Tablet Fish Tank Virtual Reality: a Usability Study

Sirisilp Kongsilp
Faculty of Engineering
Thammasat University
Pathumthani, Thailand

Mintra Ruensuk
School of Creative Design Engineering
Ulsan National Institute of Science and Technology
Ulsan, South Korea

Matthew N. Dailey
Computer Science and Information Management
Asian Institute of Technology
Pathumthani, Thailand

Takashi Komuro
Graduate School of Science and Engineering
Saitama University
Saitama, Japan

Algorithm 1 Face tracking algorithm

procedure FACETRACKING
  if a face was not detected in the previous frame **then**
    use *Haar Feature-based Cascade Classifier* to search for a face in the frame.
    if face found **then**
      $camShiftCounter = 0.$
      return new face's size and position
    end if
  else
    $searchPosition \leftarrow lastKnownFacePosition.$
    $searchArea \leftarrow lastKnownFaceSize \times 2.$
    use *Haar Feature-based Cascade Classifier* to search for a face in the *searchArea* at the *searchPosition* of the frame.
    if face found **then**
      $camShiftCounter = 0.$
      return new face's size and position
    else
      use *Camshift* to approximate the face region in the *searchArea* at the *searchPosition* of the frame.
      $camShiftCounter = camShiftCounter + 1.$
      if $camShiftCounter <= 5$ **then**
        return new face's size and position
      else
        $camShiftCounter = 0.$
        return no face detected
      end if
    end if
  end if
end procedure

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.

Tablet Fish Tank Virtual Reality: 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. For motion parallax, following previous studies [FN11, Rek95]. See Algorithm 1 for more detail. We used the Unity game engine to develop the application, and ran it on an iPad Air (model number A1474). The application operates in four view modes: Normal 2D (2D), Head-coupled display (HCD), Anaglyph 3D (Anaglyph), and Combined view mode (Combined).

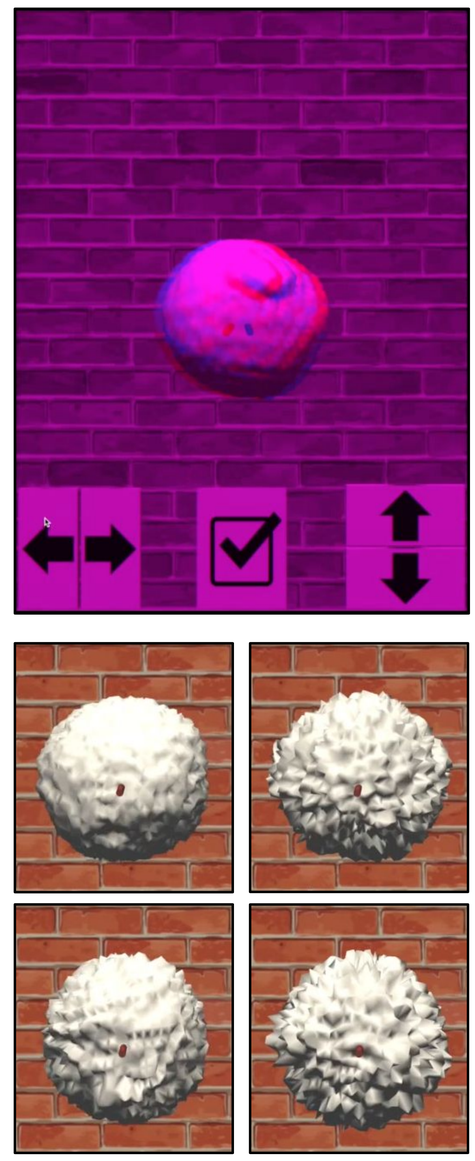

Figure 1: *Test application and the four levels of noise.*

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]. 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. 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. There were 20 random trials for each participant (1 view mode × 4 difficulty levels × 5 repetitions, giving 20 trials). When the participant completed the task, the researcher immediately asked the participant to answer the Simulation Sickness Questionnaire (SSQ) [KLBL93], followed by the Presence Questionnaire (PQ) [WS98], and compare the two view modes and give his or her preference for each view mode.

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.

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**◊ (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◊ (0.62)	74.43** (40.83)	74.80 (11.67)

*, ⋆ and ◊ indicate statistically significant differences between two means in the same table column.

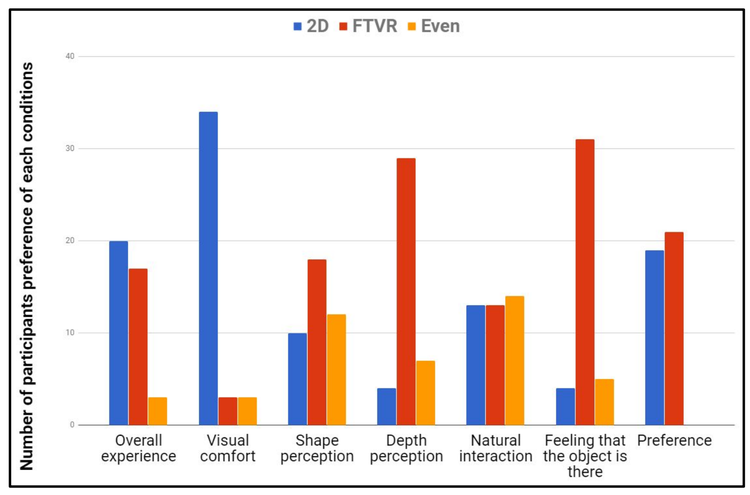

Figure 2: *Users' preference between the Normal 2D and the Combined view modes along the seven dimensions.*

Discussion & Conclusion: The first question is How effective is tablet FTVR? 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. 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.

Reference:

- [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