Navigating in Virtual Environments with 360° Omnidirectional Rendering

Jérôme Ardouin* ESIEA - INSA - Inria
Anatole Lécuyer† Inria - IRISA
Maud Marchal‡ INSA - IRISA - Inria
Eric Marchand§ Université de Rennes 1 - IRISA - Inria

ABSTRACT

Typical field-of-view (fov) of visual feedback in virtual reality applications is generally limited. In some cases, e.g. in videogames, the provided fov can be artificially increased, using simple perspective projection methods. In this paper, we design and evaluate different visualization techniques, inspired by the cartography domain, for navigating in virtual environments (VE) with a 360° horizontal fov of the scene. We have conducted an evaluation of different methods compared to a rendering method of reference, i.e. a perspective projection, in a basic navigation task. Our results confirm that using any omnidirectional rendering method could lead to more efficient navigation in terms of average task completion time. Among the different 360° projection methods, the subjective preference was significantly given to a cylindrical projection method (equirectangular). Taken together, our results suggest that omnidirectional rendering could be used in virtual reality applications in which fast navigation or full and rapid visual exploration are important. They pave the way to novel kinds of visual cues and visual rendering methods in virtual reality.

Index Terms: H.5.2 [Information Interfaces and Presentation]: User Interfaces—user-centered design; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual Reality

1 INTRODUCTION

Fov of immediate visual feedback is generally constrained when navigating in VE. Typical visual rendering in recent video games relies on classical perspective projection, providing a horizontal field of view usually around 108°. This can become frustrating in some contexts or applications: for instance when exploring the VE and aiming at targets under a strong time constraint. It can lead to many rotations (yaw) to search and locate 3D objects.

In the literature, various technical solutions have been proposed to provide an extended field of view through non planar projection [3, 2, 8, 11, 6]. Surprisingly, the current techniques do not take advantage of the projection methods developed in the cartography field [10] or in omnidirectional computer vision [1, 4]. In these domains, numerous methods have been proposed to display and apprehend a large set of (360°) information onto a 2D plane, such as for sailors or pilots. Also, to the authors' best knowledge, no user evaluation has been conducted about how the visual information is perceived in the case of full 360° rendering.

Therefore, in this paper, we propose to investigate the design and evaluation of lateral 360° visual fov, for the purpose of real-time navigation in VEs. We notably propose to use projection methods inspired by the cartography field (see Figure 1), and to exploit them for real-time omnidirectional rendering in VE. As a second contribution we have conducted an experiment on the use of such omnidirectional visual feedback in a navigation task, to compare it with current state-of-the-art visual feedback.

*e-mail: ardouin@esiea-ouest.fr
†e-mail: alecuyer@irisa.fr
‡e-mail: mmarchal@irisa.fr
§e-mail: marchand@irisa.fr

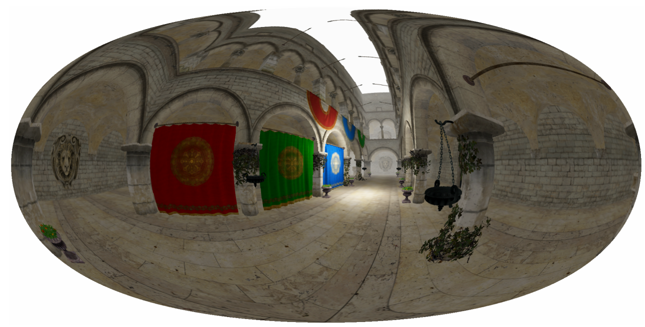

Figure 1: First-person navigation in a VE with a 360° omnidirectional rendering using a Hammer projection method rendered in real-time.

In the remainder of this paper, related work in omnidirectional real-time rendering is presented in section 2. Our approach and the different projection methods are detailed in section 3. The user study is described in section 4. The paper ends with a general discussion and a conclusion.

2 RELATED WORK

Traditional approaches for real-time rendering are based on the pinhole camera model. This model relies on planar projections. To increase the user's fov of the scene, the distance between the center of projection and the projection plane is decreased. Problems arise for wide fov (>120°) for which planar projection makes the visual information very distorted and difficult to be interpreted by the user. A second major problem of planar projection is that it cannot handle fov superior or equal to 180° due to simple geometric limitations. In such rendering approaches, the user accesses visual information from his surroundings by rotating the point of view. The full 360° fov is therefore accessed sequentially.

Augmented or virtual reality systems provide various solutions for a user to access the visual representation of his surroundings. [7] proposes a zooming method between the Augmented Reality view and an egocentric panoramic view. Cave and HMD coupled with a head tracker enable the user to scan the virtual surroundings in a way very similar to reality [3, 2]. In various situations, planar projection cannot satisfy the application needs, for example when a non-planar display surface such as those of domes or hemispherical displays, is used [2].

Omnidirectional rendering of the VE can also be used on traditional (i.e., non surrounding) displays like monitors. In this case, the goal is to give a VE representation that provides more information than classical approaches, trying to better match the real human fov (instead of the traditional 75°x108° fov of recent videogames) or extending fov up to a full omnidirectional representation (360°). Various approaches can solve the technical problem of computing real-time non planar projection. Oros [8] has proposed a geometrical approach in which vertices are processed before planar projection. However, this method, cannot handle full 360° and it requires highly tessellated geometry (a common limitation of geometry-based approaches) [12]. If the final projection uses a single center of projection [12, 11], an image-based approach is an interesting alternative. In this case, the VE is first rendered to 6 offline buffers

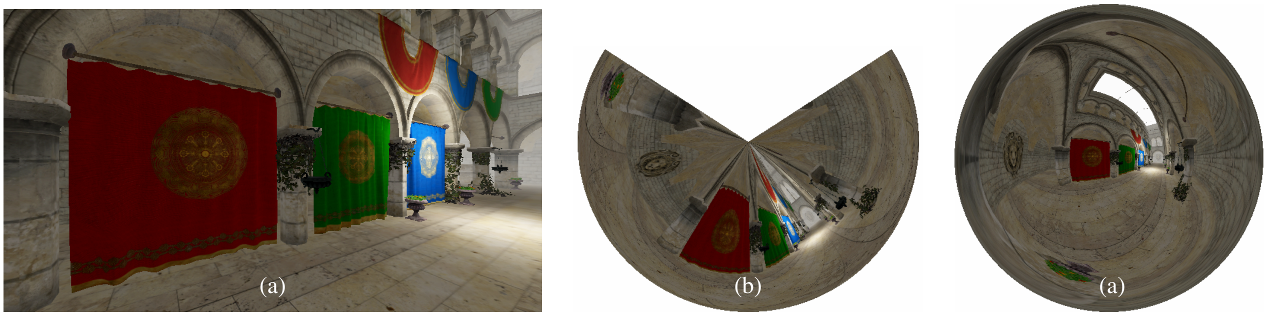

Figure 2: (a) Perspective projection, (b) Albers conic Projection, (c) Azimuthal equidistant projection.

with 6 standard planar projections [5]. The final image is then generated by using a lookup function in conjunction with the desired projection equation. This accommodates modern GPU particularly well as this kind of operation has been largely optimized (notably for environment cube mapping). Van Oortmerssen tested this in two modified versions of the famous first-person shooter game Quake: FisheyeQuake and PanQuake [13]. However, Lorenz et al. [6] have stressed the fact that this rendering approach does not benefit from improvements of modern hardware rasterizers such as anti-aliasing or anisotropic filtering. This approach also fails in handling projections with multiple centers. The general method of handling the projection consists here in an approximation with multiple planar projections [11, 6].

3 Cartography Projection Methods for Real-time 360° Omnidirectional Rendering

Some methods reviewed in the related work are already able to extend the fov up to 360° by using intuitive formulation and implementation of the mapping equation. But there are actually other projections methods in the cartography field that have not yet been applied to real-time rendering. In this section, we propose to adapt relevant projection methods from cartography to real-time 3D rendering. In our study, we have intentionally only considered the projection methods able to provide a full lateral 360° fov. We want here to provide the maximum visual information, at least in the lateral direction. But, of course, the whole approach could also apply to a 180° fov rendering, which would match the physiological characteristics of human vision.

The mathematical formulation of the problem consists in mapping all the space directions onto a plane. Mapping one direction of 3D space is equivalent to mapping a point on a unit sphere onto this plane (due to the unique representation of a direction vector in 3D space and of the points lying on the unit sphere). This mapping problem has been widely studied by mathematicians and cartographers [10]. Various mapping methods have been proposed, using different developable surfaces (cylinder, cone, plane) and exhibiting particular geometric properties on area, distance or angles (eg. equal area, equidistant or conformal property). In the cartography literature, these projection equations are expressed as functions that map a point from the unit sphere to one point onto the plane. Although these functions are not necessarily bijective, the considered ones have to. Since we work on the final image, the reciprocals of these functions need to be used (the points of coordinate (x,y) in the final image plane is mapped to its corresponding direction/point on the unit sphere). Also, the direction in space is usually not expressed as a cartesian vector but as polar coordinates (λ,ϕ). The corresponding cartesian vector is simply deduced using equation (1):

$$\begin{cases} X &= \sin\phi\cos\lambda \\ Y &= \sin\lambda \\ Z &= -\cos\phi\cos\lambda \end{cases} \quad (1)$$

3.1 Experimented Projections

The chosen projection methods have been selected to feature interesting properties like equidistance or equality of area and to cover the main kinds of projection used in cartography field. The three main kinds of developable surface (plane (e.g. Azimuthal projection), cylinder and cone) are considered. The mathematical formulations used are not given but can be found easily in references [10].

Perspective projection. The first projection method is the traditional perspective projection (Figure 2.a). It is not meant for 360° and omnidirectional rendering, and will serve here as a reference for future comparisons with the other projections. As it is natively supported by graphic hardware, the reciprocal form of the projection is not needed, and it can be implemented by using traditional, forward rendering. In our experiment, the vertical fov is set to 75° (equal to the standard upper limit found in video games). The aspect ratio of the rendering viewport determines the horizontal fov. With a 16:10 aspect ratio, the computed horizontal fov is about 102°.

Equirectangular projection (cylindrical). The equirectangular projection (Figure 3) is the most widespread projection, used for instance in panoramic photography. In the cartography field, it is well appreciated for its clarity and simplicity. For full representation of surroundings, it leads to a rectangle of 2:1 aspect ratio that fits well modern screens with 16:10 or 16:9 aspect ratios. It has the equidistant property.

Hammer (azimuthal). The Hammer projection (Figure 1) maps the sphere to an ellipse with 2:1 proportion. It has the equal area properties. Compared to equirectangular projection it is therefore less subject to shape distortion.

Albers equal area conic projection (conical). Among the different conic projection methods, the Albers Conic projection (Figure 2.b) has the property to preserve area and the projection preserves well the shapes between the two parameterized parallels. However, it can be noticed that it is not possible to represent the full surroundings as the pole representation fails. But a full 360° horizontal fov can still be obtained.

Azimuthal Equidistant. The Azimuthal equidistant projection (see Figure 2.c) is particularly interesting as it can be found in real world. Some fisheye lens have indeed this mapping function. But it is physically limited to fov just above 180°. With a synthesized 3D image, the implementation can generate a full omnidirectional representation which is not possible in real world. The projection maintains a constant radial scale from the central point. The resulting images are free of distortions at the central point, but a huge deformation occurs at the borders.

3.2 Implementation for Real-time Rendering

In order to have a flexible framework, an image-based approach has been adopted [13, 2]. It can be decomposed in two main steps:

1. The whole scene is rendered into 6 textures and stored using a cube map. In our implementation, the whole scene is rasterized 6 times with the appropriate perceptive projection matrices.

2. A full screen quad is rendered with a specially developed fragment shader. The shader implements the cartography projection equations to associate one coordinate in the viewport to one direction in space. This direction is then used as cube map texture coordinate to address the cube map rendered at step 1, giving the final pixel color. Cube map coordinates decoding provides an efficient way to associate one direction on the unit sphere to the rendered view.

The whole application prototype has been developed in C language and OpenGL API. GLSL was the adopted language for our shaders implementation.

3.3 Performances

Although pure performance was not the topic of our study, our implementation has been benchmarked. The target platform is Windows 7, running on Intel i3 2120 running at 3.3GHz with 4Go RAM and an NVidia GeForce 560 GTX graphic board. During the benchmark, vertical synchronization was disabled.

Table 1: Performances of our implemented algorithms with two different scene complexities (unit: frames per second).

Projection method	Simple VE	Rich VE
Perspective	1365	322
Equirectangular	898	73
Hammer	812	73
Albers Equal Area Conic	914	72
Azimuthal Equidistant	940	73

Table 1 shows the observed frame rates for the different projections and two different scenes displayed in Figures 3 and 4. According to the figures, the complexity of the used projection equation has minor impact on the performance. The difference between the two scenes is explained by the overhead in geometry complexity (32 triangles in the simple VE, versus 178324 in the rich VE). The significant difference between the perspective projection and the 360° projections is directly related to the implementation of the cube map generation (the scene is rasterized 6 times in six 1024×1024 textures). This could be significantly improved by using layered rendering and a proper geometry shader [9]. Figure 3 shows the additional visual information available compared to the standard perspective projection and to the real human binocular vision (for the equirectangular method).

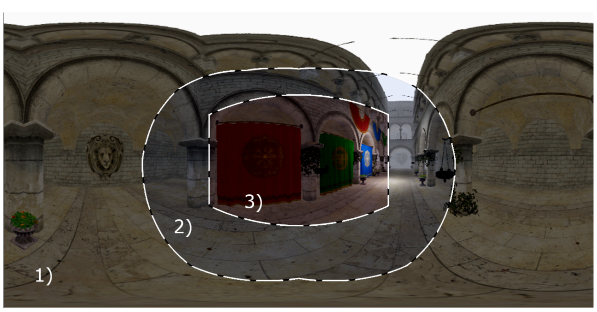

Figure 3: Comparison of the fov provided by 1) equirectangular projection, 2) human natural binocular vision (light blue), 3) classic perspective projection (light red) with 16:9 aspect ratio (75°×107.5°).

4 Evaluation

The objective of our evaluation was to assess if performing a task of object collection in a VE could be enhanced by providing a 360° view of the environment. We have compared our different projection methods against the traditional perspective projection with a standard fov of $75° \times 102°$. The quantitative evaluation was completed with a subjective questionnaire.

Population. 15 participants, aged from 21 to 44, took part in the experiment. All of them had at prior experience in video gaming.

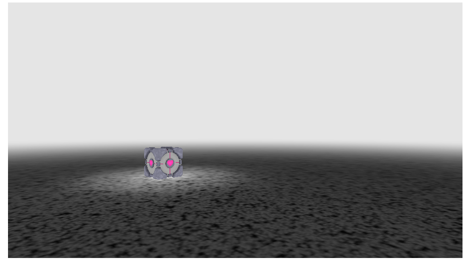

Figure 4: Simple VE used for the basic navigation task.

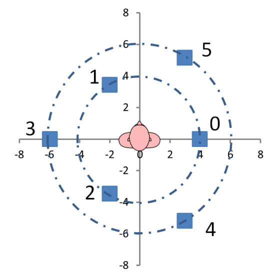

Figure 5: Top view of boxes positions used for the object collection task. The user faces the +y axis.

Experimental apparatus. A gamepad was used as a navigation interface with a simplified version of the common method found in video games: left analog stick is used to move forward and backward, right analog stick to rotate left and right. Strafe (lateral translation) and vertical aiming were locked. The evaluation software was running with the configuration described in section 3.3, completed with a 22 inches LCD desktop monitor. The panel provided a resolution of 1680×1050 with $16:10$ aspect ratio.

Procedure. The evaluation was split in two parts. The first part was aiming at collecting quantitative data on user performance in a box collection task. This part of the evaluation took place in a neutral VE, with only a floor textured with a noise pattern (Figure 4). The participant had to collect 6 boxes, one after another, with a position randomly chosen among the location described by Figure 5. The 5 projection methods described in section 3.1 were randomly tested. Each box was collected 3 times leading to a combination of 90 trials for each participant (5 methods×6 positions×3 trials). The position of the user was reinitialized after each trial.

The second part of the evaluation consisted in a subjective evaluation of the proposed method. A richer VE was used for this part (Atrium Sponza Palace). The model has been modified by removing few parts to break its symmetry (Figure 2.a). The participant was asked to fill a subjective questionnaire. During the questionnaire, the user was free to switch from one projection method to another. Short sequences of the two parts of the evaluation are given in the accompanying video.

Collected Data. For the first part, the time to collect each box is recorded. For the subjective questionnaire we used 6 criteria for the 5 projection methods scored with a 7 point scale. These criteria were: (1) Visual fatigue, (2) Ease of movement in the VE, (3) Aesthetic of the rendering, (4) Strangeness of the rendering, (5) Realism of the rendering, (6) Global appreciation.

Results. Concerning the completion times (see table 2), we performed a Shapiro test that rejected the normality hypothesis on the data distribution. Thus, we used a non-parametric Friedman test for differences among the different projections . Post-hoc comparisons were performed using Wilcoxon signed-rank tests with a threshold of 0.05 for significance. Reported p-values are adjusted for multiple comparisons. We found that the time needed to reach a box differed significantly across the 5 projections ($\chi^2 = 5.92$, $p < 0.001$). Post-hoc analysis revealed that it was significantly faster to perform the task with the equirectangular projection compared to perspective projection ($p < 0.001$) and conic projection

Table 2: Completion time results (mean completion times to reach the boxes in seconds and standard deviation).

Projection method	all boxes	box 0	box 1	box 2	box 3	box 4	box 5
Perspective	3.14 (1.39)	3.64 (1.25)	1.52 (0.36)	3.30 (0.92)	3.99 (1.59)	4.18 (0.98)	2.24 (0.45)
Equirec-tangular	2.69 (1.56)	2.15 (0.69)	1.69 (1.59)	3.08 (0.80)	2.83 (1.14)	3.64 (0.74)	2.76 (2.63)
Hammer	2.61 (1.38)	2.20 (1.11)	1.54 (0.71)	3.16 (1.02)	2.79 (1.90)	3.53 (0.80)	2.47 (1.46)
Conic	3.07 (1.38)	2.87 (2.71)	2.09 (2.03)	3.43 (1.12)	3.10 (1.73)	4.25 (1.72)	2.66 (1.22)
Azimuthal	3.04 (2.70)	2.69 (2.14)	1.44 (0.61)	3.90 (1.91)	3.44 (3.53)	4.48 (4.17)	2.12 (0.40)

($p = 0.02$). The task was also performed faster with the Hammer projection than with perspective projection ($p < 0.001$) and conic projection ($p = 0.001$). Finally the task was performed faster with the azimuthal projection compared to the perspective projection ($p = 0.001$). Taking each box independently, we found further results. Boxes indexes are given in Figure 5 for reference. We found particularly significant effect for the projection condition for boxes #3 and #4 ($\chi^2 = 5.70, p < 0.001$ and $\chi^2 = 4.61, p < 0.001$ respectively). The time needed to reach the box was significantly higher for the perspective projection compared to almost all other projections ($p < 0.001$ for box #3 for all the projections, $p < 0.001$ for box #4 for equirectangular and Hammer projections, $p = 0.02$ for box #4 for azimuthal projection).

Concerning the subjective questionnaire, we performed also a Friedman test on the differences between the different projections. We found a significant effect for all the criteria: visual fatigue ($\chi^2 = 3.29, p = 0.009$), ease of displacement ($\chi^2 = 4.12, p < 0.001$), aesthetic rendering ($\chi^2 = 5.62, p < 0.001$), strangeness ($\chi^2 = 6.03, p < 0.001$), realism ($\chi^2 = 5.90, p < 0.001$) and global appreciation ($\chi^2 = 5.90, p < 0.001$). Post-hoc analysis results are summarized in table 3.

Table 3: Post-hoc analysis for the subjective questionnaire. The criterion name in a cell means that there is a significant effect between the two conditions, the best one is in the raw.

>	Equirec.	Hammer	Conic	Azimuthal
Persp.	Aesthetic Strangeness Realism Global app.	Aesthetic Strangeness Realism Global app.	Visual fatigue Aesthetic Strangeness Realism Global app.	Ease of disp. Aesthetic Strangeness Realism
Equirec.			Aesthetic Realism Global app.	Ease of disp.
Hammer			Strangeness Realism Global app.	

Discussion. Overall, the equirectangular, Hammer and azimuthal projection improved the performances of the user by reducing the overall time to collect the boxes. The boxes #3 and #4, which were out of the field of view of the perspective projection method, were particularly faster to reach for the user when an omnidirectional projection was used. The Albers conic projection did not perform as well as the others 360° projection methods. The particular, pie shaped, distortions induced by this projection could explain the results. Investigation on the learning time for this projection against the others could be of interest.

The results for the subjective questionnaire show that the perspective projection was globally preferred over the others. User's familiarity with this kind of projection, faced in everyday life (videogames, photography), could explain this preference as no learning effort is necessary to get comfortable with it. Among the 360° projection method, the equirectangular and Hammer projection got the preference of the users. In scenarios where 360° omnidirectional vision is recommended, it would therefore be preferable to use them.

5 Conclusion

In this paper, we have designed and tested real-time omnidirectional rendering methods for a navigation task. We were inspired by projection methods designed by cartographers and mathematicians. These equations initially designed for mapping the earth, can indeed be easily adapted to omnidirectional real-time rendering with an image-based approach. They were applied to a VR interactive navigation context with a 360° lateral fov. Our implementation runs smoothly even with complex and realistic 3D scenes.

We have conducted a user study on the influence of such omnidirectional vision for a simple navigation task in virtual reality. Our results show a significant improvement in performance (time needed to collect 3D objects in the virtual scene) when using any omnidirectional projection. The participants were able to localize and reach the targets more rapidly with a 360° lateral fov. Among the different omnidirectional projection methods, a subjective preference was found for the equirectangular one. Such omnidirectional rendering could therefore notably be used in virtual reality applications in which rapid exploration is important, or when maximum visibility is required.

Future work could first focus on testing these techniques with other values of fov (e.g. 180° laterally). Second, more evaluations of the different projection methods with other tasks (object manipulation) or contexts (3D navigation without gravity) would be of interest. Other studies on the influence of the properties of the projection (equidistant/conformal/equal area) on user perception and user's spatial cognition (wayfinding, mental map) should also be carried out.

References

[1] J. Barreto and H. Araujo. Issues on the geometry of central catadioptric image formation. In *CVPR*, volume 2, pages 422–427, 2001.

[2] P. Bourke. Spherical mirror: a new approach to hemispherical dome projection. In *GRAPHITE*, pages 281–284, 2005.

[3] C. Cruz-Neira, D. J. Sandin, and T. A. DeFanti. Surround-screen projection-based virtual reality: the design and implementation of the cave. In *SIGGRAPH*, pages 135–142, 1993.

[4] C. Geyer and K. Daniilidis. A unifying theory for central panoramic systems and practical applications. In *ECCV*, pages 445–461, 2000.

[5] N. Greene. Environment mapping and other applications of world projections. *IEEE Computer Graphics and Applications*, 6(11):21–29, 1986.

[6] H. Lorenz and J. Döllner. Real-time piecewise perspective projections. In *GRAPP*, pages 147–155, 2009.

[7] A. Mulloni, H. Seichter, A. Dünser, P. Baudisch, and D. Schmalstieg. 360° panoramic overviews for location-based services. In *ACM CHI*, pages 2565–2568, 2012.

[8] D. Oros. A conceptual and practical look into spherical curvilinear projection, 2002.

[9] S. Patidar, S. Bhattacharjee, J. Singh, and P. Narayanan. Exploiting the shader model 4.0 architecture. Technical report, IIIT Hyderabad, 2006.

[10] J. Snyder. *Flattening the Earth: Two Thousand Years of Map Projections*. University of Chicago Press, 1997.

[11] M. Trapp and J. Döllner. A generalization approach for 3d viewing deformations of single-center projections. In *GRAPP*, pages 163–170, 2008.

[12] M. Trapp, H. Lorenz, and J. Döllner. Interactive stereo rendering for non-planar projections of 3d virtual environments. In *GRAPP*, pages 199–204, 2009.

[13] W. van Oortmerssen. Fisheyequake. http://strlen.com/gfxengine/fisheyequake/.