https://doi.org/10.2352/ISSN.2470-1173.2019.3.SDA-639

Evaluating the effect of stereoscopic display crosstalk on simulated remote vision system depth discrimination

Eleanor O'Keefe[1], Charles Lloyd[2], Tommy Bullock[3], Alex Van Atta[1], and Marc Winterbottom[3]
[1]KBRwyle, Beavercreek, OH, USA
[2]Visual Performance, LLC, Ellisville, MO, USA
[3]U.S. Air Force School of Aerospace Medicine, OBVA Laboratory, Wright-Patterson AFB, OH, USA

Abstract

As the use of virtual and augmented reality increases, it is important to understand how these technologies affect user performance. The introduction of stereoscopic remote vision system (RVS) technology in air refueling tankers means that the performance and level of visual fatigue of aircrew using stereoscopic displays are important operational factors to consider. Crosstalk occurs due to incomplete separation of the two images projected to the two eyes in a stereoscopic display and can degrade depth perception and cause discomfort and fatigue. A substantial amount of previous research has described measurement of crosstalk, compared crosstalk for different display technologies, and examined the effect of crosstalk on viewing discomfort. Additional research has examined the effects of crosstalk on stereoscopic image quality and on the magnitude of perceived depth from disparity and from monocular occlusions. The research described in this report shows that stereoscopic display crosstalk can substantially degrade depth discrimination under viewing conditions simulating a hyperstereoscopic RVS and could clearly be a performance limitation for tasks requiring accurate depth perception such as air refueling.

Introduction

The use of remote vision system (RVS) technology has increased in both industrial and military operations including aerial refueling, telesurgery, and bomb deactivation and disposal. While these stereoscopic displays enable three-dimensional (3D) perception on a two-dimensional screen, there are artifacts present in these displays that can limit task performance. Crosstalk is one such artifact, which occurs due to incomplete separation of the images projected to the two eyes. Perceptually this results in the appearance of a double image, or ghosting (see Figure 1) [1]. While most stereoscopic displays have a low level of crosstalk at the design eye-point (usually the center of the screen), crosstalk can increase near screen edges, with head-roll and with viewing angle. Therefore, not only is it important to use a display that has low crosstalk levels, it is important to consider where the user will be viewing stimuli on the display and if the user could be outside of the design eye-point, as both can cause an increase in crosstalk, potentially resulting in degraded task performance.

The results of a number of published studies demonstrate that increasing levels of stereoscopic display crosstalk degrade performance along several dimensions. These include discomfort and fatigue, image quality, and estimation of depth magnitude. User ratings of discomfort and fatigue increase as crosstalk levels are increased [2-5]. Several authors have recommended that crosstalk should be below about 5 to 10%. For example, Kooi and Toet [2] recommended that crosstalk levels remain under 5%. Nojiri et al. found a tolerance limit of 5-10%, and Chen et al. recommend a crosstalk level of less than 5.8%. Other research [4] has shown that while participants were able to make correct depth alignment judgments at up to 5% crosstalk, workload at these levels increased by 14% according to the well-accepted NASA Task Load Index (TLX) [6].

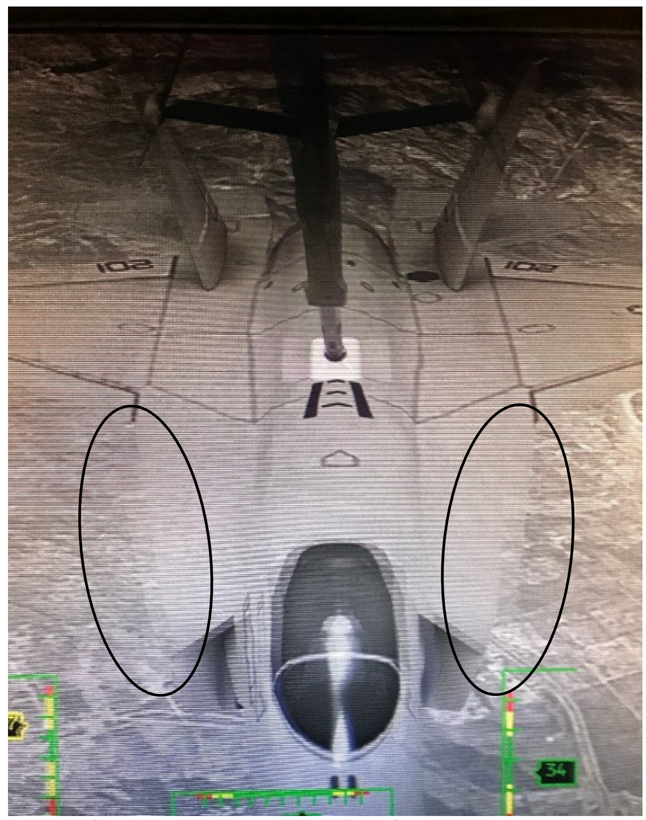


Figure 1. Crosstalk caused by leakage of light from one eye's view into the other results in a double image, or ghosting, which is particularly noticeable in the circled regions shown here.

Ratings of stereoscopic image quality decrease with increasing crosstalk levels, although the recommendations are more variable [7-9]. At the low end, it is recommended there be no more than 2-5% crosstalk to avoid a decrease in image distortion [9]. However, Huang et al. [7] recommend less than 10% system crosstalk (crosstalk divided by co-location image contrast) if no perceptible degradation in image quality is desired, and Wilcox and Stewart [8] found that crosstalk levels of 10% were reached before any significant reduction of image quality ratings was found.

Depth magnitude estimations are also reduced with increasing crosstalk levels, [4,10-12]. In Pala et al. [4], depth magnitude estimation accuracy dropped 8% at 5% crosstalk. Tsirlin and her

Cleared, 88PA, Case # 2019-0280, 23 Jan 2019.

colleagues, who tested both single stimuli and natural scene stimuli, recommend crosstalk levels of 4% or lower [10-12]. These authors also found that crosstalk affects thin stimuli at smaller disparities than thick stimuli, and that crosstalk not only affects depth magnitude estimations made from binocular disparity cues but also from monocular cues. While these authors [4,10-12] studied the effects of crosstalk on depth magnitude estimation, most research concerning crosstalk appears to focus on perceived image quality and discomfort. We could not find additional papers describing evaluations of the effect of crosstalk on depth discrimination performance specifically.

The ability to reliably discriminate fine differences in depth is of great importance to RVS users, including aerial refueling operators (AROs). During refueling operations, just prior to making contact, AROs must accurately discriminate the distance between the boom nozzle and the surface of the receiver aircraft in order to make contact with the receptacle without damaging the receiver. Accurate depth discrimination is critical to ensure that contacts do not take an excessive amount of time (dangerous for aircraft low on fuel) or lead to contacts made with excessive force. Given the speeds and mass of the refueling aircraft, incorrect judgment of distance could lead to a serious mishap.

For an air refueling task, depth discrimination may be more important than accurate depth estimation, scaling, or magnitude estimation [13]. That is, discriminating depth between the nozzle and receptacle is probably more important than judging the absolute distances. Furthermore, due to the way the RVS is designed, the AROs cannot make use of other cues to depth (e.g., known size, motion parallax); thus, depth from disparity may be critically important.

The purpose of this evaluation was to quantify the expected increase in depth discrimination threshold produced by increasing crosstalk, using a hyperstereoscopic display. The level of crosstalk produced by the display was controlled by changing the vertical position of the display relative to the participant's eyes, which emulated what happens when users move their heads above or below the design eye-point for a passive polarizing stereoscopic display.

METHOD

Participants

Eleven people (eight male, three female) participated in this evaluation, with ages ranging from 22 to 60 years (median age 33 years). Participants were recruited from a study pool comprised of volunteers from the Wright-Patterson Air Force Base (Dayton, Ohio) area. All participants provided written informed consent before participating. The study was conducted in accordance with the Wright-Patterson Air Force Research Laboratory Institutional Review Board protocol number FWR20130074H.

Equipment and Virtual Camera Configuration

A ViewSonic V3D231, model VS14136 passive stereoscopic flat panel display (ViewSonic, CA, USA), viewed through passive circularly polarized 3D glasses was used in this study. The peak white luminance of this display was 172 cd/m^2 (50 fL). The peak luminance of the display with the terrain background used in the experiment was 101.7 cd/m^2 (30 fL) and 41.7 cd/m^2 (12 fL) when viewed through the 3D glasses.

The active area of this 23-inch display was 51 x 32 cm (20 x 11.25 in). This display has a total of 1920 x 1080 pixels and is spatially interlaced with micro-polarizers and glasses blocking every other vertical line. Thus, each eye sees 1920 x 540 pixels when viewed through the glasses.

When viewed from a distance of 0.84 m (33 in), the horizontal field of view (FOV) of the display was 33.7 deg and the vertical FOV of the display was 19.3 deg. For this viewing distance, the spatial frequency of the high contrast horizontal lines that could be seen in each eye was 26 cycles/deg. The high spatial frequency cutoff for human vision is 50-60 cycles/deg; thus, the black, interlaced horizontal lines were clearly visible.

The 3D virtual hyperstereoscopic environment was generated using Laminar Research's X-Plane image generation software. The experimental task stimuli were inserted as billboard models in X-Plane above a virtual terrain database. The two instances of X-Plane were run on two separate PCs, each incorporating Intel i7 processors and NVidia Quadro K4200 video cards with Quadro Sync. Two Black Magic DVI Extenders and an AJA Video Multiplexer were used to combine the top-and-bottom input of the two video channels into an interlaced stereoscopic image at 24 fps.

The virtual cameras within the image generators were configured such that they were spaced much wider than the average interpupillary distance (i.e., hyperstereoscopic). Additionally, the simulated RVS cameras were configured such that geometric distortions were introduced into the 3D image [14,15].

The motorized adjustable height table used for setting the relative height of the display was a NextDesk Up^{TM} (XDesk, TX, USA), which included a digital display of the height setting to an accuracy of 0.1 in.

A Cedrus RB-730 (Cedrus, CA, USA) response box was used to collect participant responses for the depth discrimination task.

Single point luminance measurements were acquired with a Minolta CS-200 chromameter (Konica Minolta, NJ, USA) with a 1-deg spot size.

Experimental Task

In this evaluation, each participant completed a series of approximately 30 depth discrimination threshold measurements for table heights ranging from 30 to 37 in, which produced a wide range of crosstalk levels. The order of the crosstalk levels used was randomized for each participant to distribute any time-dependent effects (e.g., practice and fatigue) without biasing the effect of crosstalk level.

Each depth discrimination threshold was measured using a two-alternative psychophysical test similar to the "dual ring" stereo acuity test in the Operational Based Vision Assessment (OBVA) Lab's Automated Vision Test (AVT) Battery [17]. The dual ring test has been shown to correlate with stereoscopic operational-based tasks [18,19]. On each of the 35 trials per threshold, a pair of concentric rings was presented on screen. The outer ring was always presented at the same depth and so used as a reference ring. The inner ring was set at a distance randomly either nearer to or farther from the participant than the outer reference ring as shown in Figure 2. Participants responded on each trial using either a "near" or a "far" button on the response box. The PSI method [20] was implemented to estimate depth thresholds. Participants completed one practice block at one table height before performing the full experiment. Auditory feedback indicating both correct and incorrect responses was provided during the practice block. Instructions, practice, and the 30 threshold measurements required approximately 1.5 hours per participant. Participants were invited to rest between each threshold measurement as needed.

Cleared, 88PA, Case # 2019-0280, 23 Jan 2019.

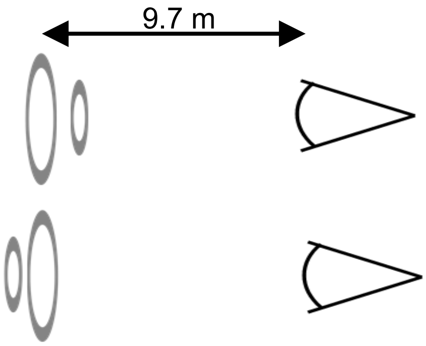


Figure 2. Illustration of the appearance of the dual-ring stimuli in 3D to during the depth discrimination task.

The rings always appeared above the horizon as shown in Figure 3. The contrast of the rings against the sky background was 0.9 (contrast ratio = 1.9). The outer ring distance was fixed at the camera convergence distance; thus, the right and left eye images of the outer ring always coincided at the screen and were never seen as a double image in the presence of crosstalk. Significant crosstalk produced a double image of only the inner ring.

The diameter of the outer edge of the outer ring was 0.508 m (20 in) in object space, which subtended 2.97 deg from the camera point of view. The average gap between the inner and outer rings was 0.24 deg from the participant point of view.

The background image was an aerial view over the Seattle-Tacoma International Airport area. The average luminance of the scene was 91.7 cd/m^2 (27 fL) without glasses and 38 cd/m^2 (11 fL) with glasses.

Figure 3. Example of the dual ring stimuli shown against the background used in the evaluation.

During the evaluation, the height (38 in) and distance (33 in) of the participant's eyes were fixed using a head/chin rest clamped to a small table as shown in Figure 4. The relative height of the display was changed using an adjustable height table with a digital display of the height setting. This provided more precise control over observer viewing position, and therefore crosstalk, than could be achieved by adjusting the position of the observer's head.

Figure 4. Depth discrimination experiment apparatus. Participant eye height remained fixed using a head/chin rest attached to the small table in the foreground. The height of the monitor relative to the eyes was controlled using the adjustable height table supporting the display.

Crosstalk Measurements

Crosstalk was measured as a function of the height and distance of the user's eyes relative to the display. In total, 125 measurements were made for each eye spanning 5 viewing distances (24 to 40 in), 5 vertical positions on the display surface (top to bottom), and 5 vertical positions of the photometer (-5 to + 5 in) relative to the surface normal of the display. Three luminance measurements were required for each crosstalk measurement; thus, a total of 750 luminance measurements were made per eye using the CS-200 chromameter with a 1-deg spot size. The apparatus used for measuring crosstalk is shown in Figure 5.

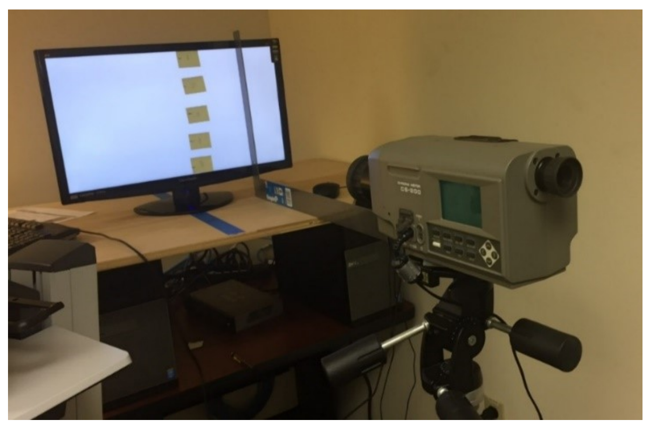


Figure 5. Apparatus used for measuring crosstalk on the ViewSonic display. Measurements were made with either the left or right lens of the polarized 3D glasses positioned in front of the chromameter lens.

Cleared, 88PA, Case # 2019-0280, 23 Jan 2019.

Luminance measurements were taken using three separate test patterns: (1) a white image presented to the right eye and a black image presented to the left eye, (2) a black image presented to the right eye and a white image presented to the left eye, and (3) a black image presented to both eyes. The right or left lens of the polarized glasses was placed over the Minolta CS-200 chromameter lens during measurements for each specific eye view. Crosstalk was defined and calculated as described in previous research [21] using the following equations:

$$OCT_{RL} = \frac{O_{GL}-O_{BL}}{O_{WL}-O_{BL}} \quad \text{and} \quad OCT_{LR} = \frac{O_{GR}-O_{BR}}{O_{WR}-O_{BR}} \qquad (1,2)$$

where:

OCT_{RL} = the observed crosstalk measured in the left eye view coming from the pattern shown to the right eye

O_{GL} = the observed luminance measured through the left eye view when the black image is shown to the left eye and the white image is shown to the right eye

O_{BL} = the observed luminance measured through the left eye view when the black image is shown to both eyes

O_{WL} = the observed luminance measured through the left eye view when the white image is shown to the left eye and the black image is shown to the right eye

The same definitions are true for the equation on the right, although switched with regard to what image is shown to each eye and measured through the right eye view. Total crosstalk was calculated by taking an average of the crosstalk measured from left and right eye views.

These measurement data were used to create a multi-dimensional mathematical model that summarized crosstalk as a function of viewing distance, vertical position on the screen, and vertical position of the eyes relative to the display surface normal. This model was fit to 96 of the 125 measurements for which the measured crosstalk was no greater than 50% ($R^2 = 0.990$, $p < 1e^{-95}$, N = 96). The root-mean-square error (RMSE) was 1.16% across all 96 measurements and was 0.56% for the data covering the range of 0 to 10% crosstalk.

The form of this 3D crosstalk model is illustrated in Figures 6 and 7. In Figure 6, the vertical head position was fixed at 1.9 in above the display surface normal and crosstalk is plotted as a function of vertical position on the display and viewing distance. This surface shows that for this display crosstalk is minimized over the vertical extent of the display when the viewing distance is 24 in. With increasing viewing distance, crosstalk increases at the top and bottom of the screen. In Figure 7, the viewing distance was fixed at 33 in and crosstalk is plotted as a function of vertical head position and vertical position on the display. This surface reveals that for the 33 in viewing distance crosstalk is minimized when the observer's eyes are about 1.9 in above the surface normal of the display.

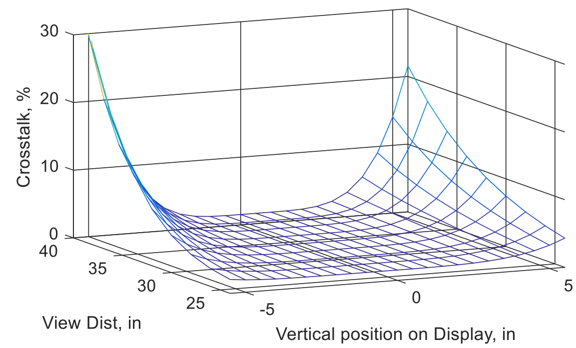


Figure 6. Crosstalk plotted as a function of vertical position on the display and viewing distance with the vertical head position fixed at 1.9 in above the surface normal of the display.

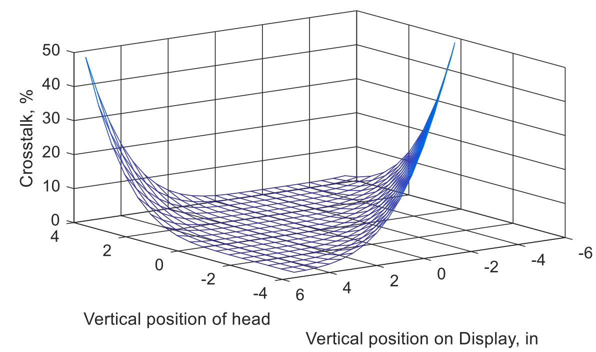


Figure 7. Crosstalk plotted as a function of vertical position on the display and vertical head position with viewing distance fixed at 33 in.

For the present evaluation, a 3D model of crosstalk is not required because the viewing distance was fixed at 33 in and stimuli were presented only at the center of the display. With these two constraints applied, the 3D model collapses to a one-dimensional model of crosstalk as a function of the height of the display relative to the eyes. To ensure crosstalk was accurately characterized, crosstalk was re-measured at 10 table heights for the fixed viewing distance of 33 in and at the center of the display only. These points fit the model, three of which are shown in Figure 8, which plots the table heights used in this study. Therefore, this model was used in the present evaluation to determine the magnitude of crosstalk produced as a function of the adjustable height of the table on which the display was mounted.

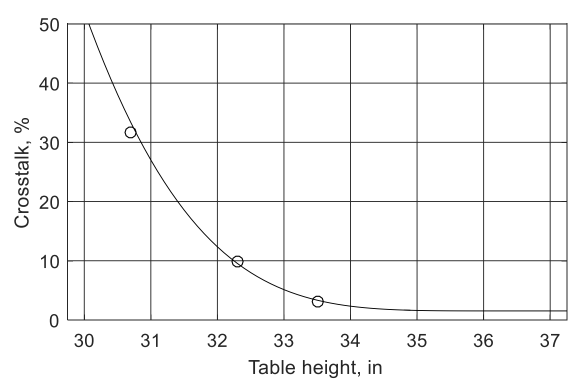


Figure 8. Crosstalk at the center of the display as a function of table height with viewing distance fixed at 33 in.

Cleared, 88PA, Case # 2019-0280, 23 Jan 2019.

RESULTS

Approximately 30 discrimination thresholds were measured for each of 11 participants. There were differences in the mean threshold across all table heights for individual participants, ranging from a low of 0.46 cm to a high of 5.0 cm. The mean threshold for one of the participants was 25 times higher than the other 10 participants, thus, these data were excluded from further analyses.

For each table height, depth discrimination threshold was measured up to three times and the median of the repeated measurements was used for subsequent analyses. The effect of participant was removed by subtracting the threshold at each table height by the mean threshold at all table heights for each individual participant. The mean threshold at all table heights across all participants was then added to these values. These data are plotted in Figure 9. For those conditions with low thresholds (table heights ≥ 34 in), the grand mean threshold across the 10 participants was 1.26 cm.

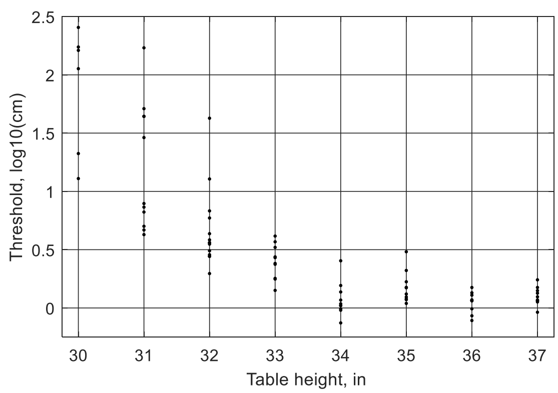


Figure 9. Threshold measurements for 10 participants plotted as a function of table height.

To more directly show the effect of crosstalk on depth discrimination, the table heights were converted to crosstalk level using the calibration curve shown in Figure 8. The results are plotted in Figure 10, which shows the means (points) and 95% confidence intervals (diamonds) for these data. A regression model fit to these data produced an $R^2 = 0.77$ ($p < 1e^{-26}$, N = 81) and is also shown in the figure:

$$\text{Thresh} = 10 \wedge [0.0544 + 0.5828 * \log10(\text{Cross})^2] \quad (3)$$

where Thresh is the depth discrimination threshold measured in cm and Cross is the stereoscopic crosstalk in percent.

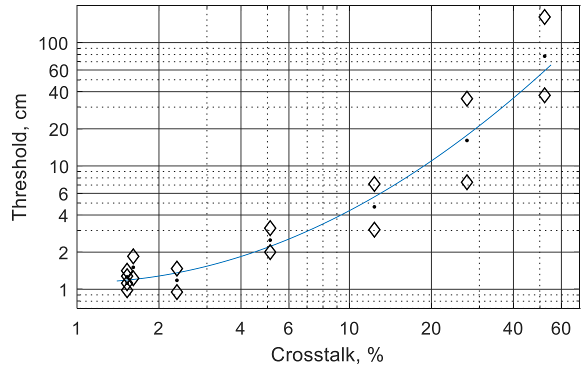


Figure 10. Mean depth discrimination thresholds (points) and 95% confidence intervals (diamonds) as a function of crosstalk percentage.

While there was no significant difference between the four lowest depth discrimination thresholds where crosstalk is 2.5% or less, a two-sample t-test between the thresholds at 2.5% and the 5% and above crosstalk showed a significant difference between depth discrimination thresholds ($p = 1e^{-4}$). In addition, a two-sample t-test between the four lowest thresholds and the threshold at 5% crosstalk also showed a significant difference ($p = 2.5e^{-8}$).

The pooled standard deviation of the $\log_{10}$ threshold for the crosstalk levels below 2.5% is 0.121. Using this value and the regression model, thresholds are reliably increased (95% confidence) from the floor of this function when crosstalk levels are 4.5%. Here, the depth discrimination threshold is 2.00 cm, which is 59% higher than the floor of the function of 1.26 cm. The average depth discrimination threshold was doubled when crosstalk increased to 6% and was 5 times greater when crosstalk was increased to 14%.

CONCLUSION AND RECOMMENDATIONS

Depth discrimination thresholds were significantly elevated with increasing crosstalk. The lowest amount of crosstalk that caused a significant increase in depth discrimination was 4.5%. Therefore, it is recommended that the maximum crosstalk across a display used for stereoscopic viewing be no greater than 4.5%, as exceeding this level reliably increases depth discrimination thresholds.

This level is in close agreement with previous research on crosstalk and depth magnitude estimation [11-13], which recommends crosstalk no higher than 4%. It is also consistent with the studies that found significant increases in discomfort and fatigue around crosstalk levels of 5% [2-5].

Suppliers and users of 3D displays should take steps to ensure that crosstalk does not exceed approximately 4.5% to maintain performance and comfort, particularly for applications such as air refueling and telesurgery where highly accurate depth discrimination is absolutely critical. For stereoscopic display systems where viewing position may vary, steps should be taken to ensure that users do not adopt a viewing position that increases crosstalk. This may be challenging for stereoscopic systems that must accommodate users of widely varying height, such as ARO's using a remote vision system. Finally, maintainers of stereoscopic display systems should take steps to ensure that crosstalk does not exceed approximately 4.5% as the system ages.

Cleared, 88PA, Case # 2019-0280, 23 Jan 2019.

ACKNOWLEDGMENTS

The authors would like to thank the Air Force Life Cycle Management Center (AFLCMC) for their support of this research. The views expressed are those of the authors and do not necessarily reflect the official policy or position of the Air Force, the Department of Defense, or the U.S. Government. This document has been approved for public release (Case # 88ABW-2019-0280, 23 Jan 2019). This work was supported by U.S. Air Force contract FA8650-12-D-6280 to Wyle Laboratories and was funded by the 711th Human Performance Wing/U.S. Air Force School of Aerospace Medicine and the Air Force Life Cycle Management Center.

REFERENCES

[1] A. Woods. "Crosstalk in stereoscopic displays: a review," Journal of Electronic imaging 21(4), 04092, 2012.

[2] F. Kooi & A. Toet. "Visual comfort of binocular and 3D displays," Displays vol. 25, pp. 99–108, 2004.

[3] Y. Nojiri, H. Yamanoue, A. Hanazato, M. Emoto, & F. Okano. "Visual comfort/discomfort and visual fatigue caused by stereoscopic HDTV viewing," in Proc. SPIE: Stereoscopic Displays and Virtual Reality Systems XI, 2004.

[4] S. Pala, R. Stevens, P. Surman, & A. Beaumont. "Optical crosstalk and visual comfort of a stereoscopic display used in a real-time application," in Proc. SPIE: Stereoscopic Displays and Virtual Reality Systems XIV, 2007.

[5] L. Chen, Tu, Y., Liu, W., & Li, Q. "Investigation of crosstalk in a 2-view 3D display," Soc. Inform. Display Symp. Digest Tech. Papers 39(1), 1138–1141, 2008.

[6] S. G. Hart & L. E. Staveland. (1988) Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock and N. Meshkati (Eds.) Human Mental Workload. Amsterdam: North Holland Press.

[7] K. Huang, J. Yuan, C. Tsai, W. Hsuah, & N. Wang. "A study of how crosstalk affects stereopsis in stereoscopic displays," in Proc. SPIE: Stereoscopic Displays and Virtual Reality Systems X, 2003.

[8] L. Wilcox & J. Stewart. "Determinants of perceived image quality: Ghosting versus brightness," in Stereoscopic Displays and Virtual Reality Systems: Proc. SPIE, 2003.

[9] P. Seuntiëns, L. Meesters, & W. Ijsselsteijn. "Perceptual attributes of crosstalk in 3D images," Displays, vol. 26, pp. 177–183, 2005.

[10] I. Tsirlin, L. M. Wilcox, & R. S. Allison. "The effect of crosstalk on depth magnitude in thin structures," in Stereoscopic Displays and Applications XXII: International Society for Optics and Photonics, 2011a.

[11] I. Tsirlin, L. Wilcox, & R. Allison. "The effect of crosstalk on the perceived depth from disparity and monocular occlusions," IEEE Trans. Broadcast, vol. 57, no. 2, pp. 445–453, 2011b.

[12] I. Tsirlin, R. Allison, & L. Wilcox. "Crosstalk reduces the amount of depth seen in 3D images of natural scenes," in Stereoscopic Displays and Applications XXIII: Proc. SPIE, 2012.

[13] C. J. Lloyd & S. G. Nigus. "Effects of stereopsis, collimation, and head tracking on air refueling boom operator performance," in Proc. of the Image Society Annual Conference, Scottsdale, Arizona, 2012.

[14] D. B. Diner & M. VonSydow. "Stereo depth distortions in teleoperation," 1988.

[15] A. J. Woods, T. Docherty, & R. Koch. "Image distortions in stereoscopic video systems," in Stereoscopic displays and applications IV: International Society for Optics and Photonics, 1993.

[16] C. J. Lloyd, M. Winterbottom, E. O'Keefe, & T. Bullock. "Predicting depth discrimination performance under hyper-stereoscopic display condition," in International Society for Optics and Photonics: Degraded Environments: Sensing, Processing, and Display, 2018.

[17] J. Gaska, M. Winterbottom, E. Shoda, E. O'Keefe, & A. Van Atta. "Operational based vision assessment: automated vision test research and reliability. Wright-Patterson AFB (OH): U.S. Air Force School of Aerospace Medicine; 2018. Technical Report AFRL-SA-WP-TR-2018-0019.

[18] M. Winterbottom, C. Lloyd, J. Gaska, L. Williams, E. Shoda, & S. Hadley. "Investigating aircrew depth perception standards using a stereoscopic simulation environment," in Electronic Imaging: Stereoscopic Displays and Applications XXVIII, 2017.

[19] M. Winterbottom, C. Lloyd, J. Gaska, L. Williams, E. Shoda, & S. Hadley. "Investigating the importance of stereo displays for helicopter landing simulation," in IMAGE Society Annual Conference Proceedings, 2016.

[20] N. Prins & F. Kingdom. (2018). Applying the model-comparison approach to test specific research hypotheses in psychophysical research using the Palamedes toolbox. Frontiers in Psychology, 9.

[21] M. Weissman & A. Woods. "A simple method for measuring crosstalk in stereoscopic displays," in Proc. SPIE: Stereoscopic Displays and Applications XXII, 2011.

Cleared, 88PA, Case # 2019-0280, 23 Jan 2019.

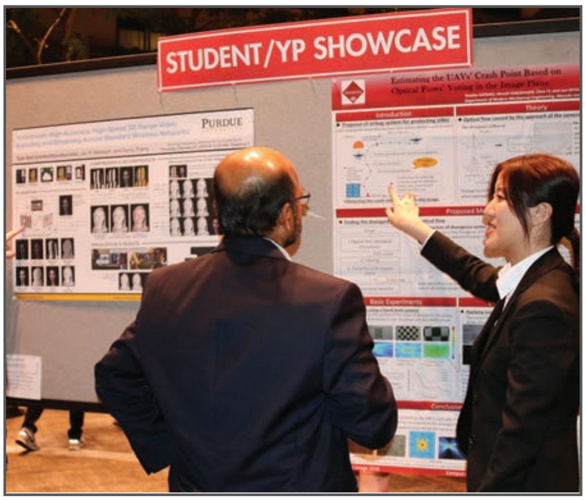
JOIN US AT THE NEXT EI!

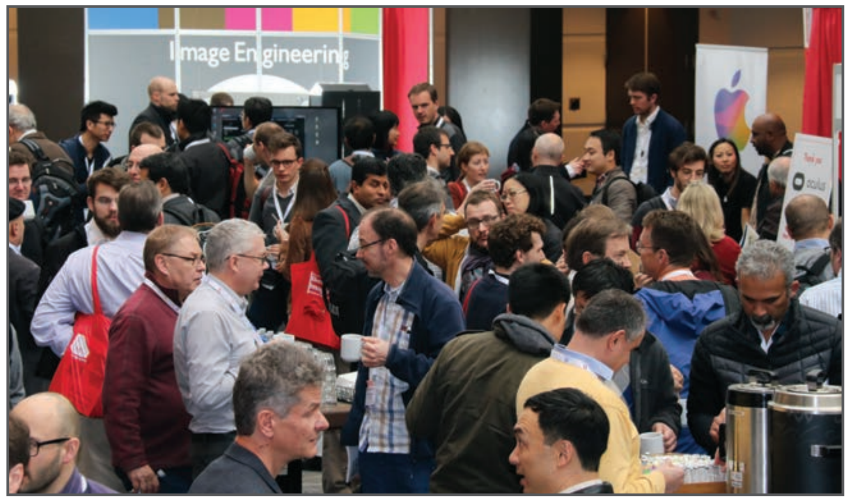
IS&T International Symposium on

Electronic Imaging

SCIENCE AND TECHNOLOGY

Imaging across applications . . . Where industry and academia meet!





• SHORT COURSES • EXHIBITS • DEMONSTRATION SESSION • PLENARY TALKS •
• INTERACTIVE PAPER SESSION • SPECIAL EVENTS • TECHNICAL SESSIONS •

www.electronicimaging.org

IS&T
imaging.org