Making the qualitative quantitative—a discussion of the specification of visual systems

R. R. Shannon
Optical Sciences Center
University of Arizona
Tucson, Arizona 85721

Virtually all optical systems have some possible interpretation as a visual optical system. Often though, a detector such as a vidicon or photographic film intercedes to record the image. In such cases several quantitative methods have been devised to specify the required aberration content and fabrication tolerances for the optics used in such systems. Direct visual systems are the oldest of optical systems, but generally lack a good quantitative basis for tolerance specification. This condition has been accepted because the human eye is a remarkably adaptive organ, which will accept imagery that appears to be unacceptable by many of the accepted quantitative measures. On the other hand, the eye is bothered by such defects as scratches, digs and other cosmetic errors that are not consequential to many other optical systems.

In this paper, I will discuss some of the attempts to obtain a quantitative tolerance or specification for an optical system. This survey is not exhaustive, because there is apparently current work going on that is not reported in the scientific literature. I will also attempt to evaluate the extent to which a quantitative standard can be developed and the extent to which it can be expected to be useful.

The basic advantage, and problem, with the eye as a detector, is that it is a dynamic, multichannel device. The dynamics of the eye involve a constant scanning action which couples the eye to the device using it. The scanning nature requires that certain limitations be placed on the nature of the imagery to avoid fatigue, or erroneous response due to the fixation of the eye on a number of areas within the field where objectionable defects reside, to the exclusion of areas where the eye should be free to carry out a constant dynamic search.

The eye is a color receptor, but with a separate set of receptors for each color band. The eye is thus capable of ignoring, or detecting chromatic errors, depending almost on the mood of the user or observer. To state what level of error is detectable by the user of a visual instrument is relatively easy. To state what is acceptable, is a more difficult problem.

Attempts to determine acceptable levels of performance have generally been based on determination of the detectable error, and then a modification of this level to the extent necessary to attain the goals of cost or acceptability by a specific class of user. Over the years, an apparently semi-infinite base of experimental data on vision has been developed. Unfortunately, the nature of this data base is such that very little of it is applicable to the determination of the interaction of the eye with visual optical systems. As a result, most of the specifications that have stood the test of time in many regards but cannot be stated in unequivocal physical quantities to determine a tolerance level that is always acceptable. The statement, "there must be a better way", is often expressed in such situations.

There are two types of quantitative data that are of importance in describing a visual system and its interaction with the eye. The two involve the specific effect of the relation between the image quality of the system and it's effect on the eye, and the effect on the dynamics viewing through the system with one or both eyes.

The first problem involves the interaction of the eye with the exit pupil of the system being used, including the effect of defocus and aberrations. A certain level of defocusing can be accepted by the eye, as long as the required accommodation is in the direction of accommodating to view an object closer to the eye. The sign convention describing this direction of accommodation is not universal, but is generally referred to as a plus diopteric error in the entering ray bundle. In the case of an axial focus error this is correctable, but in the case of field curvature the problem is more complex. Since the eye is scanning across the field of view, it will be constantly accommodating to maintain the image in best focus, and the degree of allowable field curvature is determined by the allowable dynamics of the search process, and ultimately by the allowable level of visual fatigue.

The problem of defocus has been discussed in several texts on optical design, although not in detail. As a guide, up to two diopters of field curvature can be tolerated in
most types of systems, although this varies according to application. The case of aberrations, in particular spherical aberration, coma and astigmatism have only been dealt with recently in a set of measurements carried out as a PhD dissertation study by Dr. M. Giles, at the Optical Science Center. The data is collected and described in the Journal of the Optical Society of America, Vol. 67 p.634, 1977. The conclusion is that the eye accommodates to the best focus as would be predicted by finding the minimum rms. wavefront error, and the allowable amount of residual aberration depends upon the allowable resolution value, expressed in cycles per degree. Several tables and graphs that are useful in this regard are given in the cited paper. These may be used to set the allowable aberration at a given point in the field of view, and used as a guide in determining the allowable range of change of the aberrations across the field of view.

The second type of tolerance that is often required in a visual system is that of the allowable amount of random deviation from perfection that can be accepted. This is not as amenable to a single experimental determination as the aberration relations that were measured by Giles, since the errors may take many forms. Therefore, only an average statement can be made about the allowable level.

One example of this is the specification on the random "visual distortion" that is acceptable in an optical system. The allowable level must be determined by using samples of the type of system and determining which levels are acceptable, and which are not. Here the newer techniques of testing of optical systems can be used to quantify the system error, and provide a method of determining whether a system is acceptable or not.

The problem is that of determining the acceptable level of "visual distortion" in a visor. The problem starts here. The meaning of the word distortion is taken differently by different disciplines. The optical designer looks at distortion as a specific functional error or aberration which is a consequence of the choice of the specific configuration of the system. To the designer, distortion is a specific and calculable physical quantity. In a global sense the visual scientist will consider the curvature of stylized lines as perceived, as an error arising from the eye's inability to perceive distortions. However, the geometric optics of the eye necessarily produce distortion in the retinal image and the observer achieves control of such distortion by adaptation within the neural visual processing. The tolerance on such an effect is clearly more complex than simply stating a detectable level.

The extent to which the cost of an optical system can be reduced, or made salable for other reasons is often intimately related to the allowance of a significant amount of such error in the product. The leading common example of this is the curved windshield found on most modern automobiles. The driver would be appalled to see a photograph of the metric errors or distortion errors which occur in the windshield, but actually adapts to the effect, in most cases. Finally the term "visual distortion" has been extended by some workers to encompass any level of angular mapping error which exists in the optical system, including small scale irregularities on the optical surfaces. This becomes a problem, for there is an intimate relation between the size of such irregularities and the presence of other visual effects that may occur, such as degradation of acuity, either dynamic or static. The presence of irregularities in different portions of the field will also cause an effect on the binocular vision using the visual system. Finally, when all logic is exhausted, the medical man will plead that the effect should be sufficiently small that "visual fatigue" does not occur.

The question now is, can a quantitative base be built which will respond to all of the above, and be acceptable to the engineering and the vision community?

The remainder of this paper will cover a particular topic in this regard, and discuss its interpretation. As an example, I will use a problem which I recently encountered without discussing the nature or the details of the application. The intention is to show how current technology should be used in quantifying visual data.

Military specification MIL-V-43511A, dated 30 September, 1976, and issued by the U.S. Army, addresses the question of the specifications for Polycarbonate visors for flyer's helmets. Included in this document is a set of qualitative values for acceptable and unacceptable levels of optical distortion. These visual, qualitative standards are applicable only to one method of testing, and obligate the testing authority to carry out a visual inspection of the masks of interest. Newer methods of testing rely on electro-optical devices for the acceptance and rejection of production quantities of optical masks. Conversion of these qualitative comparison standards to acceptable quantitative standards for use with such automated test and inspection devices is important.

This note is addressed to conversion of the standard to a quantitative level. The examples of interest are shown in Fig. 1, which is a copy of figure one of the referenced document. The apparent fringe patterns shown are moire' patterns resulting from the
application of a specific configuration of the "Ann Arbor" tester shown in Fig. 2. The latter figure is a copy of figure three of the referenced document and illustrates the setup used in the test process. The use of a 240 mm. focal length lens and the 60 line per inch grating in the tester leads to the conclusion that the deviation from straightness of the observed fringes of the amount of fringe spacing indicates a local slope error of .0098 radians, or about 1 prism diopeter in the return wavefront. When the double pass nature of the system is taken into account, the shift of one fringe indicates a local error of .5 prism diopeter.

When this conversion between fringe straightness and ray deviation is noted, the statistics of the error on the mask can be calculated. This can most readily be done by use of an existing computer program, called FRINGE, written by Mr. John Loomis at the Optical Science Center. This program is intended to operate with interferometric data in which the fringe data represent optical path error. The conversion mentioned above permits use of the analytic features of the program, with the understanding that one fringe represents a slope change of .005 radians, or an error of 1 prism diopeter locally in the observed area of the mask. The area observed at a given time on the mask is found from Fig. 2 to be 43.6 mm., of 1.7 inches in diameter.

The process of computation used was to scan the fringe patterns in the usual method, using a digitizing tablet, and then running the resulting fringe data through the usual fringe analysis procedure.

The output consisted of several interesting numerical outputs, plus a contour map of the residual errors. The numerical values of interest are shown in Fig. 3, and consist of the rms values, the maximum, the minimum and the peak-to-peak span of the values of the ray divergence in units of prism diopeters. The first 5 sets of data are for the acceptable masks, and the last four for the unacceptable masks. There is a definite difference between the two sets. This leads to the following conclusion. The boundary between acceptable and unacceptable should be set at .25 prism diopeter peak-to-peak variation over a 2.3 square inch area, or .055 rms prism diopeter variation over the same area. The small standard deviation of the two sets of data suggests that simply stating the acceptable level of variation of the error over a given area constitutes a sufficient specification for the error. There may be a need to refine this by adding some statements about the allowable size of isolated errors, but this may not be required, as such an error is noted in Fig. 11.

Figures 4 through 11 show the contour maps derived from reduction of the fringe data. The separation between contours in each of the cases represents a change of .05 prism diopeters in the refractive error of the mask. In each case the diameter of the pattern represents 1.7 inches on the mask.

If the error is considered to be randomly distributed, then the allowable rms value can be evaluated as .25 prism diopeter change per 2.3 square inches, and scaled directly with the area being evaluated. The peak error should be scaled with caution, and represents the most likely the allowable peak deviation of the dioptric power over a 1.7 in diameter on the surface. The constant ratio of 4.2 for peak to rms value suggests that either evaluation could be considered.

It is hoped that these observations will be of use in setting future quantitative specifications for visual products of the type discussed.

These considerations have considered only a few of the possible errors that may occur in a system which involves the eye. Nothing has been said here about color effects, vignetting due to the motion of the eye in a dynamic situation in and out of the exit pupil, or the effects in binocular of binocular vision. It is clear that some of the newer techniques of measuring and specifying optics can be used to develop quantitative specifications that may be tested in use to enable realistic tolerances to be applied to a visual optical system.
VISORS DISTORTION STANDARDS

ACCEPTABLE

1  2  3

4  5

UNACCEPTABLE

6  7

8  9

Fig. 1

Fig. 2 Distortion Tester

Table of computed values for the nine examples.

<table>
<thead>
<tr>
<th>Example</th>
<th>rms</th>
<th>max</th>
<th>min</th>
<th>span</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptable</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>.033</td>
<td>.06</td>
<td>-.10</td>
<td>.16</td>
</tr>
<tr>
<td>2</td>
<td>.05</td>
<td>.13</td>
<td>-.09</td>
<td>.21</td>
</tr>
<tr>
<td>3</td>
<td>.05</td>
<td>.08</td>
<td>-.11</td>
<td>.19</td>
</tr>
<tr>
<td>4</td>
<td>.04</td>
<td>.12</td>
<td>-.09</td>
<td>.21</td>
</tr>
<tr>
<td>5</td>
<td>.04</td>
<td>.08</td>
<td>-.10</td>
<td>.18</td>
</tr>
<tr>
<td>Unacceptable</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>.08</td>
<td>.15</td>
<td>-.19</td>
<td>.25</td>
</tr>
<tr>
<td>7</td>
<td>.07</td>
<td>.14</td>
<td>-.14</td>
<td>.28</td>
</tr>
<tr>
<td>8</td>
<td>.11</td>
<td>.23</td>
<td>-.18</td>
<td>.41</td>
</tr>
<tr>
<td>9</td>
<td>.06</td>
<td>.08</td>
<td>-.19</td>
<td>.27</td>
</tr>
</tbody>
</table>

Acceptable

rms = .043 +/-.006
span = .19 +/-.02

Unacceptable

rms = .076 +/-.018
span = .328 +/-.065

Fig. 3