Adjusting Stereoscopic Parameters by Evaluating the Point of Regard in a Virtual Environment

Abstract

Despite the growth in research and development in the area of virtual reality over the past few years, virtual worlds do not yet convey a feeling of presence that matches reality. This is particularly due to the difference in visual perception of flat images as compared to actual 3D. We studied the impact of two parameters of the stereoscopic configuration, namely, the inter-camera distance (ICD) and the presence of a depth of field blur (DOF blur). We conducted an experiment involving 18 participants in order to evaluate this impact, based on both subjective and objective criteria. We examined six configurations which differed in the presence or absence of DOF blur and the value of the ICD: fixed and equal to the anatomical interpupillary distance, fixed and chosen by the participant, or variable, depending on the depth of the viewer's point of regard (POR). The DOF blur and variable ICD require the use of an eye tracking system in order to be adjusted with respect to the POR. To our knowledge, no previously published research has tested a gaze-contingent variable ICD along with dynamic DOF blur in a Cave Automatic Virtual Environment. Our results show that the anatomical and variable ICD performed similarly regarding each criterion of the experiment, both being more efficient than the fixed ICD. Besides, as with earlier similar attempts, the configurations with DOF blur obtained lower subjective evaluations. Although mainly not significant, the results obtained by the variable ICD and DOF blur are likely due to a noticeable delay in the parameters update. We also designed a new methodology to objectively compare the geometry and depth rendering, based on the reproduction of the same scene in the real and virtual setups, and then on the study of resulting ocular convergence and angular deviation from a target. This leads to a new comparative criterion for the perceptual realism of immersive virtual environments based on the visual behavior similarity between real and virtual setups.

Keywords: Virtual reality, Stereoscopic parameters, Inter-camera distance, Depth of field blur, Point of regard tracking

1. Introduction

Because of its immersive nature [1], virtual reality (VR) is a powerful tool with many applications, ranging from entertainment to more practical uses, such as training through simulations or in medical practice. These practical applications do however call for a strong sense of immersion and perceptual realism[1], in order to bring the behaviors expressed during an immersive experience as close as possible to those seen in a real setup. For example, clinical research in psychiatry used ocular behavior in VR in order to assess phobic avoidance or sexual preferences [2][1][3]. Expressing the same behaviors in immersion as in real life could thus enhance the validity of such an assessment procedure, and the use of virtual environments and avatars instead of real scenes and people. VR is often used in combination with stereoscopic displays, which can intensify the feeling of immersion [4].

Immersion in a stereoscopic environment can be visually constraining, particularly due to a conflict between accommodation (adjustment of the eye lenses to focus at the observed depth) and eye convergence. While these two phenomena are cross-linked in normal viewing conditions, in a Cave Automatic Virtual Environment (CAVE) the viewer always focuses at the screen level regardless of the vergence angle, which leads to a conflict [5][6][7][8] as can be seen on Fig. 1. Lambooij et al. [7] cite as other common causes of visual fatigue: rapid vergence movements that stress this conflict, excessive binocular parallax (leading to diplopia), and unnatural blur intensities (causing ambiguous or erroneous depth perception).

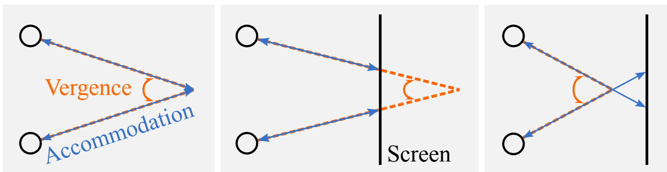

Figure 1: (Left) Cross-link between accommodation and vergence angle in normal viewing conditions. (Middle and right) In virtual environments, a conflict arises between the amount of accommodation and the vergence angle

Stereoscopic parameters play an important role in visual comfort. First, linked parameters (cameras' orientation relative to each other and distance between them, see Fig. 2 (top left)) will impact the disparity between left and right images, and thus influence the amount of perceived depth [9], as well as visual comfort. Nevertheless, there is no single optimal inter-camera distance (ICD) setting that minimizes the fusion effort and leads to optimized depth perception for varying viewing circumstances and depth ranges [10]. Indeed, the incidence of

[1] In the scope of this article, we do not designate by perceptual realism the graphical realism, mainly provided by the shapes and textures of the elements of the virtual environment, but the closeness between the viewer's perception of the virtual environment and an identical real one.

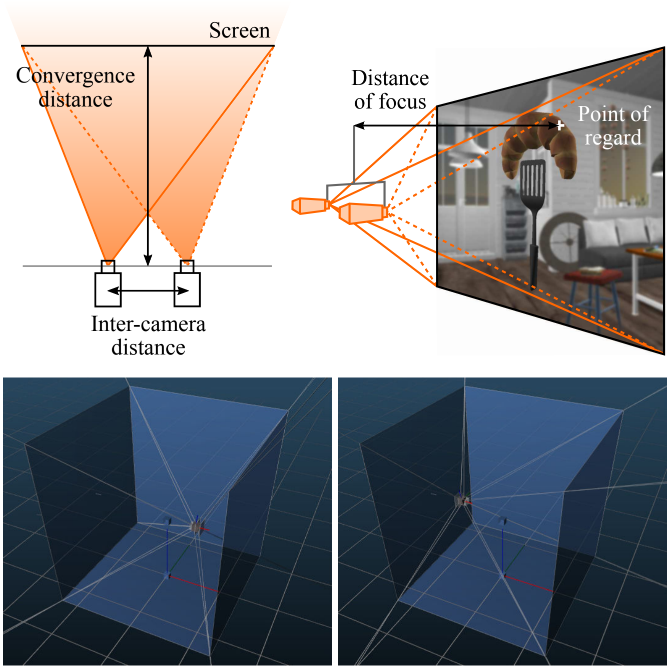

Figure 2: (Top left) In a typical CAVE environment, cameras are parallel with asymmetric view frustums. The distance of convergence used is the one between the viewer and the screen. (Top right) The distance of focus is used to adjust the DOF blur. (Bottom) Dynamic frustums adjustment to the position of the viewer's head

ICD is intrinsically linked to the interpupillary distance (IPD) of each individual [7]. The known methods to choose the best possible ICD value will be developed in Sec. 2.1. We decided to investigate an approach still unexplored which brings together several solutions, namely, the variable ICD. It adapts the ICD in real time based on the current sampled depth, given by the 3D point of regard (POR), and the viewer's preferences, as ICD will linearly vary between three values chosen by the viewer for three predetermined POR depths.

We also used the POR in order to adapt a second parameter: a depth of field blur (DOF blur) that simulates the one that occurs in natural vision due to the accommodation phenomenon. In a CAVE, this physiological blur is not correlated to the distance of the observed object as the eyes always focus at the screen level. Moreover, adding a synthetic DOF blur requires to monitor the POR in order to adapt the blur location accordingly (see Fig. 2 (top right)). When blur is not added to an immersive system, the resulting images are unnatural with only sharp objects. While it is true that DOF blur has already been studied in other papers, Hillaire et al. [11] emphasize the fact that the processing capacity was only recently acquired to compute this interactive blur in real time, and that we still need to evaluate its impact on viewers' performance and subjective preferences. We hypothesized that by leading the virtual scene perception closer to reality, blur could increase the feeling of immersion.

Finally, we were interested in quantifying the perceptual realism of the simulation, not only based on subjective evaluations, but also on objective criteria. The binocular eye tracking system we use provides sight directions and vergence information, which allow us to link objective measurements to the realism of the geometry rendering of the scene (depths, positions, relative distances and sizes). To the best of our knowledge, previous experiments dealing with ocular movements induced by VR were only based on qualitative considerations [12][13][14]. We thus designed an innovative test that compares eye movements while viewing a real scene and its virtual copy. The vergence similarity, which can be quantified, provides indications as to whether the depth is perceived similarly in both virtual and real cases, and whether the ocular behavior of the viewer remains the same. Our main contributions can be summarized as follows: (1) we investigate a new value for ICD which varies in real time with respect to the POR and the user's choices; (2) we evaluate six stereoscopic configurations in an experiment involving 18 participants, and (3) we design a new methodology to quantify and compare simulation realism.

2. Related Work

Although many papers have covered stereoscopic parameters involved in VR quality, we did not find a study that combines adjusted ICD according to the POR depth with an interactive DOF blur. This paper builds upon the following areas of previous research: (1) stereoscopic parameters adjustment, (2) methods for estimating the 3D POR, and (3) evaluation criteria to compare stereoscopic configurations.

Regarding the configuration of the cameras system, we used the typical setup of a CAVE, i.e. parallel cameras with asymmetric frustum as presented in Fig. 2. The image-shifted and converged system configurations are out of scope of the parameters we investigated in our approach, we refer the reader to [15][16][17] for analysis using these configurations.

2.1. Inter-camera Distance

When dealing with immersion quality, the ICD is the first factor that comes out as it strongly impacts both visual comfort and depth perception. In the experiment of Best [18], ICD values of 5.0 cm and 7.4 cm significantly increased visual fatigue compared to 6.3 cm, which is the average adult IPD. Indeed, an ICD value too high results in an excessive binocular parallax that notably leads to eye-straining visual artifacts like diplopia [7]. Moreover, inadequate ICD values lead to an under- or over-estimation of depth judgment [10].

Using an ICD different from the IPD seems to be contradictory with the search of realism in the depth and distance rendering, as it implies a space deformation. However, "disparity depth cue is a highly flexible depth enhancement, rather than the primary determinant of 3D space perception" [19], as other depth cues intervene in the brain's processing of the stereoscopic images. Ware [19] noted that his participants could comfortably work with ICD larger than their IPD, and suggests that it might not be mandatory to make both values match.

2.1.1. Fixed ICD

If the ICD value is fixed, it is often determined after several attempts by the scene creator and used as is for every viewer, which only ensures that stereoscopic images suit the creator's binocular vision [9]. The viewer may also be allowed to choose, thus taking into account his subjective preferences and indirectly considering his anatomical IPD. Nevertheless, Wann et al. [14] noticed that their participants adjusted the ICD so that images merge where their eyes did converge, "regardless of the apparent location within the virtual world". Furthermore, the ICD chosen was systematically smaller than the anatomical IPD. Finally, the ICD can be fixed according to the depth interval the viewer would need to sample, as recommended by Wann et al. [14]. According to them, at a close working distance, an ICD which reduces fusion effort is significantly smaller than the IPD.

However, using a fixed ICD is tantamount to assuming that the viewer's vergence remains constant, that is, the depth sampled in the scene is fixed [14]. Since that is generally not the case, a dynamic adjustment is therefore needed.

2.1.2. Dynamic Adjustment

The ICD value can be modified in real time depending on the viewer's position and movements, the task he has to perform, where he is looking at, or just so that the disparities values stay more comfortable. The ICD adjustment can be combined with a modification of the scene such as scaling. Jones et al. [9] described a method for the real-time calculation of the positions of the virtual cameras that minimize distortions as the viewer's head moves. They diverged from the current notion that cameras must follow eye positions, and instead used a proportional relationship when head movements are parallel to the screen. They also used a method of Depth Range Control that maintains a fixed perceived depth distribution when the viewer moves towards or away from the screen.

It is worth noting that ICD dynamic adjustment is not necessarily aimed at searching for perceptual realism, it may also be used to only enhance depth perception [10] or increase visual comfort. For instance, Ware [19] artificially scaled the scene in order to optimize disparities for depth discrimination and to reduce visual discomfort due to the accommodation-vergence conflict. These adjustments force the viewer to adapt his depth judgments continuously, in particular when size cues remain unchanged while depth cues vary [10]. However, given that depth perception is dominated by other cues such as motion parallax, the ICD changes can remain unnoticed, especially if they occur gradually over a few seconds [19].

In our research, for the purpose of increasing perceptual realism, we test the adjustment of the ICD as a function of the position of the POR. Wann et al. [14] argue that the ICD should vary at the same time as the vergence angle, as the latter introduces small changes in the IPD due to the eye rotation centers located 5-6 mm behind the nodal point of the optical system [20]. For example, the IPD decreases by 1 mm between focal distances at infinity and at 30 cm, leading the viewer to perceive depth differently. Several related works that adapted the ICD to the viewer's POR improved the performance of the stereoscopic configuration. Celikcan et al. [21] obtained better ratings in terms of perceived depth, visual comfort and image quality with their technique of Dynamic Attention-Aware Disparity Control. Yet, they did not use eye tracking, but rather estimated the POR based on a scene content analysis. Kulshreshth and LaViola Jr [22] also proposed a POR-contingent algorithm that dynamically adjusts ICD and convergence distance to projection surface. Their results on depth judgment tasks indicate that the adjusted parameters provide enhanced depth perception compared to fixed parameters, even in scenes with an object close to the camera. Finally, Bernhard et al. [23] found that gaze-controlled adjustment can "lower fusion time for large disparities", the fusion time being used as a discomfort measurement. In addition, they recommend to apply this adjustment in a personalized manner. This is what we aim to do in our approach, by letting the participant choose three ICD values while observing objects at three predefined depths. The ICD will then be computed in real time based on the POR, by means of a linear interpolation between these three values, which, to our knowledge, is a solution that has never been tested in a CAVE.

2.2. Depth of Field Blur Influence

The second parameter we investigate in order to improve the configuration quality is the simulation of a DOF blur. We assume that the DOF blur can increase perceptual realism and depth perception as accommodation blur is used by the brain in normal vision as a depth clue [11].

In a non-stereoscopic experiment, Hillaire et al. [24] compared three conditions of navigation: (1) without blur, (2) with blur computed considering that the POR stays in a centered rectangular area, and (3) with blur computed using eye tracking. The first two conditions obtained similar results, but the last one was significantly better in terms of fun and immersion. Other researches suggest that blur can significantly reduce visual fatigue [6][25] or discomfort [26]. Indeed, even if it does not stimulate accommodation, it might alleviate the accommodation-vergence conflict by producing a "natural relationship between retinal image blur and binocular disparity", as well as enhance depth perception [5]. Also, Nagata [27] observed that a blurred background increased the limits of fusion, i.e. the depth range for which stereoscopic images can be fused without experiencing diplopia, although their experiment only involved three participants. This is likely due to the fact that "the limits of fusion increase as a result of the decreased spatial frequency" in natural vision [7].

However, other related works show more mixed results. Another non-stereoscopic study by Hillaire et al. [11] revealed that blur had a significant negative effect on the performance, and no effect on the subjective evaluation of realism, fun, perception of depths and distances, and feeling of immersion. However, they did not use an eye tracking system, and the blur appeared disturbing when participants explored the image outside the area where the POR was assumed to stay. For the stereoscopic case, the experiment of Brooker and Sharkey [28] did not reveal performance improvement when the blur is computed using eye tracking. Vinnikov and Allison [5] also reported that adding a dynamic DOF in stereoscopic conditions did not enhance depth

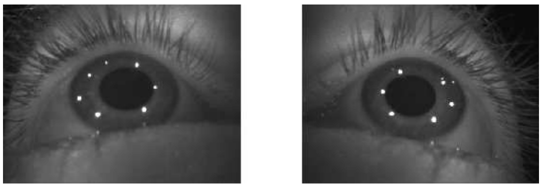

Figure 3: Corneal reflections of infrared emitting diodes

impressions and reduced image quality and viewing comfort. Finally, the participants of Duchowski et al. [26] expressed a non-significant dislike towards DOF blur. These poor results are generally likely attributed to the noticeable delay in DOF update, Duchowski et al. [26] adding the spatial innaccuracy of the eye tracker. Brooker and Sharkey [28] suggested to further evaluate the real-time DOF blur effect in a virtual environment, which is one of the objectives of our study.

2.3. Point of Regard 3D Position

In typical eye tracking systems, sight directions are determined using the geometric relationship between the center of the pupil and corneal reflections [29] produced by infrared emitting diodes (see Fig. 3). However, in order to get the 3D POR, the depth along these directions must be determined. This requires greater technical effort [30], for example, using a binocular system, as in the methods reviewed below.

2.3.1. Inferring from Vergence Movements

This category of techniques takes advantage of the fact that the POR depth variation is accompanied by an ocular convergence movement. The vergence angle varies from about 14° when the POR moves from infinity to a distance of about 25 cm, which is "the nearest distance for comfortable convergence", and from about 36° when it moves to the closest convergence point [13]. Nearly 70% of the vergence angular variation occurs within a range of 1 m. Moreover, it has been shown that stereoscopic stimuli induce adequate convergence movements despite accommodation-vergence conflict [13][12][14], lending legitimacy to this method even in virtual environments. However, due to the noise level in raw eye tracking data and the nonlinear relationship between the vergence angle and the POR depth, its accuracy remains poor. For example, Daugherty et al. [13] measured the vergence angle while displaying a target on a plane at three increasing depths. They obtained a higher angle for the back plane than for the middle one (respectively around 0.82 for the front, 0.26 for the middle, and 0.30 for the back after normalization). Duchowski et al. [12], who used the vergence angle to compute the POR depth, had to apply a filter and a least squares fit in order to counter the significant noise level, as well as wrong depth judgments by the participants. They even noticed significant errors in the monoscopic case, with their estimated average depth not being located at the screen level, but between 10 cm and 20 cm in front of it.

A similar idea consists in estimating the POR depth by evaluating the distance between pupil centers. Kwon et al. [29] tested this approach for a gaze-dependent application, i.e. an application in which interactions with virtual objects are made through glances. They obtained a good accuracy, with 95.7% of successful object selection over 30 trials, but their targets were placed in a discrete partition of the virtual space. Indeed, due to the noise level, the techniques based on vergence movements are more adapted to scenes composed of discrete depth levels than to rich environments. They also noted experimentally that the function of theoretical variations $f(IPD) = depth$ is not linear, making it necessary to rely on a calibration phase to define the IPD range for each plane, for each participant.

2.3.2. Intersection of Lines of Sight

A geometric approach consists in considering that the POR is located at the intersection of the two lines of sight, or at the closest point to both of them. Again, errors and noise in the estimation of sight directions strongly impact the accuracy. To improve this estimation, Essig et al. [31] developed a calibration method based on a Parameterized Self-Organizing Map (PSOM), which is a type of artificial neural network. The PSOM is trained in the calibration phase during which the participants look at 27 markers. Then, during the test phase, the PSOM corrects the measured sight vectors while the participants gaze at 16 markers. The distance between the estimated POR and the actual marker position gives the accuracy measurement. They obtained a global average error of 2.78 cm, noting that most of the error is observed for the z-coordinate (respectively 0.52 cm, 0.82 cm, and 2.53 cm, for x, y and z), which highlights the fact that the depth information is the most difficult to estimate.

2.3.3. Intersection with Scene Geometry

Finally, the 3D POR can be determined by intersecting one or both lines of sight with the virtual scene, the assumption being that the first object intersected is the object of attention [30]. This requires knowledge of the scene geometry (as is the case in virtual environments) and that sight directions and eye positions be expressed in virtual world coordinates, by combining eye tracking with head tracking and a calibration phase. Pfeiffer [30] listed some limitations of this method, such as when the target object for one eye is hidden by an element in the view of the other eye. On the other hand, he underlines its advantages, with the most important being that the POR estimation obtained is fairly accurate.

Thus, while a significant variety of techniques were investigated to obtain the 3D POR, many of them were unfortunately proven to be inaccurate in the presence of the current level of noise in eye tracking data. After an unsatisfactory attempt at intersecting the lines of sight, we therefore decided to base our approach on intersecting the scene geometry.

2.4. Comparative Criteria

Finally, we review the most common criteria used to compare stereoscopic configurations.

2.4.1. *Subjective Evaluations*

In several research works [22][21][32][11], the authors asked their participants to grade each experimental configuration on various criteria using a 5-point or a 7-point Likert scale, from "bad" to "excellent". Blum et al. [32] however allowed the participants to place their mark at any location, then measured and converted the results to a scale between 0 and 100. In a similar way, we used a 7-point Likert scale ranged from "very negative" to "very positive", then converted the marks to a scale between 0 and 1. Regarding the criteria, apart from the "visual comfort" which has been discussed in Sec. 2.1 and Sec. 2.2, we used the categories of "realism", "fun", "perception of depths and distances", and "feeling of immersion" used by Hillaire et al. [11].

2.4.2. *Limits of Fusion*

We designate as limits of fusion the distance between the viewer and the POR below which stereoscopic images cannot be merged. This happens when the parallax becomes excessive probably because of the too great dissociation between accommodation and convergence [8][17]. This distance shall not be confused with the limits of Panum's fusional area [6][7], which also contribute to fusion range but represent the distance between the horopter and the boundaries of Panum's area. We used the limits of fusion as another criterion linked to visual comfort. Indeed, the larger the depth range of easy fusion, the larger the variety of scenes that can be displayed while avoiding discomfort due to diplopia.

Jones et al. [9] reported an experiment in which the limits of fusion for a simple scene were generally located between 30 cm and 50 cm in front of the screen, and between 2 m and 20 m beyond it. Woods et al. [17] used the following protocol to determine the limits of fusion. The participant had to look at a 4 cm diameter donut. They increased the horizontal parallax between stereoscopic images "in the crossed (out of the screen) or uncrossed (into the screen) directions" until the viewer lost stereoscopic fusion, or they decreased it until he could fuse images. Each measurement was realized at least three times. The results revealed a great inter-participant variability, and suggest that the depth range increases with an extended exposure to stereoscopic systems.

As seen in Sec. 2.2, other research suggest that a DOF blur can increase the fusion range, most likely because it removes high spatial frequencies and might alleviate the accommodation-vergence conflict, which both have a negative effect on the limits of fusion.

2.4.3. *Vergence Angle*

Duchowski et al. [12] stated that stereoscopic stimuli induce ocular responses similar to those caused by a real scene. However, they did not quantify the similarity, and instead shifted the experimental data to match the expected values. We assume that the correlation rate between ocular movements caused by real and virtual stimuli, especially vergence angles, can be used to objectively measure the performance of the configurations. We thus designed a test to monitor eye movements while the participant looks at a real scene or at its reproduction in VR. This allows us to measure the similarity between virtual and real stimuli, as well as to compare the realism of virtual configurations in an innovative way.

In conclusion, many research works focused on the quality of the immersion provided by immersive VR environments. They covered the impact of ICD and DOF blur. In this paper, we introduce an approach that combines POR-contingent ICD and DOF blur. Based on a comparison among various methods to determine the POR, we come up with a solution involving a mixture of calibration and lines of sight intersection. We then evaluated our approach by conducting an experiment based on several comparative criteria, including an innovative one that confronts VR and reality.

3. Methodology

3.1. *Apparatus*

We conducted our experiment in an iCUBE-4 of 3.1 m × 2.4 m × 2.4 m with projection on the three walls and the ground. The eight projectors, two per screen, display at a resolution of 1280 px × 720 px and a frame rate of 120 Hz.

The optical head tracking system (OptiTrack) relies on eight infrared cameras that capture the reflection of infrared lights over markers placed on any monitored object. This system's specifications indicate a submillimeter precision and a latency of 8.3 ms. Positions and orientations are transmitted by a VRPN server to the software that controls the projections (MiddleVR) and to our own software, which uses the collected tracking data for the real-time update of the virtual cameras' parameters. The binocular eye tracking is performed by the Eye Tracking Glasses (ETG) 2.0 designed by SMI. This device includes two cameras operating at 60 Hz, located in the spectacle frame and directed towards the eyes. The ETG are USB connected and provided along with iViewETG software responsible for image processing. It operates using a sophisticated model of ocular behavior based on the method of center-corneal reflections (PCCR) cited in Sec. 2.3, and behaves as a server transmitting the ocular tracking data. The ETG provides four 3D vectors in local coordinates: two sight base points, which correspond to the center of the eyeballs, and two sight directions which connect base points to the centers of the pupils. The specifications indicate an accuracy of 0.5° within depth boundaries (40 cm - infinity) and a range of 80° horizontally and 60° vertically. The ETG are combined with Volfoni active stereoscopic glasses (see Fig. 4). Remarks about the number of frames per second (FPS), and thus the latency of the whole system, are given in Sec. 4.4.

3.1.1. *Calibration*

The server software supplied with the ETG, iViewETG, allowed us to perform either a 1-point or a 3-point calibration. They consisted in asking the participant to gaze at either one or three points in their visual field, while the experimenter clicks in iViewETG on these same points, on the image filmed by the scene camera located in front of the ETG at the nose bridge

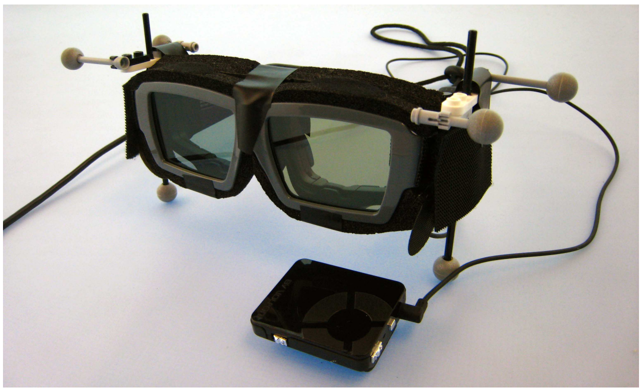

Figure 4: Eye Tracking Glasses combined with Volfoni active stereoscopic glasses. The markers are placed on the ETG so that head's and ETG's reference frames match, allowing us to collect sight directions in local (ETG) and global (CAVE room) coordinate systems

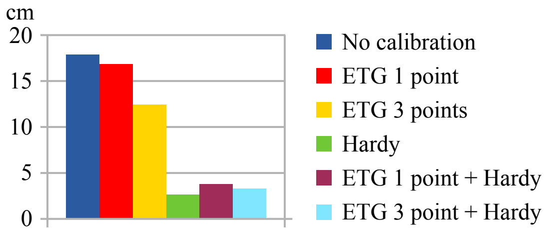

Figure 5: Average distance error obtained with different types of calibration during preliminary tests performed on one participant, for targets at 2 m

level. We chose not to use these options, both because the imprecision results are slightly higher with them (see Fig. 5), and because using it requires a manipulation of the active stereoscopic glasses part in our glasses assembly. During this operation, if the glasses move with respect to the viewer's head, or if the tracking markers are shifted, the imprecision increases.

We thus carried out another type of calibration during the experimentation, based on Hardy's Multiquadric [33], with a rectilinear grid of nine markers. After we set up the ETG on the head, the calibration was achieved by asking the participant to stare successively at nine markers placed on a virtual rectilinear grid and displayed in a random order. The imprecision measurements were carried out on a 25-marker rectilinear grid, and corresponded to the average distance between each computed POR and the current marker during a 2 s recording.

Results revealed an error of 6.47 cm on average for all the participants for targets at 2 m, corresponding to a 1.85° visual angle. Although the specifications of the ETG indicate that it works identically with contact lenses, a one-way ANOVA test showed that wearing lenses leads to a significantly higher POR imprecision ($F=13.43$, $p=0.002$).

3.1.2. 3D POR Determination

Preliminary tests performed on one participant included trials to determine the POR by intersecting the lines of sight, as described in Sec. 2.3.2. Yet, because the y-coordinate often diverges slightly in the sight direction vectors given by the ETG, this intersection usually appeared to be at the center of the eyes. Moreover, other errors and noise present in the measurements led to poor accuracy. Thus, we decided to intersect the lines of sight with the scene and to consider that the POR is located at the intersection, as described in Sec. 2.3.3. Given that the ETG provides binocular data, we tested four ways to determine the POR:

(1) using the left sight direction only (starting at the left base point), and correcting it with calibration;

(2) using the right sight direction only (starting at the right base point), and correcting it with calibration;

(3) using the average of the left and right directions before calibration (starting at the center of the base points), and correcting it with calibration;

(4) using the average of the corrected left and right directions (starting at the center of the base points).

The results of imprecision tests showed that better results were achieved using method (3), with respective average distance errors of 5.0 cm, 2.7 cm, 2.6 cm, and 3.1 cm. Furthermore, the average distance between each pair of the four PORs, obtained independently from the two sight directions, can be used as a measure of the estimation reliability: the smaller this distance, the smaller the uncertainty of our POR estimation, as it implies that binocular measurements agree with one another.

Due to the noise and jittering observed in the data collected, in particular the fact that some vectors were occasionally directed backwards, we tested two techniques of smoothing, one by averaging through several images (five and ten) the successive positions of the POR, the other using the Unity SmoothDamp method. Both of these techniques resulted in perceived latency for the authors in the monitoring of the POR, which was inconvenient for the real-time adjustments performed, that must fit the quick phenomenons of saccades and eyes accommodation. Indeed, saccades are performed at a rate of three to five times per second [34], the travel time of the eye being about 10 ms to 100 ms [35]. We thus decided, contrary to some related works [19][26], to keep the data noisy but reactive, and to adapt as quickly as permitted by the tracking systems the variable ICD and DOF blur.

3.2. Participants

Eighteen individuals (15 male and 3 female), between the ages of 22 and 49 (average 28), volunteered to participate in our experiment. Five of them wore contact lenses. They all had a significant experience with 3D or virtual reality, having seen at least four 3D films (maximum of 30, average of 11) and 14 of them having played video games (8.6 h/week on average). They all indicated that they usually had no difficulties perceiving depth at movie theatres. Each participant filled out a short pretest questionnaire, providing among other things the aforementioned demographic information and immersion tendency.

ICD Blur	Anatomical	Fixed	Variable
Without	B-AICD	B-FICD	B-VICD
With	B+AICD	B+FICD	B+VICD

Table 1: The six configurations tested during the experiment

Figure 6: Absence (left) and presence (right) of the DOF blur used for the experiment

Figure 7: Scene used during the experiment

3.3. Procedure

As seen in the litterature review, the ICD parameter strongly impacts the visual comfort and depth perception, thus the whole performance of a configuration. We decided to test three different approaches for this parameter: the anatomical IPD of each participant (AICD) measured by the ETG, a fixed single ICD value chosen by the participant (FICD), or a variable ICD that is dynamically linearly interpolated between three values (VICD). These three values correspond to ICDs chosen by the participant during the start-up phase for three predefined POR depths (0.4 m for VICD-Near, 1 m for VICD-Middle, and 2.5 m for VICD-Far). In order to avoid reaching aberrant values, we keep the ICD constant when the depth of the POR is inferior to 0.4 m or superior to 2.5 m. Regarding the DOF blur, we used the Depth of Field Scatter of Unity [36] shown in Fig. 6. We used the parameter "Focus on Transform" that automatically determines the focal distance using a target object in the scene. We defined as target a transparent sphere that follows the POR. We selected the blur intensity value based on our own feeling of a realistic DOF blur and those of one participant during a preliminary test. Only one blur intensity was tested to keep a reasonable number of configurations, thus limiting the duration of the experiment. The six tested configurations varied in ICD value and DOF blur presence. Their names are summarized in Table 1 for the rest of the paper.

3.3.1. Start-up

The experiment comported a first phase to perform the calibration based on a virtual grid of nine markers, as described in Sec. 3.1.1. This calibration was carried out for each participant in order to adapt the correction of the vectors to his anatomical characteristics. Then, the participant had to manually tune the ICDs that would be used later for FICD and VICD, with objects displayed at three predefined distances (0.4 m for VICD-Near, 1 m for FICD and VICD-Middle, and 2.5 m for VICD-Far). Using a mouse wheel, he could move closer (until reaching the monoscopic configuration) or move apart (without limits) the two virtual cameras, until reaching a satisfactory value considering his visual comfort and depth perception. For this last criterion, we indicated the objects' distances using a tape measure, which allowed the participant to ascertain that his perception matched the actual size and distance. He could also use the general depth cues given by the other objects of the scene (see Fig. 7) in order to confirm his choice. Every choice was performed three times for each predefined distance, and averaged to set the final ICD value.

Each phase of the experiment described below is linked to a comparative criterion, and required participants to perform specific tasks. The start-up lasted around 18 min on average, followed by 15 min of navigation (including the time to rate), 14 min to determine the limits of fusion, and 11 min for the vergence comparison phase, for a total duration of 57 min on average. The order of presentation of the configurations followed a Latin square rotation, in order to avoid habituation or visual fatigue bias.

3.3.2. Navigation

This phase aimed at evaluating the configurations based on subjective criteria. It consisted in navigating in a virtual scene along a predefined path, then rating the configurations according to one's impressions. We defined a fixed 90 s navigation path in the scene with a dual objective: (1) to show objects of interest with a wide depth range, and (2) to reduce the variability inter- and intra-participant, with the only remaining differences between the six navigations being the current configuration and the gaze path. After each navigation, the participant was invited to give grades in five criteria: (1) visual comfort, (2) rendering realism, (3) fun, (4) depth and distance perception, and (5) sense of immersion. The continuous rating scale ranged from "very negative" to "very positive", as presented in Sec. 2.4.1.

3.3.3. Limits of Fusion

This phase of the experiment was designed to determine the viewer's limits of fusion with each configuration, thereby obtaining an objective indication of their effectiveness in terms

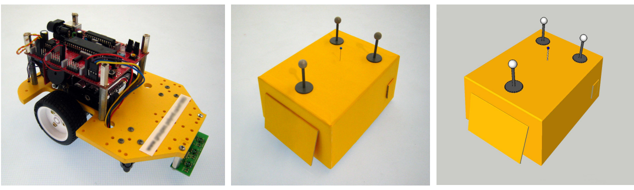

Figure 8: Real and virtual robots

of merging near virtual objects. Limits of stereoscopic fusion were determined as in [17] with two scenarios: increasing and decreasing disparity. For the first one, objects were placed at 40 cm, a distance that allowed all the participants to fuse the images. Using a wireless mouse, the viewer had to bring the objects closer until experiencing diplopia. In the second case, objects were initially placed extremely close to the participant, at 7 cm, and he had to move them away until fusing the images. These two scenarios were repeated three times for each configuration and averaged to obtain the limit. As this phase was prone to induce eye fatigue, the participant was allowed to pause for a few seconds between each measure.

3.3.4. Vergence Comparison

In addition to the participant's subjective assessment of the realism of the distances, his ocular behavior, and in particular the vergence movements, give objective indications on his perception of 3D. These movements naturally accompany the inspection of a real scene when the POR changes from a depth to another. In the virtual case, the computation of the stereoscopic images or their display may imply a space deformation, leading to under- or over-evaluation of the distance that will affect the angle of vergence. On the contrary, if the depth information is well rendered, the binocular vergence movements should reproduce those that whould have been performed with an identical real scene. Thus, the goal of this last phase was to compare the six configurations on the ocular behavior they induce when staring at a 3D moving target, using a real moving target as a reference. To ensure that the POR follows the same path, we guided the participant's eyes using "identical" stimulus. We used a robot capable of following a line on a table. A box overhung the robot, and was topped with a very small target that the participant was required to follow during the entire course (see Fig. 8). In order to reproduce the robot's movements, the position and orientation of the real robot were tracked, saved, and interpolated for moving the virtual robot, requiring the "real configuration" to be performed first. As the robot's path can sligthly change between each tour, because of its groping search in real time of the trace with the aid of detectors, the recording is done for every participant. The detailed setups are shown on Fig. 9.

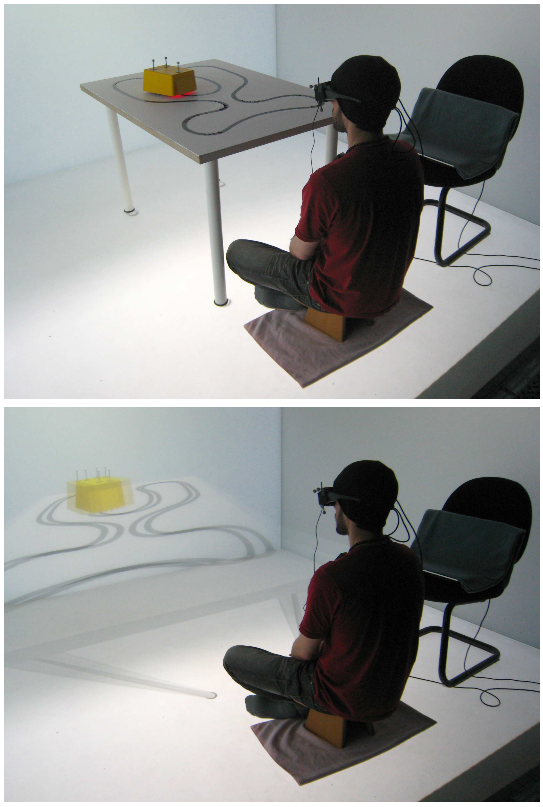

Figure 9: Real and virtual setups for the vergence comparison phase

4. Results and Discussion

4.1. Subjective Ratings

Fig. 10 summarizes the ratings assigned to each configuration during the navigation phase. At first glance, it appears that none of the configurations stood out positively in every criterion: B+AICD presents a better median for comfort and immersion, B-AICD for realism and fun, and B-FICD for depth and distance perception. However, we can note that B+FICD always got the lowest medians. We carried out ANOVA tests (one-way and two-way) in order to outline statistical impacts of the configuration, the ICD value and the presence of DOF blur on the results. They did not reveal significant effects, except for the blur, which worsened the depth and distance perception ($F=4.91$, $p=0.034$). We can however highlight some interesting observations.

Regarding the ICD parameter, we note that the order AICD $\geq$ VICD $\geq$ FICD occurs in 8 cases out of 10, which suggests that AICD was preferred. In addition, we found the AICD/FICD ratio to be significantly related to the results obtained by the B-FICD configuration in the comfort, fun, and depth and distance perception criteria ($p=0.002$, $p=0.012$, $p=0.040$ as respective correlation probabilities with linear regressions), with an FICD close to the anatomical one being preferred.

Concerning the DOF blur, we note that the configuration without blur was preferred over its counterpart in 12 cases out of 15, showing as in several earlier attempts [5][26][11][28] that the participants disliked the addition of a DOF blur. In an attempt to explain this trend, and the significant negative effect on the depth and distance perception criterion, we performed an in-depth study of visual inspection. Extracting saccades and fixations revealed that the fixations were significantly

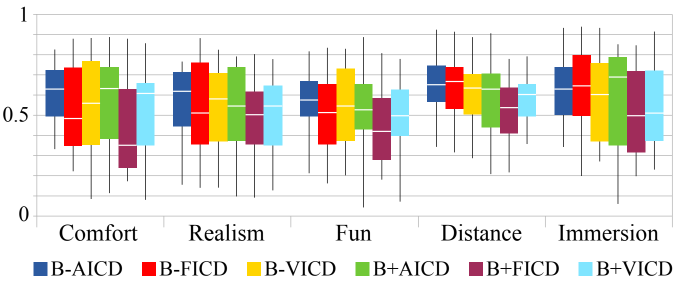

Figure 10: Ratings on the five criteria obtained by each configuration, from "very negative" (0) to "very positive" (1). The rectangle is delimited by the first and third quartiles, and the white line shows the median

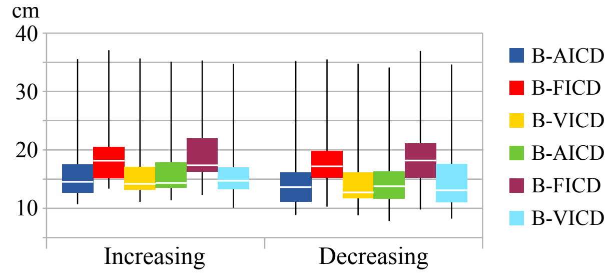

Figure 11: Closest distances of fusion. The average values are 17.11 cm for the increasing scenario, and 16.24 cm for the decreasing one, the limits of fusion per se being located between the results obtained for each scenario

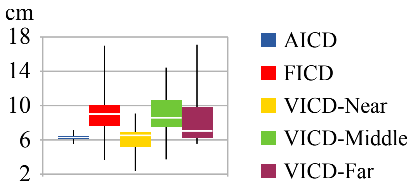

Figure 12: Anatomical ICD, and ICD values chosen during setup. The participants' AICDs were between 5.51 cm and 7.17 cm (6.28 cm on average), whereas their choices for ICD were between 2.38 cm and 17.11 cm. We note that the chosen values for VICD-Near are close to the AICD's

fewer (F=25.5, p<0.01) and longer (F=16.9, p<0.01) with blurred configurations. The saccades also lasted longer (F=236, p<0.01), but their average angular distances were not significantly different (F=1.02, p=0.316), which supports the idea of a slowdown in the visual inspection of the scene. The average number of inspected objects also significantly decreased (F=12.2, p<0.01). It appeared that during our experiment, the average number of FPS was different between the blurred (9.4 FPS) and non-blurred (13.0 FPS) configurations. We assume that the participants noticed the lag and adapted, consciously or unconsciously, their visual inspection speed to alleviate the delay in parameters update.

In summary, B-AICD obtained higher ratings in general and a higher average rating, although not statistically significant. It suggests that on a subjective basis, the anatomical ICD is preferred and the presence of a DOF blur is detrimental, likely due to the insufficient update speed of the POR-contingent parameters.

Besides, in order to evaluate the accommodation-vergence conflict, we computed the absolute difference between the distances of accommodation and of convergence (see Fig. 1) during the navigation phase. We obtained average distances of 1.30 m for the accommodation, 3.55 m for the convergence, and 2.36 m of absolute difference between them. By comparing the ratings with the absolute difference for each participant, for each configuration, it seems that the accommodation-vergence conflict did not influence the subjective evaluations. We took as statistical measure the distance correlation, which was between 0.20 and 0.27 for the five criteria.

4.2. Limits of Fusion

This phase aimed to compare the ease of stereoscopic fusion allowed by each configuration, by determining the closest fusion distance for every participant. In the results summarized in Fig. 11, we notice that with the configurations using FICD, participants had more difficulties fusing close objects, with the limits being greater. A two-way ANOVA revealed that the ICD parameter indeed had a significant effect (for increasing disparity scenario F=3.93, p=0.023, for decreasing disparity scenario F=4.74, p=0.011). By comparing B-FICD and B+FICD to all other configurations, we determined that these two were significantly worse in terms of limit of fusion than the others (a paired t-test gave a p-value<0.01 for increasing and decreasing scenarios). This can be explained by the fact that FICD is higher than the other ICD values (see Fig. 12), as the configurations using VICD actually use VICD-Near when the POR is located at a depth of under 0.4 m. With a higher ICD, the horizontal disparity between left and right images is more significant, leading to a greater degree of difficulty fusing them. We hence recommend, when scenes with close objects are to be displayed with a fixed ICD, to limit the maximum value allowed for the viewer's choice.

The DOF blur did not seem to have any influence, contrary to what suggested the litterature review [27]. We assume that this is due to the fact that for the limits of fusion reached by most participants, the objects took an important part of the screen, the blurred background thus being a minor portion of the image.

4.3. Visual Behavior Comparison Using a Real Scene

Unlike the two previous criteria, this one relies on a real setup used as a reference to compare the virtual configurations, in order to provide insights on their effective perceptual realism. We took two measurements related to the visual path: (1) the vergence angle and (2) the angular deviation from a target.

The vergence angle is a good indicator of the efficiency of the ICD. Indeed, an inappropriate ICD leads to unrealistic disparities between images, and thus to an under- or over-estimation of the target's depth by the viewer, which in turn affects his eyes convergence. We considered the global vergence angle, computed directly from the two corrected sight directions, as well as the eyes rotation toward each other by using the angle between the projection of the sight directions on the horizontal plane.

The angular deviation from a target refers to the angle between the vector from the center of the eyes (center of the sight base points) to the POR and the vector from the center of the eyes to the target position. Therefore, it reflects the angular distance between the location where the participant is looking according to our calculations (the POR), and the location where he is actually looking (the target). This angular deviation is thus an indicator of the accuracies of: the tracking systems, the calculations used to correct the sight direction vectors and to compute the POR, and finally the geometry rendering with the virtual configuration. Similarly to the vergence, we considered the angular deviation as well as its horizontal projection.

The more these two measurements reproduce the ones recorded in the real case, the more the virtual configuration can be considered realistic, as it induced a realistic depth perception whatever the difference between the ICD and the viewer's anatomical IPD. Regarding the deformation inherent to the CAVE, as it was identical for all the configurations, we assume the results differences can only be attributed to the ICD values and the presence of DOF blur.

Before performing the analysis, vergence and angular data were manually processed to remove outliers produced by eye tracking. Indeed, when infrared reflections cannot be reliably detected and tracked for the PCCR, for example due to eyelashes or blinks [37], the ETG deliver invalid values for positions and sight direction vectors. Right before or after these periods, we observe brief peaks with amplitudes too high to correspond to vergence movements, which we attribute to the fact that only some diodes were reflected, the rest being hidden by the eyelid. We manually removed these peaks, thus replacing them by a linear interpolation between the previous and next reasonable values.

4.3.1. Vergence Comparison

Fig. 13 shows, for a representative participant, the curves of his vergence angle as a function of time for the real and virtual configurations. The robot performed two laps, passing closer to the participant around times 15 s and 55 s. The distance between the robot and the participant is also plotted on Fig. 13. We observe that the variations were similar for the real and virtual cases, thus demonstrating, as noticed in the litterature review, that stereoscopic images induce realistic ocular movements. Compared to previous works, gathering data for the real case allows us to make deeper qualitative and quantitative observations. For example, considering the general shape of the curve, we note that on average the vergence angle variations had a smaller amplitude in the virtual case (this is particularly visible during the first 10 s of Fig. 13 (bottom)).

As for the quantitative comparison, for each of the virtual configurations, we computed the average distance between its curve and the one obtained in the real case (see Fig. 14). ANOVA tests (one-way and two-way) revealed no significant effect of the configuration, the ICD value or the presence of DOF blur. However, we note that based on the median, the most realistic configuration is B-AICD, followed by B+AICD when considering the global vergence, and by B+VICD when considering the horizontal vergence.

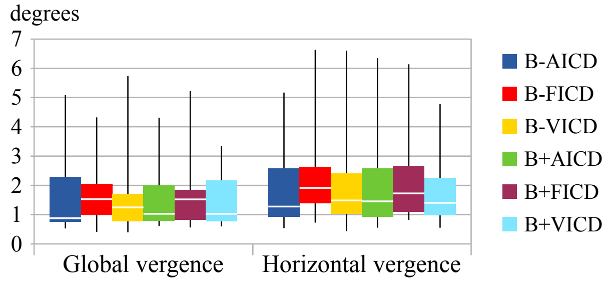

Figure 14: Average distances of global and horizontal vergence angles observed in the virtual cases with respect to the real one

Regarding the ICD parameter, AICD and VICD performed similarly, although AICD obtained vergence angles slightly closer to those measured during the real case. This is likely attributed to the large range of values taken by VICD during this phase. Indeed, as the distance between the robot and the participant took values on average from 0.37 m to 1.24 m, VICD varied mainly between VICD-Near et VICD-Middle, which, as shown on Fig. 12, was close to FICD and even higher than VICD-Far. Besides, B-FICD and B+FICD, which shared the same ICD value, led to greater differences of vergence angle, and a paired t-test revealed that these configurations were significantly further from the real configuration when considering horizontal vergence ($p=0.0257$). As for the limits of fusion criterion, this result can be attributed to the FICD higher values that lead, for close objects, to great disparities which triggered ocular behavior far from those witnessed in the real case.

The presence of DOF blur did not seem to have an influence on the configurations' performance. We observed that adding blur led to worse results when combined with AICD and better ones with VICD, in a non-significant way.

We also computed the relative deviations, which better reflect the significance of the distance between the curves, by dividing the vergence distances with the average vergence angles measured during the real case (see Table 2). These rates confirm the previous results, AICD and VICD obtaining similar lower deviation from the real case than FICD, and the presence or abscence of DOF blur not showing any influence. We note that, although the vergence angle followed the same variations in both virtual and real cases, the differences in amplitude led to a relative deviation of up to 38.5% (and 71.8% horizontally) with respect to the real case. With regard to these high deviations, it has to be taken into account that the average global vergence angle for the real configuration was of 5.1°. Thus, the higher relative deviation corresponds to 2.0°, which is similar to the inaccuracy of 1.85°measured after the calibration phase.

4.3.2. Angular Deviation from a Target

Fig. 15 presents, for a representative participant, the curves of the horizontal angular deviation as a function of time, for the real and averaged virtual configurations. We first observe that the angular deviation was around zero degree for the real configuration, indicating that the POR direction we computed coincided with the vector between the eyes and the position of the target monitored by the tracking system. The angular devi-

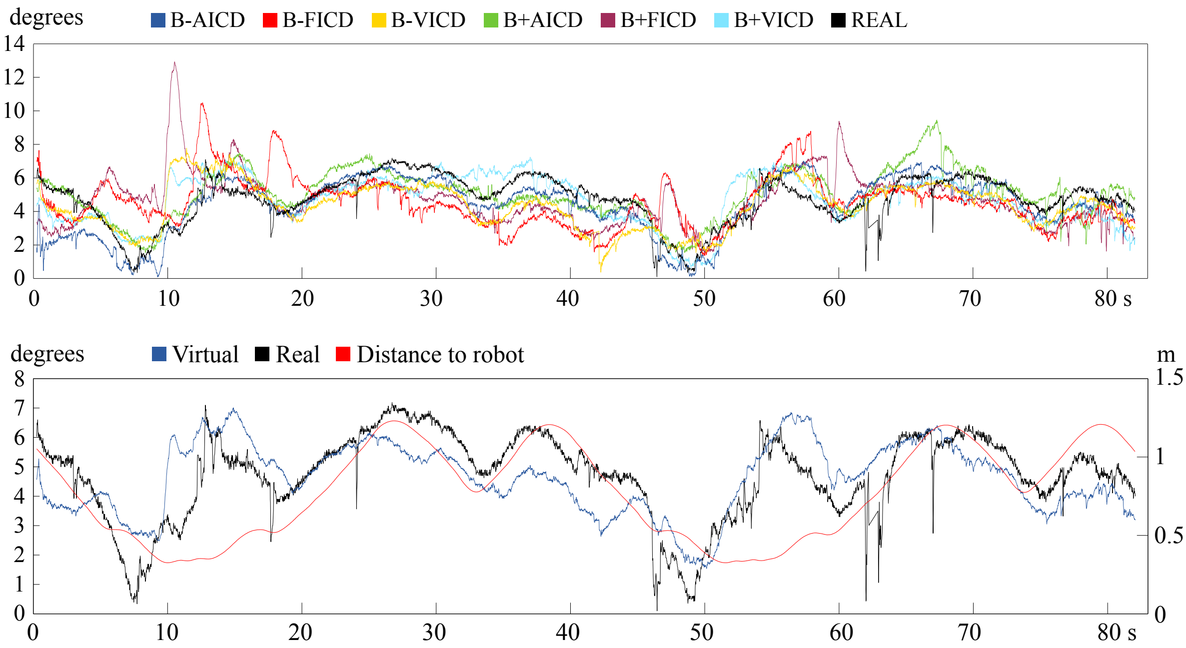

Figure 13: (Top and bottom) Global vergence angle measured for real and virtual configurations for a representative participant. In the bottom, the curves for the six virtual configurations where averaged in order to highlight more clearly the difference of vergence angle between the virtual and real cases, and put in parallel with the distance to the robot

	Global Vergence	Horizontal vergence
B-AICD	31.5%	51.2%
B-FICD	38.5%	71.8%
B-VICD	36.0%	60.6%
B+AICD	32.3%	53.9%
B+FICD	36.5%	67.7%
B+VICD	31.9%	57.6%

Table 2: Relative deviation with the average global and horizontal vergence angles measured during the real case

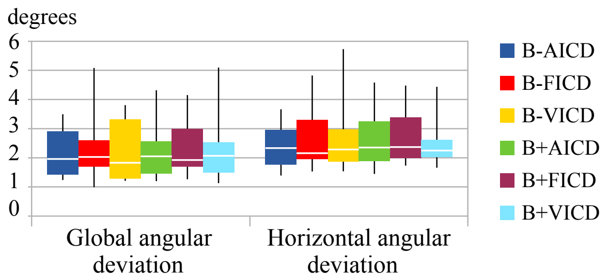

Figure 16: Average distances of global and horizontal angular deviations observed in the virtual cases with respect to the real one. **N.B.** These results do not include one of the participants, for whom measurements showed large errors (an angular deviation greater than 120° over 50% of the time with the real and B-VICD configurations)

ation was greater in the virtual cases, which indicates that the robot's position in the CAVE coordinates (where the viewer saw it) was not the same as its position in the virtual world coordinates. This corresponds to a deformation of space coordinates, probably not only due to the stereoscopic parameters of the configurations but also to the display calibration. While being a limitation in quantifying the absolute realism of a given configuration, the deformation induced by the CAVE setup should not impact the configuration results relative to each other, since it remains the same during the experiment.

Fig. 16 summarizes, for each configuration, the distance between the angular deviations from the target compared to those obtained with the real configuration. ANOVA tests (one-way and two-way) revealed that the configuration, the ICD value, or the presence of DOF blur did not have a statistically significant impact. However, we note that the configurations obtaining the lowest median, i.e. the smallest distance with respect to the real case, were B-VICD, followed by B+FICD. When considering the horizontal angular deviation, B-FICD had the smallest median, followed by B+VICD. It is worth noting that the FICD performed similarly to the others ICD values for this criterion. High stereoscopic disparities therefore seem to only influence the vergence angle and not the gaze direction.

4.4. Remarks and Future Work

As can be observed in the results, none of the configurations stood out significantly from the others as being most effective, implying that the benefits of a variable ICD or the presence of a DOF blur are hard to grasp. As in previous works with DOF blur [28][26] presented in Sec. 2.2, we assume that the delay in the adjustment of these POR-contingent parameters, followed by the images update, must have been perceivable by the participants and detrimental to such adjustments. Indeed, although we did not compute the total latency, which takes into account all the tracking systems' latency, we measured the number of FPS which gives an indication of the maximum system reactivity. As given in Sec. 4.1, the average FPS during the navigation

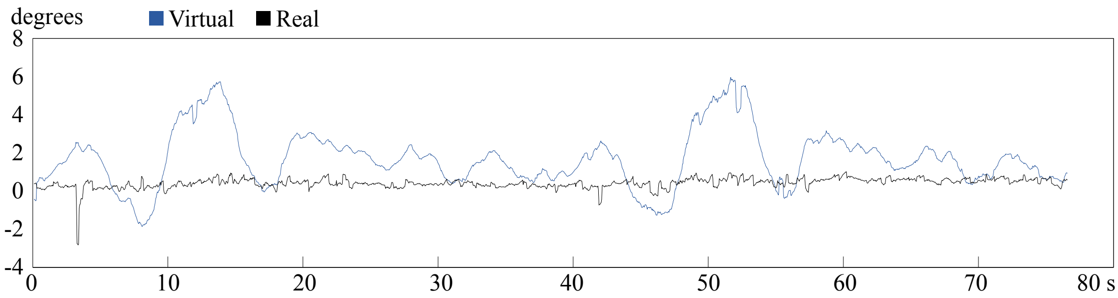

Figure 15: Horizontal angular deviation from a target measured for the real and averaged virtual configurations

phase was 13.0 FPS without DOF blur and 9.4 FPS with it. It is likely that such performance affected the user subjective appreciations, although none of them reported VR sickness. However, while being a limitation to the significance of the results regarding the perceptual realism of the configurations, we assume their relative comparison is still relevant. The number of FPS was almost the same during the limits of fusion determination (12.8 FPS in average), which we attribute to the geometrical complexity of the scene used during these two phases of the experiment. On the other hand, we reached 35.1 FPS in average during the visual behavior comparison thanks to the robot scene, thus probably lowering the impact of the update time.

In order to further investigate the adjustment of stereoscopic parameters, other values for variable ICD and blur intensity could be tested. First, if the setup requires a fixed ICD, the poor results obtained by B-FICD and B+FICD would suggest that the participant not be allowed to select this value, but rather to compare a set of predefined ICDs and select the top performer. If a dynamic adjustment is allowed, a similar test to the one used to compare visual behaviors, i.e. reproducing a real scene in VR, can also be used to determine the optimum ICD values. Indeed, after recording the viewer's vergence angles in the real case, the ICD would be modified in real time while the participant observes the virtual scene until his vergence angles are identical to those recorded. A neural network would be trained with such a procedure in order to determine the best ICD values during the real experiment.

Concerning the DOF blur, only one intensity was tested in order to keep a reasonable number of configurations and thus experiment duration. However, some research suggest that individual tuning is needed to find the appropriate value [5][11], which could also vary according to the accommodation distance. Moreover, several blur algorithms could be evaluated on how closely they replicate the natural viewing blur, using for instance similar methods to the ones proposed for the ICD.

5. Conclusion

This project was originated from a desire to measure and improve the effectiveness of stereoscopic rendering in a VR environment, particularly with respect to the visual comfort during extended exposure, the perceptual realism and the feeling of presence. Our experiment compared six configurations, different in their ICD values (anatomical, fixed or variable) and the presence or absence of a DOF blur, two parameters known for their influence on the criteria cited above. The variable ICD and the DOF blur relied on the viewer's POR in order to be adjusted in real time. For the ICD, the goal was to adapt the amount of disparity, which influence the sense of depth, according to the depth level on which the viewer chooses to focus. The blur aimed to reproduce the DOF blur that occurs in natural vision. Results showed that VICD and AICD were similarly efficient regarding each phase of the experiment. In particular, B-AICD and B+VICD led to more realistic vergence angles, and B-VICD obtained the smallest distance to the real angular deviations from a target (considering the medians). On the other hand, the FICD led to significantly higher fusion difficulties, and to greater differences of horizontal vergence angles as compared to those measured in the real case. The subjective ratings revealed that an FICD closer to the AICD led to higher ratings. The presence of DOF blur did not significantly impact the ability of fusion or the similarity of ocular behavior between real and virtual. On the other hand, the participants expressed a dislike in their subjective ratings towards blurred configurations, which was significant in the depth and distance perception criterion. This experiment highlighted the difficulty of obtaining benefits from varying the ICD or adding a DOF blur, even though the technical limitations encountered and described in Sec. 4.4 may have counteracted the advantages of these real-time adjustments of parameters based on the POR.

As part of this project, we also designed an innovative methodological framework to objectively compare the stereoscopic configurations. The test is based on a recording of the viewer's eye movements while he stares at real and virtual moving targets. The comparison between ocular behaviors, particularly the vergence angle, provides an interesting and quantifiable measure of the perceptual realism and the geometry rendering. It could be used in the future as an evaluation criterion: a greater correlation between the eye responses with the real and virtual configurations indicates an improved perceptual realism of the depth rendering. It can also be used to find the optimum stereoscopic parameters, by modifying them until the vergence corresponds to that measured in the real case.

This comparison method would be a step forward to insure the correspondence between visual behavior in real and virtual setups, and might help increase the feeling of immersion. This is of particular interest to research in psychiatry, when tracking gaze behavior is used as source of diagnosis [2][1][3]. Moreover, in this context, the POR recording provides additional tools for the clinician, such as visual paths, hit maps, measurements of fixations and saccades, which can help in the interpretation of the visual behavior in terms of attention.

[1] Renaud P, Rouleau JL, Granger L, Barsetti I, Bouchard S. Measuring sexual preferences in virtual reality: A pilot study. CyberPsychology & Behavior 2002;5(1):1–9.

[2] Renaud P, Chartier S, Rouleau JL, Proulx J, Trottier D, Bradford JP, et al. Gaze behavior nonlinear dynamics assessed in virtual immersion as a diagnostic index of sexual deviancy: preliminary results. Journal of Virtual Reality and Broadcasting 2009;6(3).

[3] Renaud P, Bouchard S, Proulx R. Behavioral avoidance dynamics in the presence of a virtual spider. Information Technology in Biomedicine, IEEE Transactions on 2002;6(3):235–43.

[4] Hardless G, Meilinger T, Mallot H. Virtual reality and spatial cognition. In: International Encyclopedia of the Social Behavioral Sciences. Elsevier Science; 2015, p. 133–7.

[5] Vinnikov M, Allison RS. Gaze-contingent depth of field in realistic scenes: The user experience. In: Proceedings of the Symposium on Eye Tracking Research and Applications. ACM; 2014, p. 119–26.

[6] Moreau G. Visual immersion issues in virtual reality: a survey. In: Graphics, Patterns and Images Tutorials (SIBGRAPI-T), 2013 26th Conference on. IEEE; 2013, p. 6–14.

[7] Lambooij M, Fortuin M, Heynderickx I, IJsselsteijn W. Visual discomfort and visual fatigue of stereoscopic displays: a review. Journal of Imaging Science and Technology 2009;53(3):30201–1.

[8] Rushton SK, Riddell PM. Developing visual systems and exposure to virtual reality and stereo displays: some concerns and speculations about the demands on accommodation and vergence. Applied Ergonomics 1999;30(1):69–78.

[9] Jones GR, Lee D, Holliman NS, Ezra D. Controlling perceived depth in stereoscopic images. In: Photonics West 2001-Electronic Imaging. International Society for Optics and Photonics; 2001, p. 42–53.

[10] Milgram P, Krüger M. Adaptation effects in stereo due to on-line changes in camera configuration. In: SPIE/IS&T 1992 Symposium on Electronic Imaging: Science and Technology. International Society for Optics and Photonics; 1992, p. 122–34.

[11] Hillaire S, Lécuyer A, Cozot R, Casiez G. Depth-of-field blur effects for first-person navigation in virtual environments. Computer Graphics and Applications, IEEE 2008;28(6):47–55.

[12] Duchowski AT, Pelfrey B, House DH, Wang R. Measuring gaze depth with an eye tracker during stereoscopic display. In: Proceedings of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization. ACM; 2011, p. 15–22.

[13] Daugherty BC, Duchowski AT, House DH, Ramasamy C. Measuring vergence over stereoscopic video with a remote eye tracker. In: Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications. ACM; 2010, p. 97–100.

[14] Wann JP, Rushton S, Mon-Williams M. Natural problems for stereoscopic depth perception in virtual environments. Vision research 1995;35(19):2731–6.

[15] Allison RS. The camera convergence problem revisited. In: Electronic Imaging 2004. International Society for Optics and Photonics; 2004, p. 167–78.

[16] Stelmach LB, Tam WJ, Speranza F, Renaud R, Martin T. Improving the visual comfort of stereoscopic images. In: Electronic Imaging 2003. International Society for Optics and Photonics; 2003, p. 269–82.

[17] Woods AJ, Docherty T, Koch R. Image distortions in stereoscopic video systems. In: IS&T/SPIE's Symposium on Electronic Imaging: Science and Technology. International Society for Optics and Photonics; 1993, p. 36–48.

[18] Best S. Perceptual and oculomotor implications of interpupillary distance settings on a head-mounted virtual display. In: Aerospace and Electronics Conference, 1996. NAECON 1996., Proceedings of the IEEE 1996 National; vol. 1. IEEE; 1996, p. 429–34.

[19] Ware C. Dynamic stereo displays. In: Proceedings of the SIGCHI conference on Human factors in computing systems. ACM Press/Addison-Wesley Publishing Co.; 1995, p. 310–6.

[20] Bennett A, Francis J. The eye as an optical system. The eye 1962;4:101–15.

[21] Celikcan U, Cimen G, Kevinc EB, Capin T. Attention-aware disparity control in interactive environments. The Visual Computer 2013;29(6-8):685–94.

[22] Kulshreshth A, LaViola Jr JJ. Dynamic stereoscopic 3d parameter adjustment for enhanced depth discrimination. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. ACM; 2016, p. 177–87.

[23] Bernhard M, Dell'mour C, Hecher M, Stavrakis E, Wimmer M. The effects of fast disparity adjustment in gaze-controlled stereoscopic applications. In: Proceedings of the Symposium on Eye Tracking Research and Applications. ACM; 2014, p. 111–8.

[24] Hillaire S, Lécuyer A, Cozot R, Casiez G. Using an eye-tracking system to improve camera motions and depth-of-field blur effects in virtual environments. In: Virtual Reality Conference, 2008. VR'08. IEEE. IEEE; 2008, p. 47–50.

[25] Leroy L, Fuchs P, Moreau G. Real-time adaptive blur for reducing eye strain in stereoscopic displays. ACM Transactions on Applied Perception (TAP) 2012;9(2).

[26] Duchowski AT, House DH, Gestring J, Wang RI, Krejtz K, Krejtz I, et al. Reducing visual discomfort of 3d stereoscopic displays with gaze-contingent depth-of-field. In: Proceedings of the ACM Symposium on Applied Perception. ACM; 2014, p. 39–46.

[27] Nagata S. The binocular fusion of human vision on stereoscopic displaysfield of view and environment effects. Ergonomics 1996;39(11):1273–84.

[28] Brooker JP, Sharkey PM. Operator performance evaluation of controlled depth of field in a stereographically displayed virtual environment. In: Photonics West 2001-Electronic Imaging. International Society for Optics and Photonics; 2001, p. 408–17.

[29] Kwon YM, Jeon KW, Ki J, Shahab QM, Jo S, Kim SK. 3D Gaze Estimation and Interaction to Stereo Display. IJVR 2006;5(3):41–5.

[30] Pfeiffer T. Measuring and visualizing attention in space with 3d attention volumes. In: Proceedings of the Symposium on Eye Tracking Research and Applications. ACM; 2012, p. 29–36.

[31] Essig K, Pomplun M, Ritter H. A neural network for 3D gaze recording with binocular eye trackers. The International Journal of Parallel, Emergent and Distributed Systems 2006;21(2):79–95.

[32] Blum T, Wieczorek M, Aichert A, Tibrewal R, Navab N. The effect of out-of-focus blur on visual discomfort when using stereo displays. In: Mixed and Augmented Reality (ISMAR), 2010 9th IEEE International Symposium on. IEEE; 2010, p. 13–7.

[33] Zachmann G, et al. Virtual Reality in Assembly Simulation: Collision Detection, Simulation Algorithms, and Interaction Techniques. Fraunhofer-IRB-Verlag; 2000.

[34] Lorenceau J. Cursive writing with smooth pursuit eye movements. Current Biology 2012;22(16):1506 –9.

[35] Shebilske WL, Fisher DF. Understanding extended discourse through the eyes: How and why. In: Groner R, Menz C, Fisher DF, Monty RA, editors. Eye Movements and Psychological Functions: International Views. Hillsdale, NJ: Lawrence Erlbaum Associates; 1983, p. 303–14.

[36] Unity Manual depth of field. `https://docs.unity3d.com/460/Documentation/Manual/script-DepthOfFieldScatter.html`; 2014. Accessed: 2014.

[37] Holmqvist K, Nyström M, Andersson R, Dewhurst R, Jarodzka H, Van de Weijer J. Eye tracking: A comprehensive guide to methods and measures. OUP Oxford; 2011.