

**SANAL SAHNELERDE SEYİRCİ KONFORUNUN İÇERİK
ANALİZİ İLE GELİŞTİRİLMESİ VE GERÇEKLİK
TUTMA ETKİSİNİN AZALTILMASI**

**ENHANCING USER EXPERIENCE OF 3D DISPLAYED
VIRTUAL SCENES WITH CONTENT ANALYSIS AND
MINIMIZING EFFECTS OF VIRTUAL REALITY
SICKNESS**

Emre Avan

PROF. DR HAŞMET GÜRÇAY

Tez Danışmanı

Submitted to Institute of Sciences of Hacettepe University
as a Partial Fulfillment to the Requirements
for the Award of the Degree of Master of Science
in Computer Animation And Game Technologies

September 2019

To my beloved wife...

YAYINLANMA FİKRİ MÜLKİYET HAKKLARI BEYANI

Enstitü tarafından onaylanan lisansüstü tezimin/raporumun tamamını veya herhangi bir kısmını, basılı (kağıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma iznini Hacettepe üniversitesine verdiğimi bildiririm. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak, tezimin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanım hakları bana ait olacaktır.

Tezin kendi orijinal çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanması zorunlu metinlerin yazılı izin alarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Yükseköğretim Kurulu tarafından yayınlanan “*Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge*” kapsamında tezim aşağıda belirtilen koşullar haricince YÖK Ulusal Tez Merkezi / H. Ü. Kütüphaneleri Açık Erişim Sisteminde erişime açılır.

- Enstitü / Fakülte yönetim kurulu kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren 2 yıl ertelenmiştir.
- Enstitü / Fakülte yönetim kurulu gerekçeli kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren 6. ay ertelenmiştir.
- Tezim ile ilgili gizlilik kararı verilmiştir.

02 / 10 / 2019

Emre Avan

ABSTRACT

Providing a depth-rich Virtual Reality (VR) experience to users without causing discomfort remains to be a challenge with head-mounted displays (HMDs), which enforce strict measures on stereoscopic camera parameters. However, these measures often lead to an unimpressive VR experience with shallow depth feeling. We propose the first method ready to be used with existing consumer HMDs for automated stereoscopic camera control in virtual environments (VEs). Using radial basis function interpolation and projection matrix manipulations, our method makes it possible to significantly enhance user experience in terms of overall perceived depth while maintaining visual discomfort on a par with the default arrangement. We also introduce the first immersive interface for authoring a unique 3D stereoscopic cinematography for any VE. We conducted a user study that demonstrates the benefits of our approach in terms of superior picture quality and perceived depth. We also investigated the effects of using depth of field (DoF) with our approach and observed that DoF degrades experience.

ENHANCING USER EXPERIENCE OF 3D DISPLAYED VIRTUAL SCENES WITH CONTENT ANALYSIS AND MINIMIZING EFFECTS OF VIRTUAL REALITY SICKNESS

Emre Avan

Master of Science, Department of Computer Animation and Game Technologies

Supervisor: Prof. Dr. Haşmet Gürçay

Co-Supervisor: Dr. Öğrt. Üyesi Ufuk Çelikcan

September 2019, 51 pages

Keywords: Virtual Reality, Computer Graphics, User Experience, Radial Basis Function, Head Mounted Displays

ÖZET

Derinlik algısının kullanıcıya rahatsızlık verilmeden artırılması, çiftli (stereo) 3 boyutlu (3B) gözlükler için halen çözüm bekleyen önemli bir sorundur. Derinlik, kullanıcı konforunun korunması amacıyla stereo 3B gözlüklerde, olabildiğince düşük seviyede tutulmaktadır ve bu sebeple, sanal ortamların etkileycilik oranı azalmaktadır. Bu sorunu çözmek amacıyla, otonom çalışarak çiftli kamera parametrelerini kontrol eden bir sistem sunmaktayız. Radyal bazlı fonksiyon yardımıyla stereo kamera parametrelerini interpolate ederek projeksiyon matrisini güncelleyen sistemimizi kullanıcının sanal sahne deneyimini belirgin seviyede arttırmıştır ve bu sırada kullanıcının konfor seviyesini kötüleştirmemiştir. Ayrıca, sanal sahnelerdeki stereo parametrelerinin belirlenmesi için geliştirmiş olduğumuz sistem sayesinde, sahne yaratıcılarının kolaylıkla sanal ortamdaki derinlik algısını oluşturmaya yardımcı olduk. Yaptığımız çalışmayı, kullanıcı testleriyle test etmiş ve sunduğumuz sistemin başarılı olduğunu göstermiş bulunmaktayız. Ayrıca yapay alan derinliği algoritmalarını sunduğumuz sistemle birlikte test ederek ve yapay alan derinliğinin kullanıcı deneyimini kötü yönde etkilediğini belirtmekteyiz.

SANAL SAHNELERDE SEYİRCİ KONFORUNUN İÇERİK ANALİZİ İLE GELİŞTİRİLMESİ VE GERÇEKLIK TUTMA ETKİSİNİN AZALTILMASI

Emre Avan

Yüksek Lisans, Bilgisayar Animasyonu ve Oyun Teknolojileri

Tez Danışmanı: Prof. Dr. Haşmet Gürçay

Eş Danışman: Dr. Öğrt. Üyesi Ufuk Çelikcan

Eylül 2019, 51 sayfa

Anahtar Kelimeler: Sanal Gerçeklik, Bilgisayar Grafîği, İzleyici Konforu, Radyal Bazlı Fonksiyonlar, Sanal Gerçeklik Gözlüğü

ACKNOWLEDGEMENTS

First and foremost, I would like to express my gratitude to my supervisor Prof. Dr. Haşmet Gürçay and co-supervisor Dr. Öğrt. Üyesi Ufuk Çelikcan. Without their guidance and assistance, it would not be possible to complete this thesis.

Besides, I am grateful to Prof. Dr. Tolga Çapın for his advice with the development of the method.

I also would like to thank to TUBITAK 1001 program members Berk Cebeci and Alper Özkan for their help.

I am grateful to my friends Umut, Orçun, Alper, Tuğçe, Kerem, Mert, Ahmet, Melike, Umutcan, Mehmet Ali and all the people who have shared their valuable time.

Finally, I am grateful to my family, who have supported me throughout my years of study. This accomplishment would not have been possible without their love.

This study is part of the project “Enhancing the User Experience of 3D Displayed Virtual Scenes” supported by TUBITAK 1001 program (Project No. 116E280).

CONTENTS

ABSTRACT.....	i
ACKNOWLEDGEMENTS.....	iv
CONTENTS.....	v
FIGURES.....	vii
TABLES	viii
SYMBOLS AND ABBREVIATIONS.....	ix
1. INTRODUCTION AND MOTIVATION	1
2. BACKGROUND	5
2.1. Stereo Vision and Stereoscopic Rendering.....	5
2.2 Radial Basis Functions.....	8
3. PREVIOUS WORKS	10
4. RBF INTERPOLATION-BASED STEREOSCOPIC CAMERA CONTROL.....	14
4.1. VR Authoring	14
4.2 VR Experience.....	18
5. EVALUATION	27
5.1 System.....	27
5.2 Scenes..	27
5.3 DoF Blur Implementation.....	28
5.4 Subjects.....	28
5.5 Test Procedure	29
5.5.1 Preset Experience vs. Default Experience	29
5.5.2 Self-Authored Experience vs. Default Experience	30
5.5.3 Preset Experience with DoF blur vs. Preset Experience.....	30
5.6 Evaluation Criteria.....	31
5.7 Results.....	32
6. CONCLUSION.....	42
7. REFERENCES	43

EK 1 – Questionnaire	47
EK 2 – Academic Research Papers based on this work.....	49
Curriculum Vitae.....	51

FIGURES

Figure 1 An overview of our approach.....	3
Figure 2 Cross coupling between accommodation and convergence.....	6
Figure 3 Difference between Asymmetric Frustum (on the left) and Toe-in (on the right).....	7
Figure 4 Projection representation is shown above.....	8
Figure 5 Flow diagram of our RBF interpolation-based stereoscopic camera control approach.	14
Figure 6 Users can interact with the system using the hand controllers of HTC Vive..	15
Figure 7 The graphical user interfaces in VR Authoring:.....	16
Figure 8 preview of a sample camera path.	17
Figure 9 RBF representation.....	19
Figure 10 Asymmetric projection of Plano-Stereoscopic and Head Mounted Displays.....	21
Figure 11 Projection matrix representation for the Plano-Stereoscopic Displays.....	22
Figure 12 Projection with 0,064m <i>Dia</i> and 0.2m <i>Dzp</i>	25
Figure 13 Screenshot with default projection with IPD of 0,064m.....	25
Figure 14 Projection with 0,064m <i>Dia</i> and 6m <i>Dzp</i>	25
Figure 15 Projection with 0,15m <i>Dia</i> and 6m <i>Dzp</i>	26
Figure 16 Sample frames from the indoor scene	27
Figure 17 Sample frames from the outdoor scene	28
Figure 18 Sample frame from the outdoor scene with the added DoF blur	31
Figure 19 Sample frame from the outdoor scene without DoF blur	31
Figure 20 Individual session ratings collected with the questionnaires	33
Figure 21 Individual session ratings collected with the questionnaires	34
Figure 22 Individual session ratings collected with the questionnaires.	35
Figure 23 Preset Experience vs. Default session ratings collected	38
Figure 24 Preset Experience vs. Default session ratings collected.	39
Figure 25 Self-Authored Experience vs. Default session ratings collected	39
Figure 26 Self-Authored Experience vs. Default session ratings collected	40
Figure 27 Preset Experience with DOF Blur vs. Preset Experience session ratings collected.....	40
Figure 28 Preset Experience with DOF Blur vs. Preset Experience session ratings collected.....	41

TABLES

Table 1 Relative maximum and minimum depth value chart of the indoor scene	37
Table 2 Relative maximum and minimum depth value chart of the outdoor scene	38

SYMBOLS AND ABBREVIATIONS

Symbols

D_{ia}	Interaxial Distance
D_{zp}	Zero-Parallax Distance
w_j	Constant for Radial Basis Function
α	Small Angle of the View Frustum
β	Big Angle of the View Frustum
f	Far Plane
n	Near Plane
t	Top Plane
b	Bottom Plane
l	Left Plane
r	Right Plane

Abbreviations

VR	Virtual Reality
HMD	Head-Mounted Displays
VE	Virtual Environment
DoF	Depth of Field
FoV	Field of View
ACC	Accommodation and Convergence Conflict
RBFs	Radial Basis Functions
CoC	Circle of Confusion
GUI	Graphical User Interfaces
IPD	Interpupillary Distance

1. INTRODUCTION AND MOTIVATION

In recent years, Virtual Reality (VR) has become more commonplace with recent advances in hardware technology that have led to the production of consumer appropriate head-mounted displays (HMDs), such as HTC Vive and Oculus Rift. A wide field-of-view (FOV) HMD that immerses a user in computer-generated virtual worlds is a key enabling technology to VR applications.

The immersive nature of HMDs creates a strong presence illusion, where users perceive virtual environments (VEs) as real and not mediated through technology. Therefore, it finds a myriad of applications in gaming, entertainment, simulation and training, defense, education, and other fields. Support from the makers of the commercially available HMDs with extensive software development kits has resulted in an unseen and rapidly expanding ecosystem of such applications specifically designed for VR. As VR has been increasingly accessible and popular, comfortable, high-quality stereo 3D has become an important and timely requirement for real-time VR applications. There are still issues that remain to be resolved in order to provide a thoroughly realistic and comfortable experience to VR users. Immersive VEs can be visually constraining. The foremost contributing factor to visual discomfort is the accommodation-convergence conflict (ACC), which arises due to the dissonance between accommodation, adjustment of the eye lenses to focus at the observed depth, and eye convergence [23, 27, 36, 41]. While these two cues are cross coupled in normal viewing conditions, in stereoscopic displays, the viewer always focuses at the screen level regardless of where eyes actually converge, which leads to ACC. Due to the ensuing discomfort, users commonly report symptoms such as eye-strain, nausea, dizziness and headaches after using HMDs for extended periods. Enhancing user experience and perceived depth together without invoking discomfort has been a major challenge in stereoscopic content production.

To create stereo vision for VR applications, there are two main parameters that need to be set. These are interaxial separation, which is the distance between the two cameras, and their convergence distance in the VE. These stereoscopic camera parameters play a major role in the VR experience as they produce the disparity between left and right

images, and therefore impact the amount of perceived depth [16], as well as visual comfort. Nonetheless, no single setting exists that minimizes the fusion effort and leads to optimized depth perception for varying viewing circumstances and depth ranges [26].

Although it is possible to set stereo camera parameters freely in a controlled environment, for interactive applications, a control system is needed to supervise depth. Therefore, in current commercially available HMDs, stereo camera parameters are set to be fixed at the API level, such that, they are kept constant no matter how the scene contents change or the depth composition of the scene varies. This is a convenient solution to avoid issues regarding visual discomfort, however it reduces the depth perceived, and, in turn, level of immersion of the users.

In this work, we aim to enhance user experience in VR with consumer HMDs in terms of overall perceived depth without sacrificing picture quality or visual comfort. Addressing the challenges of changing depth composition dynamically while maintaining visual discomfort to a minimum, we propose a new method for automated stereoscopic camera control in VEs. To the best of our knowledge, this is the first such method ready to be used with existing consumer HMDs without additional hardware requirements such as embedded eye trackers or focus-adjustable lenses.

Our proposed method demands a stereoscopic cinematography, that is, a particular arrangement of stereoscopic camera settings consisting of a series of waypoint locations with designated camera angles and stereoscopic camera parameters. The stereoscopic parameters associated with these locations constitute a set of scattered data. Using radial basis function interpolation on this small set of data, our method produces a smooth surface fit of parameters for the rest of a given VE and provides automated stereoscopic camera control by continuous projection matrix manipulations using the fitted parameters.

We also introduce a VR interface for creating the required stereoscopic cinematography. Our survey of the relevant literature indicates that this is the first immersive interface that can readily be used with all commercially available HMDs for authoring a unique 3D stereoscopic cinematography. With the proposed interface, users can author unique depth narratives for any given VE directly from the first-person perspective exactly as the VE will be experienced using the same HMD that it will be experienced with. While it can be used by VR content creators to design signature depth narratives, the easy-to-use interface lets even the most novice users to quickly create their own unique VR experiences.

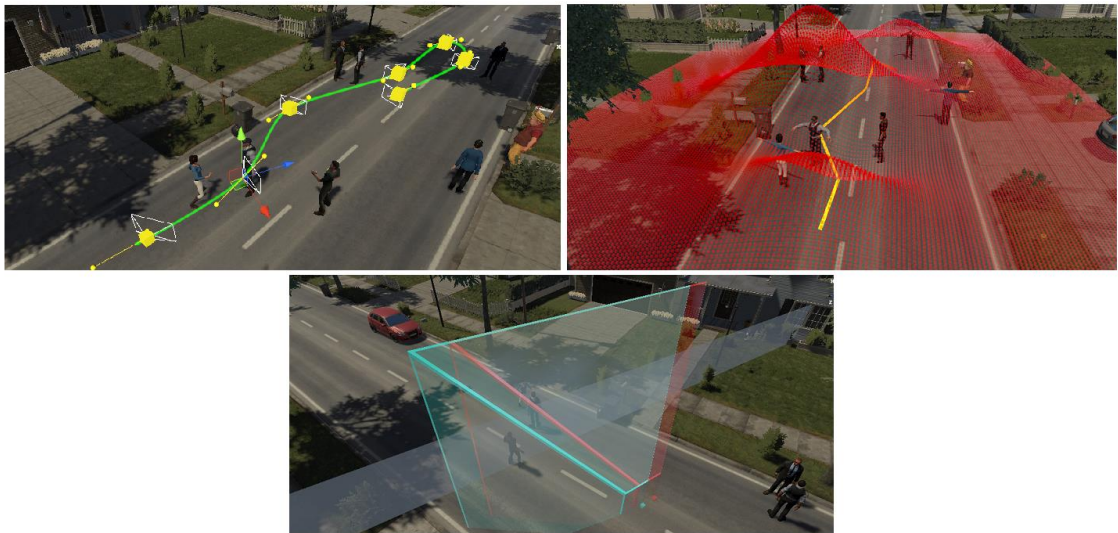


Figure 1 An overview of our approach. Clockwise from the top left: (1) The user creates a unique design of stereoscopic cinematography, which is composed of a series of waypoint locations with designated camera angles and stereoscopic camera parameters. (2) Then, the method calculates a smooth camera path and interpolates the stereoscopic camera parameters for the rest of the scene using radial basis functions. (3) Finally, during the VR experience, automated stereoscopic camera control is realized on the path by projection matrix manipulations using the fitted surfaces of parameters.

As the highly subjective nature of stereoscopic camera control necessitates, we evaluated our method in comparison to the default HMD settings in different configurations. The results illustrate that our method is able to significantly enhance user experience in terms of overall perceived depth while also boosting picture quality and maintaining visual

discomfort on a par with the default arrangement. We also analyzed whether using Depth-of-Field blur in addition to our method improves user experience further and found that, as in some similar earlier studies, the addition of our Depth-of-Field blur implementation was seen as a degraded experience notably in terms of visual comfort and similar in terms of depth perception.

The remainder of this work is organized as follows. In Sections 2 and 3, we briefly give some background information on the components of our work and present an overview of previous works on the subject matter. Our proposed method is elaborated in Section 4. Then, in Section 5, the details of our user evaluation study are given and its results are illustrated and discussed. Finally, Section 6 concludes the work.

2. BACKGROUND

2.1. Stereo Vision and Stereoscopic Rendering

Interaxial separation and *zero-parallax distance* are the two major parameters in stereoscopic rendering. For the binocular vision, there are two virtual cameras rendering the scene with a slight distance between each other. This distance is called interaxial separation and lets the human visual system to create depth perception. Both cameras need to be parallel to each other to avoid vertical asymmetry which results in discomfort [44]. Secondly, oculomotor muscles enable eyes to converge into a plane which is called *zero-parallax plane*, also known as convergence plane, and the distance between this plane and the cameras is called zero-parallax distance.

If the object is far behind the zero-parallax plane, it has positive disparity and appears as inside the display. For an object that is at a closer distance than the zero-parallax plane, it appears as in front of the display and this situation is called negative disparity. As the zero-parallax distance gets closer to the virtual cameras, image for the left eye shifted to left and right for the right eye for the HMDs. For the objects that have negative disparity, fatigue may occur more easily than the ones with positive disparity since oculomotor muscle needs to contract in order to rotate eyes inwards to focus with negative disparity. Therefore, developers and designers carefully controls the scene composition and cinematography in order to keep users' eyes rested as much as possible [13].

When looking at the screen, viewer's eyes converge according to the depth of the object in the scene while they are focused on the display. In the real-world, accommodation and convergence systems are cross-coupled, which means that eyes both converge and accommodate at the same position as seen in the Figure 2. Crystalline lens of the eye make accommodation possible by refracting the light inwards or outwards. In stereoscopic display systems, on the other hand, coupling between accommodation and convergence is broken causing ACC that can induce negative effects, e.g., eye-strain and headache.

While ACC is not a severe issue for mediums like 3D TV or cinema display, where disparity often falls within the comfort zone as the VE is placed at a large enough distance

from the eyes, for VR, the screen is placed at very close distance. Moreover, objects that are close to the virtual camera have to be rendered, which may also cause discomfort [45].

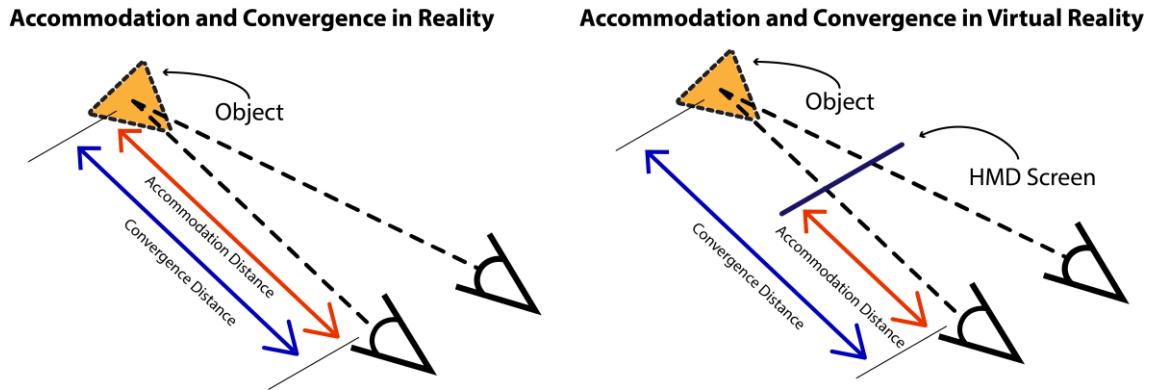


Figure 2 Cross coupling between accommodation and convergence is broken in stereoscopic display systems such as HMDs.

In order to solve this issue which can be seen in Figure 2, focus object needs to be in stereoscopic comfort zone. This zone is defined by Percival [33] to be $1/3$ diopter distance from each side (negative disparity and positive disparity) to the accommodation distance. To provide the salient VE contents within stereoscopic comfort zone, the perceived depth should be limited [43].

Typically, in plano-stereoscopic displays like 3D TV or cinema screen, where disparity often falls within the comfort zone as viewers observe the images at a large distance, image distortion is not a severe issue. But this is not the case for HMDs since the display is placed at a very close proximity to the eyes. Unnatural stereo images can cause conflicting cross-coupled interactions between the crystalline lens and oculomotor nerve systems which result in discomfort [30] and can induce VR sickness. Moreover, selected depth should be changed smoothly. Sudden changes in zero-parallax distance between near and far regions can be disorienting [42].

There are two common methods rendering the scene with a given convergence distance. One is stereo camera rotation, which is a relatively easier to develop solution than the former method for disparity range adaptation. This method is also known as toe-in stereo

and it is widely used for stereoscopic contents that is viewed at stereoscopic displays. In this method, both virtual cameras are rotated according to the zero-parallax plane distance. However, rotating cameras creates distorted images which is called Keystone distortion [4]. Vertical asymmetry occurs at the corners of the screen, since virtual cameras look at different axis which is shown at in Figure 3. HMDs such as HTC Vive and Oculus Rift also use the asymmetric projection matrix where both cameras look outwards in order to shift the scene contents inwards.

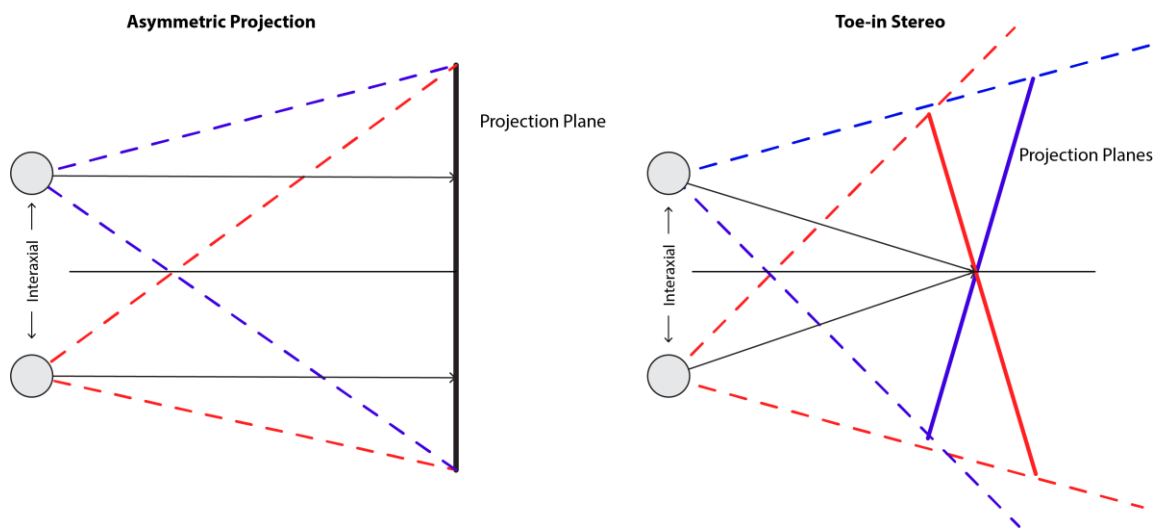


Figure 3 Difference between Asymmetric Frustum (on the left) and Toe-in (on the right) camera setups.

The other method is called asymmetric projection [4], also known as off-axis projection, which is used by current consumer HMDs. With this method, the projection matrix is calculated to make cameras have an asymmetric frustum, such that, one side of the view frustum has more area to be rendered, making one side of the view frustum more condensed and the other side expanded. Since the vertical asymmetry is avoided, keystone distortion does not occur in asymmetric projection. Moreover, this projection method is used in HMDs such as HTC Vive and Oculus Rift.

It is also stated by Oculus Rift developers website [46] that center of the lenses that are used in VR headsets is not always the center of the screen. By matching them, undesired

defects like aberration can be minimized. Moreover, there is more pixels to be rendered since the centers will be matched.

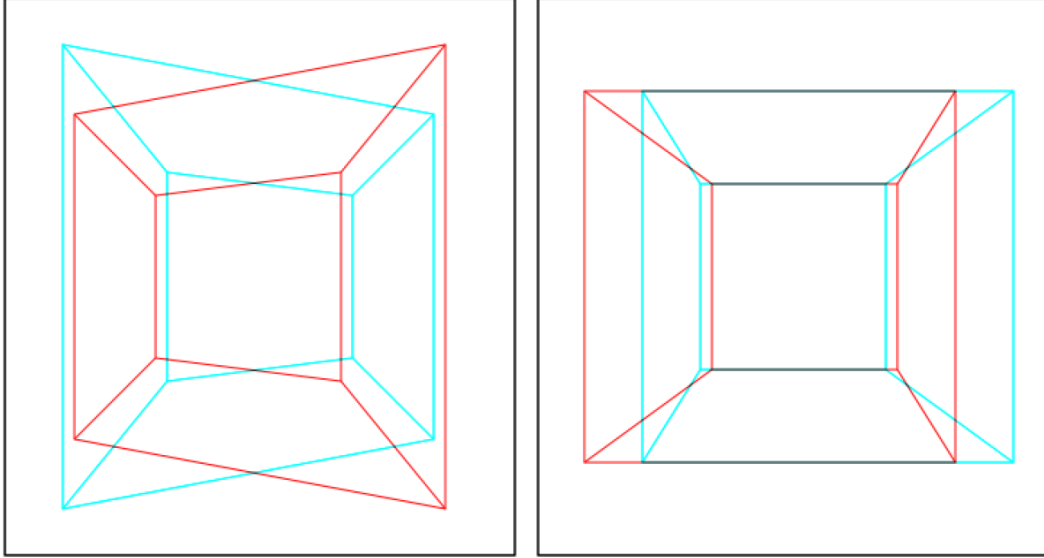


Figure 4 Projection representation is shown above. Left image shows the Toe-In projection, while right image shows Asymmetric Projection [48].

2.2 Radial Basis Functions

Radial Basis Functions (RBFs) are identified as one of the most accurate and stable methods of solving problems involving scattered data interpolation [11, 34, 35]. It is possible to obtain an implicit surface representation approximating the scattered data using RBFs.

RBF is a radially symmetric function centered around a point x_c . For a collection of scattered data $\mathbb{X} = \{x_1, x_2, \dots, x_n\}$ of n distinct data points and a matching collection of values y_1, y_2, \dots, y_n sampled from an unknown function f s.t. $y_i = f(x_i)$ we can select an RBF Φ and a group of centers $\{x_{c_1}, x_{c_2}, \dots, x_{c_m}\}, m \in \mathbb{N}$, to form a basis $\{\Phi(\|\cdot - x_{c_1}\|), \Phi(\|\cdot - x_{c_2}\|), \dots, \Phi(\|\cdot - x_{c_m}\|)\}$. Then, this basis can be utilized to construct an approximation \tilde{f} of f , such that there will be one basis function $\forall i \{1, 2, \dots, n\}$ where x_i is

the center x_c of that function. Thus, the approximate function \tilde{f} can be assembled with those n RBFs as

$$\tilde{f}(x) = \sum_{j=1}^n w_j \Phi(\|x - x_j\|)$$

with constant w_j .

RBFs have been utilized in various other areas of computer graphics. Carr et al. used RBFs to reconstruct smooth, manifold surfaces from point-cloud data [7]. In the work by Noh et al., RBFs were utilized for creating deformations of polygonal models to produce animations by localized real-time deformations [28]. Weissmann and Salomon used RBFs in gesture recognition for VR applications [37]. Yet, to the best of our knowledge, there has not been a work making use of RBFs for stereoscopic rendering. In this work, we will introduce RBF interpolation based disparity refinement approach that continuously adjusts stereoscopic camera parameters according to user input.

3. PREVIOUS WORKS

The main concern of content creators while designing a virtual environment for stereo displays is to keep the salient contents within the comfort zone where user can experience the scene without discomfort while maintaining the depth perception. In VEs where the camera is in constant motion, a control system is needed to keep the perceived depth range within the visual comfort zone. Lang et al. [24] proposed a stereoscopic camera controller for post-production disparity range adjustment while Oskam et al. [31] introduced a controller for real-time disparity range adaptation by optimizing stereoscopic camera rendering parameters. Celikkan et al. [8] suggested another real-time optimization based controller that addresses the interaction of binocular depth perception and object saliency by factoring viewer's individual stereoscopic disparity range into the optimization. However, these methods were developed for plano-stereoscopic displays and have not been adapted to work with HMDs.

In the early works, focus adjustable lenses for dynamic convergence or monovision are proposed to solve ACC. Focus adjustable lenses allow users to accommodate accurately and minimize decoupling between accommodation and convergence. However, focus adjustable lenses are still in development and it is not possible to simply use them as add-ons with current consumer HMDs. Monovision, on the other hand, uses two different focal point lenses to let one eye to focus near while the other focuses far distances. According to the studies [32, 18], while monovision reduces fatigue, eye irritation and headache, viewer comfort is not improved.

There have been various works suggesting the use of Depth-of-Field (DoF) blur in stereo vision systems as a remedy for visual discomfort. DoF is a well-known effect in photography and cinematography. It is used for increasing disparity between focus object and background or foreground to enhance user experience. %This effect occurs because of the nature of the lenses (e.g., in human visual system or camera lenses). %Lens allows light to converge on the camera sensor (or retina of the eye) but the convergence region will change depending on the sensor size, the distance between the lens and the sensor and the distance between the lens and the light source. At a certain distance from the lens, the light will converge to a single point on the sensor (or retina), while other distances

will converge on a region called Circle of Confusion (CoC). In computer graphics, a pinhole camera model is used for rendering. In this model every object in the scene looks sharp independent from its distance to the camera. Therefore, to imitate depth of field effect, blur must be calculated. After the scene is rendered into a 2D texture, every pixels contain information about only one object, therefore it is possible to retrieve their depth value and calculate their convergence point on the sensor. According to this information, algorithm re-renders pixels inside the CoC area by averaging their values into values of neighboring pixels. With DoF blur effect, regions that are far away from the focus plane are blurred to simulate how human visual system perceives objects in full sharpness only within some range round the focal distance.

O'Hare et al. [29] suggested that DoF blur can be used to supplement perceived depth. It is also noted that the effect would only be useful when the blur gradients do not cause discomfort themselves. If blurred areas contain objects that are salient for the viewer, it will degrade viewer's experience and cause even more discomfort. Koulieris et al. [20] also suggest the use of DoF rendering towards achieving goals such as realistic sense of depth and adjusting the perceived scale.

According to Carnegie and Rhee [6], unlike plano-stereoscopic displays, people rather use head motions instead of eye motions for spatial movement and tend to look at the center of the display while experiencing VEs with HMDs. Therefore, it is suggested that using DoF blur to keep the center of the screen in focus will reduce discomfort for many users. In their study, participants felt less discomfort while experiencing the scene with DoF blur.

On the other hand, Lagnbehn et al. [25] reported that blur is not crucial for the human visual system to extract depth and motion information according to the findings of their study. They state that the HMDs' narrow FOV and low resolutions are the chief reasons behind this.

It is also found that there is a trade-off between visual comfort and sense of presence when blurring is employed; and, therefore, an optimization scheme is advised [30]. Conti et al. [9] studied the impact of DoF blur with varying interaxial camera distance. They examined configurations in the presence or absence of DoF blur and found that, as with some earlier similar attempts [5, 10], the configurations with DoF blur obtained lower subjective evaluations.

At any rate, to be able to use DoF blurring approach effectively, it is imperative to have, or at least estimate, user's instant gaze location which is not possible without using a high-precision eye tracker embedded inside the HMD or accurately predicting the most salient region of the frame in real time.

When available, gaze data can also be used for direct disparity manipulations in order to improve visual comfort by reducing ACC, frame violation, and crosstalk [15], and enhance depth perception [17]. However, with the existing commercial stereo displays and HMDs, it is not yet possible to find where the user's eyes converge. Therefore different methods have been proposed to solve ACC where the convergence plane is estimated to be.

Based on the assumption that objects with a significant contrast to their surrounding regions are more likely to draw attention, Lei et al. [47] proposed a disparity control method that combines depth information with a regional contrast-based saliency model.

To find the zero-parallax distance, Shertyuk and State [38] suggest that human eyes will mostly maintain convergence on the hand as a contact point when performing direct object access and manipulations. With that assumption, they find the target object's position and rotate the cameras accordingly. Summer [39] uses a similar camera rotation approach positioning the zero-parallax plane where the target object at the center of the screen is. Koulieris et al. [18], on the other hand, designed a machine learning gaze predictor based on game variables and eye tracking data from a video game. After finding

the zero-parallax plane, off-center projection matrix shifts that plane into the comfort zone.

While some gaze prediction-based methods have been shown to be effective at improving VR experience of users within some restricted settings, each has certain shortcomings such as limited usability, low accuracy or low frame rate. While equipping HMDs with eye trackers can solve most of these issues by foveated rendering, they are not available with current consumer HMDs, with the exception of HTC Vive Pro Eye, which remains to be a niche product mainly due to its considerably higher price. To the best of our knowledge, our work presented in this work is the first that does not require gaze detection or prediction and can readily be used with existing consumer HMDs for automated stereoscopic camera control in VEs towards enhancing the VR experience.

4. RBF INTERPOLATION-BASED STEREOSCOPIC CAMERA CONTROL

The overview of our RBF interpolation-based stereoscopic camera control approach is shown in Figure 5. The approach consists of two parts, namely, VR Authoring and VR Experience. While VR Experience is the main part where the automatic camera control takes place, it requires an arrangement of stereoscopic camera settings, a *depth layout*, that is created in VR Authoring. Users can promptly start their VR Experience with one of the preset depth layouts or use VR Authoring to create their own depth layouts.

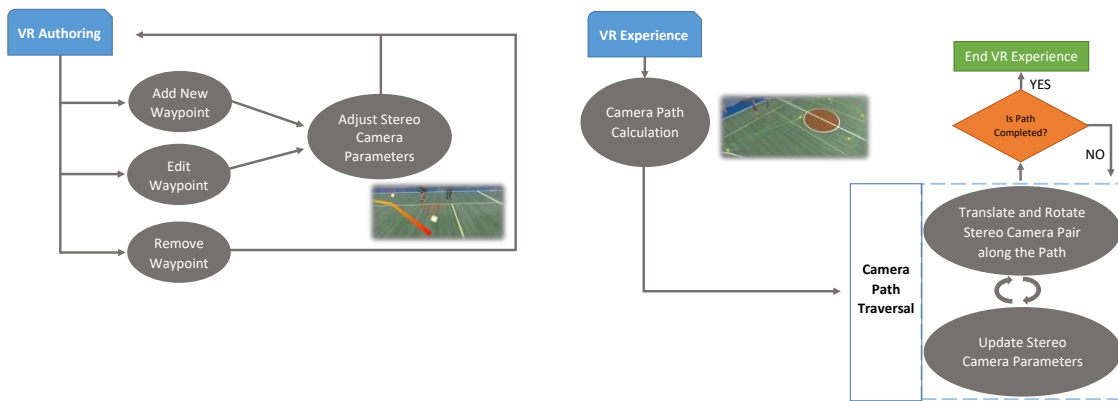


Figure 5 Flow diagram of our RBF interpolation-based stereoscopic camera control approach.

4.1. VR Authoring

In VR Authoring, the user designs the *depth narrative* of the VR experience by creating a depth layout. A depth layout is a unique design of stereoscopic cinematography that consists of a path for the stereo camera pair to follow with designated camera angles and stereoscopic camera parameters. A depth layout is composed of a series of waypoints. By placing a waypoint, the user defines a position and an angle for the stereo camera pair and sets the stereoscopic camera parameters.

During authoring, the user is placed in the VE that they are creating a depth layout for. This way, they can tailor a depth narrative to that VE from the first-person perspective in the same immersive setting using the same display (HMD) viewers will experience it

with. This aspect of our approach constitutes a significant improvement over the traditional stereoscopic editing paradigm where users are bound to work with a two-dimensional interface on a two-dimensional display.

The user can roam the VE and place a waypoint anywhere in it freely. Movement in the VE is realized by either materially moving in the physical space or virtually teleporting within the VE. With our editing in first-person view paradigm, position and angle of the stereo camera pair are controlled by the user via HMD in the same way a cameraman operates their camera in live shooting. Upon locking the position and view angle for a waypoint, the user adjusts the stereoscopic camera parameters with the corresponding dials on the hand controllers Figure 6. Interaxial-distance between stereo cameras t_c can be in between $0m$ to $0.30m$ and convergence distance Z_c can be $0m$ to $10m$. Those bounds are defined in order to protect users' depth perception in the authoring mode since after path is created, user will experience the scene after a short break.

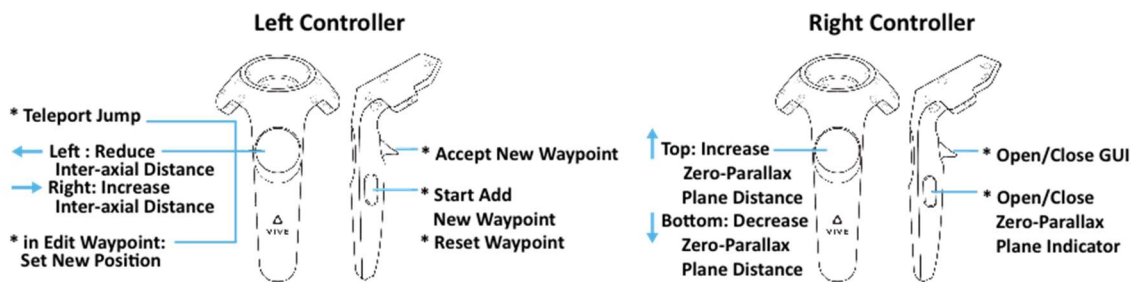


Figure 6 Users can interact with the system using the hand controllers of HTC Vive. The printed version of this figure is shown to the subjects before the testing starts to have them familiarized with the interface. Hand controller images courtesy of HTC Vive [1].

We developed the physical and graphical interfaces for VR Authoring with minimality and ease-of-use as our primary design precepts. Figure 6 shows the graphical user interfaces (GUI) of VR Authoring. At the start of authoring, in accordance with our design precepts, no GUI is displayed to the user so that they can start and continue adding waypoints with minimal interruption until they decide to finish authoring.

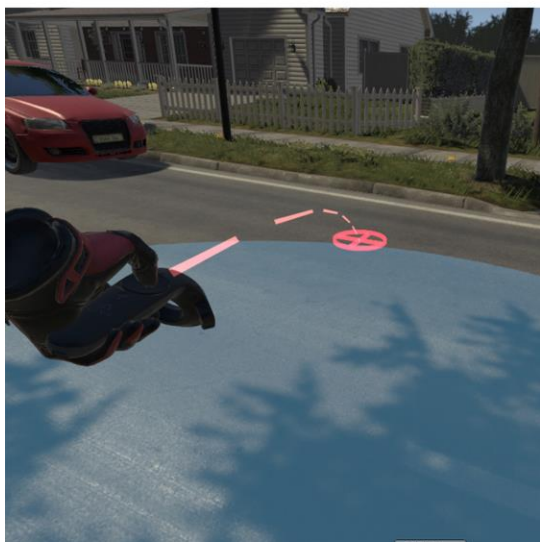
At any point during authoring, the user can reveal the GUI main menu Figure 7(a) by pressing the open/close GUI button on the right hand controller Figure 6 and choose to remove the last added waypoint, edit existing waypoints or finish authoring by selecting the corresponding button. If Edit Waypoint is selected, a submenu is shown Figure 7(b). With the submenu, the user can move between existing waypoints to select one for readjusting its position, view angle or stereo camera parameters. The only difference in editing an existing waypoint than adding a new one is that the user now can change the position of an existing waypoint within a limited radius around the current position Figure 7(c).



(a)



(b)



(c)



(d)

Figure 7 The graphical user interfaces in VR Authoring: (a) main menu (b) submenu for Edit Waypoint (c) the user is restricted to changing the position of an existing waypoint

within a certain radius (d) close-up of a waypoint marker with its saved zero-parallax distance and interaxial separation hovering above

The user can define as many waypoints as they like. After each waypoint is saved, the user can see the actual progress by the preview path with gradient colored lines Figure 8 in-between the placed waypoints markers, which are indicated by cubes and have the saved stereo camera parameters hovering over them Figure 7(d).



Figure 8 preview of a sample camera path.

When the user is content with the depth layout, they can exit VR Authoring by pressing Finish Authoring button in the main menu. Then, the user can switch to the VR Experience mode to preview the layout and then return back to authoring for further additions, removals or adjustments of waypoints.

If play mode is selected, screen goes to black and there will be a short break of 15 seconds in order to relieve users' eye muscles.

4.2 VR Experience

VR experience will start with reading the given stereo camera parameters and waypoints. In this mode, there will be two stereo cameras, one for the left eye and one for the right eye. Unity enables developers to render the scene into a single eye without any issue, we used that method since it will not be possible to increase or decrease interaxial separation other than the interpupillary distance which is hardware bound.

In this mode, our method continuously updates stereoscopic camera parameters, interaxial separation of the stereo camera pair (D_{ia}) and zero-parallax distance (D_{zp}), while following the camera path. For smooth camera translations, the piecewise linear path designated in the selected depth layout is converted to a Bezier curve using waypoints as the control points. Spherical linear interpolation is used to rotate the camera pair from one waypoint's view angle to the next one's. Head motions coming from the HMD are omitted in order to have the viewers experience the stereoscopic cinematography exactly as authored.

Before the experience starts, for each stereoscopic camera parameter, a separate set of RBF interpolation weights $w_j, j = 1, \dots, m$ are first found by solving the the set of m equations

$$F(x_i) = \sum_{j=1}^m w_j \Phi(\|x_i - x_j\|), i = 1, \dots, m$$

where m is the number of waypoints, x_i is the position of the waypoint i , $\Phi(\cdot)$ is the designated basis function, and $F(x_i)$ is the value of that parameter set for the waypoint i during authoring.

During the experience, for an arbitrary point on the camera path given by the two-dimensional coordinates x , the interpolated value $F(x)$ of the parameter is computed by

$$F(x) = \sum_{j=1}^m w_j \Phi(\|x - x_j\|)$$

with respect to the computed weights w and the chosen basis function $\Phi(\cdot)$. We use inverse multiquadric basis function $\Phi(r) = 1/\sqrt{r^2 + r_0^2}$ with $r_0 = 2$ in our approach. Inverse multiquadrics are found to provide excellent approximations even when the number of centers m is small [11].

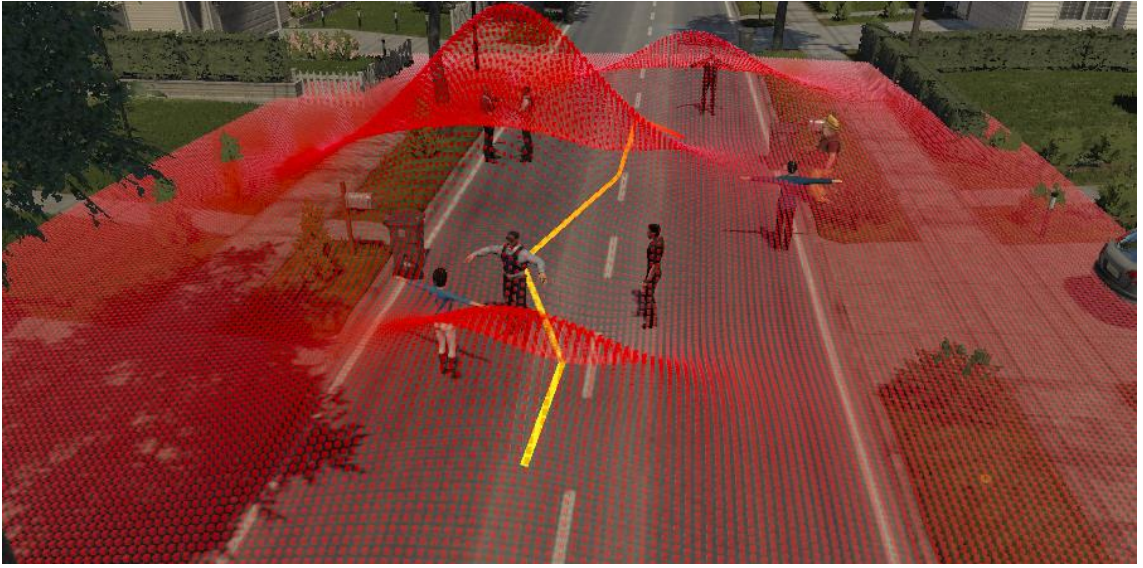
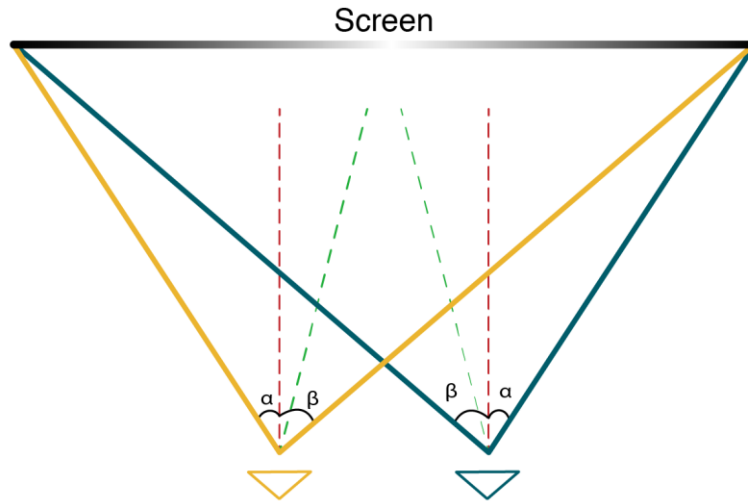


Figure 9 RBF representation. Higher red dots means that bigger interaxial or zero parallax distance

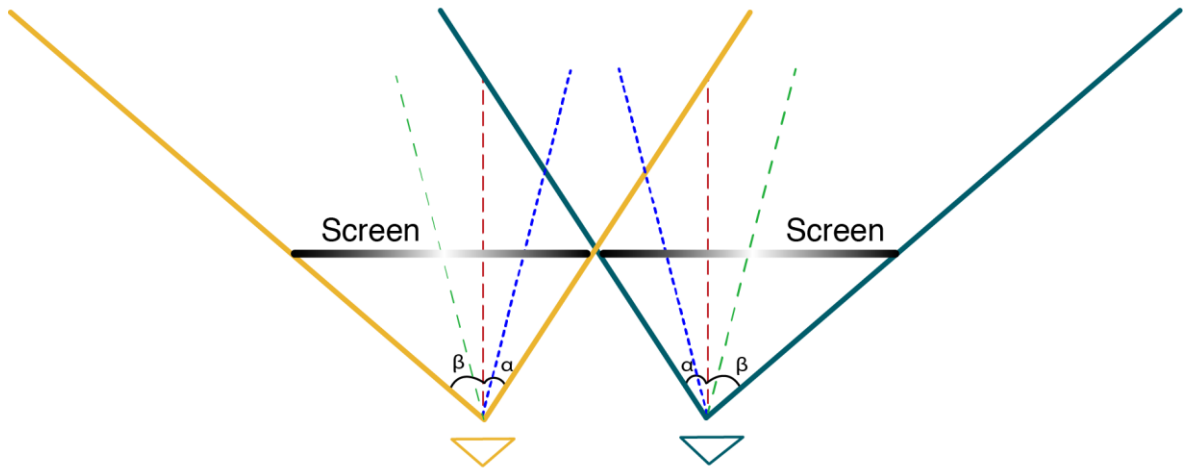
Then, the projection matrices for the two cameras are modified with the updated parameters D_{ia} and D_{zp} based on dynamic projection matrix calculation method for stereo 3D [18]. According to this method, both cameras will have asymmetric projection matrices that together create a single viewing plane at the selected zero-parallax distance. This is suitable for plano-stereoscopic displays since there is only one display for eyes to converge and focus. For the HMDs, on the other hand, there is a dedicated display for each eye rendering using the method as is infeasible. Figure 10 (a) shows the view frusta of the cameras from top-down point of view. Here, when α is larger than β , it triggers eyes to rotate inwards. As the difference between α and β gets larger, the viewer's eyes rotate more and more inwards, which creates the feeling of looking at a nearby object. Leveraging this mechanism, in our approach, the projection matrices for the left and the

right eye are swapped, so that, the cameras are directed more outwards as the zero-parallax plane comes closer and less outwards as the plane goes farther to direct viewer attention properly. As it is seen in Figure 10 (b) there are independent screens for both eyes, therefore to fuse into single image, focus object needs to be placed in a position where users' eyes converge. In order enable convergence, objects needs to be placed in the virtual scene accordingly (i.e. not in a position where user needs to diverge their eyes to focus).

Contraction of the eye muscles let human to understand depth of the object. To increase depth perception of the virtual environment, we use this phenomena and change the convergence angle of the objects according to our RBF calculations.



(a) Stereoscopic Rendering in Plano-Stereoscopic Displays



(b) Stereoscopic Rendering in Head Mounted Displays

Figure 10 Asymmetric projection of Plano-Stereoscopic and Head Mounted Displays as seen from top-down view. Dashed green line represents the true half angle between left and right planes while dashed red line separates α and β angles. Blue line on the (b) represents the users' focus point on the scene.

To change convergence angle, we need to update projection matrix of the virtual cameras. Projection matrix can be defined as:

$$\begin{bmatrix} \frac{2n}{r-l} & 0 & \frac{r+l}{r-l} & 0 \\ 0 & \frac{2n}{t-b} & \frac{t+b}{t-b} & 0 \\ 0 & 0 & \frac{-(f+n)}{f-b} & \frac{-2fn}{f-n} \\ 0 & 0 & -1 & 0 \end{bmatrix}$$

Equation 1

where l, r, n, f, t and b denote left, right, near, far, top and bottom plane positions of the view frustum, respectively. $(\frac{r+l}{r-l})$ creates horizontal off-axis asymmetry while $(\frac{t+b}{t-b})$ creates vertical off-axis asymmetry [3]. In our approach, we alter the horizontal asymmetry and leave the vertical asymmetry (i.e., t and b values) which is due to the inherent vertical asymmetry of the HMD lenses, unchanged.

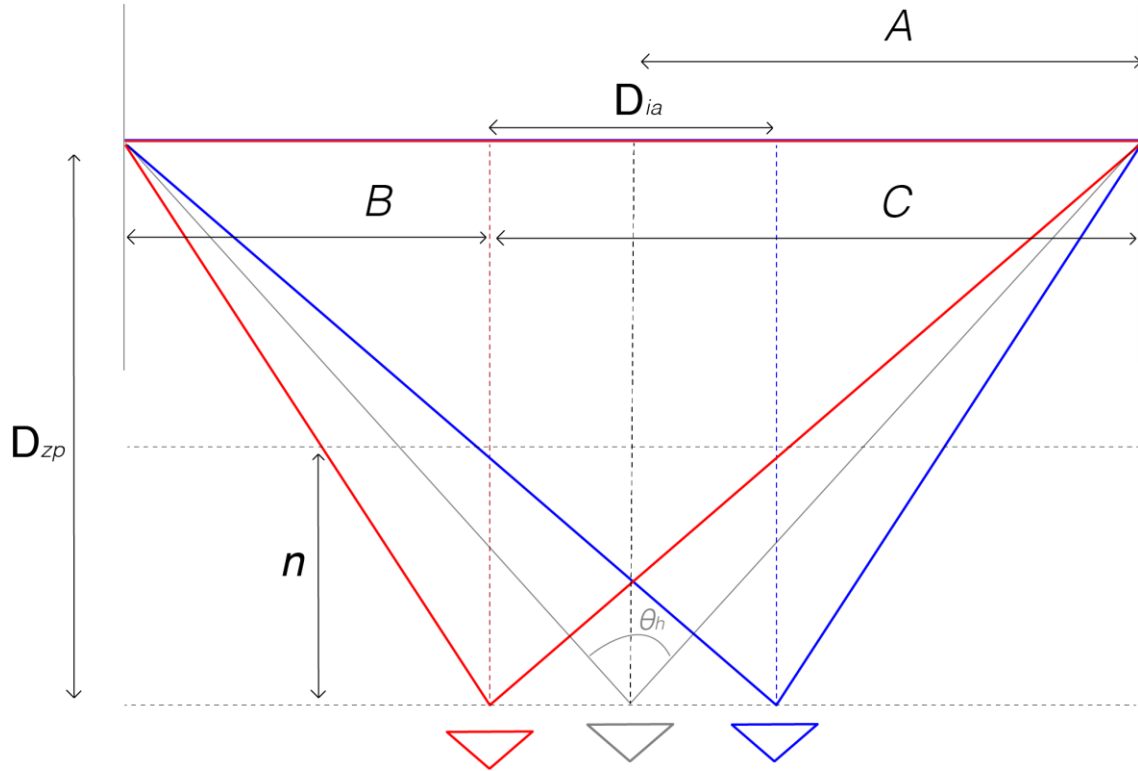


Figure 11 Projection matrix representation for the Plano-Stereoscopic Displays. θ_h means horizontal field of view angle.

Symmetric frustum is used to calculate l and r , then half of D_{ia} is added to the expanded side (l to the left virtual camera and r to the right virtual camera) and subtracted from the

condensed side at the zero-parallax distance. After that Thales theorem is used to calculate left and right planes. For stereoscopic systems, at the convergence, both cameras see the same plane, so the only difference between the condensed side and the expanded side is D_{ia} . Hence, in our approach, l and r are calculated as follows. A , the half of the frusta's horizontal width at D_{zp} , is found as

$$A = D_{zp} q \tan\left(\frac{\theta}{2}\right)$$

where θ is the vertical FOV of HTC Vive which is 111.6346° and q is the aspect ratio of the screen is 0.9495075. At D_{zp} , the condensed side's horizontal width B is half of D_{ia} narrower than A and the expanded side's horizontal width C is half of D_{ia} wider than A , i.e.,

$$B = A - \frac{D_{ia}}{2}$$

and

$$C = A + \frac{D_{ia}}{2}$$

Since we swap the left and the right view frusta, l and r values are exchanged, i.e., for the left frustum, they are given as

$$l = -C \frac{n}{D_{zp}} \quad r = B \frac{n}{D_{zp}}$$

and for the right frustum, they are given as

$$l = -B \frac{n}{D_{zp}} \quad r = C \frac{n}{D_{zp}}$$

The new l and r values per view frustum are then plugged in P given in Equation 1.

Unless additional measures are taken, an important drawback of this method is its sensitivity to D_{zp} . Diplopia will occur at large D_{zp} values and small D_{zp} values will cause eye strain and fatigue [13, 21] In order to remedy this by defining a suitable range for fusion, our system restricts D_{zp} in a dynamic range according to immediate D_{ia} that lets the difference between α and β Figure 10 to change within -10° and $+1^\circ$

According to projection matrix horizontal asymmetry value of the Steam VR stereo projection matrix, $\frac{r+l}{r-l} = -0.05624$ for the left eye and 0.05624 for the right eye, and our matrix calculation method, it can be deduced that virtual cameras of the HTC Vive have a 85.73° asymmetry which means that both of them are 4.27° off from front vector.

To find a suitable area, interaxial distance between stereo cameras will be given as an input to our frustum calculator. We need to restrict our approach with a system that decides the maximum and the minimum D_{zp} . This solution is implemented into our system because, without any physical changes in the HMDs, it will not be possible to solve diplopia issues at far zero-parallax distances.

To solve this issue, we calculated the zero-parallax distance and asymmetric angles of the HTC Vive and, while take D_{ia} into consideration, our system restricts the asymmetry according to the minimum and the maximum distances. To find the minimum and the maximum distances, we found the actual zero-parallax distance and asymmetric angles of HTC Vive.

Default projection of the HTC Vive can be seen at Figure 13. Other figures (Figure 12 Figure 14 Figure 15) shows the projections of the HTC Vive with given parameters. As the zero-parallax distance reduces, object gets closer to the corner of the screens. As the distance increases, object go farther, and if the calculated distance is past, double vision is occurred.

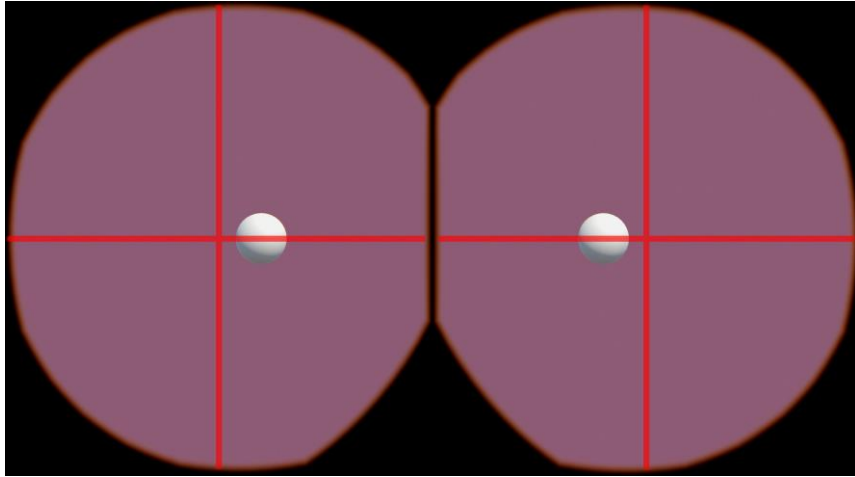


Figure 12 Projection with 0,064m D_{ia} and 0.2m D_{zp}

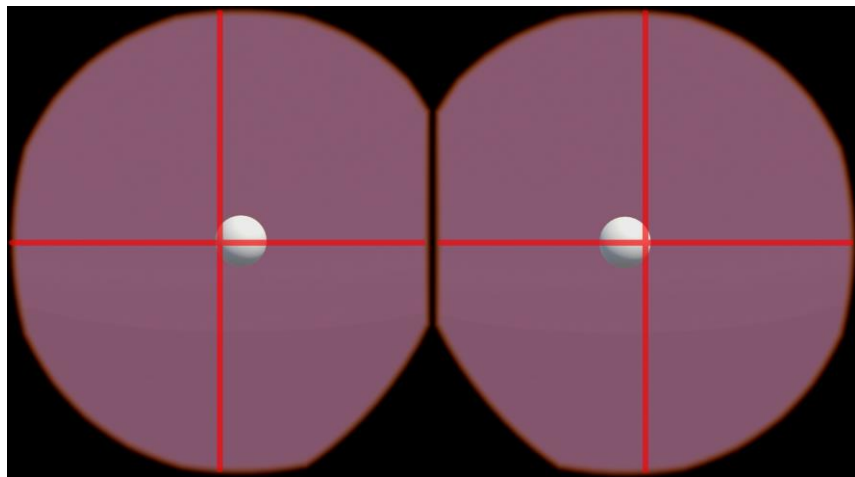


Figure 13 Screenshot with default projection with IPD of 0,064m

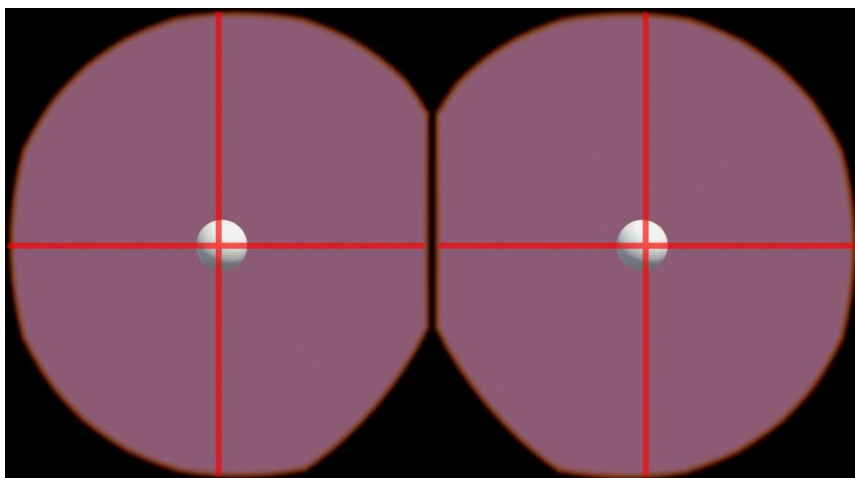


Figure 14 Projection with 0,064m D_{ia} and 6m D_{zp}

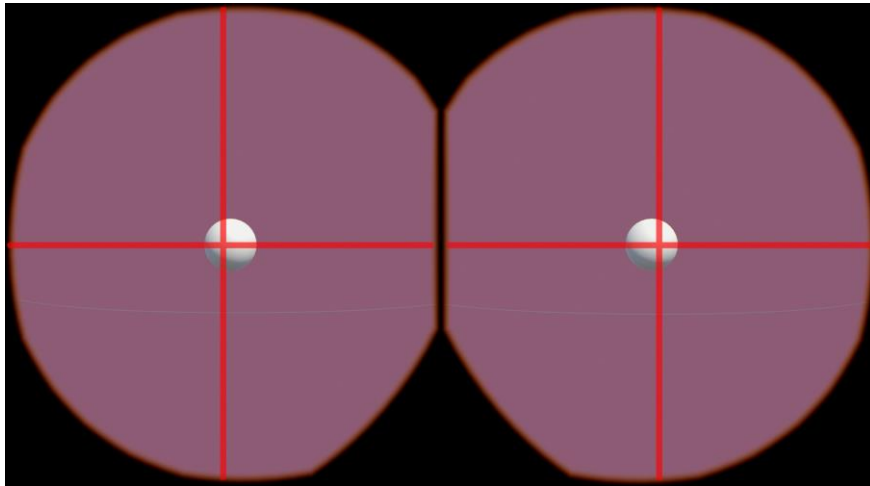


Figure 15 Projection with 0,15m D_{ia} and 6m D_{zp}

5. EVALUATION

5.1 System

The VR framework was realized using Unity graphics development engine. For the RBF-based interpolation, we used the library written in C++ by John Burkardt [2] as a plugin to the framework. During the user evaluations, HTC Vive VR setup was used with a laptop having 3.9 GHz Intel i7 CPU, 16 GB RAM and NVIDIA GTX 1070 GDDR5 8GB GPU. Frame generation rate during the tests varied between 89 and 91 FPS.

5.2 Scenes

We conducted our evaluation using two different scenes. One of the scenes takes place indoors in a school gym with multiple animated athletes and sports fans Figure 16. The other one is an outdoor scene with multiple animated human characters that are situated along a suburban residential street Figure 17.



Figure 16 Sample frames from the indoor scene



Figure 17 Sample frames from the outdoor scene

Maximum render depth is considerably higher in the outdoor scene (see Table 1 and Table 2). An often omitted aspect in similar studies, this two-tiered design has the purpose of assessing the impact of available depth range on viewer's perception, e.g., viewer's focus may change more abruptly or frequently between far and near regions when more depth is available and this may cause visual fatigue and eye-strain if the evaluated setting does not offer proper compensation.

5.3 DoF Blur Implementation

In our user study, we also test the impact of employing DoF blur on visual discomfort. We used Unity's Post Processing Stack V2 for DoF blur which applies scatter as gather algorithm [14] with multisample anti-aliasing. The focal length and aperture values were set to $75mm$ and $f/5.4$, respectively, in our tests. Focal distance of DoF is dynamically matched to zero-parallax distance (see Figure 18 and Figure 19).

5.4 Subjects

We recruited 10 subjects (4 females, 6 males, of ages between 24 and 32 with an average of 28.1) for the tests.

All subjects were volunteers and were not compensated for their participation in any form. Before the tests, each volunteer was assessed for proper stereoscopic visual acuity and those who failed did not participate in the evaluation.

The subjects had adequate technical background to use the system. They were not informed about the purpose of the experiment. Only 3 subjects had extensive experience with VR and one of those had done VR application development in the past.

5.5 Test Procedure

The subjects were informed about the tasks involved and the key points that they will rate prior to the tests with written and verbal instructions. While authoring was carried out standing, the subjects were seated while they were experiencing the scenes for the sake of comfort.

In VR systems, the initial positioning of the two virtual cameras, which compute the images for the two eyes, is perturbed by the interpupillary distance (IPD) setting. This setting is adjusted on the HMD according to the viewer's own IPD, the distance between the centers of their pupils. When the IPD is set incorrectly, perception errors and eye strain may follow [12]. Therefore, at the beginning of the tests, the IPD of a subject is measured using a digital pupilometer and set the IPD value of the HMD correspondingly.

During the tests, our approach was evaluated by the subjects in pairs of sessions for each scene in three different configurations, as given below. Order of the configurations was randomized during the tests.

5.5.1 Preset Experience vs. Default Experience

First, the subjects experienced each of the two scenes once in a session with our approach via a preset depth layout (*Preset Experience*) and once in another session with Vive's fixed stereoscopic camera parameters (*Default Experience*). During the Default Experience session, the virtual stereo camera pair follows the same path, with the same

camera positions and viewpoints, as the one that was created in the preset depth layout picked for our approach. Order of sessions was randomized at run time and the subjects were not informed about the order.

5.5.2 Self-Authored Experience vs. Default Experience

The second configuration is similar to the first one, except this time, the subjects first used VR Authoring to create their own depth layout. They were asked to design two depth layouts in total, one for each VR scene. While the subjects were free to shape their layouts according to their liking, they were told to keep the threshold to a minimum of 5 for the number of waypoints to create within a layout.

Afterwards, they again experienced each scene in two sessions, once with our approach via their self-authored depth layout (*Self-Authored Experience*) and once with the Default Experience setting, in a random order.

5.5.3 Preset Experience with DoF blur vs. Preset Experience

In the last configuration, the subjects experienced the two scenes using our approach only, once without DoF blur (Preset Experience, as in the first configuration) and once with DoF blur (*Preset Experience with DoF blur*), via the same preset depth layout (see Figure 18 and Figure 19). Order of sessions was randomized again without informing the subjects.



Figure 18 Sample frame from the outdoor scene with the added DoF blur starting at zero-parallax distance



Figure 19 Sample frame from the outdoor scene without DoF blur

5.6 Evaluation Criteria

Following each session, the subjects were asked to evaluate the session in terms of image quality, perceived depth and visual comfort in a 5-point Likert scale with the labels “bad”,

“poor”, “fair”, “good”, and “excellent”. These criteria, which are detailed below, are frequently resorted to for perceptual assessment of stereoscopic contents [40].

Image Quality: expresses the overall visual quality of the displayed content as perceived by the user. Since our approach dynamically modifies the degree of horizontal asymmetry of the frusta, proper fusion of the resulting left and right images and proper scaling of the scene contents in these images should be established by validating the image quality.

Perceived Depth: denotes the apparent depth of the displayed content perceived by the user. With stereoscopic camera control methods aiming at a dynamic depth narrative, as the stereo camera parameters change over time, so does the perceived depth. Accordingly, providing a method that brings upon a feeling of realistic depth is essential for an enhanced VR experience since it contributes a great deal to user immersion.

Visual Comfort: is to measure the subjective feeling of visual comfort of the user. Improperly set stereoscopic camera parameters cause visual discomfort in the form of eye strain, which can lead to adverse side effects including visual fatigue, nausea and headaches and yield a dissatisfactory VR experience. Hence, first of all, it is vital to ensure that a proposed stereoscopic camera control method does not invoke visual distress.

Once both sessions and their respective individual assessments are done, the subjects were then asked to evaluate the two sessions vis-a-vis each other. This time, in addition to the previous three, they were also to make a comparison in terms of overall quality.

5.7 Results

Figure 20 illustrates the results of the individual session evaluations for image quality, depth perception and visual comfort for all three configurations. The user preferences in percentages in the three configurations are given in Figure 23, Figure 24, Figure 25, Figure 26, Figure 27, Figure 28.

Preset Experience vs. Default

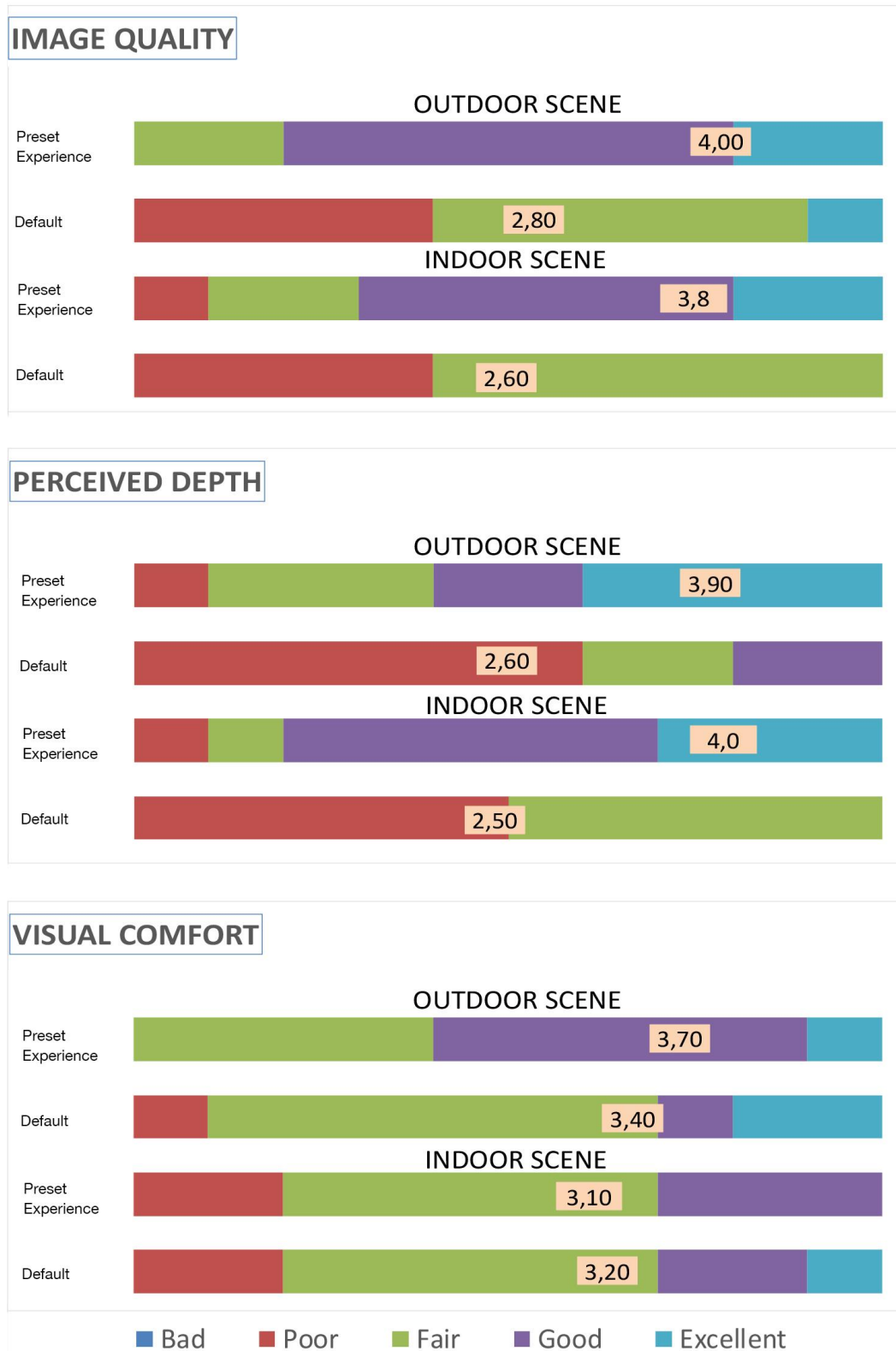


Figure 20 Individual session ratings collected with the questionnaires and their averages are given for each scene per configuration. The averages are indicated in rectangles.

Self-Authored Experience vs. Default

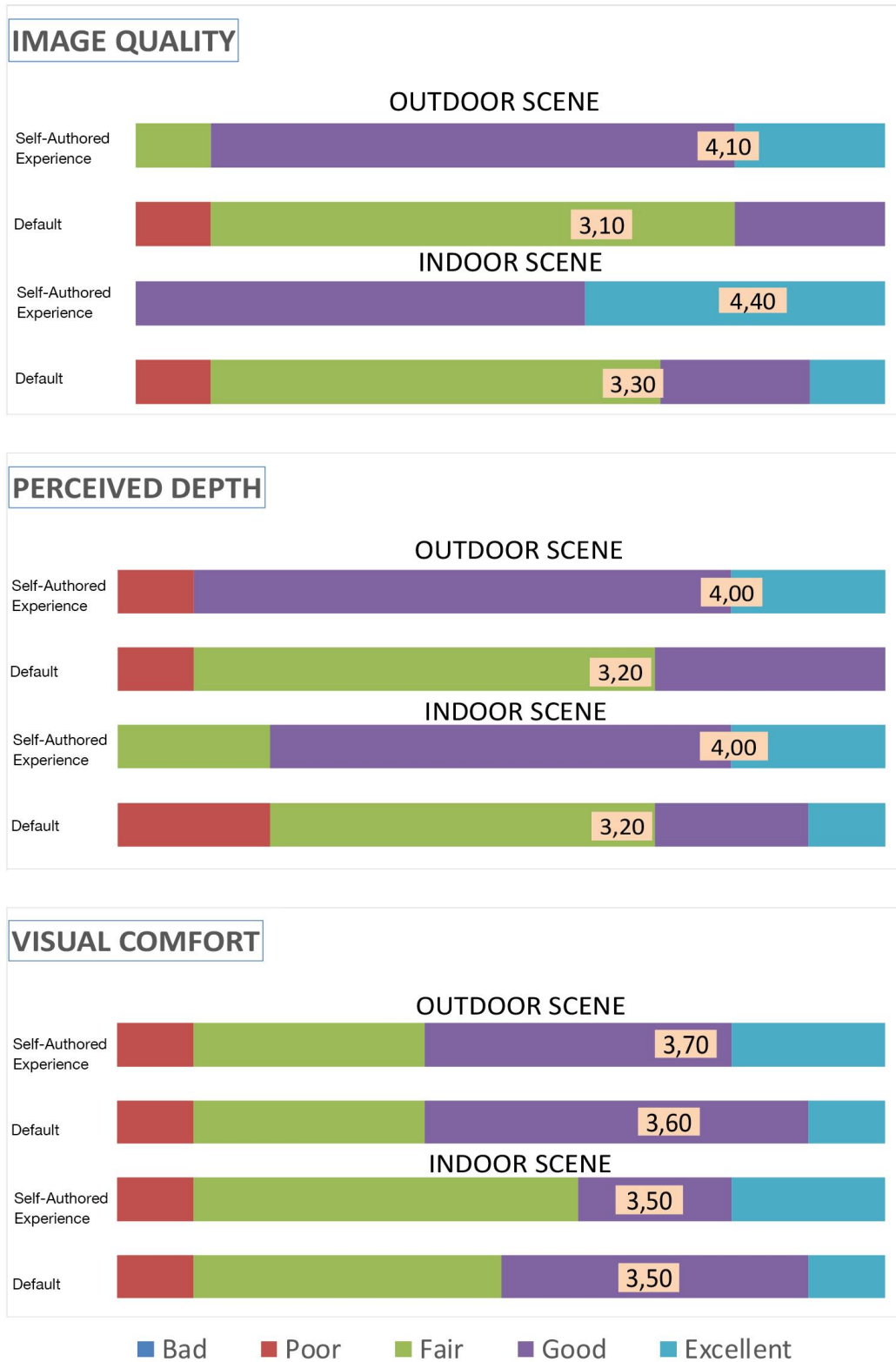


Figure 21 Individual session ratings collected with the questionnaires and their averages are given for each scene per configuration. The averages are indicated in rectangles.

Preset Experience with DOF Blur vs. Preset Experience

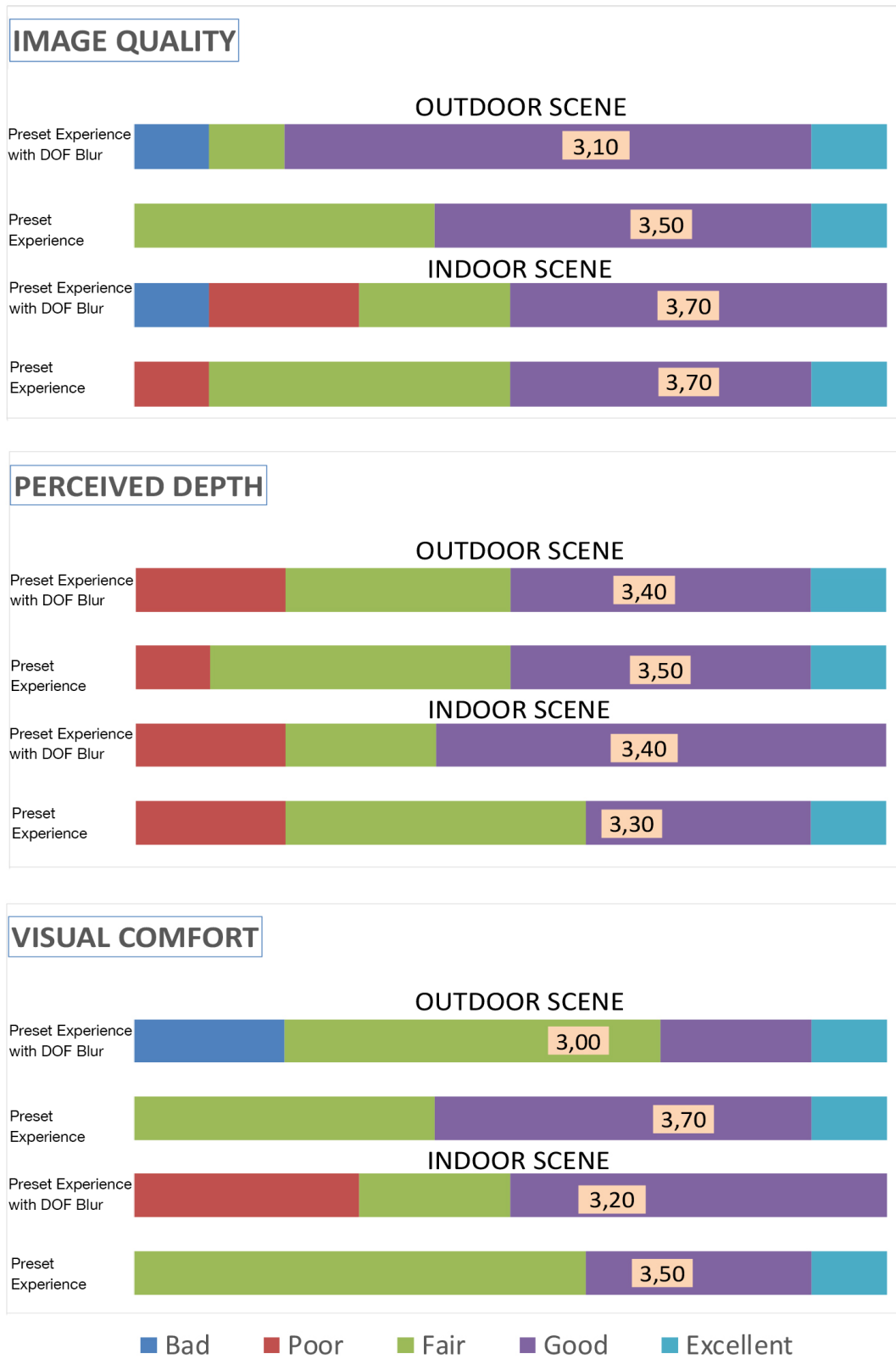


Figure 22 Individual session ratings collected with the questionnaires and their averages are given for each scene per configuration. The averages are indicated in rectangles.

For the first two configurations where our approach is evaluated against the Default Experience setting, it is seen that, in both Preset Experience and Self-Authored Experience settings, our approach was rated significantly higher on average in terms of image quality and perceived depth in both scenes. 9 subjects, for both settings and both scenes, indicated their preference of our approach to the Default Experience in terms of image quality. Similarly, in the outdoor scene, 7 subjects preferred our approach to the Default Experience in terms of perceived depth with both the Preset and the Self-Authored Experiences. For the indoor scene, 7 subjects favored the Preset Experience and 9 subjects favored the Self-Authored Experience in terms of perceived depth.

In terms of visual comfort, again for the first two configurations, our approach was rated only slightly better in the outdoor scene with both settings. In the indoor scene, however, it was rated slightly worse with the Preset Experience and the same with the Self-Authored Experience. It is seen that while the Default Experience was favored in terms of visual comfort to the Preset Experience by 2 and 3 of the subjects in the outdoor and indoor scenes, respectively, 4 and 5 of the subjects preferred it to their own Self-Authored Experiences in the outdoor and indoor scenes, respectively. The subject with VR application development background, on the other hand, was among the group who found their own Self-Authored Experience more visually comforting. These findings imply that novice VR users may still have a hard time when they first start authoring VR experiences for their ideal visual comfort level, however easy the interface is to grasp and use. Since our approach updates projection matrix, subjects' eye muscles contracts and relax a lot more than naive method which creates more depth in the scene while causes fatigue in some of the subjects because of the experience level of the users.

The questionnaire item that queries the overall preference of the subjects garnered responses demonstrating that our approach was considerably well-received with both the Preset Experience and the Self-Authored Experience settings. The breakdown of the results show that while the Self-Authored Experience was preferred by more subjects (9 in both scenes) than the Preset Experience (8 in both scenes), more subjects regarded the Preset Experience as "much better" with respect to the Default (5 in the outdoor scene and 6 in the indoor scene) than they did the Self-Authored Experience (3 in both scenes).

The results for the third configuration, which facilitates to evaluate the impact of our DoF blur implementation in combination with our approach, show that the combination was generally found to degrade both image quality and visual comfort. While, more subjects indicated their overall preference towards the session with DoF blur, the average ratings of the session without DoF blur are higher in all three evaluation criteria for the outdoor scene and similar for the indoor scene.

Table 1 and Table 2 present a sample of *depth charts* that contrast the depth distributions resulting from our approach with the results of the default arrangement in the two scenes. That is, the pair of charts for each scene is obtained by following the same camera path using the default settings once and once using our approach. On the charts, minimum and maximum depth values along the camera path are given with respect to the HMD display. It is seen that while the default arrangement constrains the zero-parallax plane to a fixed short distance to the stereo camera pair leading to a very narrow band of negative disparity region, our approach allows a dynamic yet smooth depth narrative in both negative and positive disparity regions.

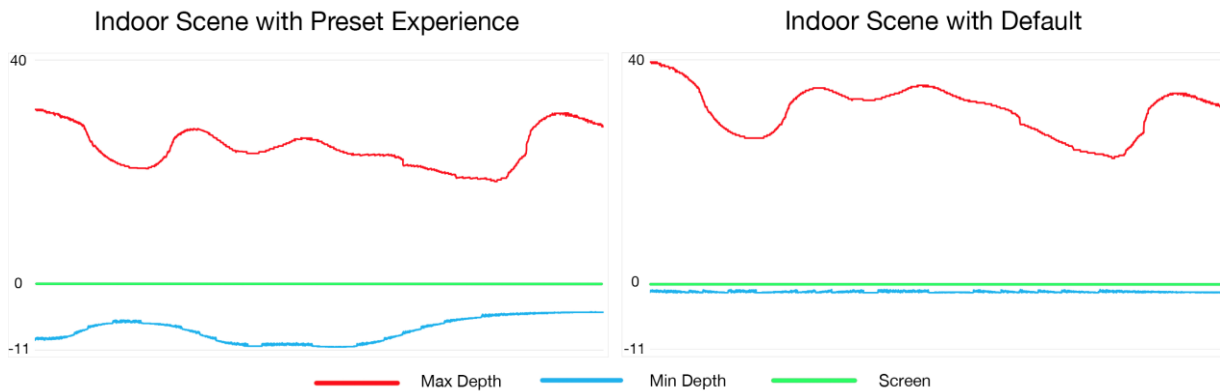


Table 1 Relative maximum and minimum depth value chart of the indoor scene

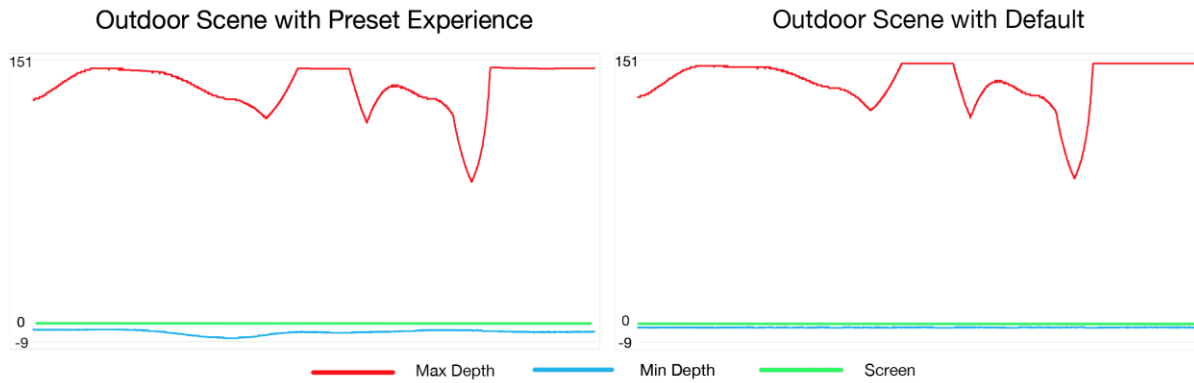


Table 2 Relative maximum and minimum depth value chart of the outdoor scene

Figure 23 Figure 24 Figure 25 Figure 26 Figure 27 shows results of the user preferences in percentages in the three configurations. For the Preset Experience 20% of the subjects in both indoor and outdoor scenes didn't preferred our method. For the Self-Authored Experience setting on the other hand, 88.89% of the subjects prefers edited parameters than the Default Experience setting.

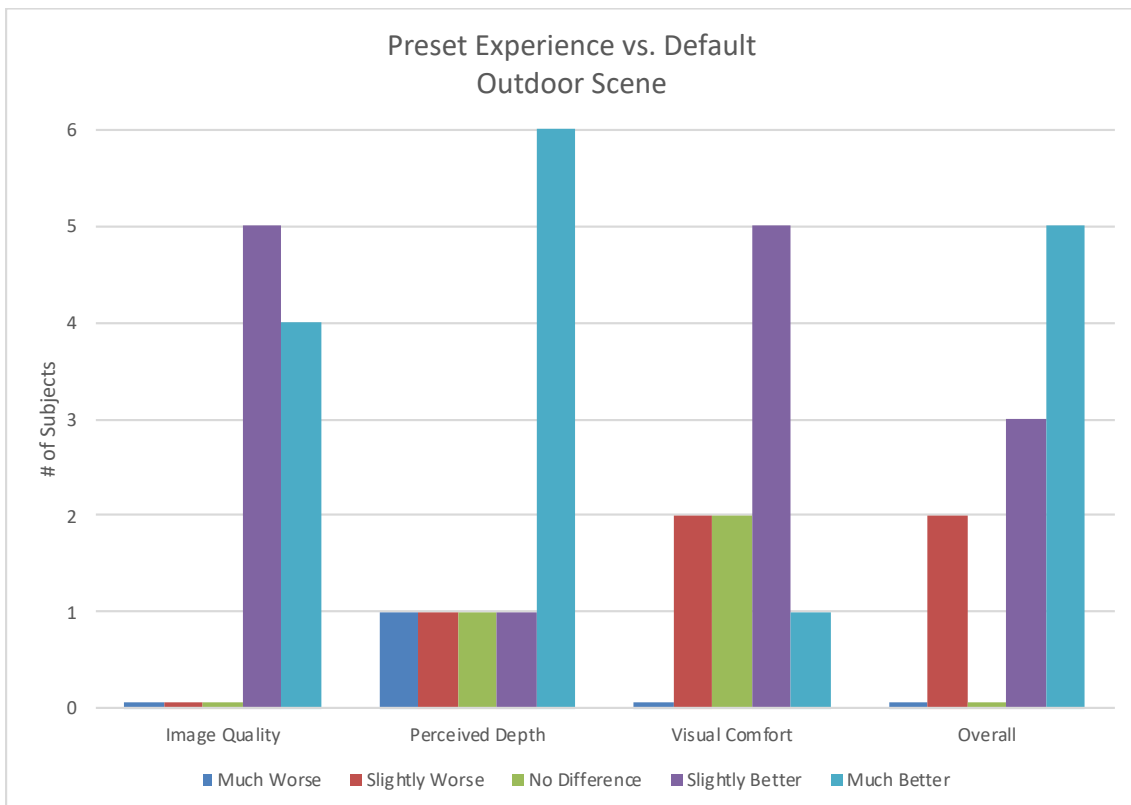


Figure 23 Preset Experience vs. Default session ratings collected with the questionnaires and its averages are given for Outdoor Scene.

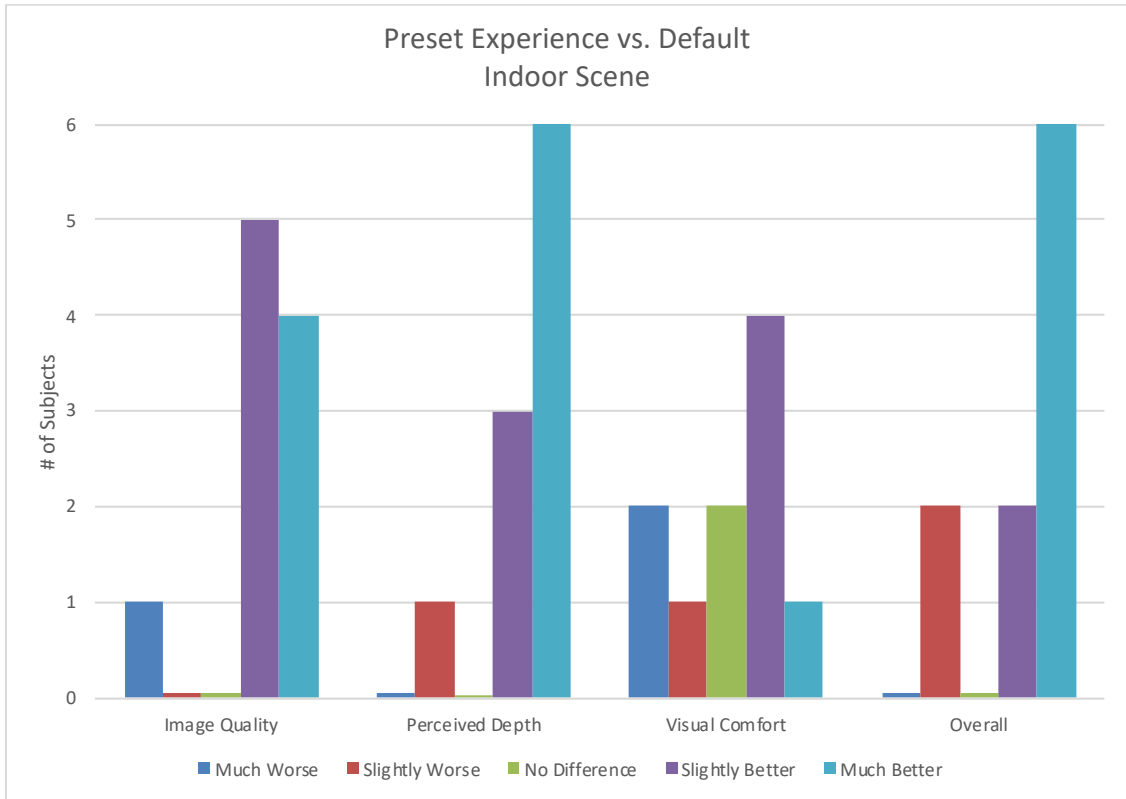


Figure 24 Preset Experience vs. Default session ratings collected with the questionnaires and its averages are given for Indoor Scene.

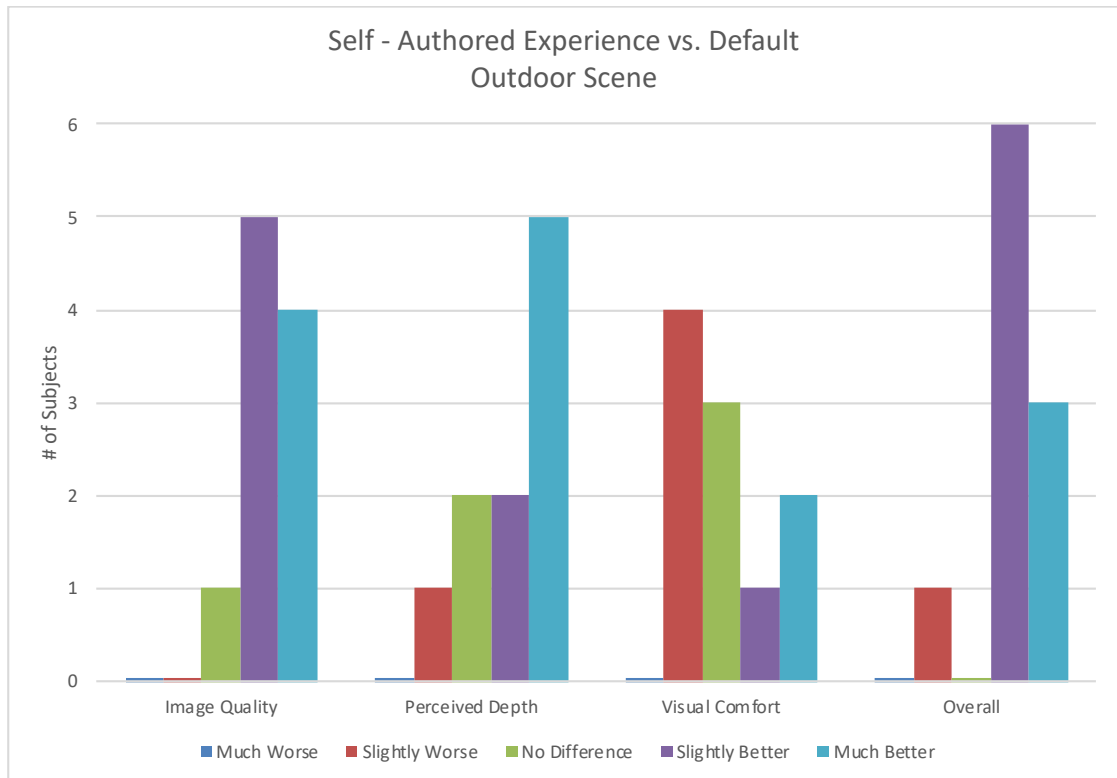


Figure 25 Self-Authored Experience vs. Default session ratings collected with the questionnaires and its averages are given for Outdoor Scene.

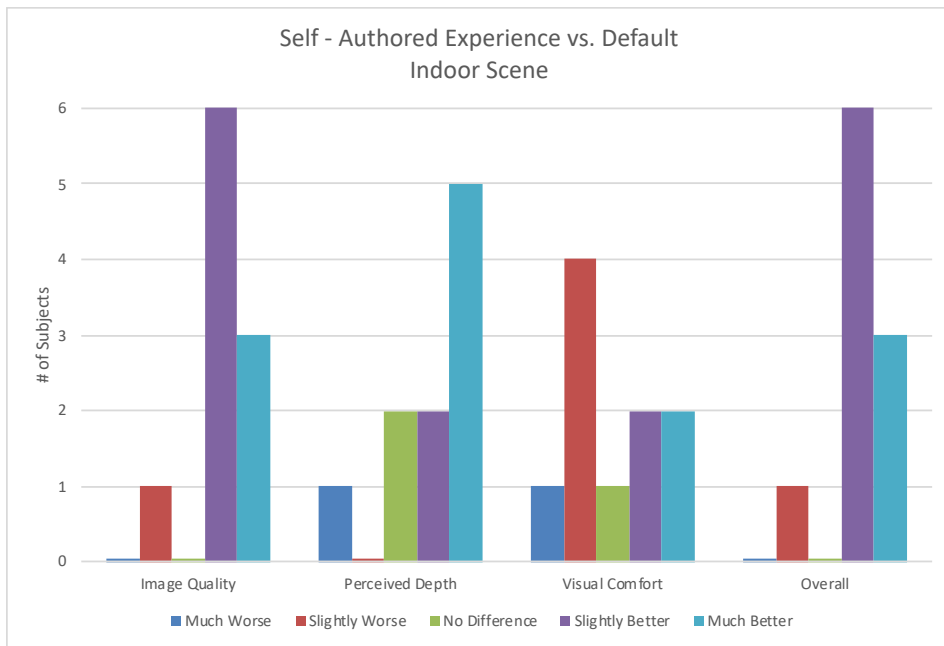


Figure 26 Self-Authored Experience vs. Default session ratings collected with the questionnaires and its averages are given for Indoor Scene.

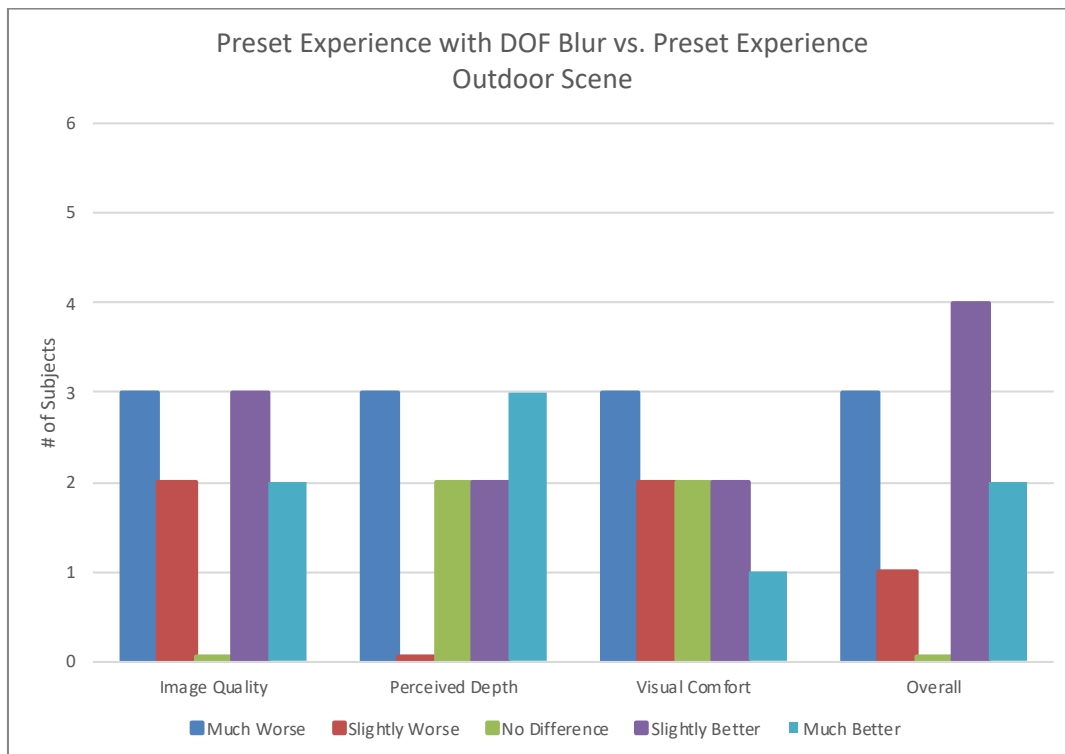


Figure 27 Preset Experience with DOF Blur vs. Preset Experience session ratings collected with the questionnaires and its averages are given for Outdoor Scene.

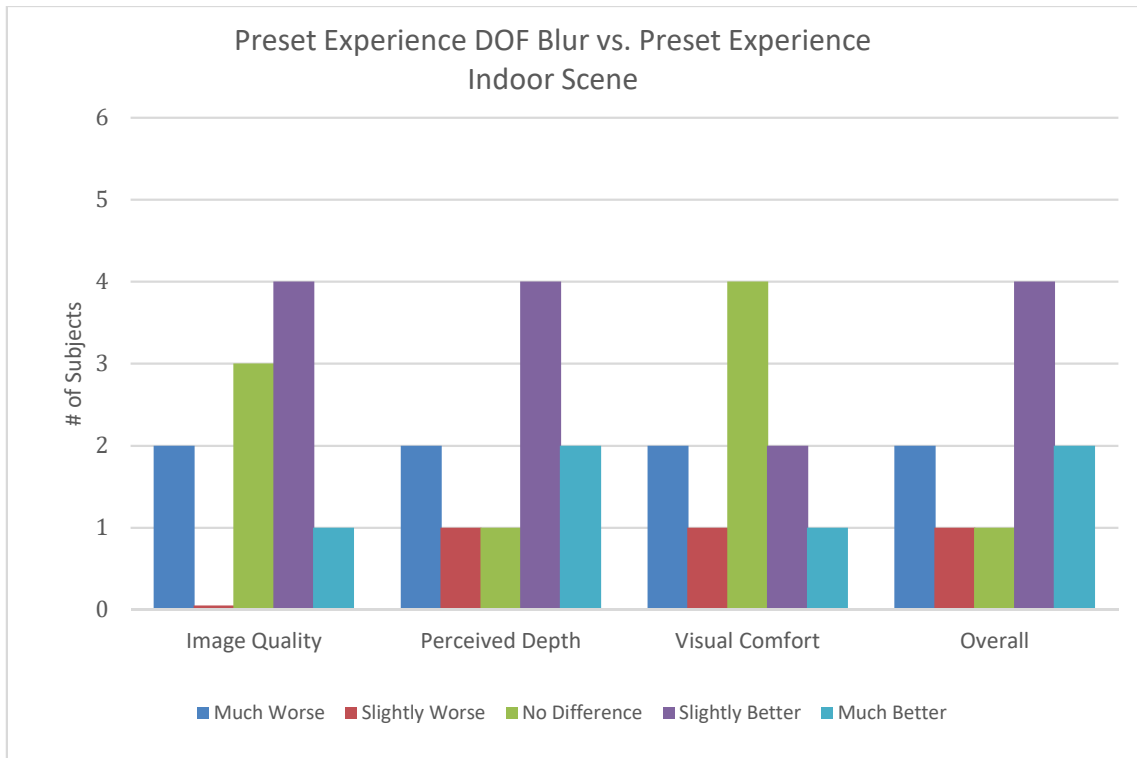


Figure 28 Preset Experience with DOF Blur vs. Preset Experience session ratings collected with the questionnaires and its averages are given for Indoor Scene.

One subject stated that scene contents looked smaller in our approach hence the slightly worse and much worse ratings in every setting in every scene. Since our approach updates inter-axial distance between cameras, the content in the scene can be perceived as smaller than they are if distance is bigger than the IPD and as bigger than they are if distance is smaller than the IPD [22].

6. CONCLUSION

In this work we have described an effective method of real-time automated stereoscopic camera control in VEs towards providing users with impressive VR experiences that are rich in depth. We have also introduced an immersive design interface for authoring unique VR experiences to be used with the proposed method. We believe both of these novel contributions that are ready to be used with existing consumer HMDs will stimulate new directions in stereoscopic rendering research. Moreover, we have addressed the issues with the stereoscopic parameter extrema that may cause undesired effects and proposed a dynamic limiting scheme.

The results of the user evaluation study demonstrate that our method is able to enhance user experience, as intended, especially in terms of overall perceived depth and image quality. The method is also seen to slightly improve the visual comfort in the scene with wider depth range and to keep it at similar levels in the scene with narrower depth range. The use of DoF blur added to our proposed method did not help to improve the experience further, since the majority of the user assessments were in accord with some earlier studies [5, 9, 10]. Surely, a further standalone study with higher number of configurations and larger sample size in which users can be grouped as first time VR users, VR enthusiasts and VR developers would be beneficial for a more thorough analysis.

According to the outcomes of our tests, we believe that, our proposed method is a suitable solution for increasing depth perception in pre defined environments. While it is observed that using our approach in a non interactive setting where the user experiences the scene with a defined camera path set in the selected depth layout lead to acceptable results, we believe further improvements to the method are crucial for proper adaptation to use in interactive VR applications.

7. REFERENCES

- [1] About the VIVE controllers, (n.d.).
https://www.vive.com/nz/support/vive/category_howto/about-the-controllers.html (accessed August 30, 2019).
- [2] Radial Basis Function Interpolation in 2D, (n.d.).
http://people.sc.fsu.edu/~jburkardt/cpp_src/rbf_interp_2d/rbf_interp_2d.html (accessed January 14, 2015).
- [3] OpenGL Projection Matrix, (2013) 1–8.
http://www.songho.ca/opengl/gl_projectionmatrix.html (accessed August 30, 2019).
- [4] R.K. Andrew J. Woods Tom Docherty, Image distortions in stereoscopic video systems, 1915 (1993). doi:10.1117/12.157041.
- [5] J.P. Brooker, P.M. Sharkey, Operator performance evaluation of controlled depth of field in a stereographically displayed virtual environment, in: *Stereosc. Displays Virtual Real. Syst. VIII*, 2001: pp. 408–417.
- [6] K. Carnegie, T. Rhee, Reducing Visual Discomfort with HMDs Using Dynamic Depth of Field, *Virtual Real. Softw. Technol.* September- (2015) 34–41.
- [7] J.C. Carr, R.K. Beatson, J.B. Cherrie, T.J. Mitchell, W.R. Fright, B.C. McCallum, T.R. Evans, Reconstruction and representation of 3D objects with radial basis functions, in: *Proc. 28th Annu. Conf. Comput. Graph. Interact. Tech.*, 2001: pp. 67–76.
- [8] U. Celikkan, G. Cimen, E.B. Kevinc, T. Capin, Attention-aware disparity control in interactive environments, *Vis. Comput.* 29 (2013) 685–694.
- [9] J. Conti, B. Ozell, E. Paquette, P. Renaud, Adjusting stereoscopic parameters by evaluating the point of regard in a virtual environment, *Comput. Graph.* 69 (2017) 24–35.
- [10] A.T. Duchowski, D.H. House, J. Gestring, R.I. Wang, K. Krejtz, I. Krejtz, R. Mantiuk, B. Bazyluk, Reducing Visual Discomfort of 3D Stereoscopic Displays with Gaze-contingent Depth-of-field, in: *Proc. ACM Symp. Appl. Percept.*, ACM, New York, NY, USA, 2014: pp. 39–46. doi:10.1145/2628257.2628259.
- [11] R. Franke, Scattered data interpolation: Tests of some methods, *Math. Comput.* 38 (1982) 181–200.
- [12] B. Gerschütz, M. Fechter, B. Schleich, S. Wartzack, A Review of Requirements and Approaches for Realistic Visual Perception in Virtual Reality, in: *Proc. Des. Soc. Int. Conf. Eng. Des.*, 2019: pp. 1893–1902.

- [13] S.-H. Guan, Y.-C. Lai, K.-W. Chen, H.-T. Chou, Y.-Y. Chuang, A Tool for Stereoscopic Parameter Setting Based on Geometric Perceived Depth Percentage, *IEEE Trans. Circuits Syst. Video Technol.* 26 (2016) 290–303.
- [14] J.-P. Guertin, Derek, High performance Non-Linear motion blur, (2015).
- [15] P. Hanhart, T. Ebrahimi, Subjective evaluation of two stereoscopic imaging systems exploiting visual attention to improve 3D quality of experience, in: *Stereosc. Displays Appl.* XXV, 2014: p. 90110D.
- [16] G.R. Jones, D. Lee, N.S. Holliman, D. Ezra, Controlling perceived depth in stereoscopic images., in: 2001.
- [17] P. Kellnhofer, P. Didyk, K. Myszkowski, M.M. Hefeeda, H.-P. Seidel, W. Matusik, GazeStereo3D: Seamless disparity manipulations, *ACM Trans. Graph.* 35 (2016) 68.
- [18] G.A. Koulteris, G. Drettakis, D. Cunningham, K. Mania, Gaze prediction using machine learning for dynamic stereo manipulation in games, in: *2016 IEEE Virtual Real.*, 2016: pp. 113–120.
- [19] G.A. Koulteris, G. Drettakis, D. Cunningham, K. Mania, Supplemental material: Gaze Prediction using Machine Learning for Dynamic Stereo Manipulation in Games, (n.d.).
- [20] G.-A. Koulteris, B. Bui, M. Banks, G. Drettakis, Accommodation and Comfort in Head-Mounted Displays, *ACM Trans. Graph.* 36 (2017) 1–11.
- [21] G. Kramida, Resolving the vergence-accommodation conflict in head-mounted displays, *IEEE Trans. Vis. Comput. Graph.* 22 (2016) 1912–1931.
- [22] V.R. LA, VRLA Summer Expo 2016: How VR Works A Perceptual Point of View, (2016).
<https://www.youtube.com/watch?v=iJ0TV2jgNoc&feature=youtu.be&t=14m20s> (accessed August 30, 2019).
- [23] M. Lambooi, M. Fortuin, I. Heynderickx, W. IJsselsteijn, Visual discomfort and visual fatigue of stereoscopic displays: A review, *J. Imaging Sci. Technol.* 53 (2009) 30201.
- [24] M. Lang, A. Hornung, O. Wang, S. Poulakos, A. Smolic, M. Gross, Nonlinear disparity mapping for stereoscopic 3D, *ACM Trans. Graph.* 29 (2010) 75:1--75:10.
- [25] E. Langbehn, T. Raupp, G. Bruder, F. Steinicke, B. Bolte, M. Lappe, Visual blur in immersive virtual environments: does depth of field or motion blur affect distance and speed estimation?, in: *Proc. 22nd ACM Conf. Virtual Real. Softw. Technol.*, 2016: pp. 241–250.

- [26] P. Milgram, M. Krüger, Adaptation effects in stereo due to on-line changes in camera configuration, in: *Stereosc. Displays Appl. III*, 1992: pp. 122–134.
- [27] G. Moreau, Visual Immersion Issues in Virtual Reality: A Survey, in: *2013 26th Conf. Graph. Patterns Images Tutorials*, 2013: pp. 6–14. doi:10.1109/SIBGRAPIT.2013.9.
- [28] J. Noh, D. Fidaleo, U. Neumann, Animated deformations with radial basis functions, in: *Proc. ACM Symp. Virtual Real. Softw. Technol.*, 2000: pp. 166–174.
- [29] L. O’Hare, T. Zhang, H.T. Nefs, P.B. Hibbard, Visual discomfort and depth-of-field, *Iperception*. 4 (2013) 156–169.
- [30] H. Oh, J. Kim, J. Kim, T. Kim, S. Lee, A.C. Bovik, Enhancement of visual comfort and sense of presence on stereoscopic 3D images, *IEEE Trans. Image Process*. 26 (2017) 3789–3801.
- [31] T. Oskam, A. Hornung, H. Bowles, K. Mitchell, M. Gross, OSCAM-optimized stereoscopic camera control for interactive 3D, in: *SA’11 Proc. 2011 SIGGRAPH Asia Conf.*, 2011: p. 189.
- [32] N. Padmanaban, R. Konrad, T. Stramer, E.A. Cooper, G. Wetzstein, Optimizing Virtual Reality for All Users through Gaze-Contingent and Adaptive Focus Displays, *Proc. Natl. Acad. Sci.* 114 (2017) 2183–2188.
- [33] A.S. Percival, The relation of convergence to accommodation and its practical bearing, *Ophthal. Rev.* 11 (1892) 313–328.
- [34] M.J.D. Powell, The theory of radial basis function approximation in 1990, *Adv. Numer. Anal.* 2 (1992) 105–210.
- [35] M.J.D. Powell, Radial basis functions for multivariable interpolation: a review, in: *Algorithms Approx.*, 1987: pp. 143–167.
- [36] S.K. Rushton, P.M. Riddell, Developing visual systems and exposure to virtual reality and stereo displays: some concerns and speculations about the demands on accommodation and vergence, *Appl. Ergon.* 30 (1999) 69–78. doi:[https://doi.org/10.1016/S0003-6870\(98\)00044-1](https://doi.org/10.1016/S0003-6870(98)00044-1).
- [37] R. Salomon, J. Weissmann, Using Radial Basis Function Networks for Hand Gesture Recognition, in: R.J. Howlett, L.C. Jain (Eds.), *Radial Basis Funct. Networks 2 New Adv. Des.*, Physica-Verlag HD, Heidelberg, 2001: pp. 37–58. doi:10.1007/978-3-7908-1826-0_2.
- [38] A. Sherstyuk, A. State, Dynamic eye convergence for head-mounted displays, (2010) 43–46.
- [39] R. Summer, *The Impact of Dynamic Convergence on the Human Visual System in Head Mounted Displays.*, Victoria University of Wellington, 2017.

- [40] I.T. Union, Subjective methods for the assessment of stereoscopic 3DTV systems, *Recomm. ITU-R BT. 2021* (2015).
- [41] M. Vinnikov, R.S. Allison, Gaze-contingent Depth of Field in Realistic Scenes: The User Experience, in: *Proc. Symp. Eye Track. Res. Appl.*, ACM, New York, NY, USA, 2014: pp. 119–126. doi:10.1145/2578153.2578170.
- [42] M. Wang, X.-J. Zhang, J.-B. Liang, S.-H. Zhang, R.R. Martin, Comfort-driven disparity adjustment for stereoscopic video, *Comput. Vis. Media*. 2 (2016) 3–17.
- [43] F. Zilly, M. Muller, P. Kauff, R. Schafer, STAN - An assistance system for 3D productions: From bad stereo to good stereo, in: *Electron. Media Technol. (CEMT)*, 2011 14th ITG Conf., 2011: pp. 1–6.
- [44] F.S.R.R.T.M. Lew B. Stelmach Wa James Tam, Improving the visual comfort of stereoscopic images, 5006 (2003). doi:10.1117/12.474093.
- [45] G. Kramida, Resolving the vergence-accommodation conflict in head-mounted displays, *IEEE Trans. Vis. Comput. Graph.* 22 (2016) 1912–1931.
- [46] Tech Note: Asymmetric Field of View FAQ, (n.d.). <https://developer.oculus.com/blog/tech-note-asymmetric-field-of-view-faq/> (accessed September 7, 2019).
- [47] J. Lei, S. Li, B. Wang, K. Fan, C. Hou, Stereoscopic Visual Attention Guided Disparity Control for Multiview Images, *J. Disp. Technol.* 10 (2014) 373–379. doi:10.1109/JDT.2014.2312648.
- [48] VR Expert to Oculus Rift Devs: Make Sure You’re Doing 3D Right, (n.d.). <https://www.roadtovr.com/vr-expert-to-oculus-rift-devs-make-sure-youre-doing-3d-right/> (accessed August 30, 2019).

EK 1 – Questionnaire

Tips for Viewing 3-D Stereograms

1. Pick a spot on the picture (the middle seems to work best) and just stare at it.
2. Allow your eyes to relax, don't just stare AT the image, try to stare THROUGH it. You'll notice your eyes will go slightly out of focus. This is normal.
3. Keep staring, don't give up, once you begin to see the first image, it gets much easier.

Our Tips

1- We ask you to judge the quality and comfort and sense of depth

2- For quality:

- Stereo image introduce ghosting effects, and other aspects like sharpness
- consider these when voting

3- For Depth:

- Stereo images offer an enhanced sense of depth as compared to 2D.

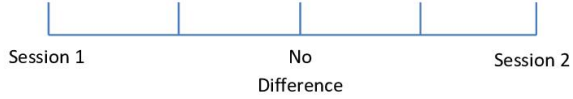









-4- For Comfort:

- sensation of (dis)comfort that can be associated with the viewing of stereoscopic images
- improperly displayed stereoscopic images could be a serious source of discomfort.

User :

Date:

Time:

TEST # :	
Which session provided better picture quality?	
Grade picture quality of Session 1	
Grade picture quality of Session 2	
Which session offered more depth?	
Grade depth of Session 1	
Grade depth of Session 2	
Which session was more comfortable to watch?	
Grade comfort of Session 1	
Grade comfort of Session 2	
Which session did you prefer?	

EK 2 – Academic Research Papers based on this work

We have submit a journal paper with the name of “*Enhancing User Experience in Virtual Reality with Radial Basis Function Interpolation Based Stereoscopic Camera Control*” to *IEEE VR 2020* and it is being reviewed at the moment this thesis is written.