The Benefits of Third-Person Perspective in Virtual and Augmented Reality?

Patrick Salamin
Virtual Reality
Laboratory
Ecole Polytechnique Fédérale,
Lausanne

patrick.salamin@epfl.ch

Daniel Thalmann
Virtual Reality
Laboratory
Ecole Polytechnique Fédérale
Lausanne

daniel.thalmann@epfl.ch

Frédéric Vexo
Virtual Reality
Laboratory
Ecole Polytechnique Fédérale
Lausanne

frederic.vexo@epfl.ch

ABSTRACT
Instead of the reality in which you can see your own limbs, in virtual reality simulations it is sometimes disturbing not to be able to see your own body. It seems to create an issue in the proprio-perception of the user who does not completely feel integrated in the environment. This perspective should be beneficial for the users. We propose to give the possibility to the people to use the first and the third-person perspective like in video games (e.g. GTA). As the gamers prefer to use the third-person perspective for moving actions and the first-person view for the thin operations, we will verify this comportment is extendable to simulations in augmented and virtual reality.

Categories and Subject Descriptors
J.4 [**Social and behavioral sciences**]: Computer applications – *Psychology*

General Terms
Measurement, Performance, Experimentation, Human Factors, Verification

Keywords
Presence, immersion, proprio-perception, exocentric perspective, distance evaluation

1. INTRODUCTION
We first assume that if the third-person perspective is preferred for some actions in the video games, it is because gamers get some benefits with this view. During the simulations in virtual and augmented reality, we propose to offer the possibility to the users to switch between the first- and the third-person perspective. We noticed the gamers usually prefer the third-person perspective while moving whereas the first-person one seems to be preferred for the actions which need more precision like sniping or some thin manipulations with the hands. We will verify these benefits are also interesting in virtual and augmented reality simulations.

We noticed that even if we have perfect conditions (perfectly built environment, best hard- and software), the user does not really feel embed in the environment. We could compare this with a psychological problem: the perception of oneself. The user does not see him-/her in the environment. We propose to use a possibiliy of switch between both perspectives to test which view is preferred for several situations. We decided to make our experimentations in the reality instead of a virtual environment with virtual camera to allow the use of this switch in the augmented reality, too.

In the case of third-person game like Colin McRae Rally '04, when a gamer is taking a curve with the car, he/she also often turns his/her game lever even if he/she knows he/she is not turning the car steering-wheel of the game. The presence notion as M. Slater defined it in [6] seems to be increased with this perspective. We want to test if it could be the same in virtual and/or augmented reality applications and which perspective better immerses the user in the simulation. On one side, (with the third-person perspective) it can be disturbing to turn the head with a HMD and always being looking at the same place with an avatar turning its head. On the other side (first-person perspective), not to see oneself when we evolve in a virtual (even in a real) environment seems to prevent from a good immersion. We will test both vision approaches and discuss the immersion quality of each.

We will make a small state of the art to highlight the current problems of simulations with a HMD in virtual and augmented reality. We propose then some experiments with both first- and third-person perspective to verify the benefits obtained with this exocentric viewpoint and the interest for the user to be able to switch from a perspective to the other.

2. RELATED WORKS
The third-person perspective appeared in the video games a few years ago. As written in [1], this perspective seems to be preferred in action games while the avatar is running in galleries. It would provide a more global view of the environment to the user. We think both perspectives are important and we verify if the preference for the third-person perspective is also available during the simulations in virtual and/or augmented reality.

We can see in [4] that there is no significant difference between ego- and exocentric vision concerning the global judgment to perform an action such as researching an object and going to take it in an unknown environment. But in their case, it was the WYSIWIS paradigm ("What You See Is What I See") that was

explored. They were two people looking at the same scene. The first one could see the target and guide the second who did not see it but who had to catch it. In our case, there is only one person who is acting.

In [10], some problems of coordination between the hands and the eyes are highlighted. Moreover, as we are working with a single camera, the tester has the same picture on both screens of the HMD. This lack of stereo vision adds again some troubles to evaluate the distances. We assume it could be partially compensated by the third-person perspective that increases the presence and the field of view.

In further experiments we could use two cameras to provide stereo vision to the tester, but this would load down our system. This could increase the risk of breaks in presence (BIP) as defined in [7]. It has been shown in [8] that people better evaluate the distance to a target when they must walk to it because they will have to provide an effort to reach it. In our case, we amplify this effort with our heavy system. Wearing a heavy system would be a way to avoid the underestimation of the distances?

It is shown in [12] that a small resolution combined with a bad quality of the image can affect the tester judgment. We decided then to use a HMD Kaiser with a resolution of 1024 per 768 at 60Hz. It hat been shown in [5] that the limited field of view of a HMD is not the cause of distance underestimation. We will then see if the user is perturbed and if he/she better estimates the distances in the third-person perspective. Moreover it has been shown in [11] that this underestimation of the distances is linear and overall valuable for virtual environments. So, even if it is also written in [9] that people usually underestimate the distances with a HMD, people used to work with this device (like gamers with video games) should be able to compensate for this problem of distance.

In a psychological point of view it can be disturbing to see oneself moving but in another side, as said in [2], watching one's body in the environment (augmented or virtual reality) is important to feel in the simulation. The third-person perspective would reinforce the user immersion because our tester would see him-/herself and not a character controlled with buttons in the HMD.

As we want to see one's top of the body in the universe, we need to have a following camera like in the video games and its location is a huge problem depending on the movements of the user. In the action games like Tomb Raider, the video must always stay in the environment being able to view the avatar even if the character is backed up against a wall. This means that sometimes the point of view changes and becomes in front of the user (e.g. the aforesaid case). We always need to have at least around 100 cm between the camera and the tester head to provide a global vision of the scene. Moreover we have to raise the height of the camera to allow the user to see the objects in front of him on the ground. During the simulations we would have to be careful with the walls and the ceiling (e.g. door framework).

3. Experiments
We made the experiments with eight people (six males) from 23 to 27 years old. Six of them were used to play video games and five had already worn a HMD.

3.1 Hardware setup
To perform these experiments we first wanted to use the digital camera Canon MVX45i. But it was too heavy and leaded to "balance effects". Moreover, due to its weight, we were not able to fix it on the HMD in front of the head. We finally decided to work with a radio color mini-spycam (wide-field of view) providing a video flow in PAL format (450 lines). It only weights a few grams and can thus be easily fixed on the HMD.

Figure 1. Description of the used hardware and screenshot.

We switch the PAL signal provided to the computer into two the VGA output to send them to the HMD Kaiser ProView XL50 providing a resolution of 1024 per 768 at a refresh of 60Hz.

We use a rigid backpack because the bars of aluminum fixed to it must not oscillate during the tester moves. We use a swiveling pivot point to plug the camera at 80cm behind and 60cm upper the user eye position with an orientation of 7 degrees in direction to the bottom from the horizontal. We have a field of view of 60 degrees which means the tester can see his/her shoulders, head and objects in front of him/her at a distance larger or equal to 1.5m corresponding to two footsteps. Concerning the first-person perspective, we simply plugged the camera on the HMD on front of the eyes in the center.

3.2 Experiments presentation
Our experiment is composed of six steps. We want to check which perspective is preferred. Every test will be performed with both perspectives. We begin with the adaptation test and we randomly choose the next step we perform (also for the perspective).

We first make an accommodation step consisting in walking in an already known room without obstacles. After this, there are five experiments which are randomly chosen: walk through a gallery of almost 50 meters composed curves with obstacles, go and open a door, put a ball in a cup of coffee, receive and send a rolling ball with the feet and finally with the hands.
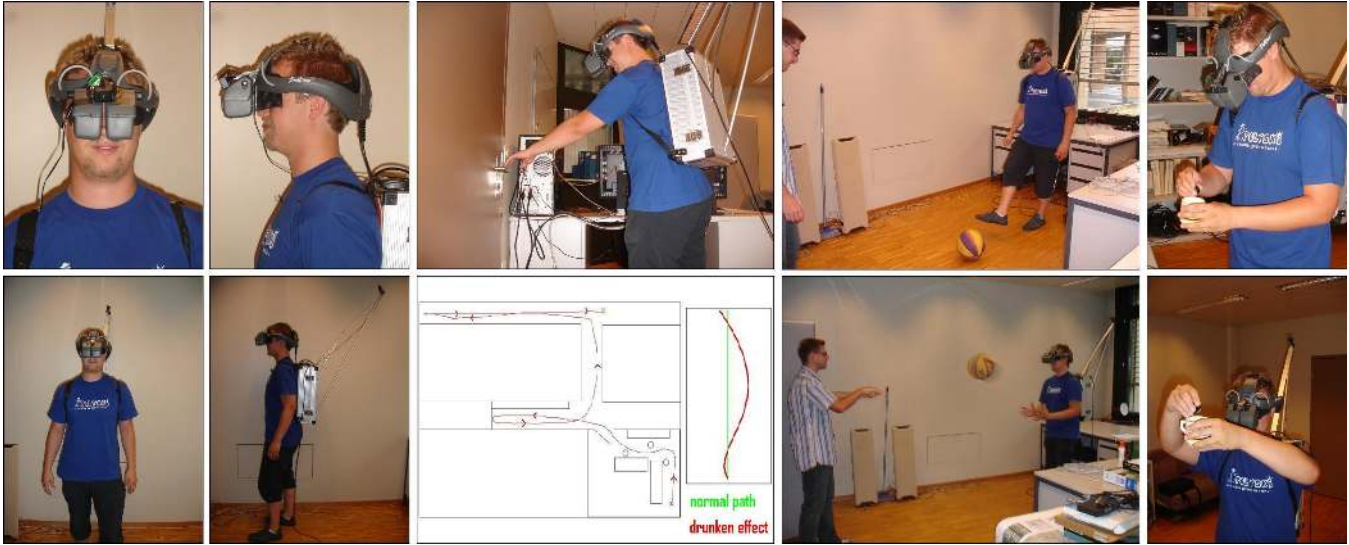
Figure 2. On the left, view of a tester using 1ˢᵗ and 3ʳᵈ perspective view. On the right, examples of some experiments.

4. RESULTS

In this part of the paper we present the results obtained for each experiment (one step per paragraph) with every perspective: third- and first-person with the HMD. For the experiments, we first randomly choose if we begin with the first- or the third-person perspective. We begin with the adaptation experiment and continue with the others (also randomly chosen).

4.1 Presentation of the experiments results

The main goal of this first experiment is to help the tester getting used with the current proposed perspective. Walking in a room without any obstacles and not going to close to the walls seems to be very easy with every perspective. After less than two minutes, every tester seemed to be fine with every perspective (120 seconds for the third-person view instead of 90).

After the tester get used with the perspectives we ask him to go out of a room (path represented by a picture above) where some desktops and dustbins oblige the tester to use roundabout ways to avoid to stumble against them. Note that with our system, he/she must bend his/her knees for height reasons while passing through the door border. This experiment showed us that the third- is preferred to the first-person perspective while the users need to walk in a gallery. Even if they must memorize the place of the closest obstacles, they follow with more facility a straight line and do not feel seasick at the end of the step. There were a few collisions with both perspectives and they used 200 seconds with the third-person view instead of 180.

Another step is to stay at two meters from a door they must open. They easily take the handle and open the door. Our tester do not need more much time to catch the handle and open the door (a fraction of second at maximum) but they usually seem to overvalue the distance to the door when they went to it. However, if the handle had been lower or out of their field of view, most of the testers said they would have had to fumble it for a moment. We could then say that walking action and distances evaluation without stereo vision is easier to do with the third-person perspective while target actions or hand manipulations such as opening a door can be better performed at the first-person perspective.

The next step of our experiments happens in an office with a desktop on which stay a ball and a cup. The testers have to take the ball and put it into the cup. We did not tell them anything about the way they have to use to perform this action because we did not think it would change from a perspective to another one (right pictures above). As we predicted it, there is a problem with target actions and hand manipulations when it happens at low height. In this current case, the objects were not in the field of view for the third-person perspective. Our participants needed then more effort and time to perform this task with the third-person view.

Until now, we only presented experiments with static environments. We will now test the interaction with an external people and mobile elements. Two people at a distance of three meters are passing a ball. At the beginning we will test if our testers can receive a ball which another person passes them with his foot on the ground. After they get this football pass, they have to make a pass to the other person. We can remark they seem to better anticipate the ball location while they are using the third-person perspective. It could be due to the field of view (the bounds) which is more common with the real eye field of view? Nevertheless, they better perform with the third-person perspective but it is easier for them to prepare the ball with their foot when they use the first-person perspective (even if some tester were unbalanced). The switch seems to be really interesting in this kind of actions.

This step is similar enough with the previous one. The main difference is that in this case we are playing with the hands instead of the feet. Moreover the ball is flying in the air instead of rolling on the ground. To perform this action, both seem to be equivalent. This result may unfortunately not be significant due to the fact that no one succeeded this step but some participants said that they preferred the third-person perspective while the ball did not come exactly in front of them.

4.2 Presentation of the analysis tools and global overview of the obtained results

During these experiments we gave a questionnaire to the testers to know exactly how they felt during the several experimentations, how hard it was, why they used this way to perform the action, etc. For most of the questions, they only had to cross a case where appeared numbers from 0 to 10. Number 0 was meaning the worst and 10 the best. There also always were some blank lines to become their comments.

After receiving the questionnaires of every tester, we compare the averages of the results obtained with the first- and the third-person perspective. Despite of what we should have thought, the results for the first-person perspective vary between 3 and 7, which is not so bad. But those obtained during the experiments with the third-person perspective were from 6 to 9, which is really better. The main advantages of the first-person perspective were the smaller adaptation time and the possibility to use more common gesture to catch a close object. But in most of the situations, the users widely prefer the third-person perspective, e.g. walking, evaluating the distance, opening a door, playing with a ball (with hands and feet).

5. CONCLUSION

During this research we wanted to show that it could be very useful and interesting to be able to switch from one perspective to the other depending on the current action, as in the video games. We made these experiments in the reality with a physically-built system to be able to use it in virtual applications but also in augmented reality.

Based on the obtained results, we can say that some actions require the first-, others need the third-person perspective and some accept both. Third-person perspective is usually preferred (verily mandatory) for displacement actions and interaction with moving objects while the first-person view is required when we need to look down or just in front of us for hand manipulations with immobile objects. Another interesting point is that users generally better evaluate the distances, anticipate and extrapolate the trajectory of mobile objects when they use the third-person perspective. It is mainly due to the larger field of view provided by the position of the camera for this perspective. The user can thus better appreciate the situation and the distance because he/she has landmarks (the position of his/her head and hands) and he/she can see what happens on his/her close right and left sides.

We conclude that both perspectives are needed during the simulations composed of varying actions (like the video games) and that the switch between them could be very useful.

6. FURTHER WORKS

It could be interesting to make some test in Virtual and Augmented Reality environments. We also could add a second camera to provide stereo vision to the user. Moreover, using and lighter and wireless stuff would really increase the user comfort and thus his/her immersion.

For the third-person perspective, we could use a camera on an articulated arm following the user head rotations with the help of remote-controlled servo controllers.

Finally, it could be interesting to make some measures of the brain activity in collaboration with the Brain Mind Institute (BMI) [3] within the framework of the experimentations.

7. ACKNOWLEDGMENTS

8. REFERENCES

[1] III Richard Rouse. What's your perspective? *SIGGRAPH Comput. Graph.,* 33(3):9–12, 1999.

[2] Benjamin Lok, Samir Naik, Mary Whitton, and Jr. Frederick P. Brooks. Effects of handling real objets and self-avatar fidelity on cognitive task performance and sense of presence in virtual environments. *Presence: Teleoper. Virtual Environ.,* 12(6):615–628, 2003.

[3] Brain Mind Institute, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland *http://bmi.epfl.ch*

[4] Huahai Yang and Gary M. Olson. Exploring collaborative navigation:: the effect of perspectives on group performance. In *CVE '02: Proceedings of the 3rd international conference on Collaborative virtual environments,* pages 135-142, New York, NY, USA, 2002. ACM Press.

[5] Joshua M. Knapp and Jack M. Loomis. Limited field of view of head-mounted displays is not the cause of distance underestimation in virtual environments. *Presence,* 13(5)572–577, 2004.

[6] Mel Slater. Do avatars dream of digital sheep? Virtual people and the sense of presence. In *VR*, pages 3 – , 2002

[7] Mel Slater. Presence and the sixth sense. *Presence: Teleoper. Virtual Environ.,* 11(4):435–439, 2002.

[8] Michael M. Popp, Edna Platzer, Matthias Eichner, and Marion Schade. Walking with and without walking: perception of distance in large-scale urban areas in reality and in virtual reality. *Presence: Teleoper. Virtual Environ.*, 13(1):61–76, 2004.

[9] Peter Willemsen, Mark B. Colton, Sarah H. Creem-Regehr, and William B. Thompson. The effects of head-mounted display mechanics on distance judgements in virtual environments. In *APGV '04: Proceedings of the 1st Symposium on Applied perception in graphics and visualization*, pages 35–38 , New York, NY, USA, 2004. ACM Press.

[10] Roland Arsenault and Colin Ware. The importance of stereo and eye coupled perspective for eye-hand coordination in fish tank vr. *Presence,* 13(5):549–559, 2004

[11] Ross Messing and Frank H. Durgin. Distance perception and the visual horizon in head-mounted displays. *ACM Trans. Appl. Percept.,* 2(3):234–250, y 05.

[12] William B. Thompson, Peter Willemsen, Amy Ashurst Gooch, Sarah H. Creem-Regehr, Jack M. Loomis, and Andrew C. Beall. Does the quality of the computer graphics matter when judging distances in visually immersive environments? *Presence,* 13(5):560–571, 2004