

Article

Motion Instruction Method Using Head Motion-Associated Virtual Stereo Rearview †

Yoshihiko Nomura *, Hiroaki Fukuoka, Ryota Sakamoto and Tokuhiro Sugiura

Graduate School of Engineering, Mie University, Tsu 514-0111, Japan; h-fukuoka@int.mach.mie-u.ac.jp (H.F.); sakamoto@int.mach.mie-u.ac.jp (R.S.); sugiura@cc.mie-u.ac.jp (T.S.)

* Correspondence: nomura@mach.mie-u.ac.jp; Tel.: +81-59-231-9894

† This paper is an extended version of our paper published in the 10th International Conference on Pervasive Technologies Related to Assistive Environments (PETRA), Island of Rhodes, Greece, 21–23 June 2017.

Received: 27 October 2017; Accepted: 30 November 2017; Published: 30 November 2017

Abstract: In recent years, virtual reality (VR) technologies have been increasingly used for teaching motion skills to learners. In this paper, the authors employed a VR assistive system for teaching motion skills to learners by the use of an inertial sensor-embedded head-mount-display (HMD). As a step of the development, we studied a motion instruction method using “Head Motion-Associated virtual stereo Rearview (HMAR in short)”, and conducted a study on pose-recognition under a time-consuming vision-restricted condition. Under this condition, subjects were to ensure their remembrance only by vision and taking enough time, and not by using proprioception. The time consuming condition is considered to be antithetical to the instantaneous less time consuming condition, and is expected to contribute to deepening the understanding of the effect of the HMAR. In the experiment, reference poses are displayed to learners with the use of a VR system. In the system, the learners observe the virtual stereo rearview via HMD, and perceive and reproduce the displayed reference poses. Here, the virtual stereo camera that is assumed to observe the reference avatar is associated with the learner’s head motion. The virtual stereo camera is moved around the reference avatar away from the back of the avatar’s head in accordance with the head-rotating motion. As the HMAR was compared with two representative ordinary methods, i.e., a key-switched rearview (KSR) and a mouse-associated rearview (MAR), the elapsed time of the HMAR showed significantly smaller variance, although did not show any significant difference in the mean.

Keywords: pose instruction; pose recognition; pose reproduction; vision; VR; HMD; rearview; head motion; sense of agency

1. Introduction

Motion skills in sports such as swimming and pitching balls, and in cultural activities such as dancing and calligraphy, are to be mastered through learning processes. In order to learn the skills effectively, the advice of experts is desirable. However, it is sometimes difficult for us to obtain this expert advice, and, therefore, learners should usually practice in their own ways by the use of tutorial books, pictures, and videos.

Recently, virtual reality technologies (VR techs) have grown popular. For example, for the use of the head-mounted display (HMD), Bowman et al. tried to develop guidelines to choose an appropriate display for a particular virtual environment (VE) system by comparing head-mounted displays with a workbench display and a foursided spatially immersive display [1]. Most recently, smart glasses have also been employed for various augmented reality (AR) applications [2]. The VR techs come into use for motion skill instruction by some researchers. For example, Swan et al. incorporated an HMD-based augmented reality technology into a 3D perceptual task, and studied 3D perceptual characteristics [3].

Covaci et al. proposed an assistive system for basketball free-throw training [4]. An ideal locus of the thrown basketball is superimposed in a third person perspective view (3PP views). It is interesting that motions can be instructed not only by the motions themselves, but also by the motion-resulted changes in the environments. It also suggests the importance of 3PP views. Salamin et al. [5,6] quantified the effect of exposure to the 3PP and the first-person perspectives (1PP) in virtual reality-based training. They attached a TV camera to the learner's back, and compared a spatial cognitive sensitivity of the thrown ball locus prediction between the 3PP and the 1PP. As a result, the 1PP was reported to be superior to the 3PP. Recognition in a field far from the observer was examined in [4], while a field near the observer was examined in [5]. The difference between the far field and the near field is considered to be a reason for the contradictory results.

As for the head-attached camera configuration, Pomés et al. also proposed another type of head attached virtual stereo camera (they referred to it as "head-tracked stereo HMD"), and they reported that the body-ownership increased more in the case of the synchronous provision of visual and tactile stimuli, than in the other case of the asynchronous provision of the two stimuli [7]. Differently from this paper, their work was based on the condition that the learner's head was not rotated, although the observing direction is a crucial factor for view-based pose recognition tasks. By contrast, Hoover et al. reported a motion-perception characteristic using an interesting experiment [8]. In their study, various views such as "behind the view" (referred to as rearview in this paper), "mirror view", and "direct view" were employed. In addition to these views, the views were, furthermore, inverted horizontally and/or vertically. However, their work was based on a simple 1-degree-of-freedom (1-DOF) motion, i.e., a single hand/finger oscillating motion, and did not examine multiple body parts as in this paper.

In this work, a vision-based pose recognition task for the reference avatar's whole-body was taken up as a research theme. Here, the 1PP where the observer's viewpoint coincides with the reference avatar's viewpoint as in the abovementioned works was not employed. Instead, another type of 1PP was employed as follows. The observer's viewpoint, i.e., the position of the virtual stereo camera that observes the reference avatar, is associated with the observer's head motion. That is, in accordance with the observer's head-rotating motion, the virtual stereo camera is rotated around the reference avatar, away from the back of the reference avatar's head. The concept of the head motion-associated virtual stereo camera was originally proposed by some of the authors of this paper [9]. In their work, virtual images of a reference avatar were presented together with the other virtual images of an observer. The virtual stereo camera was rotated around the reference avatar and the observer, and is called "Head Motion-Associated virtual stereo Rearview (HMAR in short)". The observer is superimposed on the reference avatar. Then, some strategies for indicating movements to correct the observer's pose were presented, but a comparative study, as is presented in this paper, was not carried out with respect to the head-motion association. Another result of a pilot study of the system was recently presented by the authors [10], where the with/without-observer's head-association data were compared. This work is an extended version of the previous works [9], devoting intensive experiments and employing two representative ordinary modes that do not employ the observer's head-motion association to a changing view angle. That is, an instruction system based on HMAR was again rebuilt by using an up-to-date HMD-based VR assistive system in order to improve the motion-skill instruction performance, and an intensive pose-recognizing experiment was conducted to examine perceptual characteristics of the proposed system.

Here, the motion learning can be achieved through several processes. It starts from the recognition process, and, via the remembrance process, ends in the reproduction process. Therefore, the learning system should be evaluated with respect to the depth of the learning processes. The deepest task is to perform all three processes, which would be very time-consuming. However, the simplest task is to perform the first recognition. For example, the same or different tasks, and left or right tasks: e.g., the mental rotation-related studies employed the same or different discrimination tasks. Shepard et al. originally found interesting characteristics with respect to the mental rotation of three-dimensional objects [11], and, furthermore, extended their study to the effects of the

dimensionality of objects [12]. Amorim et al. studied the effectiveness of “body analogy” [13]. Rigal examined right-left orientation, mental rotation, and perspective-taking issues by employing a right-left recognition and identification test [14]. This would be instantaneous and consume much less time. Although both kinds of tasks should be evaluated, this paper examined the former task. That is, by taking enough time, subjects ensure their remembrance under the condition that they are allowed to use only vision and not proprioception. The time-consuming condition is considered to be antithetical to the instantaneous and less time-consuming condition as explained above, and is expected to contribute to deepening the understanding of the effect of the HMAR.

2. Proposal of Head Motion-Associated Virtual Stereo Rearview (HMAR)

2.1. System Configuration

This section describes our system for improving the performance of vision-based pose-perceptual characteristics using an HMD (Oculus Rift DK2, Oculus VR, Inc., Irvine, CA, USA). The system is employed to associate the stereo HMD view with the learner’s head-rotation, which is measured by HMD-embedded inertial sensors. It is called “Head Motion-Associated virtual stereo Rearview” (HMAR) in this paper, and is explained in detail in the following. The virtual two cameras are assumed to be a stereo camera in one united body, and to be at the point of several tens of centimeters away from the back of the learner’s head. The learner’s head rotation is measured by HMD-embedded inertial sensors, and, based on the measurement, the virtual stereo rearview of a reference avatar is synthesized in a computer graphics (CG) way by using a software, Unreal Engine 4 (Epic Games, Inc., Raleigh, NC, USA), and the virtual stereo rearview is displayed in the HMD. Some examples of the presented virtual stereo rearview are shown in Figure 1c. Camporesi et al. showed that the use of avatars and user-perspective stereo vision improved the quality of produced motions and the resemblance of replicated motions [15]. Yu et al. used two different tasks. One task was related to object-centered 3-D transformation such as mental rotation, and the other task was related to perspective transformations in which people adopt the egocentric perspective of another person. As a result, similar to Camporesi et al., they suggested that observers were particularly sensitive to the presence of a human head and body in comparison with the other block and teapot figures, and that these human body features allowed observers to quickly recognize and encode the spatial configuration of a figure [16].

Here, note that two modes of human perception can be affected with respect to the avatar motion. One mode is an avatar-centered perception, and is shown in Figure 1a. The other mode is an observer-centered perception, and is shown in Figure 1b. They are explained as follows.

- Avatar-centered perceptual mode (see Figure 1)

In accordance with the observer-head rotation, e.g., a rotation in a counterclockwise (CCW) direction, the observer recognizes their body swinging in a CCW direction around the avatar’s body axis.

- Observer-centered perceptual mode (see Figure 1b)

In accordance with the observer-head rotation, e.g., a rotation in a counterclockwise (CCW) direction, the observer recognizes their body as stationary, while the virtual stereo camera (VSC) rotates in a clockwise (CW) direction around the camera center. Along with the CW rotation of the VSC, the target avatar has also swung around the VSC in the CW direction, the viewpoint of which coincides with the observer. Here, it is noted that, while swinging around the VSC, the target avatar does not perform any autorotation. It is noteworthy that this mode can be more easily realized under no background constitution as proposed in this paper, which is considered to be effective to suppress the VR sickness, i.e., the visually induced motion sickness [17–19].

The two perceptual modes as described above, the avatar-centered perceptual mode and the observer-centered one, deal with perceptual concepts similar to the egocentric and exocentric strategy presented in the paper by Tan et al. [20]. That is, the avatar-centered perceptual mode and the egocentric strategy imagine rotating the observer's body within the environment, while the observer-centered perceptual mode and the exocentric strategy imagine rotating the environment around the observer's body. In either case of the two perceptual modes, the observer can observe the avatar from various directions by rotating the observer's head. Yet, the observer was advised to take the latter observer-centered perception since the former perceptual mode causes VR sickness due to the incoherence between the actual and the perceptual motion of the observer, highlighting an important issue to consider when conducting such experiments.

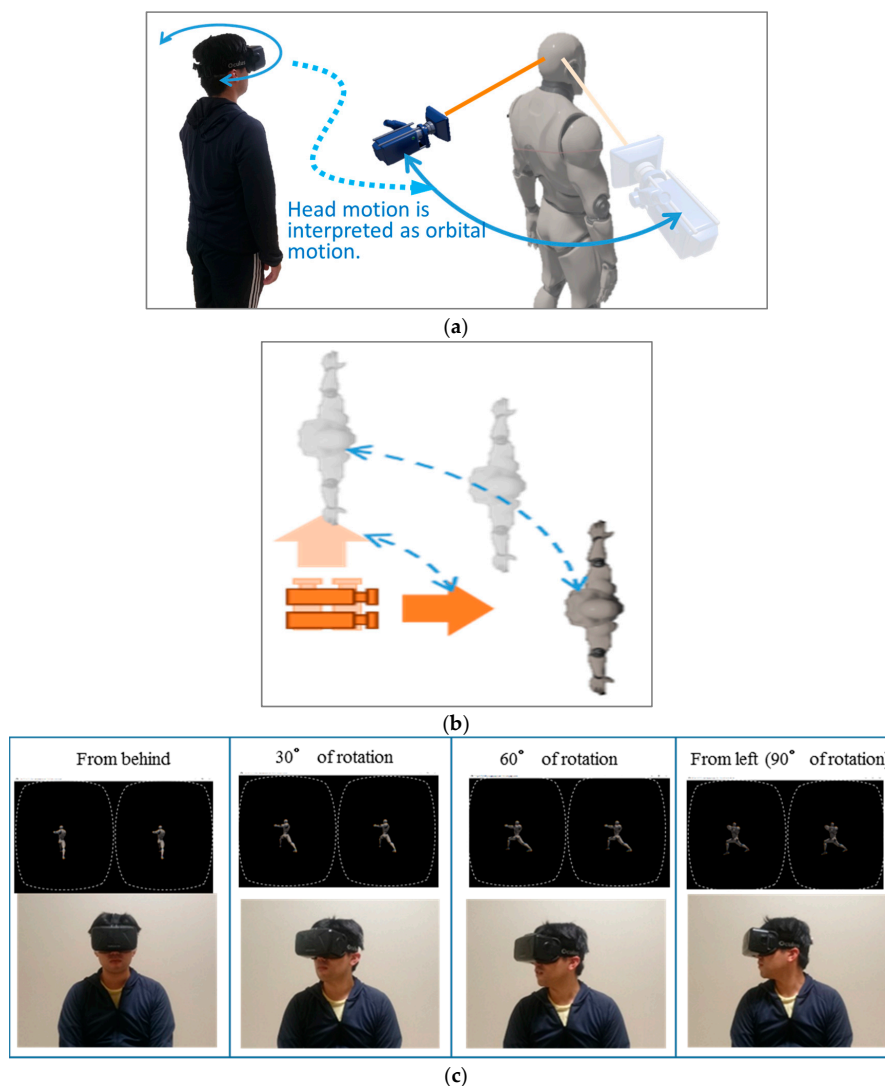


Figure 1. Head motion-associated virtual stereo rearview (HMAR) can be understood in the following two modes. (a) The avatar-centered perceptual mode: in this scheme, the camera corresponding to the observer's viewpoint is understood as the camera, i.e., the observer is swung around the avatar's body axis in accordance with the observer's head rotation. (b) The observer-centered perceptual mode: in this case, the observer feels as if their body is stationary, and just rotates their head, while the avatar is swung around the observer without autorotation. (c) Upper photographs show an example of the head-mount-display (HMD) stereo rearview. Lower photographs show an HMD-wearing subject. The stereo review was continually presented along with the learner's head rotation in the HMAR system.

2.2. Expected Effects

The observer's motion can be sensed through their vestibular sensations when the observer rotates their head in the HMAR. During the rotation, the presented virtual stereo images are varied, and the virtual stereo views are observed through the observer's visual sensation. As for a quantitative study with respect to HMD view-based pose reproduction, Roosink et al. developed a "virtual mirror" that displayed increased and decreased feedback on trunk flexion movements of a realistic full-body avatar [21]. They indicated an unbiased perception in the HMD view-based pose reproduction, which showed reliance on both visual and proprioceptive feedback. As for visuo-vestibular conflicts, Macaуда et al. employed questionnaire and skin temperature measures to assess illusory self-identification with either a mannequin or a control object [22]. They suggested an importance of congruence of the visual sense with the vestibular sense as in the HMAR.

Thus, we can expect the following effects in pose recognition tasks. That is, the head motion-associated view helps the observer feel the "sense of agency (SA)" of an avatar. The word "SA" means a mental sense that "I am able to control the avatar motion freely" [23]. In this work, the avatar is swung around the observer.

3. Pose-Recognition Experiment

3.1. Experimental Method

3.1.1. Experimental Conditions

We conducted a psychophysical experiment on a pose recognition task where the reference avatar alone was shown in an HMD, but not the observer themselves. When the other avatar reflecting the observer is also shown in the displays, another interesting psychophysical phenomena such as the body ownership illusion can occur, as in Maselli et al. [24]. The body ownership contribution was also studied for real-time mirror reflections of motor actions in an immersive virtual environment by Gonzalez-Franco et al. [25]. The influence of self-avatars on 3-D spatial judgment requiring absolute distance perception was also investigated by Mohler et al. [26].

Twelve healthy young males aged 22 to 24 participated in the experiment. They had no experience with yoga exercises. Some of the subjects had experiences with HMDs when having played 3-D games, but their experiences were very limited so that they were not accustomed to HMDs. Therefore, all the subjects were practically regarded as novices with the HMD.

As for the modes of presenting stereo views in the HMD, the following three kinds of rearview modes were introduced for comparison.

1. Head Motion-Associated virtual stereo Rearview (HMAR) (see Section 2).
2. Key-Switched virtual stereo Rearview (KSR): One of the eight discrete views that were taken from the eight viewpoints at 45° intervals in 360° around the target avatar and displayed in the HMD. It was switched on after the other upon a subject's key-pressing operation, as shown in Figure 2. This instructional mode is regarded as the most representative ordinary method.
3. Mouse-associated virtual stereo Rearview (MAR): The virtual stereo camera (VSC) is rotated around the reference avatar in accordance with the mouth-dragging operation by the observer. The stereo view that is assumed to be taken by the VSC is displayed in the HMD, as shown in Figure 3. This instructional method is regarded as an elaborated ordinary method.

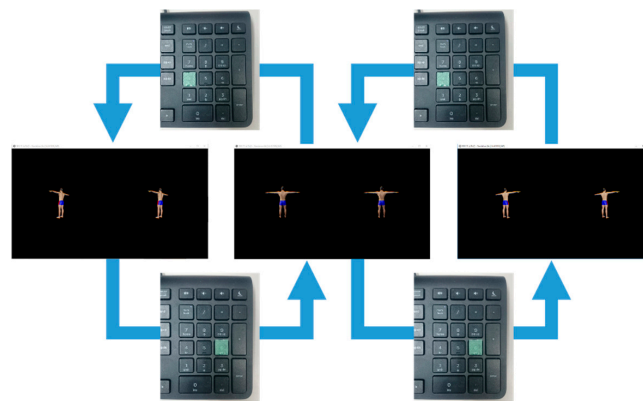


Figure 2. Key-switched virtual stereo Rearview (KSR): In the case of pressing the “6-key”, the displayed stereo-view pair is switched from the present stereo-view pair to the other stereo-view pair rotated by 45° in the counterclockwise direction. In the other case of pressing the “4-key”, it is switched from the present to the other stereo-view rotated by 45° in the clockwise direction.

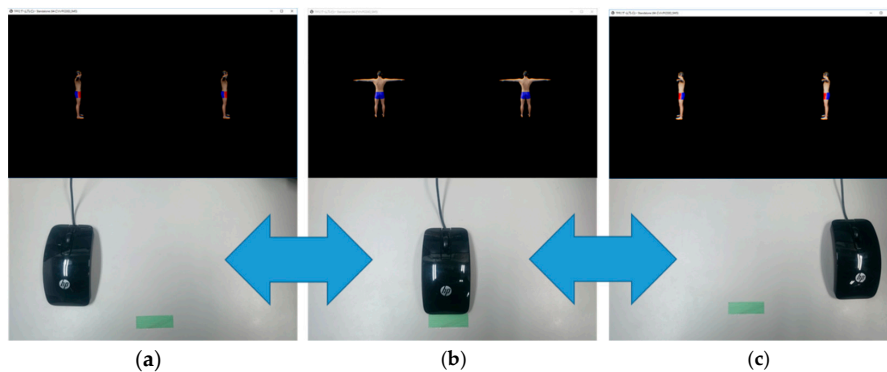


Figure 3. Mouse-associated virtual stereo rearview (MAR): In the case where the mouse is dragged toward the right, the avatar displayed in HMD is rotated in the counter clockwise (CCW) direction around the avatar trunk axis (e.g., from (b) to (c) in Figure 3). In the other case where the mouse is dragged toward the left, the avatar displayed in HMD is rotated in the clockwise (CW) direction around the avatar trunk axis (e.g., from (b) to (a) in Figure 3).

As for the poses presented in the HMD, a total of 18 poses were introduced from “yoga” positions. Some of them were slightly altered to generate more difficult ones. They were classified into three categories, based on twisting and crossing motions.

1. Plain poses (see Figure 4): This category of pose is considered to be recognized only by one view.
2. Twisted poses (see Figure 5): Upper trunk is bent and/or twisted. It is considered to be necessary to look at the avatar from multiple directions in order to recognize the poses.
3. Crossed poses (see Figure 6): Some of the avatar’s limbs are crossed. It is also considered to be necessary to look at the avatar from multiple directions in order to recognize the poses, as in the twisted poses. In addition, positional relationships between limbs should be studied further.

The difficulty is considered to increase in the order of the plain, twisted, and crossed poses.

The experiment was designed by the following protocol. There were three main factors to be examined. The first factor was the view-presenting mode. The number of this factor level was three. The second factor was the pose category. The number of this factor level was also three. The third factor was the specific pose defined for each of the three pose categories. The number of this factor level was six. The last factor was the subject. The number of this factor level was twelve. Then, using a part of a 36-runs orthogonal array [27], the combinations of the factor levels for the four factors

were determined by considering the counterbalance among the factor levels in each of the factors. As a result, 108 runs in total (nine runs in each of the 12 subjects) were carried out in this experiment. The orders of the view-presenting modes were also counterbalanced among the subjects.

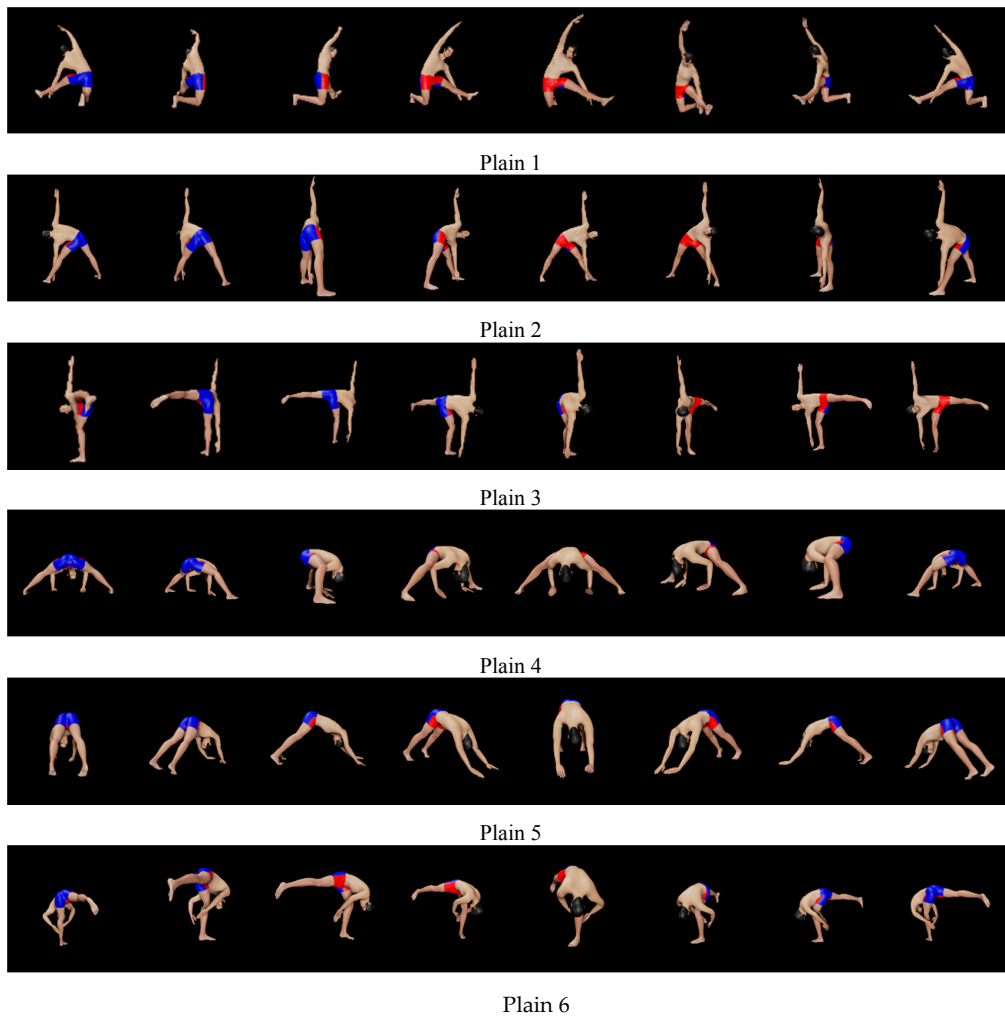


Figure 4. Six plain poses.

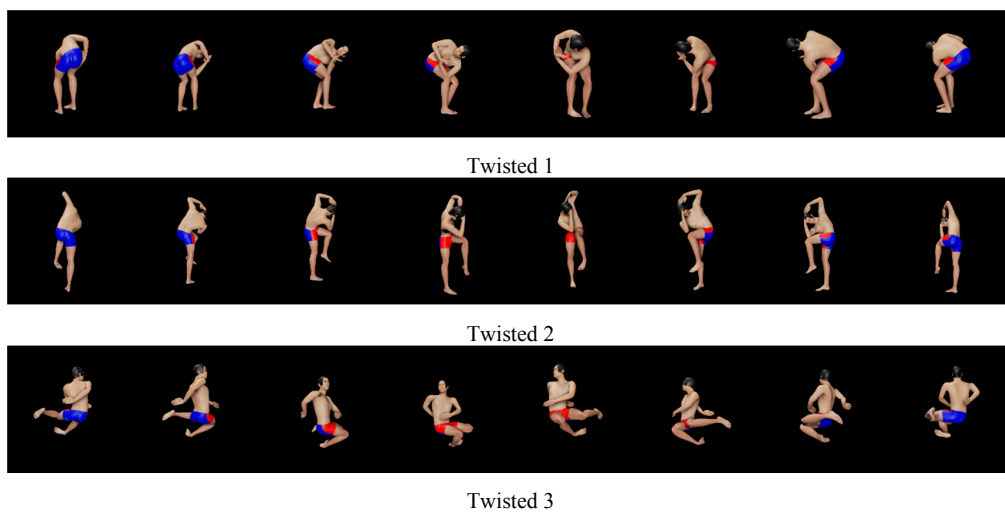


Figure 5. Cont.

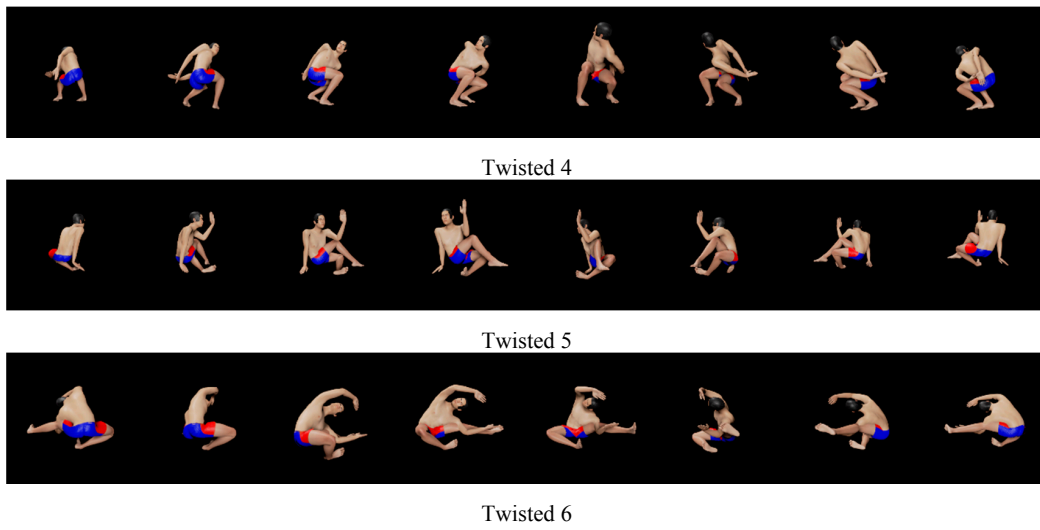


Figure 5. Six twisted poses.

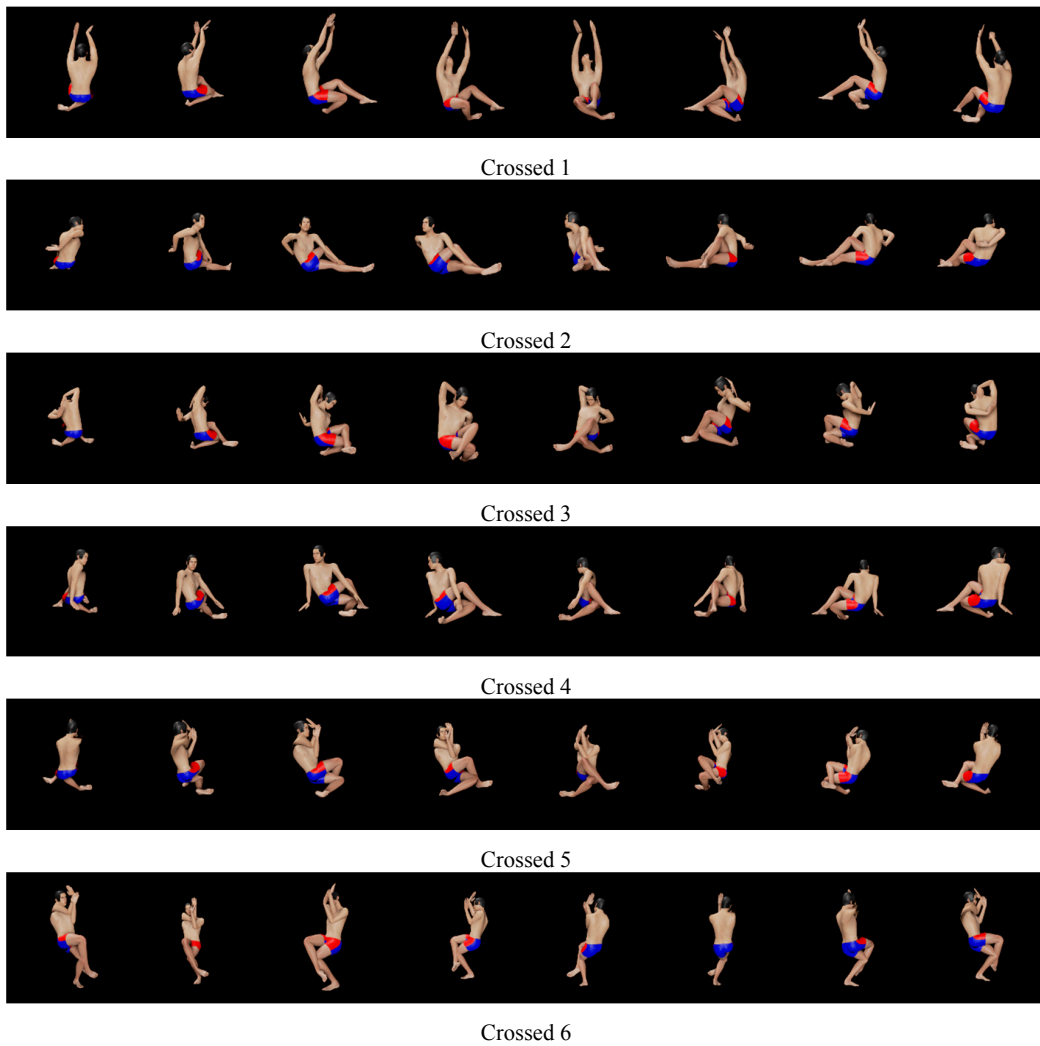


Figure 6. Six crossed poses.

3.1.2. Experimental Procedure

Every run was performed using the following procedure.

1. Sitting on a chair, the subject wore an HMD, and closed his eyes.
2. Experimenter decided upon one of the above explained three view-presenting modes and an avatar pose. Then, the experimenter instructed the subject to perceive the avatar's pose by using the instructed view-presenting mode.
3. Triggered by the notice of the experimenter, the subject opened his eyes. Then, employing the instructed view-presentation mode, the subject made an effort to recognize and store the avatar's whole pose in their memory as early as possible.
4. Just after the subject finished the above process, they said "Yes", instantly stood up, and reproduced the avatar's pose by themselves.
5. The elapsed time was measured as the overall evaluation value of the pose recognition, storage, and reproduction performance. The shorter the elapsed time is, the more the subjects are assumed to get many pieces of information for pose-recognition effectively and to store them.

3.2. Experimental Results

Figure 7 shows an example of the presented pose (Figure 7a) and the reproduced poses by 12 subjects (Figure 7b–m). As shown in Figure 7m, though a rare case, subjects sometimes took left-and-right reversed poses in pose reproduction. The causes of the pose-reproduction failure were due to misperception and forgetfulness. The misperception-related failures are the left-and-right and front-and-back reversions in the crossed poses, and the clockwise-and-counterclockwise reversions in the twisted poses. The ratio of the reproduction failure due to forgetfulness was about four percent, and that due to misperception was six percent. However, there was no significant difference between them as a result of a statistical test on the ratio difference. Next, the failure occurrences due to misperception were very few, and the number of failures was only 2, 1, and 1 for the HMAR, MAR, and KSR, respectively. Therefore, we were not able to confirm significant differences between them as a result of the same statistical test as described above. In the time-consuming conditions presented in this work, it is considered difficult to make the differences obvious, but we could not deny the possibility of finding a difference by conducting many large sample size experiments. Thus, since such pose-reproduction failures rarely occurred, we shall examine the elapsed time in the following. It is considered that, regardless of whether there are failures in pose reproductions or not, the reproduction processes will be similar. Therefore, there would be no appreciate differences between the elapsed time with-failure and that without-failure, and the elapsed time data with failure as well as without failure were also applied to the analysis.

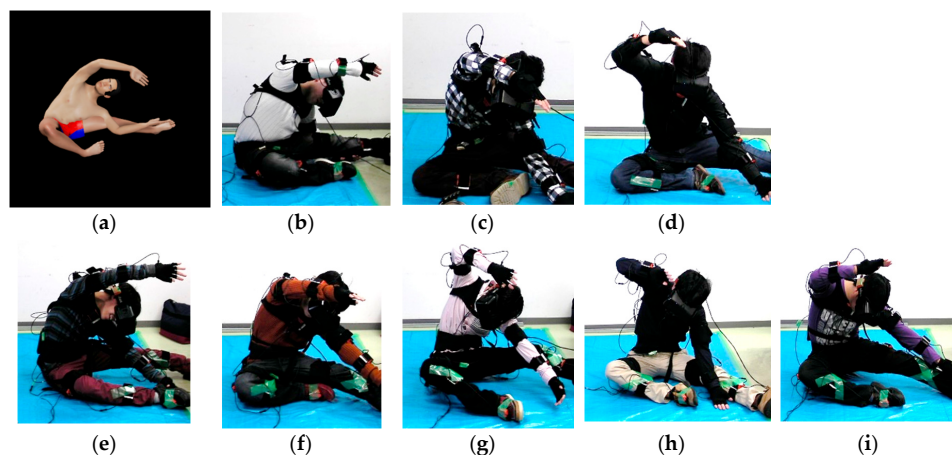


Figure 7. *Cont.*

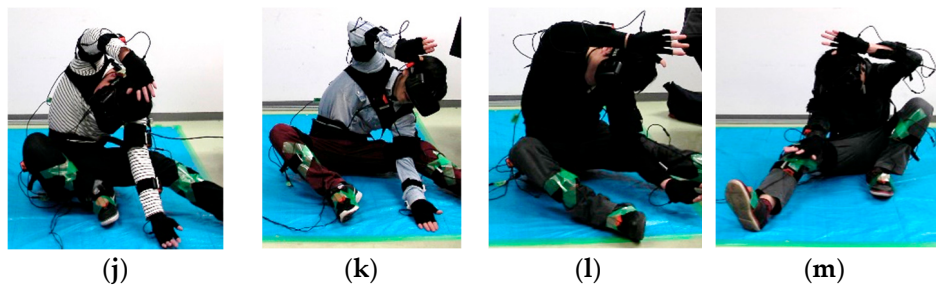


Figure 7. A presented pose and its reproduced poses: (a) Presented; (b–l) successfully reproduced; (m) left and right reversed.

The factor effect on the subjects and the pose categories are shown in Figures 8 and 9, respectively. The individual subject differences and the pose-dependent differences are very large. Therefore, cancelling the factor effects with respect to the individual subject, and the 18 poses, the means and the standard deviations of the elapsed time were plotted for each of the three viewing schemes as shown in Figure 10. By removing the global mean effect in Figure 10, and by changing the representation of the error bars from the standard deviations into the standard errors, the view-presenting mode factor effect is shown in Figure 11.

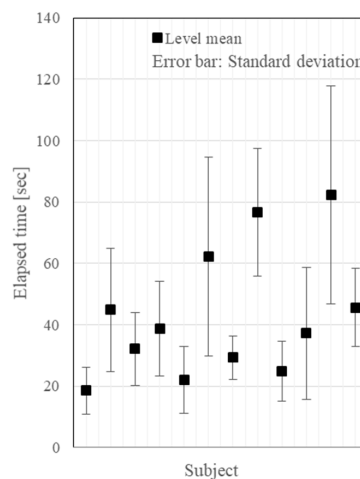


Figure 8. Experimental results of the pose-reproduction. The subject factor effect.

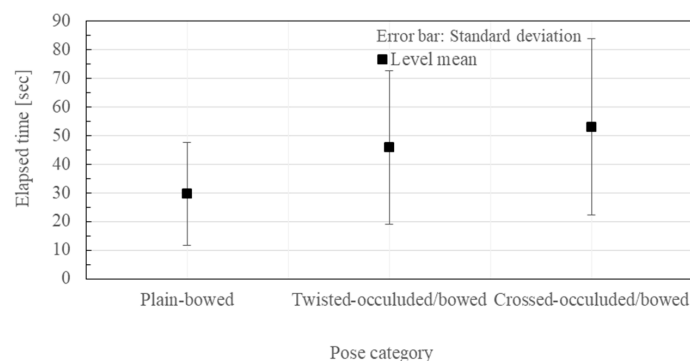


Figure 9. Experimental results of the pose-reproduction. The pose category factor effect.

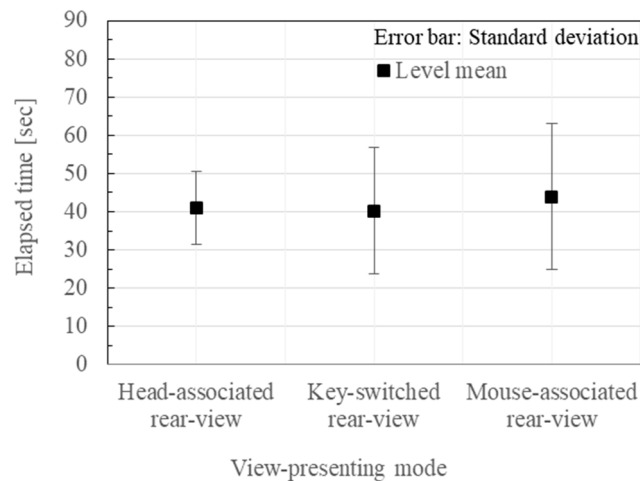


Figure 10. Experimental results of the pose-reproduction. The elapsed time for the three pose-presenting modes where the effects of the subject and the pose factor together with the global mean were removed, but the effect of the global mean was not removed.

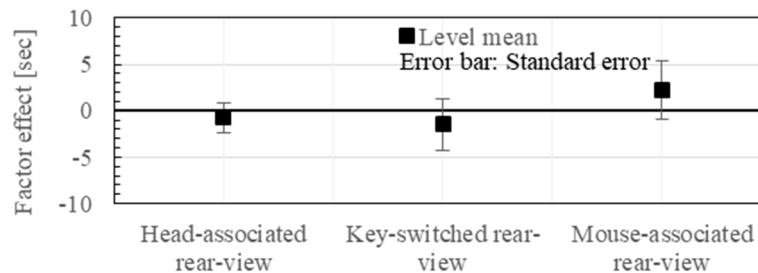


Figure 11. Experimental results of the pose-reproduction. The pose-presenting mode factor effect on elapsed time where the effects of the subject and the pose factor together with the global mean were removed.

3.3. Discussion

The authors have applied some statistical tests to the experimental results of Figure 11. As a result of the *t*-test with the Bonferroni multiple-comparison correction, there were no significant differences among the population means of the three view-presenting modes. However, as a result of the *F*-test with the Bonferroni multiple-comparison correction, the variance of the Head Motion-Associated virtual stereo Rearview (HMAR) was significantly smaller in comparison to either the key-switched rearview (KSR) or the mouse-associated rearview (MAR) under the significant level of 0.1%: the test statistics of $F = 4.00$ and 2.93 for KSR and MAR, respectively. Generally speaking, worse cases sometimes result in fatal results, and the stability that is represented by the variance is principally one of the important performance indices. Considering this principle, the significant difference with respect to the variance shows an advantage of the HMAR mode.

However, the former result of no difference in the population means was considered to stem largely from the following matter. In this experiment, the subjects were instructed not to move their body, but were permitted to operate only one motion in Section 3.1.2 Step 2: The permitted motion was rotating their head in the HMAR mode, pressing with their right finger in the KSR mode, and dragging their right-hand in the MAR mode. This kind of vision-restricted constraint increases the mental burdens for subjects in order to recognize and, especially, to store the body part poses in Section 3.1.2 Step 3. The mental burdens would result in a much longer elapsed time.

However, there were differences in the perceptual performances of the presenting modes, i.e., the presenting-mode factor effect on the elapsed time is considered to be momentary and much less

than the overhead elapsed time, i.e., the global mean of the elapsed time. The much longer elapsed time overwhelms the presenting-mode factor effect of the elapsed time.

The time-consuming condition, as was employed in this paper, is antithetical to the less time-consuming conditions. We can deepen our understanding of the view-presenting modes by examining both the time-consuming conditions and the less time-consuming conditions.

Therefore, different from the perceptual condition presented in this paper, which required participants to take time and to ensure their remembrance only by vision, another perceptual condition should also be examined. That is, an instantaneous perceptual condition, which would be important for dynamic motion instructions and would include a short elapsed time, is desired for further investigation. For example, while observing the reference avatar, subjects can move their limbs one once: Under this condition, loads for remembrance would be markedly decreased, and the presenting-mode factor effect on the elapsed time would show clearer differences.

4. Conclusions

A method for vision-based pose/motion-perception was proposed. It associates HMD views with the observer's head-rotation measured by HMD-embedded inertial sensors, and is called "Head Motion-Associated virtual stereo Rearview" (HMAR): The viewing direction of the virtual stereo camera, which is assumed to be set at a back point of the reference avatar's head, is controlled by the observer themselves via the HMD-embedded inertial sensors. Thus, HMD views are changed in accordance with the subject's head-rotating movements.

The authors have presented an experimental result with HMAR employing 12 subjects in a vision-restricted condition in which subjects ensured their remembrance only by vision and by taking enough time, and not by using proprioception. Comparing the elapsed time by the proposed HMAR with those by two representative ordinary methods, i.e., a key-switched rearview (KSR) and a mouse-associated rearview (MAR), the proposed HMAR showed significantly smaller variance, although equivalent in the mean.

In the future, the authors are directed to conduct another pose/motion perceptual experiments with the HMAR. In the experiments, instantaneous less-time consuming responses are to be precisely measured, and the performance of the HMAR is to be made clear from both the time consuming and the less-time consuming responses.

Acknowledgments: This work was supported by KAKENHI (Grant-in-Aid for scientific research (B) 19H02929 from Japan Society for the Promotion of Science (JSPS)).

Author Contributions: Yoshihiko Nomura conceived and designed the experiments, analyzed the experimental data, and wrote the paper; Hiroaki Fukuoka made the programs and performed the experiments; Ryota Sakamoto supported the programming; Tokuhiro Sugiura wrote the paper. All authors have read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The authors declare that the founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

1. Bowman, D.A.; Datey, A.; Farooq, U.; Ryu, Y.; Vasnaik, O. Empirical Comparisons of Virtual Environment Displays. Available online: <https://vtechworks.lib.vt.edu/bitstream/handle/10919/20060/TR1.pdf?sequence=3> (accessed on 30 November 2017).
2. Lee, L.H.; Hui, P. Interaction Methods for Smart Glasses. *arXiv* **2017**, arXiv:1707.09728.
3. Swan, J.E.; Jones, A.; Kolstad, E.; Livingston, M.A.; Smallman, H.S. Egocentric depth judgments in optical, see-through augmented reality. *IEEE Trans. Vis. Comput. Graph.* **2007**, *13*, 429–442. [[CrossRef](#)] [[PubMed](#)]
4. Covaci, A.; Olivier, A.H.; Multon, F. Visual perspective and feedback guidance for VR free-throw training. *IEEE Comput. Graph. Appl.* **2015**, *35*, 55–65. [[CrossRef](#)] [[PubMed](#)]

5. Salamin, P.; Tadi, T.; Blanke, O.; Vexo, F.; Thalmann, D. Quantifying effects of exposure to the third and first-person perspectives in virtual-reality-based training. *IEEE Trans. Learn. Technol.* **2010**, *3*, 272–276. [[CrossRef](#)]
6. Salamin, P.; Thalmann, D.; Vexo, F. The benefits of third-person perspective in virtual and augmented reality? In Proceedings of the ACM Symposium on Virtual Reality Software and Technology, Limassol, Cyprus, 1–3 November 2006; pp. 27–30.
7. Pomés, A.; Slater, M. Drift and ownership toward a distant virtual body. *Front. Hum. Neurosci.* **2013**, *7*, 1–11. [[CrossRef](#)] [[PubMed](#)]
8. Hoover, A.E.; Harris, L.R. The role of the viewpoint on body ownership. *Exp. Brain Res.* **2015**, *233*, 1053–1060. [[CrossRef](#)] [[PubMed](#)]
9. Sakamoto, R.; Yoshimura, Y.; Sugiura, T.; Nomura, Y. External skeleton type upper-limbs motion instruction system. In Proceedings of the JSME-IIP/ASME-ISPS Joint Conference on Micromechatronics for Information and Precision Equipment: IIP/ISPS joint MIPE, Tsukuba, Japan, 17–20 June 2009; pp. 59–60.
10. Nomura, Y.; Fukuoka, H.; Sakamoto, R.; Sugiura, T. Motion Instruction Method Using Head Motion-Associated Virtual Stereo Rearview. In Proceedings of the 10th International Conference on Pervasive Technologies Related to Assistive Environments, Island of Rhodes, Greece, 21–23 June 2017; pp. 56–58.
11. Shepard, R.N.; Metzler, J. Mental rotation of three-dimensional objects. *Science* **1971**, *171*, 701–703. [[CrossRef](#)] [[PubMed](#)]
12. Shenna, S.; Metzler, D. Mental rotation: Effects of dimensionality of objects and type of task. *J. Exp. Psychol. Hum. Percept. Perform.* **1988**, *14*, 3–11.
13. Amorim, M.A.; Isableu, B.; Jarraya, M. Embodied spatial transformations: “Body analogy” for the mental rotation of objects. *J. Exp. Psychol. Gen.* **2006**, *135*, 327–347. [[CrossRef](#)] [[PubMed](#)]
14. Rigal, R. Right-left orientation, mental rotation, and perspective-taking: When can children imagine what people see from their own viewpoint? *Percept. Mot. Skills* **1996**, *83*, 831–842. [[CrossRef](#)] [[PubMed](#)]
15. Camporesi, C.; Kallmann, M. The Effects of Avatars, Stereo Vision and Display Size on Reaching and Motion Reproduction. *IEEE Trans. Vis. Comput. Graph.* **2016**, *22*, 1592–1604. [[CrossRef](#)] [[PubMed](#)]
16. Yu, A.B.; Zacks, J.M. How are bodies special? Effects of body features on spatial reasoning. *J. Exp. Psychol.* **2016**, *69*, 1210–1226. [[CrossRef](#)] [[PubMed](#)]
17. Sharples, S.; Cobb, S.; Moody, A.; Wilson, J.R. Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display systems. *Displays* **2008**, *29*, 58–69. [[CrossRef](#)]
18. Lo, W.T.; So, R.H. Cybersickness in the presence of scene rotational movements along different axes. *Appl. Ergon.* **2001**, *32*, 1–4. [[CrossRef](#)]
19. Rebenitsch, L.; Owen, C. Review on cybersickness in applications and visual displays. *Virtual Reality.* **2016**, *20*, 101–125. [[CrossRef](#)]
20. Tan, D.S.; Gergle, D.; Scupelli, P.; Pausch, R. Physically large displays improve performance on spatial tasks. *ACM Trans. Comput. Hum. Interact.* **2006**, *13*, 71–99. [[CrossRef](#)]
21. Roosink, M.; Robitaille, N.; McFadyen, B.J.; Hébert, L.J.; Jackson, P.L.; Bouyer, L.J.; Mercier, C. Real-time modulation of visual feedback on human full-body movements in a virtual mirror: Development and proof-of-concept. *J. Neuroeng. Rehabil.* **2015**, *12*, 2. [[CrossRef](#)] [[PubMed](#)]
22. Macaudo, G.; Bertolini, G.; Palla, A.; Straumann, D.; Brugger, P.; Lenggenhager, B. Binding body and self in visuo-vestibular conflicts. *Eur. J. Neurosci.* **2015**, *41*, 810–817. [[CrossRef](#)] [[PubMed](#)]
23. David, N.; Newen, A.; Vogeley, K. The “sense of agency” and its underlying cognitive and neural mechanisms. *Conscious. Cognit.* **2008**, *17*, 523–534. [[CrossRef](#)] [[PubMed](#)]
24. Maselli, A.; Slater, M. The building blocks of the full body ownership illusion. *Front. Hum. Neurosci.* **2013**, *7*, 1–15. [[CrossRef](#)] [[PubMed](#)]
25. Gonzalez-Franco, M.; Perez-Marcos, D.; Spanlang, B.; Slater, M. The contribution of real-time mirror reflections of motor actions on virtual body ownership in an immersive virtual environment. In Proceedings of the IEEE Virtual Reality Conference (VR), Waltham, MA, USA, 20–24 March 2010; pp. 111–114.

26. Mohler, B.J.; Creem-Regehr, S.H.; Thompson, W.B.; Bühlhoff, H.H. The effect of viewing a self-avatar on distance judgments in an HMD-based virtual environment. *Presence Teleoper. Virtual Environ.* **2010**, *19*, 230–242. [[CrossRef](#)]
27. Available online: http://support.sas.com/techsup/technote/ts723_Designs.txt (accessed on 29 November 2017).



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).