A Computational Model for Perception of Stereoscopic Window Violations

Steven Poulakos
Disney Research Zurich
ETH Zurich

Rafael Monroy
ETH Zurich

Tunc Aydin
Disney Research Zurich

Aljoscha Smolic
Disney Research Zurich

Markus Gross
Disney Research Zurich
ETH Zurich

***Abstract*—Creating a computational model for stereoscopic 3D perception is a highly complex undertaking. As one step towards this goal, this paper investigates stereoscopic window violation artifacts, which often interfere with artistic freedom and constrain the comfortable depth volume. Window violations need to be compensated for in most 3D feature movies. Currently this is done in an ad-hoc manner due to a limited understanding of the problem. In this work, we present a model predicting problematic window violations that are visually disturbing. The model parameters were defined through psychophysical experiments on simple stimuli. Then the model was calibrated and validated on real, complex stereoscopic images. Finally, we present a system to provide visualization of problematic stereoscopic window violations as well as details for how to correct them.**

I. Introduction

The stereoscopic window violation is a problematic artifact influencing perceived stereoscopic 3D quality. It occurs when depth perception from stereopsis and occlusion depth cues are inconsistent due to interactions with the stereoscopic window border. Fortunately, window violations are not always problematic. Making that assessment and compensating, if necessary, requires time-consuming expert input. We present a computational model based on perceptual measurements to identify when a window violation is problematic and a system to automatically correct it.

The Stereoscopic Window represents the virtual window through which stereoscopic depth is perceived. Its importance has been recognized since stereoscopic pictures were first made [1]. The stereo window is defined by the parallax of the lateral edges of the two images projected on the screen as represented in line segment p_1p_2 of Figure 1. At zero parallax, the window is perceived to be at the plane of the screen.

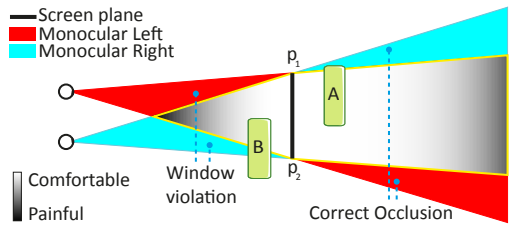


Fig. 1. Stereoscopic window violations can occur in the monocular viewing regions located in front of the screen. Features from Object B are occluded by the screen edge p_2 behind it.

This research was supported by ETH Research Grant TH-23/04-3 and also by the European Commission under the Contract FP7-ICT-287723 REVERIE.

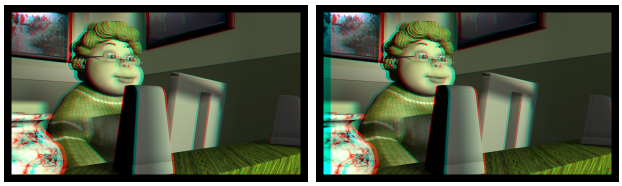
Stereoscopic window violations occur when a scene element meets two conditions: (1) it must be presented with disparity nearer than the stereo window, and (2) it must collide with the lateral image border as shown in Figure 1. In the real-world, this problem would never exist. An object in front of the window is visible to both eyes. When objects in front of the stereo window are only visible to one eye, the viewer perceives the image border to occlude the missing information. Two depth cues are in conflict: disparity provided by the visible stereo features and the depth ordering from occlusion. Humans are most sensitive to occlusion in identifying proper depth order [2], therefore violating it has the potential to produce a disturbing visual conflict.

Window violations are difficult to avoid. The range of stereoscopic depth that is comfortable to view is limited to a region both in front of and behind the screen [3]. Maximizing this zone of comfort requires placing scene elements in front of the screen (negative parallax), increasing the likelihood that a window violation will occur.

Floating Windows are a common solution to remove stereoscopic window violations. It is produced by applying an asymmetric mask to the left and right eye images. Figure 2 provides an example scene with and without floating windows applied. Floating windows remove features that should be visible in the two eyes. The black border to the left of the image appears to float in front of the vase. This preserves the expected depth ordering of occluding elements.

Floating windows have been applied in feature cinematic films as a tool to utilize a larger depth volume, regain artistic control and remove artifacts that distract attention from the story [4]. Digital cinematography makes it easier to vary floating windows between shots, animate them within shots, and place them to correct only the problematic violations. However, these are all time-consuming tasks requiring expert input that can be especially difficult to apply in real-time applications. Our research can automate these operations.

Other solutions include global disparity shift to push objects behind the stereo window, at the risk of exceeding the comfort zone. Local disparity warping methods have been demonstrated to move only a conflicting object behind the image plane [5]. Blur near the image border has been proposed to reduce visibility of the violation [6]. Existing solutions are often applied in an empirical, ad-hoc manner. We aim to formalize this through a computational model of perception.



(a) Window Violation (b) Floating Windows

Fig. 2. (a) Stereoscopic image with window violation. (b) Window violation removed with floating windows. Note the asymmetric mask on the left.

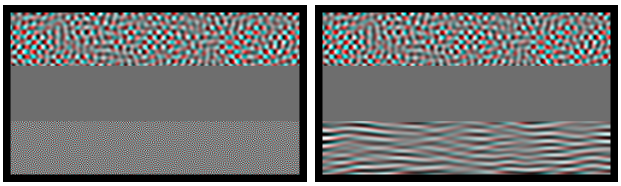
II. Perceptual Modeling

Human vision is better at distinguishing two objects when their relative difference in color or luminance is large. This difference can be expressed in terms of *contrast* [7]. The inverse value of the minimum contrast required for detection is called *contrast sensitivity*. The change in contrast sensitivity is a function of spatial frequency and is modeled by the well-known *contrast sensitivity function* [7].

We are inspired by image quality assessment metrics. Some metrics are based on low-level representations of the human visual system (HVS) including contrast sensitivity and visual masking [8]. Perceived contrast distortion metrics, such as the *visible difference predictor* (VDP) [9], are designed for detecting near threshold differences, and extensions [10] are capable of expressing suprathreshold difference in JND (just noticeable difference) units. Another quality metric is the use of *structural similarity index metric* (SSIM) [11], which exploits the structural information from a scene. Higher level visual equivalence models provide metrics to determine when perceived image changes do not result in a perceived change in image quality [12], [13]. These suprathreshold models do not predict the probability of detection, but rather assess if a difference is significantly visible. These considerations influence the performance of an image quality metric [14].

Stereoscopic Visual Processing also has a contrast sensitivity correlate. Frisby and Mayhew's demonstrated a correlation between stereopsis sensitivity and contrast detection sensitivity as a function of spatial frequency [15]. Their findings show that the shape of contrast detection and stereopsis are similar, although with a shift representing a decrease in stereopsis sensitivity. It is possible to perceive the disparate features, but not achieve stereopsis. The CSF correlation with stereopsis has led to models of perceived depth of frequency and magnitude changes in disparity [16]. In contrast, our work is focused on detecting disturbing window violations.

We assess if the visual system can find a viable depth interpretation when point correspondences do not exist. Marr and Poggio [17] developed a cooperative algorithm for extracting disparity information. They suggest that perceiving the disparity of an object involves finding a continuous and smooth matching of point correspondences. Mitchison and McKee [18], [19] later observed that strong edges can cause a stable, yet incorrect correspondence match of stereoscopic stimuli. The visual system favors strong edges at the expense of misinterpreting the fine texture detail. This implies that the image border can provide a strong edge biasing the interpretation of scene elements to be perceived as if they are behind the stereoscopic window, and free of violation.



(a) Main experiment (b) Orientation experiment

Fig. 4. Example stimuli. (a) Main model experiment with no orientation filtering. The top stimulus is spatial frequency condition level three (SF3) at 1.16 cpd, and the bottom stimulus is SF4 at 4.65 cpd. (b) Orientation mixing experiment. The top stimuli is the All Pass (AP) condition. The bottom stimuli is the Horizontal (H) condition. Both stimuli are spatial frequency condition SF3 (1.16 cpd). Example provided in anaglyph. The black border is provided by the black border of the HDTV display device in a darkened room.

We are guided by two key concepts: (1) Stereopsis sensitivity has a CSF-like behavior, and (2) strong edges of the image borders can influence the depth interpretation of a window violation. We hypothesize that a scene element in window violation will be problematic when it is represented by strong visible contrast. The following sections present our construction and validation of a computational model for the perception of disturbing stereoscopic window violations. We isolate four dominant variables to develop our perceptual model: contrast magnitude, spatial frequency, orientation, and disparity. Our model as represented in Figure 3 produces a binary classification of disturbing window violation.

III. Model Experiments

Stimuli - Each stimuli trial was composed of two textured planes, presented in stereoscopic depth (see Figure 4). The stimuli was shown on a 50" Panasonic 3D plasma TV (TX-P50VT20E) in a darkened room. Subjects were seated 2 meters away from the display wearing active shutter stereoscopic glasses. The maximum pixel disparity of each plane is -50px, corresponding to an angular disparity of -0.8°. The height of each plane was 350px (angular height of 5.9°) and their width was 1920px (angular width of 30.75°).

The stimulus planes were textured with filtered, random dot stereograms to control the experiment conditions of contrast magnitude, orientation and frequency. We presented 15 combinations of spatial frequency and contrast. Five spatial frequency levels were investigated spanning the range from 0.14 cpd to 18.6 cpd. The possible levels of maximum contrast were 0.21, 0.63 and 1.35. The orientation condition was created using a fan filter applied in the frequency domain [20]. Disparity was applied to each plane.

Procedure - The experiments were implemented as a two-alternative forced choice (2AFC) procedure. Human subjects were paid, naive to the experiment and had normal or corrected-to-normal visual acuity and stereo acuity. Subjects were first introduced to the concept of window violations through presentation of real-world examples. The experiment task was to compare the two planes in the stimuli image (Figure 4) and choose which of the two looked less disturbing or annoying. They were instructed to base their assessment on the image regions close to the left and right vertical screen borders, where window violations are expected to occur. The experiment duration was approximately 40 minutes.

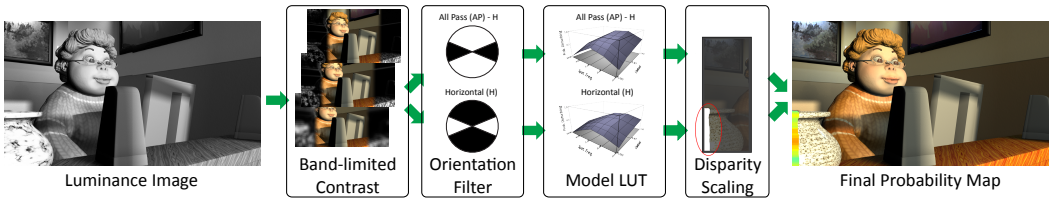


Fig. 3. Pipeline of our computational model. A luminance image is decomposed into band-limited contrast and contrast orientation channels. We then apply our predictive model and additional disparity scaling. Results are combined using winner-take-all producing the final probability map.

Evaluation - Our **main experiment** observed the influence of contrast and spatial frequency on the preference of window violations. Six subjects participated and evaluated the balanced combinations of five spatial frequency levels and three contrast levels resulting in 225 trials per experiment. No orientation filtering was applied on these stimuli. Disparity was fixed at -50 pixels. Figure 4(a) presents an example.

We performed Two-factor Analysis of Variance (ANOVA) with repeated measures to analyze the influence of spatial frequency and contrast on viewer preference. There was a strong statistical significance for within-subjects main effects of contrast ($F(2, 30) = 367.32$, $p < .001$) and spatial frequency ($F(1.32, 19.89) = 100.43$, $p < .001$) using Greenhouse-Geisser adjusted degrees of freedom. Both factors significantly influence window violation preference. We also analyzed the Bonferroni-adjusted pairwise comparison between the levels of each condition [21]. We observed a statistically significant difference in preference between all contrast levels ($p < .01$) and nearly all spatial frequency levels.

The responses collected from each session are normalized to represent the likelihood for each combination to be preferred. Figure 5(a) demonstrates a monotonic preference for lower contrast as well as a preference for high and low spatial frequency conditions. The plot also shows behavior similar to the familiar contrast sensitivity function for luminance supporting the assumption that a disturbing window violation is significantly influenced by contrast magnitude and frequency.

The **orientation mixing experiment** was created by filtering the RDS textures applied to the two stimuli planes using a segmented fan filter. Through a series of experiments, we identified a significant orientation effect when the stimuli presented a narrow orientation band. Our orientation experiment enabled comparison of different stimuli orientations as shown in Figure 5b. This experiment was conducted with six new subjects who had not participated in the main experiment. The maximum contrast and disparity condition levels were held constant. Spatial frequency and orientation was balanced resulting in 400 trials per experiment. A narrow orientation filter was used (one 22.5° segment from an eight segment fan filter) to produce the following orientation conditions: All-Pass (no orientation filtering), Horizontal, Diagonal and Vertical orientations.

The narrow orientation tuning enabled us to directly observe a significant orientation effect ($F(3, 15) = 14.29$, $p < .001$) in addition to reproducing a significant spatial frequency effect ($F(4, 20) = 4.06$, $p < .05$). The results are represented in Figure 5(b). The horizontal condition was the only condition to show an insignificant influence ($p = .125$) on spatial frequency. The spatial frequency preference curve was much flatter. Pairwise comparison reveal no significant preference between spatial frequency levels for the horizontal condition.

All Pass, Diagonal and Vertical stimuli orientation conditions had similar mean preference scores. However, the Horizontal condition was significantly preferred more than the All Pass and Diagonal orientations ($p < .05$) and a trend for preference over Vertical ($p = .066$). The mean preference for the Horizontal condition was 30% higher than all conditions and 32% higher than All Pass.

The **disparity mixing experiment** was conducted through balanced presentation of disparity and spatial frequency conditions. The maximum contrast level was used and no orientation filtering applied. Four levels of negative disparity were compared within the experiment: 6, 12, 25 and 50 pixels (ranging from -0.1° to -0.8° angular disparity). The experiment involved 400 trials per experiment.

The effect of the disparity condition was significant ($F(3, 15) = 58.35$, $p < .001$). The results are represented in Figure 5(c). Pairwise comparison revealed a significant preference between disparity levels for all except the two smallest, which were nearly significant ($p = .058$). There was also a significant interaction between disparity and spatial frequency ($F(2.77, 13.83) = 8.52$, $p < .001$). Smaller disparities showed a flatter CSF-like preference.

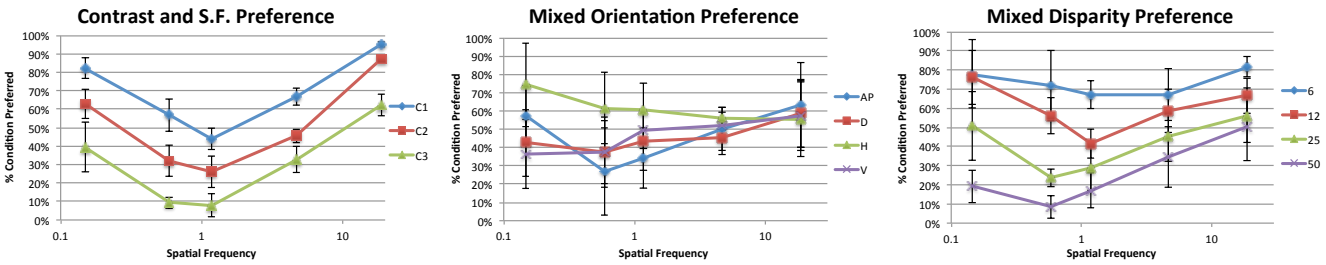
IV. Computational Model

Experimental data guided creation of our computational model depicted in Figure 3. We first process the image content from RGB to perceptually meaningful luminance units. Perceived luminance depends on the spectral emission properties of the display. Our display was characterized using a Photo Research PR-730 spectroradiometer. The obtained spectral emission curves for the three channels were applied to the color matching function of the *XYZ* colorspace [22] to obtain the luminance image described by the Y component. The resulting image describes the amount of $\frac{cd}{m^2}$ emitted by the display on a per-pixel basis.

Contrast and spatial frequency information is extracted from a complex image using Peli's [23] decomposition into band-limited versions by applying cosine-log filters. A cosine-log filter centered at 2^i cycles/picture is defined as:

$$G_i(u,v) = G_i(r) = \frac{1}{2}(1 + cos(\pi log_2 r - \pi i)), \qquad (1)$$

where u and v are the horizontal and vertical spatial frequency coordinates respectively and r is one of the polar spatial frequency coordinates defined as $r = \sqrt{u^2 + v^2}$.



(a) Main Experiment (b) Orientation Experiment (c) Disparity Experiment

Fig. 5. (a) The mean likelihood for each combination of contrast and spatial frequency to result in a more preferred window violation. Participants preferred low and high spatial frequency and had a monotonically increasing preference for lower contrast. Contrasts C1, C2 and C3 correspond to .21, 0.63 and 1.35 respectively. (b) Orientation mixing experiment presented all paired comparisons of four orientation conditions and five spatial frequencies. Spatial frequency effects are not as clear as our other experiments. However, the mean preference for Horizontal orientation condition was 30% greater. (c) Disparity mixing experiment presented all paired comparisons of four disparity conditions and five spatial frequencies. There is a monotonically increasing preference for smaller disparity. Error bars represent 95% confidence intervals.

The contrast of the ith band is computed as:

$$c_i(x,y) = \frac{|a_i(x,y)|}{L'}, \tag{2}$$

where c_i is the ith contrast image, a_i is the band-limited luminance image and L' is the mean luminance of the luminance image. L' is motivated by its use in VDP-related metrics [9].

We construct a lookup table (LUT) based on data collected in the main experiment (Figure 5(a)) to represent the likelihood a combination of contrast and spatial frequency is perceived to be more disturbing than the other conditions. The LUT is a single matrix, with 15 entries in total: one for each combination of spatial frequency and contrast condition level. The values were normalized resulting in a matrix describing the likelihood for each combination to be visually disturbing.

Bilinear interpolation of the LUT is applied between sample points. This results in one probability map for each spatial frequency band. Bands are combined together using a winner-take-all approach, such that for each pixel we use the maximum probability value across all probability maps. This approach is motivated by the independent-channel hypothesis in which disparity discrimination is influenced by the largest active spatial frequency channel [24].

As visualized in Figure 3, the LUT is further modulated by a scaling coefficient to account for orientation effects. We apply a scaling factor, $s_o = 0.7$, to our LUT for only the narrow-band horizontal components. All other orientation features (AP-H) have a scaling factor, $s_o = 1.0$.

The disparity scaling factor, s_d, is computed by shifting the linear fit from the disparity experiment such that our maximum disparity tested, 50 pixels, is represented by $s_d = 1$. Our primary LUT already represents the probability disturbing for 50 pixel disparities. The result is the following disparity scaling function: $s_d = 0.1603 \log d_{max} + 0.3598$. Since disparity is undefined in window violation regions, we set d_{max} equal to the maximum window violation size for all pixels of the given row.

Since we are only interested in the regions in window violations, we prune the probability map using a disparity map of the stereoscopic pair removing regions that are not in contact with the borders or do not have negative disparities. This gives a window violation detection mask as shown in the Disparity Scaling component in Figure 3 as well as in results Figure 8.

The probability map per orientation channel, k, is expressed as follows:

$$P^k(x,y) = \max_n(P_0^k(x,y), ..., P_N^k(x,y))\ s_o^k\ s_d, \tag{3}$$

where the maximum per frequency band, P_N, is modulated by orientation and disparity scaling coefficients. We then apply the maximum orientation channel, H or AP-H, to produce a final per pixel probability map:

$$P(x,y) = \max_k(P^k(x,y)). \tag{4}$$

The next section describes model validation and thresholding for binary classification.

V. Validation Experiments

We conducted a subjective study using real-world and computer-generated images to validate our model and obtain a measure regarding its performance. Since our goal is to predict whether a window violation will be perceived as disturbing, we asked subjects to look at stereoscopic images and label where a window violation was disturbing.

Stimuli - Stimuli consisted of 95 stereoscopic images: screen captures taken from stereoscopic movies (including live action and computer-generated imagery) and in-house produced computer-generated imagery. The stereoscopic images presented window violations as large as 50 pixels in width. The images created in-house described a similar scene, but with varying object texture and scene composition.

Procedure - The experiments used the same setup as described in Section III with an additional computer monitor for user input (Figure 6). Our subjective methodology was similar to the one presented by Aydin et al. [25] to evaluate HDR video tone mapping. Eleven subjects participated and were instructed to look at the stereoscopic image in the 3D TV and localize the regions on the lateral borders perceived as disturbing or annoying. The 2D monitor presented the left half of the left image together with the right half of the right image overlayed with a grid. This approach allowed all pixels representing window violations to be visible in one image.

(a) Stereoscopic image (b) 2D with grid overlay

Fig. 6. Validation and Calibration Experiment. (a) Stereoscopic Image (presented in paper as anaglyph). (b) The user interface for grid-based selection of problematic window violations.

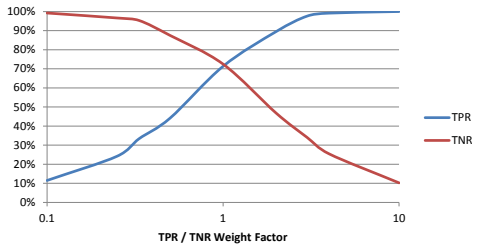


Fig. 7. Resulting TPR and TNR after varying their respective importance.

Each cell had a size of 25 pixels and users were asked to label cells containing disturbing window violations.

Evaluation - The collected data was averaged per stereoscopic image across all 11 subjects to provided ground truth labeled data. The model prediction and ground truth were both quantized to a cell resolution of 50 pixels. To account for labeling error and inconsistency, we defined sensitivity for the ground truth at 60% of subject agreement. To measure the performance we computed 7-fold cross-validation between the ground truth and model prediction. The cross-validation is binned into training and test data sets to find a threshold value for our model that minimizes the absolute difference between the true positive ratio (TPR) and true negative ratio (TNR). The resulting threshold of our cross-validation gave an average TPR: 71% and TNR: 72%.

Sensitivity of the system can be tuned to different applications. A floating window generator may be optimized for a high TPR in order to not miss disturbing violations at the cost of correcting some of those that are not disturbing. A quality assistance system may be optimized more towards TNR, predicting when it is not necessary to intervene with high reliability. Figure 7 illustrates the performance when we vary the relative importance. For instance, if we set the true positives to be twice as important as the true negatives we observe a TPR: 89% and TNR: 47%.

VI. Results

Figure 8 illustrates our results by providing anaglyph image as well as visualization of the ground truth and model prediction. The middle column (Raw data) uses the Matlab JET colormap to represent the level of subject agreement in labelling window violations. Dark red and dark blue represent greatest and least agreement, respectively. The same colormap is used to represent the model prediction of a disturbing window violation. Dark red shows a higher probability disturbing. The right column shows the comparison of the thresholded ground truth data and thresholded model prediction. The threshold for the ground truth is set to 60% agreement, while the threshold used for the model is set to the value obtained from the cross-validation: 0.4394. The regions in red represent positives and blue negatives.

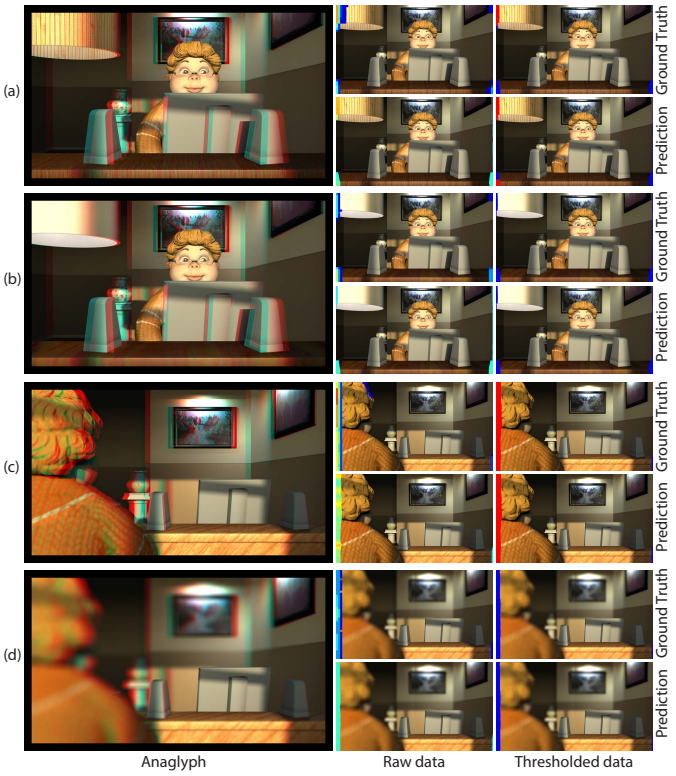
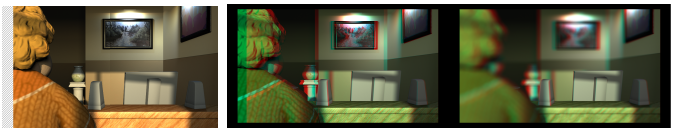


Fig. 8. Illustration of results: anaglyph image, ground truth and model prediction. Raw data column uses JET colormap to represent level of agreement (Ground Truth) or probability disturbing (Prediction). Thresholded data represent result of binary threshold: Red is positive (disturbing) and blue is negative (not disturbing).

Our model identifies most of the problematic areas. Figure 8(a) shows how our model correctly labeled the hanging lampshade to be disturbing, however, it incorrectly labeled a portion of the desk as problematic. Figure 8(b) is correctly predicted to not be problematic. This was accomplished by reducing texture detail in both the lampshade and desk. Similarly, Figure 8(c) and 8(d) show how disturbing window violations can be removed using depth-of-field. Contrast is reduced in the regions farther from the focal plane. Both the thresholded ground truth and model prediction data agree the disturbing window violation is removed.

Perceptual models by their nature are limited to certain number of factors. Our model currently assumes a black border around the image and does not explore the effect of matching the luminance of both the window violation and image border. We also do not directly consider the luminance of features in window violation, however, this is partially represented in assessing contrast of the luminance image. Our model also does not consider the influence of disparity frequency.



(a) Visualization (b) Floating Windows

Fig. 9. Applications. (a) Disturbing window violations are visualized by a zebra pattern. (b)-left Automatic floating windows to remove disturbing window violation. (b)-right is not predicted to be problematic due to depth of focus blur. No floating windows are applied.

VII. Applications

We present two applications of our model: visualization and the automatic removal of disturbing window violations. The stereo-matching algorithm developed by Werlberger [26] is used to identify regions of undefined disparity near the lateral borders.

Our visualization application is intended to support the detection of disturbing window violations. We adopt a zebra patterning approach commonly used to represent regions of an image that are overexposed. Visualization is provided in two modes. First, it identifies images containing disturbing window violations. Second, it provides a visual representation of the disturbing region by displaying an animated zebra pattern as shown in Figure 9(a). A user parameter is available to adjust sensitivity to disturbing window violations. This visualization could be included in a computational stereoscopic camera system [27] or stereoscopic analyzers [28]–[30] to provide real-time detection of disturbing window violations. It could also be applied as a quality assurance step before releasing stereoscopic content.

Floating windows effectively resolve window violations. However, it can be difficult to determine how to apply them, especially during live-capture. Our prediction can help reduce the uncertainty. Figure 9(b) shows two results of our automatic floating window generator: one requires the use of floating windows and the other does not. In Figure 9(b)-left, we apply a vertical asymmetric crop based on the largest window violation width. Because our system localizes regions that are disturbing, more elaborate floating windows can preserve pixels not in window violation.

VIII. Conclusion

We have demonstrated the development and application of a computational model for the perception of stereoscopic window violations. We presented a method of measuring window violation preference as a function of luminance contrast magnitude, spatial frequency, orientation and disparity. Our data fits the expectation of a CSF-like sensitivity function for stereopsis. The model was calibrated and validated using viewer input from real stereoscopic images. It can successfully detect user-labeled disturbing window violations. We present two important applications: visualization and automatic floating window correction of disturbing window violations.

References

[1] R. Spottiswoode and N. L. Spottiswoode, *The Theory of Stereoscopic Transmission*. University of California Press, Berkeley & Los Angeles, 1953.

[2] J. E. Cutting and P. M. Vishton, "Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth," in *Handbook of perception and cognition*. Academic Press, 1995, vol. 5, ch. Perception, pp. 69–117.

[3] T. Shibata, J. Kim, D. M. Hoffman, and M. S. Banks, "The zone of comfort: Predicting visual discomfort with stereo displays," *Journal of Vision*, vol. 11, no. 8, pp. 1–29, 2011.

[4] R. Neuman, "Bolt 3d: A case study," *Proc SPIE*, vol. 7237, 2009.

[5] M. Lang, A. Hornung, O. Wang, S. Poulakos, A. Smolic, and M. Gross, "Nonlinear disparity mapping for stereoscopic 3d," *ACM Trans. on Graphics (Proc. SIGGRAPH)*, vol. 29, no. 3, 2010.

[6] L. Lipton, *Foundations of the Stereoscopic Cinema: A Study in Depth*. Van Nostrand Reinhold Company Inc., 1982.

[7] P. G. J. Barten, *Contrast Sensitivity of the Human Eye and Its Effects on Image Quality*. SPIE - The International Society for Optical Engineering, 1999.

[8] A. B. Watson and J. A. Solomon, "Model of visual contrast gain control and pattern masking," *J. Opt. Soc. Am. A*, vol. 14, no. 9, pp. 2379–2391, Sep 1997.

[9] S. Daly, "Visible differences predictor: an algorithm for the assessment of image fidelity," *Proc SPIE*, vol. 1666, 1992.

[10] R. Mantiuk, K. J. Kim, A. G. Rempel, and W. Heidrich, "HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions," *ACM Trans. Graph.*, vol. 30, no. 4, pp. 40:1–40:14, Jul. 2011.

[11] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, Apr. 2004.

[12] G. Ramanarayanan, J. Ferwerda, B. Walter, and K. Bala, "Visual equivalence: towards a new standard for image fidelity," ser. SIGGRAPH '07. ACM, 2007.

[13] J. Krivánek, J. A. Ferwerda, and K. Bala, "Effects of global illumination approximations on material appearance," in *ACM SIGGRAPH 2010 papers*, ser. SIGGRAPH '10. ACM, 2010, pp. 112:1–112:10.

[14] M. Cadik, R. Herzog, R. Mantiuk, K. Myszkowski, and H.-P. Seidel, "New measurements reveal weaknesses of image quality metrics in evaluating graphics artifacts," in *SIGGRAPH Asia 2012*, 2012.

[15] J. P. Frisby and J. E. W. Mayhew, "Contrast sensitivity function for stereopsis," *Perception*, vol. 7, pp. 423–429, 1978.

[16] P. Didyk, T. Ritschel, E. Eisemann, K. Myszkowski, and H.-P. Seidel, "A perceptual model for disparity," *ACM Trans. on Computer Graphics*, vol. 30, no. 4, pp. 96:1–96:10, 2011.

[17] D. Marr and T. Poggio, "Cooperative computation of stereo disparity," *Science*, vol. 194, no. 4262, 1976.

[18] G. J. Mitchison and S. P. McKee, "Interpolation in stereoscopic matching," *Nature*, vol. 315, 1985.

[19] S. P. McKee and G. J. Mitchison, "The role of retinal correspondence in stereoscopic matching," *Vision Research*, vol. 28, no. 9, 1988.

[20] A. B. Watson, "The cortex transform: rapid computation of simulated neural images," *Comput. Vision Graph. Image Process.*, vol. 39, no. 3, pp. 311–327, Sep. 1987.

[21] D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*. Chapman & Hall/CRC, 2007.

[22] Y. Ohno, "Cie fundamentals for color measurements," *Proceedings of IS&T NIP 16 Conference*, 2000.

[23] E. Peli, "Contrast in complex images," *J. Opt. Soc. Am. A*, vol. 7, no. 10, pp. 2032–2040, Oct 1990.

[24] D. Marr and T. Poggio, "A computational theory of human stereo vision," *Proceedings of the Royal Society of London. Series B, Biological Sciences*, vol. 204, no. 1156, 1979.

[25] T. O. Aydin, M. Cadík, K. Myszkowski, and H.-P. Seidel, "Video quality assessment for computer graphics applications," *ACM Trans. Graph.*, vol. 29, pp. 161:1–161:12, December 2010.

[26] M. Werlberger, T. Pock, and H. Bischof, "Motion estimation with non-local total variation regularization," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, San Francisco, CA, USA, June 2010.

[27] S. Heinzle, P. Greisen, D. Gallup, C. Chen, D. Saner, A. Smolic, A. Burg, W. Matusik, and M. Gross, "Computational stereo camera system with programmable control loop," *ACM Trans. Graph.*, vol. 30, pp. 94:1–94:10, August 2011.

[28] F. Zilly, M. Muller, P. Eisert, and P. Kauff, "The stereoscopic analyzer — an image-based assistance tool for stereo shooting and 3d production," in *Image Processing (ICIP), 2010 17th IEEE International Conference on*, sept. 2010, pp. 4029 –4032.

[29] A. Smolic, S. Poulakos, S. Heinzle, P. Greisen, M. Lang, A. Hornung, M. Farre, N. Stefanoski, O. Wang, L. Schnyder, R. Monroy, and M. Gross, "Disparity-aware stereo 3d production tools," in *Visual Media Production (CVMP), 2011 Conference for*, Nov 2011, pp. 165–173.

[30] Cel-Soft, "Cel-scope3d stereoscopic analyser," January 2012, available from http://www.cel-soft.com/celscope3d/.