A DISPLAY SPECIFICATIONS

Table 2. Varjo Aero: relevant specifications.

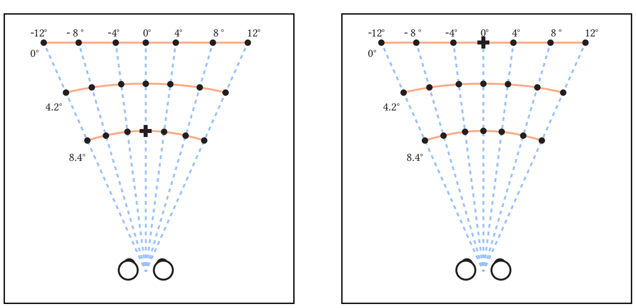
Resolution	Frequency	Peak Luminance
2880 × 2720	90 Hz	150 cd/m^2
Focal Distance	**FoV**	**Supported IPD**
0.85 m	134° (diagonal)	59 – 71 mm
Eye Tracker	**Frequency**	**Accuracy**
	200 Hz	< 1°

B PSYCHOPHYSICAL STUDY CONDITIONS

Calibration of maximum vergence amplitudes. The closest depth at which majority of user study participants could fuse a stereo image in VR was approximately $d_{min} = 0.4$ m. Depth, d, and vergence angle coordinates, α_v, have an inversely proportional relationship,

$$\alpha_v = \arctan\left(\frac{w_{IPD}}{2d}\right), \tag{9}$$

which varies from person to person depending on their IPD, w_{IPD}. This relationship, and the fact that there are no negative vergence angle coordinates, effectively limits the range of vergence gaze movement amplitudes, $\Delta\alpha_v$, a user study participant can make. Crucially, since the IPD, w_{IPD}, of participants varied, and we couldn't foresee the IPDs of all future user study participants, we could not determine the maximum vergence angle coordinate, α_v^{max}, by applying Equation (9) naively. Therefore, to ensure consistency across different participants, we selected the most conservative value of maximum vergence angle coordinates by minimizing Equation (9) under the constraints of $d > d_{min} = 0.4$ m, and $w_{IPD} > w_{IPD}^{min} = 59$ mm — the minimum IPD supported by the HMD. Then, applying these edge conditions to Equation (9), we get our maximum vergence angle coordinate of $\alpha_v^{max} = 8.4°$.



(a) divergent (b) convergent

Fig. 11. *Study conditions.* All visualized conditions originate at a + sign (near for divergent, far for convergent conditions), and target · signs. Leftward and rightward saccades are treated as equivalent in data analysis, but there are equal number of leftward and rightward conditions implemented.

Implementation of Study Conditions. We construct three isovergence circles for each $\alpha_v^{init}+\Delta\alpha_v$, starting with the smallest. As established earlier, this circle must be at least d_{min} away from the observer. Therefore we pick the first isovergence circle to be $d^{(0)} = d_{min}$ away, which corresponds to a vergence angle coordinate equal to

$$\alpha_v^{(0)} = \arctan\left(\frac{w_{IPD}}{2d^{(0)}}\right). \tag{10}$$

The following circles are constructed by adding the $\Delta\alpha_v$ to $\alpha_v^{(0)}$:

$$\alpha_v^{(i)} = \alpha_v^{(0)} + \Delta\alpha_v^{(i-1)}, \text{ for } i \in \{1, 2\}, \tag{11}$$

where $\Delta\alpha_v^{(i-1)}$ is the $i-1$th condition among vergence conditions.

Equipped with the isovergence circles with angles $\{\alpha_v^{(i)}\}$ for $i \in \{0, 1, 2\}$, we can select the initial fixation point for all divergent and convergent gaze motions to be at coordinates

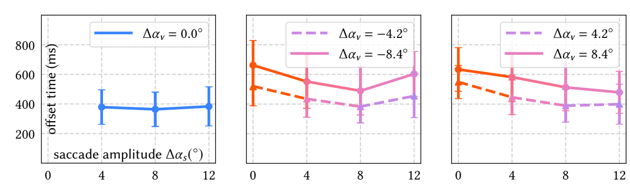
$$\begin{aligned}(\alpha_v^{init,\ div}, \alpha_s^{init,\ div}) &= (\alpha_v^{(0)}, 0°)\\ (\alpha_v^{init,\ conv}, \alpha_s^{init,\ conv}) &= (\alpha_v^{(2)}, 0°),\end{aligned} \tag{12}$$

respectively. Originating from a given fixation point, the rest of the condition locations are found as

$$(\alpha_v, \alpha_s) = (\alpha_v^{init} + \Delta\alpha_v, \alpha_s^{init} + \Delta\alpha_s), \tag{13}$$

where $\Delta\alpha_v$ and $\Delta\alpha_s$ correspond to the specific experimental condition of interest. The resulting grid of conditions are visualized in Figure 11.

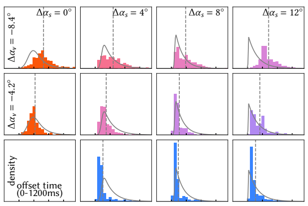
C EXPERIMENT RESULTS



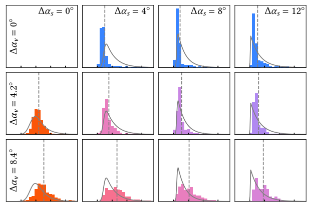
(a) saccade (b) divergent (c) convergent

Fig. 12. *Aggregated mean offset time of studied conditions across all participants with error bars.* This is a version of Figure 4 with std error bars as a more detailed visualization. See Figure 4 for further details.

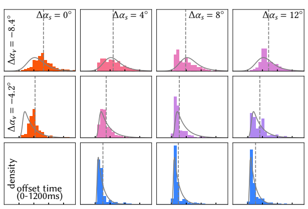
D ABLATION STUDY HISTOGRAMS



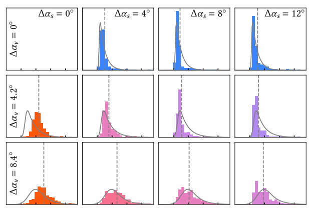
(a) **SAC**, divergent



(b) **SAC**, convergent



(c) **VER**, divergent



(d) **VER**, convergent

Fig. 13. *Histograms vs. predicted distributions of ablation models.* Predicted distributions by the ablation models are compared to measured data from psychophysical study. Ablation model **SAC** was trained using only saccade amplitude information from the study data, while **VER** only used vergence amplitude information. Since either model does not have full information that distinguishes individual conditions within a single column and row respectively, the models make the same predictions across multiple conditions within this histogram visualization. Thus, in (a)/(b) the model makes the same prediction within the same columns, while in (c)/(d) the model makes the same predictions within the same rows.

E FULL STATISTICAL VISUALIZATION OF USER STUDY

user	D	p
U_1	0.2	1.0
U_2	0.3	.79
U_3	0.2	1.0
U_4	0.1	1.0
U_5	0.2	1.0
U_6	0.1	1.0
U_7	0.1	1.0
U_8	0.2	1.0
U_9	0.1	1.0
U_{10}	0.2	1.0
U_{11}	0.2	1.0
U_{12}	0.2	1.0

Fig. 14. *Visualization of all participants, conditions and scenes of the user study Section 4.3.* X-axis indicates time (0-1000ms). Y-axis indicates density. Each color of the stacked bars indicates each condition: blue/red/green represents $\mathbf{C_s}: \Delta\alpha_v = 0°$ / $\mathbf{C_m}: \Delta\alpha_v = 7°$ / $\mathbf{C_l}: \Delta\alpha_v = 10.5°$. The inset numbers are the corresponding K.S. test results for each user across all conditions and scenes. Note that the discrepancy between eye travel distance ($\mathbf{C_s} < \mathbf{C_m} < \mathbf{C_l}$) and landing times ($\mathbf{C_m} < \mathbf{C_l} < \mathbf{C_s}$) share statistical significance across individuals.