A study on presence quality and cybersickness in 2D, smartphone, and VR

Saleh Saeed[1], Unsang Park[1*]
[1] Department of Computer Science and Engineering, Sogang University
Seoul, 04107 South Korea
[e-mail: saleh@sogang.ac.kr, unsangpark@sogang.ac.kr]
[*]Corresponding author: Unsang Park

Received February 9, 2021; revised December 19, 2021; revised March 28, 2022; revised May 7, 2022; accepted June 1, 2022; published July 31, 2022

Abstract

Recent improvements in technology have increased the consumption of virtual reality (VR) contents on immersive displays. The VR experience depends on the type of displays as well as the quality of VR contents. However, research on the impacts of VR content quality on VR experience and comparisons among different types of immersive display devices are lacking. In this study, VR contents created with our VR framework, are provided to participants on conventional two-dimensional (2D) immersive displays and VR headset. The geometric alignment of VR contents is improved with the addition of two calibration modes (i.e. preprocessing and straightening). The subjective feelings of presence and cybersickness experienced by participants while consuming VR contents created by our framework and commercial solutions are recorded in the form of questionnaires. The results of this study indicate that the improvements in VR quality lead to a better presence and less cybersickness in both conventional 2D displays and VR headset. Furthermore, the level of presence and cybersickness increases in VR headsets as compared to conventional 2D displays. Finally, the VR content quality improvements lead to a better VR experience for our VR framework as compared to commercial solutions.

Keywords: Cybersickness, Immersion, Perception, Presence, Virtual Reality.

This research was partly supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2020R1F1A1072332), and Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.2022-0-00377, Development of Intelligent Analysis and Classification Based Contents Class Categorization Technique to Prevent Imprudent Harmful Media Distribution).

http://doi.org/10.3837/tiis.2022.07.010 **ISSN : 1976-7277**

1. Introduction

The methods for the creation and consumption of multimedia have been constantly evolving over the last few decades. Primitive methods for capturing visual imagery involve a basic camera, with a standard lens, based on the pinhole camera model. Such cameras produce images and videos with low resolutions, low frame rates, and small field-of-view (FoV). With advancements in hardware and software techniques, the resolution and frame rates of the captured multimedia contents have been improved. The use of a wide-angle lens, for example, a fisheye lens, paved the way for higher quality and large FoV images and videos. The high-quality images and videos cannot be rendered on a primitive display device which can only display low-quality content. Recent advancements in display technologies enabled the presentation of high-quality content in an immersive environment. In an immersive environment, multimedia contents with up to 4K resolutions, covering hemispherical to full spherical FoV, and 60 frames per second can be presented in such a way that users can interact with the environment and feel a presence. Experiencing this type of content is called Virtual Reality (VR).

1.1 Virtual Reality

VR is defined as the interactive experience of a user with a 3D world in a simulated environment such that the user feels a presence in the 3D world. The 3D world can be created by transforming the real-world scene, generating an artificial scene, or a mixture of both real and artificial scenes. Recent advances in technology have led to the increased usage of VR in different applications such as gaming, multimedia, medical [1,2,3], or simulated training [4], etc. The use of VR depends on the quality of VR contents, immersive display devices, and user comfort. VR contents are usually captured by a single camera with an omnidirectional ultrawide angle lens or a group of pre-calibrated cameras mounted on a rig. In the case of single cameras with an omnidirectional lens, the extreme radial distortions in the captured frames degrade the VR experience. In a multicamera pre-calibrated rig, each camera has a wide-angle fisheye lens that can capture the hemispherical scene. The two hemispherical scenes can be viewed independently or can be stitched to form a single full spherical video stream using software-based or hardware-based geometric transformations. The stitching process may add geometric or photometric inconsistencies along the seam of the stitching boundary. The inconsistencies are particularly more prominent when the object-in-focus is spatially close to the camera rig. A recent increase in the commercial availability of single and multiple camera devices has caused an increase in VR multimedia consumption in the form of 3D cinema or YouTube VR. The VR contents can be consumed on two types of devices i.e. 2D displays and head-mounted displays (HMDs). The 2D displays include wide/ultrawide monitors, smartphones, and multi-monitor/multi-projector setups. The VR contents are presented on wide/ultrawide 2D monitors and VR software is used for interaction. A VR app can be used on a smartphone along with a GPS sensor to consume VR contents. In the case of multi-monitor/multi-projector setups, VR contents are presented after calibrating the display devices. Multi-monitor and multi-projector setups are typically used for training purposes. These setups are complex and expensive to use. Therefore, the use of VR in commercial multimedia applications is mostly limited to using a wide/ultra display or a smartphone. In the last decade, many high-quality HMDs [5] have been developed which range from simple Google Cardboard based headsets [6] to more advanced devices such as Oculus Rift, HTC Vive, and Sony PlayStation VR.

1.2 Presence quality and cybersickness in Virtual Reality

The technological developments and increased usage of VR have led to the study of factors involving a comfortable VR experience. In a typical VR environment, users may experience increased discomfort in the form of nausea, blurriness, disorientation, etc. The discomfort experienced in a VR environment is called cybersickness. The cybersickness experienced by users may lead to reduced enjoyability and performance in VR. The reduction of enjoyability and performance can be described in terms of the subjective feeling of presence. Presence is defined as the deceptive feeling of being located at a certain place while interacting with a simulated environment. The complex relation between presence and cybersickness is difficult to understand. There are various studies [7-17] that model the relation between the presence and cybersickness in a VR environment [7]. In general, the presence and cybersickness have an inverse relationship. The VR environments are generally designed so that the presence experienced by users is increased while keeping the cybersickness at lower levels. The presence and cybersickness are dependent on many factors. The cybersickness levels can be related to the low, semi and fully immersive levels in a VR environment. A study [8] on cybersickness at the three immersive levels shows that the cybersickness is higher in more immersive environments. In another study [9], the relation between cybersickness and repeated exposure to the VR environment may end up in the reduction of cybersickness for certain VR contents. In [10], the cybersickness and presence experienced in a VR environment are found to have a temporal relationship. In [11], the presence and cybersickness are shown to be related to the dynamic FoV, where the restriction of dynamic FoV can reduce the cybersickness and improve the presence in a VR environment. In [13], a study on the relationship between presence and cybersickness experienced by the user's in a virtual environment and the corresponding real environments show that the users perceive a positive correlation for both presence and cybersickness between the two environments. Subjective assessment of presence and cybersickness can be carried out by using the presence and sickness questionnaire [18,19].

In previous studies [20-27] the user's subjective behavior in an immersive environment is extensively studied in terms of the quality-of-experience (QoE) (i.e. presence or cybersickness) with respect to different quality metrics. The quality of VR contents is a broad term that may relate to the specifications of a video stream (i.e. the resolution, framerate, FoV, etc.), the type of camera movements (i.e. fixed, horizontal, or vertical movement), the number of moving objects in foreground or background, etc. The previous studies [23, 24, 26] evaluate the VR experience by improving these qualities and subjectively measuring the feelings of immersion and cybersickness. However, these studies are lacking in terms of covering different quality improvement metrics. Furthermore, these studies do not provide analysis of QoE in relation to the geometric and photometric quality improvements at hardware level. In this paper, we provide a novel study on the QoE in relation to the geometric and photometric quality improvements in an end-to-end VR system. In order to provide a comparison between cybersickness and presence concerning the quality of VR contents, we provide immersive multimedia contents on three different levels of immersion i.e. an immersive 2D screen, a VR capable smartphone, and a VR headset. The VR contents are created by stitching video streams [28,29,30] from multiple cameras attached to a rig. We use two different methods to improve the geometric alignment [31,32] of stitched video streams i.e. preprocessing [33] and straightening. In preprocessing, we mitigate the misalignment by removing redundant and outlying scale-invariant features [34,35,36,37]. In straightening, we use line features to straighten out VR streams perpendicular to the camera's z-axis in such a way that the wavy effects are removed. The VR contents created by preprocessing and straightening are provided

to participants and their cybersickness and presence response is compared using the presence and sickness questionnaire [18,19]. Participants are also provided with VR contents from two commercial solutions, i.e. PTGui [38], and Hugin [39]. The comparison between cybersickness and presence in relation to PTGui and Hugin's VR contents are also measured. Finally, we relate the three immersion levels with the quality of VR contents in comparison to presence and cybersickness. In this work, our main contributions are improvements in VR experience by improving the quality of VR contents, providing a better experience as compared to commercial solutions, and analysis of VR quality impacts on different kinds of immersive displays.

2. Proposed Method

In this section, firstly, we will discuss our VR framework [40] for capturing and rendering high-quality contents. Secondly, we will illustrate the experimental setup for processing presence quality and cybersickness feedback from participants.

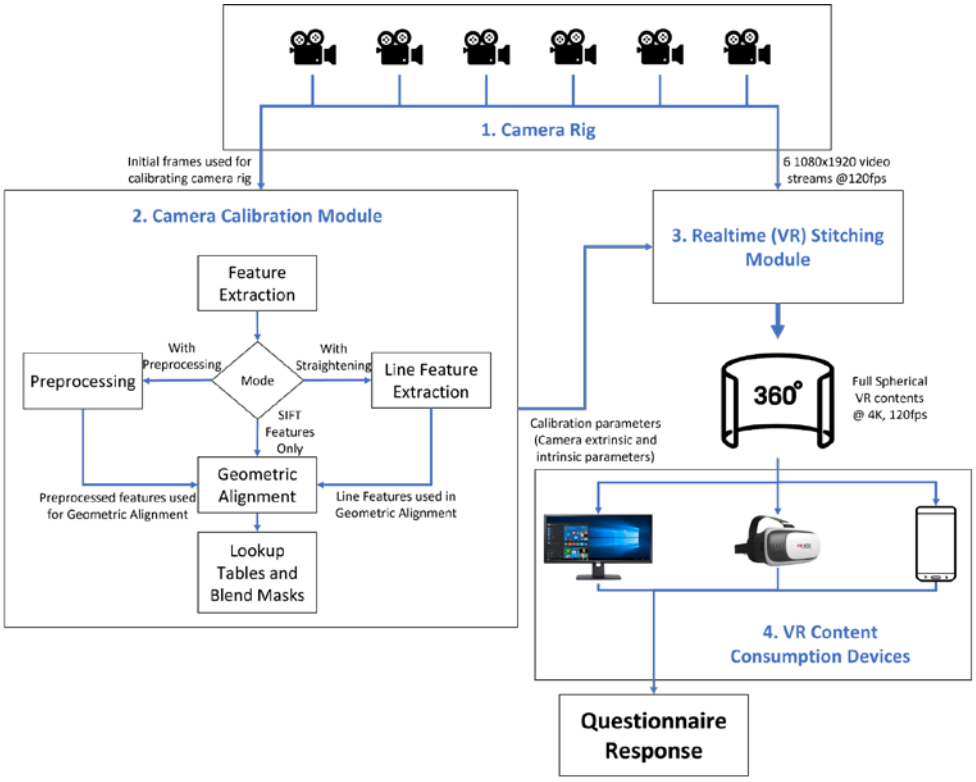


Fig. 1. An end-to-end VR content creation and consumption framework to analyze presence quality and cybersickness.

2.1 VR Framework

The experimental setup uses a VR framework [40] with four main components that are based on traditional panorama stitching and rendering techniques [41-49]. The first component is the camera rig, the second component is the camera calibration module, the third component is the real-time stitching module and the fourth one consists of VR content consumption devices.

The output of the fourth component is used to assess the participant's questionnaire response in terms of presence and cybersickness ratings. The working flow of the VR framework is shown in **Fig. 1** and described in detail in this section.

The camera rig consists of six full HD cameras positioned in portrait mode. Each camera is equipped with a fisheye lens that can capture video streams with horizontal FoV of ~75 degrees and vertical FoV of ~180 degrees. There is approximately a 25% overlap between adjacent cameras. The cameras are mounted on a rig with the same rotational axis. It is assumed that there is no translation between the centers of the cameras and the scene geometry, as captured from these cameras, follows a homographic relationship. Furthermore, each camera has a frame rate of up to 120 fps and the frames are assumed to be synchronized. The group of six cameras covers a full spherical view. The camera rig is attached to a processing unit that can estimate the calibration parameters for each camera. The frames from each camera are shown in **Fig. 2**.

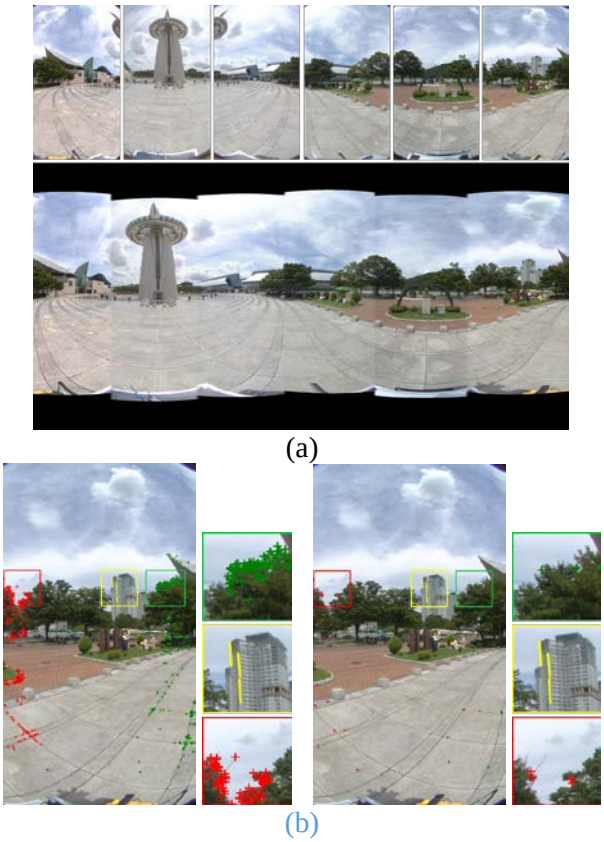


Fig. 2. (a) VR content (bottom) stitched with six camera frames (top). (b) SIFT (red + and green +) and line (yellow lines) features, where (left) shows features without preprocessing and (right) shows features with preprocessing. Note the removal of erroneous features from leaves (in local areas) which may cause misalignments. The features in left and right overlap areas (with neighboring camera images) are represented by red + and green + signs.

The second component of the VR framework (i.e. the camera calibration module) is used to estimate the camera calibration parameters. There are two types of calibration parameters i.e. camera intrinsic parameters and camera extrinsic parameters. The intrinsic parameters consist of focal length, principal point, and scaling size. The intrinsic parameters are assumed to be

the same for each camera. The extrinsic parameters consist of rotational values along three dimensions. The extrinsic parameters are different for each camera. The intrinsic and extrinsic parameters are denoted by β. Initial frames from the camera rig are used in the calibration process. The calibration process starts with the transformation of fisheye frames into equirectangular frames. Scale-invariant features (for example SIFT), as shown in **Fig. 2b** left image, are extracted from each equirectangular frame and matched with features from other frames. The features are matched using the K-nearest-neighbor algorithm and geometrically verified with RANSAC. The geometrically verified feature matches are used to create a graph tree of overlapping cameras. The graph tree consists of nodes and edges. Each node represents a camera and the nodes with overlapping cameras are connected with edges. Each edge is assigned a weight equal to the number of matched features between the two nodes. The node and edge connections in the graph tree are used in a Levenberg Marquardt based bundle adjustment algorithm. The bundle adjustment algorithm estimates the camera calibration parameters by mapping each pair of matched features in a sphere such that the spherical distance between matching features is minimized. For a pair of overlapping camera images a & b, the l^{th} matching 2D features $x_{l,a}$ and $x_{l,b}$ are mapped in a sphere in 3D as $X_{l,a}$ and $X_{l,b}$, where $l = 1,..L_f$ and L_f is the total number of matched features between the overlapping camera images. The spherical distance e_f between $X_{l,a}$ and $X_{l,b}$ is minimized using (1),

$$\boldsymbol{e_f} = argmin_{\beta}\left(\frac{1}{L_f}\sum_{l=1}^{L_f} \sin^{-1}\frac{\|X_{l,a}\times X_{l,b}\|}{\|X_{l,a}\|.\|X_{l,b}\|}\right) \quad (1)$$

After estimation of camera calibration parameters, lookup tables (LUTs) are created that contain backward projection information, from VR panorama domain to camera domain in the form of quadrilaterals. For each corner of quadrilateral in the VR panorama domain, there is a corresponding corner of quadrilateral in the camera domain. Each camera has a separate LUT. The LUT transformation leaves abrupt photometric inconsistencies along the stitching boundary. The stitching boundary transition is made seamless with the application of blending masks.

The third component of the VR framework (i.e. real-time stitching module) creates a real-time VR stream by stitching frames from the camera rig using the LUTs and blend masks. The LUTs and blend masks, obtained from the camera calibration module, contain direct transformation information between the camera domain and VR panorama domain. We use OpenGL APIs for texture mapping. The texture mapping process can render camera contents on VR panorama canvas in real-time. The stitched VR contents can be saved into a file, visualized with VR display devices, or transmitted over a network.

The fourth component of the VR framework (i.e. VR content consumption devices) consists of three types of immersive displays. The first display device that we have used in our experiments is a panoramic ultrawidescreen monitor. The monitor (i.e. Dell Ultrasharp U2913WM) is ideal for immersive and panoramic multimedia visualizations. The display has a size 73.7cm, aspect ratio 21:9, resolution 2560x1080. Its IPS panel has a viewing angle of 178°/178° and a refresh rate of 60Hz. The second display device is a VR Box headset. The VR Box headset is a low cost VR head mounted display that is designed on the principles of Google Cardboard VR headset [6]. The VR Box headset uses smartphone for displaying VR contents. The headset is ideal for visualizing full spherical VR contents. The third display device is a smartphone (i.e. Samsung Galaxy Note 4) that is bundled with a VR visualization application and an AGPS sensor. The smartphone has a super AMOLED display with 16M colors. Its size is 14.48cm and resolution is 2560x1440 pixels. We use the same smartphone device with the VR headset for visualization.

The second component of the VR framework is directly responsible for improving the quality of stitched VR contents. We use two methods to improve VR content quality. The two methods can be used in conjunction with the camera calibration module. The two methods are preprocessing and straightening.

(a) (b)

Fig. 3. Contents stitched with VR framework using: (a) scale-invariant features (without preprocessing of features); (b) preprocessed features. Notice the local regions showing improvements in geometric alignment for contents stitched with preprocessed features.

(a) (b)

Fig. 4. Contents stitched with VR framework using: (a) scale-invariant features only (without line features); (b) line and scale invariant features. Notice the wavy effect (shown with red wavy lines in (a)) is removed in (b) with the use of line and scale-invariant features.

2.1.1 Preprocessing

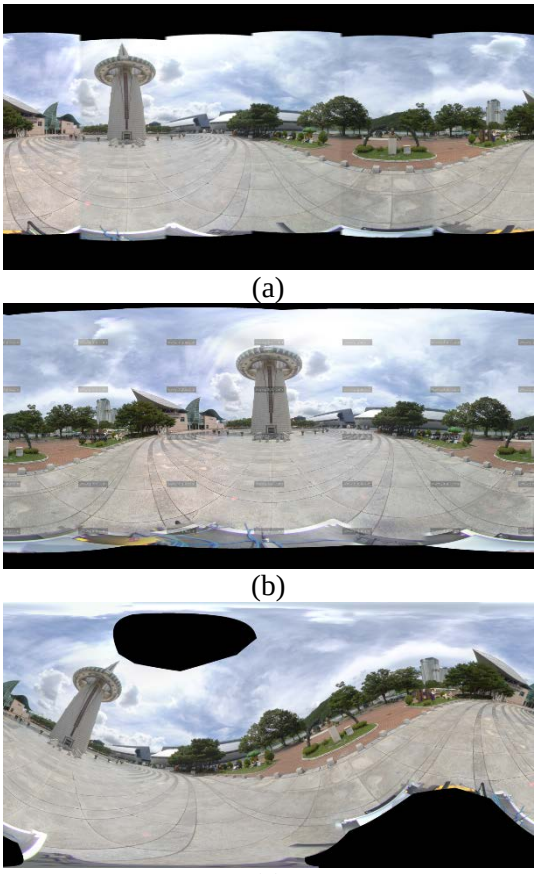
In the camera calibration module, the matched feature pairs [50] are geometrically verified. The verification method (i.e. RANSAC) ensures homography based geometric consistency among matched feature pairs. It tends to find a majority cluster such that all pairs of matched features follow the same geometric relation. The majority cluster may lead to the inclusion of a small number of defective pairs of matched features. The removal of these defective pairs of matched features may improve the geometric alignment of VR contents. Moreover, the majority clusters introduce weightage in a bundle adjustment process. As a result, the estimation of camera parameters may favor improved geometric alignment between a pair of cameras that has a majority cluster of matched features and vice versa. Therefore, we use a preprocessing method to remove defective pairs of matched features and to distribute feature pairs in a majority cluster uniformly. **Fig. 3** shows VR contents stitched without (**Fig. 3a**) and with (**Fig. 3b**) preprocessing method. For a pair of overlapping camera images a & b, the geometric alignment is estimated using (1) with l^{th} preprocessed feature pair $X_{l,a}$ and $X_{l,b}$, where $l = 1,..L_p$ and $L_p \leq L_f$ is the total number of preprocessed feature pairs between a & b.

2.1.2 Straightening

The camera calibration module uses bundle adjustment to estimate camera parameters. In bundle adjustment, pairs of matched features are mapped inside a sphere. The spherical distance between matched features is minimized by varying the camera intrinsic and extrinsic parameters. The minimum spherical distance between matched features is achieved at estimated camera parameters. Since the distance is minimized inside a sphere, the bundle adjustment process does not take the alignment axis into account. This results in wavy stitching of VR contents as shown in **Fig. 4a**. The wavy effect can be removed with the use of line features, as shown in **Fig. 2b** with **yellow** lines. The vertical texture (i.e. door, building, wall, etc.) is used to extract line features. For a camera image with line features, the geometric alignment is estimated using (2) with l^{th} line feature $X_{l,1}$ and $X_{l,2}$, where $l = 1,..L_l$, $X_{l,1}$and $X_{l,2}$ are the top and bottom x coordinates of line feature X_l and L_l is the total number of line features in the image. The spherical distance e_l between $X_{l,1}$ and $X_{l,2}$ is minimized using (2),

$$\boldsymbol{e_l} = argmin_{\beta}\left(\frac{1}{L_l}\sum_{l=1}^{L_l}\left\|X_{l,1} - X_{l,2}\right\|\right) \quad (2)$$

The bundle adjustment process is performed jointly with scale-invariant feature pairs in ($\boldsymbol{e_f}$) from (1) and line features in ($\boldsymbol{e_l}$) from (2). The line features keep the VR contents perpendicular to the z-axis of camera. VR contents stitched with scale-invariant and line features are shown in **Fig. 4b**.



(a) (b) (c)

Fig. 5. VR content stitched with; (a) Our (b) PTGui (c) Hugin.

2.2 Commercial Solutions

Most of the commercially available VR content creation solutions are non-configurable. These solutions are made of either one or two cameras. The cameras are fixed on a stand. These solutions are hardcoded with pre-calibrated LUTs. It is desirable to design a configurable VR content creation solution so that the number of cameras and the output device types can be configured in different scenarios. In this regard, our framework is reconfigurable in a way that we can attach multiple cameras and obtain calibration on the run. Moreover, we can render VR contents on different types of devices. We compare the VR framework with two reconfigurable VR content creation solutions (i.e. PTGui and Hugin). We extract frames from video streams captured by our camera rig. The frames are used to stitch VR contents using our, PTGui and Hugin software. We provide these VR contents, as shown in **Fig. 5**, to participants to assess their presence quality and cybersickness feedback. The VR contents stitched with our, PTGui, and Hugin are shown in **Fig. 5a, 5b** and **5c**.

Fig. 6. VR contents captured with (a) camera rig (left), stitched with real-time stitching system (right) and viewed on (b) smartphone (left) and 2D display (right) using VR viewing applications.

3. Experiments and Results

We asked 10 participants, with a mean age of 28.7 years (SD = 6.41), to take part in this study. The participants were university students with diverse educational backgrounds. Each participant had a normal or corrected vision and was unfamiliar with the workings of the VR framework and experimental setup. The participants performed experiments with prior consent. None of the participants had prior experience of multimedia consumption in VR. The experimental setup consists of an end-to-end VR framework. The VR framework can capture VR contents with the help of six Full HD cameras and calibrate on the run with a portable camera rig, as shown in **Fig. 6(a) (left)**, attached to a processing unit, as shown in **Fig. 6(a)**

(right). The VR framework is written in C++ using OpenCV and OpenGL APIs. The processing unit is equipped with a 4 core CPU clocked at 3.6GHz and operated with Windows 10 operating system. For presence and cybersickness feedback experiments, we recorded 14 datasets at different locations. The datasets contain diverse textures with different foreground and background conditions. We recorded the datasets prior to performing the experiments. The experiments were performed in a room set up with virtual conditions. We used three VR content consumption devices i.e. an ultrawide two-dimensional display (2D), a VR headset (VR), and a VR capable smartphone (smartphone), as shown in **Fig. 6(b)**. The 2D and smartphone are used in monoscopic mode where a single VR content is directed to both eyes of the participants with the help of VR viewing applications such that the visual area of the contents is traversed through keyboard, mouse or touch gestures. The VR is used in stereoscopic mode such that the VR contents are transformed by VR viewer application into a stereo format and two separate images are rendered for left and right eyes. The visual area of VR contents is traversed with the help of AGPS and gyroscopic sensor built into the VR (smartphone for VR rendering). The hardware specifications of each device are discussed in section 2.1. The participants observed VR contents on 2D at a distance of 100cm and with a viewing angle of 15 degrees. Participants were seated in front of the display and provided with controlling devices (i.e. a mouse and a keyboard). The VR contents are displayed on 2D using PTGui VR Viewer. The movement of contents in the VR headset was controlled by an AGPS sensor and the VR contents were manually changed after each experiment. In the case of the smartphone, we used the VR Media Player - 360° Viewer to display VR contents. The VR application is developed by Poppolab and it is controlled with touch or Galaxy S pen. Participants placed the smartphone at a distance of approximately 75cm and used it in both landscape and portrait modes. The participants spent approximately 1, 15, and 5 minutes per dataset during the visualization process on 2D, VR, and smartphone. The participants were asked to take a short break while shifting from one display device to another. Each participant was provided with VR contents from 14 datasets recorded at 120 fps. The participants were asked to view multimedia contents interactively, visually analyze the geometric consistency and stitching quality of objects in VR content. The three VR content consumption devices (i.e, 2D, Smartphone, and VR headset) are tested with the same procedures under which each participant is provided with 14 datasets to visualize on each display. The participants can select the type of display and the VR contents in a random order. The VR contents provided to participants were stitched with three calibration modes i.e. preprocessing, straightening, and preprocessing with straightening. In case of preprocessing mode, the processing time for calibrating the camera rig was 169.3 seconds without preprocessing and 53.34 seconds with preprocessing. In case of straightening mode, the processing time for calibrating the camera rig was 169.3 seconds without straightening and 173.1 seconds with straightening. For Our VR framework, the processing time for calibrating the camera rig using both preprocessing and straightening modes together resulted in 56.65 seconds. The processing time for PtGui and Hugin was 6.94 seconds and 46.77 seconds. After the session, participants were asked to briefly explain the VR contents. Typically, a session consisted of four steps. In the first step, the VR framework was calibrated and configured. In the second step, each participant was given a tutorial to get familiar with the experimental setup. In the third step, participants consumed VR contents to perform experiments. In the final step, the participants briefed their experience with VR contents. They were asked to record their responses by answering questionnaires. In this study, we have used the igroup presence questionnaire (IPQ) [18] and the simulator sickness questionnaire (SSQ) [19]. We have used questions REAL1, REAL2, SP2, SP5, INV2, INV3, and INV4 from the IPQ and General Discomfort, Fatigue, Headache,

Eye Strain, Difficulty Focusing, Difficulty Concentrating, and Blurred Vision from the SSQ. The IPQ scores are collected on 7-point Likert scale (-3 (Fully disagree), -2, -1, 0(moderate), 1, 2, 3 (Fully agree)) and the SSQ scores are collected on 4-point Likert scale(None (0), Slight (1), Moderate (2), Severe (3)). Each session lasted approximately 4 hours. In each session, VR contents were created using the three calibration modes. We designed three experiments for each calibration mode. In the first experiment, VR contents were stitched from 14 datasets using preprocessing calibration mode (Section 2.1.1). A slight improvement was observed in the geometric alignment of VR contents stitched with preprocessed features as compared to VR contents stitched without preprocessed features, as shown in **Fig. 3**.

Table 1. A quantitative comparison of average presence scores (IPQ) for different calibration modes using three immersive displays.

VR Calibration Method	**Avg. Presence Scores[1]**			**Friedman Test Scores[2]** ($p < 0.05$)		**Post hoc Test [5]** Z ($p < 0.017$)		
	2D	**Smart phone**	**VR**	**Q Score**	$\chi^2(2)$[6]	**2D-Smart phone**	**2D – VR**	**Smart phone -VR**
Without Preprocessing	3.17 *(1.14)*	3.93 *(0.80)*	4.79 *(0.69)*	18.20 ($p < .000$)	11.39	-2.68 ($p = .007$)	-2.75 ($p = .006$)	-2.76 ($p = .006$)
With Preprocessing	3.29 *(1.07)*	4.03 *(0.74)*	**4.81** *(0.56)*	18.20 ($p < .000$)	11.31	-2.69 ($p = .007$)	-2.76 ($p = .006$)	-2.76 ($p = .006$)
Without Straightening	2.21 *(1.12)*	3.04 *(0.89)*	4.03 *(0.86)*	20.00 ($p < .000$)	13.11	-2.76 ($p = .006$)	-2.77 ($p = .006$)	-2.86 ($p = .004$)
With Straightening	3.46 *(1.01)*	3.96 *(0.50)*	**4.96** *(0.50)*	16.80 ($p < .000$)	11.89	-1.62 ($p = .106$)	-2.75 ($p = .006$)	-2.86 ($p = .004$)
Our[3]	3.56 *(0.99)*	4.17 *(0.87)*	**5.01** *(0.47)*	19.05 ($p < .000$)	12.15	-2.62 ($p = .009$)	-2.76 ($p = .006$)	-2.76 ($p = .006$)
C.S.[4] PTGui	3.41 *(1.01)*	3.97 *(0.94)*	4.89 *(0.53)*	18.20 ($p < .000$)	11.93	-2.65 ($p = .008$)	-2.76 ($p = .006$)	-2.76 ($p = .006$)
C.S.[4] Hugin	2.81 *(1.11)*	3.34 *(1.09)*	4.33 *(0.90)*	12.95 ($p = .001$)	08.56	-2.24 ($p = .002$)	-2.61 ($p = .009$)	-2.65 ($p = .008$)

[1] higher IPQ scores are better. The standard deviations from average IPQ scores are added in *(brackets)* [2]The Friedman scores test differences between 2D, Smartphone and VR. [3] Both preprocessing and straightening modes are used to create VR contents. [4] C.S. stands for commercial solutions. Best results are represented in **bold** for each calibration mode. [5] The post hoc tests are performed using Wilcoxon signed rank test and the significance level is obtained after Bonferroni correction. The post hoc test reveal that except for VR contents created with straightening and viewed on 2D-vs-Smartphone, all other results showed significant difference in VR experience. [6] The chi-square values have degrees of freedom equal to 2.

The participants interactively visualized the contents on three types of content consumption devices. The aim of this experiment is to find if the quality improvement is making the VR experience better and whether the impact of quality improvement is different for VR

consumption on different types of content consumption devices. In the second experiment, the VR contents were stitched from 14 datasets using straightening calibration mode (Section 2.1.2). The wavy effect in VR contents, as shown in **Fig. 4**, is removed with the use of straightening calibration mode as compared to the VR contents stitched without straightening calibration mode. The aim of this experiment is to investigate if the participants see any improvement in VR experience with the use of straightening mode and if there is an impact of straightening on VR consumption on different types of content consumption devices. In the third experiment, 14 datasets were used to compare the stitching quality of our, PTGui, and Hugin's VR contents, as shown in **Fig. 5**. The third calibration mode (i.e. preprocessing with straightening) is used to create VR contents in our framework. The aim of this experiment is to assess the participant's VR experience for our VR framework as compared to commercial solutions (Section 2.2).

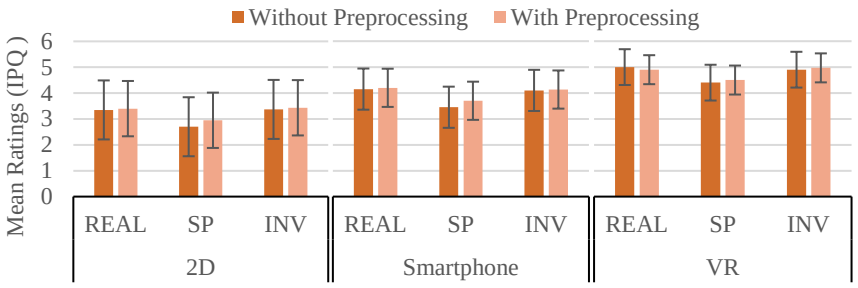
The IPQ and SSQ questionnaire responses to different VR contents are analyzed in this section. The mean ratings for IPQ at different immersion levels and statistic results of Friedman test are included in **Table 1**. The IPQ subscale mean ratings for different VR contents at three immersion levels are shown in **Fig. 7**. The mean ratings and statistic results of Friedman test for SSQ at different immersion levels are added in **Table 2**. The SSQ subscale mean ratings for different VR contents at three immersion levels are shown in **Fig. 8**.

3.1 Impact of VR content quality and display devices on presence

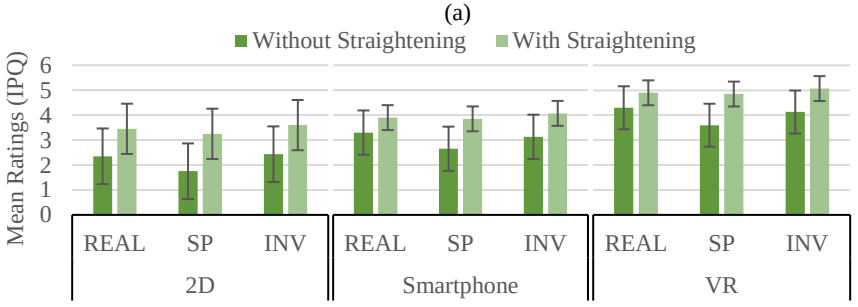
The Friedman tests on IPQ data reveal that participants experience significant difference of presence on 2D, Smartphone and VR displays. Moreover, the improvements in VR content quality, with the use of preprocessing and straightening calibration modes, resulted in a higher presence on all the three types of display devices. Finally, the comparison of our VR contents with contents created by commercial solutions showed a higher presence because of improved VR content quality.

To test the presence levels experienced by users on three types of displays, we hypothesize that user's experience same levels of presence on 2D, Smartphone and VR displays when provided with VR contents created with and without preprocessing. All user's experience all the display devices without any subgroupings. The Q scores obtained from IPQ ratings (with and without preprocessing: $Q = 18.20, p < 000$) show that the hypothesis is incorrect, and the users experience significant differences in the presence among 2D, Smartphone and VR displays. The highest average value of presence (i.e. 4.81, *SD*=0.56) for preprocessing calibration mode was obtained when VR contents were created with preprocessing and visualized with VR display. An increase of 3.65%, 2.48%, and 0.42% in average values of presence was observed for 2D, smartphone, and VR for preprocessing calibration mode. The small increase in the average presence for VR as compared to 2D conventional displays indicates that the geometric misalignment is more prominent in the case of conventional 2D displays and the distance of objects from the viewer is perceived to be greater in VR than in 2D displays. The increase in average presence values coincides with a decrease in standard deviation from mean values for both the preprocessing calibration mode and the type of display device. The average IPQ subscale values reveal that involvement (INV) (for 2D), Realism (REAL) (for Smartphone), and INV (for VR) are the biggest contributing factors for presence in case of VR contents created with preprocessing. Whereas the average IPQ subscale values reveal that INV (for 2D), REAL (for Smartphone), and REAL (for VR) are the biggest contributing factors for presence in case of VR contents created without preprocessing. The use of preprocessing resulted in an increase of 1.75%, 1.19%, and 1.41% in average values of INV (2D), REAL (Smartphone), and INV (VR). However, REAL (for VR) is decreased by 2%

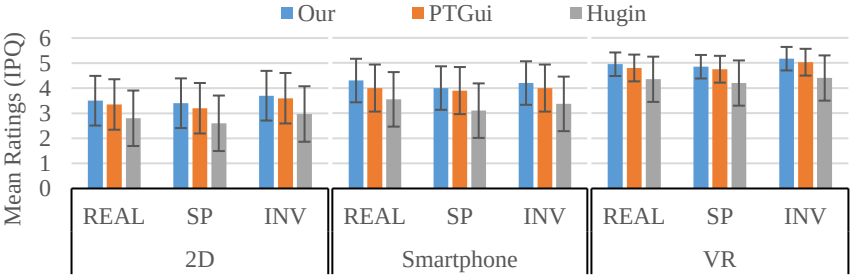
with the use of preprocessing. Overall, the IPQ subscales indicate higher average values of REAL and INV as compared to the spatial presence (SP), as shown in **Fig. 7 (a)**. The average presence values of REAL, SP, and INV were highest for VR as compared to smartphone and 2D.



(a)



(b)



(c)

Fig. 7. Mean ratings at subscales of IPQ i.e. Realism (REAL), Spatial Presence (SP) and Involvement (INV). The mean ratings of IPQ subscales are obtained by visualizing VR contents on 2D, Smartphone and VR. The VR contents are created with (a) Preprocessing calibration mode; (b) Straightening calibration mode; (c) Our VR framework and commercial solutions. The error bars correspond to standard deviation from mean ratings. (Higher values are better).

To test the presence levels experienced by users on three types of displays, we hypothesize that user's experience same levels of presence on 2D, Smartphone and VR displays when provided with VR contents created with and without straightening. All user's experience all the display devices without any subgroupings. The Q scores obtained from IPQ ratings (without straightening: $Q = 20.00, p < .000$ and with straightening: $Q = 16.80, p < .000$) show that the hypothesis is incorrect, and the users experience significant differences in the presence among 2D, Smartphone and VR displays. The highest average value of presence (i.e. 4.96, *SD*=0.50) for straightening calibration mode was obtained when VR contents were

created with straightening calibration mode and visualized on VR display. In the case of straightening calibration mode, an increase of 36.13%, 23.23%, and 18.75% in average values of presence was observed for 2D, smartphone, and VR. A bigger increase of average presence values for conventional 2D displays indicate that straightening is more effective in these displays as compared to VR display. The increase in average presence values coincides with a decrease in standard deviation from mean values for both the straightening calibration mode and the type of display device. The average presence values at the subscales of IPQ reveal that INV, REAL, and REAL are the biggest contributing factors for presence in the case of 2D, smartphone, and VR for VR contents created without straightening. However, the removal of wavy effects with straightening resulted in changing the biggest contributing factors for presence. In the case of VR contents created with straightening, the subscales of IPQ reveal that INV, INV, and INV are the biggest contributing factors for presence in the case of 2D, smartphone, and VR. The use of straightening resulted in an increase of presence by 32.00% (2D), 23.10% (smartphone), 18.54% (VR) for INV, 31.88% (2D), 15.39% (smartphone), 12.25% (VR) for REAL and 46.15% (2D), 31.17% (smartphone), 25.77% (VR) for SP. The average presence values of INV, REAL, and SP were highest for VR as compared to smartphone and 2D, as shown in **Fig. 7 (b)**.

In the case of VR contents created with our VR framework, and commercial solutions (i.e. PTGui and Hugin), the presence levels experienced by users on three types of displays are tested by hypothesis that user's experience same levels of presence on 2D, Smartphone and VR displays when provided with VR contents created with our framework and commercial solutions. All user's experience all the display devices without any subgroupings. The Q scores obtained from IPQ ratings (Our VR framework: $Q = 19.05, p < .000$, PtGui: $Q = 18.20, p < .000$ and Hugin: $Q = 12.95, p = .001$) show that the hypothesis is incorrect, and the users experience significant differences in the presence among 2D, Smartphone and VR displays. The highest average value of presence (i.e. 5.01, *SD*=0.47) was obtained when VR contents were created with our VR framework and consumed on VR display. Overall, an increase in average presence by 4.21% (2D), 4.80% (smartphone), and 2.40% (VR) was observed for VR contents created by our VR framework as compared to PTGui. Comparing VR contents created by our framework with Hugin, an increase in average presence by 21.07% (2D), 19.90% (smartphone), and 13.57% (VR) was observed. The improvements in VR content quality affect the presence scores on conventional 2D displays more than the VR display. The increase in average presence values coincides with a decrease in standard deviation from mean values for all the VR solutions as well as the type of display devices. The average presence values at the subscales of IPQ reveal that INV, REAL, and INV are the biggest contributing factors for presence in the case of 2D, smartphone, and VR for our VR framework and commercial solutions. A comparison of VR contents created from our framework and PTGui revealed an increase of presence by 2.70% (2D), 4.76%(smartphone), 2.71% (VR) for INV, 4.29% (2D), 6.98%(smartphone), 3.03% (VR) for REAL and 5.88% (2D), 2.5%(smartphone), 2.06% (VR) for SP while consuming VR contents on 2D, smartphone and VR. A comparison of VR contents created from our framework and Hugin resulted in an increase of presence by 19.73% (2D), 19.76% (smartphone), 14.89% (VR) for INV, 20.00% (2D), 17.44% (smartphone), 12.12% (VR) for REAL and 23.53% (2D), 22.50% (smartphone), 13.40% (VR) for SP. The average presence values of INV, REAL, and SP were highest for VR as compared to smartphone and 2D, as shown in **Fig. 7 (c)**.

Table 2. A quantitative comparison of average cybersickness scores (SSQ) for different calibration modes at three immersive displays.

VR Calibration Method	Avg. Cybersickness Scores [1]			Friedman Test Score[2] ($p < 0.05$)		Post hoc test [5] Z ($p < 0.017$)		
	2D	Smart phone	VR	Q Score	$\chi^2(2)$[6]	2D-Smart phone	2D-VR	Smart phone-VR
Without Preprocessing	6.36 *(2.92)*	20.20 *(5.84)*	35.16 *(9.49)*	20.00 ($p < .000$)	13.43	-2.76 ($p = .006$)	-2.76 ($p = .006$)	-2.76 ($p = .006$)
With Preprocessing	**4.49** *(2.24)*	17.20 *(4.79)*	32.91 *(8.98)*	20.00 ($p < .000$)	13.21	-2.78 ($p = .006$)	-2.76 ($p = .006$)	-2.76 ($p = .006$)
Without Straightening	12.34 *(6.04)*	22.44 *(6.26)*	39.64 *(12.08)*	15.65 ($p < .000$)	11.55	-2.34 ($p = .019$)	-2.75 ($p = .006$)	-2.61 ($p = .009$)
With Straightening	**3.37** *(2.62)*	17.95 *(4.36)*	37.40 *(6.26)*	20.00 ($p < .000$)	13.19	-2.76 ($p = .006$)	-2.76 ($p = .006$)	-2.76 ($p = .006$)
Our [3]	**5.24** *(1.83)*	14.59 *(6.57)*	28.05 *(7.53)*	17.15 ($p < .000$)	12.05	-2.62 ($p = .009$)	-3.10 ($p = .002$)	-2.55 ($p = .001$)
C.S. [4] PTGui	7.48 *(2.90)*	19.07 *(8.44)*	32.54 *(9.62)*	15.65 ($p < .000$)	10.72	-2.61 ($p = .009$)	-2.75 ($p = .006$)	-2.40 ($p = .016$)
C.S. [4] Hugin	17.95 *(6.86)*	25.06 *(13.28)*	40.77 *(10.37)*	11.25 ($p = .003$)	08.01	-1.84 ($p = .066$)	-1.84 ($p = .066$)	2.21 ($p = .027$)

[1] Lower SSQ scores are better. The standard deviations from average SSQ scores are added in *(brackets)* [2]The Friedman scores test differences between 2D, Smartphone and VR. [3] Both preprocessing and straightening modes are used to create VR contents. [4] C.S. stands for commercial solutions. Best results are represented in **bold** for each calibration method. [5] The post hoc tests are performed using Wilcoxon signed rank test and the significance level is obtained after Bonferroni correction. The post hoc test reveal that except for VR contents created without straightening (viewed on 2D-vs-Smartphone) and contents created using Hugin (view on all three displays), all other results showed significant difference in VR experience. [6] The chi-square values have degrees of freedom equal to 2.

3.2 Impact of VR content quality and display devices on cybersickness

The Friedman tests on SSQ data reveal that participants experience significant difference of cybersickness on 2D, Smartphone and VR displays. There was a slight reduction in cybersickness level with the use of VR contents created with preprocessing and straightening calibration modes. The comparison of our VR contents with contents created by commercial solutions showed a lower cybersickness because of improved VR content quality.

To test the cybersickness levels experienced by users on three types of displays, we hypothesize that user's experience same levels of cybersickness on 2D, Smartphone and VR displays when provided with VR contents created with and without preprocessing. All user's experience all the display devices without any subgroupings. The Q scores obtained from SSQ ratings (with and without preprocessing: $Q = 20.00, p < .000$) show that the hypothesis is incorrect, and the users experience significant differences in the cybersickness among 2D,

Smartphone and VR displays. There is higher reduction in cybersickness, with the use of preprocessed VR contents, for conventional 2D displays as compared to VR. This reduction of cybersickness can be explained in terms of the reduction in average cybersickness ratings for 2D (29.40%), smartphone (14.85%), and VR (6.40%). In the case of preprocessing calibration mode, the lowest cybersickness was observed with the use of VR contents created with preprocessing and visualized on 2D display whereas the highest cybersickness was observed with the use of VR contents created without preprocessing and visualized on VR. The reduction of average cybersickness ratings in **Table 2** coincides with an increase in standard deviation from mean values for both the preprocessing calibration mode and the type of display devices. The average cybersickness values at the subscales of SSQ reveal that Oculomotor (O) is the biggest contributing factor for cybersickness for all three types of displays and both with and without preprocessing. The contribution of O values in overall cybersickness is followed by Nausea (N) and Disorientation (D). The use of preprocessing resulted in a decrease in cybersickness ratings of O by 25.00% (2D), 14.29% (smartphone), 4.76% (VR), N by 30.82% (2D), 7.69% (smartphone), 9.52% (VR), and D by 50.00% (2D), 33.33% (smartphone), 10.00% (VR), as shown in **Fig. 8 (a)**.

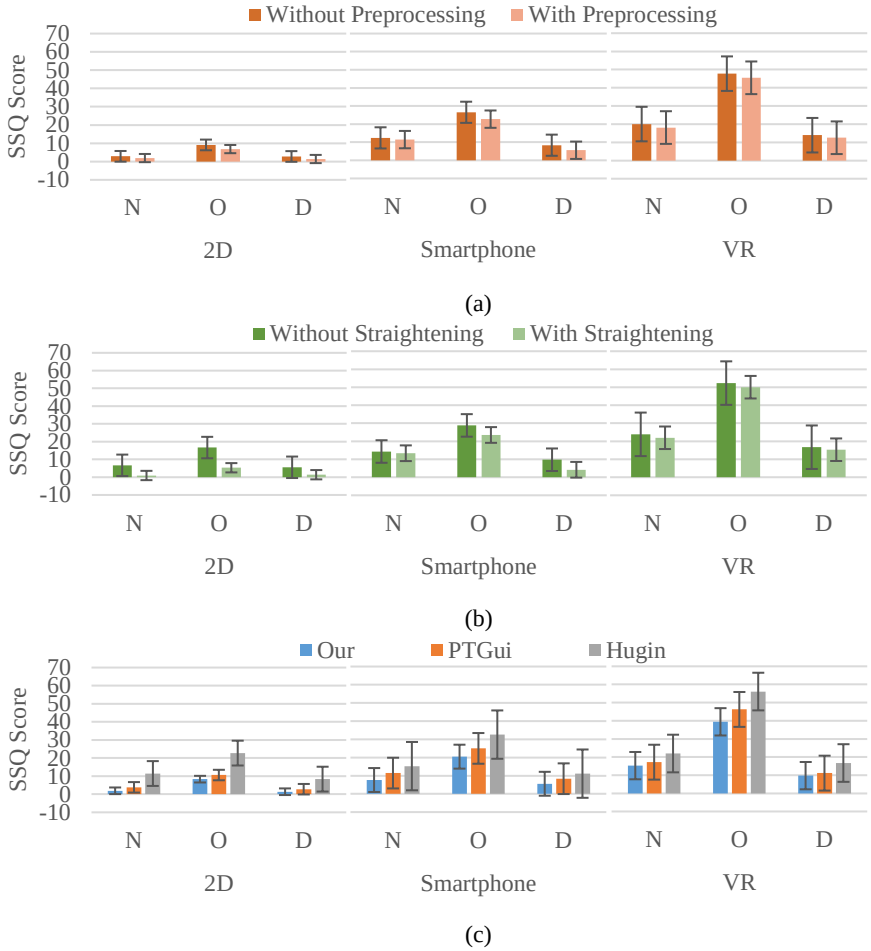


Fig. 8. Mean scores at subscales of SSQ i.e. Nausea (N), Oculomotor (O) and Disorientation (D). The mean ratings of SSQ subscales are obtained by visualizing VR contents on 2D, Smartphone, and VR. The VR contents are created with (a) Preprocessing calibration mode; (b) Straightening calibration mode; (c) Our VR framework and commercial solutions. The error bars correspond to standard deviation from mean ratings. (Lower values are better).

To test the cybersickness levels experienced by users on three types of displays, we hypothesize that user's experience same levels of cybersickness on 2D, Smartphone and VR displays when provided with VR contents created with and without straightening. All user's experience all the display devices without any subgroupings. The Q scores obtained from SSQ ratings (without straightening: $Q = 15.65, p < .000$ and with straightening: $Q = 20.00, p < .000$) show that the hypothesis is incorrect, and the users experience significant differences in the cybersickness among 2D, Smartphone and VR displays. . The use of straightening results in the reduction of average cybersickness ratings for 2D (72.69%), smartphone (20.01%), and VR (5.65%). In the case of straightening calibration mode, the lowest cybersickness was observed with the use of VR contents created with straightening and visualized on 2D display whereas the highest cybersickness was observed with the use of VR contents created without straightening and visualized on VR. The reduction of average cybersickness ratings in **Table 2** coincides with an increase in standard deviation from mean values for both the straightening calibration mode and the type of display devices. The average cybersickness values at the subscales of SSQ reveal that O is the biggest contributing factor for cybersickness for all three types of displays and both with and without straightening. The contribution of O values in overall cybersickness is followed by N and D. The use of straightening resulted in a decrease in cybersickness ratings of O by 68.18% (2D), 18.42% (smartphone), 4.35% (VR), N by 85.71%(2D), 6.67% (smartphone), 8.00% (VR), and D by 75.00% (2D), 57.14% (smartphone), 8.33% (VR), as shown in **Fig. 8 (b)**. In the case of VR contents created with our VR framework, and commercial solutions (i.e. PTGui and Hugin), the cybersickness levels experienced by users on three types of displays are tested by hypothesis that user's experience same levels of cybersickness on 2D, Smartphone and VR displays when provided with VR contents created with our framework and commercial solutions. All user's experience all the display devices without any subgroupings. The Q scores obtained from SSQ ratings (Our VR framework: $Q = 17.15, p < .000$, PtGui: $Q = 15.65, p < .000$ and Hugin: $Q = 11.25, p < 0.004$) show that the hypothesis is incorrect, and the users experience significant differences in the cybersickness among 2D, Smartphone and VR displays. The highest average value of cybersickness (i.e. 40.77, *SD*=10.37) was obtained when VR contents were created with Hugin and consumed on VR display whereas the lowest average value of cybersickness (i.e. 5.24, *SD*=1.83) was obtained when VR contents were created with our VR framework and consumed on 2D. Overall, a decrease in average cybersickness by 29.95% (2D), 23.49% (smartphone), and 13.80% (VR) was observed for VR contents created by our VR framework as compared to PTGui. Comparing VR contents created by our framework and Hugin, a decrease in average cybersickness by 70.81% (2D), 41.78% (smartphone), and 31.20% (VR) was observed. The reduction in cybersickness scores and improvements in VR content quality are more effective on conventional 2D displays as compared to VR. The average cybersickness values at the subscales of SSQ reveal that O is the biggest contributing factor for cybersickness when VR contents are created with our VR framework and commercial solutions and visualized on three types of displays. The contribution of O values in overall cybersickness is followed by N and D. A comparison of VR contents created from our framework and PTGui reveal a decrease in cybersickness by 21.43% (2D), 18.18%(smartphone), 14.75% (VR) for O, 50.00% (2D), 33.33%(smartphone), 11.11% (VR display) for N and 50.00% (2D), 33.33%(smartphone), 12.50% (VR) for D. A comparison of VR contents created from our framework and Hugin resulted in a decrease in cybersickness by 63.33% (2D), 37.21% (smartphone), 29.73% (VR) for O, 83.33% (2D), 50.00% (smartphone), 30.44% (VR) for N and 83.33% (2D), 50.00% (smartphone), 41.67% (VR) for D, shown in **Fig. 8 (c)**. The average cybersickness values of O, N, and D were highest for VR as compared to smartphone and 2D.

The quantitative relationship between the VR calibration quality and the user response in the form of cybersickness and presence is analyzed in **Table 3**. The VR calibration quality is measured using geometric and photometric metrics. The geometric metric used for quality measurement is average projection error in pixels. The projection error is calculated by projecting SIFT and line features in a sphere and the corresponding spherical distance between matching features is measured in pixels. The photometric metric used for quality measurement is called structural similarity index (SSIM). The SSIM is measured by photometrically comparing a region from the original image with the corresponding region in reconstructed VR content. The results from **Table 3** show that the quantitative improvements in VR calibration directly correlate with the improvements in the presence and a reduction in the cybersickness scores. In the next section, the results from section 3 are discussed in detail.

Table 3. A quantitative comparison of VR content quality with presence (IPQ) and cybersickness (SSQ).

Calibration mode	**Quantitative Results**			**Presence (IPQ)[3] and cybersickness (SSQ)[4] scores on different display types**					
				2D		**Smartphone**		**VR**	
	Avg. Proj. Error (px)[1]	**SSIM (%)[2]**	**Processing Time (seconds)**	**IPQ**	**SSQ**	**IPQ**	**SSQ**	**IPQ**	**SSQ**
Without Preprocessing	2.38	78.14	169.3	3.17	6.36	3.93	20.20	4.79	35.16
With Preprocessing	2.27	83.01	53.34	3.29	4.49	4.03	17.20	4.81	32.91
Without Straightening	2.38	78.14	169.3	2.21	12.34	3.04	22.44	4.03	39.64
With Straightening	2.28	83.22	173.1	3.46	**3.37**	3.96	17.95	4.96	37.40
Our [5]	**2.26**	**84.03**	56.65	3.56	5.24	4.17	14.59	**5.01**	28.05
C.S. PTGui [6,7]	2.34	78.46	**06.94**	3.41	7.48	3.97	19.07	4.89	32.54
C.S. Hugin [6]	5.94	71.32	46.77	2.81	17.95	3.34	25.06	4.33	40.77

[1] Average projection error in pixels (lower is better). [2] Percentage values of structural similarity index. (higher is better). [3] Higher IPQ scores are better. [4] Lower SSQ scores are better. [5] Both preprocessing and straightening modes are used to create VR contents. [5] C.S. stands for commercial solutions. (**Bold** numbers represent best results). [7] C.S PTGui uses GPU for processing.

4. Discussion

To analyze the effects of the quality of VR contents and the immersion levels of displays on the participant's feedback (i.e. presence and cybersickness), multiple experiments are performed using 14 diverse datasets. First, the quality of VR contents, created from the datasets, is improved with the help of preprocessing and straightening calibration modes. The participant's QoE (i.e. presence and cybersickness) is measured in experimental setups similar to previous studies [20-27]. Second, a comparison of the presence and cybersickness ratings of participants towards VR contents created from our VR framework [40] and commercial solutions [38, 39] is performed. Third, the VR QoE is compared among conventional 2D displays and VR headset in a setup similar to [16,51]. Overall, the results of this research show that the quality of VR contents and the immersion level of displays directly influence the QoE of participants.

4.1 Effects of improved content quality on VR experience

The quality of VR contents is a broad term that may relate to the specifications of a video stream (i.e. the resolution, framerate, FoV, etc.), the type of camera movements (i.e. fixed, horizontal, or vertical movement), the number of moving objects in foreground or background, etc. The previous studies [23, 24, 26] evaluate the VR experience by improving these qualities and subjectively measuring the feelings of immersion and cybersickness. This study is novel regarding the quality of the VR contents such that the VR contents are improved in terms of geometric alignment of camera rig at the hardware level. The VR experience is measured by comparing the subjective ratings of presence and cybersickness using preprocessing and straightening calibration modes. The results from questionnaire data show that the presence increased for both preprocessing and straightening calibration modes irrespective of the type of display, in connection with previous research [26]. The relative increase in the presence, with respect to the quality improvements, was higher for conventional 2D immersive displays as compared to VR. However, the increase in presence for VR headset shows that the quality of VR contents does influence the feelings of presence in a VR environment. The increase in presence values is reported to be higher for straightening mode as compared to the preprocessing mode. A higher relative increase in presence for straightening mode shows that the vertical movements and upright alignment of objects in VR contents [26] play a significant role in increasing presence. The questionnaire results show that the cybersickness level of participants was reduced with the improvements in the quality of VR contents, in connection with previous studies [23,24,26]. With both preprocessing and straightening calibration modes, there was a reduction in cybersickness for all types of displays [23, 24]. There was a relatively higher reduction in cybersickness, with respect to the quality improvements, for conventional 2D immersive displays as compared to VR headset. The straightening mode was more effective for conventional 2D immersive displays as compared to VR. The reduction in cybersickness with the use of preprocessing and straightening modes was similar for VR headset. In other words, the vertical movements and upright alignment of objects cause less cybersickness in a VR headset as compared to conventional 2D immersive displays [23,26]. Overall, the cybersickness level for conventional 2D displays and VR headset was not very strong, and the users were satisfied with the VR experience.

4.2 Comparison of our VR framework with commercial solutions

The previous studies on presence and cybersickness in a VR environment either use VR contents from the same software on different HMDs [22,27] or use variations in VR contents created from the same software on a single HMD [20,23,24,25,26,51]. This work is novel such that VR contents are created from the same datasets, but different software's and the analysis is performed on three types of displays. The results show that the improvements in VR quality improve the presence and reduce cybersickness irrespective of the type of display. The same datasets are used to create VR contents with our VR framework [40] and commercial solutions (i.e. PTGui [38] and Hugin [39]) and provided as supplementary materials. The quality improvements in our framework result in the higher presence and lower cybersickness as compared to the commercial solutions in connection with the improvements in the geometric alignment reported in [40].

4.3 Comparison of VR experience on three types of immersive displays

The questionnaire data show that both presence and cybersickness faced by users are influenced significantly by the type of display device used in a VR environment. In general,

the use of the VR headset improved the presence ratings, irrespective of the VR content quality, in connection with previous research [16,27]. In the case of conventional 2D immersive displays, the results show a higher presence for smartphone as compared to a 2D. On the contrary, the cybersickness level increased significantly for the VR headset as compared to conventional 2D immersive displays, in connection with previous research [27]. In the case of conventional 2D immersive displays, the cybersickness ratings for 2D were lower than the cybersickness ratings for smartphone. The increased cybersickness level for the VR headset can be explained by the increase in immersion level, visualization in stereoscopic mode and interaction with the virtual environment. From the results in section 3.2, it can be concluded that the participants suffer cybersickness in VR display for all cases with headaches, eyestrain and blurred vision as the most common symptoms. The high cybersickness can be attributed to the use of stereoscopic visualization in VR as compared to the monoscopic visualization in 2D and smartphone. Overall, the cybersickness ratings for VR headset were not very strong and participants were satisfied with the VR experience, in connection with previous research [20,21,22,25].

The present research can be improved in the future by dealing with many limitations such as the number of participants, the variations in VR contents, and the diversity of display devices. Moreover, the experimental setup can be enhanced in a crossover manner to mitigate the carryover effects [27] while switching display devices. Although a framerate of 120fps in our VR contents is sufficient for removing any VR framework induced cybersickness or loss in the presence [51], further experiments using VR contents created with different framerates can be performed. Despite the above-mentioned limitations, the present research analyzes the impacts of VR content quality and the types of display devices on the presence and cybersickness in a VR environment, efficiently.

5. Conclusion

The baseline study provides an analysis of factors that influence the VR experience. First, an extensive analysis of the impacts of VR content quality on the VR experience is provided. The VR content quality is improved with the help of preprocessing and straightening calibration modes in our VR framework. Second, the impacts of different types of display devices on the VR experience is studied. Finally, a comparison of VR experience is performed using VR contents from our VR framework and commercial solutions. The overall results show a higher presence and increased cybersickness for VR headset as compared to immersive 2D displays. The improvements in VR content quality result in an enhanced presence and reduced cybersickness irrespective of the type of display. The improvements in VR contents with our VR framework, as compared to commercial solutions, result in a better VR experience.

References

[1] M. G. Maggio, A. Naro, G. L. Rosa, A. Cambria, P. Lauria, L. Billeri, D. Latella, A. Manuli, and R. S. Calabrò, “Virtual Reality Based Cognitive Rehabilitation in Minimally Conscious State: A Case Report with EEG Findings and Systematic Literature Review,” *Brain Sci.*, vol. 10, no. 7, pp. 414, 2020. Article (CrossRef Link)

[2] I. Cortés-Pérez, F. A. Nieto-Escamez, and E. Obrero-Gaitán, “Immersive Virtual Reality in Stroke Patients as a New Approach for Reducing Postural Disabilities and Falls Risk: A Case Series,” *Brain Sci.*, vol. 10, no. 5, pp. 296, 2020. Article (CrossRef Link)

[3] I. Valori, R. Bayramova, P. E. McKenna-Plumley, and T. Farroni, “Sensorimotor Research Utilising Immersive Virtual Reality: A Pilot Study with Children and Adults with Autism Spectrum Disorders,” *Brain Sci.*, vol. 10, no. 5, pp. 259, 2020. Article (CrossRef Link)
[4] R. S. Calabrò, and A. Naro, “Understanding Social Cognition Using Virtual Reality: Are We still Nibbling around the Edges?,” *Brain Sci.*, vol. 10, no. 1, pp. 17, 2020. Article (CrossRef Link)
[5] J. Ratcliff, A. Supikov, S. Alfaro, and R. Azuma, “ThinVR: Heterogeneous microlens arrays for compact, 180 degree FOV VR near-eye displays,” *IEEE Trans. Vis. Comput. Graphics*, vol. 26, no. 5, pp. 1981-1990, 2020. Article (CrossRef Link)
[6] T. Chen, L. Xu, X. Xu, and K. Zhu. "Gestonhmd: Enabling gesture-based interaction on low-cost vr head-mounted display," *IEEE Trans. Vis. Comput. Graphics*, vol. 27, no. 5, pp. 2597-2607, 2021. Article (CrossRef Link)
[7] S. Weech, S. Kenny, and M. B. Cowan, "Presence and cybersickness in virtual reality are negatively related: a review," *Front. Psychol.*, vol. 10, pp.158, 2019. Article (CrossRef Link)
[8] S. Martirosov, M. Bureš, and T. Zítka, "Cyber sickness in low-immersive, semi-immersive, and fully immersive virtual reality," *Virtual Reality*, vol. 26, no. 1, pp. 15-32, 2022. Article (CrossRef Link)
[9] S. Palmisano, and R. Constable, "Reductions in sickness with repeated exposure to HMD-based virtual reality appear to be game-specific," *Virtual Reality*, pp. 1-17, 2022. Article (CrossRef Link)
[10] S. Thorp, A. S. Ree, and S. Grassini, "Temporal Development of Sense of Presence and Cybersickness during an Immersive VR Experience," *Multimodal Technol. Interact.*, vol. 6, no. 5, pp. 31, 2022. Article (CrossRef Link)
[11] J. Teixeira, and S. Palmisano, "Effects of dynamic field-of-view restriction on cybersickness and presence in HMD-based virtual reality," *Virtual Reality*, vol. 25, no. 2, pp. 433-445, 2021. Article (CrossRef Link)
[12] M. Slater, and S. Wilbur, “A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments,” *Presence: Teleoperators Virtual Environ.*, vol. 6, no. 6, pp. 603-616, 1997. Article (CrossRef Link)
[13] M. Magalhães, M. Melo, M. Bessa, and A. F. Coelho, "The Relationship Between Cybersickness, Sense of Presence, and the Users’ Expectancy and Perceived Similarity Between Virtual and Real Places," *IEEE Access*, vol. 9, pp. 79685-79694, 2021. Article (CrossRef Link)
[14] Y. Ryan, C. Khoo-Lattimore, and L. E. Potter, “Virtual reality and tourism marketing: Conceptualizing a framework on presence, emotion, and intention,” *Curr. Issues Tour.*, vol. 24, no. 11, pp. 1505-1525, 2021. Article (CrossRef Link)
[15] J. J. Cummings, and J. N. Bailenson, “How Immersive Is Enough? A Meta-Analysis of the Effect of Immersive Technology on User Presence,” *Media Psychol.*, vol. 19, no. 2, pp. 272-309, 2016. Article (CrossRef Link)
[16] J. Freeman, J. Lessiter, K. Pugh, and E. Keogh, “When presence and emotion are related, and when they are not,” in *Proc. of 8th International Workshop Presence (PRESENCE)*, pp. 213–219, London, UK, 21–23 September 2005.
[17] G. Makransky, T. S. Terkildsen, and R. E. Mayer, “Adding immersive virtual reality to a science lab simulation causes more presence but less learning,” *Learn. Instr.*, vol. 60, pp. 225-236, 2019. Article (CrossRef Link)
[18] T. Schubert, F. Friedmann, and H. Regenbrecht, “The experience of presence: Factor analytic insights,” *Presence Teleoperators Virtual Environ.*, vol. 10, no. 3, pp. 266-281, 2001. Article (CrossRef Link)
[19] R. S. Kennedy, N. E. Lane, S. Kevin, and M. G. Lilienthal, “Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness,” *Int. J. Aviat. Psychol.*, vol. 3, no. 3, pp. 203-220, 1993. Article (CrossRef Link)
[20] B. Kjell, E. Dima, T. Qureshi, M. Johanson, M. Andersson, and M. Sjöström, “Latency impact on Quality of Experience in a virtual reality simulator for remote control of machines,” *Signal Process.: Image Commun.*, vol. 89, pp. 116005, 2020. Article (CrossRef Link)

[21] E. Majed, H. J. Zepernick, Y. Hu, T. M. C. Chu, and V. Sundstedt, "Evaluation of Simulator Sickness for 360° Videos on an HMD Subject to Participants' Experience with Virtual Reality," in *Proc. of IEEE Conf. Virtual Reality 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 477-484, Atlanta, GA, USA, 22-26 March 2020. Article (CrossRef Link)
[22] G. Gregor, H. Lu, and J. Guna, "Effect of VR technology matureness on VR sickness," *Multimedia Tools Appl.*, vol. 79, no. 21, pp. 14491-14507, 2020. Article (CrossRef Link)
[23] S. Kim, S. Lee, and Y. M. Ro, "Estimating VR Sickness Caused By Camera Shake in VR Videography," in *Proc. of IEEE International Conf. Image Proc. (ICIP)*, pp. 3433-3437, Abu Dhabi, UAE, 25-28 October 2020. Article (CrossRef Link)
[24] S. Ashutosh, S. Fremerey, W. Robitza, and A. Raake, "Measuring and comparing QoE and simulator sickness of omnidirectional videos in different head mounted displays," in *Proc. of 9th International Conf. Quality Multimedia Experience (QoMEX)*, pp. 1-6, Erfurt, Germany, 31May–2June 2017. Article (CrossRef Link)
[25] B. Kjell, M. Sjöström, M. Imran, M. Pettersson, and M. Johanson, "Quality of experience for a virtual reality simulator," in *Proc. of IS&T Int'l. Symp. on Electronic Imaging: Human Vision and Electronic Imaging*, pp. 1-9, 2018. Article (CrossRef Link)
[26] M. S. Anwar, J. Wang, S. Ahmad, A. Ullah, W. Khan, and Z. Fei, "Evaluating the Factors Affecting QoE of 360-Degree Videos and Cybersickness Levels Predictions in Virtual Reality," *Electron.*, vol. 9, no. 9, pp. 1530, 2020. Article (CrossRef Link)
[27] M. Lieze, V. C. Jelle, B. Deforche, V. D. W. Nico, M. Mario, and D. V. Dyck, "Using virtual reality to investigate physical environmental factors related to cycling in older adults: A comparison between two methodologies," *J. Transp. Health*, vol. 19, pp. 100921, 2020. Article (CrossRef Link)
[28] S. Saeed, R. Hafiz, A. Rasul, M. M. Khan, Y. Cho, U. Park, and J. Cha, "A unified panoramic stitching and multi-projector rendering scheme for immersive panoramic displays," *Disp.*, vol. 40, pp. 78-87, 2015. Article (CrossRef Link)
[29] K. W. Park, Y. J. Shim, M. J. Lee, and H. Ahn, "Multi-Frame Based Homography Estimation for Video Stitching in Static Camera Environments," *Sensors*, vol. 20, no. 1, pp. 92, 2020. Article (CrossRef Link)
[30] N. Zhu, "Simulation analysis of spherical panoramic mosaic," *Signal Process.*, vol. 158, pp. 190-200, 2019. Article (CrossRef Link)
[31] P. F. McLauchlan, and A. Jaenicke, "Image mosaicing using sequential bundle adjustment," *Image Vis. comput.*, vol. 20, no. 9-10, pp. 751-759, 2002. Article (CrossRef Link)
[32] C. Kanzow, N. Yamashita, and M. Fukushima, "Levenberg-marquardt methods with strong local convergence properties for solving nonlinear equations with convex constraints," *J. Comput. Appl. Math.*, vol. 173, no. 2, pp. 321-343, 2005. Article (CrossRef Link)
[33] W. Cheng, K. Chen, W. Lin, M. Goesele, X. Zhang, and Y. Zhang, "A Two-stage Outlier Filtering Framework for City-Scale Localization using 3D SfM Point Clouds," *IEEE Trans. Image Process.*, vol. 28, no. 10, pp. 4857-4869, 2019. Article (CrossRef Link)
[34] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91-110, 2004. Article (CrossRef Link)
[35] M. Brown, and D. G. Lowe, "Automatic panoramic image stitching using invariant features," *Int. J. Comput. Vis.*, vol. 74, no. 1, pp. 59-73, 2007. Article (CrossRef Link)
[36] R. Szeliski, "Image alignment and stitching: A tutorial," *Found. Trends Comput. Graph. Vis.*, vol. 2, no. 1, pp. 1-104, 2007. Article (CrossRef Link)
[37] Y. Cho, D. Kim, S. Saeed, M. U. Kakli, S. H. Jung, J. Seo, and U. Park, "Keypoint Detection Using Higher Order Laplacian of Gaussian," *IEEE Access*, vol. 8, pp. 10416-10425, 2020. Article (CrossRef Link)
[38] PTGui. [Online]. Available: https://www.ptgui.com/ (accessed on 08 June 2022).
[39] Hugin-Panorama photo stitcher. [Online]. Available: http://hugin.sourceforge.net/ (accessed on 08 June 2022).

[40] S. Saeed, M. U. Kakli, Y. Cho, J. Seo, and U. Park, "A High-Quality VR Calibration and Real-Time Stitching Framework Using Preprocessed Features," *IEEE Access*, vol. 8, pp. 190300-190311, 2020. Article (CrossRef Link)
[41] J. Cubelos, P. C. López, J. Gutiérrez, and N. García, "QoE analysis of dense multiview video with head-mounted devices," *IEEE Trans. Multimedia*, vol. 22, no. 1, pp. 69-81, 2020. Article (CrossRef Link)
[42] R. A. Rolin, J. Fooken, M. Spering, and D. Pai, "Perception of Looming Motion in Virtual Reality Egocentric Interception Tasks," *IEEE Trans. Vis. Comput. Graphics*, vol. 25, no. 10, pp. 3042-3048, 2019. Article (CrossRef Link)
[43] T. Rhee, S. Thompson, D. Medeiros, R. dos-Anjos, and A. Chalmers, "Augmented Virtual Teleportation for High-Fidelity Telecollaboration," *IEEE Trans. Vis. Comput. Graphics*, vol. 26, no. 5, pp. 1923-1933, 2020. Article (CrossRef Link)
[44] B. S. Kim, K. A. Choi, W. J. Park, S. W. Kim, S. J. Ko, "Content-preserving video stitching method for multi-camera systems," *IEEE Trans. Consum. Electron.*, vol. 63, no. 2, pp. 109-116, 2017. Article (CrossRef Link)
[45] W. Jiang, and J. Gu, "Video stitching with spatial-temporal content-preserving warping," in *Proc. of IEEE Conf. Comput. Vis. Recognit. Workshop*, pp. 42-48, Boston, MA, USA, 2015. Article (CrossRef Link)
[46] Y. Nie, T. Su, Z. Zhang, H. Sun, and G. Li, "Dynamic video stitching via shakiness removing," *IEEE Trans. Image Process.*, vol. 27, no. 1, pp. 164-178, 2018. Article (CrossRef Link)
[47] J. Kang, J. Kim, I. Lee, and K. Kim, "Minimum error seam-based efficient panorama video stitching method robust to parallax," *IEEE Access*, vol. 7, pp. 167127-167140, 2019. Article (CrossRef Link)
[48] M. U. Kakli, Y. Cho, and J. Seo, "Minimization of parallax artifacts in video stitching for moving foregrounds," *IEEE Access*, vol. 6, pp. 57763-57777, 2018. Article (CrossRef Link)
[49] T. Bertel, N. D. Campbell, and C. Richardt, "MegaParallax: Casual 360 Panoramas with Motion Parallax," *IEEE Trans. Vis. Comput. Graphics*, vol. 25, no. 5, pp. 1828-1835, 2019. Article (CrossRef Link)
[50] J. Jo, J. Seo, and J. D. Fekete, "Panene: A progressive algorithm for indexing and querying approximate k-nearest neighbors," *IEEE Trans. Vis. Comput. Graphics*, vol. 26, no. 2, pp. 1347-1360, 2020. Article (CrossRef Link)
[51] F. Hofmeyer, S. Fremerey, T. Cohrs, and A. Raake, "Impacts of internal HMD playback processing on subjective quality perception," in *Proc. of IS&T Int'l. Symp. on Electronic Imaging: Human Vision and Electronic Imaging*, pp. 219-1–219-7, 2019. Article (CrossRef Link)

Saleh Saeed received his BS degree in electrical engineering from the University of Engineering and Technology, Lahore, Pakistan, in 2011 and his MS degree in electrical engineering from the National University of Sciences and Technology, Islamabad, Pakistan, in 2015. From 2014 to 2015, he was a Research Assistant at Vision, Imaging, and Signal Processing Lab, NUST, Islamabad, Pakistan. He is currently pursuing his PhD at Sogang University, Seoul, Rep. of Korea. His research interests include image processing, computer vision, and machine learning.

Unsang Park received his BS and MS degrees in materials science from Hanyang University, Seoul, Republic of Korea, in 1998 and 2000, respectively. He received MS and PhD degrees in computer science from Michigan State University, MI, USA, in 2004 and 2009, respectively. He is currently an associate professor in the Department of Computer Science and Engineering at Sogang University, Seoul, Rep. of Korea. His research interests include pattern recognition, image processing, computer vision, and machine learning.