

# Embodied Interaction using Non-Planar Projections in Immersive Virtual Reality

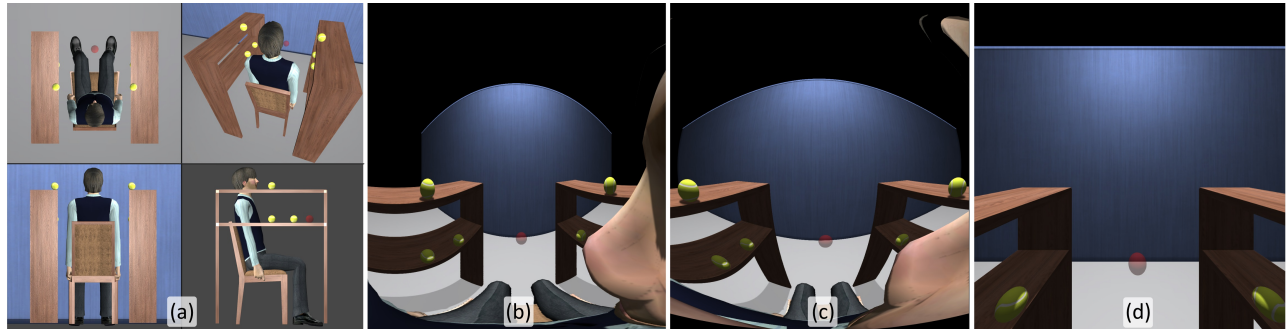
Henrique G. Debarba<sup>\*1</sup>

Sami Perrin<sup>†</sup>

Bruno Herbelin<sup>‡2</sup>

Ronan Boulic<sup>§1</sup>

<sup>1</sup>Immersive Interaction Group, <sup>2</sup>Center for Neuroprosthetics  
École Polytechnique Fédérale de Lausanne, Switzerland



**Figure 1:** Orthogonal and perspective views of the virtual environment used in the experiment (a). Render capture of the left eye for Equirectangular (b), Hammer (c) and Perspective (d) projections respectively. Notice that it is possible to see the virtual body (including the nose) with the non-planar projections.

## Abstract

In this paper we evaluate the use of non-planar projections as a means to increase the Field of View (FoV) in embodied Virtual Reality (VR). Our main goal is to bring the virtual body into the user's FoV and to understand how this affects the virtual body/environment relation and quality of interaction. Subjects wore a Head Mounted Display (HMD) and were instructed to perform a selection and docking task while using either Perspective ( $\approx 106^\circ$  vertical FoV), Hammer or Equirectangular ( $\approx 180^\circ$  vertical FoV for both) projection. The increased FoV allowed for a shorter search time as well as less head movements. However, quality of interaction was generally inferior, requiring more time to dock, increasing docking error and producing more body/environment collisions. We also assessed cybersickness and the sense of embodiment toward the virtual body through questionnaires, for which the difference between projections seemed to be less pronounced.

**CR Categories:** H.1.2 [Models and Principles]: User/Machine Systems—Human factors I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual reality;

**Keywords:** non-planar projections, sense of embodiment, immersive interaction

\*e-mail: henrique.galvandebarba@epfl.ch

†e-mail: sami.perrin@epfl.ch

‡e-mail: bruno.herbelin@epfl.ch

§e-mail: ronan.boulic@epfl.ch

## 1 Introduction

Even with the recent rise of affordable HMD, which brought VR back into the popular imaginary while approaching the mass market, delivered FoV is still a lot inferior than human capability. Oculus DK2, which is arguably the most popular current model offers a maximum of  $\approx 106^\circ$  FoV, even though in practice most users experience something around  $90^\circ$  as the FoV also depends on the eye/lenses and lenses/screen distances. On the other hand, human FoV is  $\approx 180^\circ$  horizontally, and  $\approx 135^\circ$  vertically (with a downward bias). This allows us to be visually aware of our body at all times—the fact we have one as well as its posture—, which is not the case while using an HMD and controlling a virtual body.

In this paper we explore the use of non-planar projections to address this limitation, i.e. showing more of the virtual body and environment at the cost of altering the projection we experience in the natural world. To assess this issue, we realised an experiment that evaluates how one feels and performs in a selection and docking task while controlling a virtual body in a cluttered virtual environment. Our implementation and choices of non-planar projections notably relies on cartography, which has long studied means to represent the surface of the Earth on a plane. Here we compare the regular Perspective projection ( $\approx 106^\circ$  vertical FoV), with the Hammer ( $\approx 180^\circ$  vert. FoV) and Equirectangular ( $\approx 180^\circ$  vert. FoV) non-planar projections (Figure 1). During the development phase, we tested several non-planar projections and selected Hammer and Equirectangular, being consistent with the positive results presented in [Ardouin et al. 2013]. Equirectangular is a cylindrical projection with a property of equidistance, while Hammer is an azimuthal projection with a property of equal-area. More information on these projections can be found in [Kennedy and Koop 1994].

In order to assess whether seeing the virtual body more often increase subject's identification with it, we evaluate the reported sense of Embodiment with a questionnaire considering: (a) sense of agency, i.e. feeling of control over the virtual body; and (b) sense of body ownership, i.e. feeling that a virtual body is one's own body. The questionnaire was based on the one presented in [Slater et al. 2010]. A simulation sickness questionnaire (SSQ) was also administered [Kennedy et al. 1993].

## 2 Related Work

Non-planar projection have been studied as alternatives to perspective projection in order to increase the FoV of a camera (i.e. perspective is limited to  $< 180^\circ$ ), and/or to keep better proportionality for information:rendering area ratio (i.e. large FoV perspective will render most of the information in a small region at the center of the image). This allows for the presentation of more environmental information with reduced effort (e.g. less camera movements). Early works have approached non-planar projections on conventional displays, while more recent work have also explored this matter in an HMD and augmented reality context.

[Glaeser and Gröller 1999] proposed the use of such mappings as an alternative to perspective rendering when a wide FoV is required. They argue that the distortions introduced to the image by a non-planar projection are less detrimental than the ones introduced by a wide FoV perspective projection when rendering in a desktop. On a similar direction, [Ardouin et al. 2013] evaluates how subjects perform in a navigation task for various  $360^\circ$  non-planar projections while the VE is viewed on a 22 inches screen. They found advantages and favoured the usage of Hammer and Equirectangular projections. Furthermore, [Mulloni et al. 2012] explores how different panoramic images in a computer screen affect subjects ability to find and correctly point to objects at their surroundings.

Closer to our context, a few studies also approached non-planar projections rendered by HMDs. [Ardouin et al. 2012] proposed a system that delivers  $360^\circ$  of horizontal FoV to the user. The image is captured by a camera from the top of the users head and is fed to the HMD, providing easy and intuitive control of the point of view. However, the project is mostly conceptual and the evaluation was solely based on user impressions. [Orlosky et al. 2014] brings a deeper study using a pair of cameras and  $233^\circ$  fisheye lenses to evaluate perception of objects in the periphery of vision

To the best of our knowledge, no past work explored the use of non-planar projections in an immersive VR setup and from the perspective of embodied interaction.

## 3 Projection Implementation

The experiment was developed with the Unity game engine. To obtain the non-planar projection images, we render a  $360^\circ$  image using six  $90^\circ$  perspective cameras. We then map the 6 rendered images into a high density cube mesh. This cube mesh is then mapped into a sphere by normalising each vertex distance from the centre of mass of the mesh. Finally, the resulting spherical mesh is transformed into a plane with a cartographic projection equation (latitude and longitude coordinates into x and y coordinates). This approach is similar to the one employed by [Bourke 2009] to project on a hemispherical surface.

## 4 Experiment

### 4.1 Virtual Environment and Task

The VE consisted of a chair, on which the virtual body is seated. Additionally, a pair of two levels shelves were placed along each side of the chair, where targets could appear at 6 predefined positions (2 targets within and 4 beyond a  $90^\circ$  FoV while looking forward). The target has the shape of a tennis ball with 6.7 cm diameter. An additional docking volume is shown in front of the virtual body, it has the same size as the targets and is rendered with transparency. An aerial view is show in Figure 3.

The task consisted of reaching and docking targets with the dominant hand (as indicated by the subject), in each trial: a target appears; the subject searches for it; then (s)he moves the end effector in order to intersect the target; (s)he selects the target by holding a trigger button and translates it toward the semi-transparent docking volume; finally, (s)he releases the target by releasing the trigger button; the next target appears. Subjects were asked to perform the task as fast and accurately as they could. There were a total of 6 different targets, which were repeated 4 times, for a total of 24 *trials* in each *block*. Subjects performed a *block* for each tested projection, yielding a total of 3 *blocks*. If no interpenetration between target and docking volume occurs (i.e. docking error bigger than 6.7 cm), the trial is marked as failed and has to be repeated by the end of the block.

### 4.2 Tracking and motion

A *PhaseSpace ImpulseX2* with 14 cameras is used to track the position of 4 LED markers attached to hands and elbows. To retrieve hands orientation and allow for input selection, a pair of *PS Move* controllers are used. They communicate with a *Playstation 3*, which uses the software *Move.me* to stream the controllers data to our program. The arms of the virtual body are driven by inverse kinematics, which defines a posture relying on the 6 degrees of freedom of the tracked hands as well as the position of the tracked elbows. LED markers were also added to the chair for a calibration step (match real and virtual chairs). The subjects were asked (and assumed) to keep their posture and avoid moving trunk and legs – a predefined sitting posture is applied to the virtual body for these body parts. For the experiment, a gender matching virtual body was used, and its height was scaled to match the height of the subject.

An *Oculus DK 2* is used as display, the Oculus also provides drift-free orientation based on its inertial sensors and optical tracking. This orientation is used to rotate the virtual camera and the head of the virtual body. No position tracking was used, thus the point of view rotated around a predefined pivot position in the virtual body neck. The setup is shown in Figure 2.

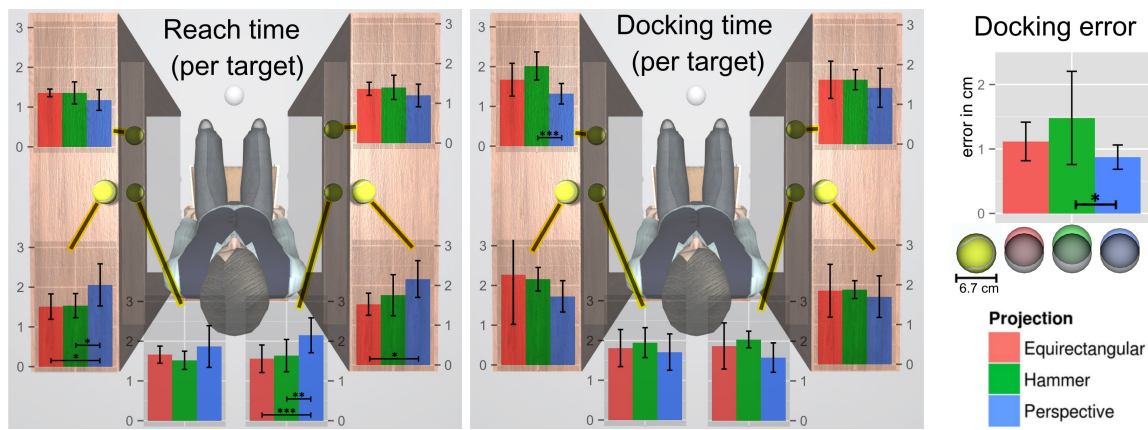


Figure 2: Overview of the system setup.

### 4.3 Analysis and results

The experiment followed a within-subject design with random projection order. A total of 6 subjects aged from 17 to 25 participated on the experiment (1 female). *Projection* was the main factor to be controlled. *Target position* was also treated as a factor for the time related responses.

Time and precision: we consider reach time (RT) – time until the selection of a target –, dock time (DT) – time from selection to dock –, dock error (DE) – error in cm between predefined docking position and actually docked position. These values are computed by



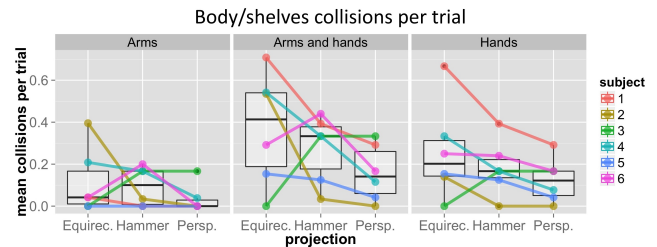
**Figure 3:** Time to reach (left) and dock (center) in seconds, and docking precision (right) in centimeters. Reach time with non-planar projections was significantly shorter for most of the lateral targets. Perspective presented time advantage for docking, as well as smaller docking error. Error bars indicates the estimated standard deviation. Significance results were computed with Tukey HSD test. ‘\*’, ‘\*\*’ and ‘\*\*\*’ indicates  $p < .05$ ,  $p < .01$  and  $p < .001$  respectively.

taking the median of the successful trials for each combination of *subject*, *target position* (except for DE) and *projection*. Subject 2 had difficulty to adapt to the non-planar projections and is not considered for RT, DT and DE analyses. Subject 2 had to perform 104 trials to successfully dock a total of 48 targets (i.e. docking error of less than 6.7 cm) for Equirectangular and Hammer projections together. Other subjects had to repeat a maximum of 4 trials during the whole experiment.

For RT and DT, statistical analysis was carried with two-way ANOVA with *projection* and *target position* as independent variables. RT presented an interaction between *projection* and *target position* ( $F(10, 40) = 3.9, p < 0.001$ ), indicating a tradeoff for these factors. For DT only target position yield a significant difference ( $F(5, 20) = 6.848, p < 0.001$ ). We further analysed the effect of *projection* to RT and DT per *target position* using Tukey HSD test. For DT, difference was significant for one of the frontal targets in favour of Perspective as compared to Hammer. For RT, the difference was significant for 3 of the lateral targets in favour of Equirectangular as compared to Perspective, and for 2 of the later targets in favour of Hammer as compared to Perspective (details presented in Figure 3). Further analysis with Tukey HSD to the mean head movement per trial (in radians) shows that subjects performed significantly less head turns with Equirectangular ( $M = .57, SD = .11$ ) and Hammer ( $M = .69, SD = .16$ ) projections as compared to Perspective ( $M = .83, SD = .10$ , with  $p < .001$  and  $p < .011$  respectively). Suggesting that subjects took advantage of the increased FoV, which in turn led to a decrease in search time.

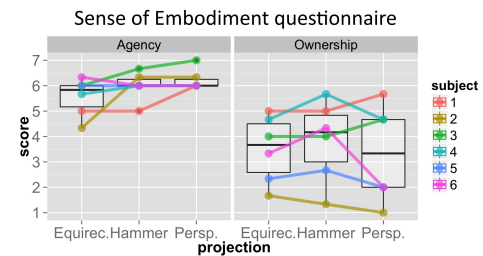
To analyse DE we used Tukey HSD corrected for multiple comparisons with *projection* as the only independent variable, which has shown that Perspective performed significantly better than Hammer. Figure 3 shows the RT, DT and DE means and standard deviations for each projection.

Collisions with environment: we assess subjects’ understanding of the environment and its relation with the virtual body by considering the mean of collisions per trial (MCol) between virtual body and shelves while performing the task. We compute MCol for the arms (upper arm and forearm), hands, and arms/hands together. If simultaneous collisions happens, only one collision is considered. Analysis was carried with Wilcoxon signed rank test with Holm-Bonferroni correction. The increase in collisions with Hammer and Equirectangular projections were not statistically significant



**Figure 4:** Collisions of the body with shelves per subject. Everyone but subject 3 presented less body/shelves collisions while using perspective projection.

when compared to Perspective projection ( $p = 0.059$  for corrected  $\alpha = .0167$  and  $p = .156$  for corrected  $\alpha = .025$  respectively). MCol per subject for each *projection* is shown in Figure 4.



**Figure 5:** Reported sense of agency and sense of body ownership.

Sense of embodiment: we assess the senses of Agency (AG) and Body Ownership (BO) with a questionnaire. Agency score is the mean of 3 questions, whether the subject felt: to be in control of the VB, not to be in control of the VB, and the VB was responsive to his/her movements. Ownership score is the mean of 3 questions, whether the subject felt: that the VB was his/her body, that the VB was not his/her body, to be wearing VB clothes. Negative questions had their score inverted before taking the mean score.

Statistical analysis was carried with Wilcoxon signed-rank test and Holm-Bonferroni correction. No significant difference was found for both, agency and ownership (smaller p-value at  $p = .134$  for

corrected  $\alpha = .0167$ ). As the virtual body can be seen the whole time, we were expecting an increase with non-planar projections. However, these results suggest that the non-planar projections had no consistent effect to the reported sense of embodiment (Figure 5).

Cybersickness: was assessed with the simulation sickness questionnaire (SSQ). The mean score for Equirectangular, Hammer and Perspective projections were 1.83, 1.67 and 1.5 respectively. No statistically significant difference was found with Wilcoxon signed-rank test and Holm-Bonferroni correction (smaller p-value at  $p = .572$  for corrected  $\alpha = .0167$ ).

## 5 Discussion

On the one hand, our results indicate that the large FoV non-planar projections may increase reaching performance when interacting with targets beyond the perspective FoV, leading to less head rotations. On the other hand, the time needed to dock as well as the docking precision were reduced, even though the difference was not always statistically significant.

Additionally, the amount of collisions between the virtual body and environment obstacles have increased with the non-planar projections. Although the difference was not significant, we expected that the visual feedback of the body would improve subject's perception of the virtual body/environment relation, which was clearly not the case. Even with the reduced FoV, the spatial model one creates in perspective projection seems to be more effective than the more constant – but distorted – visual feedback that the large FoV non-planar projections provide. This did not prevent subjects from succeeding in the proposed experiment, but may point a poor translation of structural visual information into movement planning. Alternatively, the distorted feedback may negatively affect the recalibration of proprioception, by integrating sensorial input into incorrect postural information.

Finally, the non-planar projections were not detrimental to the reported senses of agency and of body ownership. They also did not elicit stronger cybersickness as compared to Perspective. With these results, and given that the performance gain was limited, we believe that non-planar projections could be used for quick inspection of the environment. In such scenario, the user would switch from perspective to a non-planar projection and back as an alternative to rotating the head. This would allow for the efficient gathering of the structural information of the environment, as demonstrated in our time to reach results.

Furthermore, we were initially expecting to find differences between Hammer and Equirectangular projections. More specifically, we had the expectation that Equirectangular projection would provide more accurate structural information, given that it preserves some straight lines (Figure 1b). Such information could be useful to prevent collisions and increase docking precision. As both non-planar projections presented very similar performance, this hypothesis could not be verified. As a matter of fact, our results are in line with [Ardouin et al. 2013], in which navigation performance and subjective evaluation ranked Equirectangular and Hammer projections to be very close as compared to the other projections the authors have tested.

## 6 Conclusion and Future Work

We evaluated non-planar projections in the context of a selection and docking task with a virtual body in IVR. Our results demonstrate some advantages, such as the reduction in search time. Other aspects, such as the SSQ and the Embodiment questionnaire responses seem not to be affected by the change in projection method.

On the other hand, quality of interaction seems to be consistently lower, with increased time to dock and reduced docking precision. Finally, the spatial understanding of the virtual body/environment relation seems impaired as all but one subject presented increase in the mean collisions per trial with the non-planar projections.

Furthermore, there are other basic 3D user interaction tasks that should be considered in the future, such as navigation and finer manipulation. It might be the case that non-planar projections are less suitable for fine manipulation of objects (e.g. involving the fingers). In addition, SSQ scores could be altered in a navigation task, given that change in visual flow promoted by non-planar projections would be coupled with additional forms of movement.

Finally, we point to the fact that subjects only used each projection for 2 to 5 minutes in our experiment. A long term adaptation might play a strong role on performance, and no related work explored this venue yet.

## Acknowledgements

We would like to thank the reviewers for their time and valuable opinion and suggestions. This research is supported by the SNF project No. 140910.

## References

- ARDOUIN, J., LÉCUYER, A., MARCHAL, M., Riant, C., AND MARCHAND, E. 2012. Flyviz: a novel display device to provide humans with 360 vision by coupling catadioptric camera with hmd. In *Proceedings of the 18th ACM symposium on Virtual reality software and technology*, ACM, 41–44.
- ARDOUIN, J., LÉCUYER, A., MARCHAL, M., AND MARCHAND, E. 2013. Navigating in virtual environments with 360 omnidirectional rendering. In *3D User Interfaces (3DUI), 2013 IEEE Symposium on*, IEEE, 95–98.
- BOURKE, P. 2009. idome: Immersive gaming with the unity game engine. In *Proceedings of the Computer Games & Allied Technology*, 265–272.
- GLAESER, G., AND GRÖLLER, E. 1999. Fast generation of curved perspectives for ultra-wide-angle lenses in vr applications. *The Visual Computer* 15, 7-8, 365–376.
- KENNEDY, M., AND KOOP, S. 1994. Understanding map projections. *GIS by ESRI, Environmental System Research Institute*.
- KENNEDY, R. S., LANE, N. E., BERBAUM, K. S., AND LILIEN-THAL, M. G. 1993. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology* 3, 3, 203–220.
- MULLONI, A., SEICHTER, H., DÜNSER, A., BAUDISCH, P., AND SCHMALSTIEG, D. 2012. 360 panoramic overviews for location-based services. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 2565–2568.
- ORLOSZY, J., WU, Q., KIYOKAWA, K., TAKEMURA, H., AND NITSCHKE, C. 2014. Fisheye vision: Peripheral spatial compression for improved field of view in head mounted displays. In *Proceedings of the 2Nd ACM Symposium on Spatial User Interaction*, ACM, New York, NY, USA, SUI '14, ACM, 54–61.
- SLATER, M., SPANLANG, B., SANCHEZ-VIVES, M. V., AND BLANKE, O. 2010. First person experience of body transfer in virtual reality. *PloS one* 5, 5, e10564.