Reactive Displays for Virtual Reality

G S S Srinivas Rao*
Indian Institute of Technology Kanpur

Neeraj Thakur†
Indian Institute of Technology Kanpur

Vinay Namboodiri‡
Indian Institute of Technology Kanpur

ABSTRACT

The feeling of presence in virtual reality has enabled a large number of applications. These applications typically deal with 360° content. However, a large amount of existing content is available in terms of images and videos i.e 2*D* content. Unfortunately, these do not react to the viewer's position or motion when viewed through a VR HMD. Thus in this work, we propose reactive displays for VR which instigate a feeling of discovery while exploring 2*D* content. We create this by taking into account user's position and motion to compute homography based mappings that adapt the 2*D* content and re-project it onto the display. This allows the viewer to obtain a more richer experience of interacting with 2*D* content similar to the effect of viewing through the window at a scene.

We also provide a VR interface that uses a constrained set of reactive displays to easily browse through 360° content. The proposed interface tackles the problem of nausea caused by existing interfaces like photospheres by providing a natural room-like intermediate interface before changing 360° content. We perform user studies to evaluate both of our interfaces. The results show that the proposed reactive display interfaces are indeed beneficial.

Keywords: Virtual Reality, Reactive Displays

Index Terms: H.5.1 [Information interfaces and presentation]: Multimedia Information SystemsArtificial— augmented, and virtual realities H.5.2 [Information interfaces and presentation]: User Interfaces—Graphical user interfaces

1 INTRODUCTION

Reactive displays have been thought of and adopted since the very early days of computing [7]. The term 'Reactive Display' refers to a display that can react to a user's interactions (gesture, sensor, motion etc.) by changing its graphical content. The aim since early 70s has been to allow for reactive displays to promote better interaction with the user.

Recently, virtual reality (VR) [13] has been on the rise. A large number of virtual reality headsets with various capabilities have been coming up for both desktop as well as mobile systems. This has created a huge demand for 360° content. But, the amount of 360° content available on the web is very limited. Most of the content that is available, is in the form of 2*D* videos and images. This content when viewed in VR does not respond to the user's head motion or position and thus is not as interesting or engrossing as a 360° video viewed in VR. Capturing 360° content is also a problem in many cases because it requires careful placement of camera rigs at various locations. This process is expensive as well as time consuming. Therefore, a simple method for converting existing videos or images i.e. 2*D* content to an intuitive VR like experience is desirable. The ability to use existing content on a large scale in an appealing, immersive manner could benefit in further increasing the

*e-mail: gsssrao@iitk.ac.in
†e-mail:stneeraj@iitk.ac.in
‡e-mail:vinaypn@iitk.ac.in

Figure 1: Some of the existing 2D interface for images[1]. Top: Fisheye View, Bottom:Equirectangular Projection

widespread usage of VR. VR systems always had a wide range of applications in education, entertainment and industrial sectors [13], however, lack of sufficient content has been a drawback. Thus, we aim to enable the usage of widely available content through an appealing reactive VR interface to overcome this drawback.

In this paper we propose a method to convert normal 2*D* content to a reactive display in VR that takes into account user's position/motion and adapts itself accordingly. We use a simple homography [6] based warping to achieve this effect which is computationally efficient and works real-time. We also present a 3D interface to go through the 360° content that is more natural than a traditional photosphere app. We demonstrate the functionality using a demo application built in Unity3D [3] for Google Cardboard [2].

A number of other interfaces developed in past have tried to provide similar functionality. Prominent among these has been the work by Lee *et al.* [9]. In their work, the authors used projected patterns to anticipate the head movement of users to obtain reactive virtual reality displays for 3*D* content. This idea has been extended to human level interactive displays of large multi-projector immersive interfaces [14]. A number of interfaces have also been proposed based on panoramas. Tompkin *et al.* [12] created new interface tools to browse through video collection based on spatio-temporal index of the scene by embedding videos into panorama. Dalvandi *et al.* [4] demonstrated that creating a virtual tour with panoramic videos provides a better sense of presence in comparison to normal videos and slideshows. Pece *et al.* [10] showed that for complex spatio-temporal reasoning tasks, use of HMD based interfaces improved performance over tablet displays. Homography mappings have also been used for various applications. In the work by Tomioka *et al.* [11], the authors used homography mappings to render approximate user-perspective images to develop a see-through hand-held display for augmented reality. In contrast to these approaches we aim to use widely available low cost VR

[1]`https://theta360.com/en/gallery/`

HMDs to provide a reactive display for $2D$ content. Considerable research has gone into enabling interaction with $2D$ content using motion sensing devices [8, 5]. Our system has more modest aims, to enable the viewing of existing $2D$ content in a seamless fashion with a sense of immersion and depth. In this work we show that using simple geometric principles this can be achieved without requiring additional hardware with existing low (or high) cost VR HMDs.

With the development of reactive displays arises the question of the design of a suitable interface to view the 360° content. In this project, we also address this problem by proposing a simple 'window in a room' based viewing metaphor that can be used for browsing through 360° content. This is useful as it avoids switching between handheld and HMD mode of mobile devices for selecting and browsing through existing content by providing a natural VR interface. Even the existing VR interfaces for viewing 360° content have some limitations. When using a VR HMD, the viewer gets an immense feeling of presence. Therefore, changing images like a slide-show without any transition can cause nausea and is undesirable. Even the 2D interfaces for these images like fish-eye view and equi-rectangular projections (like in Figure 1) either produce distortions or don't cover the complete image. We compare the proposed reactive display interface against existing interfaces and obtain quantitative user ratings that demonstrate the efficacy of the proposed interface. The proposed method for obtaining reactive displays can result in further research into more such interface metaphors being developed that can yield better user experience.

The rest of the paper is organized as follows. Section 2 cover the theory behind the reactive displays. Section 3 covers the proposed interface of the system. Section 4 covers evaluations and discussion followed by Section 5 describing conclusion.

2 Reactive Display Theory

In this section we explain the theory behind the reactive displays. We propose two different types of reactive displays: one type for $2D$ content and another type for 360° content. We explain the theory behind each of the displays in the subsequent subsections.

2.1 Reactive Displays for $2D$ content

Currently, the approach to view $2D$ images and videos in VR is to have a virtual theatre, gallery or a room (like in Facebook Spaces [1] and Oculus Home) where the viewer can select and view the content in an undistorted manner in VR. The viewer feels this experience to be less immersive and interesting as compared to viewing 360° content in VR where he/she can rotate and look at different parts of the content. We therefore propose a reactive display interface to view visual content in VR in an interactive way. Here, by the word 'reactive display' we mean 'a display that reacts and adapts its content based on user's position and motion'.

We use homography [6] based transformation for adapting the contents of the reactive display. To explain this consider Figure 2 where we have a camera facing a 3D scene. The image plane of this camera is displayed as a plane surface in the Figure. Now, this camera moves from its initial position a to position b, by undergoing translation (t) and rotation (represented by rotation matrix R). If we represent the normal of image plane as $\hat{n}$ and the magnitude of distance between camera position to plane by d then the homography matrix H_{ba} for transformation from position a to b is given by the equation:

$$H_{ba} = R - t \cdot n^T / d \tag{1}$$

This matrix can be used to relate the position of a pixel on an image captured by camera a to its position on image captured by camera b by the following equation:

$$^{a}p_i = K_a \cdot H_{ba} \cdot K_b^{-1} \cdot {}^{b}p_i \tag{2}$$

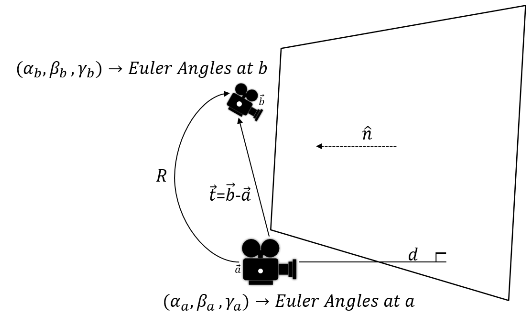

Figure 2: Illustration of the planar homography transformation

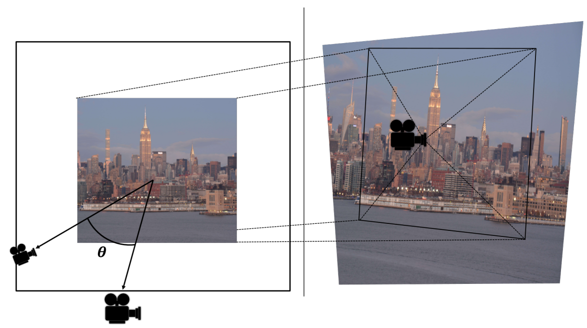

Figure 3: Left: The scene view with camera facing the reactive display Right:A camera looking at the plane undergoing homography

Here, K_a and K_b are the camera intrinsic matrices of camera a and camera b respectively. Thus, this transformation remaps the plane based on the position and viewing direction of the camera. Since this transformation is a planar homography, it is valid for all points lying in the same planar surface and for all other points we have approximations. The approach thereby provides novel view of an existing scene from a different viewpoint by approximating the scene in terms of a planar homography transformation. However, as can be seen from the user study, this is still perceived to be near-realistic 3D. In our work we ignore these approximations by assuming the images to be relatively large. This transformation can also cause the final image to not be rectangular because of the rotation matrix. Since, for our effect we only consider a small rectangular region of the final image (like in Figure 3), this does not cause any problems.

Next, we explain the use of this transformation. Consider Figure 3 where the camera on the left (corresponding to the user) in the scene view undergoes a motion in virtual world. The camera on the right is a separate camera in the virtual world looking at a image-plane which undergoes homographic transformation. The contents seen by this right camera are rendered onto the reactive display on the left. Thus, corresponding to the user's motion and position in virtual world (i.e. position and motion of the left camera), we can render the final image on the reactive display (on the left) by applying homography transformations to the image-plane on the right. But, the experience we are aiming to obtain is a window effect, where the user should feel as if they are viewing through the window at a scene. The benefit of this experience is that it incites a feeling of discovery when looking at an image and is also more natural because of pseudo 3D effect perceived. So, in the homography Equation 1 we change the matrix R to correspond to only a rotation of angle θ (where θ corresponds to the angle between the center of display and two positions of user). This assumption is based on the fact that in the real world, when looking at a window, window's content does not change as we move our eyes. Only a change in our position results in a change in window's content.

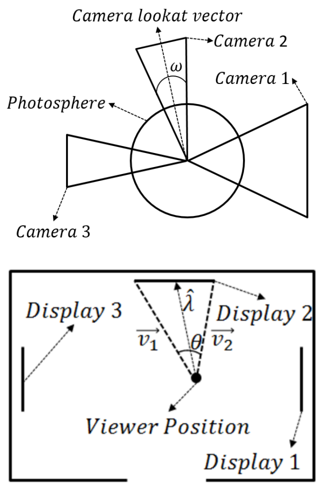

Figure 4: Illustration of the reactive display configuration for 360° image. The figure depicts the constrained collection of reactive displays.

2.2 Reactive Displays for 360° content

Presently, 360° content are viewed in VR using photospheres in a slideshow manner. While viewing photospheres in VR, one may feel disconnect or nausea on suddenly changing the image of the photosphere. Further, even the existing $2D$ interfaces like in Figure 1 are not good enough as they either produce distortions or don't cover the complete image. Thus, we feel that there is a need for a better $3D$/VR interface for accessing 360° content in an intuitive way. The recent commercial solutions like Facebook Spaces [1] explore to solve these problems by providing the viewers with an intermediate room where they can select between small spherical balls before viewing them in a complete photosphere experience. In contrast to these approaches, we propose as well as evaluate a room-like interface with a constrained collection of reactive displays to view 360° content.

In Figure 4 (bottom) we have the top view of a room with three reactive displays. When the viewer moves in the room, for each display we have a θ value (angle between the viewer and the two edges of the display) corresponding to it and a constrained set of angular values for the other two windows. Now, separately in the virtual world the 360° image/video is textured over an inverted normal sphere and 3 cameras are instantiated to the center of it (Figure 4 (top)). Image/Video captured by these cameras are rendered onto the corresponding displays in the room. As the viewer moves in the room, the parameters λ (i.e. the vector joining the viewer's position to center of display) and θ change accordingly as shown in Figure 4. We propose a relationship of this parameter to the parameter ω (camera's field of view) and camera lookat vector (dependent on ω) of cameras inside the sphere by the relation:

$$camera.FOV = \angle\omega = f \cdot \angle\theta \tag{3}$$

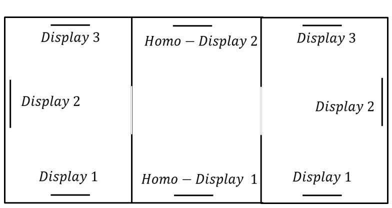

Figure 5: Top View of the application scene as developed in Unity

Figure 6: Figure showing the experimental setup used. Left: Google Cardboard(HMD), Xbox Controller and Android Smartphone Right: A subject wearing the setup

Here, f (field of view factor) is a parameter which is high for images with faraway objects and small for images with nearby/close objects. In this interface, the viewer can now walk around the gallery and decide which 360° content to view. Finally, on the trigger of a button he can jump into a photosphere experience and look at the whole 360° image if required. On triggering the button once more, the viewer can return back to the gallery.

3 Interface

In this section we describe our interface developed for a Google Cardboard application built using Unity3D game engine. The interface metaphor adopted is that of a 'window in a room' metaphor to view both existing 2D content as well as 360° content. The interface is representative and was used to evaluate the benefit of reactive displays. The scene consists of an art gallery having 3 rooms (Figure 5). 'Homo-Display' corresponds to reactive display for normal content and 'Display' corresponds to reactive display for 360° content. Two corner rooms have displays for 360° content while the center one has displays for images and videos. The interface with the reactive display is shown through a video in the supplementary material.

Each of these displays can access content from a url, so they can be connected to webcam as well. Whenever a user moves into one of the rooms, the displays adapt themselves as per the user's motion following the theory explained in the previous section. The homography transformation of pixels are done using a vertex shaders. Finally, since Unity is cross-platform, it can be exported to Android or even to a Web app. The interface developed is prototypical and meant to illustrate and evaluate the concept of reactive displays. This can be further explored and improved on. For example, as the number of 2D or 360° images increase, we can first provide the users with a simple 2D interface on phone with thumbnails to select the content they are interested in and then procedurally generate the 3D interface with reactive displays as per the number of images selected.

4 Evaluations and Discussion

4.1 Experimental Setup and Participants

For conducting user study, an experimental setup consisting of Android Smartphone, Google Cardboard (HMD) and Xbox Controller was used. The complete setup is shown in Figure 6. Here, we used a USB cable for connecting Xbox Controller but our application can run with wireless controllers as well and can be easily ported to Google Daydream. For navigation in the virtual space, rotation was obtained from HMD (gyroscope) and translation was obtained from the input to the thumb-stick of the Xbox Controller.

For the user study, we selected 10 participants (all of which were Engineering students) with a mean age of 20.6 years (SD = 1.075). 5 of the participants had prior experience using VR apps and HMDs

Table 1: MOS and t-test comparison between displays for $2D$ content based on questions in Table 2. Here, R indicates 'Reactive Display Interface' and NR indicates 'Non-Reactive Display Interface'. Significant level α = 0.05. Two sided t-critical = 2.1010 (for 10 subjects with Degrees of Freedom = 18)

Evaluation	Q1		Q2		Q3		Q4		Q5		Q6	
	R	NR	R	NR	R	NR	R	NR	R	NR	R	NR
MOS	4.8000	3.6000	4.7000	3.1000	4.0000	2.9000	4.5000	2.9000	4.3000	2.3000	4.4000	3.2000
SD	0.4216	0.9661	0.4830	0.7379	1.0541	0.5676	0.7071	0.8756	0.4830	1.3375	0.5164	0.9189
t-value	4.1295		7.2363		3.9727		5.2372		6.0000		4.1295	
p-value	0.0026		0.0000		0.0032		0.0005		0.0002		0.0026	

while 5 others were first time users. All the subjects were fluent in english and understood the experimental protocol well.

The participants were provided with necessary instructions and were asked to try out the following 4 VR apps:

1. **Reactive Display Interface for $2D$ Content**: An art gallery having reactive displays for $2D$ content (images and videos) and reactive displays for 360° images.

2. **Non-Reactive Display Interface for $2D$ Content** : An art gallery having normal displays (i.e. without homography just like portraits) for $2D$ content.

3. **Reactive Display Interface for 360° Content**: An art gallery with different rooms for each 360° image. Each room has 3 reactive displays like the setup shown in Figure 4. On triggering the button, the user can jump into photosphere experience and on triggering the button once more, he/she returns back to the art gallery.

4. **Non-Reactive Display Interface for 360° Content**: A VR photosphere app where the 360° images changes on triggering the button of Xbox Controller. The user is located in the center of photosphere and cannot move.

All the participants were given sufficient time to test all the apps to prevent any time/order based confounding effects. Please refer to the supplementary video to view the clips about the VR apps and the experience. In the video we show a monoscopic view but all the applications we developed have a stereoscopic view (i.e. used with Google Cardboard). We mainly chose landscapes and cities (refer the supplementary video) as the type of images used in our evaluations. Based on the participants' experiences, two users studies were done to draw a comparison between reactive and non-reactive display interfaces for $2D$ content and 360° content respectively. The next three subsections cover details and inferences obtained from these studies.

4.2 User Study for $2D$ Content Interfaces

This user study was conducted with 10 participants described in the previous section once they tried out VR apps 1 and 2. Each of the participant was asked to rate the 6 questions in Table 2 with a Likert scale rating where 1 indicated 'Strongly Disagree' and 5 indicated 'Strongly Agree'. Each of these correspond to a different attribute and some of these were picked from the presence questionnaire. From the scores thus obtained, Mean Opinion Scores were obtained and a two sample t-test analysis was carried out (see Table 1).

Since the p-values for all the questions are less than $\alpha = 0.05$ (which is equivalent to t-value being greater than t-critical=2.1010), we can reject the null hypothesis. Thus, from the MOS scores (see Figure 7) we can infer that the participants felt the reactive display (R) interface was more enjoyable (Q1), natural (Q2) and had a greater feeling of presence (Q6) than the non-reactive display (NR) interface. We can also infer that the participants remembered more about the $2D$ content on using the R interface (from MOS scores of Q3). We observed that in general, the participants spent more time examining the $2D$ content when using the R interface. These are probably because of the feeling of discovery instigated in the participants when using it. From the MOS scores of Q5 and Q2 we can infer that the homography approximation method seems to work well to generate the window effect. Thus, from this study we can conclude that the reactive displays for $2D$ content are indeed beneficial.

Table 2: Likert scale rating questions for $2D$ content

Q#	Question Statement
Q1	I **enjoyed** the experience
Q2	I felt that experiences in the virtual environment seem **consistent** with real world experience
Q3	I felt that I **remembered** more about the picture/video after using this interface
Q4	I was able to **examine** the image/video closely and from multiple viewpoints
Q5	I felt that the image/video was in **3D**
Q6	I was completely **engrossed** in the virtual environment experience.

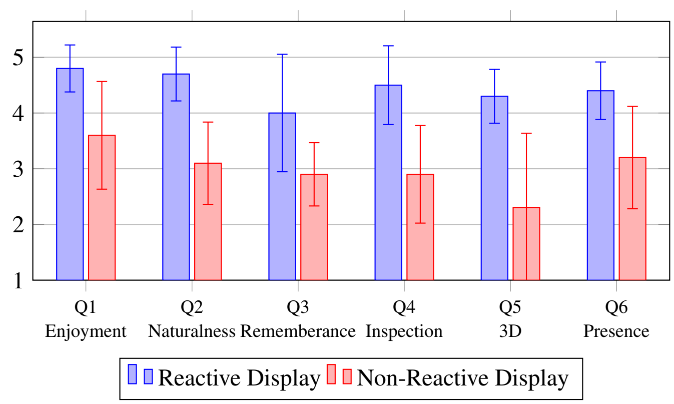

Figure 7: Likert scale MOS ratings for $2D$ content displays, error bars indicate +/-SD

4.3 User Study for 360° Content Interfaces

This user study was conducted with 10 participants described in the previous section once they tried out VR apps 3 and 4. Each of the participant was asked to rate the 6 questions in Table 4 with a Likert scale rating where 1 indicated 'Strongly Disagree' and 5 indicated 'Strongly Agree'. Each of these correspond to a different attribute and some of these were picked from the presence questionnaire.

Table 3: MOS and t-test comparison between displays for 360° content based on questions in Table 4. Here, R indicates 'Reactive Display Interface' and NR indicates 'Non-Reactive Display Interface (Simple Photosphere)'. Significant level α = 0.05. Two sided t-critical = 2.1010 (for 10 subjects with Degrees of Freedom = 18)

Evaluation	Q1		Q2		Q3		Q4		Q5		Q6	
	R	NR	R	NR	R	NR	R	NR	R	NR	R	NR
MOS	4.9000	3.7000	4.6000	3.7000	4.0000	2.9000	4.4000	2.8000	4.4000	3.0000	4.4000	3.3000
SD	0.3162	0.6749	0.6992	1.3375	0.9428	0.8756	0.5164	0.7888	0.5164	0.8165	0.5164	0.9487
t-value	6.0000		3.2504		3.9727		9.7980		5.2500		3.1608	
p-value	0.0002		0.0100		0.0032		0.0000		0.0005		0.0115	

From the scores thus obtained, Mean Opinion Scores were calculated and a two sample t-test analysis was carried out (see Table 3).

Since the p-values for all the questions are less than $\alpha = 0.05$ (which is equivalent to t-value being greater than t-critical=2.1010), we can reject the null hypothesis. Thus, from the MOS scores (see Figure 9) we can infer that the participants felt the reactive display (R) interface was more enjoyable (Q1) than the non-reactive display (NR) interface. From MOS scores of Q5 and Q6 we can infer that both the R and NR were natural and engrossing. We can also infer that the participants remembered a lot more about the 360° content on using the R interface (from MOS scores of Q3). We observed that in general, the participants spent more time examining the 360° content when using the R interface. These are probably because of the feeling of discovery instigated in the participants when using it. Though, the MOS score for NR in Q2 (=3.7) is not much lower than R (=4.6), we can see that the SD for NR is a big number (=1.3375). Thus, we can infer that introducing a room as an intermediary was successful in slightly lessening the nausea. From the MOS scores of Q5 and Q6 we can also conclude that the proposed method for changing FOV seems to work well to generate the window effect. Thus, from this study we can conclude that the reactive displays for 360° content are indeed beneficial.

4.4 Preference

While conducting both the user studies, we also asked the participants to vote for reactive display or non-reactive display or both the interfaces for their next time usage. Figure 8 shows a stacked bar plots of the preference votes for $2D$ as well as 360° content. From the plot we can infer that in general the participants preferred reactive display interface over the non-reactive one. The single vote preferring non-reactive display was because of its ease of use. We can also see that the preference votes for reactive displays for 360° is greater than for $2D$ content. This is probably because the 360° reactive displays tackle nausea.

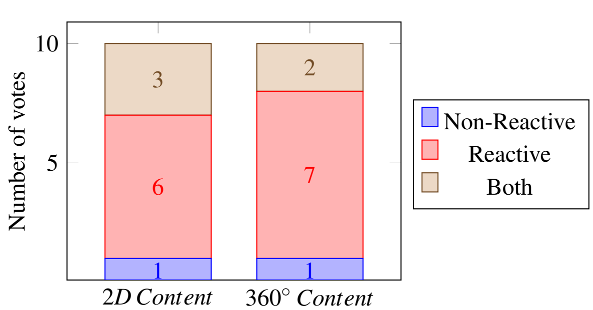

Figure 8: Preference votes

Table 4: Likert scale rating questions for 360° content

Q#	Question Statement
Q1	I **enjoyed** the experience
Q2	I did not experience **nausea** after using this interface.
Q3	I felt that I **remembered** more about the scene after using this interface
Q4	I was able to **examine** the scene closely and from multiple viewpoints
Q5	I felt that the experience was **natural**
Q6	I was completely **engrossed** in the virtual environment experience.

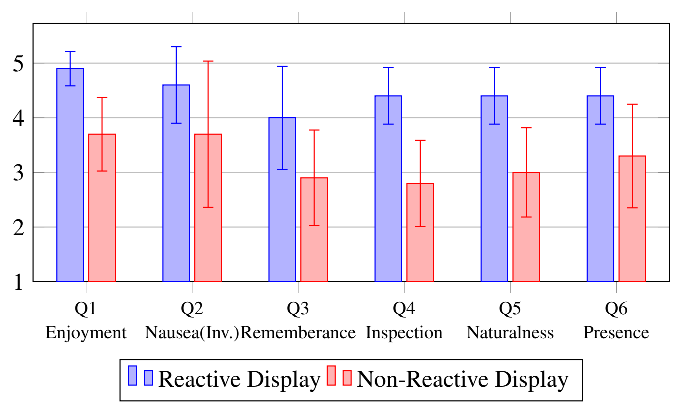

Figure 9: Likert scale MOS ratings for 360° content displays, error bars indicate +/-SD

5 Conclusion

We present in this paper, a method of making reactive display interfaces for VR which work with $2D$ as well as 360° content. This provides a means for interactively viewing existing $2D$ images/videos and 360° content in an interesting manner. The interface metaphor adopted is that of a 'window in a room' to view both existing 2D content as well as 360° content. From our experience and user studies, we feel that the proposed method will make the users feel much more comfortable than the currently prevailing VR display techniques for viewing visual content.

Our contributions include conceptualizing the idea of providing a robust, real-time geometry based approach that makes media more interact-able, developing a cross-platform application (i.e. works with iOS, Android, etc) for demonstrating the idea and validating it based on user studies.

REFERENCES

[1] Facebook spaces. `https://www.facebook.com/spaces`. [Online; accessed 26-June-2017].

[2] Google cardboard. `https://vr.google.com/cardboard/`. [Online; accessed 26-June-2017].

[3] Unity3d. `https://unity3d.com/`. [Online; accessed 26-June-2017].

[4] A. Dalvandi, B. E. Riecke, and T. Calvert. Panoramic Video Techniques for Improving Presence in Virtual Environments. In S. Coquillart, A. Steed, and G. Welch, editors, *Joint Virtual Reality Conference of EGVE - EuroVR*. The Eurographics Association, 2011.

[5] K. Hald. Low-cost 3dui using hand tracking and controllers. In *IEEE Symposium on 3D User Interfaces, 3DUI 2013, Orlando, FL, USA, March 16-17, 2013*, pages 205–206, 2013.

[6] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second edition, 2004.

[7] J. D. Joyce and M. J. Cianciolo. Reactive displays: Improving man-machine graphical communication. In *Proceedings of the November 14-16, 1967, Fall Joint Computer Conference*, AFIPS '67 (Fall), pages 713–721, 1967.

[8] J. Laviole and M. Hachet. Papart: Interactive 3d graphics and multi-touch augmented paper for artistic creation. In *IEEE Symposium on 3D User Interfaces 2012, 3DUI 2012, Costa Mesa, CA, USA, March 4-5, 2012*, pages 3–6, 2012.

[9] J. Lee, S. Hudson, and P. Dietz. Hybrid infrared and visible light projection for location tracking. In *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology*, UIST '07, pages 57–60, 2007.

[10] F. Pece, J. Tompkin, H. Pfister, J. Kautz, and C. Theobalt. Device effect on panoramic video+context tasks. In *Proceedings of the 11th European Conference on Visual Media Production*, CVMP '14, pages 14:1–14:9, New York, NY, USA, 2014. ACM.

[11] M. Tomioka, S. Ikeda, and K. Sato. Approximated user-perspective rendering in tablet-based augmented reality. In *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 21–28, Oct 2013.

[12] J. Tompkin, F. Pece, R. Shah, S. Izadi, J. Kautz, and C. Theobalt. Video collections in panoramic contexts. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*, UIST '13, pages 131–140, New York, NY, USA, 2013. ACM.

[13] J. Vince. *Virtual Reality Systems*. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 1995.

[14] T. C. Yapo, Y. Sheng, J. D. Nasman, A. Dolce, E. Li, and B. Cutler. Dynamic projection environments for immersive visualization. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR Workshops 2010, San Francisco, CA, USA, 13-18 June, 2010*, pages 1–8, 2010.