EFFECTS OF DISPLAY PIXEL PITCH AND ANTIALIASING ON THRESHOLD VERNIER ACUITY

Charles J. Lloyd[1], Marc D. Winterbottom[2], James P. Gaska[2], and Logan A. Williams[2]

[1]Visual Performance LLC, Ellisville, MO, USA
[2]Operational Based Vision Assessment Laboratory, USAF School of Aerospace Medicine, Wright-Patterson AFB, OH, USA

ABSTRACT

The term "eye-limited" resolution (ELR) has seen significant use in recent years within the simulation training and related industries. Results of a literature review revealed several distinct definitions of ELR and a range of estimates of the pixel pitch required to achieve it. When asymptotic visual task performance is used as the basis of ELR, relatively consistent results are obtained for practical tasks such as target identification and orientation detection range. The results of nine published evaluations indicate the pixel pitch that produces 90% of peak performance is in the range of 0.5 to 0.93 arcmin with a median estimate of .7 arcmin. However, a number of authors have asserted that a much finer pixel pitch may be required if observers are to achieve eye-limited performance on hyperacuity tasks such as Vernier acuity. Given that resolution is a primary driver of the performance, cost, and complexity of training display systems, this assertion was tested in the present evaluation.

Performance on the Vernier acuity task was predicted using an observer model and was also measured using five high acuity human subjects who viewed 20 combinations of pixel pitch and antialiasing filter width. The human performance data confirmed the expectation that with sufficient antialiasing a 7.5 arc second Vernier acuity threshold can be obtained with a pixel pitch of 1.6 arcmin. However, a much smaller pixel pitch was required to obtain that level of performance without antialiasing. Based on the results of the research presented here, and our previous work, we conclude that the combination of pixel pitch of approximately 0.7 arcmin and sufficient antialiasing supports eye-limited task performance, even for tasks involving hyperacuity, such as Vernier acuity and stereo acuity. This conclusion is relevant for identifying visual system requirements for training and human factors research.

INTRODUCTION

The results of previous research[1-8] have shown that the pixel pitch required to achieve 90% of asymptotic performance on several acuity-bound visual tasks ranges between 0.5 and 0.93 arcmin with a median estimate of 0.7 arcmin. However, this level of performance is obtainable only if the images are relatively free of spatial sampling artifacts[8, 9] which generally requires the use of sufficient antialiasing. The consistency of these results is good considering the large differences in the tasks and data analyses used across the studies. These results are also consistent with the pixel pitch of the highest resolution versions of the Apple iPhone, iPad, and MacBook Pro which range between .76 and .87 arcmin when viewed from their design distances. This pixel pitch is fine enough that Steve Jobs declared[10] the resolution of the iPhone 4 was "comfortably over the limit" of the human retina.

While many appear to agree that ELR is achievable with a pixel pitch of approximately 0.7 arcmin, a number of authors have asserted that a much finer pixel pitch (e.g., a few arc seconds) may be required if observers are to achieve eye-limited performance for hyperacuity tasks[11-14] such as Vernier acuity. In these papers, the assertions are based on theoretical considerations and the authors do not provide or cite empirical data that indicate a much finer pixel pitch is required. Given the obvious consequences of this assertion in terms of display system cost and complexity, we designed the present study to directly measure the effect of display resolution on Vernier acuity threshold. In the current study we measured performance for many combinations of pixel pitch and antialiasing for the purpose of revealing and quantifying the trade space for these design variables

Presented at the **IMAGE 2015** Conference
Dayton, Ohio – July 2015

METHOD

The collection of Vernier acuity data from human subjects is tedious, time consuming, and therefore expensive. Typically, this type of threshold data has enough variance that the measurements must be repeated many times and averaged together in order to acquire stable estimates. Given the large amount of data required to characterize the design trade space, and the limited budget allocated for hiring and managing subjects, a model-centric approach was taken for this evaluation.

With this approach the Vernier acuity task was designed and the display system set up, calibrated, and tested with human observers prior to the modeling exercise. Next, the stimuli that would ordinarily be presented to the human observer were presented to an observer model that examined each image and computed the positon of the top line segment relative to the bottom. This calculation was performed using the horizontal centroids of the line segments computed as per Equation 1.

1) $C = \Sigma \text{ lum x dx} \;/\; \Sigma \text{ lum dx}$

where x is the horizontal distance across the display surface (arc seconds) and lum is the luminance of the line and background as a function of distance. For each trial the difference between the centroids of the two line segments was computed and a random sample from a Gaussian distribution (mean = 0) was added to the difference. If the difference plus noise was greater than zero, the top line was considered to be to the left of the bottom line. The standard deviation of this noise source was the only free parameter in the model that was fit to the empirical data. This parameter was set to produce a Vernier acuity threshold floor of 7.5 arc seconds. All other aspects of the experiment (e.g., use of Quest[15] to control stimulus magnitude) were held constant between the observer model and the human subjects.

Using the observer model, the time required to measure a single threshold was less than one second. Thus, it was practical to make many repeated measurements of hundreds of design variable combinations. With the human subjects, measurement of a single threshold required about 3 minutes, thus, we could afford to collect far less data with the humans. The human performance data were used to test the hypothesis that the observer model was a valid representation of human performance for this task.

Equipment and Software

The Vernier acuity stimuli were presented to subjects using an Optima HD25e single chip DLP projector with a native addressability of 1920 x 1080 pixels. The projector contained a 240w arc lamp and was rated at 2000 lumens. The projector was positioned 29.5 inches from a wall mounted screen and the zoom was set to produce an image that was 49 x 27 cm (19.25 x 10.75 in) for a pixel pitch of 0.25 mm (0.010 in). Observers were seated at a distance of 4.37 m (172 in) from the screen which produced a native pixel pitch of 0.20 arcmin. At this distance the FOV of the projected image was 6.5 x 3.6 deg.

The projection screen was the same size as the image and was painted with flat grey primer with a reflectance of 0.44. All of the lights in the room were turned on which produced an average wall luminance of 25 fL. Two 1500 lumen compact fluorescent flood lamps were positioned 1.5 m (5 ft) from the front wall of the room to increase the luminance of the wall surrounding the screen to 69 fL.

The stimuli were created and presented using the MATLAB software from The Math Works. The electro-optical response function (gamma) of the projector was measured in the presence of the ambient lighting using a Minolta CS-100 meter. A response correction equation was employed which linearized the response of the display and ambient lighting. The peak luminance of the display was 470 fL and the luminance of the black lines in the stimulus was 33 fL for a maximum contrast ratio of 14:1.

Experimental Procedure

Subjects made their responses using a Flight Stick Pro game controller. On each trial the observer indicated the position of the upper line relative to the lower line. To indicate their responses subjects moved the controller to the left or right and pressed the trigger. The next stimulus presentation was initiated by the subject using a thumb button on the controller. Thus, the evaluation was paced by the observer who was allowed to take a break at any time. On each trial the pair of line segments was presented for a duration of 1.5 sec. The Quest procedure (PsychToolbox implementation) was used to control stimulus presentation and compute the thresholds. Each threshold was measured using a series of 37 trials.

The Vernier acuity stimuli consisted of two dark vertical line segments on a light background presented with a maximum contrast of 14:1 luminance contrast ratio. Line widths (half-maximum) were set to 2.4 arcmin for all conditions. The length of each segment was 40 arcmin and the gap between the segments was 2.4 arcmin. Figure 1 provides examples of two of the stimuli used in the evaluation. The lines presented in the left panel of the figure are 12 pixels wide (pitch = 0.2 arcmin) and the lines in the right panel are 2 pixels wide (pitch = 1.2 arcmin). The antialiasing filter width for both of these stimuli was set to 1.2 pixels. The use of antialiasing is more apparent in the right panel due to the much larger pixel size.

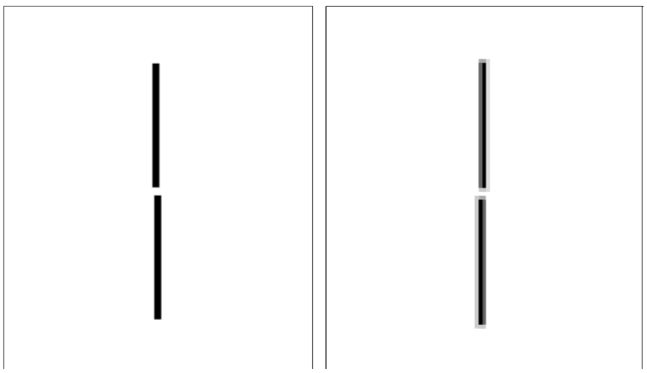

Figure 1. Two examples of the Vernier acuity stimuli with a line width of 2.4 arcmin and an antialiasing filter width of 1.2 pixels. Pixel pitch for the left panel is 0.2 arcmin and pixel pitch for the right panel is 1.2 arcmin.

For the first evaluation employing the observer model, all combinations of 30 levels of display pitch (0.1 to 3.0 arcmin) and 13 levels of antialiasing filter width (0 to 1.2 pixels) were used. Each of these experimental conditions was repeated 10 times for a total of 3900 measurements. The observer model was also used to measure 99 replicates of the 20 combinations of pixel pitch and antialiasing filter width that were used in the human performance evaluation.

For the second evaluation involving the human subjects, a total of 220 thresholds were measured for the 20 combinations of pixel pitch and antialiasing filter width for an average of 11 thresholds per experimental condition.

Five male subjects with a mean age of 28 years, participated in the evaluation. Subjects were invited to participate based on the results of clinical tests of visual acuity and stereopsis. All observers had better than 20/20 visual acuity and normal stereopsis and contrast sensitivity.

On arrival at the laboratory each observer was shown the experimental apparatus and the experimenter read the instructions and demonstrated the procedure. After signing an informed consent form the observers practiced the procedure. Data collection for each threshold required an average of 2 minutes to collect and a break of approximately one minute was provided between each threshold. Subjects took longer breaks after every ten thresholds. Including instructions, practice, data and rest breaks, data collection for each subject required about 2 hours.

RESULTS

Observer Model

The results of the first evaluation are presented in Figure 2 which shows Vernier acuity thresholds (arc seconds) as a function of pixel pitch (arcmin) and antialiasing filter width (pixels). From this figure it is readily apparent that antialiasing is expected to have a very large effect on performance for this task. When a 1.2 pixel wide antialiasing filter is applied, a threshold of 7.5 arc seconds is obtained for pixel pitches up to about 1.5 arcmin. Between 1.5 and 3 arcmin, the threshold slowly rises to about 10 arc seconds. In sharp contrast, when antialiasing was not used, the threshold rises rapidly with increasing pixel pitch. At a pitch of 0.1 arcmin, an 8 arc second threshold was obtained and the threshold increases to about 130 arc seconds at a pixel pitch of 3 arcmin.

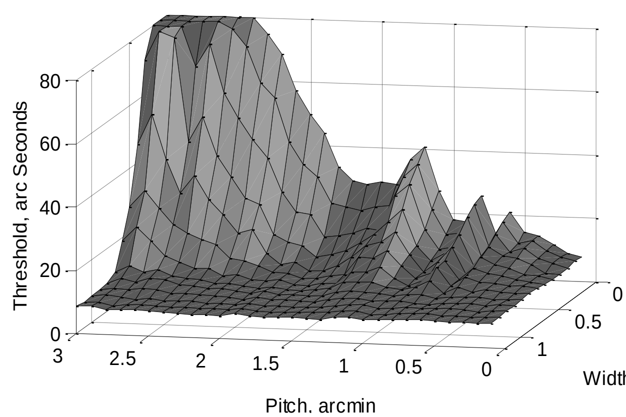

Figure 2. Vernier acuity threshold as a function of pixel pitch and antialiasing filter width (pixels) for the observer model. Line width = 2.4 arcmin.

When little or no antialiasing is performed, the threshold curve is clearly non-monotonic with peaks occurring where the line width of the Vernier acuity stimulus (2.4 arcmin) is equal to integer multiples of the pixel pitch (e.g., 2.4, 1.2, 0.8, 0.6 arcmin). Distinct minima are present in the threshold function when the line width is midway between integer multiples of the pixel pitch.

Human Performance

The results of the second evaluation with human subjects are presented in Figure 3 (black lines) which shows the measured thresholds as a function of the 20 combinations of pixel pitch and antialiasing filter width. The standard errors of the mean for these data are well described as:

2) SEM = 0.25 + 0.071 * threshold

Thus, the confidence intervals for the human performance estimates clearly increase with increasing threshold. These human performance data are consistent with previous research showing that Vernier acuity thresholds in the range of 4 to 10 arcsec[16, 17] can be obtained on sampled display systems when sufficient antialiasing is applied. The results of two-sample T-tests indicated no significant differences ($p > 0.10$) between the thresholds for the four smallest pixel pitches when the antialiasing filter width was set at 1.2 pixels. The threshold obtained at the 2.4 and 3.0 arcmin pitch conditions are significantly higher ($p < 0.01$) than the thresholds for the four smaller pitch conditions. These results confirm the expectation that Vernier acuity thresholds do not rise significantly until the pixel pitch is greater than 1.6 arcmin when sufficient antialiasing is used.

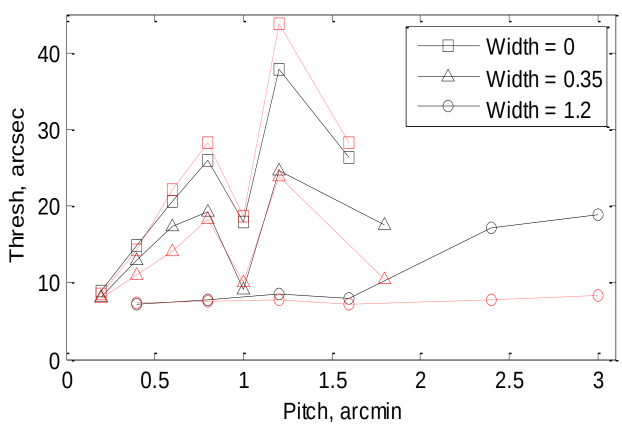

Figure 3. Human Vernier acuity thresholds (black) and observer model predictions (red) as a function of pixel pitch and antialiasing filter width (pix). Line width = 2.4 arcmin.

Correlation with Model

From Figure 3 it is apparent that the overall pattern of results obtained from the subjects is very similar to the observer model with the exception of the largest three pitch conditions. When the pixel pitch is greater than or equal to 1.8 arcmin, the humans did not perform as well as predicted by the observer model ($p < 0.025$, 10 df). The correlation between the human performance and the modeled data is $R^2 = 0.849$ ($p < 0.001$, 18 df). When the three largest pitch conditions are removed, the correlation between the human performance and the modeled data is $R^2 = 0.977$ ($p < 0.001$, 15 df).

In the human performance evaluation, the width of the (pre-antialiased) lines was held constant at 2.4 arcmin. With this strategy, the process of antialiasing and sampling at the pitch of the simulated display system held the total amount of contrast energy constant across all 20 experimental conditions. For most of the stimulus conditions, the peak contrast ratio of the stimuli was 14:1. For those few stimulus conditions for which the pixel pitch was greater than about half the line width, and the wider antialiasing filters were used, two qualitative changes to the stimuli occurred. Figure 4 illustrates these changes for the two largest pitch conditions (2.4 and 3.0 arcmin) for which the antialiasing filter width was set to 1.2 pixels. Comparison of Figures 4 and 2 reveals that the gap between the upper and lower lines becomes less distinct and that the peak contrast of the stimuli is reduced. These qualitative changes to the stimuli are not accounted for in our simple observer model and may explain why the model fails to predict performance for these conditions.

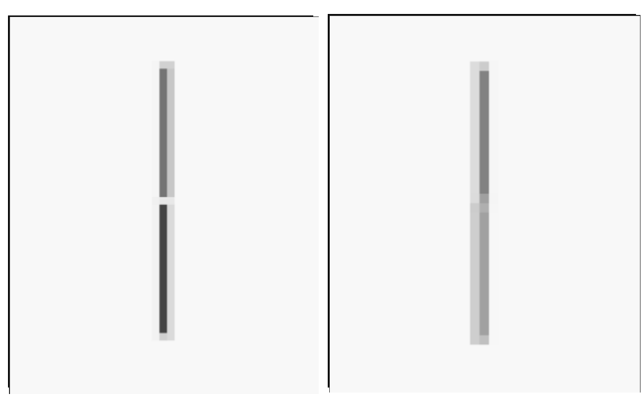

Figure 4. Examples of the Vernier acuity stimuli for the largest two pitch conditions. Left panel: pitch = 2.4, CR = 4.8. Right panel: pitch = 3.0, CR = 2.5.

Critical Antialiasing Filter Width

The results of this evaluation indicate the observer model is a reasonable approximation of human performance for pixel pitches of less than 1.8 arcmin. Thus, using the data plotted in Figure 2 it is possible to estimate the minimum antialiasing filter width required to avoid degrading Vernier acuity performance, as a function of pixel pitch. However, the data from this surface apply only for a line width of 2.4 arcmin. The specific pattern of peaks and valleys that occur in this surface appear to be related to the line width. To test this assertion, the same computational experiment was performed using different line widths and in each case the peaks in the surface occurred where the line width of the Vernier acuity stimulus was equal to integer multiples of the pixel pitch. For example, Figure 4 shows the surface for a line width of 1.9 arcmin with distinct peaks at 1.9 / 2 and 1.9 / 3 arcmin.

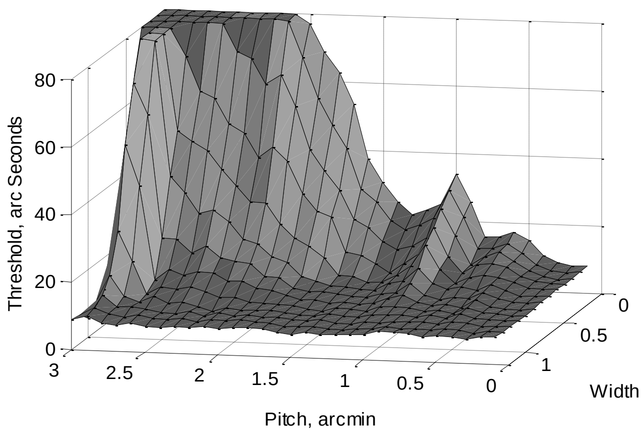

Figure 4. Vernier acuity threshold as a function of pixel pitch and antialiasing filter width (pixels) for the observer model. Line width = 1.9 arcmin.

For the purpose of providing more general design guidance, an estimate of the minimum filter width is needed that is independent of the specific line width used for the Vernier acuity evaluation. This estimate was derived by computing the response surfaces for five different line widths ranging from 1.2 to 3.4 arcmin and considering the worst case performance for each line width. The surface shown in Figure 5 was computed using the maximum of the five surfaces computed for each line width. The curve on the floor of this plot shows where the Vernier acuity threshold increases by approximately 10% above the floor of the combined surface.

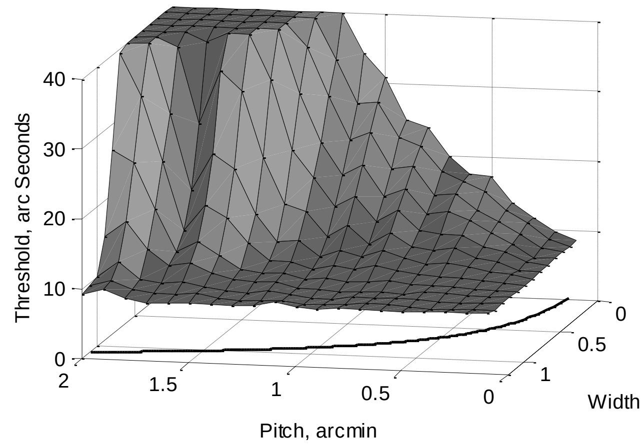

Figure 5. Vernier acuity threshold as a function of pixel pitch and antialiasing filter width (pixels) for the observer model. Line widths range from 1.2 to 3.4 arcmin.

The equation for the curve indicating the critical antialiasing filter width required to avoid degraded Vernier acuity tasks performance is:

3) $Wc = 0.92 + 0.62 * \log10(p - 0.12);$

Where Wc is the half-maximum width of Gaussian filter (pixels), and p is the pixel pitch (arcmin). This equation is considered valid for pixel pitches no greater than 1.6 arcmin and line widths no less than twice the pixel pitch.

CONCLUSIONS AND RECOMMENDATIONS

- The threshold data collected from five observers were highly correlated with the responses of the observer model. Thus, the model appears to be a valid representation of human performance for this task for a pixel pitch of less than 1.8 arcmin.
- As expected, the presence of spatial sampling artifacts seriously degrades Vernier acuity task performance:
  - When sufficient antialiasing was performed, eye-limited Vernier acuity thresholds (e.g., 7.5 arc seconds) were obtained with a pixel pitch as large as 1.6 arcmin.
  - When no antialiasing was performed, the pixel pitch required to achieve eye limited thresholds is less than 0.20 arcmin.
- When sufficient antialiasing is performed, the pixel pitch required (e.g., 1.6 arcmin) to achieve eye-limited Vernier acuity thresholds is significantly larger than the

- pitch required to achieve eye limited performance on other acuity-bound tasks (e.g., 0.7 arcmin) similar to simulation training tasks. Thus, performance on the Vernier acuity task is not a primary determinant of the resolution requirement for simulation training display systems.
- In general, performing sufficient antialiasing is far less expensive than decreasing the display pixel pitch enough to compensate for the detrimental effects of spatial sampling artifacts.

ACKNOWLEDGMENTS

The views expressed in this paper are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government. This paper has been cleared for public release 6/24/2015.

BIOGRAPHIES

Dr. Charles J. Lloyd is president of Visual Performance LLC where he addresses research and development challenges relating to training display system design, requirements, metrics, and measurements. He has 28 years of experience in display systems and applied vision research at such organizations as Honeywell's Advanced Displays Group, The Lighting Research Center, BARCO Projection Systems, FlightSafety International, and the Air Force Research Laboratory. Charles has published more than 75 papers in this arena.

Dr. Marc Winterbottom is a senior research psychologist supporting the Operational Based Vision Assessment Laboratory at the U.S. Air Force School of Aerospace Medicine, Wright-Patterson AFB, Ohio. His most recent research focuses on U.S. Air Force vision standards and modernization of vision screening practices. His prior research focus at the Warfighter Readiness Research Division, Mesa, Arizona, was on visual perception, particularly as it related to display technologies for simulation and training applications. He received an MS in Human Factors Psychology from Wright State University (2000) and a BA degree in Psychology from Purdue University (1996). He was awarded a DoD SMART Scholarship in 2010 and recently completed a PhD in Human Factors Psychology at Wright State University.

Dr. James P. Gaska received a Ph.D. degree in Biopsychology from Syracuse University and from 1981 to 2000 worked at UMASS Medical School where he used single cell electrophysiological and computational techniques to explore and model the representation of the world in the visual cortices of primates. From 1995 to present, Dr. Gaska worked as contractor to the Air Force and used vision science and visual performance models to aid in the design of camouflage, laser eye protection and flight simulation technologies. He is currently Senior Scientist for OBVA at the USAF School of Aerospace Medicine at WPAFB.

Dr. Logan Williams is a senior research engineer at the U.S. Air Force School of Aerospace Medicine at Wright-Patterson AFB, Ohio, and currently leads technology development for the Operational Based Vision Assessment laboratory. Previously, he has led multiple lines of research in various fields such as human effectiveness, immersive environments, visual display system design, and distributed simulation for aircrew training. He has served as the lead systems engineer for F-16, A-10, and KC-135 aircrew training systems and has two decades of experience in analog and digital circuit design, networked control systems, optical & electro-optical system design, computer programming, and physics-based modeling and simulation. He has earned a PhD in Electro-Optics, ME and BS degrees in Electrical Engineering, as well as a BS in Physics.

REFERENCES

[1] Lloyd, C. J., Winterbottom, M., Gaska, J., and Williams, L. (2015) A practical definition of eye-limited display resolution. Proceedings of the SPIE Defense, Security, and Sensing conference, Baltimore.

[2] Lloyd, C. J., Williams, L., and Pierce, B. (2011) A model of the relative effects of key task and display design parameters on training task performance. *Proceedings of the IMAGE 2011 Conference*, The IMAGE Society, Phoenix, AZ.

[3] Gaska, J. P., Winterbottom, M. D., Sweet, B., and Rader, J. (2010) Pixel size requirements for eye-limited flight simulation. Proceedings of the IMAGE Society Annual Conference.

[4] Hogervorst, M. A., Bijl, P., & Valeton, J. M. (2001, September). Capturing the sampling effects: a TOD sensor performance model. In *Aerospace/Defense Sensing, Simulation, and Controls* (pp. 62-73). International Society for Optics and Photonics.

[5] Lloyd, C. J., Williams, L., and Pierce, B. (2011) A model of the relative effects of key task and display design parameters on training task performance. *Proceedings of the IMAGE 2011 Conference*, The IMAGE Society, Phoenix, AZ.

[6] Masaoka, K., Niida, T., Murakami, M., Suzuki, K., Sugawara, M., and Nojiri, Y. (2008) Perceptual limit to display resolution of images as per visual acuity.

Human Vision and Electronic Imaging XIII, SPIE Vol. 6806

[7] Baxter, B. and Corriveau, P. (2005) PC display resolution matched to the limits of visual acuity. Journal of the Society for Information Display, 13 (2)

[8] Spencer, L., Jakobsen, M., Shah, S., and Cairns, G. (2013) Minimum required angular resolution of smartphone displays for the human visual system. Journal of the Society for Information Display, 21 (8)

[9] Lloyd, C. J. (2012) Effects of spatial resolution and antialiasing on stereoacuity and comfort. Proceedings of the AIAA Modeling and Simulation Conference, Minneapolis, MN.

[10] Jobs, S. (2010) Keynote Address at World Wide Developers Conference, San Francisco. http://events.apple.com.edgesuite.net/1006ad9g4hjk/event/index.html

[11] Holmes, R. (1997) Head-mounted display technology in virtual reality systems. In L. W. MacDonald & A. C. Lowe (Ed) Display Systems Design and Applications. Wiley, New York.

[12] Clapp (1987) Visual Simulation. Annual Simulation Symposium

[13] Hopper, D. G. (2000) 1000 X difference between current displays and capability of human visual system: payoff potential for affordable defense systems. Cockpit Displays VII. Proceedings of the SPIE Vol. 4022.

[14] Wright, S. L. (2002, April). IBM 9.2-megapixel flat-panel display: Technology and infrastructure. In SPIE Proceedings (Vol. 4712, pp. 24-34).

[15] Watson, A. & Pelli, D. (1983). QUEST: a Bayesian adaptive psychometric method. Perception & Psychophysics, 33 (2), pp. 113-120.

[16] McKee, S. and Westheimer, G. (1978). Improvement in vernier acuity with practice. Perception & Psychophysics, 24 (3), pp. 258 -262.

[17] Sullivan, G., Oatley, K., and Sutherland, N. (1972). Vernier acuity as affected by target length and separation. Perception and Psychophysics, 12 (5), pp. 438 – 444.