Perceptual Image Quality and Telescope Performance Ranking

Joshua K. Lentz and James E. Harvey
College of Optics and Photonics/CREOL, University of Central Florida
Orlando, FL, USA 32816-2700

Kenneth H. Marshall and Joseph Salg
Damewood Optical Maintenance Laboratory, Space Lift Range Systems Contract
Patrick Air Force Base, FL, USA 32925

Joseph B. Houston
Houston Research Associates
Saratoga, CA, USA 95070-3444

ABSTRACT

Launch Vehicle Imaging Telescopes (LVIT) are expensive, high quality devices intended for improving the safety of vehicle personnel, ground support, civilians, and physical assets during launch activities. If allowed to degrade from the combination of wear, environmental factors, and ineffective or inadequate maintenance, these devices lose their ability to provide adequate quality imagery to analysts to prevent catastrophic events such as the NASA Space Shuttle, Challenger, accident in 1986 and the Columbia disaster of 2003. A software tool incorporating aberrations and diffraction that was developed for maintenance evaluation and modeling of telescope imagery is presented. This tool provides MTF-based image quality metric outputs which are correlated to ascent imagery analysts' perception of image quality, allowing a prediction of usefulness of imagery which would be produced by a telescope under different simulated conditions.

Keywords: Perceptual Image Quality, MTF, metric, telescope, image simulation

1. INTRODUCTION

Monitoring of launch vehicles through optical imaging and tracking is performed throughout ascent to improve the safety of range personnel and astronauts. Images are monitored both in real-time and post launch to determine if any damage or unexpected events occurred during the operation which would compromise human safety. The optical imaging and tracking is performed using a fleet of ascent cameras and telescopes, the quality of which must be high to ensure that imagery is adequate for analysts to recognize and identify abnormal events.

Many of the Launch Vehicle Imaging Telescopes currently in use on the Eastern Range were designed in the 1960s and built in the 1970s and have degraded over time through various mechanisms. The maintenance performed on the Eastern Range telescopes until 2006 has been inadequate at best. The Space Lift Range System Contract, a cooperative and contractual partnership between ITT Corporation and L-3 Communications, has subcontracted and worked with CREOL/UCF over the past four years to develop an LVIT maintenance and modeling program. This Telescope Interferometric Maintenance Evaluation (TIME) Tool* provides an improved process of maintaining the telescopes used on the Eastern Range.

One component of this maintenance aid is an interactive software program system to model degraded images based on interferometric test data. The software also calculates MTF curves, OTF-based perceptual image quality metrics, and common metric outputs including wavefront errors and Strehl Ratio. A key to the modeling software is the link between image analysts' opinions of image quality, and perceptual image quality metrics. Perceptual testing was conducted with

Current Developments in Lens Design and Optical Engineering XI; and Advances in Thin Film Coatings VI,
edited by R. Barry Johnson, Virendra N. Mahajan, Simon Thibault, Proc. of SPIE Vol. 7786
77860D · © 2010 SPIE · CCC code: 0277-786X/10/$18 · doi: 10.1117/12.862833

analysts and optical laboratory personnel to establish "rankings" of image quality that can be directly translated to the quality of an optical device. Details of the perceptual testing, analysis and results of the ranking process are presented.

* Patent Pending

1.1 Maintenance Process

The process developed for maintaining the Eastern Range LVIT can be summarized with a flow chart (Figure 1). The process begins with the in-situ full-aperture interferometric testing of a telescope. This step includes the relative alignments of both the telescope and the interferometer.

After interferometric analysis is complete, the results are compared with the original optical prescription and test data to determine the source of errors which result in aberrations detected by the interferometry.

Next, the image simulation software developed for the maintenance process is used with the interferometry results to simulate a degraded extended image, model MTF curves, calculate various metrics both perceptual and non-perceptual in nature, and assign a telescope ranking based on its current condition. The interferometric test results, error sourcing results, and image simulation software outputs including the simulated imagery are included in a maintenance record for the telescope.

The performance ranking (numerical ranking 1 (low) to 5 (high), listed in Table 1) can be used as a guide in determining what if any telescope maintenance is necessary. Generally, a telescope with a ranking of 5 will be accepted for immediate use. Telescopes with rankings of 3 or 4 may be rejected or conditionally accepted for particular tasks. Telescopes with rankings of 1 or 2 will be rejected from service. Telescopes rejected from service may be repaired, refurbished, overhauled, or replaced depending on the condition, and type/source of errors.

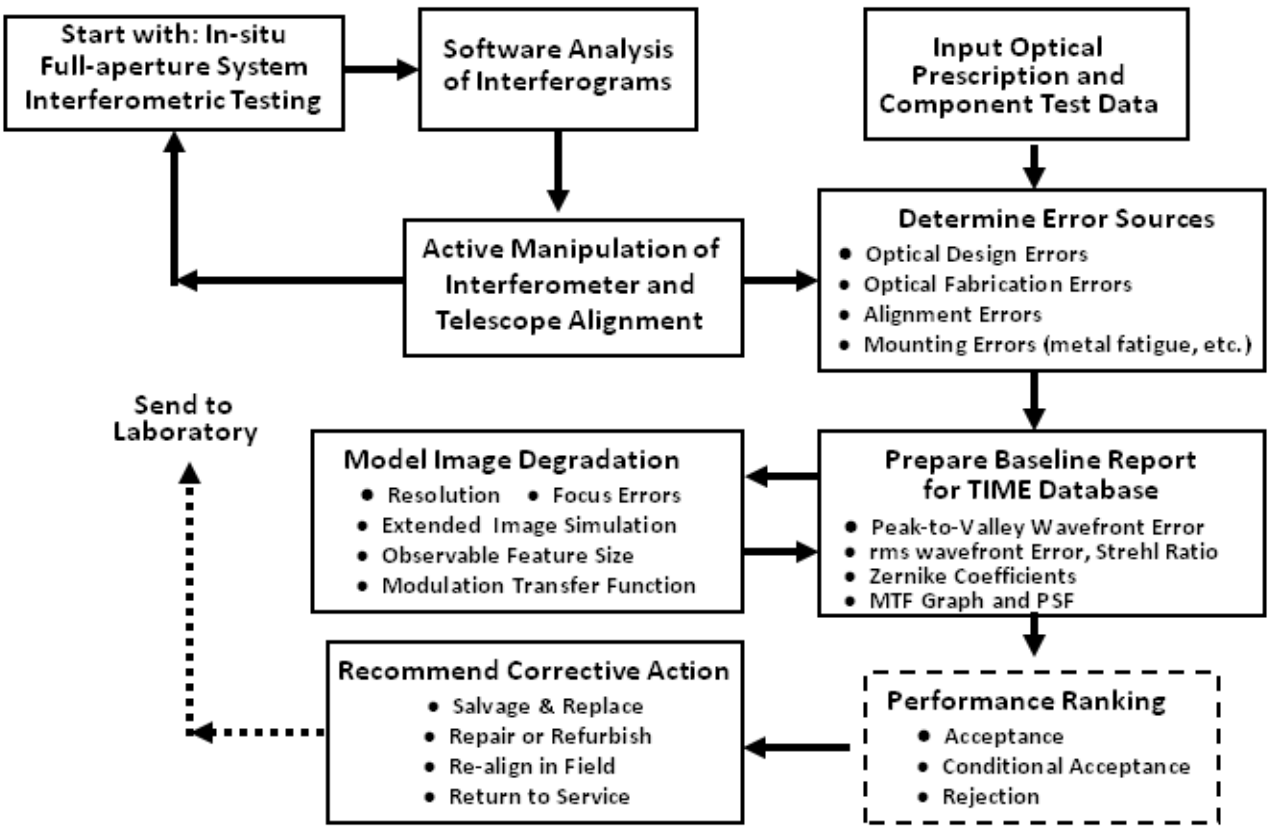

Figure 1. Flow chart of the maintenance process developed for the Eastern Range Launch Vehicle Imaging Telescopes.

Table 1. Category descriptors, numerical rankings, and corresponding color codes chosen for the telescope performance capability based on image quality provided by the telescope.

	Excellent	Good	Fair	Poor	Unusable
Numerical Ranking	5	4	3	2	1
Corresponding Color Code	Blue	Green	Yellow	Orange	Red

1.2 Simulation Software

The two main elements of the simulation software developed for assisting in the maintenance process are the extended image simulation and the performance ranking of telescopes based on the telescope interferometry results. Simulations apply linear systems theory to effectively convolve a Point Spread Function (PSF) with an extended object, represented by a bitmap picture file. To prevent the introduction of interpolation and/or compression errors, the "image" does not undergo magnification. Instead, the convolution is performed in the object plane and is displayed or saved as such with the "image" output bitmap equal in size to the "object" input bitmap.

Several assumptions are made in the software which limits its effectiveness. They are: (1) ideal image sampling (2) ideal atmospheric conditions (3) no mechanically-induced degradations (4) no motion blur effects.

1.3 Perceptual Image Quality Metrics

Of the plethora of image quality metrics in existence, only those few OTF-based metrics are rigorous enough to incorporate the effects of diffraction, optical aberrations, etc. Further limiting the available metrics is the perceptual requirements placed on the metric by the very nature of the ranking task. The metrics considered for this application and defined below (Eq. 1-3) are the Subjective Quality Factor (SQF)[1], the Square Root Integral (SQRI)[2], and the base metric of the Targeting Task Performance (TTP)[3] metric.

$$SQF = \frac{\int_{10cyc/mm}^{40cyc/mm} \int_0^{2\pi} MTF(u,\theta)\, d\theta\, d\log(u)}{\int_{10cyc/mm}^{40cyc/mm} \int_0^{2\pi} d\theta\, d\log(u)} \tag{1}$$

$$SQRI = \frac{1}{\ln(2)} \int_0^{u_{\max}} \sqrt{\frac{MTF(u)}{CTF(u)}}\, d\log(u) \tag{2}$$

$$TTP = \int_{u_{low}}^{u_{cut}} \sqrt{\frac{C_{TGT} MTF(u)}{CTF(u)}}\, du \tag{3}$$

In Eq. 1-3, u is a spatial frequency, u_{max} is the maximum displayed spatial frequency, CTF is the Contrast Threshold Function (the inverse of the Contrast Sensitivity Function) of the human eye, u_{low} and u_{high} are spatial frequency limits of integration which may vary by application and are discussed in the results, and C_{TGT} is the average target contrast.

The normalization of the SQF is not the normalization relative to the diffraction limit of performance, but rather, the normalization relative to the ideal performance whereby diffraction is not present. The classical SQF then has no advantage over either the SQRI or TTP, but it does present a significant disadvantage—it does not include the perceptual effects induced by the human eye, through the CTF. An updated SQF which applies a CSF-weighted MTF in the

integral rather than a simple band-limited integral was proposed by Imatest[4], but perceptual testing verifying the correlation to observer opinions of image quality has not been presented. For the reasons above, the SQF is considered to be a poor candidate for use in ranking the quality of imaging telescopes.

2. METRIC SELECTION

Both the SQRI and TTP in diffraction-limited forms were applied for comparison of results. Both metrics were used in a two-dimensional form applying the rigorous MTF(u_x,u_y) in the calculations. As a result of the normalization process, constants are irrelevant and the SQRI's claim to be in Just Noticeable Difference (JND) units fails.

2.1 Normalized Metrics

Using normalized metrics (normalized to the diffraction-limited value) for ranking provide additional information about a telescope. When a telescope is modeled in the simulation software, a telescope ranking based on its condition is provided, as well as the metric value and the breakpoints for each ranking category. It can be immediately seen in some cases if a particular ranking is even possible for a given telescope. As an example, suppose an 18 inch diameter telescope with some aberrations is modeled in the software for a 30 mile object distance. The ranking provided is 3, and the metric value is 72. Then consider that the ranking breakpoints (the values separating adjacent rankings) are 118, 88, 60, and 21. One can immediately see that even if the telescope were overhauled and brought to diffraction-limited performance (metric value of 100), rank 5 performance could never be achieved for that aperture diameter at that range. This provides information to laboratory personnel for deciding if anything would be gained by working on the optical system.

2.2 Contrast Threshold Function

The CTF follows the definition of Mannos[5] with a scaling term introduced to force the contrast threshold to unity at a retinal frequency of 48 cyc/deg. The CTF was assumed circularly symmetric for simplicity, and considering the limitations in the software (no sampling, atmosphere, vibration, motion, glare, etc.), and the variation in observation distances, this was considered an adequate approximation. Further, the CTF was not selected for a particular brightness condition but chosesn to represent a typical image use environment which may vary significantly. A 2/3 inch square format detector is assumed with a 10 inch displayed image size. The display magnification is then adjusted between 16 and 50 to provide a constant mM product (*m* is the telescope magnification and *M* is the display magnification).

2.3 Object Range Effects

The appropriate plane for metric evaluation is the object plane[6]. This allows the metric value to reflect object range variations. However, if the metrics are normalized by their diffraction-limited metric value, the object distance parameter is removed, as seen from a simple change of variables (C.O.V.).

$$TTP = \int_{u_{low}}^{u_{cut}} \sqrt{\frac{C_{TGT}\,MTF(aRu)}{CTF(bRu)}}\,du \rightarrow C.O.V \quad w = Ru \rightarrow \int_{Ru_{low}}^{Ru_{cut}} \sqrt{\frac{C_{TGT}\,MTF(aw)}{CTF(bw)}}\,\frac{1}{R}\,dw \tag{4}$$

$$SQRI = \int_{u_{low}}^{u_{cut}} \sqrt{\frac{MTF(aRu)}{CTF(bRu)}}\,\frac{1}{u}\,du \rightarrow C.O.V \quad w = Ru \rightarrow \int_{Ru_{low}}^{Ru_{cut}} \sqrt{\frac{MTF(aw)}{CTF(bw)}}\,\frac{1}{w}\,dw \tag{5}$$

In this normalized case, there is no advantage to using the object plane MTF over the natural image plane MTF. The lost object distance dependence can be recovered using an additional perceptual test as seen in section 4.

3. PARAMETRIC RANKING ASSUMPTIONS

Some assumptions have been made in the telescope ranking process to reduce the amount of perceptual testing needed. The first assumption is that all modeling parameters (aperture diameter, focal length, obscuration ratio, aberrations, object range) are independent in the sense that a single multi-parameter surface describing the entire process is separable into curves varying in a single parameter. Several parameters require additional discussion.

3.1 Obscuration Ratio

All of the obscuration ratios for LVIT are from 0.3 and 0.35 inclusive. It is assumed that within that variation, the obscuration ratio has NO discernable effect on perceived image quality. It is fully expected that obscuration ratios significantly outside that range will have noticeable perceptual variations.

3.2 Focal Length

Although the focal length affects the MTF cutoff spatial frequency, it has no effect on perceptual image quality[6]. Also inherent in this assumption is that imagery will be magnified significantly such that it is not eye-limited.

3.3 Wavefront Aberrations

Individual wavefront aberrations are expected to have the same perceptual effect on image quality if they have the same metric value. That is not to say that different aberrations have the same affect on image quality, but rather that the overall quality or usefulness of images is determined by the metric values, and not the aberration present. It is assumed that the telescope operator always maximizes the image quality and that this optimum image quality corresponds to aberrations which are balanced to minimize RMS wavefront error. Finally, all synthetic images generated for the perceptual testing varied in level of (balanced) third order spherical aberration only.

4. RANKING PROCESS

Telescope ranking is a multiple step process involving perceptual testing and analysis which is dependent on the assumptions provided above. It is important to note that the only parameters which are assumed to yield changes in perceptual image quality under the imaging telescope conditions are wavefront aberrations, aperture diameter, and object range.

4.1 Perceptual Testing

The perceptual testing conducted to find the image quality breakpoints was of the categorical sort test format in which single images were presented to observers who were asked to "rank" the quality of the image. Options were "Excellent", "Good", "Fair", "Poor", and "Unusable". Although the "Unusable" category is difficult to define and varies substantially from one observer to the next, the difficulty is insignificant since rank 1 and 2 telescopes both get rejected. Images were presented on a single computer to a single observer at a time. Images were randomly ordered and a ranking was required before continuing. Two tests were conducted—the first to determine the breakpoints as a function of aperture diameter, the second for the functional dependence of object range. For each data point on such a curve, approximately 8 images varying in aberration magnitude were necessary. For diameter variation, 9 diameters were tested with a total of 70 images. For range variation, 7 object ranges were tested with a total of 55 images. Each data set was presented to the observers twice.

All testing was conducted in the same room (identical ambient lighting) with maximum brightness on the monitor, and monitor resolution set at 1140×900. Observers were allowed to adjust the screen angle to accommodate different observer heights. Observers who normally use corrective lenses were asked to use them during the testing. Images were 768×900 pixels, chosen to be familiar to the observers.

4.2 Observers

Observers selected for the perceptual testing were individuals experienced with image quality and the application of the LVIT. A total of 7 observers participated in the test with diameter variation, and 6 observers participated in the object range variation test.

5. RESULTS

The perceptual test data was analyzed using Class II, Condition B outlined by Torgerson[7] and Engeldrum[8]. Further, the assumption that trials and observers can be considered equivalent was applied. Metric values were calculated using the following forms of the SQRI and TTP:

$$SQRI = \frac{\int_{2\,cyc/mm}^{u_{cutoff}} \int_{2\,cyc/mm}^{u_{cutoff}} \sqrt{\frac{MTF(u_x,u_y)}{CTF(u_x,u_y)}}\, d\log(u_x)\, d\log(u_y)}{\int_{2\,cyc/mm}^{u_{cutoff}} \int_{2\,cyc/mm}^{u_{cutoff}} \sqrt{\frac{MTF_{diff}(u_x,u_y)}{CTF(u_x,u_y)}}\, d\log(u_x)\, d\log(u_y)}\,, \tag{6}$$

$$TTP = \frac{\int_{2\,cyc/mm}^{u_{cutoff}} \int_{2\,cyc/mm}^{u_{cutoff}} \sqrt{\frac{MTF(u_x,u_y)}{CTF(u_x,u_y)}}\, du_x\, du_y}{\int_{2\,cyc/mm}^{u_{cutoff}} \int_{2\,cyc/mm}^{u_{cutoff}} \sqrt{\frac{MTF_{diff}(u_x,u_y)}{CTF(u_x,u_y)}}\, du_x\, du_y}\,, \tag{7}$$

where MTF_{diff} is the diffraction-limited MTF, u_{cutoff} is the optical cutoff spatial frequency, the CTF is as given above, and C_{TGT} from Eq.3 cancels from the normalization process.

5.1 Data Set 1: Aperture Diameter

Results of the analysis for the first data set (variation in aperture diameter) are shown in Figure 2 and Figure 3. The SQRI boundary (breakpoint) curves (Figure 2) make sense intuitively in that higher aberration magnitudes are tolerable for larger aperture diameters. The zero SQRI asymptotes also make sense as a true zero value of the metric requires extreme levels of aberration (a result of unit MTF values at zero spatial frequency). Further, it becomes increasingly difficult to achieve high quality imagery for smaller apertures. Note that an SQRI value of 100 corresponds to the diffraction limit, indicating that for diameters below approximately 17 inches, excellent quality imagery can never be obtained at this object distance. Note that this figure is valid only for the object distance of 15 miles, but the curves are assumed independent of object distance, and thus, a simple scaling of the values makes the figure valid for other ranges.

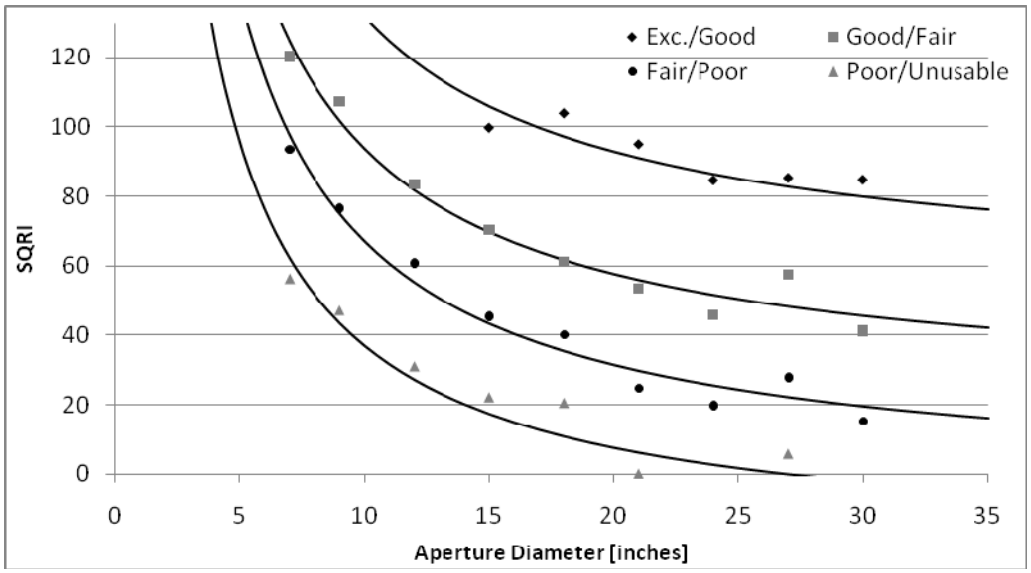

Figure 2. SQRI boundary values resulting from categorical sort testing for aperture diameter trends. Solid curves are inversely proportional to the aperture diameter with excellent R^2 values.

Although the SQRI boundary curves are intuitive, the TTP boundary curves of Figure 3 for aperture diameter variation are not. The linear curves resulting from the analysis indicate that lower quality images will no longer exist as higher aperture diameters are encountered. This is concluded from the zero crossings of the curves and is clearly not the case encountered in reality since badly aberrated optics will create bad imagery regardless of the size of the aperture.

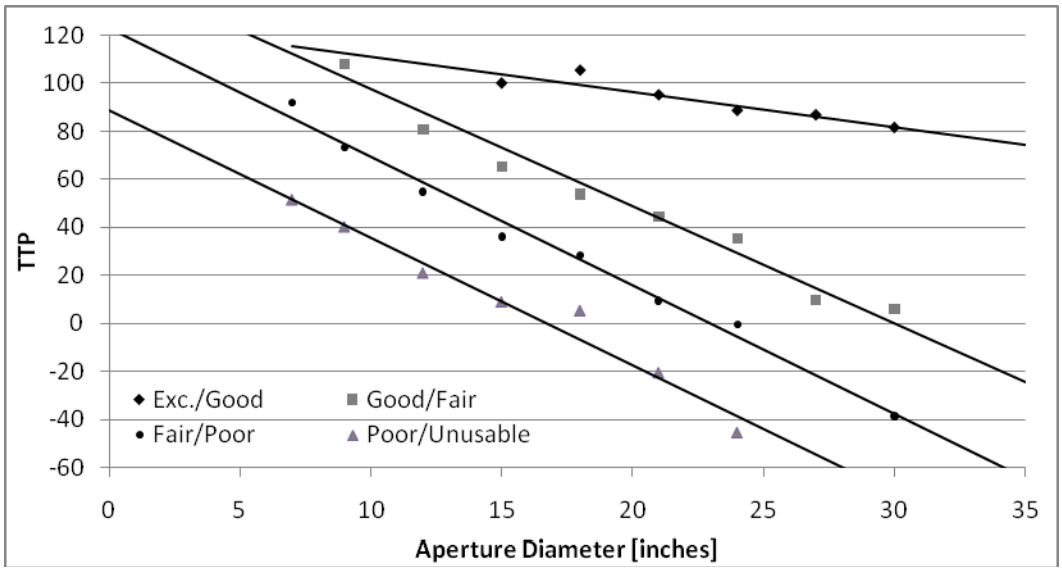

Figure 3. TTP boundary values resulting from categorical sort testing for aperture diameter trends. Solid curves are proportional to the aperture diameter with excellent R^2 values.

5.2 Data Set 2: Object Distance

Results of the analysis for the second data set (variation in object range) are shown in Figure 4 and Figure 5. The data appears to be linear in nature (of the form shown in Eq.6) for both metrics, as evidenced by the high mathematical fit values shown in Table 2. It is plausible that the boundary curves are linear in object range but level off to an asymptote as object ranges increase. However, for the object ranges simulated and tested, a linear relationship appears to be at least an excellent approximation.

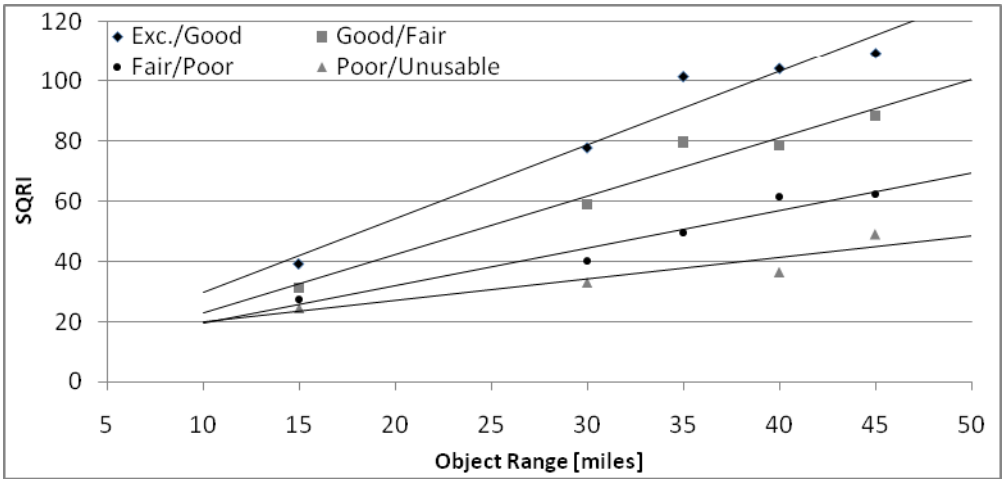

Figure 4. SQRI boundary values resulting from categorical sort testing for object range trends. Solid curves are proportional to the object range with excellent R^2 values.

Figure 5. TTP boundary values resulting from categorical sort testing for object range trends. Solid curves are proportional to the object range with excellent R^2 values.

The R^2 values were used as a measure of the quality of fit to the data. R^2 values obtained for Figures 2-5 are listed in Table 2.

Table 2. R^2 fit values for the ranking boundary curves shown in Figure 2 through Figure 4.

	Exc./Good	Good/Fair	Fair/Poor	Poor/Unusable
SQRI Inverse Diameter	0.641	0.967	0.968	0.868
SQRI Linear Range	0.954	0.959	0.949	0.863
TTP Linear Diameter	0.863	0.974	0.989	0.968
TTP Linear Range	0.941	0.956	0.971	0.908

6. CONCLUSION

From the previous section, the intuitive advantage of the SQRI metric over the TTP metric was illustrated. For that reason, further analysis was conducted only for the SQRI. Note that although the TTP metric made little intuitive sense in this application, it by no means implies that the TTP is poorly correlated to perceptual image quality. It must be considered that the TTP was applied here in a modified (normalized to diffraction limit) form.

After attempting data fitting with various curve types, it was found that the data had highest fit values for curves of the form found in Eq.8, namely an inverse diameter curve. Upon visual inspection, it is immediately apparent that the Excellent/Good data do not fit the curve as well as the two middle boundaries. This stems from an observer tendency to shy away from ranking images in the extreme positive category. Similarly, the Poor/Unusable data is not quite as good a fit to the curve as found in the middle categories, this time a result of the diverse understanding of the "Unusable" category. The middle boundary curves however, have excellent visual and mathematical fits (Table 2) to the data. It is expected that all boundary curves have the same mathematical relationship (inversely related to the aperture diameter in this case) as occurs here.

For constants C_i and aperture diameter D in inches,

$$SQRI_{bound}(D) = \frac{C_1}{D} + C_2 \tag{8}$$

For constants C_i and object range R in miles,

$$SQRI_{bound}(R) = C_3 * R + C_4 . \tag{9}$$

With the one-dimensional parametric variation boundary curves shown in Figure 2 and Figure 4, two-dimensional parametric boundary surfaces can be created of the following form:

$$SQRI_{bound}(R, D) = k_1\left(\frac{k_2}{D} + k_3\right)\left(\frac{R}{15\,miles}\right), \tag{10}$$

with k_i constants. The values determined for the constants of Equation 7 are shown in Table 3.

Table 3. Constants of Eq.7 found from the data.

	Exc./Good	Good/Fair	Fair/Poor	Poor/Unusable
k_1	0.8812	0.8971	0.7259	0.461
k_2	780	723.71	713.79	589.55
k_3	54	21.382	-4.219	-21.954

It is important to note that for a given object range and aperture diameter, a constant SQRI (or other OTF-metric) implies constant image quality (averaged over a group of observers). This is not the case for parametric variations as can be seen from the data above. The boundary curves are curves of constant image quality as a function of the parameter being varied. Further, the curves determined were for a particular metric, namely the SQRI, normalized to its diffraction-limited value. The relationship of the curves to the parameters may change drastically for other metrics (as seen with the TTP) and the above results should not be expected under different assumptions.

7. SUMMARY

Perceptual image quality testing combined with well-established psychometric scaling techniques can be implemented to find image quality category boundary curves as a function of image quality parameters such as aperture diameter and object range. This allows the prediction of image quality for a considerable variation of long range imaging telescope conditions such as those typically encountered in the use of LVIT. These predictions allow more informed maintenance decisions to be made regarding the condition and expected performance of the LVIT, increasing range safety by improving telescope selection, and decreasing financial loss encountered through unwarranted maintenance. The TIME Tool Optical Evaluation and Maintenance software package is presently being prepared for distribution to qualified optical laboratories and organizations for qualification purposes.

REFERENCES

[1] Granger, E. M., Cupery, K., "An optical merit function (SQF), which correlates with subjective image judgments," Photog. Sci. Eng. 16, 221-230 (1972)).

[2] Barten, P. G., "Evaluation of subjective image quality with the square-root integral method," J. Opt. Soc. Am., 2024-2031 (1990)C. Jones, Director, Miscellaneous Optics Corporation, interview, Sept. 23, 2008.

[3] Vollmerhausen, R., Jacobs, E., Driggers, R., "New metric for predicting target acquisition performance," Opt. Eng., 43(11) 2806-2818 (2004).

[4] http://www.imatest.com/docs/sqf.html/

[5] Mannos, J. L., and Sakrison, D. J.., "The Effects of a Visual Fidelity Criterion on the Encoding of Images," IEEE Transactions on Information Theory, Vol. IT-20, NO.4, July 1974.

[6] Lentz, J. K. and Harvey, J. E., "Focal Length Invariance of Perceptual Image Quality for Long Range Imaging Applications," Technical Digest, OSA Topical Meeting on Imaging Systems (IS), Tucson, AZ (June 2010).

[7] Torgerson, W. S., [Theory and Methods of Scaling] (John Wiley & Sons Inc., NY, 1967).

[8] Engeldrum, P. G., [Psychometric Scaling] (Imcotek Press, MA, 2000), p.129.