A MODEL OF THE RELATIVE EFFECTS OF KEY TASK AND DISPLAY DESIGN PARAMETERS ON TRAINING TASK PERFORMANCE

Charles J. Lloyd
L-3 Communications, Link Simulation & Training, Mesa, AZ

Logan Williams, Byron Pierce
711 HPW/RHA, Air Force Research Laboratory, Mesa, AZ

ABSTRACT

This paper describes the design and initial validation of a model of the relative effects of key task and display design parameters on training task performance. The display system design parameters include pixel hold time, pixel pitch, luminance, contrast, and noise. The task and observer parameters include target size and range, angular velocity (of the image), target contrast, and observer capability (e.g., acuity).

This model was developed as part of the Immersive Display Evaluation and Assessment Study (IDEAS) program for the USAF. The model was designed to be used by display system acquisition professionals who must develop defensible display system requirements that are based on the ability of the display system to support the training planned for the system. The model was also designed for those display systems professionals in the supplier community who wish to steer their product designs in the direction of maximum utility to the USAF.

The initial validation results indicate the model accurately summarizes the findings from several published evaluations that employed tasks generally representative of flight simulation training. The model was then used to make specific predictions of the effects of 220 combinations of four parameters (hold time, velocity, pitch, and luminance) on the range at which pilots could reliably identify fighter sized aircraft. A formal evaluation of the effects of these parameters was conducted and the results of this evaluation are compared with the predictions. A high correlation between model predictions and the results of this evaluation was obtained. The details of this evaluation are presented in a companion paper published at this conference.

INTRODUCTION

This paper describes a model developed for the purpose of characterizing the effects of a number of important task and display system design parameters on performance for simulation training tasks mediated by display resolution and observer acuity. Details of the model are provided along with the results of initial validation tests based on published data.

A companion paper at this conference[13] describes two laboratory validation experiments in which 420 combinations of five model parameters were exercised over a wide range of levels. The variables and their ranges were selected to cover the training display system design envelope expected to be of interest to the Air Force over the next few years.

An additional paper, to be published at the I/ITSEC 2011 conference[22], will illustrate how the model can be used within a decision support system (DSS) designed for acquisition professionals who prepare requirements and select suppliers and for product planners who would like to maximize the value of their product offerings.

Model Design Goals

The proposed model was developed under the Immersive Display Evaluation and Assessment Study (IDEAS) program of the Air Force Research Laboratory. The overriding goal of this program is to develop defensible requirements for training display systems that are based on the ability of these devices to support specific training needs. A general strategy used to achieve these goals is to collect data and develop models that quantitatively relate training task performance goals to well defined and measurable attributes of display systems. A long term goal of the program is to package these data and models into a decision support system (DSS) that can be used by both acquisitions professionals and the supplier community.

Early in the IDEAS program discussions were held with the customer to refine the scope for the initial phase of what is expected to be a multi-year program. High on the list of challenges to be addressed were the display parameters that mediate motion-induced blurring. The recent introduction of LCD, LCoS and DLP projectors into the simulation training industry revealed significant challenges with motion-induced blurring that were not present with the CRT-based projectors that had been used for several decades. As of early 2010, no motion related metrics or design recommendations could be found that had been validated using task performance.

Model Origins

The model described in this paper has a long history traceable to the image quality metric development work of many researchers. A good introduction to the early image quality research can be found in Biberman[4]. Side-by-side evaluations of the most important of dozens of metrics can be found in reports by Beaton and Task[2, 20]. More thorough discussions of the capabilities and limitations of the conventional metrics of the time can be found in Snyder and Lloyd[14, 16, 17]. Barten provides a more recent overview of the perceptually weighted metrics based on the modulation transfer function (MTF) while emphasizing the square root integral (SQRI) metric he developed and validated[1].

Of the dozens of image quality metrics developed over the years, the modulation transfer function area (MTFA) stands out as one of the most capable and well-tested. The MTFA metric was so successful in predicting task performance that it was selected as the acceptance standard for video display terminals by the American National Standards Institute and the Human Factors Society (HFS, 1988)[10].

The MTFA metric was first introduced by Charman and Olin in 1965 for use in evaluating photographic images[7]. While Charmen and Olin originally called their metric the threshold quality factor (TQF), the many researchers who followed have refere to it as the MTFA.

At the heart of the MTFA and related metrics is the measurement and computation of two related functions that are compared to determine the overall capability of the display-observer system. The MTF describes the display and the contrast threshold function (CTF) describes the observer.

Display Characterization

The capability of the display system under test is quantified by displaying an impulse signal (e.g., single line or edge) and measuring a line (or edge) spread function (LSF) that characterizes the ability of the display system to reproduce the impulse (see Figure 1).

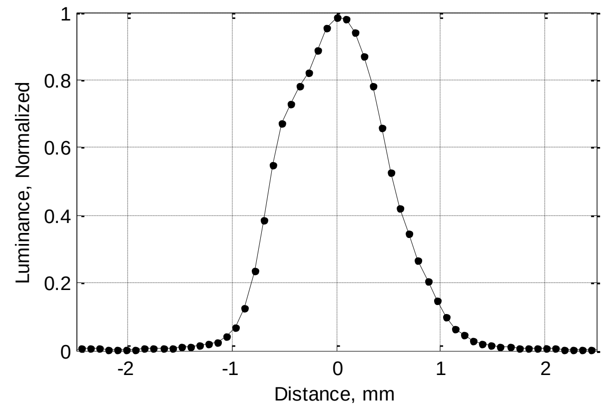

Figure 1. Measured line spread function (LSF) for a single pixel wide white line on a black background for an LCoS projector setup for a pixel pitch of 0.6 mm. The half-maximum width of this LSF was 1.3 mm. The most apparent causes of line spreading for this system were mis-convergence, optical blurring, and anti-aliasing.

The MTF of the display is computed from the LSF via Fourier transform. By transforming the LSF into an MTF (see Figure 2), the measurement data can be more directly compared with the capability of the observer as expressed in the form of a CTF as described below.

Observer Characterization

The contrast threshold model selected for our effort is based largely on the contrast sensitivity model of Barton[1]. Compared with the CTF model originally used by Charmin and Olin, the more recent Barten model explicitly accounts for three parameters of great interest to the display system designer: luminance, field size (number of cycles across target), and noise (pixel level). By coupling the more comprehensive Barten model to the MTF-based methods that have performed so well, our new model has the potential of allowing us to address a wider range of display design variables. Figures 3 and 4 illustrate the effects of luminance and field size on contrast thresholds as a function of spatial frequency.

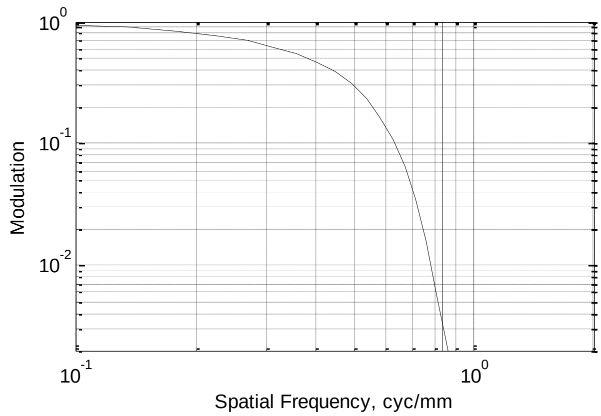

Figure 2. Modulation transfer function (MTF) computed from the LSF shown in Figure 1. The vertical line at 0.83 cyc/mm indicates the Nyquist sampling limit of the display.

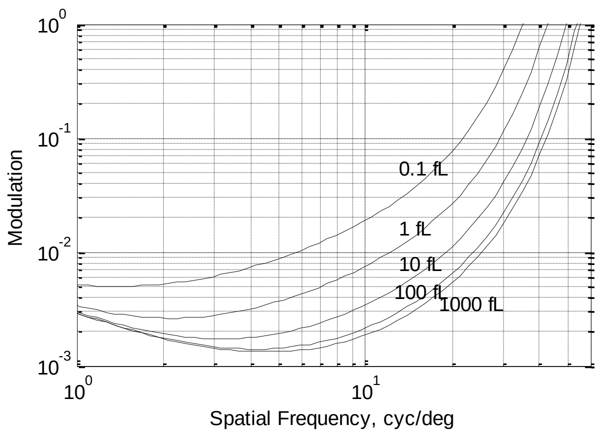

Figure 3. Contrast threshold as a function of spatial frequency and luminance for a constant field size of 6 deg. Correlation of Barten model[1] with the data of van Nes & Bouman is excellent, $R^2 > 0.95$.

Combined Display and Observer Functions

The MTFA and related metrics provide a concise and intuitively-appealing way to compare the capability of a display system with the needs of an observer. The display MTF shows the amount of modulation the display system is capable of producing at each spatial frequency or level of detail. The CTF indicates the minimum level of modulation required by the observer at each spatial frequency. The MTFA metric is computed as the area bounded by the MTF and CTF and thus it quantifies the modulation producible by the display system that is usable by the observer. In Figure 5, these calculations are illustrated for the display indicated in the Figures 1 and 2 at a luminance of 10 fL viewed from a distance of 2 m.

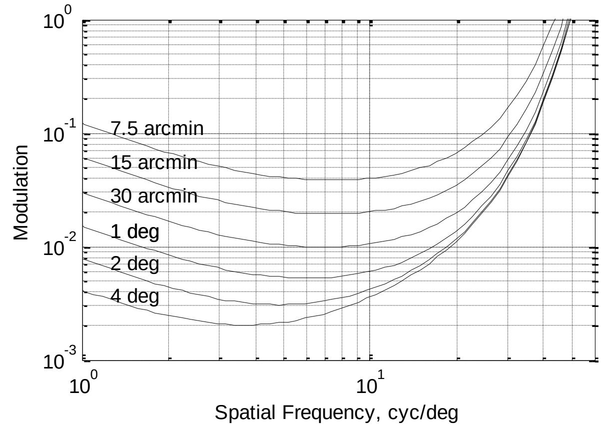

Figure 4. Contrast threshold as a function of spatial frequency and field size for a fixed luminance of 10 fL. Correlation of Barten model[1] with the data of Carlson is excellent, $R^2 > 0.90$.

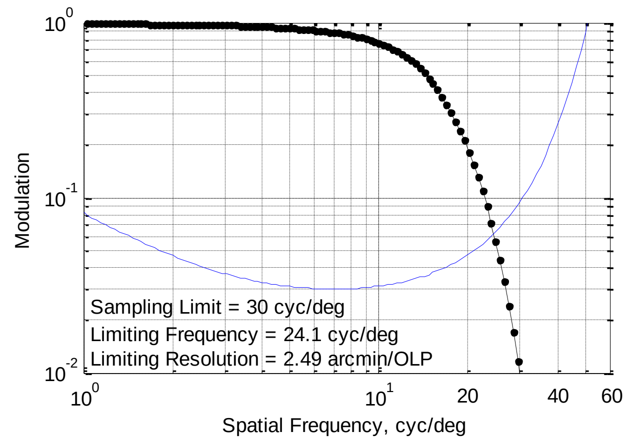

Figure 5. Display MTF and observer CTF for a viewing distance of 2 m producing an angular pixel pitch of 1 arcmin. The limiting resolution of the display-observer system is the point at which the curves cross.

In the proposed model, the limiting resolution of the display system is used to calculate the threshold angular target size which is used to determine the expected range at which targets of a specified size can be discriminated or identified. Our approach is apparently similar to the approach suggested by Streid[19]. However, the Streid paper provides few details and no discussion of implementation or validation of the approach.

Model Extensions

The model components described thus far (MTF and CTF) are identical to those used for computing the MTFA and SQRI metrics. This section describes a significant addition to these components that allows our model to address the challenge of motion induced blurring.

Motion induced blurring

Several recent papers[3, 8, 12, 18, 21, 23] describe essentially the same approach for modeling the detrimental effects of motion induced blurring. With this approach, it is assumed the observer accurately tracks the target and the spatial impulse response (SIR) function of the display system is measured or computed. The SIR can be acquired directly using a moving impulse (or edge) signal that is measured using one of the several pursuit camera techniques currently available. Alternatively, the SIR can be computed from the temporal impulse response (TIR) and velocity or from the velocity and the temporal edge response convolved with the pixel hold time. Measurement of the TIR function is generally less complex and requires less expensive equipment than direct measurement of the SIR function for moving stimuli. In our model, the net effect of the static LSF and motion induced blurring is accounted for using the product of velocity and the TIR convolved with the static MTF of the display system.

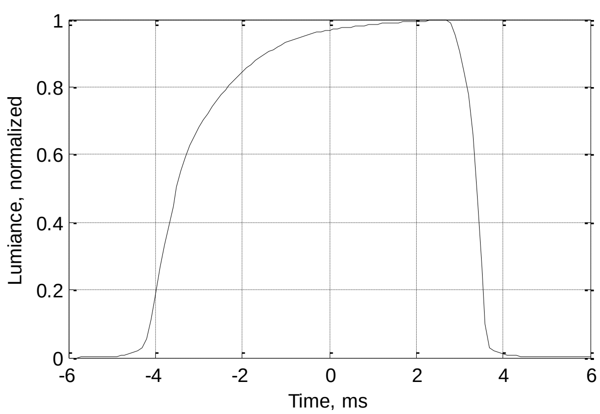

Figure 6. Measured temporal impulse response (TIR) function for LCoS projector with an external LCD motion blur reduction shutter. Hold time (half-maximum) for this example is 7 ms.

For the TIR in Figure 6 and an angular velocity of 5 deg/sec, the SIR would have the same shape but would be scaled to a half-maximum width of 5 * 0.007 * 60 = 2.4 arcmin. When viewed from a distance of 2 m the half-maximum width would be 1.4 mm which is about the same as the static LSF for this display. The static LSF is combined with the SIR by convolution. The net result of the convolution is illustrated in Figure 7 for four target velocities.

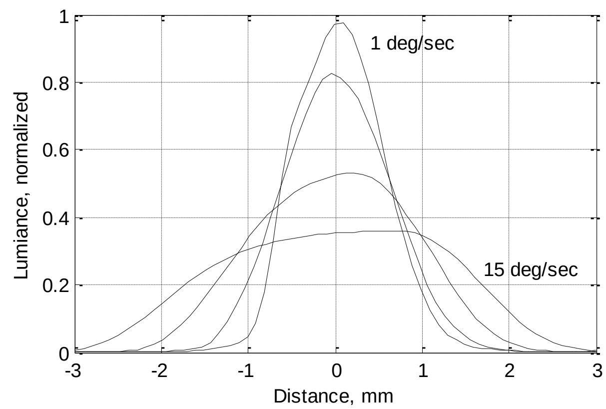

Figure 7. Effective LSFs for target velocities of 1, 5, 10, and 15 deg/sec at a viewing distance of 2 m. Analysis applies only along the direction of image motion and assumes the observer accurately tracks the target.

For low target speeds the effective LSF is nearly identical to the static LSF as can be seen by comparing the 1 deg/sec curve in Figure 7 with the curve in Figure 1. As target speed is increased to the point where the SIR is about the same width as the static LSF, the effective LSF becomes more Gaussian in shape and has a half-max width that is 20 to 25% larger than the widths of the LSF or SIR (see the 5 deg/sec curve in Figure 7). As speed increases beyond this point, the shape of the combined LSF approaches the shape of the TIR as can be seen by comparing the 15 deg/sec curve of Figure 7 with the curve in Figure 6.

Asymmetrical Blurring

Motion induced blurring is a one-dimensional phenomena that occurs only along the direction of motion. The amount of motion blurring can be much greater than the blurring perpendicular to the motion, thus, highly asymmetrical blur functions are produced for fast target motions.

To date we have found no papers that provide data or theoretical insight on how we might model the effect of highly asymmetrical blurring on task performance. In general, we expect task performance will decline monotonically with increasing target velocity. However, since the motion induced blurring occurs in only one direction, we expect the degradation due to motion to be less than it would be if the motion induced blurring were applied in both dimensions (isotropic). We thus settled for the inclusion of a parameter in our model that scales the angular velocity of the target using a power function, the exponent of which we determined empirically.

In our model, the effective blur function of the system is the product of velocity and TIR taken to a power P convolved with the static LSF of the display system.

$$LSF_{eff} = (vel \times TIR)^{P} \bullet LSF_{static}$$

The exponent P allows control over the magnitude of the degradation of task performance due to motion induced blurring. We expect a setting of 1.0 to correspond to the case where motion induced blurring occurred equally in two dimensions, as it does for typical optical blur. Reducing the exponent below 1.0 has the effect of reducing the amount of effective blurring attributable to motion. For the case of blurring applied along a single dimension, we expect performance to degrade less than if the blurring had been applied in two dimensions. Prior to conducting the laboratory evaluations, we had little basis for setting this parameter. Optimizing the model performance using the laboratory data produced an exponent of 0.66.

Scaling supra-threshold modulation

In 1985, Snyder et. al., conceded that there was evidence the modulation just above the threshold contrast curve seemed to have more influence on performance than did modulation at higher levels[17]. In other words, these authors expected the performance of the metric would be improved if a nonlinear scaling of supra-threshold contrast were applied. The non-linear scaling of supra-threshold modulation was entierely consistient with the works of Carlson and Barten who both used square root weighting[1, 6]. In 1990, Lloyd adopted a nonlinear scaling of modulation and demonstrated that his model accurately predicted supra-threshold contrast discrimination thresholds and magnitude estimates of contrast over a range of spatial frequencies[14].

In the model described in this paper, supra-threshold modulation is non-linearly scaled such that the user can specify the number of just noticeable differences (JNDs) above threshold the modulation needs to be at the crossover point that defines the limiting resolution of the display-observer system.

Generalizing Across Training Tasks

Most of the parameters discussed thus far describe attributes of the display system (pixel pitch, hold time, luminance, contrast, and noise). Once a display system is specified and designed, the settings for these parameters are relatively permanent attributes of the device.

There are four additional parameters associated with this modeling effort that are better described as attributes of training tasks: target size, velocity, contrast, and luminance ratio. Estimates of the settings of these parameters are required for the computation of expected task performance. Unlike the display parameters, the settings of the task parameters are not fixed with the display design. The inclusion of these task parameters in the model significantly increases the range of tasks the model can potentially describe.

Target contrast and relative luminance

Within the model, the computation of expected performance is dependent on *target* luminance and contrast. However, the modeling system is more immediately useful if acquisitions and design professionals are allowed to work directly with the *display* luminance and contrast. Thus, it is important to define and discriminate between these display and target level parameters.

For the example that follows, let L_d refer to the peak luminance and CR_d the maximum contrast ratio (e.g., checkerboard) of the display system. For the following example, assume L_d = 10 fL and CR_d = 8. From these quantities, we can estimate the amount of "washout" luminance created by various sources that produce the unavoidable scattered light within practical display systems. A first order approximation of the washout lighting can be computed as follows:

$$L_w = L_d / (CR_d + 1)$$

For our example, the washout lighting L_w would be 10 / 9 = 1.11 fL. Subtracting the washout luminance from the peak display luminance leaves L_{max} = 10 – 1.11 = 8.89 fL which represents the maximum luminance of the display system if washout lighting were not present.

Target contrast refers to the desired contrast of the target, that is, the contrast the model designer had assumed the display system would produce. Target luminance ratio refers to the luminance of the target relative the brightest part of the scene (peak display luminance).

For example, assume the "real world" contrast of a fighter aircraft against the sky with light haze is 2:1(CR_t = 2) and the luminance of the pale blue sky is set to 80% of the peak luminance of the display system, thus the luminance ratio of the target would be 40% of the peak display luminance. Assume also that accurate models of the aircraft and atmospheric effect are used and accurate display gamma correction is employed. With these assumptions, the luminance of the background (hazy sky) against which the target is viewed is:

$$L_b = 0.80 * L_{max} + L_w = 8.22$$

And the luminance of the target is:

$$L_t = 0.80 * L_{max} / CR_t + L_w = 4.67$$

The actual contrast of the target produced in the practical display system is $L_b / L_t = 1.76$.

Repeating this calculation for a range of display contrast ratios produces the upper curve shown in Figure 8. This curve shows display contrast has a strong effect on target contrast for low display contrast ratios. For higher display contrast ratios the target contrast is much less dependent on display contrast.

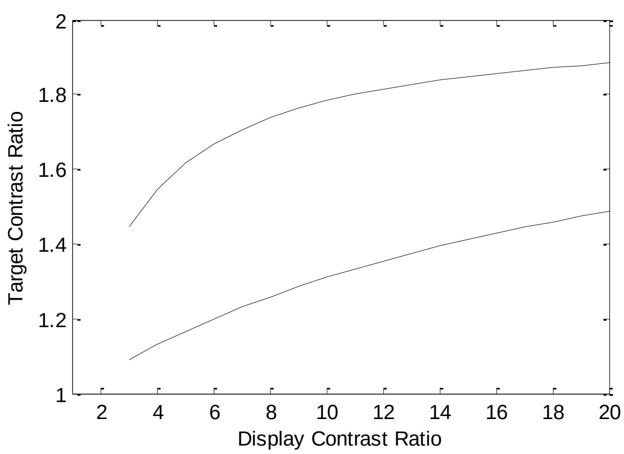

Figure 8. Effect of display contrast ratio on target contrast for targets with 2:1 contrast. Upper curve assumes the target is at 40% of the peak display luminance. The lower curve assumes the target is at 5% of the peak display luminance.

The upper curve in Figure 8 implies there would be relatively little value in increasing the display contrast above 8 or 10:1. However, practical experience with training display systems suggests that the unavoidable scattered light washes out contrast in the dark portions of the scene more than it does the bright. To assess this expectation, the calculation was repeated assuming a target luminance of 5% and a background luminance of 10% of peak display luminance. As expected, the lower curve in Figure 8 shows a stronger and more linear dependence of target contrast on display system contrast.

VALIDATION WITH PUBLISHED DATA

In the previous section the model components inherited from previous image quality metrics were shown to correlate with key data sets from the vision science literature and with the data collected in metric validation studies. Given this heritage, we would expect the model to correlate well with performance on tasks more similar to those performed in simulation training. To test this assertion the model was used to make predictions that were compared with the data from the limited number of published evaluations we could find that employed task performance measures.

Effect of display pitch

The pixel pitch of a display system is a design parameter that has a very clear effect on the cost of the system since the number of display and image generator pixels varies with the inverse of the square of pitch. Since pitch is such a strong cost driver, one would hope the proposed model would be able to predict task performance as a function of pitch. The effect of pixel pitch on task performance has been studied in two recent efforts that provide data that can be used for testing.

As part of their determination of the requirements for an "eye limited" display system for the Operational Based Visual Assessment (OBVA) program, Gaska et. al. measured the threshold angular size at which observers could just discriminate the orientation of triangular stimuli[9]. Observers viewed a 53 fL projected image from a distance of 6 m for a native pixel pitch of 0.11 arcmin. The targets were dark on a light background with a contrast ratio of 3.33. Ten pitch conditions were produced using pixel replication. The resulting threshold triangle size is plotted as a function of pixel pitch in Figure 9 (triangles). The solid curve going through these data was computed using the proposed model. One degree of freedom, the acuity of the observer, was fit to the data. The correlation between the model predictions and the data was $R^2 = 0.991$ ($p < 0.001$, 8 df).

In early 2010, Lloyd independently conducted a similar evaluation in which the threshold for the detection of Landolt C target orientation was measured as a function of pixel pitch. Using a DLP projector that produced a small rear-projected image, a 1000 fL image was produced with a native resolution of 0.15 arcmin and a contrast ratio greater than 30:1. Threshold orientation discrimination performance was measured as a function of 11 pitch conditions which were produced by pixel replication. The data from this evaluation is plotted as circles in Figure 9. The curve through these data indicates the fit of the proposed model. One degree of freedom, the acuity of the observer, was fit to the data.

The correlation between the model predictions and the data was $R^2 = 0.967$ ($p < 0.001$, 9 df).

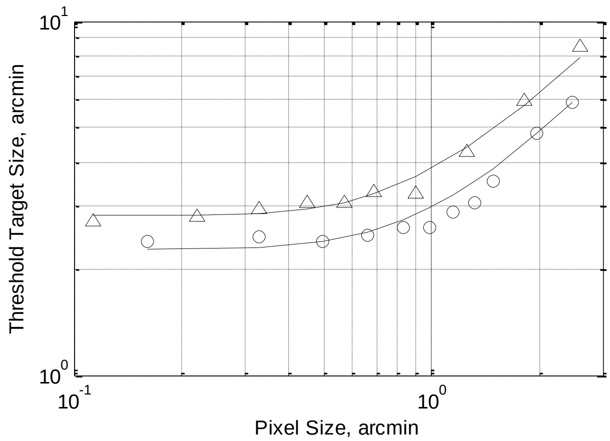

Figure 9. Threshold target size as a function of display pixel pitch for triangle and Landolt C orientation discrimination tasks. The smooth curves through the data are predictions based on the model.

For a pixel pitch greater than about 1 arcmin, these data show threshold target size is proportional to pixel pitch. In other words, the observer needs some minimum number of pixels across the target to accomplish the task. For pixel pitches below about 0.6 arcmin, pitch has no effect as performance is limited not by the display system but by the capability of the observer.

Effects of Blur and Noise

Three decades ago, Snyder, Beaton, and others at the Displays and Controls laboratory at Virginia Tech completed an extensive R&D program for the Office of Scientific Research (USAF) addressing the utility of high resolution aerial photography[15]. For this study, a set of 250 calibrated images was prepared with controlled amounts of blur and noise. A group of 15 senior photo interpreters (PIs) served as observers. The PIs completed two separate tasks using the images from this database. In the Information Extraction task, the PIs had to answer a series of specific questions regarding essential elements of information (EEIs) in each image. In the Quality Rating task, the observers used a standardized NATO rating scale which ranged from 0 (totally un-interpretable) to 9 (permits detailed analysis and interpretation).

For half the trials the images were examined on a high resolution monochrome CRT display. Since Snyder et al. provide a detailed description of the blur and noise conditions they evaluated, we can make specific predictions using our model and compare our predictions with the data they collected. In this evaluation the observers were free to select the magnification they used for viewing the images. Since the selected magnifications were not reported, we optimized this parameter in our model to maximize the fit to the data. The fitting process produced an image height of 24 deg which is about the height of the typical desktop monitor in use today. The predictions produced by our model are compared with the data collected by Snyder et al in Figure 10 below.

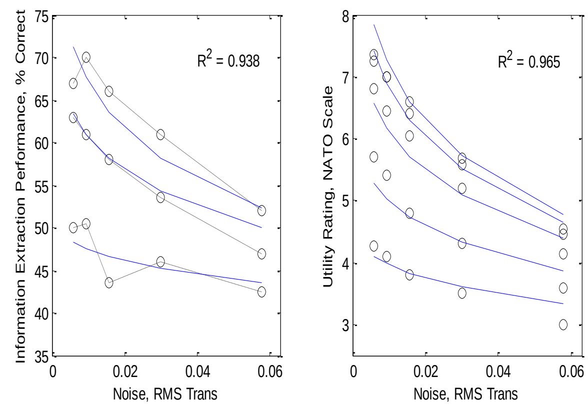

Figure 10. Mean information Extraction Performance (left panel) and mean Utility Rating for the 25 combinations of blur and noise introduced into the images. Open circles represent the data from Snyder (1983) and the blue solid lines indicate the model predictions.

The correlation between the model predictions and the data was $R^2 = 0.938$ ($p < 0.001$, 13 df) for the information extraction task and = 0.965 ($p < 0.001$, 23 df) for the utility ratings.

Effects of Resolution, Luminance, and Contrast

Of the evaluations we have identified thus far, the one that employs a task most representative of simulation training is by Kennedy et. al.[11]. In this evaluation, aircraft aspect detection range was measured as a function of 36 combinations of three practical display variables: resolution (pitch), target luminance, and background luminance (contrast).

High resolution targets were created using a target projector, the resolution (pitch) of which was varied using the zoom setting of the projection lens. The background scene was created using an independent projector which provided a means of independently manipulating background luminance and thus contrast. Target luminance was varied between 0.29 and 3 fL and background luminance was varied between 0.13 and 0.58 fL. Four levels of resolution (1.0, 1.3, 1.6, and 1.9 arcmin/OLP) were used in the evaluation.

Four observers participated in a two-alternative, forced-choice task in which they indicated the aircraft was climbing or diving. Each observer provided a threshold via a 100-trial staircase procedure for each of the 36 combinations of the independent variables. The correlation between our model predictions and the data from this evaluation was $R^2 = 0.923$ ($p < 0.001$, 34 df).

Summary

The proposed model shows strong predictive capability for each of the four task performance measures described in this section. However, this testing is considered insufficient as three of the four tasks are not particularly representative of flight simulation training. While the aspect detection range task used by Kennedy et. al., is considered representative, this study is considered a weak test of our model because the range of the design variables is unrepresentative of modern training display systems. In the Kennedy evaluation, luminance levels were very low and the contrast and contrast polarity were atypical. In the highest contrast condition, the aircraft were 23 times brighter than the sky against which they were viewed. Despite these shortcomings, this evaluation was considered practical and significant enough that Brown and Brunderman picked it up and converted it into a set of design tables for use by AF planners[5].

Model Scope

We expect the model described here to correlate with performance for simulation training tasks involving target orientation, discrimination, recognition, or identification against uncluttered backgrounds. We expect the model to cover combinations of target size and range that produce visual angles greater than a few arcmin. The model is intended for use with daylight scenes and display systems capable of producing at least a few fL and contrast ratios of at least 3 or 4:1. We expect reasonable performance for image velocities up to 40 to 50 deg/sec

The model is not expected to perform well for tasks involving visual search or target detection. Visual search and detection performance is highly sensitive to the area to be searched and the probability a target is present and may not be particularly sensitive to display resolution.

This model was not designed to address night or twilight scenes made up of very high contrast light points and objects set against very dark backgrounds.

We do not recommend the model be used for the determination of *absolute* levels of identification range. We expect absolute performance to be strongly dependent on a variety of variables not modeled here including: stimulus duration, observer capability, practice, vibration, and target characteristics. However, we expect these variables will exhibit little influence on the *relative* effects of the design variables.

IDEAS LABORATORY EVALUATIONS

To date, few published evaluations have been found that use task performance to assess the effects of the display design variables of interest. For this reason, two laboratory evaluations were designed for the purpose of generating more comprehensive data for model validation and for tuning the model parameters. These evaluations are described in the companion paper presented at this conference.

In the first laboratory evaluation, aircraft identification range was measured as a function of 220 combinations of 4 variables: pixel hold time, target velocity, pixel pitch, and luminance. The correlation between predictions made using the initial version of the model and the mean data of a group of AF pilots was statistically significant ($R^2 > 0.75$, $p < 0.001$, 109 df). After tuning the parameters of the model to these data, the correlation increased significantly ($R^2 > 0.973$, $p < 0.001$, 106 df).

In the second laboratory evaluation, identification range was measured as a function of 200 combinations of contrast, luminance, pixel hold time, target velocity, and pixel pitch. The correlation between predictions made using the model (optimized using the data from Evaluation 1) and the mean data of this group was $R^2 = 0.911$ ($p < 0.001$, 199 df). The tuning of the final model to the data of both evaluations has not yet been completed.

CONCLUSIONS

- The proposed model is an extension of the MTFA, SQRI, and JNDA image quality metrics which have been shown to produce high correlations with task performance in previous validation studies.
- The model uses the CTF model of Barten which has been shown to produce high correlations with several published data sets from the vision science literature.
- Model predictions are highly correlated with task performance data from four published evaluations employing practical tasks more similar to simulation training tasks than the vision science studies.
    - Photo-interpreter ratings of image utility
    - Photo-interpreter information extraction scores
    - Aircraft aspect detection thresholds
    - Triangle orientation detection thresholds
- A more comprehensive test of the predictive capability of the model is described in the companion paper presented at this conference.

AUTHOR BIOGRAPHIES

Dr. Charles J. Lloyd has 25 years of experience in display systems and applied vision research at such organizations as the Displays and Controls Lab at Virginia Tech, the Advanced Displays Group at Honeywell, Lighting Research Center, Visual Performance Inc., BARCO Projection Systems, and FlightSafety Int. Charles is now the Lead Scientist for the IDEAS program at the Air Force Research Laboratory (L-3 Communications) where he manages the development and validation of display system metrics for simulation training. Charles has presented 65 papers in this arena.

ACKNOWLEDGEMENTS

We wish to acknowledge the following people for their significant contributions to the IDEAS program.

- Jim Evans, ASC/ENDR
- Tim Dwyer, ASC/WNSEA
- Deke Joralmon, L-3 Communications
- Jim Basinger, Superior Technical Services
- James Gaska, L-3 Communications

REFERENCES

[1] Barton, P G. J. (2000) Contrast sensitivity of the human eye and its effects on image quality. SPIE Press.

[2] Beaton, R. J. (1984) A human-performance based evaluation of quality metrics for hard-copy and soft-copy digital imaging systems. Dissertation, Virginia Polytechnic Institute and State University, Blacksburg.

[3] Becker, M. E. (2008) Evaluation of moving-line contrast degradation without motion. Soc. for Info. Display Digest of Tech. Papers.

[4] Biberman, L. M. (1973) Perception of displayed information. Plenum Press, New York.

[5] Brown, L. and Brunderman, J. (1985) Flight simulator visual display requirements for aerial combat maneuvering. AFHRL-TR-85-39, Brooks AFB.

[6] Carlson, C. R., and Cohen, R. W. (1978) Visibility of displayed information. Arlington, VA: Office of Naval Research, Technical Report ONR-CR213-120-4F.

[7] Charman, W. N., and Olin, A. (1965) Tutorial: Image quality criteria for aerial camera systems. *Photographic Science and Engineering, 9,* 385-397.

[8] Feng, X., Pan, H., and Daly, S. (2008) Comparisons of motion-blur assessment strategies for newly emergent LCD and backlight drive technologies. Journal of the SID, 16. pp. 981-988.

[9] Gaska, J. P., Winterbottom, M., Sweet, W., and Rader, J. (2010) Pixel size requirements for eye-limited flight simulation. IMAGE Society Annual Conference, Scottsdale, AZ.

[10] HFS (1988) *American National Standard for Human Factors Engineering of Visual Display Terminal Workstations.* Santa Monica, CA: The Human Factors Society.

[11] Kennedy, R. K., Berbaum, K, S., and Collyer, S. C. (1988) Spatial requirements for visual simulation of aircraft at real-world distances. *Human Factors, 30(2),* pp. 153-161.

[12] Klompenhouwer, M. A. (2006) Comparison of LCD motion blur reduction methods using temporal impulse response and MPRT. SID Digest of Tech. Papers.

[13] Lloyd, C. J., Joralmon, D., Amann, R, Tai, C., Williams, L., and Pierce, B. (2011) Relative effects of five display design variables on aircraft identification range in daylight. *Proceedings of the IMAGE 2011 Conference,* The IMAGE Society, Phoenix, AZ.

[14] Lloyd, C. J. and Beaton, R. J. (1990) Modeling suprathreshold visual responses for image quality evaluations of color displays**.** Proceedings of the SPIE's 43rd Annual Conference. The Society for Imaging Science and Technology, Springfield, VA

[15] Snyder, H. L (1983) Quality metrics of digitally derived imagery and their relation to interpreter performance: VIII Final Report. HFL-83 -1. USAF Office of Scientific Research. Bolling AFB, Washington D.C.

[16] Snyder, H. L. (1985) Image quality: Measures and visual performance. In L.Tannas Ed., *Flat-panel displays and CRTs*, vanNostrand Reinhold, New York.

[17] Snyder, H. L. (1988) Toward the determination of electronic display image quality. In W. Rouse (Ed.) Advances *in man-machine systems research, Vol 4.* pp. 54-69. JAI Press.

[18] Someya, J. and Sugiura, H. (2007) Evaluation of liquid-crystal-display motion blur with moving-picture response time and human perception. Journal of the SID.

[19] Streid, H. (2007) Eye limited resolution displays: On the cusp? *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference.*

[20] Task, H. L. (1979) *An evaluation and comparison of several measures of image quality for television displays.* Wright-Patterson AFB, Report AMRL-TR-79-7.

[21] Watson, A. B. (2009) Comparison of motion-blur measurement methods. SID Digest of Technical Papers.

UPCOMMING REPORTS

[22] Lloyd, C. J., et. al. (2011) Towards a decision support system for simulation training display requirements. To be presented at I/ITSEC 2011.

[23] (2011) Display Measurement Standard, International Committee for Display Metrology. Society for Information Display. Release anticipated June 2011.