Crosstalk reduces the amount of depth seen in 3D images of natural scenes

Inna Tsirlin*, Robert S. Allison and Laurie M. Wilcox
Centre for Vision Research, York University, 4700 Keele st., Toronto, ON, Canada

ABSTRACT

Crosstalk remains an important determinant of stereoscopic 3D (S3D) image quality. Defined as the leakage of one eye's image into the image of the other eye it affects all commercially available stereoscopic viewing systems. Previously we have shown that crosstalk affects perceived depth magnitude in S3D displays. We found that perceived depth between two lines separated in depth decreased as crosstalk increased. The experiments described here extend our previous work to complex images of natural scenes. We controlled crosstalk levels by simulating them in images presented on a zero-crosstalk mirror stereoscope display. The observers were asked to estimate the amount of stereoscopic depth between pairs of objects in stereo-photographs of cluttered rooms. Data show that as crosstalk increased perceived depth decreased; an effect found at all disparities. Similarly to our previous experiments a significant decrease in perceived depth was observed with as little as 2-4% crosstalk. Taken together these results demonstrate that our previous findings generalize to natural scenes and show that crosstalk reduces perceived depth magnitude even in natural scenes with pictorial depth cues.

Keywords: Crosstalk, ghosting, 3D displays, perceived depth, psychophysics, depth magnitude, natural scenes

1. INTRODUCTION

Crosstalk is generally defined as the leakage of one eye's image into the image of the other eye. It is present in all commercially available stereoscopic viewing systems including polarized, time-sequential and autostereoscopic displays [1]. It is well recognized that crosstalk decreases perceived image quality and causes image distortion [2, 3]. Moreover, visual comfort decreases and perceived workload increases with increasing crosstalk [4-6]. In previous experiments we have shown that crosstalk also affects perceived depth magnitude [7, 8]. We used two disparate white bars on a black background, and measured the perceived depth between the bars as a function of disparity and degree of crosstalk. We found that as crosstalk increased beyond 4-8% (depending on disparity) perceived depth decreased significantly. Moreover, the decrease in perceived depth was more drastic at larger disparities. Further, by manipulating the width of the white bars, we found that this detrimental effect was present regardless of whether the ghost image was spatially separated from, or overlapped with, the original image. Results of another experiment showed that depth from monocular occlusions was degraded by crosstalk even more than depth from disparity [7].

In the current series of experiments, we investigate whether the preceding findings can be generalized to complex images of natural scenes. The stimuli were color S3D photographs that showed cluttered scenes containing furniture and various objects. Observers were asked to estimate the amount of stereoscopic depth between pairs of objects in the scene as we manipulated the amount of crosstalk in the images. We used two different estimation methods- a virtual measurement scale and a disparity probe. Data show that, as was the case with simple line stimuli, depth in natural scenes was dramatically affected by crosstalk. As crosstalk increased perceived depth decreased; an effect that persisted across different disparities. Moreover, statistical analysis showed that, as was the case with synthetic stimuli, in natural scenes perceived depth is reduced significantly at as little as 2-4% crosstalk (depending on disparity).

---

* itsirlin@yorku.ca; phone 1 416 736-2100 x 70430; http://www.wilcoxlab.yorku.ca/~inna

Stereoscopic Displays and Applications XXIII, edited by Andrew J. Woods, Nicolas S. Holliman,
Gregg E. Favalora, Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 8288, 82880W
· doi: 10.1117/12.906751

2. METHODOLOGY

2.1 Observers

Nine observers, two authors (IT and LW) and seven volunteers (graduate and undergraduate students), participated in the study. All observers had normal or corrected-to-normal visual acuity and good stereoacuity in that they were able to discriminate disparities at least as small as 40 seconds of arc using the Randot Stereotest. The interocular distance for each observer was measured with a Richter digital pupil distance meter.

2.2 Apparatus

The stimuli were presented using the Psychtoolbox (v. 7.0.8) package for MATLAB (v. 7.4) executed on a G5 Power Macintosh. Stimuli were viewed on a pair of CRT monitors (ViewSonic G225f) arranged as a mirror stereoscope (see Figure 1). The viewing distance was 0.6 m, the resolution of the monitors was 1280x960 pixels and the refresh rate was 75Hz. With these settings each pixel subtended 1.77 arcmin. A chin rest was used to stabilize observers' head position during testing.

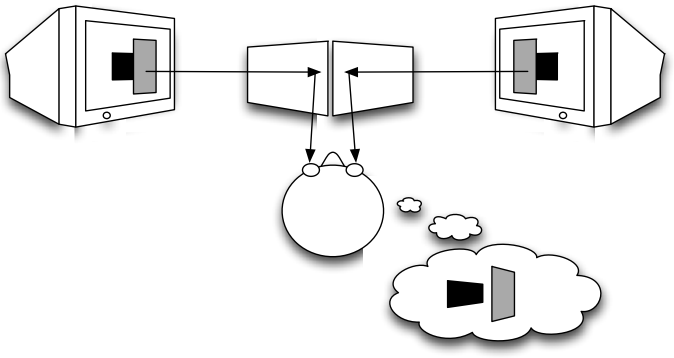


Figure 1 Mirror stereoscope. The left and right eyes' images were presented on two CRT displays. The images were then reflected from two mirrors to the observer's eyes. This type of spatial multiplexing provides crosstalk-free stereoscopic images.

2.3 Stimulus

The stimuli were two color stereoscopic photographs of natural scenes shown in Figure 2. One depicted a laboratory (Lab scene) and another a kitchen (Kitchen scene). Both photographs were taken with FUJIFILM FinePix REAL 3D W3 camera with a focal length of 6.3 mm (equivalent to 35mm on a 35mm still camera) and an inter-axial distance of 7.5 mm. The left and the right image planes of the 3D camera were parallel and the zero-parallax setting of the photographs was adjusted subsequently using horizontal image translation such that the scene contained mostly positive parallax (uncrossed disparity). The size of the Lab image was cropped to 850x478 px (25x14.1 deg) and the size of the Kitchen image to 816×638 px (24 x 18.8 deg).

In each of the photographs we selected four pairs of objects and measured the relative disparities between them. In the Lab scene the disparities were 5.3, 8.8, 12.4 and 17.7 arcmin (between green block and center of the level, green block and red car, edge of cardboard box and green-yellow block and edge of cardboard box and red car, respectively). In the kitchen scene the disparities were 10.6, 26.5, 56.6 and 65.5 arcmin (between pumpkin and cup, pumpkin and faucet, pumpkin and top drawer handle and pumpkin and toaster, respectively). We wanted to test a large range of disparities and hence selected different disparity ranges for the two images.

A vertical ruler with an adjustable cursor was positioned 88.5 arcmin to the left of the images (see Figure 4). The ruler was 17.7 deg long (23 cm) and the cursor was 7.08 arcmin wide. The cursor could be moved along the ruler using a computer mouse. The screen background was black and the ruler was light gray.

Simulated crosstalk levels were one of 0, 1, 2, 4, 8, 16 or 32%. Crosstalk was simulated by taking an attenuated version of one eye's image (intensities were multiplied by the current crosstalk value) and adding it to the other image. Since the photographs were captured with gamma encoding, to add linear amounts of crosstalk, we first linearized the images, added an attenuated version of one image to the other and then applied the display gamma function. The effect of crosstalk on the images is demonstrated in Figure 3.

Lab scene

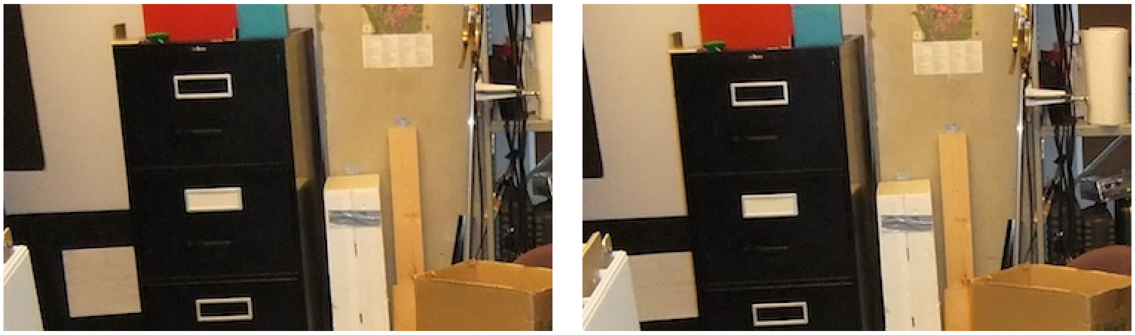
Kitchen scene

Figure 2 Stimuli used in the experiment. The top row shows the Lab scene and the bottom row shows the Kitchen scene. The stereo pairs are arranged for crossed fusion.

2.4 Procedure

The Lab scene and the Kitchen scene stimuli were tested in different sessions. Eight observers participated in the Lab scene sessions and six observers participated in the Kitchen scene sessions. Observers were asked to indicate the perceived depth between pairs of objects in the scenes using the cursor on the virtual ruler. On each trial observers were first shown a 2D image of the scene with a thick blue arrow connecting the two objects of interest. When they were ready, observers pressed a button on a gamepad and a 3D image of the scene appeared; the arrow was removed, and the virtual ruler was presented to the left of the image. They then adjusted the sliding cursor on the ruler to indicate the amount of depth they perceived between the two objects. All estimates were made relative to the base/bottom of the scale. The experimental procedure is illustrated in Figure 4. For each of the scenes, each of the 28 conditions (7 crosstalk levels x 4 disparities) was presented 10 times in random order in two sessions of 140 trials each. The experiment took place in a completely dark room and observers were free to move their eyes and had unlimited viewing time.

No crosstalk



20% crosstalk

Figure 3 Illustration of the effect of crosstalk. The top row shows a fragment of the Lab scene with 0% crosstalk. The bottom row shows the same fragment with 20% crosstalk.

Figure 4 Illustration of the experimental procedure. First screen shows the observers what relative depth they need to estimate on a 2D image (in this case the depth between the pumpkin and the toaster). After an observer presses a button, the 3D version of the image is shown along with a virtual ruler, which observers use to indicate the perceived depth magnitude.

3. RESULTS AND DISCUSSION

In the following analysis, angular disparities were converted to equivalent theoretical depth in centimeters to simplify the comparison with estimated perceived depth. We used a standard formula, which relates the disparity on the screen to theoretical depth at a known viewing distance. In the conversion we used the average inter-ocular distance of our observers (6.0 cm for Lab scene and 6.15 cm for the Kitchen scene). The predicted relative depth between the pairs of objects in the Lab scene were 0.99, 1.67, 2.37 and 3.26 cm (corresponding to 5.3, 8.8, 12.4 and 17.7 arcmin) and in the Kitchen scene 1.93, 5.08, 11.99 and 13.31 cm (corresponding to 10.6, 26.5, 56.6 and 65.5 arcmin).

Figure 5 shows mean data for both scenes. Graphs in column A show mean perceived depth magnitude as a function of crosstalk. Individual functions indicate data for different predicted depths (disparities) between pairs of objects. In the absence of an effect of crosstalk, all lines would be parallel to the x-axis, but this is clearly not the case. Instead, as crosstalk increases there is a decrease in perceived depth at all disparities. This effect can be appreciated from a different perspective in the graphs shown in column B. Here we have re-plotted perceived depth as a function of the depth in cm predicted by the disparity, now each function corresponds to a different level of crosstalk. If crosstalk had no effect then the lines on this graph would coincide. It is clear that perceived depth was reduced at crosstalk levels as low as 4%.

Since there was a relatively large difference between the perceived/predicted depth of the largest and the smallest disparities in both scenes, the magnitude of the effect of crosstalk at the smallest disparities might not be appreciable in the raw data. To examine the effects in the small disparity range more closely we normalized the data for each disparity for each observer by dividing the depth estimates for each disparity by the largest estimate obtained for that disparity. The averaged normalized data are shown in column C of Figure 5. It can be seen in this figure that depth judgments at all disparities were affected by crosstalk in a similar fashion.

These observations were confirmed by statistical analysis. All statistical analyses used alpha level of 5% and were performed using the statistical software package R. Data were analyzed separately for the Lab and the Kitchen scenes. We first analyzed the data using two-way repeated-measures ANOVAs with crosstalk and disparity as factors. The results are shown in **Table 1**. For both scenes the main effects of crosstalk and disparity were found to be statistically significant. This reflects the detrimental effect of crosstalk on perception of depth and the differences in perceived depth as a function of disparity. The interactions between disparity and crosstalk were also found to be significant for both scenes. This occurs since the difference in depth estimates for different disparities decreases with increasing crosstalk as can be seen in Figure 5 (compare the differences in perceived depth at 0% crosstalk and 32% crosstalk).

Table 1 Results of statistical analysis with factorial ANOVA

	Effect	**F-value**	**P-value**
Lab scene	**Crosstalk**	$F(6,42) = 40.3$	$p < 0.001$
	Disparity	$F(3,21) = 87.23$	$p < 0.001$
	Crosstalk x Disparity	$F(18,126) = 13.2$	$p < 0.001$
Kitchen scene	**Crosstalk**	$F(6,30) = 31.65$	$p < 0.001$
	Disparity	$F(3,15) = 28.6$	$p < 0.001$
	Crosstalk x Disparity	$F(18,90) = 10.1$	$p < 0.001$

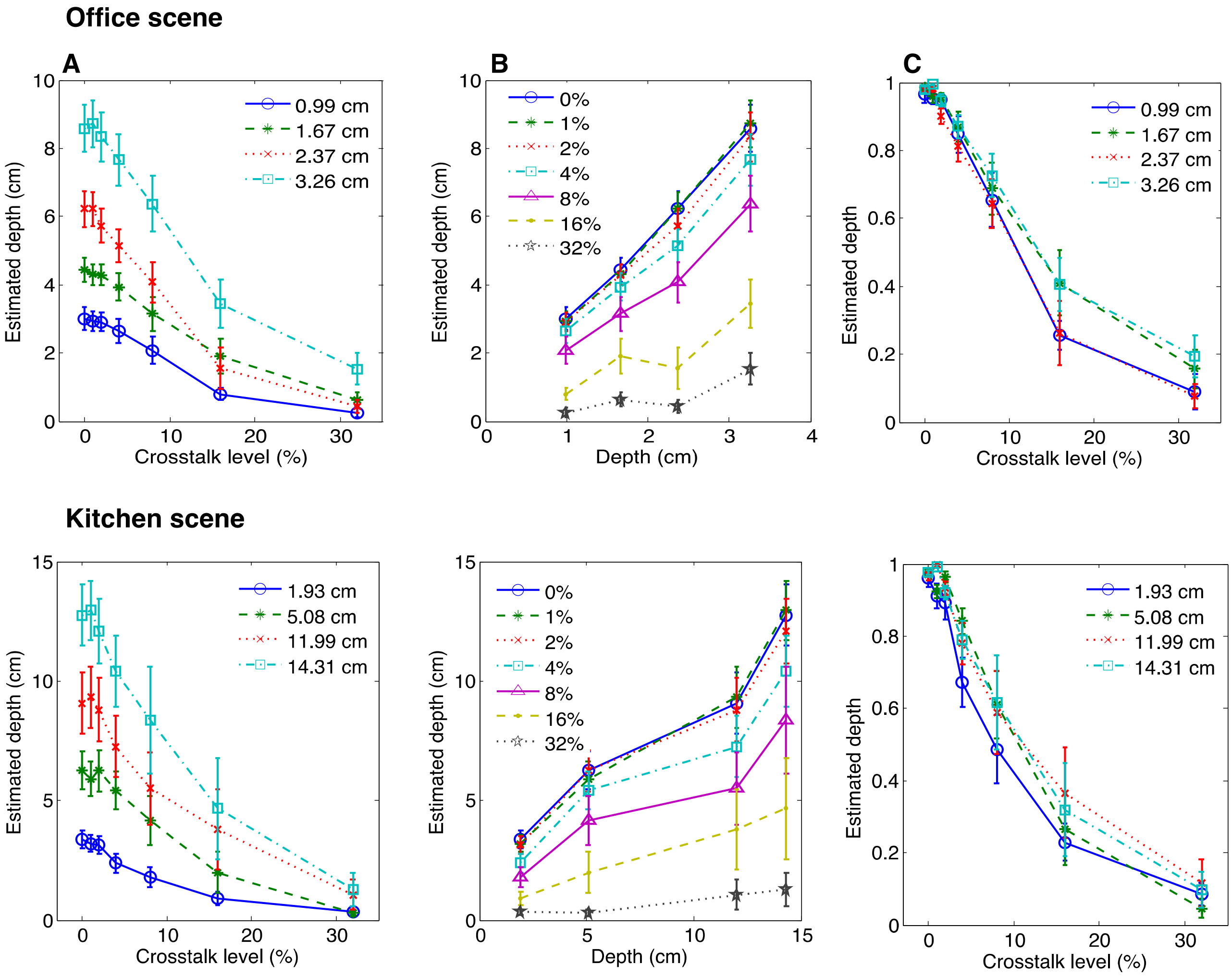


Figure 5 Mean data for the Lab scene (n=8) and the Kitchen scene (n=6). Column A: the abscissa shows the crosstalk levels and the ordinate the depth estimates. Different functions show stimuli with different disparities. The disparities are expressed in terms of the corresponding theoretical depth in cm (see text). Column B: the abscissa shows the theoretical depth corresponding to the test disparities and the ordinate shows the depth estimates. Each function represents stimuli with different crosstalk levels. Column C: the abscissa shows the crosstalk levels and the ordinate the *normalized* depth estimates (see text). Each function represents stimuli with different disparities. The error bars in all graphs indicate +/-1 standard error.

To establish at which level of crosstalk the estimated depth became significantly reduced for each disparity we compared each of the non-zero crosstalk conditions to the zero crosstalk condition using paired directional t-tests (all confirmed with a nonparametric Wilcoxon signed-rank test). **Table 2** shows at which crosstalk level perceived depth was significantly reduced in comparison to no crosstalk for each disparity in each scene. In each case, perceived depth at all subsequent (larger) crosstalk levels was significantly different from the no crosstalk condition so, for simplicity, we are not showing the results of these comparisons in the table. These data reveal that at smaller disparities, perceived depth was significantly reduced at 4% and at the larger disparities at 2%. This could indicate a stronger effect of crosstalk on depth percepts from larger disparities but it is more likely that it is simply harder to discern small changes at the smaller disparities. In support of this interpretation, the slopes in the normalized data graphs (Figure 5 column C) appear to be very similar for all disparities and in some cases smaller disparities seem to have steeper slopes than larger disparities.

Table 2 Results of statistical analysis with t-tests.

	Disparity	**Crosstalk level**	**t-test**
Lab scene	0.99 cm (5.3 arcmin)	4%	t(7)=2.03, p = 0.04
	1.67 cm (8.8 arcmin)	4%	t(7)=2.4, p = 0.023
	2.37 cm (12.4 arcmin)	2%	t(7)=3.12, p = 0.008
	3.26 cm (17.7 arcmin)	4%	t(7)=2.45, p = 0.021
Kitchen scene	1.93 cm (10.6 arcmin)	4%	t(5)=3.3, p = 0.011
	5.08 cm (26.5 arcmin)	4%	t(5)=3.6, p = 0.007
	11.99 cm (56.6 arcmin)	2%	t(5)=2.5, p = 0.026
	13.31 cm (65.5 arcmin)	2%	t(5)=2.6, p = 0.024

Note that in the Lab scene and in the two smaller disparities of the Kitchen scene, perceived depth is larger than the predicted theoretical depth. This could be due to the depth estimation method we employed. To control for this, we repeated the experiment for two observers (IT and TP) with the Lab scene using a disparity probe to estimate depth. The disparity probe consisted of two rectangles positioned side by side, which could be moved in depth away from each other by pressing buttons on a gamepad. Observers were instructed to match the stereoscopic depth between the two rectangles to the perceived stereoscopic depth of the object pairs in the Lab scene (the probe was presented simultaneously with the scene and positioned below it). The stimuli, apparatus and procedure were otherwise the same as in the original experiment except that we used two disparities (5.3 and 8.8 arcmin). As can be seen in Figure 6 both observers overestimated the disparity (depth) in all cases just as in the original experiment. Furthermore, the degree of overestimation for both observers was very similar in the two experiments: IT by a factor of ~2 and AC by a factor of ~3.

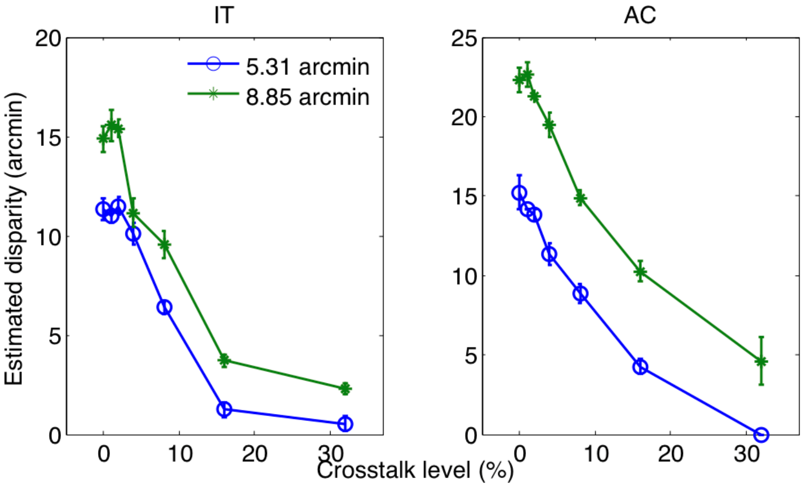


Figure 6 Individual data for two observers for the Lab scene estimated by using a disparity probe (see text). The abscissa shows the crosstalk levels and the ordinate the disparity estimates. Different lines show stimuli with different disparities. The error bars in all graphs indicate +/-1 standard error.

A more likely explanation for the overestimation of depth in the Lab and Kitchen scenes is related to the complexity of the S3D images[1]. That is, unlike our previous stimuli, these images contain multiple pictorial cues to

---

[1] Another possible source of depth overestimation in our images is the perspective distortion that occurs in photographs when they are viewed from a distance different from their center of projection [9]. However, this explanation would predict similar depth overestimation for both the Lab and the

depth such as linear perspective, texture gradients and occlusion as well as an overall implied scale. These pictorial cues likely add to the overall percept of depth between objects and may affect observers' estimates of stereoscopic depth. To probe this possibility we asked two observers AS and TP (naïve as the goals of the experiment) to estimate perceived depth between the same pairs of objects as in the original experiment with a 2D (and necessarily zero crosstalk) presentation of the Lab scene. The apparatus, stimuli and procedure were otherwise the same as in the original experiment. Although there was no stereoscopic depth in this case, observers still perceived depth between object pairs as shown in Figure 7 and this depth increased for pairs with larger relative disparity between them when viewed in S3D. The perceived depth in this case was smaller for all object pairs than that estimated by these observers in the 3D images. Thus, as one would expect, when pictorial and stereoscopic depth cues are consistent, they combine to produce the overall depth percept for the scene. Interestingly, crosstalk reduced the amount of perceived depth in the 3D images below that estimated for 2D images (compare Figures 7 and 5). This suggests that crosstalk affects depth from pictorial cues as well, but this hypothesis needs to be verified empirically.

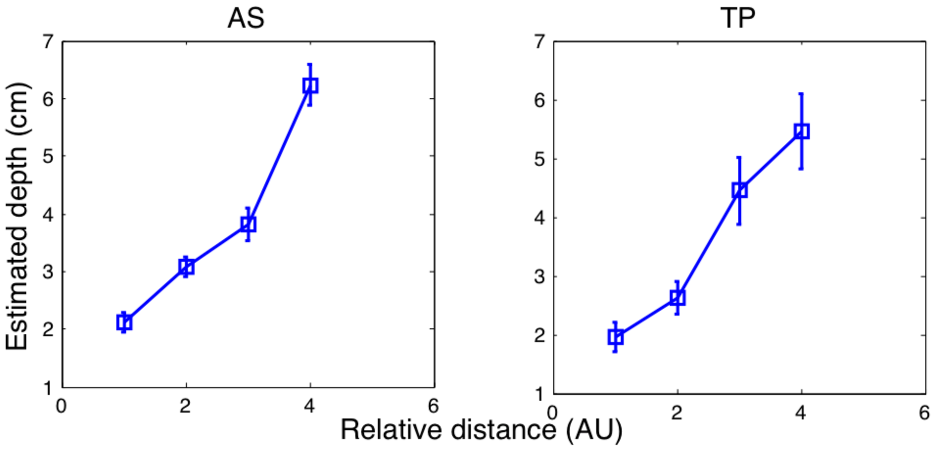


Figure 7 Individual data for the Lab scene presented in 2D with 0% crosstalk (see text). The abscissa shows the four different object pairs (relative distances) numbered 1-4 (in increasing order of disparity) and the ordinate the depth estimates. The error bars in all graphs indicate +/-1 standard error.

In the current study, increasing crosstalk caused a decrease in perceived depth in natural images in a manner very similar to that found with synthetic images in our previous experiments[7, 8]. Data were also similar quantitatively since with synthetic images there was a significant reduction in perceived depth at 4-8% crosstalk and with natural images perceived depth was significantly reduced at 2-4% crosstalk. Moreover, at 4% crosstalk for both synthetic and natural scene depth was reduce by 10-30% (depending on disparity) and at 32% crosstalk depth was virtually eliminated. However, there were several differences between the results of these studies. For synthetic images, the effect of crosstalk was more drastic at larger disparities, while in the case of natural images this trend is not as evident. It seems that crosstalk has a more consistently disruptive effect at all disparities in natural scenes. Moreover, when viewing natural images some observers spontaneously reported nausea and headaches after performing the task in S3D, which confirms previous findings that crosstalk causes discomfort in viewers [4]. We did not receive any such complaints of discomfort from observers tested with the simple synthetic stimuli. These differences suggest that the effect of crosstalk on perceived depth and on the general viewing experience might be more severe when viewing complex natural scenes instead of simple synthetic images.

Kitchen scenes, which was not the case. It is possible that perspective distortion interacts in complex ways with disparity (the disparity range in the Kitchen scene is larger), however, exploring this possibility is outside of the scope of this study.

4. CONCLUSION

The results of the experiments reported here show that our previous findings generalize to natural scenes demonstrating that crosstalk affects perceived depth magnitude even in complex natural scenes in the presence of pictorial depth cues. Moreover, this effect seems to be greater in natural scenes. Our data underscore the fact that crosstalk is a serious challenge to the quality of S3D media and has to be carefully addressed by display manufacturers. We recommend keeping crosstalk in S3D displays well under 4% in order to achieve maximum quality of stereoscopic depth and consumer satisfaction.

REFERENCES

[1] Woods, A. J., "Understanding crosstalk in stereoscopic displays," Three-dimensional systems and applications. (2010).

[2] Seuntiens, P. J. H., Meesters, L. M. J., and IJsselsteijn, W. A., “Perceptual attributes of crosstalk in 3D images,” Displays, 26, 177-183 (2005).

[3] Wilcox, L. M., and Stewart, J. A. D., "Determinants of perceived image quality: ghosting vs. brightness " Stereoscopic Displays and Virtual Reality Systems X, 5006 (2003).

[4] Kooi, F., and Toet, A., “Visual confort of binocular and 3D displays,” Displays, 25, 99-108 (2004).

[5] Lambooij, M., Fortuin, M. F., IJsselsteijn, W. A. *et al.*, "Visual discomfort associated with 3D displays," Fifth International Workshop on Video Processing and Quality Metrics for Consumer Electronics, (2010).

[6] Pala, S., Stevens, R., and Surman, R., "Optical crosstalk and visual comfort of a stereoscopic display used in a real-time application," Stereoscopic displays and virtual reality systems XIV, 6490, 1-12 (2007).

[7] Tsirlin, I., Wilcox, L. M., and Allison, R. S., “The effect of crosstalk on perceived depth from disparity and monocular occlusions,” Special Issue of the IEEE Transactions on Broadcasting: 3D-TV Horizon: Contents, Systems, and Visual Perception, 57(2), 445-453 (2011).

[8] Tsirlin, I., Wilcox, L. M., and Allison, R. S., “The effect of crosstalk on depth magnitude in thin structures,” Journal of Electronic Imaging, 21(1), (2012).

[9] Cooper, E. A., Piazza, E. A., and Banks, M. S., “Depth Compression and Expansion in Photographs,” Journal of Vision, 11(11), 65 (2011).