United States Department of Commerce
Technology Administration
National Institute of Standards and Technology

NIST Technical Note 1370

Software for Performing Gray-Scale Measurements of Optical Fiber End Faces

C.M. Wang
Dominic F. Vecchia
Matt Young
Nathan A. Brilliant

DISKETTE CARE AND HANDLING INFORMATION

NIST Technical Note 1370

Software for Performing Gray-Scale Measurements of Optical Fiber End Faces

C.M. Wang
Dominic F. Vecchia
Matt Young
Nathan A. Brilliant

Statistical Engineering Division
Computing and Applied Mathematics Laboratory
National Institute of Standards and Technology
325 Broadway
Boulder, Colorado 80303-3328

November 1994

U.S. DEPARTMENT OF COMMERCE, Ronald H. Brown, Secretary
TECHNOLOGY ADMINISTRATION, Mary L. Good, Under Secretary for Technology
NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY, Arati Prabhakar, Director

Software for Performing Gray–Scale Measurements of Optical Fiber End Faces

C. M. Wang and Dominic F. Vecchia
Statistical Engineering Division

Matt Young and Nathan A. Brilliant
Optoelectronics Division

National Institute of Standards and Technology
Boulder, Colorado 80303

The video microscope, or gray–scale, method is the most frequently used technique on the manufacturing floor for measuring critical dimensions of an optical fiber end face. Gray–scale images of optical fibers and hence their corresponding edge tables, or locus of points that represent the edge of the fiber in digitized form, can easily be contaminated by dirt or distorted by faulty cleaves. Analysis of such edge tables can be difficult. We present a method for performing optical fiber dimensional quality control which allows for end face damage and accounts for the special structure of measurement errors in fiber edge tables. The new approach adheres to the industrial standard test procedure by fitting an ellipse to the edge table to obtain geometric measurements. But, to create high breakdown resistance to outliers, a data filter based on the least–median–of–squares criterion is used. Some computational issues and a brief description of a computer program that takes the digitized image, locates and filters the edge points, and estimates the geometric parameters of interest are given. Its operation is also described.

Key words: edge points; gray–scale analysis; least–median–of–squares regression; outliers; orthogonal–distance regression

1. Introduction

The optical fiber industry, with a U.S. market that amounted to $2.1 billion in 1992 (Department of Commerce statistics), is very measurement intensive. About 20 percent of the cost of manufacturing fibers is in process and quality measurements. As more buildings, homes, and offices are wired with optical fibers, the need to accurately measure fiber geometry becomes even greater. This is because fiber networks require many connections between fibers; poor connections could result in degraded light signals if the geometry of the two connecting fibers differs. For example, a typical single–mode fiber is nearly circular with a glass core of about 10 μm in diameter surrounded by a glass cladding whose outer diameter is about 125 μm. If the cladding diameters differ, the inner cores probably would fail to align precisely when joined. For an offset of 1 μm of the cores, the loss is about 0.2 dB, or about 5 percent in transmitted power.

To minimize these coupling losses, the industry needs to measure the cladding diameter with an accuracy of 0.1 μm. To support this effort, the National Institute of Standards and Technology (NIST) has developed instruments that can make diameter measurements accurate to 0.04 μm [1]. These instruments are expensive and delicate, and are used primarily to generate Standard Reference Material (SRM) fibers by which the industry can calibrate their own measuring systems and hence are not suitable for routine industrial quality control.

The gray–scale method, which uses a video microscope in conjunction with a frame digitizer, is the most frequently used technique on the manufacturing floor since geometric parameters of interest, such as the cladding diameter and circularity, are easily and quickly measured by this method. Very briefly, the fiber is illuminated with light, and the microscope is focused onto a cleaved end face; that is, the axis of the microscope is parallel to that of the fiber and perpendicular to the end face. In our instrument, the image, an array of 512×440 pixels, is acquired and digitized. Each pixel receives a value between 0 and 255, representing the image of the fiber end in 256 gray levels. The left panel of Figure 1 shows the gray–scale image of a fiber end. Also, calibration is used to determine the correspondence between positions of pixels and distance, that is, the relationship between pixel units and micrometers. For a more detailed discussion of video microscopes dedicated to fiber geometry, see **Refs.** [2] and [3].

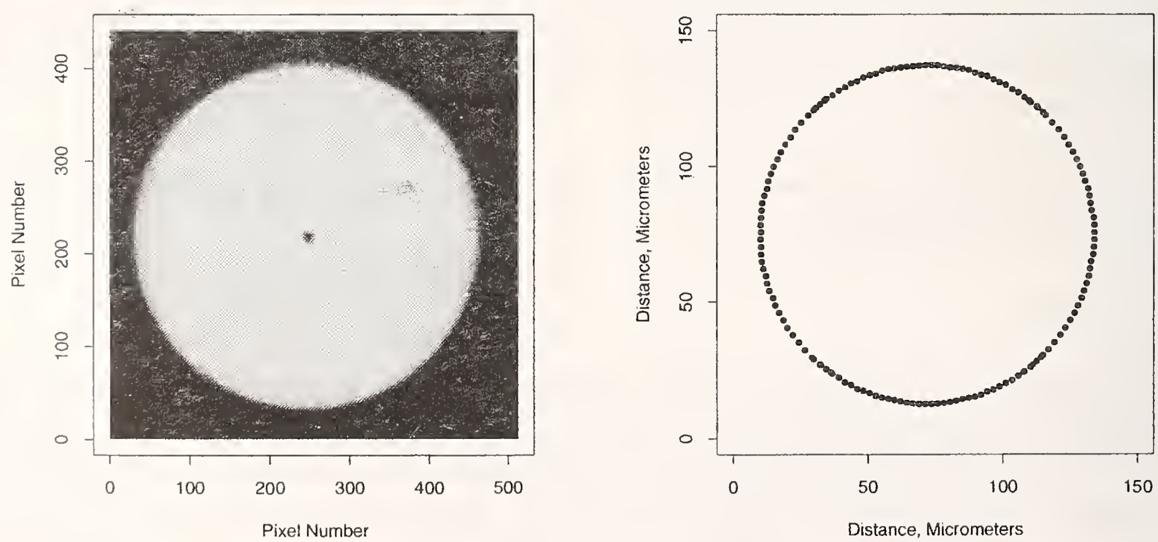

Figure 1. The left panel is the 512×440 pixel gray–scale image of a pristine optical fiber end in reverse contrast; that is, black represents the highest intensity. The right panel shows edge points (•) computed from a 64×55 image by averaging 8×8 blocks of pixels.

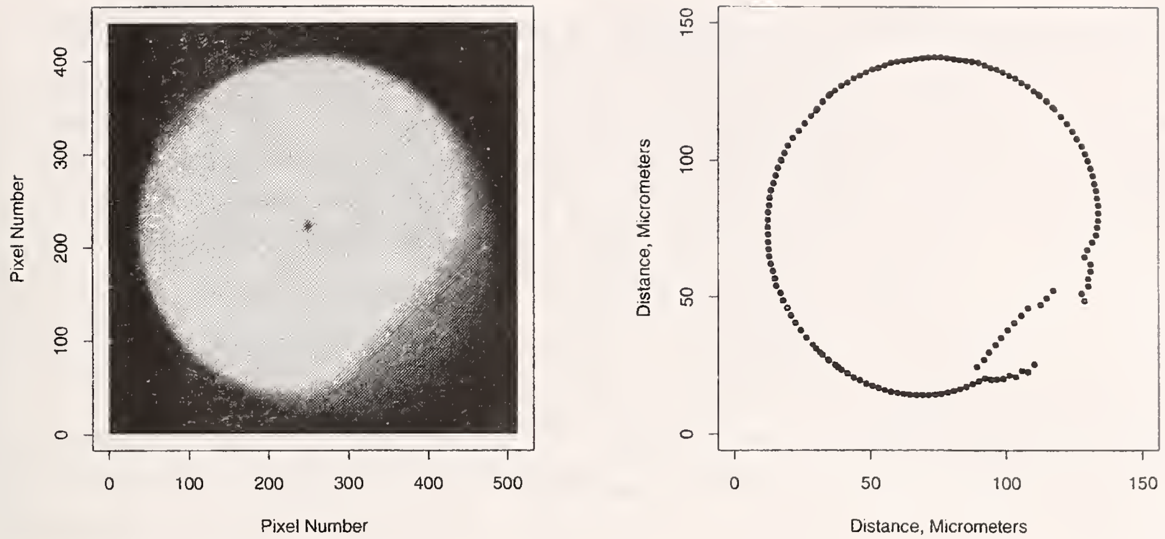

Figure 2. The left panel is the 512 × 440 *pixel gray–scale image of a damaged optical fiber end. The right panel plots edge points (•) computed from lower resolution* 64 × 55 *image.*

An edge table, or locus of points that represent the edge of the fiber in digitized form, is then obtained. The right panel of Figure 1 shows the cladding edge points for the fiber image in the left panel. The edge points were computed by estimating the inflection points of the intensity distribution across respective horizontal or vertical cross sections of the image. (This method gives a small systematic error in the calculated value of some fiber dimensions, but that is not our concern here, see Ref. [2].) Finally, an ellipse is fitted to the edge table to obtain dimensional measurements useful in quality assurance, such as cladding diameter and noncircularity, and cladding–core decentering (or concentricity error). The main reason for modeling the edge points by an ellipse, rather than a circle, is that fiber cladding noncircularity is one of the primary geometric measurements of interest, and most fibers are approximately elliptical in cross section.

Gray–scale images of optical fibers can easily be contaminated by dirt or distorted by faulty cleaves. The left panel of Figure 2 shows the gray–scale image of a fiber end that was prepared by deliberately striking it against a microscope cover slip to simulate damage routinely encountered in an industrial setting. The damage appears in the lower right corner; the upper left corner may show damage due to cleaving the fiber. The computed edge points for the damaged fiber are shown on a micrometer scale in the right panel.

Analysis of images like that in Figure 2 can be difficult unless the influence of outliers in the edge table can be eliminated from the computations. In this context, the term "outliers" refers to points that do not truly lie on the circumference of the *undamaged* fiber, but are artifacts of cleaving or imperfection in the end of the fiber and can be safely ignored.

This outlier problem prompted the Telecommunications Industry Association (TIA) to propose *Fiber Optics Test Procedure*–176 [4]. This FOTP is part of the series of *recommended standard* test procedures proposed by TIA. FOTP–176 documents guidelines for measuring optical fiber cross–sectional geometry by gray–scale analysis. The basic algorithm proposed in FOTP–176 consists of three steps: edge table filtering, edge table fitting, and parameter estimation. The algorithm calls for removing outliers before calculating the best–fit ellipse. There are, accordingly, many possible fits, depending on the criterion for omitting outliers from calculation. Furthermore, the protocols recommended in FOTP–176 depend on iterative outlier rejection rules which are difficult to automate reliably.

In this Technical Note, we present a method for filtering edge tables and allowing for end face damage. This method accounts for the special structure of "measurement" errors in fiber edge tables. The approach adheres to the industrial test procedure by fitting an ellipse to the edge table to obtain dimensional measurements. But, to create high breakdown resistance to outliers, we have implemented an edge table filter based on the least–median–of–squares (LMS) criterion [5, 6]. We also discuss certain aspects of the computation and describe an integrated program that takes the digitized image, locates and filters the edge points, and estimates the geometric parameters of interest.

2. Equation and Fitting of the Ellipse

An elliptical model for fiber edge points must be parameterized to allow for an arbitrary angular orientation. The general equation of an ellipse (assumed not to pass through the origin) is given by

$$f(x,y) = Ax^2 + Bxy + Cy^2 + Dx + Ey + 1 = 0, \tag{1}$$

where $B^2 - 4AC < 0$. Figure 3 is a graphical illustration of the elliptical fitting problem for the edge table shown in Figure 1. The ellipse in Figure 3 is more conveniently described by its center (α, β), semimajor axis M, semiminor axis m, and the angle θ formed by the major axis with the x–axis. The five parameters in the geometric representation are simple functions of the parameters in eq (1):

$$\alpha = \frac{2CD - BE}{B^2 - 4AC}, \tag{2}$$

$$\beta = \frac{2AE - BD}{B^2 - 4AC}, \tag{3}$$

$$\theta = \frac{1}{2}\tan^{-1}\left(\frac{B}{A - C}\right), \tag{4}$$

$$
\begin{aligned}
M &= \sqrt{\frac{2(\alpha^2 A + \alpha\beta B + \beta^2 C - 1)}{A + C + B/\sin 2\theta}}, && (5) \\
m &= \sqrt{\frac{2(\alpha^2 A + \alpha\beta B + \beta^2 C - 1)}{A + C - B/\sin 2\theta}}, && (6)
\end{aligned}
$$

where the definition of M and m above correspond to the geometry in Figure 3. In general, the semimajor axis is defined to be the larger of the quantities in eq (5) and eq (6) and the semiminor axis the smaller value, depending on the sign of $\sin 2\theta$. We define the mean diameter of an ellipse as the sum $M+m$, while *noncircularity* is defined as $(M-m)/(M+m)$ and is expressed as a percent. Given an edge table like that in Figure 1, the mean diameter and noncircularity are estimated by substituting appropriate parameter estimates $\hat{A}$, $\hat{B}$, $\hat{C}$, $\hat{D}$, $\hat{E}$ in eq (2) to eq (6).

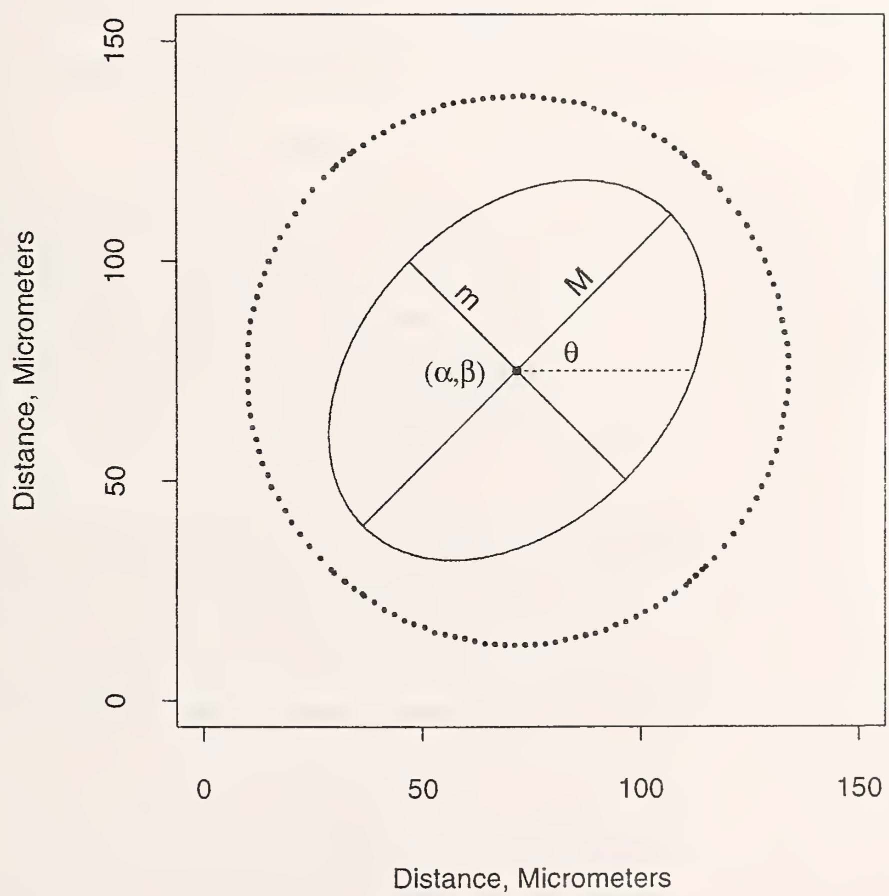

Figure 3. Optical fiber edge table with embedded ellipse to be fitted.

2.1 A "Convenient" Least–Sum–of–Squares Solution

One procedure commonly used in industrial systems (e.g., see **Ref.** [4]) is to fit eq (1) to a set of edge data $\{(X_i, Y_i), i = 1, 2, \ldots, n\}$ by computing estimates of the five elliptical parameters as the solution of the following minimization problem:

$$\min_{A,B,C,D,E} \sum_{i=1}^{n} \left(AX_i^2 + BX_iY_i + CY_i^2 + DX_i + EY_i + 1 \right)^2 . \tag{7}$$

The solution of eq (7) may be easily obtained in a one–step, closed–form equation using linear least–sum–of–squares (LSS) calculations. We will denote the computed estimates by $\tilde{A}$, $\tilde{B}$, $\tilde{C}$, $\tilde{D}$, $\tilde{E}$ and call this result the *convenient* LSS solution for the following reason.

While the solution of eq (7) is *computationally convenient*, the *pseudo–residuals* associated with eq (7) have no useful interpretation as deviations of the computed edge points from a fitted ellipse. In other words, if the terms being squared in eq (7) are interpreted, in the usual sense of regression analysis, as deviations of a dependent variable from an assumed model, then the ith residual from the convenient LSS fit would be expressed as

$$r_i = -1 - \left(\tilde{A}X_i^2 + \tilde{B}X_iY_i + \tilde{C}Y_i^2 + \tilde{D}X_i + \tilde{E}Y_i \right) . \tag{8}$$

In the terminology of regression analysis, by choosing the convenient LSS solution, we are, in effect, computing the regression equation of a dependent variable $w \equiv -1$ on the independent variables X_i^2, X_iY_i, Y_i^2, X_i, and Y_i. Forbes [7] has shown that r_i is proportional to the difference of two *areas*. One is the area of the ellipse that has center (α, β), orientation θ, and eccentricity m/M, and passes through the point (X_i, Y_i); the other is the area of the ellipse with the same center and orientation, and semimajor and semiminor axes M and m. These "residuals" are purely a mathematical convenience that have no physical significance in the usual sense of Euclidean *distance* from measured data to the fitted curve. Accordingly, we cannot expect the convenient LSS solution to have any of the "good" statistical properties of least–squares estimates in general. In fact, a simulation study, given by Vecchia, Wang, and Young [8], shows that the convenient LSS fit is not accurate for noisy data. In our application, however, it is expected to be "accurate enough" for fitting edge tables of the quality shown in Figure 1.

2.2 Orthogonal–Distance Regression

A LSS solution that properly incorporates the natural residuals in the gray–scale image application is given below. The method presented, called either errors–in–variables or orthogonal–distance regression, applies to many other model–fitting problems as well.

The orthogonal–distance LSS solution to our problem can be formulated as follows ([9], p. 238). We let $\{\mathbf{z}_i = (x_i, y_i),\ i = 1, 2, \ldots, n\}$ denote the *true*, or noise–free, edge points of the fiber. The true values are defined implicitly by eq (1); they are assumed to fall exactly on an ellipse for some values of the unknown parameters A, B, C, D, E. The observed data, however, are perturbed by measurement errors; that is, they are noisy measurements of the true edge points. If $\mathbf{Z}_i = (X_i, Y_i)$ denotes the measurement of $\mathbf{z}_i$, the observed edge points may then be described by

$$\mathbf{Z}_i = \mathbf{z}_i + \mathbf{e}_i,\ i = 1, 2, \ldots, n, \tag{9}$$

where the errors $\{\mathbf{e}_i, i = 1, 2, \ldots, n\}$ will be assumed to be independent and come from a two–dimensional distribution with mean $\mathbf{0}$ and covariance matrix $\Sigma_e = \sigma^2\mathbf{V}$, where $\mathbf{V}$ is known and nonsingular.

The measurements $\{\mathbf{Z}_i,\ i = 1, 2, \ldots, n\}$ and the equations satisfied by the true values $\{f(\mathbf{z}_i; A, B, C, D, E) = 0,\ i = 1, 2, \ldots, n\}$ (from eq (1)) jointly specify the model to be fitted to fiber images. A LSS fit of this model is obtained by minimizing the Lagrangean

$$\sum_{i=1}^{n}(\mathbf{Z}_i - \hat{\mathbf{z}}_i)\Sigma_e^{-1}(\mathbf{Z}_i - \hat{\mathbf{z}}_i)^t + \sum_{i=1}^{n}\lambda_i f(\hat{\mathbf{z}}_i; \hat{A}, \hat{B}, \hat{C}, \hat{D}, \hat{E}) \tag{10}$$

with respect to $\hat{A}$, $\hat{B}$, $\hat{C}$, $\hat{D}$, $\hat{E}$ and $\{\hat{\mathbf{z}}_i, i = 1, 2, \ldots, n\}$, where the λ_i are Lagrange multipliers.

The first sum in eq (10) is the sum of squared "statistical" distances between the actual measurements and estimated true values on the fitted ellipse, while the second sum in eq (10) constrains the estimated true values to fall on an ellipse. If errors in the measured data are small enough, the first part of eq (10) (the sum of squared residuals) becomes negligible, and the problem reduces to the convenient LSS solution in eq (7). In other words, the simple linear fit of eq (7) is strictly correct only when the data are exactly on an ellipse.

Owing to the procedure for digitizing gray–scale images, there is a special error structure that we can take advantage of when solving the optimization problem in eq (10). Specifically, the edge–finding routine obtains the edge tables by splitting the image into two sections. The first section includes the pixels from approximately −45° to 45° and 135° to 225°. In this section, the pixels are scanned *horizontally* line by line. The second section includes the pixels from 45° to 135° and 225° to 315°, and the pixels are scanned *vertically* column by column (Figure 4). Thus an edge table like that in Figure 1 can be partitioned into two groups of edge points: a collection of points $\{\mathbf{Z}_i^{(y)} = (x_i, Y_i),\ i = 1, 2, \ldots, n_y\}$ estimated by processing *columns* of pixels (fixed x–coordinate, noisy y–coordinate) and the complementary set of points $\{\mathbf{Z}_j^{(x)} = (X_j, y_j),\ j = 1, 2, \ldots, n_x\}$ obtained from *rows* of pixels (fixed y–coordinate, noisy x–coordinate). This special structure can be used to simplify eq (10). For example, if

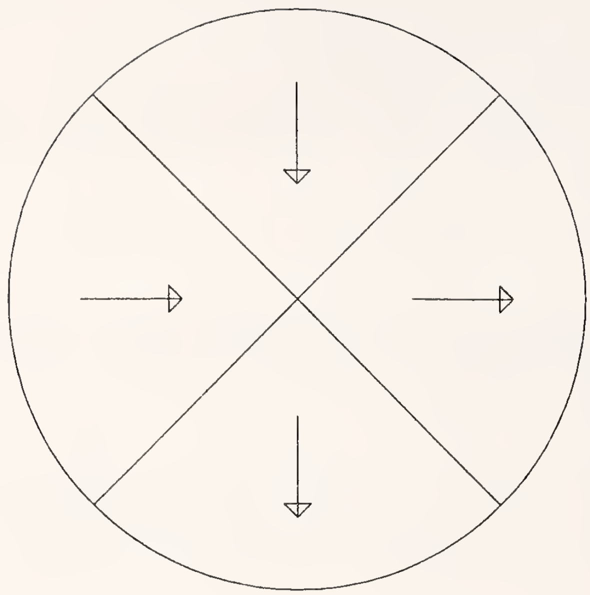

Figure 4. Directions of scanning and processing pixels.

we assume that $\Sigma_e = \sigma^2 \mathbf{I}_2$, and without loss of generality, we also assume that sequence of $n = n_x + n_y$ points in eq (9) have been arranged so that the first n_x measurements correspond to edge points with y fixed, then eq (10) reduces to

$$\sum_{i=1}^{n_x}(X_i - \hat{x}_i)^2/\sigma^2 + \sum_{i=n_x+1}^{n} (Y_i - \hat{y}_i)^2/\sigma^2 + \sum_{i=1}^{n_x} \lambda_i f(\hat{x}_i, y_i; \hat{A}, \hat{B}, \hat{C}, \hat{D}, \hat{E}) + \sum_{i=n_x+1}^{n} \lambda_i f(x_i, \hat{y}_i; \hat{A}, \hat{B}, \hat{C}, \hat{D}, \hat{E}). \quad (11)$$

The minimization is with respect to $\hat{A}$, $\hat{B}$, $\hat{C}$, $\hat{D}$, $\hat{E}$, $\{\hat{x}_i, i = 1, 2, \ldots, n_x\}$ and $\{\hat{y}_i, i = n_x + 1, \ldots, n\}$. This reduces the number of unknowns in the optimization from $5 + 2n$ in eq (10) to $5 + n$ in eq (11). We call this result the gray–scale LSS solution. Iterative methods are used to obtain the solution. Further remarks on computational issues are presented in Section 4. Moreover, in contrast to the convenient LSS approach, the gray–scale LSS solution provides approximate standard errors for the parameter estimates $\hat{A}$, $\hat{B}$, $\hat{C}$, $\hat{D}$, and $\hat{E}$. Using the relationships in eq (2) to eq (6), we can also derive the approximate standard errors for the geometric parameter estimates $\hat{\alpha}$, $\hat{\beta}$, $\hat{\theta}$, $\hat{M}$, and $\hat{m}$ by the propagation of errors technique. Specifically, if we denote the approximate variances of $\hat{A}$, $\hat{B}$, $\hat{C}$, $\hat{D}$, and $\hat{E}$, which are part of the gray–scale LSS solution, by $\hat{\sigma}_A^2$, $\hat{\sigma}_B^2$, $\hat{\sigma}_C^2$, $\hat{\sigma}_D^2$, and $\hat{\sigma}_E^2$, respectively, then the approximate variance of $\hat{\alpha}$ is given by

$$\hat{\sigma}_\alpha^2 \approx \left(\frac{\partial \alpha}{\partial A}\right)^2 \hat{\sigma}_A^2 + \left(\frac{\partial \alpha}{\partial B}\right)^2 \hat{\sigma}_B^2 + \left(\frac{\partial \alpha}{\partial C}\right)^2 \hat{\sigma}_C^2 + \left(\frac{\partial \alpha}{\partial D}\right)^2 \hat{\sigma}_D^2 + \left(\frac{\partial \alpha}{\partial E}\right)^2 \hat{\sigma}_E^2, \tag{12}$$

where the partial derivatives are evaluated at $A = \hat{A}$, $B = \hat{B}$, $C = \hat{C}$, $D = \hat{D}$, and $E = \hat{E}$. Thus

$$\begin{aligned}\hat{\sigma}_\alpha^2 \approx\ & \frac{16\hat{C}^2(2\hat{C}\hat{D} - \hat{B}\hat{E})^2}{(\hat{B}^2 - 4\hat{A}\hat{C})^4}\hat{\sigma}_A^2 + \frac{(4\hat{B}\hat{C}\hat{D} - 4\hat{A}\hat{C}\hat{E} - \hat{B}^2\hat{E})^2}{(\hat{B}^2 - 4\hat{A}\hat{C})^4}\hat{\sigma}_B^2 + \\ & \frac{4\hat{B}^2(2\hat{A}\hat{E} - \hat{B}\hat{D})^2}{(\hat{B}^2 - 4\hat{A}\hat{C})^4}\hat{\sigma}_C^2 + \frac{4\hat{C}^2}{(\hat{B}^2 - 4\hat{A}\hat{C})^2}\hat{\sigma}_D^2 + \frac{\hat{B}^2}{(\hat{B}^2 - 4\hat{A}\hat{C})^2}\hat{\sigma}_E^2.\end{aligned}$$

The approximate variances of $\hat{\beta}$, $\hat{\theta}$, $\hat{M}$, and $\hat{m}$ are similarly obtained.

Other authors have also considered the problem of fitting circles and ellipses to data, from both numerical and statistical points of view; several references, for example, are given by Mamileti, Wang, Young, and Vecchia [10].

3. Data Filter for Fiber Edge Tables

Because they are least–squares procedures, both the convenient and the gray–scale LSS will perform poorly when outliers are present. Thus the outliers must be removed before calculating the best–fit ellipse. We use *least–median–of–squares* (LMS) regression developed by Rousseeuw [5] to identify and eliminate outliers.

LMS regression is a robust method obtained by replacing the *sum* in LSS by the *median*. That is, while LSS produces parameter estimates which minimize the sum of the squares of the residuals, LMS chooses parameter estimates which minimize the median of the squares of the residuals. It can tolerate a high proportion of contaminated data without degrading the accuracy of the fitted equation. In the terminology of robust regression, the LMS solution has a *breakdown point* of 50 percent (the highest possible value), whereas the LSS method has a breakdown point of 0 percent. A single aberrant value can cause LSS to give an arbitrarily bad answer, whereas nearly half of the data could be corrupted without affecting the validity of the LMS solution. In the optical fiber application, there is no obvious way to implement the LMS criterion in the (implicit model) form of the objective function in eq (11). However, robust residuals obtained from an approximate LMS solution, based on pseudo–residuals associated with equation eq (1), are reliable indicators of bad edge points.

Specifically, the robust solution based on pseudo–residuals is

$$\min_{\tilde{A},\tilde{B},\tilde{C},\tilde{D},\tilde{E}} \text{median}\{(\tilde{A}X_i^2 + \tilde{B}X_iY_i + \tilde{C}Y_i^2 + \tilde{D}X_i + \tilde{E}Y_i + 1)^2,\ i = 1,\ldots,n\}. \tag{13}$$

The solution of eq (13) may be easily obtained using existing software for linear LMS calculations [6, 11, 12]. We call this result the convenient LMS fit for the reason given earlier in connection with the convenient LSS fit.

An edge table filter based on the convenient LMS regression consists of following steps: first obtain estimates of the regression parameters $\tilde{A}$, $\tilde{B}$, $\tilde{C}$, $\tilde{D}$, $\tilde{E}$, using the criterion in eq (13), together with a corresponding error scale estimate $\tilde{\sigma}$ ([6], p. 44). Then for each edge point (X_i, Y_i), compute the associated absolute standardized pseudo–residual

$$R_i = |\tilde{A}X_i^2 + \tilde{B}X_iY_i + \tilde{C}Y_i^2 + \tilde{D}X_i + \tilde{E}Y_i + 1|/\tilde{\sigma}, \tag{14}$$

and remove that point from the edge table if its value of R_i is larger than a threshold, such as 2.5 or 3. Figure 5 shows the edge table of the damaged fiber as well as the ellipses fitted by convenient LMS (solid curve), the analogous one–step convenient LSS (dotted curve), and the gray–scale LSS (dashed curve). The difficulty of designing an automatic outlier–rejection scheme based on standardized residuals from iterated LSS fits, as suggested by FOTP–176, is evident: some good edge points may be identified as outliers and some outliers may appear to be valid edge points. The LMS fit, by contrast, produces small (robust) residuals for virtually all good edge points and a clear indication (by large residuals) of the bad edge points. The "o" points, which have standardized residuals larger than 2.5, are outliers and are removed from the edge table.

4. Computational–Related Issues

We have developed an integrated software package for fiber geometry measurement. A brief description of the program and its operation is given in the next section. The program reads the digitized image and locates the cladding edge points for rows and columns. The edge–finding routine is flexible and can find the edge at a specified fraction of the maximum intensity or at the inflection point. We use an efficient algorithm developed by Savitzky and Golay [13] to calculate the second derivative of the intensity distribution. A detailed discussion of edge–finding and related problems can be found in the paper by Mechels and Young [2].

Having found the cladding edge table, the program locates a tentative center of the image. An area around this center is then used to scan for the core intensity contour. This step is required because the image is not necessarily located at the center of the frame.

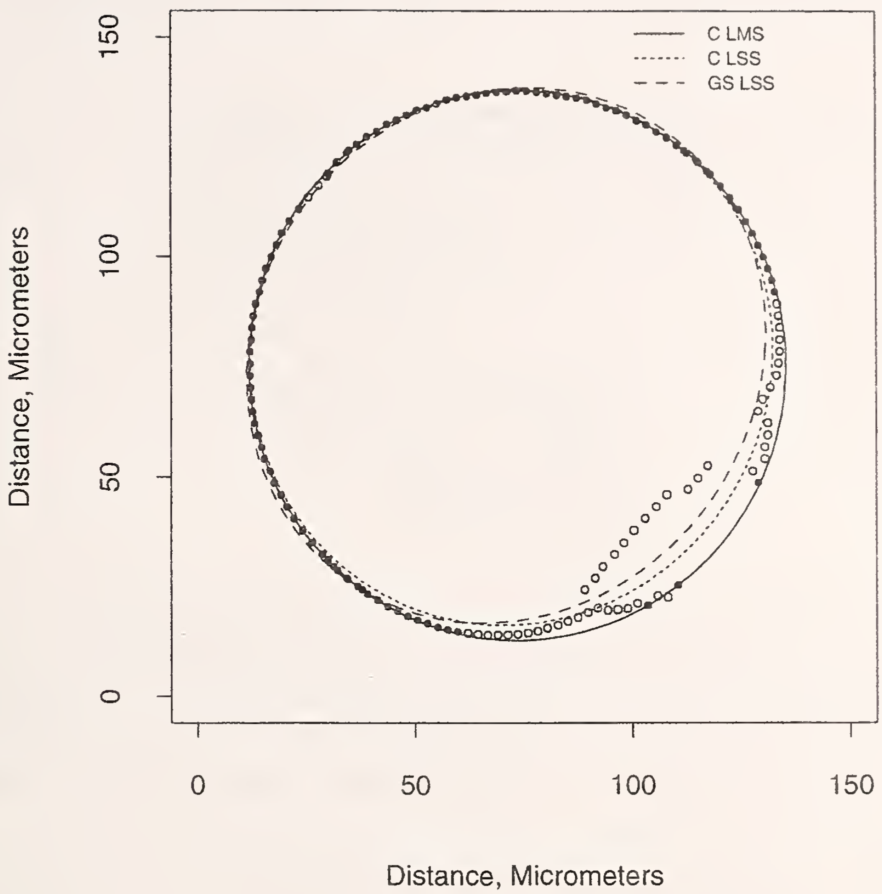

Figure 5. Ellipses fitted to the damaged–fiber edge table. The convenient LMS, the convenient LSS, and the gray–scale LSS ellipses are denoted by the solid, the dotted, and the dashed curves, respectively. The "o" points are outliers according to the LMS outlier–rejection scheme.

The program then uses the LMS fit to filter the edge tables. To fit an ellipse by LMS, we use the same procedure as the program PROGRESS described by Rousseeuw and Leroy [6]. That is, the program randomly chooses five points, fits an ellipse to these five points, and calculates the median of the squared residuals of all edge points. It performs this operation a number of times and estimates the five parameters as those corresponding to the ellipse that displays the lowest median. Since the purpose of this step is to remove the potential outliers, the number of resampling runs we used is much smaller than the 2,500 recommended by Rousseeuw and Leroy [6] for a five–parameter model. We use 250 in our program.

Next, we fit an ellipse to the remaining "good" points using the gray–scale LSS solution. The general orthogonal–distance LSS solution of eq (10) can be obtained by a Gauss–Newton type of iterative algorithm described by Fuller ([9], p. 239). A brief description of the algorithm follows. Let $\tilde{\mathbf{w}} = (\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D}, \tilde{E})'$ and $\tilde{\mathbf{z}}_i$, $i = 1, 2, \cdots, n$ be the initial estimates. Calculate

$$\hat{\mathbf{w}} = \tilde{\mathbf{w}} - \left[\sum_{i=1}^{n} \left(\frac{\partial f_i}{\partial \mathbf{w}}\right)' \left(\frac{\partial f_i}{\partial \mathbf{w}}\right) h_i^{-1}\right]^{-1} \sum_{i=1}^{n} \left(\frac{\partial f_i}{\partial \mathbf{w}}\right)' g_i\, h_i^{-1}$$

$$\hat{\mathbf{z}}_i = \mathbf{Z}_i - \left[\left(\frac{\partial f_i}{\partial \mathbf{w}}\right)(\hat{\mathbf{w}} - \tilde{\mathbf{w}}) + g_i\right] \left(\frac{\partial f_i}{\partial \mathbf{z}}\right) \Sigma_e\, h_i^{-1},$$

where

$$g_i = \left(\frac{\partial f_i}{\partial \mathbf{z}}\right)(\mathbf{Z}_i - \tilde{\mathbf{z}}_i)' + f(\tilde{\mathbf{z}}_i; \tilde{\mathbf{w}}),$$

$$h_i = \left(\frac{\partial f_i}{\partial \mathbf{z}}\right) \Sigma_e \left(\frac{\partial f_i}{\partial \mathbf{z}}\right)',$$

and use the resulting $(\hat{\mathbf{w}}, \hat{\mathbf{z}}_i)$ as initial estimates $(\tilde{\mathbf{w}}, \tilde{\mathbf{z}}_i)$ for the next iteration. The expression $\partial f_i/\partial \mathbf{w}$ is a row vector of size 5 containing the partial derivatives of $f(\mathbf{z}_i; A, B, C, D, E)$ with respect to A, B, C, D, and E evaluated at $(\tilde{\mathbf{z}}_i, \tilde{\mathbf{w}})$, and $\partial f_i/\partial \mathbf{z}$ is an 1×2 vector containing the partial derivatives of $f(\mathbf{z}_i; A, B, C, D, E)$ with respect to x and y evaluated at $(\tilde{\mathbf{z}}_i, \tilde{\mathbf{w}})$. The iterations continue until the relative change in the sum of squared residuals and $f(\tilde{\mathbf{z}}_i; \tilde{\mathbf{w}}_i)$ are within a pre-determined tolerance.

The gray–scale LSS solution is a special case of the orthogonal–distance LSS regression. Specifically, $\partial f_i/\partial \mathbf{z}$ is reduced to

$$\frac{\partial f_i}{\partial \mathbf{z}} = \left(\frac{\partial f_i}{\partial x}, \frac{\partial f_i}{\partial y}\right) = \left(\frac{\partial f_i}{\partial x}, 0\right) \quad i = 1, \cdots, n_x$$

$$= \left(0, \frac{\partial f_i}{\partial y}\right) \quad i = n_x + 1, \cdots, n,$$

and only $\hat{x}_i$, $i = 1, \cdots, n_x$, and $\hat{y}_i$, $i = n_x + 1, \cdots, n$ need to be updated in each iteration. By using initial values obtained from the convenient LSS fit, the program usually takes only a few iterations (four or five) to converge.

5. Program Description and Operation

The program is written in ANSI FORTRAN and has been tested on a 386 personal computer. (Appendix contains a list of the files necessary to perform the analyses described in this document.) The program prompts the user for the following information that is needed to carry out the analyses (see Figure 6).

- A brief description of the job.

- The data type: either the digitized image or edge points. In our installation, the image data consist of 512×440 2-byte hexadecimal values, corresponding to setting the parameters IPXLS and JPXLS to 512 and 440, respectively, in the subroutine OFIBER. To read image data other than hexadecimal format, just replace the OPEN and READ statements in the subroutine HEXGET with the appropriate input routines. For the edge table, the data consist of X and Y coordinates of edge points in micrometers. The maximum allowable number of cladding edge points is 1216, and the maximum allowable number of core edge points is 88, corresponding to setting the parameters NMOBS1 and NMOBS2 to 1216 and 88, respectively, in the subroutine OFIBER. If the input contains both the cladding and core edge tables, the first 1216 records contain the cladding edge points and the remaining 88 records contain the core edge points.

- The name of the file containing either the digitized image or the edge points.

- If a digitized image is specified as the data input, a file that contains the calibration coefficients must also be given. The calibration is used to determine the correspondence between pixel units and micrometers. The file consists of four records. The first record contains the degree of the calibration polynomial equation for the horizontal pixel. The second record contains the coefficients of the polynomial. The third record contains the degree of the calibration polynomial equation for the vertical pixel. The fourth record contains the coefficients. For example, suppose it is determined from the calibration that the relationship between the horizontal pixel number and distance (in micrometers) is

  $$X = 0.273054\,P + 9.332 \times 10^{-8}\,P^2,$$

  where P is the pixel number, and relationship between the vertical pixel number and distance is

  $$Y = 0.320220\,P + 3.9825 \times 10^{-6}\,P^2.$$

  Then the file would contain the following four records

```
2
0.273054 9.332E-8
2
0.320220 3.9825E-6
```

- If the data input is a digitized image, an adjustment factor that is used in the edge detection must also be given. If 0 is entered, the edge points are located by estimating the inflection points of the intensity distribution. If a number between 0 and 1 is entered, the program finds the edge points at that specific fraction of the maximum intensity.

- If the data input is a digitized image, it is possible to write the edge points (in micrometers) to a file.

- The selection of the method(s) of geometric parameter estimation. The input can be 1, 2, 3, or any combinations of these three digits. For example, "123" requests that all three methods be used to estimate the geometric parameters. For the gray–scale LSS method, the program assumes that the first half of the edge points have the structure of fixed y–coordinate and noisy x–coordinate.

- The name of the file to capture the output.

In this example, the digitized image stored in the file "yaw5.dat" was used, and the calibration coefficients were contained in the file "cal.dat". The edge points were found as the inflection points of the intensity distribution. It also requested that the edge tables be written to the file "yaw5edge.out". The geometric parameters were obtained using both the convenient and the gray–scale LSS methods. The output of the example, saved in the file "yaw5.out", is given in Figure 7.

The program first prints the description of the run and the method employed in edge–finding. It then prints the parameter estimates for each method requested. In this example, since the input was the digitized image, the results from both the cladding and the core data are given. The gray–scale LSS solution also provides the approximate standard errors of the estimates for the cladding data (values inside the parentheses). The estimated cladding diameter is 125.0209 μm and the noncircularity is 0.0776 percent, based on the convenient LSS solution; and the cladding diameter is 125.0217 μm and the noncircularity is 0.0772 percent, if the gray–scale LSS method is used. Thus, these two methods are in good agreement for the fiber YAW5. The standard error of the estimated cladding diameter is calculated to be 0.0247 μm. We can use the distance between the cladding and the core centers as a possible measure for the concentricity error, which is 0.243 μm based on the gray–scale LSS solution.

```
Enter fiber ID
This is an example with fiber YAW5.

Enter type of data
  0 : Raw image data
  1 : Previously saved edge table (both cladding and core)
  2 : Previously saved edge table (cladding only)
  3 : Previously saved edge table (core only)
0

Enter file name containing data
yaw5.dat

Enter file name containing calibration coeff.
cal.dat

Enter adjustment factor for edge detection
0

Amount of data to save
  0 : No save
  1 : Both cladding & core edge tables
  2 : Cladding edge table only
  3 : Core edge table only
1

Enter file name to save edge table(s)
yaw5edge.out

Enter method(s) of parameter estimation
  1 : Convenient LSS
  2 : Orthogonal-distance LSS
  3 : Gray-scale LSS
13

Enter file name to save results
yaw5.out
```

Figure 6. The input for the example run.

```
This is an example with fiber YAW5.
 Use inflection points for edge detection.
 Units are: micrometers for lengths and radians for angles.

  Method: Convenient LSS
    Data: Cladding
          Major Axis =  62.55896
          Minor Axis =  62.46192
               Angle =  -1.47398
      Center X-Coord =  70.64283
             Y-Coord =  73.75232

  Method: Gray-scale LSS
    Data: Cladding
          Major Axis =  62.55914  ( 0.01880)
          Minor Axis =  62.46261  ( 0.01605)
               Angle =  -1.46443  ( 0.01638)
      Center X-Coord =  70.64265  ( 0.00951)
             Y-Coord =  73.74964  ( 0.01040)

  Method: Convenient LSS
    Data: Core
          Major Axis =   3.68977
          Minor Axis =   3.58388
               Angle =  -1.13511
      Center X-Coord =  70.74800
             Y-Coord =  73.95351

  Method: Gray-scale LSS
    Data: Core
          Major Axis =   3.75637
          Minor Axis =   3.60715
               Angle =  -1.45926
      Center X-Coord =  70.76023
             Y-Coord =  73.96218
```

Figure 7. The output for the example run.

6. References

[1] Young, M.; Hale, P.D.; Mechels, S.E. Optical Fiber Geometry: Accurate Measurement of Cladding Diameter. Journal of Research of the National Institute of Standards and Technology 98: 203–216; 1993.

[2] Mechels, S.E.; Young, M. Video Microscope with Submicrometer Resolution. Applied Optics 30: 2202–2211; 1991.

[3] Brilliant, N.A.; Young, M. Video Microscopy Applied to Optical Fiber Geometry Measurements. Natl. Inst. Stand. Technol. Technical Note 1369; 1994.

[4] Fiber Optics Test Procedure FOTP–176: Measurement Method for Optical Fiber Geometry by Automated Gray–Scale Analysis. Telecommunications Industry Association–Electronic Industries Association, 2001 Pennsylvania Avenue, N.W., Washington, DC

[5] Rousseeuw, P.J. Least Median of Squares Regression. Journal of the American Statistical Association 79: 871–880; 1984.

[6] Rousseeuw, P.J.; Leroy, A.M. Robust Regression and Outlier Detection. New York: Wiley; 1987.

[7] Forbes, A.B. Fitting an Ellipse to Data. National Physical Laboratory (U. K.) Report DITC 95; 1987.

[8] Vecchia, D.F.; Wang, C.M.; Young, M. Outlier–Resistant Fitting of Gray–Scale Images Illustrated by Optical Fiber Geometry. Proceedings of the Measurement Science Conference, Los Angeles, CA; 1993.

[9] Fuller, W.A. Measurement Error Models. New York: Wiley; 1987.

[10] Mamileti, L.; Wang, C.M.; Young, M.; Vecchia, D.F. Optical Fiber Geometry by Gray–Scale Analysis with Robust Regression. Applied Optics 31: 4182–4185; 1992.

[11] S–PLUS User's Manual Volume 2, Version 3.0. Statistical Sciences, Inc., Seattle, WA; 1991.

[12] Hawkins, D.M.; Simonoff, J.S.; Stromberg, A.J. Distributing a Computationally Intensive Estimator: The Case of Exact LMS Regression. University of Kentucky Technical Report No. 339; 1992.

[13] Savitzky A.; Golay, M.J.E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Analytical Chemistry 36: 1627–1639; 1964.

Appendix

The accompanying disk contains the following files:

- FIBER.FOR: the Fortran source file,
- FIBER.EXE: the program that can be executed on a 386 (or better) personal computer,
- YAW5.DAT: the raw image data used in the example,
- CAL.DAT: the file containing the calibration coefficients used in the example,
- YAW5EDGE.OUT: the file containing the edge tables (output from the example),
- YAW5.OUT: the file containing the estimation results (output from the example).

Trade names are used to allow the reader to use the program effectively; no endorsement by NIST or the authors is implied.

NIST *Technical Publications*

Periodical

Journal of Research of the National Institute of Standards and Technology—Reports NIST research and development in those disciplines of the physical and engineering sciences in which the Institute is active. These include physics, chemistry, engineering, mathematics, and computer sciences. Papers cover a broad range of subjects, with major emphasis on measurement methodology and the basic technology underlying standardization. Also included from time to time are survey articles on topics closely related to the Institute's technical and scientific programs. Issued six times a year.

Nonperiodicals

Monographs—Major contributions to the technical literature on various subjects related to the Institute's scientific and technical activities.

Handbooks—Recommended codes of engineering and industrial practice (including safety codes) developed in cooperation with interested industries, professional organizations, and regulatory bodies.

Special Publications—Include proceedings of conferences sponsored by NIST, NIST annual reports, and other special publications appropriate to this grouping such as wall charts, pocket cards, and bibliographies.

Applied Mathematics Series—Mathematical tables, manuals, and studies of special interest to physicists, engineers, chemists, biologists, mathematicians, computer programmers, and others engaged in scientific and technical work.

National Standard Reference Data Series—Provides quantitative data on the physical and chemical properties of materials, compiled from the world's literature and critically evaluated. Developed under a worldwide program coordinated by NIST under the authority of the National Standard Data Act (Public Law 90-396). NOTE: The Journal of Physical and Chemical Reference Data (JPCRD) is published bi-monthly for NIST by the American Chemical Society (ACS) and the American Institute of Physics (AIP). Subscriptions, reprints, and supplements are available from ACS, 1155 Sixteenth St., NW, Washington, DC 20056.

Building Science Series—Disseminates technical information developed at the Institute on building materials, components, systems, and whole structures. The series presents research results, test methods, and performance criteria related to the structural and environmental functions and the durability and safety characteristics of building elements and systems.

Technical Notes—Studies or reports which are complete in themselves but restrictive in their treatment of a subject. Analogous to monographs but not so comprehensive in scope or definitive in treatment of the subject area. Often serve as a vehicle for final reports of work performed at NIST under the sponsorship of other government agencies.

Voluntary Product Standards—Developed under procedures published by the Department of Commerce in Part 10, Title 15, of the Code of Federal Regulations. The standards establish nationally recognized requirements for products, and provide all concerned interests with a basis for common understanding of the characteristics of the products. NIST administers this program in support of the efforts of private-sector standardizing organizations.

Consumer Information Series—Practical information, based on NIST research and experience, covering areas of interest to the consumer. Easily understandable language and illustrations provide useful background knowledge for shopping in today's technological marketplace.

Order the **above** *NIST publications from: Superintendent of Documents, Government Printing Office, Washington, DC 20402.*

Order the **following** *NIST publications—FIPS and NISTIRs—from the National Technical Information Service, Springfield, VA 22161.*

Federal Information Processing Standards Publications (FIPS PUB)—Publications in this series collectively constitute the Federal Information Processing Standards Register. The Register serves as the official source of information in the Federal Government regarding standards issued by NIST pursuant to the Federal Property and Administrative Services Act of 1949 as amended, Public Law 89-306 (79 Stat. 1127), and as implemented by Executive Order 11717 (38 FR 12315, dated May 11, 1973) and Part 6 of Title 15 CFR (Code of Federal Regulations).

NIST Interagency Reports (NISTIR)—A special series of interim or final reports on work performed by NIST for outside sponsors (both government and non-government). In general, initial distribution is handled by the sponsor; public distribution is by the National Technical Information Service, Springfield, VA 22161, in paper copy or microfiche form.