PHOTOGRAPHIC SCIENCE AND ENGINEERING • Volume 16, Number 3, May–June 1972
pages 221-230

An optical merit function (SQF), which correlates with subjective image judgments

E. M. Granger and K. N. Cupery
Eastman Kodak Company, Kodak Apparatus Division Research Laboratory,
Rochester, New York 14650

A Subjective Quality Factor (SQF) was developed as the result of a search for an objective figure of merit which could be easily calculated and directly measured in practice and which would correlate with subjective rank regardless of MTF form. A number of experiments were performed to test the quality factor for a wide variety of MTF shapes. These tests included nonsymmetric, two-dimensional images. The results of the experimental program were that SQF was able to predict image quality within normal reader error and was 0.988 correlated with the measured data. At present, the quality factor tests have been limited to physically realizable optical systems producing grain-free images on black and white film products and to a range of image quality defined by an Image Sharpness Scale (ISS).

DESCRIPTORS: **image quality · image evaluation · sharpness · optical transfer function**

I. Introduction

Ever since the introduction of the Optical Transfer Function (OTF), efforts have been made with varying success to provide an analytical bridge between the OTF and a subjective measurement of image quality. Recently, K. G. Birch[1] has provided an excellent review of these efforts. Despite extensive work, however, the problem remains that the OTF, which is generally recognized as one of the most inclusive measures of system performance, is difficult to use in practical situations and is not easily related to conventional image attributes. The following work addresses itself to this problem with the aim of providing a generally applicable OTF-derived merit function which is correlated with subjective quality judgments. This tool provides the system and design engineer with a practical method of predicting the subjective quality of an image with respect to a reference standard.

A quality reference scale commonly in use in the Kodak Apparatus Division is the Image Sharpness Scale (ISS) developed by H. F. Rogers.[2] This scale is a standard boy and dog scene with several auxiliary resolution patterns which are reproducible as a series of thirteen pictures degraded by successive amounts of defocus (Fig. 1). (Note: The resolution targets were deleted from the scene for the subjective quality studies reported in this paper.) The scale represents a useful range of quality that is of practical interest in most applications. This quality ranges from the best reasonable reproduction of the test scene to what is almost universally agreed upon as being totally unsatisfactory quality at the given magnification. As used, the scale serves as a standard by which the performance of a product under test may be specified. The facilities exist, for example, where a camera under test can be used to photograph an array of boy and dog scenes, and the resultant picture may be compared directly with the standard scale. Thus, a record can be made of product performance. Additionally, this image quality

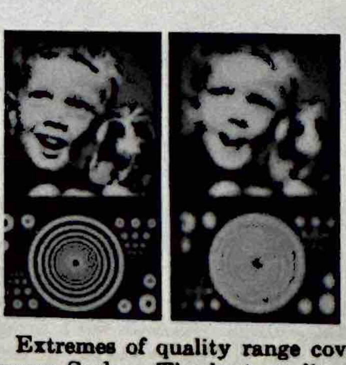

FIGURE 1. Extremes of quality range covered by the Image Sharpness Scale. The best quality print is the best print from a high quality system, which may be degraded slightly by the halftone reproduction process. Pictures are reproduced at the size used in the experiments.

Received October 15, 1971; revised January 11, 1972.

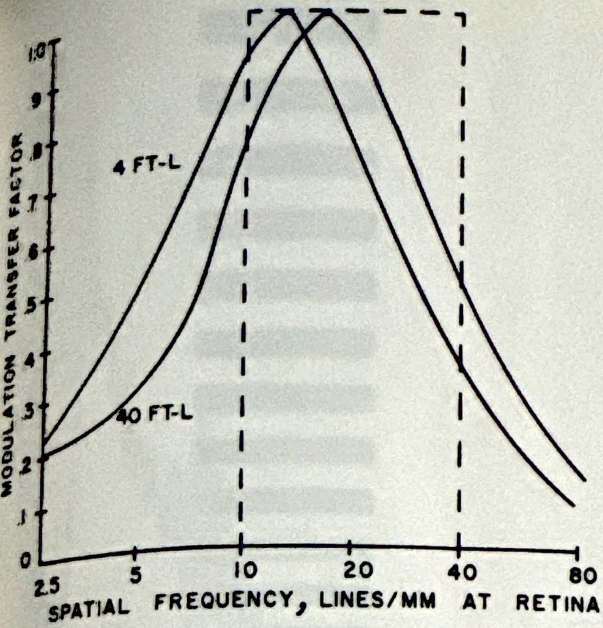

FIGURE 2. —— Eye transfer functions at 40 ft-L and 4 ft-L as determined by O. Schade (*JSMPTE*, February 1964), - - - SQF band defined in text.

scale allows various personnel to communicate information about image quality without ever visually inspecting the same print. This scale has been useful in specifying quality for all common lens defects despite the fact that it contains only defocus. However, the scale lacks an OTF relation which limits its applicability to a product test situation. Initially, our efforts were devoted to providing a merit function which would correlate subjective quality judgments of the image standard with the OTF.

At the onset, it is important to note that in all of our work we have attempted to correlate a specific OTF-related merit function with a general quality judgment by our subjective panel. We have consciously avoided the use of such terms as sharpness and resolution in our subjective testing instructions and have attempted to elicit a quality judgment from the test subjects based upon the over-all acceptability of the image. The approach, thus, has not been to see how specific image descriptors correlate with quality but rather to correlate the OTF with an undefined term "quality." Particularly in amateur photographic products, we feel that this general quality measurement is the most significant measurement that may be made as it is the direct measurement of a customer's acceptance of a hypothetical system. Thus, our instructions in the subjective testing procedures did not ask for choices or evaluations on the basis of specific image attributes but rather contained questions such as: "If you could keep only one picture, which picture would you prefer to keep?," or "Which picture was taken by the better camera?." As the boy and dog scene used in the ISS was the result of an extensive search for an aesthetically neutral scene, we have used this scene almost exclusively in our work, without any resolution targets. Test samples were usually printed at what a photographer would describe as "normal contrast" on standard Kodak papers using standard processing. Care was taken to provide and verify uniformity in processing. At the viewing distances used (30 cm), the granularity of the paper or intermediate film emulsions was not visually noticeable. It was not our intention at this point to study various emulsion or processing variables.

The range of conditions studied was limited to optical systems that are physically realizable and that might be encountered in practice. Thus, we do not consider systems with extreme phase reversal, although we have included systems subject to defocus. By making these restrictions and by confining the range of quality investigated, we will not have, of course, a universally applicable merit function. We can only agree with Birch[1] that the prospect of ever achieving such a goal is extremely dim.

We consider agreement to within single reader error (the error made by a typical subject in a single rank judgment), to be adequate from a practical standpoint and we have not attempted to reduce this error by altering weighting functions or introducing further mathematical calculations. This approach is justified, we feel, if it leads to utility in the application of a merit function.

II. SQF Definition

A. Basic Calculation

Many workers in the field of image quality definition have noticed that a perceptible difference in image quality can be obtained by changing either the scale of the point spread function or the modulation transfer function by 5 to 10%.[3] This phenomenon of just noticeable difference (JND) being related to a constant percentage change in the stimulus has been observed for many neural processes. Tasks such as judging weight and loudness of sounds follow this law known as Weber's Law. This led us to hypothesize that image quality might in some way be related to a logarithmic spatial frequency weighting of the system OTF. Specifically, image quality might correlate with the area under the system OTF on a log spatial frequency scale.

The quality of a visual image must also be related to the scale of the image on the retina. The human visual system has an MTF which peaks in the region of 10–20 cycles/mm at the retina. A typical visual system MTF[4] is shown plotted vs. log frequency plotted on Fig. 2.

Postulating an arbitrary bandpass nature for the eye, we have defined limits of integration to simulate this effect. As this represents an approximation to the true eye response, improvements, in general, can be made in the correlation between quality rank and computed rank if the "true" eye modulation transfer function is used in the calculation of quality rank. Such improvements are, however, usually less than single reader error and are obtained at the sacrifice of convenience and simplicity.

Based on these observations, a one-dimensional SQF was initially defined[5] as the integral of the system MTF (including lenses and films) between the limits of 10–40 cycles/mm when the MTF has been scaled to the retina of the observer by the magnification of the system, which includes the eye.

$$SQF = K \int_{10}^{40} \left|\tau(f)\right| d(\log f)$$
$$\text{at retina}$$

where $\tau(f)$ is the Optical Transfer Function, f is the spatial frequency, and $1/K$ is a normalizing constant obtained by performing the above integration with $\tau = 1$.

There is no reason to limit the considerations to one-dimensional MTF descriptions, as real images contain two-dimensional MTF descriptors, a two-dimensional MTF must be factored into the system. In fact, the puzzle really is why so many quality factors, using only the one-dimensional MTF, have proven quite successful in predicting quality rank for two-dimensional image structure. These successes suggest that the one-dimensional treatment includes the proper weighting function to describe the two-dimensional visual properties of the image. That is, the impression of quality is obtained by equally weighting information over all directions.

The SQF value can be obtained for a general MTF by describing the system MTF in polar coordinates and performing the following integration:

$$SQF = K \int_{10}^{40} \int_{0}^{2\pi} |\tau(f,\theta)| d(\log f) d\theta$$

where f is the spatial frequency in cycles/mm along a given azimuth θ of line structure, and K is the appropriate normalizing constant as above.

The above formula allows a simple calculation of the quality of a print which lies within the limits of the ISS quality range. Several comments are necessary about the practicality of this restriction and the role of system magnification in quality judgments. Systems which, at a specified magnification, have transfer functions which are essentially unity across the SQF bandpass will both receive equal SQF evaluations regardless of their relative performance at very high spatial frequencies. Such systems will, however, both receive equal subjective evaluations as well. Thus, it makes little sense to distinguish between such systems at this magnification. If, however, the system magnification is changed (e.g., by use of an eye loupe), differences may be seen in the image as the high frequency portion of the MTF's are moved down into the visual passband.

The same rationale is applied at the low frequency side of the SQF band. Here, however, systems are rejected on the basis of totally unacceptable image quality for most useful applications. It is characteristic of the SQF system, and of subjective evaluations, that a quality assessment is not intrinsic to the image itself but only to the image as viewed at a specified magnification. For example, a billboard image at close range (e.g., 1–2 ft) constitutes what would be called an unacceptable quality image while the same image examined at 300 ft might be described as excellent quality. Quality, as beauty, is in the eye of the beholder and the characteristics of the human visual system must serve as a reference for an objective quality factor. Systems which lie significantly outside the SQF passband at one magnification may be brought inside the passband at another magnification. It is important that the proper system magnification be specified when calculating the SQF of a system.

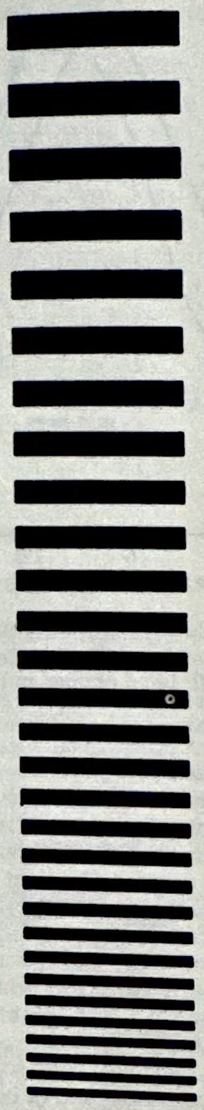

FIGURE 3. Log periodic test target.

B. Log-Periodic Target

One of our objectives was to relate a subjective scale to a convenient objective measure of image quality. This could be accomplished by measuring the MTF over the spatial frequency bandpass defined for SQF. The usual procedures require accurate mechanical operations, specialized and expensive test targets and computer reduction of the data.

To overcome these problems, a new test target was developed which has the unique property of being able to directly encode the MTF within the SQF bandpass. The target is shown in Fig. 3 and has been called the Log Periodic Target. This name was derived from the fact that the bar width and spacing width of the nth bar and space in the sequence are given by:

$$\log(W_n) = \log W_0 - nC_2$$

where W_0 and C_2 are appropriate constants. The main advantages of the test target are that it is

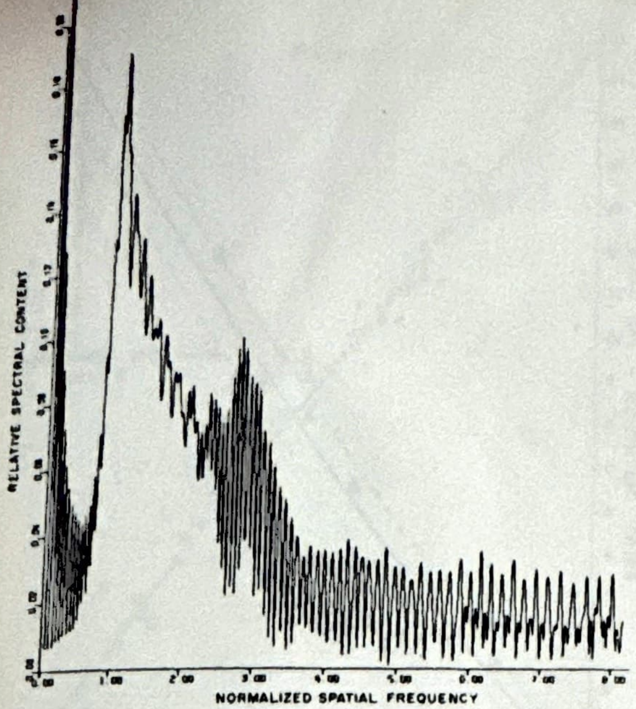

FIGURE 4. Calculated spatial frequency spectrum of log periodic target. The designed passband of the target lies between 1 and 4 on the spatial frequency scale. The frequency scale is normalized to 1 for the largest bar and space fundamental frequency.

aperiodic and that it has an easily reproduced square wave profile.

The calculated spatial frequency spectrum of the target is shown on Fig. 4. The important feature of the spectrum is that within the SQF passband the spectrum has the desired log frequency spectrum weighting. There is some power outside the passband contributed by the square edges of the test target. Under normal testing, i.e., when the lens has quality below SQF = 0.9, most of this energy is truncated by the lens MTF yielding nearly the same spectrum as if one had produced the target with a sinusoidal profile. Therefore, over the range of interest, the image of the Log Periodic Target will exhibit modulation proportional to the modulation transfer of the optical system. The principle of operation of this target is shown in Fig. 5.

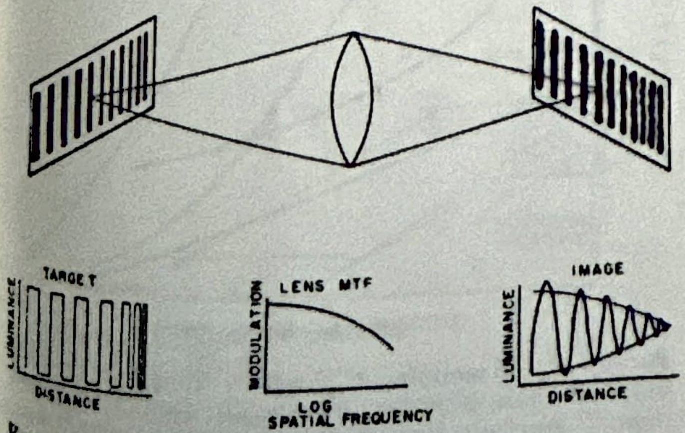

FIGURE 5. Principle of use of the log periodic target.

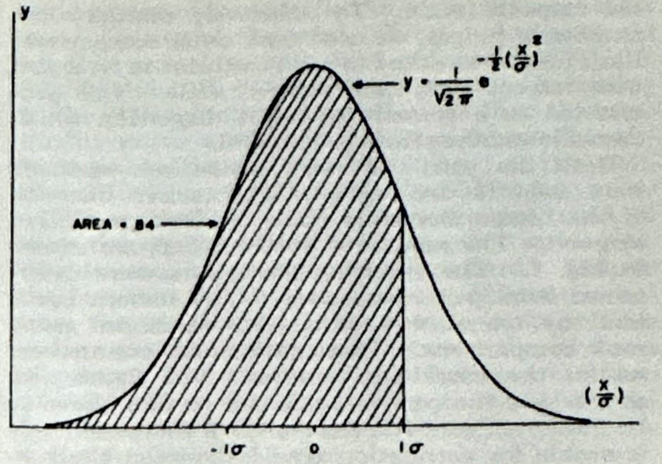

FIGURE 6. Example of a preference ratio as an area under a normal distribution.

III. Experimental Program

A. Subjective Quality Scaling

The development of a quality factor implies that the subjective response to a given set of photographs can be placed on a scale of image quality. That is, if we are to develop a mathematical correlate of image quality, we must be able to measure the subject's response to a given stimulus. In this study, we have attempted to correlate the loose term image quality (the judge's reaction) to a physical parameter, the weighted MTF of the photographic system.

The correlation study requires that we establish a linear subjective scale and a method of scaling new stimuli to the "standard" subjective scale. The method chosen for our experiments was that of paired comparison.[6] This method uses preference ratios to scale the stimuli. It is assumed that the panel of judges exhibit a normal error distribution in determining image quality. The subjective distance between two stimuli can be computed from the preference ratios. A preference ratio may be represented as the cumulative area under the normal error distribution from minus infinity to a finite value of a standardized normal distribution. The magnitude and sign of the upper integration limit is the response difference between the two stimuli. Figure 6 illustrates that a scaling distance of 1 corresponds to a preference ratio of 0.84.

The calibration of a subjective scale or the determination of the scale value of an unknown print would require that the print be compared with all the other prints on the scale. As a result of the pair comparisons judgment, estimates of the preference ratios and subjective distances can be computed for each pair. The scale value of a print or the set of prints is determined by using a least square estimation technique developed by Morrissey.[7]

One problem common to the pair comparison method is that in the case where subjective distances are large there will be no disagreement among the judges unless prohibitively large panels are used. In this case, the distance between prints cannot be scaled and other methods must be used or more prints must be added to fill in the gap in the response scale. To effectively multiply the number of judges, we used rank order comparison. Each judge was asked to rank the print to be scaled with respect to a rank ordered scale. This process led to a greater judgment dispersion which then allowed the prints to be scaled.

Both the pair and rank comparison methods were used to determine the subjective linearity of the Image Sharpness Scale (without resolution targets). The results of both scalings are shown in Fig. 7. The triangle data points were determined from pair comparison. The smooth curve and the tolerance bounds were developed using rank comparisons. These gave consistent answers within the confidence bounds. The figure also shows that the current scale can be considered to be linear. Therefore, the standard sharpness scale is usable for correlation studies between the subjective quality factor and subjective print rank. The correspondence between two scales (SQF and ISS) is defined by Fig. 7.

B. Experimental Test of SQF[8]

In order to map the region over which the SQF concept is valid, a number of experiments were conducted. These consisted first of all of providing black and white prints of the ISS scene which had been subjected to a variety of different transfer functions. In addition to simple defocus, transfer functions were used of the types shown in Fig. 8. The curves used represented the effects of a number of different films and lens effects and provide a particularly demanding test of any quality factor because of the radically different curve shapes represented by this group of functions. Additionally, results were derived from a laboratory breadboard of a test fixture which uses the log periodic test target.

Prints with these transfer functions, scaled in spatial frequency, were presented to our test panel for pair comparisons experiments. In most cases, testing was conducted with Eastman Kodak personnel at testing booths placed in high traffic corridors at several different plants in Rochester. In every case, the subject was asked for a choice only once on any given day, and each pair was presented to approximately 30 judges.

Each section that follows describes a different experiment which was conducted. The results of all the sections are summarized on Fig. 9. This plot compares the SQF value predicted on the basis of the known transfer function with the measured SQF value obtained from subjective judgments. Taking all data on Fig. 9, we show a correlation coefficient of 0.988.

Defocus experiment. The Image Sharpness Standard was produced by photographing a "boy and dog" test chart using various amounts of defocus. Since the steps of the ISS have been shown to be linear subjective steps, this gave us a well calibrated data set which we used in initial tests of the Subjective Quality Factor.

A. LaVigueur of the Kodak Apparatus Division has supplied us with measurements of the transfer functions characteristic of the 13 steps of the ISS. To obtain these measurements, sine wave targets were photographed under the conditions used to generate a set of ISS standards. Prints from these negatives were then scanned with a reflection microdensitometer to obtain the transfer function. The SQF evaluations of these prints were made by direct calculations from the resulting transfer functions. The results of these calculations are shown on Fig. 9 as the points labeled *O*.

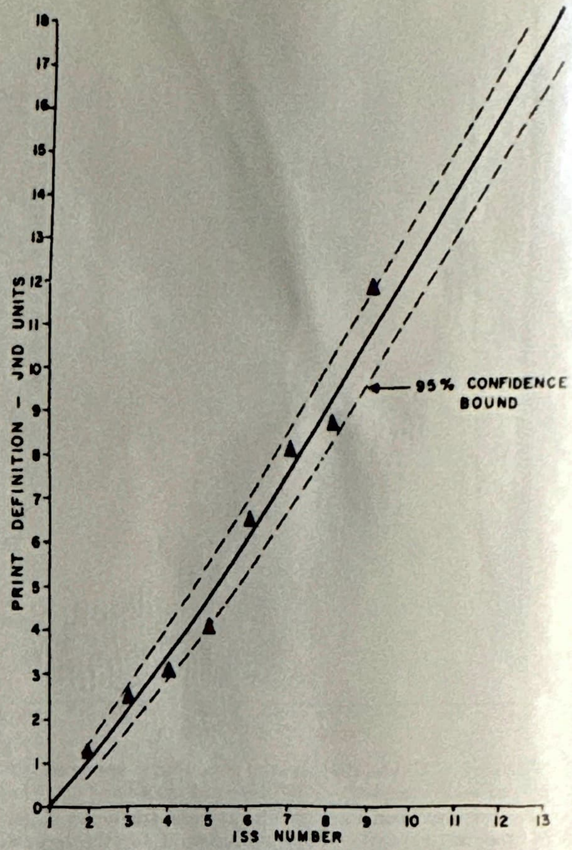

FIGURE 7. Scaling between ISS and JND scale. Confidence bounds were obtained from rank comparisons. Individual points were determined from pair comparions.

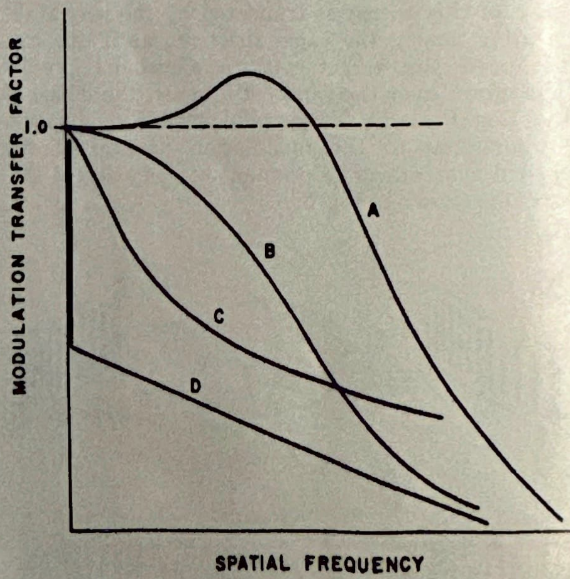

FIGURE 8. Examples of transfer functions resulting from: *A*, film adjacency effects; *B*, Gaussian degradation; *C*, centrally obstructed lens aperture; and *D*, image contrast degradation.

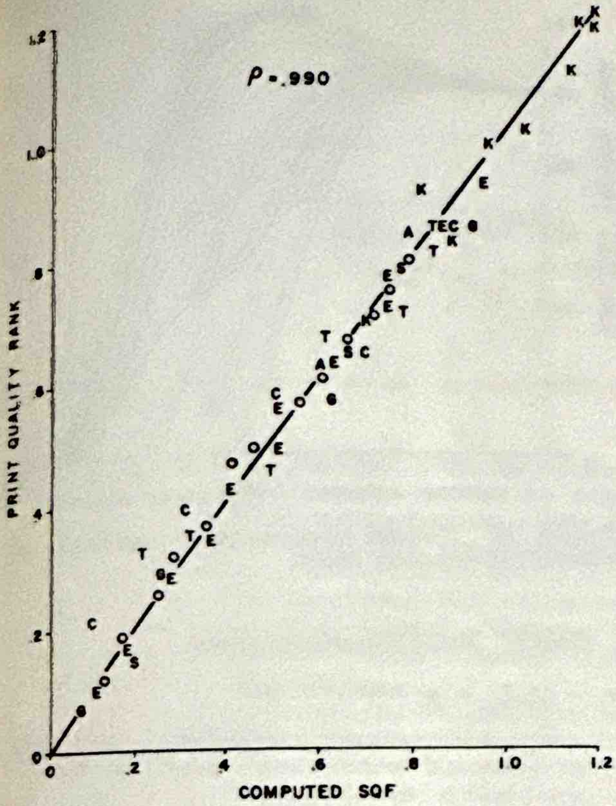

FIGURE 9. Summary of correlation between subjectively measured print quality rank and computed SQF rank. Discussion of each of the following cases is given in text. *O*—simple defocus, *C*—image contrast loss, *G*—Gaussian degradation, *S*—simulated central obstruction of lens aperture, *K,T*—film adjacency effects, *A*—astigmatism, *E*—electronic log periodic test.

Contrast loss experiments. Many situations arise where an image maintains detail but loses quality due to a loss in scene contrast. Typical of this case is photography through atmospheric haze. The effect of such a bias exposure is independent of spatial frequency and acts to uniformly decrease the modulation in the image, which may be described by a transfer function of type *D* on Fig. 8. In this example, an MTF of a diffraction limited circular aperture optical system is assumed to suffer from a bias light degradation.

The loss of contrast illustrated by the figure can be simulated by exposing the negative of the boy and dog target, then removing the negative and re-exposing the print to a bias light source. The total exposure is controlled to maintain a constant mean density of the print. Prints were produced by this procedure with bias exposures of 90, 70, 50, 30, and 10% of the total exposure. According to the SQF hypothesis, the effect on the subjective quality rating would be a reduction in quality proportional to the per cent bias light; i.e., the addition of 10% bias would reduce the SQF to 0.9.

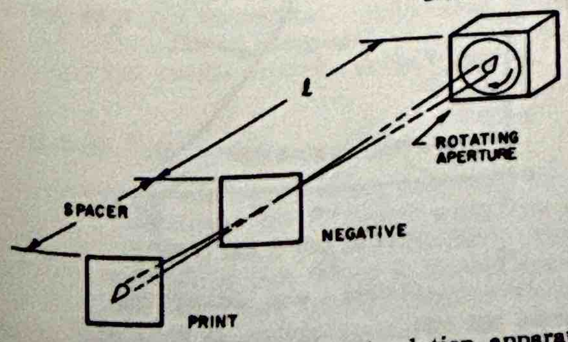

FIGURE 10. Diagram of image simulation apparatus.

The method of pair comparisons was used to rank the quality of the contrast print series relative to the image sharpness standard. Despite the fact that the degradations are different in nature, the test subjects were able to make reliable and consistent comparisons between the two sets of prints. The results of this testing are shown as points *C* on Fig. 9. Each pair was presented to approximately 30 judges and was scaled by the technique described earlier.

Analog simulation experiments. To provide a number of photographs subjected to different transfer functions, an image simulator was constructed similar to ones in use in other image quality studies[9] and shown schematically in Fig. 10. A test negative is printed onto paper in a modified printing frame containing a space of known thickness. By varying the spacer thickness or the distance to the source, the scale of the point spread function describing the printing process may be changed. By altering the shape of the rotating aperture, the shape of the point spread function may be changed. The technique has been calibrated by an edge trace technique to estimate the effects of the Callier Q-factor and other complications.

In the simulations described below, a 4 × 5 in. negative array of boy and dog photographs on a fine grain emulsion was printed onto a standard Kodabromide paper. Knowing the transfer functions of the photographic steps and the original negative, the MTF of the simulation may be obtained from the physical dimension of the apparatus. Two sets of transfer functions were used, a Gaussian scaled in width (the spatial frequency at which the MTF drops to 0.6) between 0.4 cycle/mm and 4 cycles/mm, and a skirty transfer function scaled over the same region of quality. The skirty function fell more rapidly at low frequency than a Gaussian but contained a much longer high frequency response. Pictures produced by such functions are visually characterized as being low in contrast but high in resolution when compared to a Gaussian. Figure 11 presents several transfer functions of each type.

Prints produced by this simulation technique were submitted for subjective evaluation against the sharpness scale. Representative results from this testing appear as points *G* (Gaussian) and *S* (Skirty) on Fig. 9.

Film adjacency effects. A common phenomenon in film products is a nonlinear processing of the optical image by chemical action at an exposure edge. Therefore, a steep gradient will be amplified to produce density differences greater than predicted by large area *D*-log *E* character of the film. This chemical action results in an amplification in small detail relative to large detail. This effect can be described to a good approximation by an MTF curve of the shape shown as curve *A* in Fig. 8.

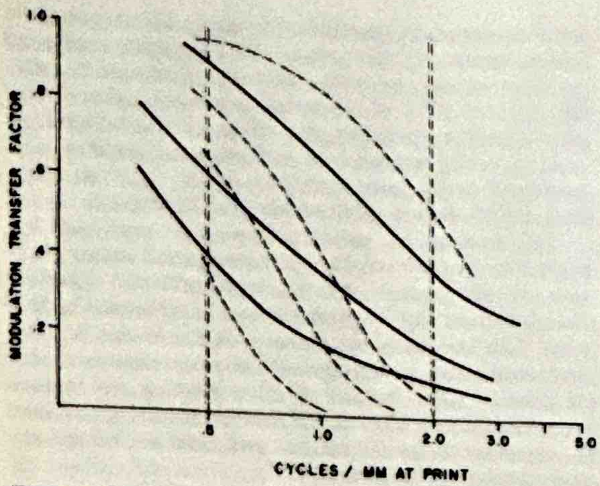

FIGURE 11. Comparison of - - - Gaussian and —— skirty transfer functions, the skirty functions simulating the effects of a centrally obstructed lens aperture.

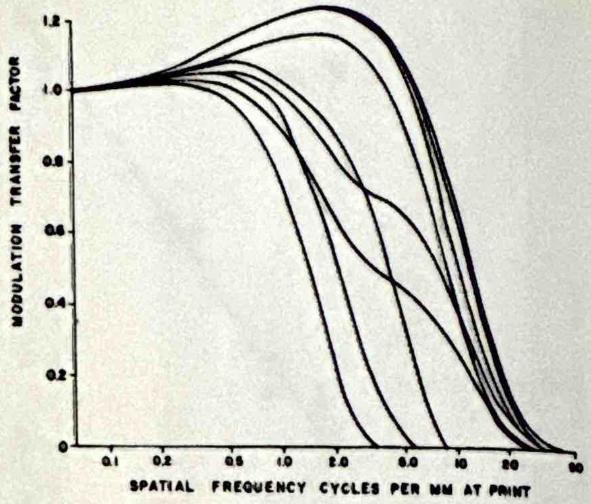

FIGURE 12. Transfer functions of transparencies exhibiting film adjacency effects.

This curve would be measured at the mean density typical of normal exposure. Since most scene detail can be represented as a small signal relative to the mean exposure level, the MTF can be used to approximate the nonlinear process. The film MTF can be cascaded with a lens MTF to yield a system transfer function.

A series of transparencies which have system transfer functions shown on Fig. 12 have been made by the Image Science Laboratory of the Kodak Research Laboratories.[10] Since the purpose of the experiment was to observe the effects of the chemical development on small areas, the large area tone reproduction was maintained constant for all prints.

These transparencies were shown to a panel of judges at a variety of viewing distances. Each transparency was judged for sharpness, and observed scale values were computed by pooling the judgments and using least square methods to estimate the average subjective scale value. The results of this scaling are shown as points K and T of Fig. 9. The points labeled K were observed at a normal viewing distance; the points labeled T were observed at a 2-in. viewing distance (monocularly with the aid of a 5 × magnifying glass).

Asymmetric transfer function experiments. So far only symmetric transfer functions have been considered. The symmetric forms are easily produced but do not test the angular integration of the SQF definition. To test the angular averaging hypothesis of this equation, a series of prints were produced by a digital image simulation technique.[11] The system used incorporates a specially modified K. S. Paul Scanner for input and output, with intermediate digital processing to subject the scanned image to an arbitrary transfer function. Accurate control may be maintained of both the tone reproduction and the OTF description, and a wide range of CTF's may be used which are not easily developed by analog methods.

In one specific experiment, two sets of prints were made to simulate an image error which is similar to astigmatism or smear. The prints created had the following OTF descriptions:

Case 1: Small Astigmatic Error

$$\tau_1 = (f_x, f_y) = e^{-(0.045/x^2+0.173/y^2)} \times \frac{\sin(0.4\,\pi f_x)}{0.4\pi f_x}\,\frac{\sin(0.4\pi f_y)}{0.4\pi f_y}$$

Case 2: Large Astigmatic Error

$$\tau_2 = (f_x, f_y) = e^{-(0.045/x^2+0.75/y^2)} \times \frac{\sin(0.4\pi f_x)}{0.4\pi f_x}\,\frac{\sin(0.4\pi f_y)}{0.4\pi f_y}$$

where f_x and f_y are spatial frequencies in cycles/mm measured at the print. These transfer functions are shown in Fig. 13. The sine x/x factors take

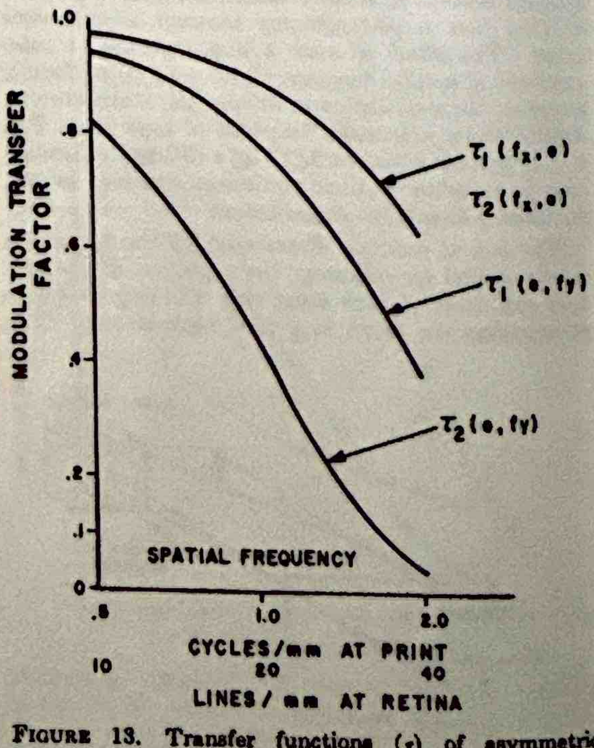

FIGURE 13. Transfer functions (τ) of asymmetric prints as a function of f_x or f_y.

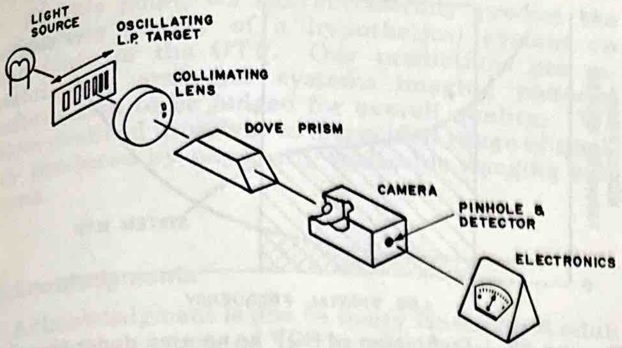

FIGURE 14. Schematic diagram of electronic log periodic test system.

into account the sampling effect of the digital simulator. Each transfer function has as its transform an elliptical point spread function where the ratio of the major to minor axis is approximately 2:1 for the better print and 6:1 in the poorer print. The computed azimuthal average SQF values were 0.80 for Case 1 and 0.62 for Case 2. These values predict Image Sharpness Standard ratings of 2.7 for Case 1 and 5.3 for Case 2.

The prints were displayed at 10 scan lines/mm producing a print cutoff frequency of 5 cycles/mm on the print. Previous work with scanned images has shown us that images scanned at this frequency do not suffer significant degradation when viewed at 30 cm despite the fact that the resolution limit of the eye would traditionally be placed at approximately 10 cycles/mm under these viewing conditions.

The prints were each ranked by seven judges and received average rankings vs. the ISS of 2.4 for Case 1 and 5.3 for Case 2. The single reader error of the ISS is approximately ±1 unit leading to an error of ±0.37 in the above rankings. These results are shown as points *A* on Fig. 9.

A good approximation to the angular average is to perform a simple average of the SQF computed in the directions of best and worst response. One might expect agreement between the two techniques since most OTF's are smooth, continuous functions as a function of angle. In Case 1, the SQF values in best and worst direction were calculated as 0.85 for f_x and 0.73 for f_y leading to an average ISS prediction of 2.9. For Case 2, the two SQF values are 0.85 and 0.43 yielding a prediction ISS number of 5.0. These results are summarized in Table I.

Simulations were also made with the same transfer functions rotated by 90° with nearly identical results, indicating that the orientation of the astigmatism was not important. Additional evidence for this treatment of astigmatism comes from the electronic test results reported in Section C.

C. Electronic Test

One of the goals in this work has been to provide an OTF-related merit factor which may be easily measured in an objective manner. Such a feature is essential if this factor is to find application in a production or quality control environment. The general principles of the log periodic test target have been discussed in Section II.B.

Table I. Astigmatic Results

Case	SQF prediction of ISS no.	Best-worst SQF prediction	Subjective test of ISS
1	2.7	2.9	2.4 ± 0.4
2	5.3	5.0	5.3 ± 0.4

The practical result of this target is the log periodic camera test device under study in our laboratory. Figure 14 is a schematic diagram of this device. Along the axis of this device lie: an oscillating log periodic target, a collimating lens which images this target to infinity, a dove prism which allows the image of the target to be rotated, and the camera to be tested. In the film plane of the camera is mounted a small pinhole and phototransistor. The imaged oscillating log periodic target sweeps back and forth across the pinhole and provides an electronic signal proportional to the luminance profile of the image. The SQF value is obtained from this signal by an electronic system which measures the peak height of the signal from each bright bar in the log periodic target. These peak values are then summed in a staircase generator and the resulting sum, properly normalized, is the measured SQF value of the image formed by the camera lens. This technique provides a good approximation of the integration required by the SQF definition. To evaluate astigmatism, the dove prism is rotated to provide the SQF value for two perpendicular orientations of the log periodic target image. The simple average of these two numbers has been found to be entirely adequate for a practical measurement.

Off-axis points may by included in the measurements by systems of beam splitters and mirrors which will simultaneously present off-axis detectors with images of the oscillating log periodic reticle. Such a system, which has been constructed in our laboratory, measures response of a specific lens system on-axis and at four additional positions in the field.

W. T. Deever of our laboratory has conducted a study to establish an experimental correlation between the SQF value predicted from the log periodic camera test equipment, and the SQF value derived from a visual evaluation and the SQF-ISS correlation described earlier. A number of cameras were measured on the log periodic camera test apparatus. These cameras included good cameras, cameras artificially degraded by defocus, and cameras rejected for manufacturing defects. The same cameras were also used to photograph the standard boy and dog scene and the resulting prints were evaluated vs. the ISS. A typical correlation for this type of experiment is shown in Fig. 15. In this plot, a total of 118 data points are shown of which approximately 94 are off-axis points. The off-axis points suffered from astigmatism in varying amounts and, as indicated earlier, the averaged best and worst direction responses were taken as the log periodic camera test evaluation in this case. Each point in the plot represents an individual reader judgment of each print and is subject to the variability characteristic of the visual evaluation. The effects of printing the photographs were found to be small

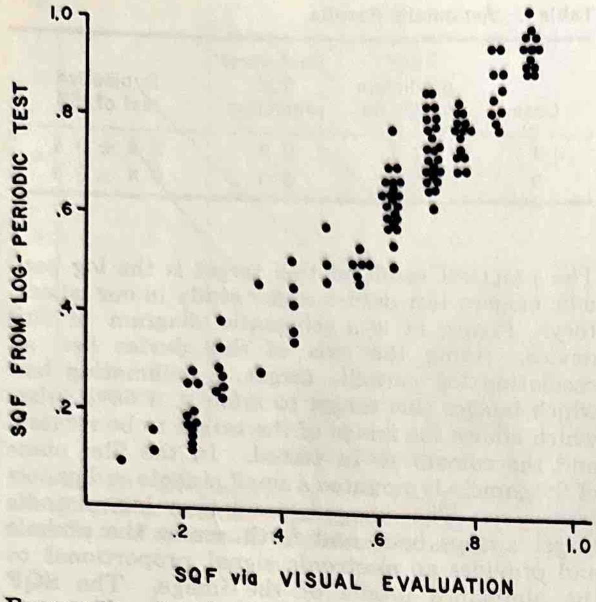

FIGURE 15. Correlation of electronic log periodic test results with visual evaluations.

(less than one ISS unit) and were factored into the system calibrations as a uniform degradation. The results given in Fig. 15 exhibit a 0.97 correlation coefficient with a straight line. The residual error may be explained on the basis of reader error, calculated on a 5% MTF change required to produce a noticeable difference in the photographic image.

Results of the visual evaluation are quantized by the steps of the ISS. Averaging the points at each level of the ISS yields the points which are shown as E on Fig. 9. Performing this average increases the correlation coefficient to 0.996.

IV. Conclusion

The work to date has included a large number of transfer functions and experimental parameters. Several important areas, however, yet remain to be studied. In particular, two photographic effects, granularity and contrast (gamma), remain to be entered into the SQF system. It is at this time our hypothesis, based on preliminary testing,

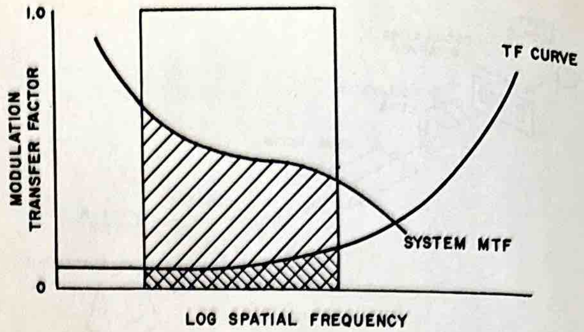

FIGURE 16. Definition of SQF as an area under transfer function and TF curves.

that a threshold factor curve (TF) may be defined similar in concept to Threshold Modulation,[12] but determined subjectively by the quality loss attributable to granularity. One would expect the curve to be similar in shape to the curve shown in Fig. 16 and the SQF value to be the normalized area between the MTF and the TF curve.

Extensive work with a single scene, a restriction necessary for a number of experimental reasons, leaves open to question whether the results derived are generally applicable to other scenes. Certainly, unusual cases may be constructed which will cause any quality factor to fail. The band nature of SQF makes it fail with scenes containing no detail whatsoever (e.g., an image of a blank wall) or unresolved detail. Both of these cases consist of scenes which have spatial frequency spectrums lying totally outside the SQF bandpass. These cases are, however, usually indeterminate in subjective tests as well. We have conducted limited testing with different scenes and have observed an effect also observed by Biederman,[13] a slight scene dependent shift for portraits. A comparison of the subjective ratings of the portrait images with those of three other scenes, for a variety of modulation transfer functions, shows a 1 JND shift for the portrait scene and no significant differences among the results obtained with the other three scenes. Figure 17 shows reproductions of the four scenes that were used.

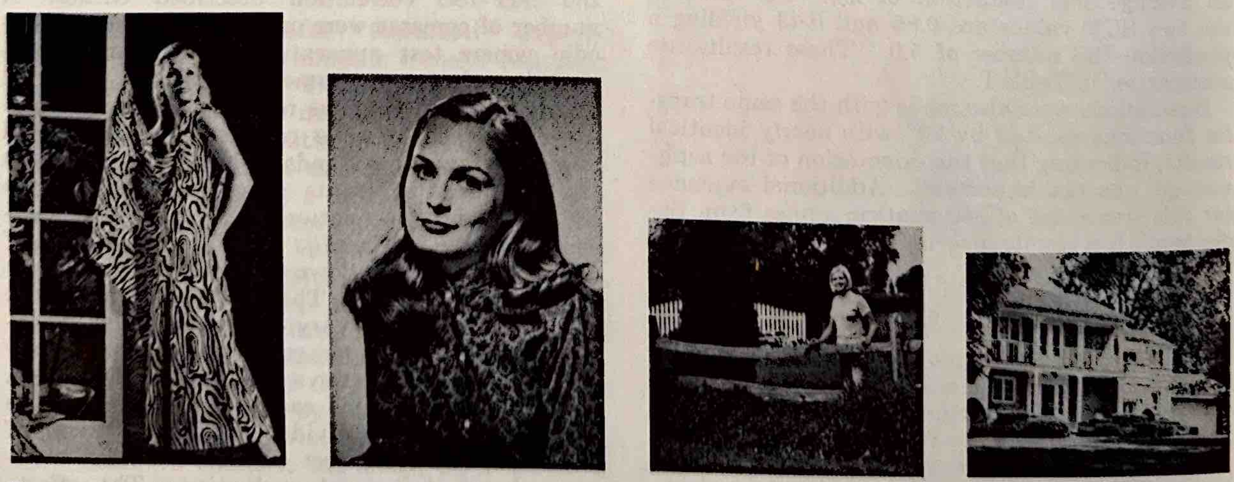

FIGURE 17. Four photographs used in scene dependence study of subjective judgments. The prints used were approximately 10 × 13 cm and were viewed at 30 cm.

At this point, we can successfully predict the subjective quality of a hypothetical system on the basis of the OTF. Our predictions are restricted to grain-free systems imaging pictorial information to be judged for overall quality. We have confined uroselves to a specified range of quality produced by physically realizable imaging systems.

Acknowledgments

Acknowledgment is due to many Eastman Kodak personnel who assisted in the work, not the least of which were the many individuals who took part in our subjective testing procedure.

Special acknowledgment is due W. T. Deever who developed a new log periodic camera test system based upon earlier work by F. W. Behnke and Dr. D. A. Warner and to R. L. Shanebrook who supervised the generation of the photographic materials used in this work.

References

1. K. G. Birch, "A Survey of OTF-Based Criteria Used in the Specification of Image Quality," Natl. Phys. Lab., Teddington, England, 1969, Clearinghouse for Fed. Sci. Tech. Info., Springfield, Va., No. N70 21982.
2. H. F. Rogers, Proceedings S.P.I.E. Photogr. Opt. Sys. Eval. Seminar, May 11–12, 1967.
3. F. Scott, *Photogr. Sci. Eng.*, **12**: 154(1968).
4. O. H. Schade, *JSMPTE*, **73**: 81(1964).
5. E. M. Granger, *J. Opt. Soc. Amer.*, **59**: 1538A(1969), (Abstract Th18).
6. H. A. David, *Method of Pair Comparisons*, C. Griffin and Co., London, 1963.
7. J. H. Morrissey, *J. Opt. Soc. Amer.*, **45**: 373(1955).
8. Previously reported results: K. N. Cupery and D. A. Warner, **59**: 1539A(1969), (Abstract Th19).
9. F. Scott, *J. Photogr. Sci.*, **12**: 139(1964).
10. G. C. Higgins, *Photogr. Sci. Eng.*, **15**: 106(1971).
11. E. M. Granger and R. S. Mickelson, *J. Opt. Soc. Amer.* **60**: 710A(1970), (Abstract TuB18).
12. T. J. Lauroesch, et al., *Appl. Opt.*, **9**: 875(1970).
13. K. Biederman, *Photogr. Korresp.*, **103**: 25(1967); **103**: 41(1967).