hands, if the Leap Motion is tracking them. However, as explained before, there is the possibility to go out of range of the Oculus camera: in this case, a warning message invites the user to step back into the tracking range.

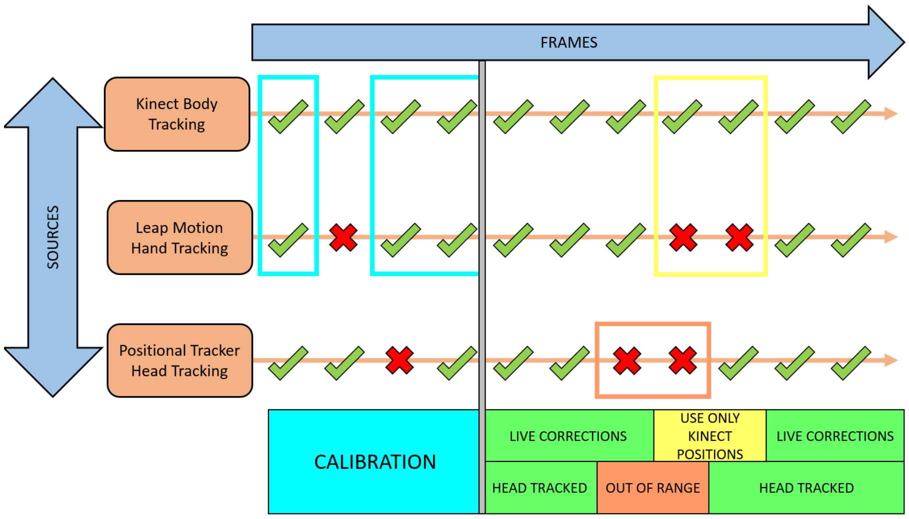


Figure 5: The system work flow: (i) at each start a calibration is performed, during which the system collects the data for the rigid transformation (only if they are available from both the Kinect and the Leap Motion); (ii) once the system is calibrated, live corrections are operated to use the best data and correctly animate the hands (if the Leap Motion is tracking them).

### 2.4.3 Adding a Mesh to the Avatar

With the aim of giving an even more realistic sense of presence in the VR, we added a human mesh to the scene as the body of the user (see Figure 6), but see (Lugrin et al., 2015).

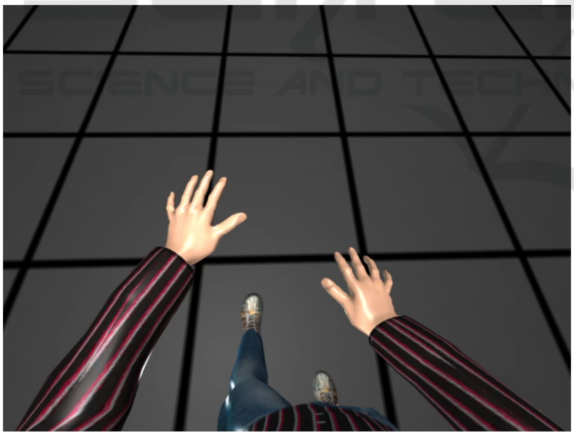

Figure 6: Adding a human mesh to the tracked skeleton.

We animated the mesh taking inspirations from the skeleton we reconstructed and scaling in a proper way to fit the user, but several problems came out. The main issue consists in the different way of animating the skeleton and the mesh. The former is animated by giving to the joints the positions taken by the Kinect and the Leap Motion and by connecting them with simple structures, like parallelepipeds.The latter is animated by giving to each bones of the skeleton linked with the mesh a correct rotation. So, the challenge was to calculate, from the first skeleton we made, the correct angles to be given to the second skeleton in order to obtain the same movements. This work required a proper mapping of the movement of the first skeleton to a set of rotations to apply to the second one, in relation to the starting position of the mesh, in particular the initial rotation of every bone of the second skeleton. Another problem involves the structure of the second skeleton and the noise which affects the Kinect data: each bones is nested with the previous one. In this way, the noise that affects a single joint, propagates to the following, altering the final result. The final result is a more realistic body, which may lead to a good feeling of presence in the virtual environment, but the positions of the hands result to be less realistic than in the previous case, thus hampering a proper interaction with the virtual objects of the scene. This is due to the propagating noise coming from the Kinect data. A solution to this problem represents one of the future steps of this work.

## 3 EXPERIMENTAL EVALUATION

To validate the quality of the proposed system, we have created an experimental session, which was attended by 14 subjects: the participants were both male and female, with ages ranging from 20 to 50, and with normal or corrected-to-normal vision. In this experiment, we recreate a simple skeleton consisting of small spheres, in place of the joints tracked by the Kinect, and parallelepipeds, in place of the bones of the subject, except for the hands, for which we used the ones of the model included in the Leap Motion

Core Assets for Unity3D. A simple skeleton for the body, composed of basic objects, allows us to better manage the data acquired by the Kinect and to provide a coarse interaction with the virtual environment. Meanwhile, a more complex but tested model for the hands guarantees us a better feeling of presence and a more accurate interaction with the virtual objects. Figure 7 shows a snapshot of the experimental session: on the left there is an image from the Oculus Rift, where it is possible to see the reconstruction of the user's hands and the virtual objects with which he/she can interact, on the right there is the acquisition from the point of view of the RGBD sensor. Moreover, a video showing the system and the experimental session is available at https://youtu.be/2UaxkyZbeLQ.

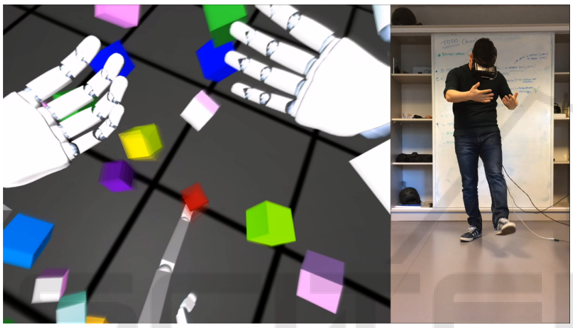

Figure 7: A snapshot of the experiment: (left) scene from the Oculus Rift, and (right) from the RGB-D camera.

After the calibration phase, carried out for each subject, the users were free to move in an area of about $1.5m^2$ and to explore the movements of the skeleton, to ensure that they were coherent with their movements; then some floating cubes were shown and the subjects were free to interact with them, by moving them with the hands and with the entire skeleton.

The aim of the experimental validation was to test the level of immersivity and the experience of the users that tested the proposed system. The sensation of presence (or level of immersion) is often measured by means of self-rated questionnaires (Gorini et al., 2011). Several examples can be found in the literature: in the UCL Presence Questionnaire participants are required to provide ratings on a seven level Likert scale (Slater et al., 1994), and in the Independent Television Company Sense of Presence Inventory (ITCSOPI) users are evaluated postexposure by providing scores on a five level Likert scale (Lessiter et al., 2001). In this paper, we have taken inspiration from the questionnaires present in the literature, and at the end of the session, users were asked to answer a questionnaire with 13 close-ended questions. Subjects rated their feelings on a 5-points Likert scale, where 1 indicated negative feelings at all and 5 indicated the most positive experience. The questions, listed below, were related to: the quality of the VR scenario (Q1-Q4), the truthfulness of the virtual body (Q5-Q9), the sense of immersion in the VR (Q10-Q13).

Q1 How often do you play videogames?
Q2 How much was the scenario immersive?
Q3 How much lag did you notice in the scenario?
Q4 How much did the visual aspects of environment involve you in the scenario?
Q5 How strongly did you feel your virtual body was truly?
Q6 How much did you feel your virtual body follow your real movement?
Q7 How much did the virtual body help you to increase the sense of presence in the environment?
Q8 How much natural was your interaction with the object?
Q9 How closely were you able to examine your hands?
Q10 How much did your experiences in the virtual environment seem consistent with your real world experiences?
Q11 How compelling was your sense of moving around inside the virtual environment?
Q12 How much did the screen resolution negatively affect your sense of presence?
Q13 Did you feel more immersed in the VR over time?

The distribution and the statistics of the responses to the 13 questions are reported in Figure 8.

In general, the feeling about the developed system

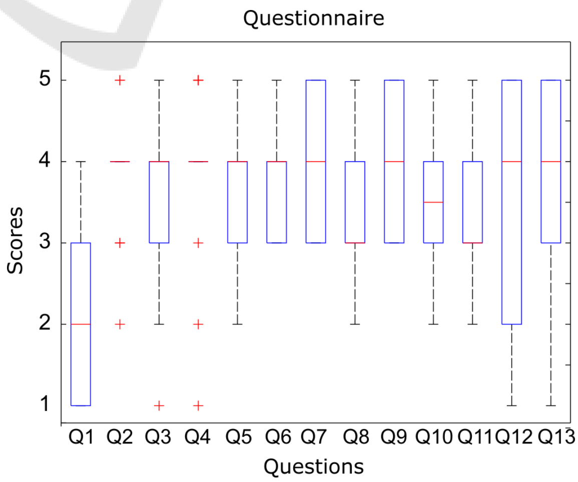


Figure 8: Questionnaire results: box plots of the scores (on the 5-point Likert scale) given by the subjects to the 13 questions. The red lines represent the median scores, the blue boxes are scores between the first and the third quartiles.

was good. Question Q2 about immersivity reported a median value of 4, together with Q5 about the positive feeling about the user's own virtual body inside the VR environment. Nevertheless, users were critic about the natural interaction with the objects (question Q8, median 3) and about the sense of moving around inside the environment (question Q11, median 3). Of course, in such a system the sense of touch is not addressed, though a tactile feedback, also by means of a sensorial substitution, might help the user.

To sum up, the goals of this experiment were: (i) to assess the performances of the proposed system, for what concerns the responsiveness of the reconstructed skeleton to the real movements of the user; (ii) to verify the effects of a full-body controllable avatar on the feeling of presence in the VR. This experiment confirms the validity of the developed system, even in its first prototypal implementation.

# 4 CONCLUSION AND DISCUSSION

In this paper, we have presented a method to insert the avatar of the user own body, which gives the visual feedback and replicates the movements of the user, in a VR environment, enjoyed through an HMD. In particular, our scope was to improve the sense of presence by using immersive devices, such as the Oculus Rift. In order to replicate the movements of both the body of the user, and of his/her hands, by also taking into account the fine details of the fingers, we have proposed to use both a Microsoft Kinect, which acquires the entire body of the users, and a Leap Motion, a low cost device used to measure and track the fingers. Such devices must be accurately registered in order to obtain stable and robust measures, with respect to a coherent reference system. To this aim, we present a 2 steps procedure to achieve such a registration: (i) a rigid transformation through a least-square SVD, from which we estimate the roto-translation to align common points acquired by the two devices; (ii) a live correction, by also taking into account the movements of the users, captured through the Oculus Positional Tracker.

We assessed the proposed VR system through an experimental session, attended by 14 subjects who answered to a self-reported questionnaire. The results show that the users have a positive feeling about the proposed system, in particular with respect to the immersivity and to the sense of presence.

# ACKNOWLEDGEMENTS

The authors would like to thank Prof. George Drettakis and Dr. Adrien Bousseau of the GRAPHDECO group, INRIA Sophia-Antipolis, France.

# REFERENCES

Ahlberg, G., Heikkinen, T., Iselius, L., Leijonmarck, C.-E., Rutqvist, J., and Arvidsson, D. (2002). Does training in a virtual reality simulator improve surgical performance? *Surgical Endoscopy and Other Interventional Techniques*, 16(1):126–129.

Ahmed, B., Yoo, J. D., Lee, Y. Y., Jang, I. Y., and Lee, K. H. (2014). Fusion of Kinect range images and Leap Motion data for immersive AR applications. *Proceedings of the Society of CAD/CAM Conference*, pages 382–387.

Beattie, N., Horan, B., and McKenzie, S. (2015). Taking the LEAP with the Oculus HMD and CAD-Plucking at thin air? *Procedia Technology*, 20:149–154.

Creagh, H. (2003). Cave automatic virtual environment. In *Electrical Insulation Conference and Electrical Manufacturing & Coil Winding Technology Conference, 2003. Proceedings*, pages 499–504.

Gorini, A., Capideville, C. S., De Leo, G., Mantovani, F., and Riva, G. (2011). The role of immersion and narrative in mediated presence: The virtual hospital experience. *Cyberpsychology, Behavior, and Social Networking*, 14(3):99–105.

Lee, D., Baek, G., Lim, Y., and Lim, H. (2015). Virtual Reality contents using the OculusRift and Kinect. *Mathematics and Computers in Sciences and Industry*, pages 102–104.

Lessiter, J., Freeman, J., Keogh, E., and Davidoff, J. (2001). A cross-media presence questionnaire: The ITC-sense of presence inventory. *Presence*, 10(3):282–297.

Lugrin, J.-L., Latt, J., and Latoschik, M. (2015). Avatar anthropomorphism and illusion of body ownership in VR. In *Virtual Reality (VR), 2015 IEEE*, pages 229–230.

Penelle, B. and Debeir, O. (2014). Multi-sensor data fusion for hand tracking using Kinect and Leap Motion. In *Proceedings of the 2014 Virtual Reality International Conference*, page 22.

Seymour, N. E., Gallagher, A. G., Roman, S. A., OBrien, M. K., Bansal, V. K., Andersen, D. K., and Satava, R. M. (2002). Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Annals of surgery*, 236(4):458.

Slater, M., Usoh, M., and Steed, A. (1994). Depth of presence in virtual environments. *Presence*, 3(2):130–144.

Sorkine, O. (2009). Least-squares rigid motion using SVD. *Technical notes*, 120:3.

Sra, M. and Schmandt, C. (2015). Metaspace: Full-body tracking for immersive multiperson Virtual Reality. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, pages 47–48.

Tecchia, F., Avveduto, G., Brondi, R., Carrozzino, M., Bergamasco, M., and Alem, L. (2014). I'm in VR!: using your own hands in a fully immersive MR system. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology*, pages 73–76.

Weichert, F., Bachmann, D., Rudak, B., and Fisseler, D. (2013). Analysis of the accuracy and robustness of the Leap Motion controller. *Sensors*, 13(5):6380–6393.