Psychometrically appropriate assessment of afocal optics by measurement of the Strehl intensity ratio

Nigel D. Haig and T. L. Williams

Several different performance criteria have been proposed for assessing the quality of visual afocal sights. Earlier research by one of the authors (Haig) has shown that a high degree of correlation exists between a subjective assessment of performance and the Strehl intensity ratio of the optical system. We discuss some of the problems in choosing an objective performance criterion for visual sights and describe equipment that has been developed for measuring the line Strehl ratio of binoculars, both on and off axes. The equipment can be modified for testing other types of visual sight. It can also be used for measuring several additional performance parameters such as the modulation transfer function, transmission, and field curvature.

Key words: Performance testing, visual sights, binoculars, Strehl intensity ratio, modulation transfer function, transmission, field curvature, computer controlled.

1. Introduction

A primary aim of the optical designer is to devise a design that translates a given specification into an optical system that most efficiently meets that requirement. This aim appears to be perfectly straightforward and innocuous, but behind that simple statement lies a minefield of problems. For example, how do we know whether the specification actually meets the requirement? How does the designer relate the design to probable performance? Even worse, how do we know if a given quality test can measure an appropriate quantity?

The chain of introduction into service or production, for afocal visual systems at least, is often incomplete but should generally correspond to the following:

(1) The user identifies a specific requirement in his/her own descriptive way.
(2) An optical physicist translates the user's perceived need into specification parameters that are consistent with the need and with one another.
(3) The optical designer sketches out a simple preliminary design layout to meet the most fundamental parameter specifications.
(4) If the design attains likely feasibility in terms of both probable performance and cost, the preliminary layout is put into a more realistic form that can be manufactured.
(5) The design is optimized against standard criteria that take account of the characteristics of the human visual system.
(6) The manufactured instrument is tested against the specified parameters with quality-control methods that are unequivocally linked to human visual and perceptual performance.
(7) Instruments that meet the test criteria are then issued for service.

Any one of these stages may be deficient in some way, but in this paper we are concerned particularly with (2), (5), and (6). In particular, (5) and (6) should be closely linked, because it is vital that the designer optimize by means of a criterion that relates closely both to the test criterion and, more important, to actual human performance. It is also important that this research is concerned only with optical instruments of relatively high quality, with performances approaching the diffraction limit. For reasons that become clear below, the proposed test technique is inappropriate for visual instruments with relatively poor optical quality.

N. D. Haig is with the Defence Research Agency, Fort Halstead, Sevenoaks, Kent, UK. T. L. Williams is with Sira, Ltd., South Hill, Chislehurst, Kent BR7 5EH, UK.

Received 23 June 1994; revised manuscript received 7 September 1994.
0003-6935/95/101728-13$06.00/0.

2. Background

The assessment of optical performance is a comparatively straightforward task when the detector assembly can be perfectly specified. For example, the (focal) optical system serving a particular CCD detector array may be tested at accurately measured conjugates, with light of a specific wavelength, by a variety of methods. Normally such methods would include optical transfer function [MTF/phase transfer function (PTF)] at some specified spatial frequency, point-spread function/line-spread function (PSF/LSF) scanning, or some form of wave-front aberration measurement, such as interferometry. All these methods may be related to one another mathematically, if care is taken in attaining the specified test conditions.

The assessment of the performance of optical instruments for visual applications is a far less analytic process. The difference lies in the form of the detector, which is now the human eye and brain. Because the optical instrument and the eye are coherently coupled, they both form two separate parts of the same instrument. Now the detector array, known as the retina, is sequestered inside the eye, hidden behind an optical system that is complete in its own right. The problem is clearly revealed as the inaccessibility to conventional measuring instruments of the retina.

Any instrument-assessment method that tests only the man-made optics, without taking account of the effects of the eye and the brain, cannot properly be said to be measuring the true performance of the instrument–eye combination. For example, the most comfortable viewing distance for the eye is, in most viewing conditions, ~1 m, whereas the optical instrument may be focused to deliver its best image at some other distance. Thus simply manufacturing an afocal visual system to deliver a plane wave front, equivalent to an image at infinity, may be inappropriate and has indeed been shown to be so for military binoculars.[1]

One technique often used to bypass this problem is to render the afocal instrument focal by the addition of a good-quality decollimator. In theory this is somehow supposed to represent the optics of the eye, so that the PSF/LSF may be scanned and the corresponding MTF computed. Now it is inconvenient on a production line to measure the MTF at all spatial frequencies and at an infinite number of image planes. It seems therefore that if the MTF is to be used to test afocal visual optics, two crucial questions must be answered:

First, if we do not know how the eye chooses to focus, how do we select the appropriate focus of the MTF apparatus?

Second, the MTF is a function of spatial frequency, so, over what range of spatial frequencies should we test?

Apart from the breathtaking assumption that the eye acts as a simple camera, there is the matter of deciding the appropriate parameters of the MTF to apply. For example, the MTF must be measured at a given back-focal distance from the decollimator to correspond with the required defocus, and the appropriate optimum spatial frequency for the measurement must be selected. We take the latter problem first. The selection of the optimum spatial frequency is little more than an arbitrary choice, because even a glance at the enormous range of human contrast-sensitivity-function curve shapes (e.g., Ref. 2) (see Fig. 1), covering just the normal span of scene illuminances, demonstrates that the entire range of resolvable spatial frequencies, over 2 orders of magnitude, contributes to the final image. Only at the lowest light levels can a relatively restricted band of spatial frequencies be said to dominate, and even this band covers 1 order of magnitude.

In tackling the associated problem of which image plane to choose, for us to make the MTF measurement, different spatial frequencies usually give peak responses at quite different axial image planes, making the combined choice of spatial frequency and image plane an almost pure lottery. Aside from this, the eye is a constantly changing system with continuous shifts in focus and position. Note that the eye requires constant movement to function; keeping the eye stationary with a static image causes the image to disappear within seconds.[3]

In addition, it was established in a series of independent measurements made on a number of gunsights taken from a production run[4] that subjective MTF and objective estimates of contrast degradation could not be reconciled in any way. The gunsights were of high optical quality, as judged by human inspectors viewing test charts. Careful measurement by experienced inspectors demonstrated that not only was there little correlation between the MTFs of supposedly similar instruments in the production run but there was almost no correlation between the MTF's measured at similar field positions on opposite sides of the axis of the same instruments[4] (see Fig. 2).

The inescapable conclusion of the gunsight measurements was that no combination of MTF measurements, at any image distance or at any single (or few) spatial frequency, could mimic the results by human

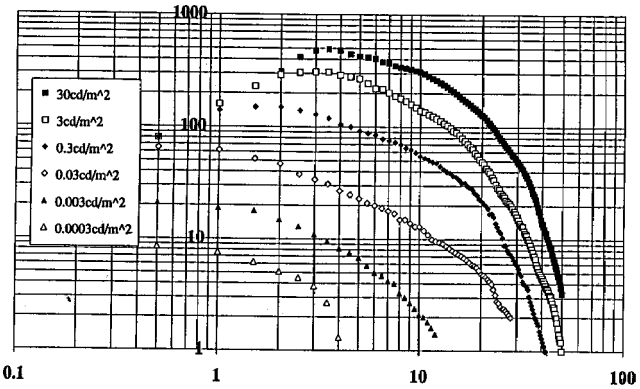

Fig. 1. Typical human contrast-sensitivity-function responses adopted from Bouman and van Nes.[2]

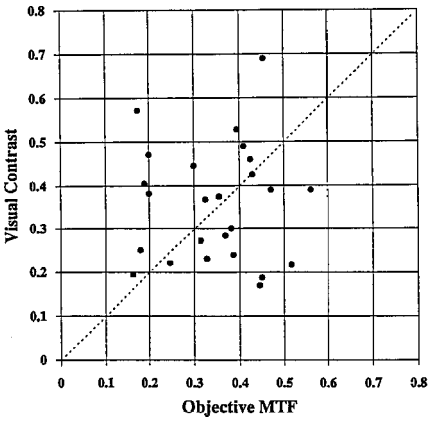

Fig. 2. Correlation of subjective/objective thresholds at 22.6 cycles/deg adopted from Burton and Haig.[4]

inspectors. The problem that remained was to discover precisely what form of metric most closely matched the performance of the human visual system.

3. New Quality Metric

The theory and results of research into a possible new afocal instrument test method were reported by Burton and Haig[5] and Haig and Burton.[6] In brief, the results of the research confirmed that MTF was inadequate and inappropriate as a representation of human visual performance. More significantly, it was shown that for optical instruments of moderate to high quality the metric that most closely corresponded to human visual discrimination was the Strehl ratio. This was found to be so by a technique that permitted not only isolated wave-front aberrations to be simulated but also combinations of aberrations. The results were unequivocal for all of them provided that the human visual system was being worked quite hard. In other words, the technique worked well for instruments operating near the diffraction limit but was not tested in conditions for which the human eye was not being pressed to perform. In practice, this means that, when an optical instrument is to be used to discriminate fine details (high-spatial-frequency structure), the Strehl ratio is the ideal and appropriate test.

The reason for qualifying the above statements was found by Haig,[7] who illustrates the PSF's for a range of wave-front aberrations. Many, if not all, PSF's develop multiple peaks, as the aberration magnitudes increase, thus rendering meaningless the concepts of both peak intensity and point contrast. As pointed out in Ref. 7, the Strehl ratio is a meaningful term only for well-corrected and properly focused systems, since it is only single-valued for pure defocus of less than 0.75 wavelengths. Thus for all practical purposes we take the general limit of application of the Strehl ratio technique to correspond with the Marechal criterion[8] of a Strehl ratio of greater than 0.8, as recommended by Hopkins[9] and Hopkins and Zalar.[10] This is why the method is applicable only to the assessment of good-quality instruments.

We now come to the precise value of the visual quality metric that will normally be applied as the assessment criterion. Clearly, because the eye regulates its own focus according to the Strehl ratio and the ratio must exceed 0.8 to remain single valued, the eye is a most sensitive discriminator. This fact was apparent to Marechal[8] and also to Kuwabara.[11] Kuwabara in particular became aware that the Strehl ratio was a most important quality metric and explored its usefulness in the presence of other forms of aberration. However, neither Marechal nor Kuwabara was able to determine the precise value of the ratio that gave rise to a measurable change in the image as perceived by the human observer. As pointed out by Hopkins[9] the image tolerance criterion suggested by Marechal of a Strehl ratio of 0.8 amounts in practice to a wave-front aberration of approximately $\lambda/4$. This is a value that is easily measurable and easily computed and that usually gives good imagery, as judged by human inspection.

Thus we now have a lower limit for acceptable visual performance, and it is necessary to undertake some careful research to discover the threshold value of the Strehl ratio. As reported by Haig and Burton,[6] the liminally discriminable upper value of the Strehl ratio was found to be 0.87, regardless of the type or combination of wave-front aberration. This means in practical terms that any afocal optical instrument capable of delivering a Strehl ratio of better than 0.87 is essentially perfect when used by a human observer with good eyesight.

Recently a critique of the applicability of the Strehl ratio, in favor of a restricted form of the MTF, was published by Mouroulis and Zhang,[12] but they compromised their results by restricting severely the range of the MTF spatial frequencies (from 5 to 24 cycles/deg) and allowing Strehl ratios of considerably less than 0.8. We have already explained why it is necessary to consider the entire range of discriminable spatial frequencies (from 0 to 60 cycles/degree) and why it is inappropriate to consider Strehl ratios of less than 0.8. Mouroulis and Zhang are still not yet able to define the conditions in which (MTF) measurements are meaningful, yet Haig and Burton[6] have demonstrated that the measurement of the Strehl ratio, provided it is kept in the physiologically meaningful band of greater than 0.8, is simple and effective. It is simple because no arbitrary decisions must be made concerning the plane of best focus or the best spatial frequencies to use, and it is effective because one can automate the measurement by allowing the test machine to move automatically to the position of best focus by monitoring the Strehl ratio.

4. Description of the Basic Strehl Measurement Equipment

4.A. Theoretical Basis of Measurement

Initial research has concentrated on measuring the line Strehl ratio (LSR) rather than the point Strehl ratio (PSR), mainly to ease the problems associated with the low light levels associated with a point rather

than a line object. However, the basic measurement method described in this paper is also applicable to the measurement of the PSR, provided appropriate changes are made to the algorithms used in the computations.

The LSR is defined as the following ratio: (peak intensity in the LSF of the test piece)/(peak intensity in the LSF of the same system if it were diffraction-limited). The peak intensity in the LSF refers to the situation in which the slit object has negligible width. In a practical situation slits with finite widths must be used to permit the passage of adequate light flux, and a relationship is required that permits one to define the LSR in terms of peak intensities measured with such slits. The relationship we used is one derived by Bryam,[13] who showed that the peak illuminance in the LSF of a diffraction-limited optical system, when the slit object has a finite width, is given by

$$I = 4I_b\gamma(zo)/\pi^2, \tag{1}$$

where I_b is the illuminance in the image of a wide slit or aperture (i.e., the illuminance if the slit object is removed) and zo is the reduced angular width of the geometric image of the slit, i.e.,

$$zo = \pi \mathrm{NA} w/\lambda, \tag{2}$$

where NA is the numerical aperture of the image-forming beam, w is the geometrical width of the slit in the image plane, and λ is the wavelength of the light; $\gamma(zo)$ is defined as

$$\gamma(zo) = \sum [(-1)^n 2^{(2n+2)} zo^{(2n+1)}/1^2 3^2 \ldots (2n-1)^2(2n+1)^3(2n+3)], \tag{3}$$

where the summation is over all n from $n = 0$.

If the peak illuminance in the measured LSF of the test piece (i.e., the aberrated image of the slit) is given by I_m, then using Eq. (1), we have

$$\mathrm{LSR} = I_m/I = I_m\pi^2/4I_b\gamma(zo). \tag{4}$$

From Eq. (4) and the other associated equations we see that, provided we know w, NA, and λ, we can determine the LSR by measuring I_m and I_b.

This forms the basis for the measurement technique described in this paper. It is in essence the measurement technique sketched by Haig and Burton.[6] They proposed using the Kuwabara[11] graticule for this purpose. Kuwabara combined on the same graticule a narrow, just resolved, slit and a uniformly illuminated aperture that is sufficiently large that the illuminance in the image is not reduced by diffraction or aberrations. In practice the use of this type of graticule is very restrictive because a range of different slit widths and spacings between the slit and aperture are necessary to accommodate the requirements of different test pieces. In addition it has several disadvantages including the possibility of light spreading from the image of the aperture to that of the slit and the need to have very high illuminance uniformity across its area. For these reasons the test method described below uses two separate graticules, one with a slit and one with a large aperture.

4.B. Line Strehl Ratio Measurement Technique

The basic arrangement for measuring the LSR is illustrated in Fig. 3. The source is a quartz tungsten halogen lamp inside a 100-mm-diameter integrating sphere. The sphere has a 25-mm output port that acts as a very uniform high luminance source. In front of the source is a heat filter followed by a five-position filter wheel. The latter has one blanked-off position and four positions in which monochromatic filters can be mounted. Immediately in front of the filters is another turret wheel that permits one of three target patterns to be positioned in front of the source (i.e., on the optical axis of the system). The three target patterns are a vertical slit, a horizontal slit of the same width, and finally a clear, relatively large, circular aperture. A facility is available for mounting a calibrated neutral-density filter in the clear aperture to adjust the luminance level so that signal levels measured at the detector are of similar magnitude for both target patterns.

The optical system being tested images one or the other of these target patterns in the object plane of a microscope objective, which in turn relays a magnified image of the test pattern into the focal plane of a silicon CCD array camera. The camera, with the microscope OG attached, is mounted on a micrometer stage that has manual x, y movements and a motorized z (focus) movement. (The overall signal-processing and control system is illustrated in Fig. 4.)

The video signal from the camera goes to a PC, fitted with a frame grabber, where the signals are processed to determine the LSR. Two measurements are needed to determine the LSR. In the first measurement one or the other of the slits is used as the test pattern, and the software processes the image of the slit captured by the camera to determine the peak signal in the resulting LSF. This is taken as the value for I_m in Eq. (4). The other measurement is made with the circular aperture, and in this case the average signal over the central part of the image of the aperture is determined. This signal is taken

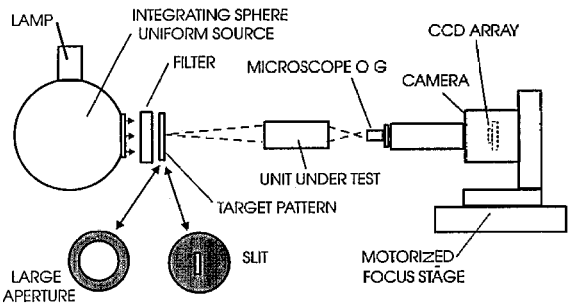

Fig. 3. Diagram of the basic arrangement for measuring the Strehl ratio.

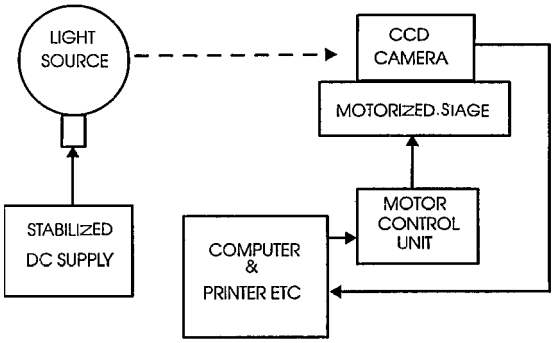

Fig. 4. Block diagram of the control and signal-processing system.

as the value for I_b in Eq. (4). In practice three measurements are actually made, the third one being to determine the dark signal from the camera by blanking-off the source. This dark-level signal is used to correct the values of I_m and I_b.

The operator enters the remaining data necessary to compute the LSR (i.e., the width of the slit and the appropriate NA) through the keyboard.

4.C. Correcting for the MTF of the CCD Camera

The performance of the detector array used to measure the peak intensity in the image of the slit itself produces a drop in the height of the peak, and a factor should be applied to a measurement to correct for this.

If C is the correction factor, LSRm is the measured Strehl ratio [by use of Eq. (4)], and LSRc is the corrected Strehl ratio, we have

$$\mathrm{LSR}c = \mathrm{LSR}m \times C. \tag{5}$$

Because the LSR is equal to the integral under the equivalent MTF curve divided by that under the MTF for the diffraction-limited system, we have

$$\mathrm{LSR}m = \int \mathrm{MTFtp}(s)\mathrm{MTFd}t(s)[\sin(\pi ws)/(\pi ws)]\mathrm{d}s \Big/ \int \mathrm{MTFdif}(s)[\sin(\pi ws)/(\pi ws)]\mathrm{d}s, \tag{6}$$

$$\mathrm{LSR}c = \int \mathrm{MTFtp}(s)[\sin(\pi ws)/(\pi ws)]\mathrm{d}s \Big/ \int \mathrm{MTFdif}(s)[\sin(\pi ws)/(\pi ws)]\mathrm{d}s, \tag{7}$$

where MTFtp(s) is the MTF of the test piece, MTFdif(s) is that of the equivalent diffraction-limited test piece, and MTFdt(s) is the MTF of the CCD array (i.e., the detector). Both the spatial frequency variable s and the slit width w must refer to the same image or object plane.

Using Eqs. (5)–(7), we see that the correction factor is given by

$$C = \int \mathrm{MTFtp}(s)[\sin(\pi ws)/s]\mathrm{d}s] \Big/ \int \mathrm{MTFtp}(s)\mathrm{MTFd}t(s)[\sin(\pi ws)/s]\mathrm{d}s. \tag{8}$$

To use Eq. (8) we must know the MTF of the test piece and that of the array. Provided that the test piece is reasonably well corrected, one can consider using the theoretical diffraction-limited MTF for the test piece instead of its actual MTF. MTFdt(s) needs to be measured, but this can be a once and for all measurement (at least for a given wavelength).

If we make this assumption (i.e., that the test piece is reasonably well corrected), we have (for an optical system with a circular pupil)

$$\mathrm{MTFtp}(s) = \{2\arccos(s'/2) - \sin[2\arccos(s'/2)]\}/\pi, \tag{9}$$

where the reduced spatial frequency

$$s' = s(\lambda/\mathrm{NA}). \tag{10}$$

NA is the numerical aperture of the test piece, and λ is the wavelength of the light. Note that all parameters must refer to the same image plane.

In most cases C differs from unity by only a few percentage points, and the error in assuming that the test piece is diffraction limited is therefore even smaller. Moreover this error tends to undercorrect rather than overcorrect the value of the LSR.

Note that, owing to the inherent anamorphic design of the CCD camera chip, the MTF of a CCD camera may be significantly different in the vertical and horizontal directions and that a different correction factor therefore applies for the LSR measured with a vertical or horizontal slit.

4.D. Software Features

The software provides two main measurement options. The first option is to make a single LSR measurement at a set focus position, whereas the other is to perform a series of measurements at a specified number of focus positions covering a specified focus range. In the latter case the process of driving the camera to the different focus positions and making the LSR measurement can be completely automatic. When the through-focus measurements are completed the software automatically fits a cubic spline curve through the measured data and from this predicts both the optimum focus position and the value of the LSR that would be obtained at that position.

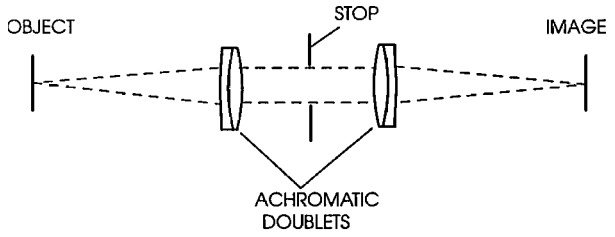

Fig. 5. Lens arrangement used for evaluating the Strehl measurement technique.

Other features of the software are as follows:

- A real-time image of the camera output is generated on the PC display.
- The operator can set up a rectangular box on the display, which defines the area of the image used for the measurements. When the slit target is vertical the width of this box defines the width over which the LSF and associated circular aperture signals are measured, whereas the height of the box defines the number of rows of the camera array that are averaged together to give the measured LSF or circular aperture signal. With a horizontal slit the significance of the height and width of the box is appropriately interchanged.
- The peak of the LSF used in calculating the LSR can be either the actual peak value of the signal or a predicted value from fitting a spline curve through all the measured points within an operator-specified percentage of the peak. This provides for better repeatability when signals are noisy.
- All measurements can be the result of averaging the signal over an operator-specified number of TV frames.

5. Validation of the Measurement Technique by Use of Special Test Pieces

To evaluate the validity of the technique experimentally, we performed measurements of LSR on several different test pieces. They were of relatively simple design and specially constructed so that the actual values of the LSR should be close to the theoretical values. The general form of these systems is illustrated in Fig. 5. A comparison of the measured and theoretical values of a through-focus LSR at a monochromatic wavelength of 546 nm for one of these test pieces is shown in Fig. 6. The difference between the two curves is easily explained by the buildup of errors from uncertainties in the values for slit width, test piece NA, theoretical computations, etc.

6. Equipment for Testing Binoculars

In view of the satisfactory results from these earlier measurements, it was decided, with the use of this technique, to build a test facility for binoculars to measure the LSR. The aim was to design a facility that would allow a range of different binoculars to be tested both on and off axis. The optical-bench arrangement is illustrated diagrammatically in Fig. 7. The source, filters, and target patterns are exactly as described in Subsection 4.B.

The target patterns are positioned at the focus of a 1500-mm, focal-length refracting collimator (two-element airspaced achromat) with an aperture diameter of 100 mm. The collimator is mounted in an angle bracket fixed to a carrier that can be moved along a length of a lathe-bed-section optical bench. The latter runs parallel to the optical axis and permits the collimator focus to be accurately set with respect to the target patterns.

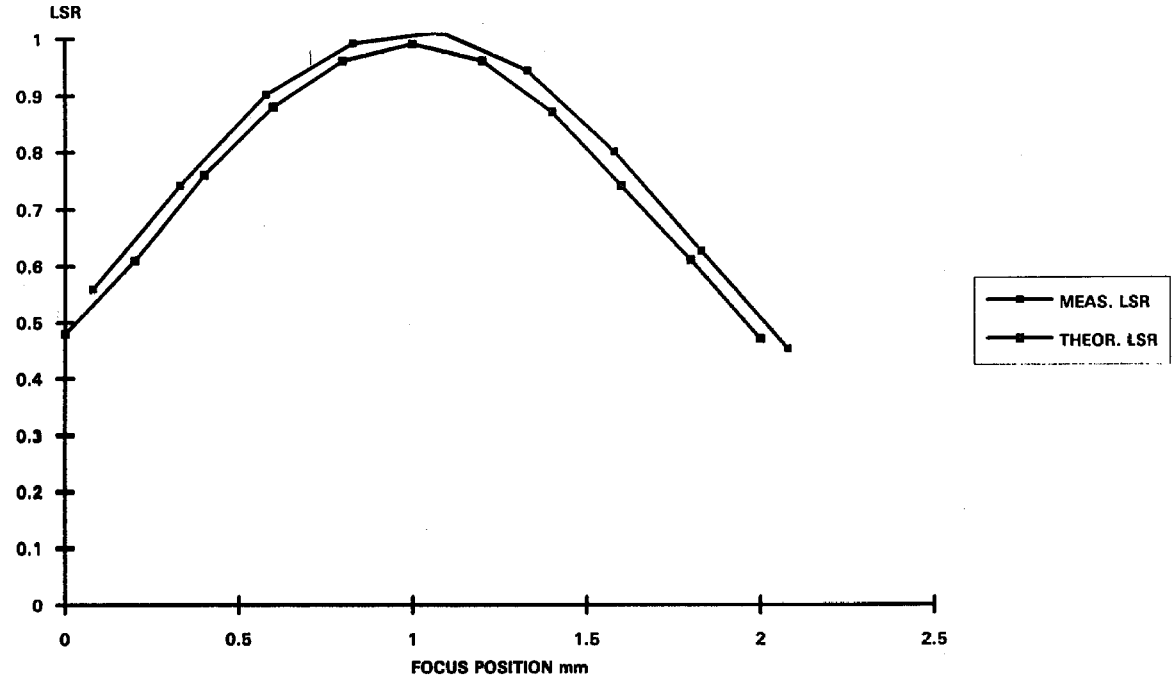

Fig. 6. Comparison of the measured and theoretical through-focus LSR for the lens illustrated in Fig. 5.

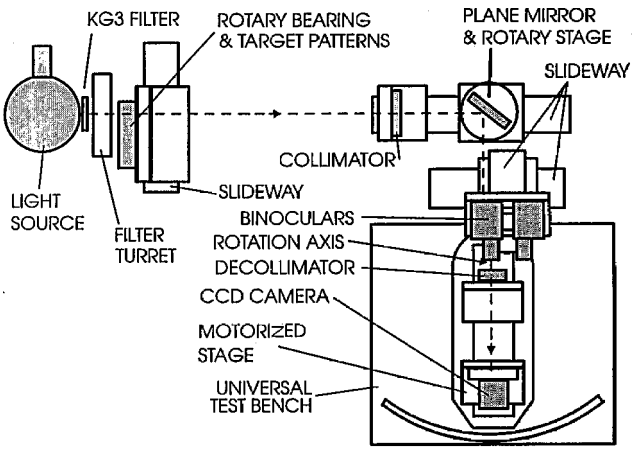

Fig. 7. Facility for measuring the LSR etc. of binoculars.

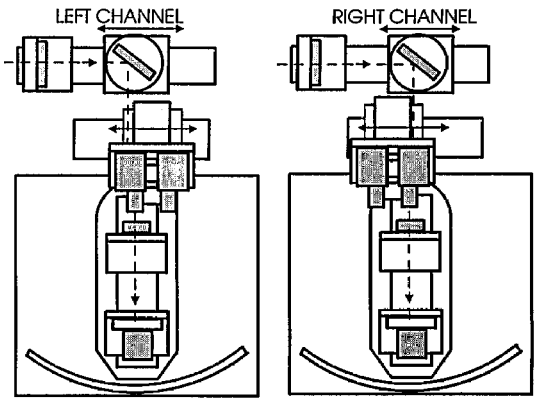

Fig. 8. Method of changing between oculars.

On the same length of optical bench is a second carrier on which is mounted a rotary stage (with a vertical axis) and a 150-mm-diameter plane mirror. The mirror is set at ~45 deg to the light beam from the collimator and directs this beam to the entrance aperture of the binocular being tested. The rotary stage has a calibrated scale so that it is possible to rotate the plane mirror to place the test pattern at any desired angular position (along a horizontal plane) in the field of view of the unit being tested. The mirror mount itself has fine tilt adjustments about two axes. Particularly important is the tilt about a horizontal axis, which allows the direction of the collimated beam to be adjusted in a vertical plane.

The provision for sliding the rotary stage and mirror assembly along the lathe bed section permits either side of a pair of binoculars to be illuminated. It also allows one to compensate for beam movement when testing off axis, although, for the relatively small field angles normally required, the beam does not move far enough to cause vignetting.

The binoculars being tested are held in a mount that comprises a plate with two stepped holes into which the front ends of the binoculars are a close fit, attached to a carrier. A second plate, with a rectangular slot in it that clears the two eyepieces, is attached to the first by two lengths of threaded rod and serves to clamp the binoculars to the first plate. The carrier fits on a lathe-bed-section slideway that runs parallel to the optical axis and that is itself attached to a second carrier running on a slideway perpendicular to the optical axis. By moving the binoculars along the latter, it is possible to bring one or the other of the binocular channels into the test position as illustrated in Fig. 8. The purpose of the slideway running parallel to the optical axis is to enable the exit pupil of the binoculars to be positioned over the axis of rotation of the universal test bench (see below).

The collimated beam emerging from the relevant eyepiece of the binocular goes to an achromatized doublet decollimating lens that either focuses an image of the test pattern directly onto the focal plane of a CCD camera or produces an aerial image that is relayed to the focal-plane array by a low-power microscope OG. In the latter case the OG is attached directly to the lens mount of the camera by a suitable tube and adapter.

The decollimating lens and the CCD camera form part of an assembly that is basically a modified version of the Sira universal test bench. The latter is a unit designed to permit MTF testing of lens objectives to be done on and off axes with a camera-bench configuration. In this case its use permits the binocular to be tested both on and off axes. It comprises a platform that is pivoted about a vertical axis at its front end and is supported by a single roller bearing, running on a flat base plate, at its rear end. This arrangement permits the platform to be swung through angles to ± 35 deg.

The decollimating lens is mounted on a carrier supported on a slideway running along the length of the platform. The mount for this lens includes a micrometer stage giving movement in the horizontal Z (focus) direction and a second micrometer stage giving adjustment in the vertical direction.

The CCD camera is mounted on the same slideway through a motorized focus stage. The camera mount includes a micrometer stage giving adjustment in the horizontal X direction (i.e., perpendicular to the focus direction), a vertical micrometer stage, and finally a rotary stage with its axis of rotation along the optical axis. The function of the latter is to permit the CCD

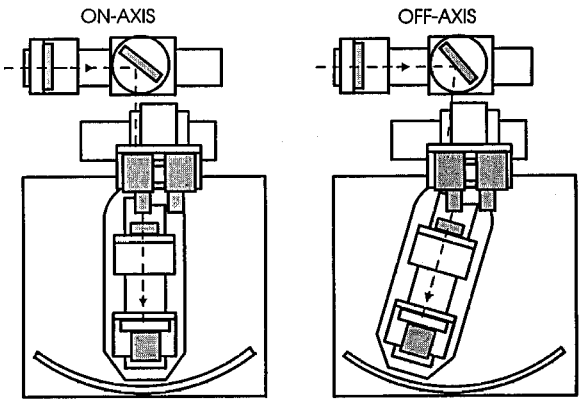

Fig. 9. Method of off-axis testing.

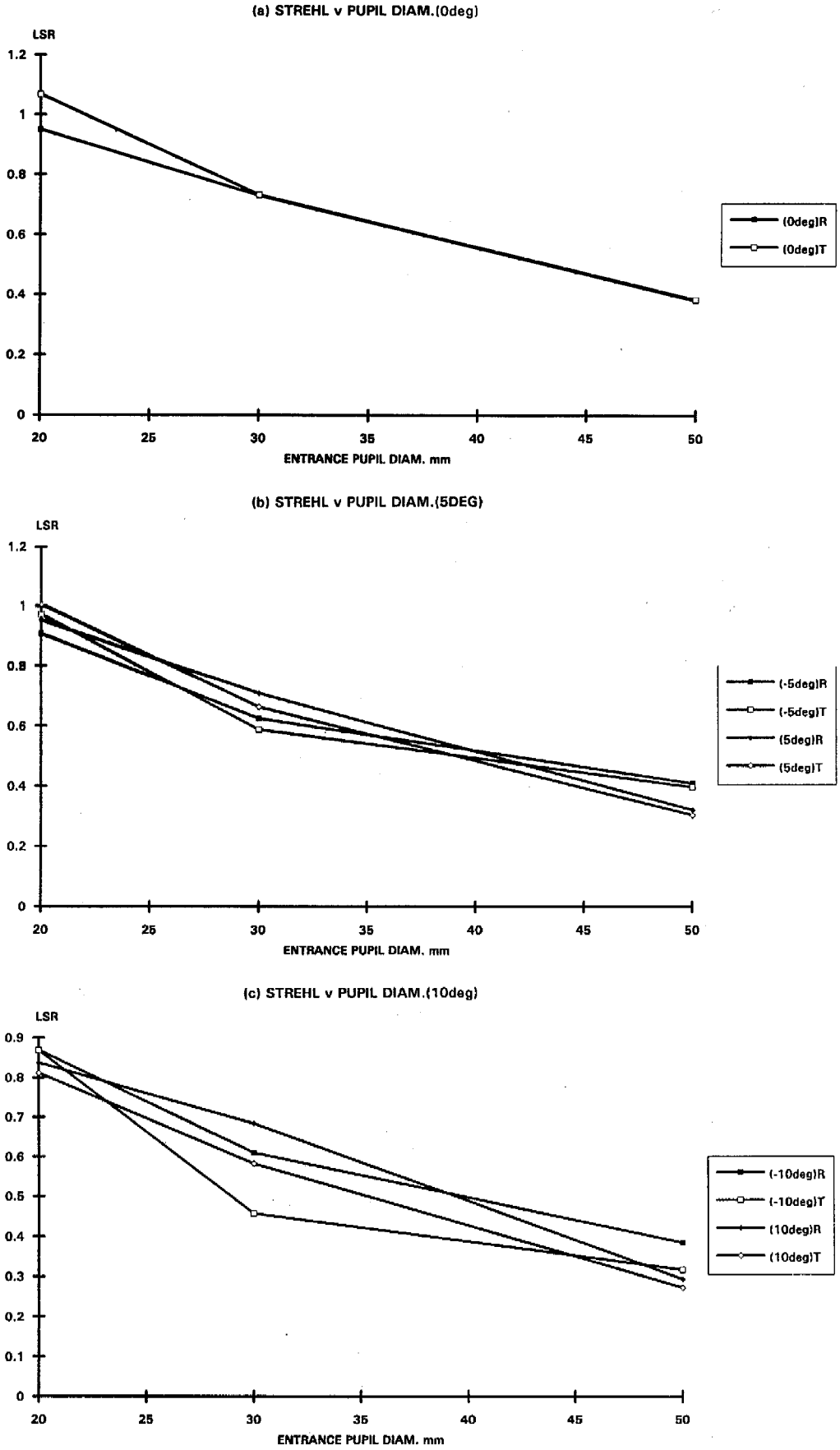

Fig. 10. LSR as a function of the entrance-pupil diameter for a pair of 10 × 50 binoculars: (a) on axis, (b) ±5-deg off axis, (c) ±10-deg off axis. *R* and *T* refer to measurements in the radial (sagittal) and tangential azimuths, respectively.

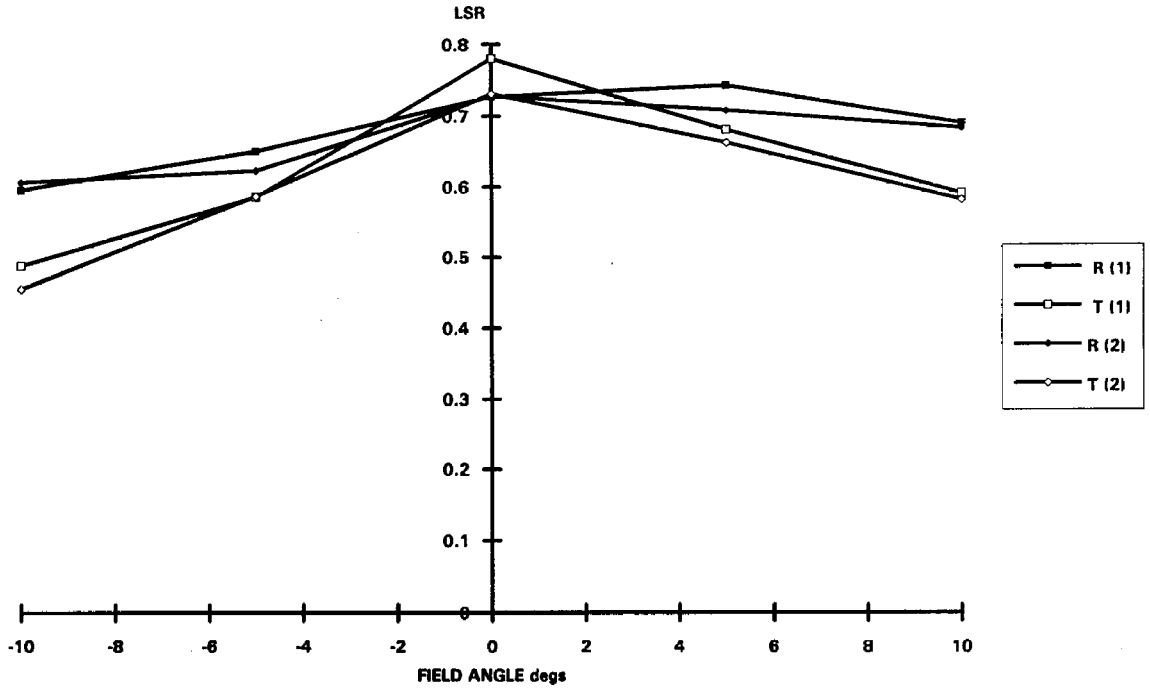

Fig. 11. LSR as a function of the image angle for a pair of 10 × 50 binoculars (entrance-pupil diameter, 30 mm).

camera to be aligned so that the pixel columns are vertical.

Note that the relevant exit pupil of the binocular being tested should be positioned over the vertical axis of rotation of the platform of the universal test bench.

The on- and off-axes testing positions are illustrated in Fig. 9.

As indicated above, the system can be used with the decollimator forming an image directly on the CCD array or through a microscope OG. Measurements were done with both arrangements; in the former case a 300-mm focal-length achromat was used, and in the latter case a 50-mm focal-length achromat was used in conjunction with a 25-mm focal-length microscope OG. The advantage of using a decollimator

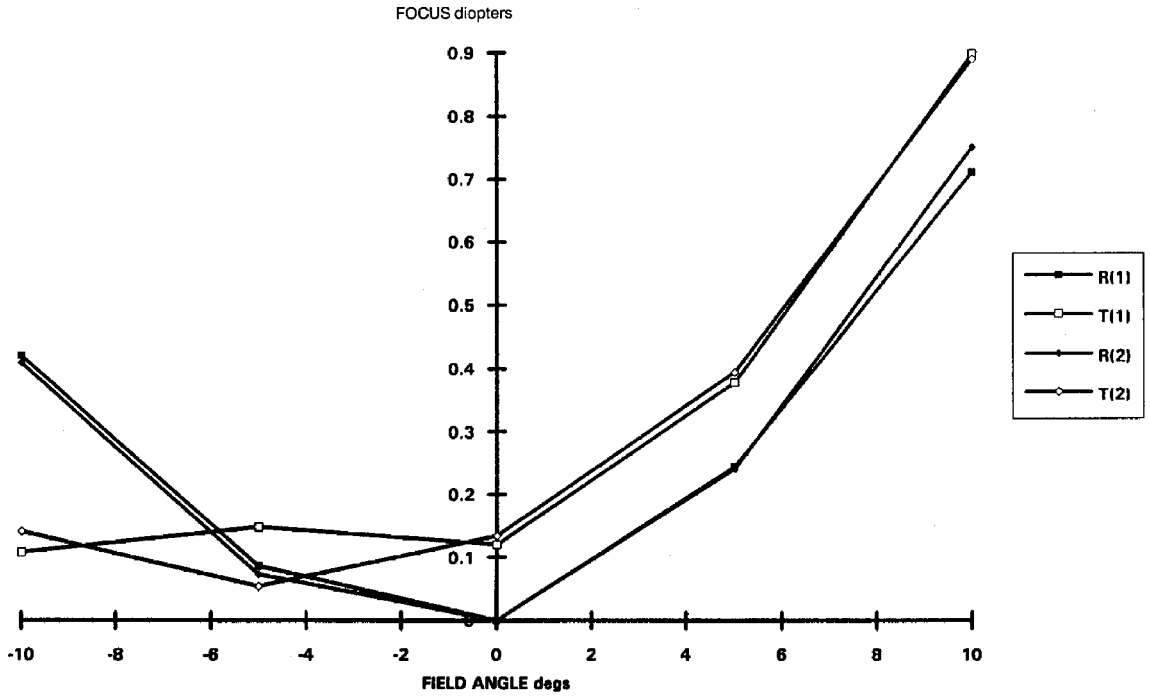

Fig. 12. Position of best focus as a function of the image angle for the measurements plotted in Fig. 11.

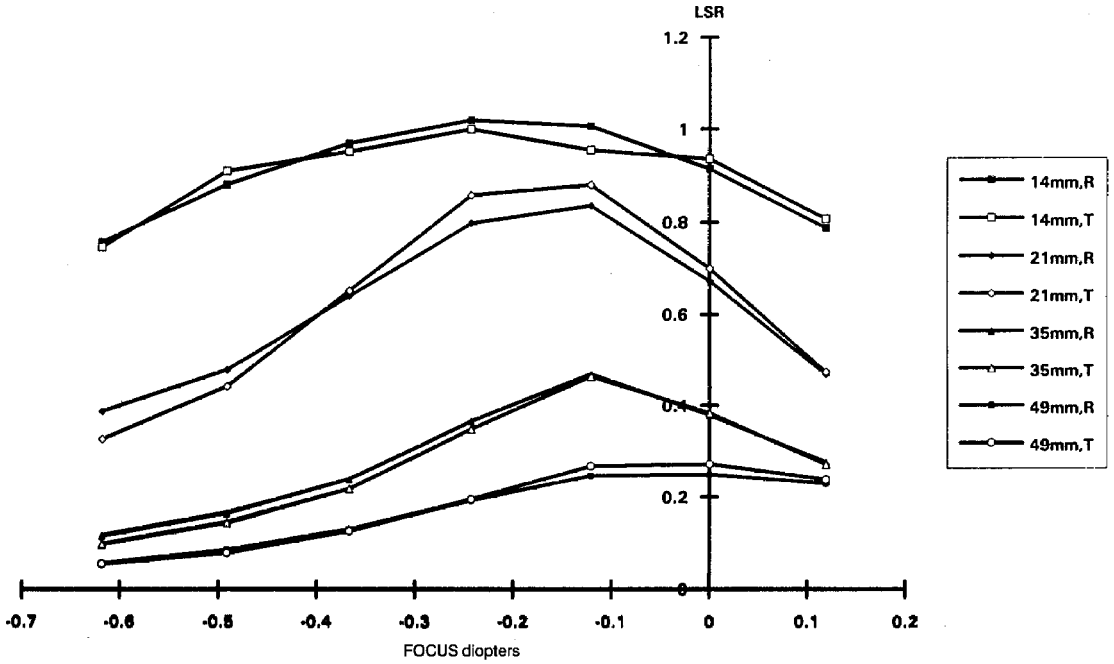

Fig. 13. Through-focus LSR for different entrance-pupil diameters for a pair of 7 × 50 binoculars.

lens on its own is that the mechanical arrangement is simplified, and, because there are fewer optical components, the possibility of aberrations that could affect the results is reduced. The disadvantage is that inconveniently large movements of the CCD camera are necessary to accommodate the focus changes that occur with most afocal systems as one goes off axis. In practice the use of a decollimator on its own was found to be workable for on-axis measurements but not for off-axis ones. The decollimator plus microscope OG proved to be a more useful general arrangement.

The control and signal-processing system is the same as that illustrated in block diagram form in Fig. 4.

7. Evaluation of the Binocular Test Facility

The equipment has been used to test a number of commercial and military binoculars including a set of six binoculars of nominally the same design. The

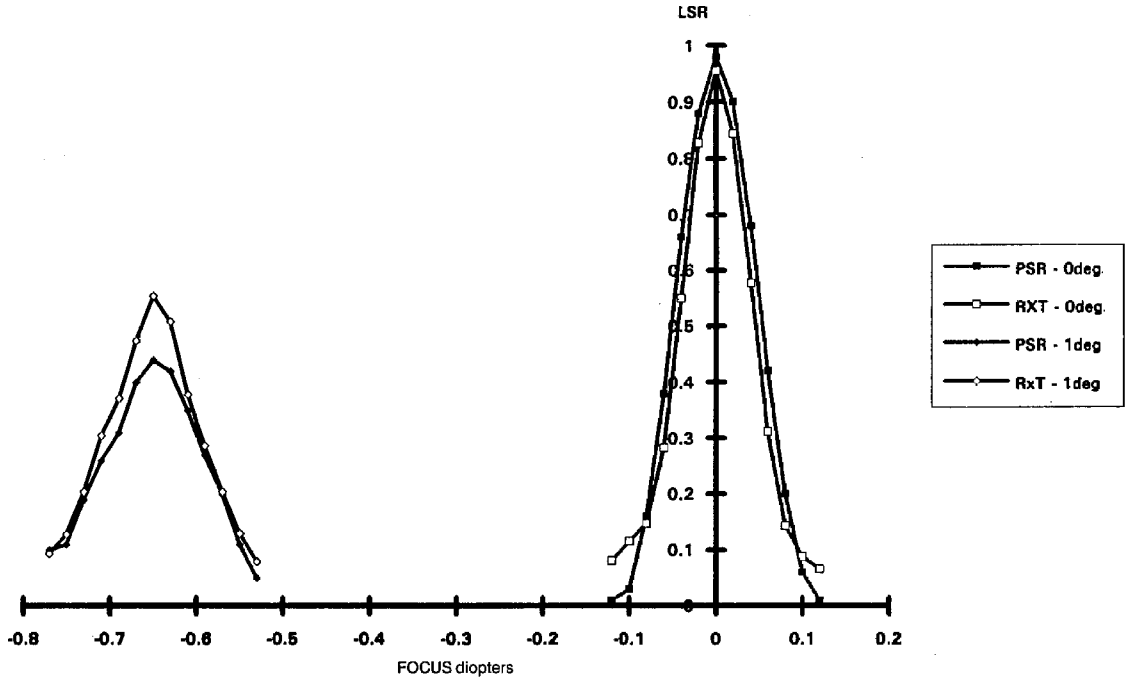

Fig. 14. Comparison of the theoretical values of the PSR and the product of the radial and tangential LSR's for a 10 × 60 telescope.

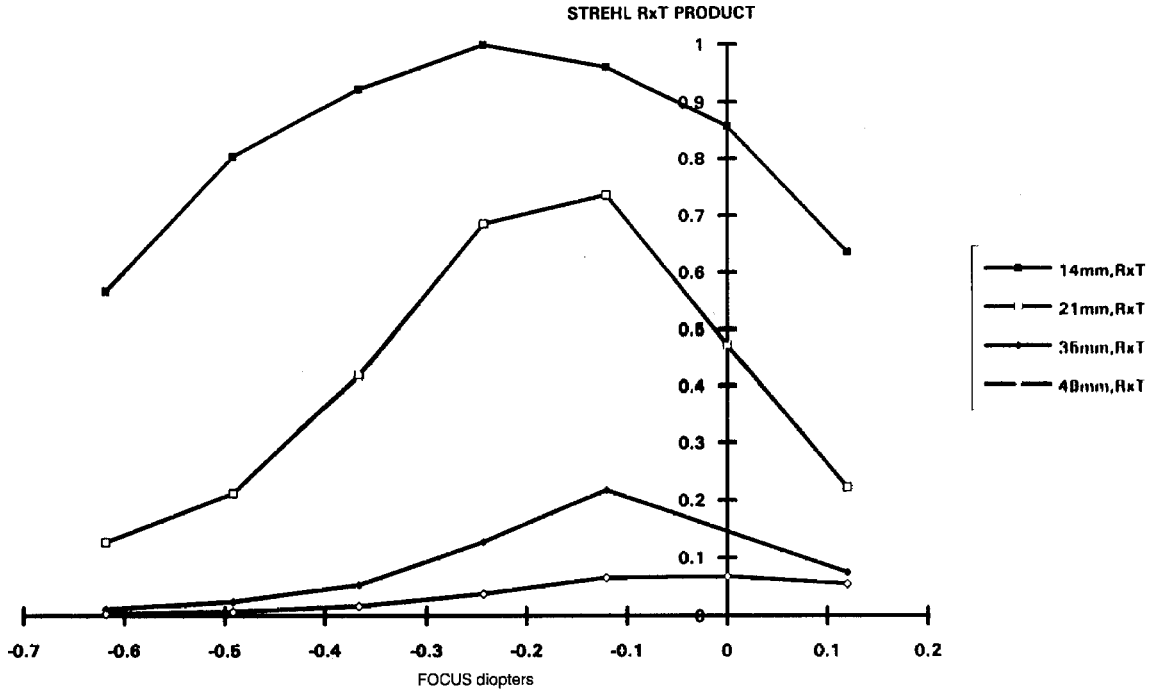

Fig. 15. Through-focus values of the product of the radial and tangential LSR's for the curves plotted in Fig. 13.

aim of these measurements was both to assess their repeatability and to obtain some background experience on the problems, if any, from use of a Strehl ratio as a means of determining the performance of binoculars.

Figure 10 shows plots of the LSR versus the entrance-pupil diameter for one ocular of a pair of 10 × 50 binoculars. The plots are for image angles of 0, ±5, and ±10 deg and were all done with a monochromatic 546-nm filter. They illustrate, as we found for most of the binoculars tested, that for exit-pupil diameters of 2 mm or less the units are close to being diffraction limited and that at full aperture they are far from being so. In most cases we found that the use of the equivalent of a 3-mm exit pupil gives a measure of performance that is sensitive

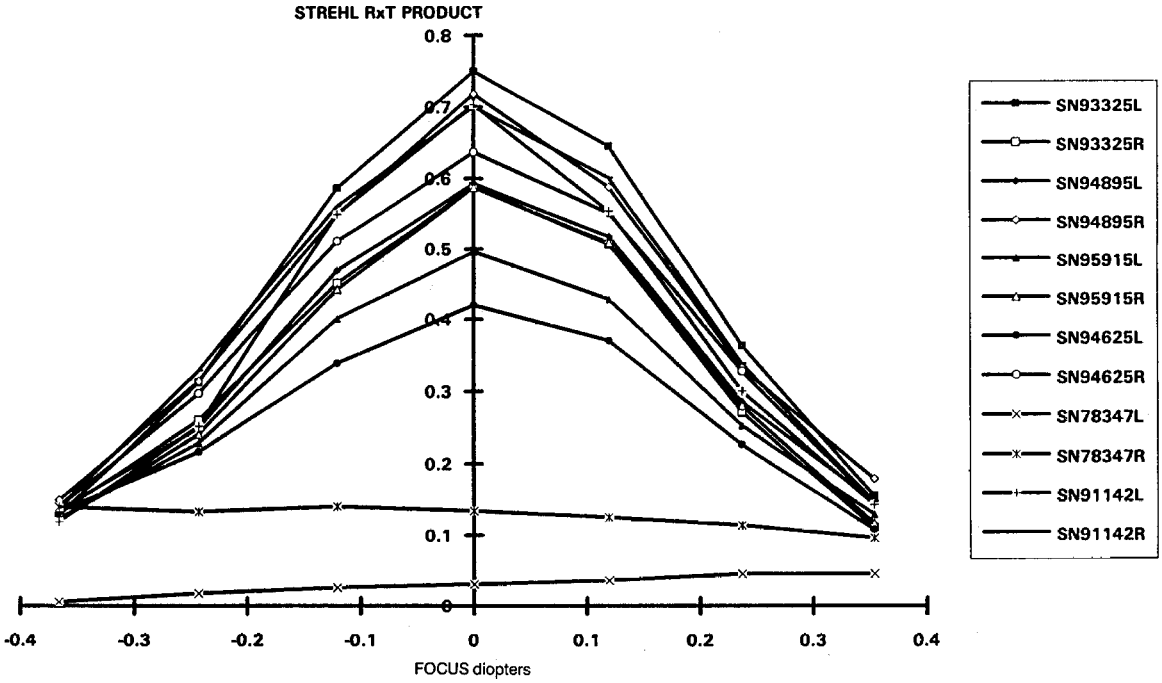

Fig. 16. Through-focus values of the product of the radial and tangential LSR's for both oculars of a set of six similar 7 × 50 binoculars. The numbers in the legend box refer to the serial numbers of the binoculars with *L* and *R* denoting the left and right oculars, respectively.

to the quality of both the design and the construction. Moreover, this is within the range of eye-pupil diameters for which critical performance is still required.

Larger exit pupils may be used when one wishes to assess the performance for low illumination levels. However, some care is required in extending Strehl-ratio measurements to situations in which imaging is relatively poor (see the comments in Section 3).

Figure 11 shows how the LSR for the same binoculars varies with the image angle. The measurements are for a 30-mm entrance pupil, equivalent in this case to a 3-mm exit pupil. Two sets of measurements are shown that were made several weeks apart with the binoculars having been removed from the equipment between measurements. The measurements are a typical example of the repeatability that is achieved in that situation. Figure 12 shows a plot of the best focus position versus the image angle for the same two sets of measurements. These measurements were made with the through-focus LSR option, so that both the peak value of the LSR and the focus position at which it occurred were available as output data.

Figure 13 shows a plot of an on-axis, through-focus LSR for a pair of 7×50 binoculars, which was one of a set of six similar binoculars used in assessing the equipment. All of these were tested at least twice (with both oculars) with a period of a few weeks between the repeat measurements on each unit and with the units removed and replaced on the equipment between measurements. The values of LSR on axis with a 21-mm entrance pupil (i.e., equivalent to a 3-mm exit pupil) were analyzed to obtain a figure for the repeatability of the measurements. The average of the absolute value of the difference between the repeat measurements on the same ocular was found to be 0.04 LSR units. Repeat measurements on the same ocular, made with the test piece remaining in position on the equipment, gave a repeatability (standard deviation) figure of 0.01 LSR units.

The actual value of the LSR obtained with this set of binoculars varied from 0.23 for the worst ocular to 0.96 for the best ocular with an average value for all oculars of 0.75 and a standard deviation of 0.19.

8. Combining Radial and Tangential LSR

Measurement of LSR in general yields different results for different slit orientations, and it is therefore necessary (as we described above) to perform measurements in at least two slit orientations (i.e., radial and tangential). Although this procedure provides useful information concerning the presence of astigmatism and other aberrations, it is an advantage from the point of view of simplifying test procedures to have a single figure of merit for each field position. An obvious way to do this is to measure the PSR directly. However, this is not necessarily an easy measurement to implement because of the relatively low light levels involved. The alternative is to combine the R and T LSR measurements in a simple meaningful manner. Various ways of doing this have been considered, the most promising appears to be to use the product of the two. The $R \times T$ product, as one might expect, relates relatively closely to the actual PSR. For example, Fig. 14 shows a comparison of the theoretical values of the LSR $R \times T$ product and the PSR for a 10×60 telescope (originally designed as an afocal test standard.[14]) The comparison is for the on-axis and the 1-deg off-axis conditions. For the on-axis case the peak values agree to within 0.025. For 1-deg off axis the agreement is not quite so good, although still reasonable, the peak values differing by 0.11. In both cases the best focus positions agree to better than 0.02 diopters.

Figure 15 shows a plot of the $R \times T$ product obtained from the LSR results plotted in Fig. 13, whereas Fig. 16 is a plot of the through-focus $R \times T$ products (an on-axis, 21-mm entrance pupil) for all the oculars in the set of six binoculars referred to above.

9. Making Transmission, Veiling Glare, and MTF Measurements

In this paper we are primarily concerned with the measurement of the Strehl ratio. However, we should mention that the binocular test facility has also been designed to make transmission, veiling glare index, and MTF measurements. The transmission and MTF measurements require no additions to the equipment other than to the software. One can determine transmission by measuring the ratio of the illuminance in the final image of the full output port of the integrating sphere, both with and without the binocular (a fixed stop must be placed at the exit-pupil position for both measurements). One can simply determine the MTF by measuring the full LSF and taking its Fourier transform. For meaningful measurements of the veiling glare index one requires the addition of a special source that provides a bright field subtending at least 90 deg at the entrance aperture of the test piece with an absorbing target at its center.

The equipment can in principle also be used for measuring angular magnification, distortion, vignetting, chromatic aberration, etc.

10. Conclusions

Earlier research by Haig and Burton[5,6] shows that the metric that most closely corresponds to human visual discrimination is the Strehl ratio. In this paper we have described a suitable technique for measuring the LSR, which is generally applicable to good-quality optical systems. We also described a new and automated test facility for testing binoculars and sights with this more appropriate and meaningful technique. If required, this system may also be used for measuring transmission and MTF and with some additions veiling glare and several other relevant performance parameters.

In this paper we have proposed the use of the product of the LSR in the radial and tangential azimuths as a means of obtaining a single-value

performance criterion for each field position. Indications are that this criterion approximates the PSR both in magnitude and focus position.

The technique described for measuring the LSR can in principle also be used to measure the PSR by use of small circular apertures instead of slits and appropriate equations for this situation. A practical implementation of such a measurement requires that the luminance of the source be increased and or the noise equivalent luminance of the CCD camera be reduced. Both options are possible and in particular the noise equivalent luminance achieved by the camera can be considerably improved when cooled CCD arrays and longer frame integration times are used.

References

1. R. Home and J. Poole, "Measurement of the preferred binocular dioptric settings at a high and low light level." Opt. Acta **24,** 97–98 (1977).
2. M. A. Bouman and F. L. van Nes, "Spatial modulation transfer in the human eye," J. Opt. Soc. Am. **57,** 401–406 (1967).
3. Troxler (1804), quoted in *The Eye,* H. Davson, ed. (Academic, New York, 1962), Vol. 2, Chap. 10, p. 197.
4. G. J. Burton and N. D. Haig, "Criteria for the testing of afocal visual instruments," in *Assessment of Imaging Systems II,* T. L. Williams, ed., Proc. Soc. Photo-Opt. Instrum. Eng. **274,** 191–201 (1981).
5. G. J. Burton and N. D. Haig, "Effects of the Seidel aberrations on visual target discrimination," J. Opt. Soc. Am. A **1,** 373–385 (1984).
6. N. D. Haig and G. J. Burton, "Effects of wavefront aberration on visual instrument performance and a consequential test technique," Appl. Opt. **26,** 492–500 (1987).
7. N. D. Haig, "The interferograms and spread functions of the sixth order aberrations," in *Assessment of Imaging Systems II,* T. L. Williams, ed., Proc. Soc. Photo-Opt. Instrum. Eng. **274,** 21–36 (1981).
8. A. Marechal, "Etude des effets combines de la diffraction et des aberrations geometriques sur l'image d'un point lumineux," Rev. Opt. Theor. Instrum. **26,** 257–279 (1947).
9. H. H. Hopkins, "The aberration permissible in optical systems," Proc. Phys. Soc. London Sect. B **70,** 449–470 (1957).
10. H. H. Hopkins and B. Zalar, "Aberration tolerances based on the line spread function," J. Mod. Opt. **34,** 371–406 (1987).
11. G. Kuwabara, "Studies on the image formed by lenses. 1. On the characteristics of an image and their quantitative representation," J. Opt. Soc. Am **45,** 309–319 (1955).
12. P. Mouroulis and H. Zhang, "Visual instrument quality metrics and the effects of coma and astigmatism," J. Opt. Soc. Am. A **9,** 34–42 (1992).
13. G. M. Bryam, "The physical and photochemical basis of visual resolving power. Part 1. The distribution of illumination in the retinal images," J. Opt. Soc. Am. **34,** 571–591 (1944).
14. T. L. Williams, M. L. Nunn, and N. P. Barton, "An afocal-system OTF test standard," Opt. Acta **25,** 1097–1111 (1978).