RENDERING HIGH QUALITY LIGHTPOINTS on FIXED MATRIX DISPLAYS

Dr. Charles J. Lloyd
Principal Staff Engineer
FlightSafety International, Visual Simulation Systems

Abstract

This paper summarizes three human factors evaluations of lightpoint quality for fixed matrix displays. Lightpoint quality was evaluated as a function of four practical design variables: lightpoint width and motion and display pitch and fill factor. The results showed that all of these variables except fill factor had strong and statistically reliable effects on lightpoint quality.

The last evaluation employed both the FAA/JAA "lightpoint size" and "Vernier resolution" tests. Results indicate a display pitch that just barely passes the lightpoint size test will clearly pass the Vernier resolution test, indicating resolution is mediated by the lightpoint size test. These evaluations reveal enough ambiguity in the FAA/JAA lightpoint size test that the 5th percentile observer may accept a pixel pitch that is 40% larger than the pitch accepted by the 95th percentile observer. A 40% difference in pitch translates into a 2:1 difference in the number of pixels, projectors, IG channels, and thus $ required for the display system. Several means of reducing this variance are proposed.

Introduction

In principle it seems possible to produce acceptable lightpoints using fixed matrix displays, assuming resolution, contrast, and luminance are high enough. Accepting this, the practical engineering questions become:

- How much resolution is required?
- To what degree do lightpoints need to be anti-aliased?
- How much brighter do lightpoints need to be above the background image?

It is assumed here that the resolution requirement should be driven by the goal that lightpoints do not change brightness, size, shape, or relative positions as they are moved with respect to the pixel structure of the display system. This goal is essentially a restatement of the primary goal of antialiasing, thus, the concepts of lightpoint size and shape, display resolution, and antialiasing are inseparable in a practical sense.

The third question drives display design decisions related to the contrast ratio and peak luminance capabilities of the display system. This research question will be addressed in a future evaluation.

Terminology In this paper lightpoint width is discussed both in terms of arcmin and number of pixels. To avoid confusion, lightpoint width expressed in arcmin will be called "widthA" while the width expressed in number of pixels will be called "widthP." Both widthA and widthP are measured from the half-maximum points of the Gaussian-like luminance distribution defining the lightpoint.

Presented at the IMAGE 2007 Conference, Scottsdale, AZ, July, 2007

Lightpoints ala Raster-Calligraphic CRT

Over the past few years the size of calligraphic lightpoints has been carefully measured in display systems employing FlightSafety as well as competing raster-calligraphic CRT projectors (Long, 2006; Lloyd et. al., 2006).

In the more competitive 3 and 5 channel collimated display systems used in the flight simulation training industry, the smallest (dimmest) lightpoints are approximately 2.5 arcmin while the largest (brightest) lightpoints are approximately 4.0 arcmin in width. These numbers apply for red and green lightpoints and for well-converged white lightpoints.

Blue CRTs are often purposefully de-focused a bit to produce wider calligraphic (and raster) blue lightpoints as this serves to stabilize color rendering for the display and the blue component of the image has almost no effect on resolution.

The half-maximum width of single pixel horizontal and vertical lines drawn with raster scanned CRTs is typically a bit larger than the width of calligraphic lightpoints. This is because the electron beam current is much higher in raster mode. While the width of raster lines is only slightly larger than calligraphic lightpoints, most would consider the effective "resolution" of calligraphic mode to be much higher than raster mode for two reasons. First, to draw fully anti-aliased lightpoints in the raster mode the half-maximum width of the lightpoint must be at least 1.5 to 2 times the pixel pitch. For a display system with a 3.2 arcmin pitch, the minimum sized raster lightpoint would be 5 to 6.5 arcmin wide. The second reason calligraphic lightpoints are so effective is their peak luminance can be *10 to 20 times higher* than the peak luminance of the raster.

Lightpoints ala JAA and FAA

Close examination of the regulations reveals the JAA/FAA "lightpoint size test" does not control lightpoint size, rather it describes the maximum *spacing* at which a horizontal string of lightpoints will be "discernable." The more stringent JAA requires a maximum spacing of 5 arcmin while the FAA requires only 6 arcmin. Given CRT calligraphic lightpoint widths are in the 2.5 to 4 arcmin range, the lightpoint "size" test has historically been easy to meet with calligraphic CRT projectors.

Lightpoints in the Real World

In general, lightpoints as observed from a commercial aircraft at cruising altitude are much smaller than can be produced by any modern display technology. From an altitude of 8000 m (27000 ft), a 15 cm (6 in) diameter street light diffuser subtends 4 arcsec. The nominal resolution often cited for human vision is about 1 arcmin which is about 15 times larger than the street light seen from altitude.

As a point of reference, a 15 cm wide light viewed at 0.52 km (1700 ft or 1/3rd mile) subtends 1.0 arcmin. The 2.5 to 4 arcmin lightpoints typical of calligraphic CRT projectors correspond to a 15 cm wide light at distances of 210 and 130 m. Thus, the size of these lightpoints would be correct only within the last few seconds of an approach and landing.

Practical Lightpoint Requirements

It is obviously important that a display system be capable of producing lightpoints that meet the JAA/FAA requirements. It is quite desirable that lightpoints be as small as calligraphic CRT lightpoints as the flight simulation training industry has three decades of experience with the calligraphic CRT.

Note that the widthA of the lightpoint (in arcmin) is *not* the variable of interest in these evaluations. Given that real world lightpoints are much smaller than can be produced on any known display system, the practical design goal is to make them "as small as possible" while maintaining acceptable image quality. The evaluations described in this paper deal with those display design variables that affect lightpoint quality.

Evaluation 1: Lightpoint Width, Display Pitch, and Fill Factor

Display Design Variables

This evaluation was designed to reveal the effects of three fixed matrix display (FMD) system design variables on three dimensions of image quality: Lightpoint widthP (pixels), Display Pitch (arcmin), and Display Fill Factor.

Lightpoint WidthP refers to the number of pixels used to create the lightpoint and is measured across the half-maximum points of the Gaussian-like profile which defines its shape. Six levels of lightpoint widthP were used in the evaluation: 0.90, 1.08, 1.27, 1.46, 1.67, and 1.90 pixels/lightpoint.

Display Pitch refers to the pixel-to-pixel spacing of the display system and is measured in arcmin. Five levels of display pitch were used: 1.00, 1.39, 1.81, 2.28, and 2.80 arcmin/pixel. The high end of this range approaches the pitches that have been sold into the simulation training market over the past few years. The center of this range approximates the pitch required for the most demanding display systems sold today, and the small end of this range approaches the pitches required to achieve "eye limited" resolution.

Fill Factor refers to the visibility of the spaces that occur between columns and rows of pixels within the projected image. This variable is sometimes called "aperture ratio" and is measured as the percentage of the pixel area (defined by pixel pitch) that transmits light. Two levels of fill factor were used: 0.78 and 0.92. The lower level corresponds to the fill factor that can be expected for light valve projectors of the transmissive, high temp pSi technology. The higher level is more representative of LCoS or DLP projectors for which the column and row conductors are positioned behind the active area of each pixel.

Experimental Design

A full-factorial, within-observer, experimental design was used for this evaluation, meaning that all combinations (6 x 5 x 2 = 60) of each of the three experimental variables was evaluated by each observer. Each observer was presented with the conditions in a different random order so that any unavoidable noise or drift in observer ratings, such as those caused by practice and fatigue, would be distributed randomly throughout the data and would not bias the results.

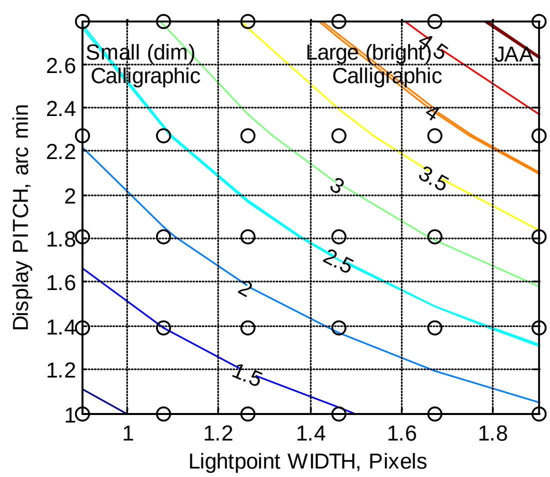

Figure 1. Design matrix showing the 30 combinations of WidthP and Pitch evaluated. Each combination was used for two levels of display Fill Factor for a total of 60 conditions. Contour lines show the resulting widthA (arcmin). The heavy contour lines show where large and small calligraphic lightpoints would fall and the upper right most contour shows the approximate JAA requirement.

Equipment and Software

Images were presented on a 53 cm diagonal (21 in) direct view color CRT of the RGB triad shadow mask design. The CRT was used to simulate the pixel structure of a FMD by using a 6x6 array of CRT pixels for each FMD pixel. Modulation produced by the inter-pixel column and row conductor lines of the FMD was simulated by darkening every 6th column and row of CRT pixels by the amount required to simulate the Fill factor of the display. The pixel pitch of the CRT was 0.32 mm/pix, thus, the pitch of the simulated FMD was 1.91 mm/pix. Viewing distance was used to control the angular subtense of the FMD pixels.

Images were prepared using the MATLAB and the Image Processing Toolbox products from The Math Works. Video images were generated using an Nvidia 7600 graphics card installed in a Windows personal computer. The display resolution was set to 1280 x 1024 pixels at a refresh rate of 60 Hz. The gamma setting on the graphics card was set to 1.

The electro-optical response of the display system was measured using a Minolta LS-100 meter. Gamma correction of the system was performed in software rather than relying on look up tables on the graphics card. A variation of Roberts method (Schreiber 1986) of adding uniform noise to the image was used to avoid gray scale banding. Linearity (luminance as a function of image level) was better than 2% at all gray levels. Peak luminance of the display was 17.8 fL, the luminance of the black background was 0.11 fL, for a CR of 160.

For each trial, a line of six lightpoints was presented and these lightpoints slowly orbited on the display while maintaining a straight line. This movement was designed so that the lightpoints moved at different speeds and the angle of the line of lights varied with respect to vertical. The angular subtense of the set of six lightpoints was held constant across all conditions at 105 arcmin. The lightpoint spacing varied from 15 to 28 arcmin along the pattern with an average of 21 arcmin.

A single orbit consisted of 129 frames, presented at frame rate 11 Hz for 11.6 seconds per orbit. Three orbits were typically presented over a 35 sec interval. Additional orbits were provided on observer request.

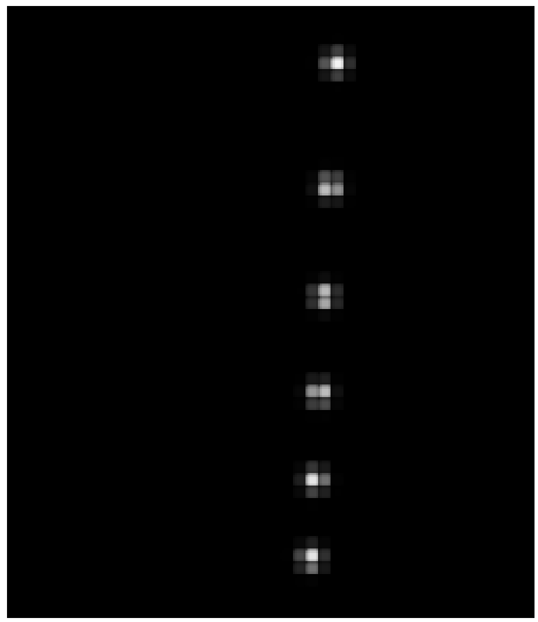

Figure 2. Sample of the line of six raster lightpoints as rendered on the FMD.

Task and Rating Scales

Observers made lightpoint quality ratings for each of the following four criteria.

<u>Pixel Grid Structure Visibility</u> Observers were instructed to attend to the visibility of the pixel grid structure and ignore the shape or movement of the lightpoints.

<u>Variation in Shape or Size</u> Observers were instructed to focus on the lightpoints and ignore the pixel grid structure where it was visible.

<u>Variation in Brightness</u> Observers were instructed to attend to any blinking, twinkling, or variation in brightness while ignoring variations in lightpoint size or shape.

Variation in Rigidity Observers were instructed to attend to the rigidity of the "bar" of lightpoints, that is, any variations in the regularity of the lightpoint spacing or line straightness.

For each criteria observers provided feedback using a rating scale with the following four categories:

1. Completely Invisible
2. Barely Invisible
3. Barely Visible
4. Clearly Visible

Observers

Six people, all employees of FlightSafety International, participated in the evaluation. All observers were male and their ages ranged from 32 to 49 years with a mean age of 39.5 years. All observers reported good distance vision. Data collection took about 1 hour per observer, including instructions and practice.

Procedure

On arrival in the lab each observer was read instructions and shown specific examples of the lightpoint attributes they were to focus on. Each observer then practiced making ratings so they could become familiar with the range of quality and rating scales.

During the evaluation observers were seated on a wheeled office chair. On each trial the observer rolled the chair to one of five distances marked on the wall. The distance between the observer and the display was used to control the Pitch of the FMD. The observer distances ranged from 2.34 m (92 in) for the largest Pitch condition to 6.55 m (258 in) for the smallest Pitch condition.

Observers viewed the lightpoints in a darkened room illuminated with floodlights on dimmers. The dimmers were set such that the wall luminance appeared to be the same brightness as the black background of the display (approximately 0.1 fL). During the evaluation observers verbalized their ratings and the experimenter recorded the ratings

Results

The rating scale data were first analyzed using the "stepwise" multiple regression modeling approach to determine the effects of the three design variables and interactions. The variables Pitch, Pitch^2, WidthP, WidthP^2, Fill, Pitch x WidthP^2, Pitch x Fill, and WidthP x Fill were used as candidate regressors in the models. Separate models were fitted for each of the four rating scales. Models were fit to the data of individual observers as well as to the mean data pooled across the observers. In no case did the model for an individual observer vary meaningfully from the model fit to the mean of the observers.

Effect of Fill Factor In general the multiple regression models consisting of the variables Pitch and WidthP and their squares and interactions fit the data with high statistical reliability ($p < 0.0001$). In none of the many models fit to these data did the factors Fill, Pitch x Fill, or WidthP x Fill have a reliable effect ($p > 0.05$) on any of the four dependent variables. Thus, the data were averaged across Fill for subsequent analyses.

Final Models After collapsing the data across fill, and observer, final models of a cumulative Gaussian (sigmoid) form were fitted using a nonlinear multi-dimensional optimizer (fminsearch in MATLAB) set up to minimize the sum of the squared errors between the raw and fitted data. The resulting model for each rating scale is shown as a contour plot in Figures 3 to 6. For all four dependent variables these models had high statistically reliability ($p < 0.0001$).

Figure 3 shows that while WidthP has a reliable effect, Display Pitch is the primary determinant of Pixel Grid Structure Visibility. At pitches above about 2.5 arcmin the pixel grid structure is readily visible when lightpoints are displayed. The pixel grid structure is invisible at display pitches of less than about 2.0 arcmin.

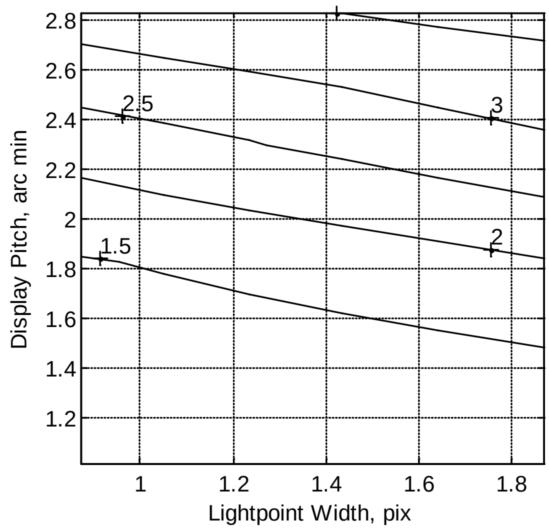

Figure 3. Effect of WidthP and Pitch on **Pixel Grid Structure Visibility**. $R^2 = 0.983$

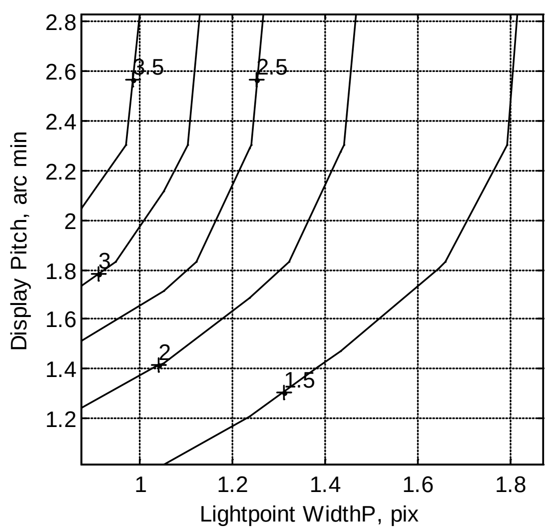

Figure 4. Effect of WidthP and Pitch on **Variation in Shape or Size**. $R^2 = 0.910$

From Figure 4 it is apparent that the visibility of variations in lightpoint Shape and Size depends on a complex interaction between display Pitch and lightpoint WidthP. When Pitch is greater than about 2.5 arcmin, where the grid structure is clearly visible, display Pitch has no effect on the visibility of variations in lightpoint size/shape. As Pitch is reduced below 2.25 arcmin, the number of FMD pixels required to produce lightpoints of constant shape and size can be reduced. This effect is likely due to the resolution limit of human vision. For very small lightpoints (e.g., <= 1 arcmin) all lightpoints look round like the optical blur function of the eye.

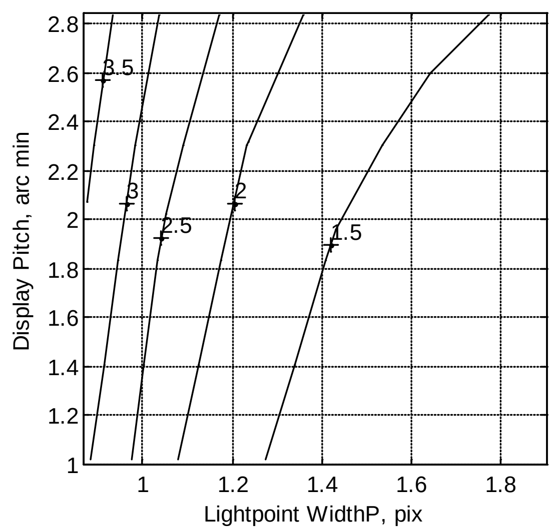

Figure 5. Effect of WidthP and Pitch on **Variation in Brightness**. $R^2 = 0.924$

The contours in Figure 5 show that the visibility of variations in brightness are strongly dependent on WidthP and more weakly on display Pitch. The WidthP at which brightness variations become visible decreases as display pitch is decreased.

The contours shown in Figure 6 are similar in shape and magnitude to that shown in Figure 5. The rigidity of the lightpoint pattern is strongly dependent on WidthP and relatively weakly dependent on display Pitch. A likely reason this interaction occurs is that the magnitude of the position errors for each lightpoint is a smaller fraction of the inter-lightpoint spacing as display Pitch decreases (inter-lightpoint distance was held constant).

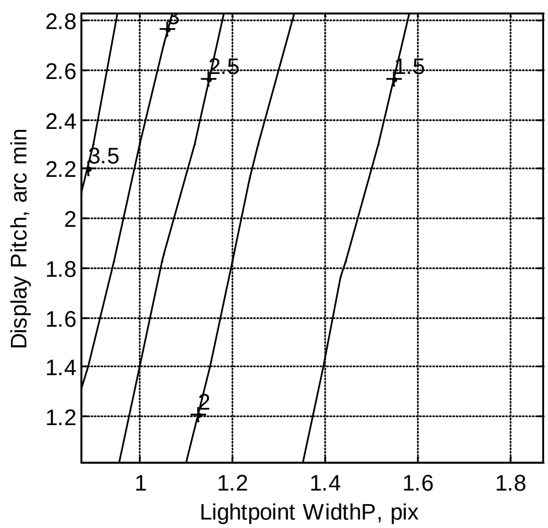

Figure 6. Effect of WidthP and Pitch on **Variation in Rigidity**. $R^2 = 0.916$

Conclusions: Evaluation 1

Over the pitch range of 2.6 to 3 arcmin, where most simulation training display systems are today, we can be confident that observers will not see spatial aliasing artifacts in raster lightpoints if the WidthP is 1.8 pix or greater (using the "1.5" contours on Figures 4 and 5). The WidthP could be set as low as 1.45 pix if we are willing to risk observers occasionally seeing faint variations in lightpoint brightness, shape, or size (using the "2" contours on Figures 4 and 5). For this range of widthP, the WidthA of the lightpoints is 4 to 5 arcmin.

For the tighter range of display pitches proposed for the first crop of FMD projectors to be used in flight simulation trainers, 2 to 2.2 arcmin, the "more confident" WidthP reduces to 1.72 pix and the "less confident" WidthP reduces to 1.4 pix (WidthA = 2.9 to 3.6 arcmin).

For display systems with a 2-ish arcmin pitch, the *size* of the alias-free "raster" lightpoints that can be produced will rival the best calligraphic lightpoints produced today. However, the peak luminance of these lightpoints will be no greater than the max luminance of the raster image. Recall that the peak luminance of calligraphic lightpoints can be 10 to 20 times higher than the max luminance of the raster image.

The reader is cautioned that this first evaluation applies for the case of rows of lightpoints that are not uniformly spaced and are no closer than about 15 arcmin. The results of the next two evaluations show that for rows of uniformly spaced (5 arcmin) lightpoints the WidthP requirement increases significantly.

Evaluation 2: Lightpoints at 5 arcmin Spacing

In this evaluation the effects of three design variables were evaluated using a task much closer to the JAA/FAA lightpoint size test.

Task and Rating Scales

Observers in this evaluation viewed strings of lightpoints that were spaced at 5 arcmin as per the JAA requirement. Figure 7 shows a sample of one of the 72 conditions evaluated.

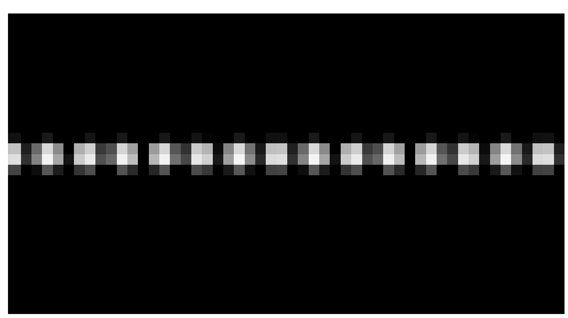

Figure 7. Sample of the horizontal line of raster lightpoints at 5 arcmin spacing as per the JAA "lightpoint size" test.

For all 72 conditions used in this evaluation, observers provided ratings of discernability using the following rating scale:

Modulation Discernability Between Lightpoints: Distinctness of individual lightpoints, regularity of their spacing, ability to count the lightpoints:

1. Clearly discernable
2. Barely discernable
3. Barely not discernable
4. Clearly not discernable

For that half of the trials that involved motion, the observers provided a second rating using the following scale:

Brightness Variation Moving Moiré patterns, pulsation, movement of stripes up and down the row:

1. Clearly stable
2. Barely stable
3. Barely varying
4. Clearly varying

Display Design Variables

Two of the design variables used in this evaluation, WidthP and Pitch, are described above in the first evaluation. The levels of these variables were adjusted to most efficiently reveal their effects on discernability and Moiré pattern visibility.

Display Pitch Seven levels of pitch were used: 1.40, 1.49, 1.59, 1.69, 1.81, 1.92, and 2.05 arcmin.

Lightpoint WidthP Four levels of lightpoint WidthP were used: 1.40, 1.60, 1.80, and 2.00 pixels.

The lightpoint WidthA that resulted from these variable settings ranged from 1.96 to 4.1 arcmin.

Lightpoint Motion For half the 72 conditions the string of lightpoints did not move relative to the pixel grid. For the other half of the conditions the string of lightpoints moved in a circular orbit with a period of about 8 seconds over a diameter of about 7 pixels. This motion was used because it effectively reveals the undesirable moving Moiré patterns that can appear in digital video.

Experimental Design

As in evaluation 1, a full factorial design was used where each observer rated all 72 conditions in a different random order so that practice and fatigue effects would not bias the data.

Equipment and Software

A Dell desktop LCD monitor that was brighter than the CRT used in the first evaluation was used in the second evaluation. The peak luminance of the lightpoints on this display was 26.6 fL and the luminance of the black background was 0.15 fL for a CR of 177.

Simulated FMD pixels were created using a 5x5 array of LCD monitor pixels for each FMD pixel. The native pitch of the monitor was 0.293 mm/pixel, thus, the pitch of the simulated FMD pixels was 1.465 mm.

From condition to condition the lightpoint spacing was fixed at 5.0 arcmin and the length of the lightpoint string was held constant at 17 cm. Thus, between 31 and 46 lightpoints were presented, depending on the display pitch (viewing distance).

Procedure

During the evaluation each observer again sat in a wheeled office chair which they moved to one of seven distances marked on the wall for each condition. The nearest distance of 2.46 m from the monitor produced a pitch of 2.04 arcmin while the furthest distance of 3.60 m produced the 1.40 arcmin pitch condition.

For the 36 conditions involving lightpoint motion the observers gave two ratings, Modulation Discernability and Brightness Variation, and the evaluator recorded the ratings. For the 36 no motion conditions the observers provided only the Discernability rating.

Observers Seven observers, all male employees of FSI, participated in this evaluation. The observer ages ranged from 32 to 62 years, mean of 43.1 years.

Results

Effect of Motion The analysis of the rating scale data indicated there was no statistically reliable difference ($p > 0.10$) in Modulation Discernability between the motion and no motion conditions. Moving the lightpoint string relative to the pixel grid did not improve or degrade the visibility of the gaps between lightpoints.

Effects of WidthP and Pitch The variables display Pitch and lightpoint WidthP both had strong and statistically reliable effects on both rating scales. The mean data, averaged across observer, motion, and a replicate, were used to fit multiple regression models for each dependent variable. Consistent results were obtained across the seven observers and the reliability of the regression models was very high ($p < 0.0001$). The R^2 and RMSE metrics are provided for each model in the figure captions below.

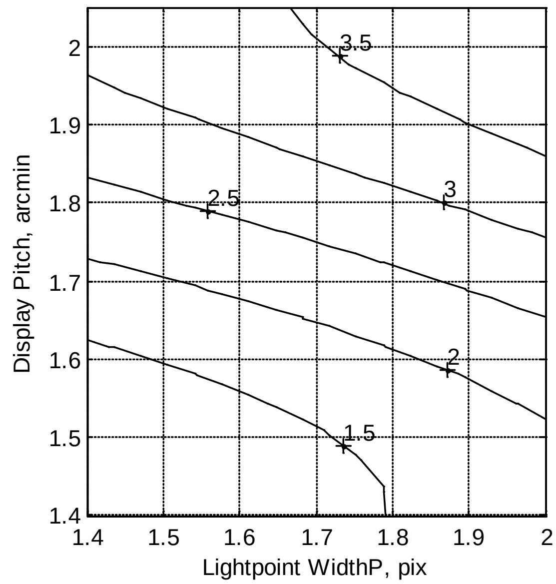

Figure 7. **Modulation Discernability** ratings as a function lightpoint WidthP and of display Pitch for lines of lightpoints with 5 arcmin spacing. $R^2 = 0.962$, RMSE = 0.19.

The shape of the contours in Figure 8 tell a story similar to that illustrated in Figure 5, that brightness variations are primarily determined by WidthP and only mildly influenced by display Pitch. Note, however, that the widths in Figure 8 are about 1.6 times larger than those in Figure 5. These data indicate observers are much more sensitive to the moving Moiré patterns that can occur with tightly and regularly spaced lightpoints than they are to the artifacts that occur with the more widely and non-uniformly spaced lightpoints used in the first evaluation.

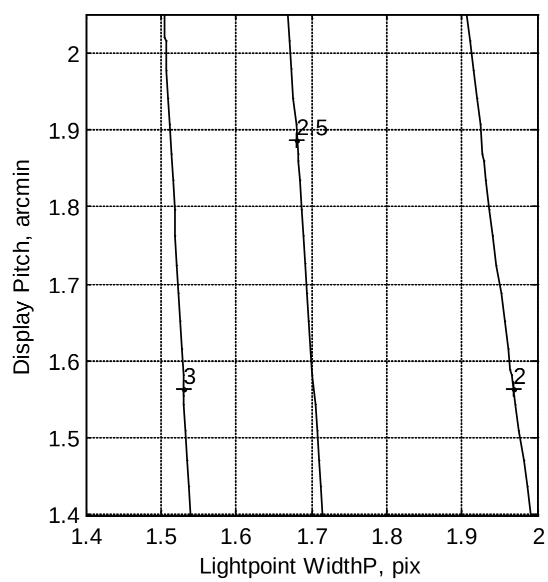

Figure 8. Ratings of **Brightness Variation** (moving Moiré patterns) occurring with movement as a function of lightpoint WidthP display Pitch and for lines of lightpoints with 5 arcmin spacing. $R^2 = 0.844$, RMSE = 0.26.

For the smaller display pitches where inter-lightpoint modulation is discernable (<= 1.7 arcmin), Figure 8 indicates moving Moiré patterns will not be visible if WidthP is larger than about 1.95 pix.

Assuming we can tolerate the occasional visibility of faint moving Moiré patterns (Contour "2.5" in Figure 8), the "1.5" contour in Figure 7 shows that we would need a display pitch of 1.52 arcmin to be confident of producing discernable modulation using the JAA test. A display pitch of 1.65 arcmin can be used if we are willing to occasionally fail the lightpoint discernability test.

Evaluation 3: JAA Lightpoint "Size" Test

The conclusion from Evaluation 2 that a display Pitch of 1.52 arcmin will be needed to reliably pass the JAA lightpoint size test has strong financial consequences, thus, this finding was tested again in Evaluation 3 to see if it could be validated.

Equipment and Software

The images used in Evaluation 3 were generated using a FSI Vital X image generator driving a SEOS Zorro 1410 projector which produced images with 1400 x 1050 pixels. The HFOV of the IG was set to 40 deg which produced an average display Pitch of 1.71 arcmin on the flat screen with the center pixel spanning 1.79 arcmin. During the evaluation each observer remained seated at a fixed distance of 1.82 m (72 in) from the screen which produced a 40 deg wide image and a Pitch of 1.79 arcmin at the screen center. The TOD was set to night, thus, the lightpoints had high contrast (e.g., CR > 10) relative to their background.

Procedure: Lightpoint Size Test

As per the JAA test, a string of 48 evenly spaced red lightpoints was presented 20 feet above the runway. The spacing of the lights was 8.30 ft for an angular spacing of 5.0 arcmin when viewed from a distance of 5700 ft.

Using the Host Emulator controls for the IG the experimenter flew the eyepoint back and forth along a 3 deg glideslope. During each trial the observer commanded "forward" or "reverse" to move the eyepoint towards or away from the runway. On each trial the starting distance from the lightpoints was randomly varied as was the speed.

The observer was instructed to move forward and backward several times in order to hone in on the distance at which the modulation between the lightpoints was "just discernable." The observer could see only the OTW image of the airport environment and did not know what the actual distances were. Each observer repeated this test four times.

Procedure: Vernier resolution Test

For this test the observers viewed the runway bars rather than the string of lightpoints. The JAA requires that the (4.0 ft) dark spaces between the white markings be discernable at 2 arcmin, corresponding to a viewing distance of 6880 ft.

As with the lightpoint test, the observer repeatedly moved forward and backward to hone in on the distance at which the dark bars were "just discernable." Again, the starting distance and speed varied randomly from trial to trial. Each observer repeated this test eight times.

Observers Six observers, all male FSI employees, participated in this evaluation. Observer ages ranged from 32 to 62 years, for a mean age of 43 years. All of the observers had field experience with the setup and conduct of the JAA/FAA lightpoint size test.

Results: Vernier resolution Test

The results of an ANOVA indicated there were significant differences among the observers (p , 0.0001) in their settings of the distances at which the runway dark bars were just discernable. The mean distance settings for each observer range from 6249 ft to 9032 ft for a maximum difference of 2783 ft. The within-observer SEM was 146 ft.

The grand mean across observers and trials was 8131 feet which corresponds to a dark gap of 1.69 arcmin or 0.945 pixels. Scaling this result by the JAA requirement of 2.0 arcmin, we determine that a display system with a Pitch of 1.79 arcmin * (2 / 1.69) = 2.12 arcmin would be on the threshold of acceptability for the Vernier resolution test.

Results: Lightpoint Size Test

The results of an ANOVA indicated there were significant differences among the observers ($p = 0.0001$) in their settings of the distances at which the lightpoints were just discernable. The mean distance settings for each observer range from 4986 to 5930 ft for a maximum difference of 944 ft. The within-observer SEM was 89 ft.

The grand mean across observers and trials was 5262 ft which corresponds to a lightpoint spacing of 5.42 arcmin or 3.03 pixels. Using this result we can estimate the Pitch at which a display system would just barely pass the lightpoint test: 1.79 * (5 / 5.42) = 1.65 arcmin.

By comparing the threshold Pitch for the Vernier resolution test (2.12 arcmin) with the smaller pitch required for the lightpoint size test (1.65 arcmin) we can see that a display system capable of passing the lightpoint size test will surely pass the Vernier resolution test.

Observer Variance

In evaluations such as those reported here we have the luxury of making precise estimates of thresholds, primarily because we can average together many estimates from each of many observers. The average standard deviation of the individual threshold pitch estimates reported in Evaluations 2 and 3 is 0.08 arcmin. Considering that 460 independent ratings were made, the standard error of the 1.65 arcmin estimate is well under 0.01 arcmin. We can be 95% confident that the mean threshold pitch for our group of observers is between 1.63 and 1.67 arcmin. However, if we wish to generalize these findings, we must consider the differences in pitch estimates made by different observers. This variance is illustrated in Figure 9 which shows a histogram of four threshold pitch estimates made by each of the 13 observers in Evaluations 2 and 3. The histogram contains both within and between observer variance but is dominated by the differences between observers (std dev = 0.14 arcmin).

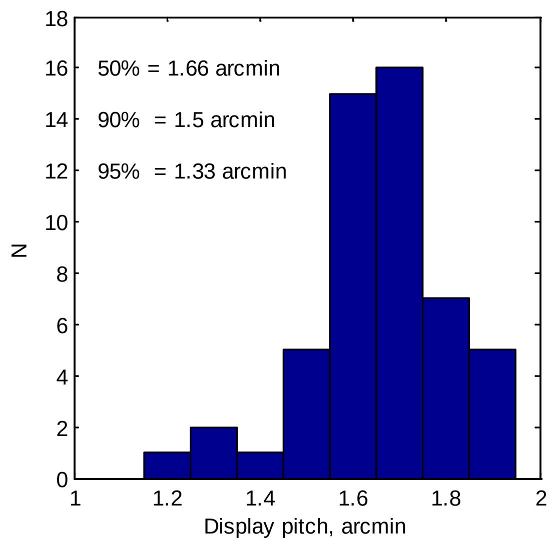

Figure 9. Histogram of “barely acceptable” pitch estimates for the 13 observers in Evaluations 2 and 3. Four estimates are shown for each for a total of 52 estimates. Most of the variance in these data (75%) is due to the differences between observers.

Interviews conducted with the observers after making their ratings revealed that a significant source of the variance between observers was the interpretation of “just discernable” modulation. Some observers thought the lightpoints in the string should appear evenly spaced and of equal size while others thought that any evidence of modulation (including the Moiré patterns?) qualified. A follow on evaluation is in the planning to determine the effect of instructions on the variance of the lightpoint size test.

Using this histogram in Figure 9 we can make more meaningful estimates of what might occur when a single inspector shows up at your facility to conduct the JAA lightpoint size test. These data indicate you have a 50% chance of a "just barely acceptable" rating if you have a pitch of 1.66 arcmin.

Personally, this author would not design a display system so close to the edge of acceptability as the cost of failing a certification test is usually quite high. A less risky design would use a pitch of 1.5 arcmin as this would give a 90% confidence of passing the lightpoint test.

The analyses in this section assume the visibility of the moving Moiré patterns was set to 2.5, meaning observers will "occasionally see faint Moiré patterns" in the lightpoint string. For those willing to field a display system for which Moiré patterns are more clearly visible, the histogram in Figure 9 can be shifted to the right by about 0.05 to 0.10 arcmin by shifting to a smaller WidthP as per Figure 8 and shifting the display Pitch up as per Figure 7.

Conclusions

The following conclusions are supported by the three evaluations described in this report.

1. The spatial aliasing artifacts occurring on FMDs are much more visible for regularly and tightly spaced lightpoints than they are for more loosely and non-uniformly spaced lightpoints.
2. The JAA "lightpoint size test" (5 arcmin spacing) is the primary determinant of the resolution requirement for FMDs for Level D training systems. The display pitch needed to pass the Vernier resolution test is 28% larger, thus the Vernier resolution test is easier to pass.
3. Given this author's tolerance to aliasing artifacts, a display Pitch of 1.50 arcmin is indicated for a 90% confidence that modulation will be "just discernable" for the JAA lightpoint size test.
4. The reader is provided the information required to make their own design trade between aliasing artifacts, display pitch, and probability of discernable modulation using the data in Figures 3-9.
5. The ratio of pitch requirements for the 5th and 95th percentile observer is 1.40 meaning that the more demanding observer will require twice as many pixels as less demanding observer.
6. No "Kell Factors" were used in the making of these analyses.

Recommendations

Those display design variables that most directly drive the cost of the system are FOV, Pitch, and perhaps luminance. FOV is typically measured using an instrument with a repeatability that is less than 0.0001% of the FOV requirement. Luminance is typically measured with a NIST-traceable instrument that is accurate to about 5%. While the relative tolerance for this meter is much larger than FOV, the expected cost of over-designing the luminance by 5% is small.

The resolution specification in the commercial flight training industry is unique in the sense that it is perhaps the biggest driver of the cost of the display system, yet we can measure this variable with a precision of only +/- 20%. In the interest of reducing this variance, and avoiding the energy wasted in arguing over the acceptability of candidate display system designs, four recommendations are provided.

1. Conduct an evaluation of the effect of instruction on the inter-observer variance for the existing JAA lightpoint size test. If the variance can be reduced by more clearly defining "discernable modulation" change the JAA standard accordingly.
2. Consider training inspectors using a stable standard of lightpoint modulation discern-

ability that holds the lightpoint quality at an acceptable level.

3. Consider using the calligraphic CRT lightpoint string as a standard against which FMDs will be evaluated. With calligraphic CRTs the lightpoints look "evenly spaced" and of "uniform size".
4. Develop an objective (measurements-based) metric of resolution that simultaneously maintains a constant and acceptable level of image quality (absence of aliasing artifacts).

References

FAA (1995) Advisory Circular, 120-40C Draft, Airplane Simulator Qualification.

JAR (2003) STD 1-A, Amend 3, 1 July 03.

Lloyd et. al., (2006) Effects of display pitch and lightpoint profile, width, and motion on JAA lightpoint size test performance. Technical Report, FlightSafety Int. VSS, Hazlewood, MO.

Lloyd, C. J, Long, J. Linn, T, and Nigus, S. (2006) Lightpoint evaluation 1: Effects of lightpoint width, display pitch, and display fill factor. Technical Report, FlightSafety Int. VSS, Hazlewood, MO.

Long, J. L. (2006) *JAR Vernier raster resolution and lightpoint size tests: Skylight 1500.* Technical Report, FlightSafety Int. VSS, Hazlewood, MO.

Schreiber, W. F. (1986) *Fundamentals of Electronic Imaging Systems.* Springer-Verlag. New York.