International Conference on Artificial Reality and Telexistence
Eurographics Symposium on Virtual Environments (2017)
T. Huang (Editor)

Poster

Tablet Fish Tank Virtual Reality: a Usability Study

Sirisilp Kongsilp[1,3], Mintra Ruensuk[2], Matthew N. Dailey[3], and Takashi Komuro[4]

[1] Faculty of Engineering, Thammasat University, Thailand
[2] Graduate School of Creative Design Engineering, Ulsan National Institute of Science and Technology, Korea
[3] Information and Communication Technologies, Asian Institute of Technology, Thailand
[4] Graduate School of Science and Engineering, Saitama University, Japan

Abstract

In this paper, we describe the development a tablet FTVR prototype that incorporates both motion parallax and stereo cues with the use of easy-to-find hardware. We also present findings of a usability study based on the prototype.

CCS Concepts

•***Human-centered computing*** → *Usability testing; Empirical studies in ubiquitous and mobile computing;*

1. Tablet Fish Tank Virtual Reality

In an effort to further expand the impact of VR technology, we developed a new implementation of an existing technique that allows widely accessible consumer-level tablets to display perspective-corrected 3D (Fish Tank Virtual Reality or FTVR) images. To assess the usability of the technique, we conducted a human study using a visual search task previously developed for desktop FTVR systems. We recorded participants' task performance, subjective level of presence, visual fatigue, and informal feedback. In this paper, we identify challenges and opportunities for the adoption of tablet-based FTVR and point toward appropriate directions for future research. To achieve tablet FTVR without any enhancement to the hardware itself, we combine Anaglyph 3D for stereopsis with head position tracking from the tablet's front camera. For stereo, we use Anaglyph 3D images, multiplexing two color-filtered images (red and cyan). The user views Anaglyph 3D images as shown in Figure 1. For motion parallax, following previous studies [FN11, Rek95], we use face tracking. The face tracking system tracks the user's face in real time using images from the tablet's front camera. We use the well-known Haar face detection cascade technique to find the face, and we track the detected face region using the Camshift tracking algorithm. We used the Unity game engine to develop the application, and ran it on an iPad Air (model number A1474). Based on the face tracking system and Anaglyph 3D, the application operates in four view modes. 1) *Normal 2D* view mode (2D): the application displays a static scene in 2D. 2) *Head-coupled display* view mode (HCD): the application shows perspective-corrected images according to the user's head position in 2D. 3) *Anaglyph 3D* view mode (Anaglyph): the application displays a static anaglyph 3D scene. 4) *Combined* view mode (Combined): the application shows an Anaglyph 3D scene according to the user's head position. The system runs at 60 fps in all view modes.

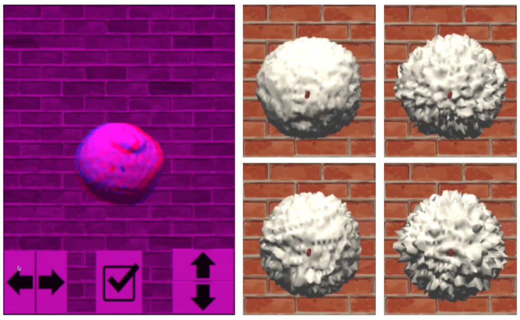

Figure 1: *Test application and the four levels of noise.*

2. Experiment

We conducted an experiment on the usability of our tablet FTVR prototype using the visual search task from the comparative study between CAVE and FTVR [DJK*06]. We recruited 40 participants (30 male and 10 female, age ranging from 17 to 31 years old). We used a 2×2 experimental design in which each participant was assigned to the *Normal 2D* group, the *Head-coupled* group, the *Anaglyph 3D* group, or the *Combined* group. To perform the task, participants had to identify the location of a rectangular bump on the surface of a noisy potato-shaped object then move it under a pole by rotating the potato using the arrow keys at the bottom of the display, as shown in Figure 1. The application records the participant's performance time and number of false identifications. To avoid ceiling effects, we made the task harder be applying four levels of noise to the potato-shaped object, as shown in Figure 1. There were 20 random trials for each participant (1 view mode

DOI: 10.2312/egve.20171377

Table 1: *Results summary. Mean and standard deviation of task performance time, error rates, SSQ scores, and PQ scores.*

	Average task performance time (second)	*Average number of error*	*SSQ score*	*PQ score*
2D	11.57 * (9.09)	0.14* (0.46)	29.50* (37.75)	76.11 (16.34)
HCD	16.43 * (15.84)	0.66$^{**\Diamond}$ (1.79)	31.42* (42.90)	70.25 (15.34)
Anaglyph	13.70 (30.85)	0.16* (0.66)	48.62 (28.21)	83.70 (16.26)
Combined	12.64 (8.64)	0.16$^\Diamond$ (0.62)	74.43** (40.83)	74.80 (11.67)

$*, \star$ and $\Diamond$ indicate statistically significant differences between two means in the same table column by a two-way ANOVA.

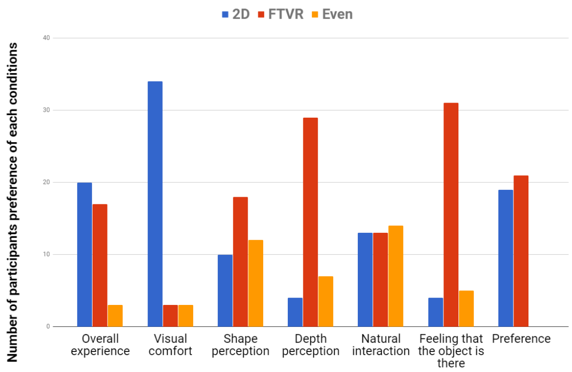

Figure 2: *Users' preference between the Normal 2D and the Combined view modes along the seven dimensions.*

× 4 difficulty levels × 5 repetitions, giving 20 trials). When the participant completed the task, the researcher immediately asked the participant to evaluate his or her level of visual fatigue with the Simulation Sickness Questionnaire (SSQ) [KLBL93], followed by the Presence Questionnaire (PQ) [WS98] to evaluate the level of presence he or she experienced. After the participant completed both questionnaires, the researcher asked the participant to use the system in the normal 2D and in the combined view modes. As the participant freely used the system, the researcher asked the participant to compare the two view modes and give his or her preference for each view mode along the seven dimensions Overall experience, Visual comfort, Shape perception, Depth perception, Natural interaction, Feeling that the object is there, and Preference.

3. Results

Here we present the results of the experiment. We dropped the data for one participant from all analyses because the time the individual took to complete the task was many standard deviations beyond the mean. The objective data are summarised in Table 1. The results of the comparison between the *Normal 2D* view mode and the combined view mode are summarized in Figure 2.

4. Discussion and Conclusion

Here we answer the underlying questions motivating this research. The first question is *How effective is tablet FTVR?* To answer the this question, we look to the results of the experiment. Although there were no statistically significant differences between the view modes for PQ scores, we suspect that this was more because of the visual discomfort from Anaglyph 3D and the front-facing camera-based tracking technique's limitations than anything else. The comparison results suggest that participants perceived depth and felt that a virtual object existed in front of them more in the *Combined* view mode, when compared to the *Normal 2D* view mode. Our findings coincide with those of Li et al. [LPWL12]. We suspect that participants were unable to perform the task better in the *Combined* view mode because of the front-facing camera-based tracking technique's limitations. This coincides with a study by Kongsilp and Dailey [KD17], who found that in desktop FTVR settings, the combination of motion parallax and stereopsis cues produces lower visual discomfort and higher subjective level of presence when compared to the stereopsis cue only. Overall, we believe that there are opportunities for tablet FTVR development. If done correctly, we believe that tablet FTVR can convey depth information and immerse users in a scene, displaying virtual objects as if they really existed in front of them. This capability would enable a new range of applications, interactions, and user experiences. The last question is *If it is useful, should we develop a new system or enhance existing devices?* We believe that it would be best to develop a new system from scratch if we absolutely require stereoscopic displays. Both polarized 3D and active shutter 3D technologies would require a fair amount of hardware changes to today's commodity tablets.

References

[DJK*06] DEMIRALP C., JACKSON C. D., KARELITZ D. B., ZHANG S., LAIDLAW D. H.: Cave and fishtank virtual-reality displays: A qualitative and quantitative comparison. *IEEE Transactions on Visualization and Computer Graphics 12*, 3 (2006), 323–330. 1

[FN11] FRANCONE J., NIGAY L.: Using the user's point of view for interaction on mobile devices. In *Proceedings of 23rd French Speaking Conference on Human-Computer Interaction* (2011), ACM, p. 4. 1

[KD17] KONGSILP S., DAILEY M. N.: Motion parallax from head movement enhances stereoscopic displays by improving presence and decreasing visual fatigue. *Displays 49* (2017), 72–79. 2

[KLBL93] KENNEDY R. S., LANE N. E., BERBAUM K. S., LILIENTHAL M. G.: Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology 3*, 3 (1993), 203–220. 1

[LPWL12] LI I. K., PEEK E. M., WÜNSCHE B. C., LUTTEROTH C.: Enhancing 3d applications using stereoscopic 3d and motion parallax. In *Proceedings of the Thirteenth Australasian User Interface Conference-Volume 126* (2012), Australian Computer Society, Inc., pp. 59–68. 2

[Rek95] REKIMOTO J.: A vision-based head tracker for fish tank virtual reality-vr without head gear. In *Proceedings of Virtual Reality Annual International Symposium, 1995. Proceedings.* (1995), IEEE, pp. 94–100. 1

[WS98] WITMER B. G., SINGER M. J.: Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and virtual environments 7*, 3 (1998), 225–240. 1

For full details of this research, please refer to https://goo.gl/pgsXSn