Optimal object and edge localization in the presence of correlated noise

Kenneth M. Hanson

Los Alamos National Laboratory
Mail Stop P912, Los Alamos, New Mexico 87545

Abstract

The ideal observer for the task of object localization in the presence of correlated noise is implemented by means of the minimum chi-squared method, which is equivalent to the maximum likelihood method when the noise is additive and normally distributed. The prewhitening approach is explored in which the noise in the data is made uncorrelated by filtering the data with the filter $S^{-1/2}$, where S is the noise power spectrum. The location of the object is then found by fitting a model of the object to the prewhitened data. A measure of the goodness of fit is proposed that is based upon the serial correlation in the prewhitened residuals, the difference between the data and the fit. These concepts are demonstrated by applying them to simulated data that possess known noise characteristics. It is shown that the accuracies predicted for the ideal observer can be achieved by the prewhitening technique.

Introduction

It was shown in earlier work[1] that the ideal observer places increased emphasis on the high-frequency components of an image when asked to perform high-order tasks such as object localization, width determination and measurement of the separation between objects. For example, in typical medical film/screen systems the ideal observer may derive significant information from the high-frequency regions where the modulation transfer function (MTF) and the noise power spectrum have dropped by more than an order of magnitude. In this paper we consider the implementation of the ideal observer for the task of localization of a known object from noisy data in which the noise correlations are known. While the specific task of localization of objects or edges given one-dimensional data is addressed, the approach taken has general application to essentially all parameter estimation tasks. The often-mentioned trick of prewhitening the noise in the data is employed to convert it to white, or uncorrelated noise. Then a standard parameter estimation procedure that assumes the noise is uncorrelated, such as minimum chi-squared or least squares, may be used. We discuss the problems associated with the prewhitening procedure and the robustness of the technique. It is demonstrated through simulation that this method achieves an accuracy in object localization that essentially matches that predicted by the formulae given in Ref. 1. It is shown that the minimum chi-squared procedure also provides a reliable estimate of the accuracy of the estimated parameters when the data have been prewhitened. Of further interest is the penalty in accuracy incurred when noise correlations are not taken into account.

Whenever a parametric model is fit to a set of data, it is important to determine whether the data are matched to within their statistical accuracy. The goodness of fit is traditionally obtained by comparison of the chi-squared value with the number of degrees of freedom, that is, the number of (independent) data points minus the number of fitted parameters. This method suffers from its reliance on accurate knowledge of the rms deviation of the noise. We discuss an improved method for determining the goodness of fit that is based upon the observed serial correlation between the residuals (deviations between the fit and the data) and the application of this method to situations in which correlations in the noise are known to exist.

Optimum estimators

It will be assumed that aside from stochastic noise an observed set of data can be adequately predicted by a known model of reality. Typically there are one or more parameters in the model that are unknown. It is the task of an estimation procedure or estimator to determine the unknown parameters. If the available data are not degraded by noise and the model is correct, it is normally possible to determine one set of the unknown parameters exactly. If the data are subjected to noise, however, the estimated values of the parameters will be uncertain to some extent. An optimum estimator is one that minimizes the error in the estimated parameters with respect to some weighting function, called a cost function, averaged over the ensemble of possible noise samples. It is desirable that the averages of the estimated parameters be their actual values, that is, that the estimator be unbiased. It is well known that the maximum likelihood estimator yields optimum results in a wide variety of circumstances.[2] Suppose a set of measurements y_i are subject to additive and normally distributed noise. The likelihood or probability density function for a particular set of y_i is[3]

$$\mathscr{L} \sim \exp\left[-\frac{1}{2} \sum_{ij} (y_i - \bar{y}_i)\, w_{ij}\, (y_j - \bar{y}_j) \right] \tag{1}$$

where the matrix W, which weights the importance of correlated noise fluctuations in two different measurements, is the inverse of the noise covariance matrix

$$[W^{-1}]_{ij} = C_{ij} = \langle (y_i - \bar{y}_i)(y_j - \bar{y}_j) \rangle \tag{2}$$

Here the angular parentheses indicate an average over the ensemble of all noise fluctuations. By the definition both C and W are symmetric. The noise covariance matrix is the Fourier transform of the noise power or Wiener spectrum. In the above two expressions, $\bar{y}_i$ is the mean value of y_i, which is also the value of y_i that would be obtained in the absence of noise. We see from Eq. (1) that the likelihood function is simply a product of Gaussians centered on the mean values $\bar{y}_i$. If the noise is uncorrelated, C is diagonal and hence so is W. Then the likelihood function would include only one Gaussian for each y_i. It is the off-diagonal elements in C arising from correlations in the noise that lead to additional Gaussian factors involving cross-products between different measurements.

Suppose that the measurements y_i are to be modelled by a function f that is dependent upon some number of parameters whose values are to be estimated from the measurements. Replacement of the $\bar{y}_i$ in Eq. (1) by values f_i predicted by the model for a particular set of parameters yields the likelihood of obtaining a particular set of y_i under those circumstances. The principle of maximum likelihood indicates that the optimum way to determine the unknown parameters is to find that set of parameters that maximizes this likelihood function. Instead of maximizing the likelihood function itself, it is convenient to minimize the negative of the logarithm of the likelihood function. Half of this is chi-squared

$$\chi^2 = \sum_{ij} (y_i - f_i)\, w_{ij}\, (y_j - f_j) \tag{3}$$

By taking the logarithm of the likelihood function, the problem has been reduced to one of finding a minimum in a quadratic function of the residuals $y_i - f_i$. It should be emphasized that the minimum χ^2 solution for the unknown parameters is optimum for Gaussian distributed, additive noise as it is identical to the maximum likelihood estimate in that situation.

When the noise is uncorrelated, W is diagonal and Eq. (3) reduces to the familiar expression

$$\chi^2 = \sum_i \frac{(y_i - f_i)^2}{\sigma_i^2} \tag{4}$$

where σ_i^2 is the noise variance for the ith measurement. Since the simplifying assumption that noise is uncorrelated is often made, this equation abounds in textbooks on data analysis and the methods for finding the values of the unknown parameters that minimize χ^2 are well documented. When f is linearly related to the parameters, the solution to the resulting linear matrix equation is easily found by the Newton-Raphson technique.[4,5] When the parametric dependence is nonlinear, the minimum chi-squared solution usually may be found by iteration, making the assumption that f is approximately linearly related to the parameters in local regions.

If the noise is stationary (has the same variance everywhere) as well as uncorrelated, χ^2 simplifies to

$$\chi^2 = \sigma^{-2} \sum_i (y_i - f_i)^2 \tag{5}$$

In this case the condition of minimum χ^2 becomes equivalent to the minimization of the sum of the squared residuals usually referred to as least squares. We see that the least squares solution to parameter estimation will yield optimum results when the noise is additive, normally distributed, uncorrelated, and stationary. Expanding Eq. (5), we find that

$$\chi^2 = \sigma^{-2} \left[\sum_i y_i^2 + \sum_i f_i^2 - 2 \sum_i y_i f_i\right] \tag{6}$$

If the index i refers to samples at an ordered sequence of evenly spaced positions, the last term of this expression is recognized as being proportional to the cross correlation between the data y_i and the predicted function f_i

$$\Phi_j = \sum_i y_i f_{i+j} \tag{7}$$

Provided that the sum over f_i^2 is unchanged as the position of the samples is shifted, which occurs when f is constant outside the observation interval, then the condition of maximum cross correlation is identical to that of minimum χ^2. Thus the cross correlator is optimal only for uncorrelated noise. It is suited to the two tasks of amplitude estimation and object localization. In the former case, j is fixed to correspond to the known location of the object. In the latter, j is varied to find the maximum value of the cross correlation. The maximum in Φ_j can be interpolated between the discrete samples to achieve sub sample-spacing accuracy. Cross correlation is not as useful for other tasks in which the parametrization is not simply related to the positional index j. It is also not optimal for correlated noise unless the noise in the data has been prewhitened, as described below.

It is worthwhile to mention the relationship between parameter estimation and the binary decision problem. Wagner[6] discussed the latter and derived the SNR for detection for arbitrary noise correlation but under the assumption that the observer treats the noise as white. Hanson[7] extended this using the prewhitening concept to obtain the SNR for the ideal observer, which properly takes into account the known noise correlation, and obtains the familiar expression for the matched filter. In the binary decision problem it is necessary to choose between two alternative functions on the basis of a set of data. The strategy is to construct a decision variable (ψ_{12} in Wagner's notation) that is the logarithm of the ratio of the likelihood functions for the two alternatives. Under the assumption of additive, normally distributed noise, this is similar in construction to that of χ^2. Expansion of ψ_{12} leads to Wagner's Eq. (6), which is closely related to our Eq. (6) in that a cross correlation term describes the dependence upon the data. If Wagner's alternative function $N^{(2)}$ is written as a Taylor series expansion of $N^{(1)}$ with respect to a parameter α, it is easily shown that ψ_{12} is proportional to $\Delta\alpha$ in the limit as $\Delta\alpha \to 0$. This allows one to interpret ψ_{12} as the signal in the SNR in that limit. This correspondence was used in Ref. 1 and similarly in Ref. 8 to determine the uncertainties in the estimation of various parameters for the ideal observer through the expedient of the expression of the SNR^2 for binary decision. Since the latter as well as the minimum χ^2 formalism are based on the same underlying theory and assumptions, it is expected that both should yield identical results. However, the generality of the minimum χ^2 approach allows one to address more complicated problems, such as the estimation of multiple parameters, than does the simple binary decision theory.

One approach to the incorporation of noise correlation in parameter estimation would be to minimize the complete χ^2 expression given in Eq. (3). Setting the derivative of Eq. (3) with respect to the positional parameter Δ to zero, identifying this derivative as that with respect to x, and using the symmetry of W, we find that

$$\sum_{ij} \left[\frac{\partial f}{\partial x}\right]_i w_{ij} (y_j - f_j) = 0 \tag{8}$$

This nonlinear equation may be solved using the standard method for solving Eq. (4). However, the nondiagonal nature of W for correlated noise complicates the solution of this problem. When the noise is white, W is diagonal and Eq. (8) amounts to an inner product over the data index of the derivative of f with the residuals. When W is not diagonal, a convolution between these two is required, lengthening the computation time.

Prewhitening approach

As an alternative approach to optimal estimation in the presence of correlated noise, we will explore the technique of prewhitening the noise[9] so that the more standard methods for minimizing Eq. (4), applicable to white noise, may be employed. The prewhitening approach may be made rigorous by invoking the singular value decomposition theorem[10] for the correlation matrix C, which is square and symmetric and given by

$$C = U^T S U \tag{9}$$

where U is an orthogonal matrix ($U^{-1} = U^T$) and S is a diagonal matrix whose elements are the singular values of C. The weight matrix W may be written as

$$W = C^{-1} = U^T \Lambda U \tag{10}$$

where Λ is also diagonal and $\lambda_k = s_k^{-1}$. Then in matrix notation χ^2 from Eq. (3) becomes

$$\chi^2 = (Y^T - F^T) U^T \Lambda U (Y - F) = (Y^T - F^T) U^T \Lambda^{1/2} U U^T \Lambda^{1/2} U (Y - F) \tag{11}$$

where $\Lambda^{1/2}$ is a diagonal matrix with entries $\lambda_k^{1/2}$. It is observed that this is an inner product

$$\chi^2 = (\tilde{Y}^T - \tilde{F}^T)(\tilde{Y} - \tilde{F}) = \sum_k (\tilde{y}_k - \tilde{f}_k)^2 \tag{12}$$

where the ~ indicates the transformation,

$$\tilde{y}_k = [U^T \Lambda^{1/2} U Y]_k = \sum_{ij} \lambda_i^{1/2} u_{ik} u_{ij} y_j \tag{13}$$

and the same for $\tilde{f}_k$. Through this transformation it is possible to reduce the more complicated expression for χ^2 appropriate to correlated noise, Eq. (3), to the form of the simpler one for uncorrelated noise, Eq. (4), with $\sigma_i = 1$. Thus, the transformed residuals clearly are uncorrelated, that is they have been prewhitened. Given a specific matrix C, the matrices U and S may be found using standard computer algorithms. By transforming the data y_i and the predicted function values f_i using Eq. (13), it is possible to simplify the form of χ^2, Eq. (12), and make use of the standard minimum χ^2 algorithms based on this simpler form.

If the noise is stationary, it is possible to simplfy the prewhitening procedure. In this case the matrix C is Toeplitz, that is, each row of C has the same entries as the one above it but shifted one element to the right. Thus the entries along diagonal lines are identical. Often it is reasonable to approximate Toeplitz matrices by circulant matrices in which each row is a right cyclic shift of the row preceding it. Then the orthogonal matrix U may be identified as the discrete Fourier transform[11] and the diagonal elements of S are the (real) Fourier amplitudes of the transformation of the first row of C or the noise power spectrum. The transformation given in Eq. (13) is identified as a filtering operation since y_i is Fourier transformed, multiplied by λ_i and inverse transformed, where the filter is $\lambda_i^{1/2} = s_i^{-1/2}$ or the reciprocal of the square root of the noise power spectrum. This was to be anticipated because it was desired to change the original noise power spectrum S to a constant, and a filtering procedure alters all spectra by the square of the filter. Note that the variance in the noise is unity after prewhitening if the above prescription is strictly followed. The prewhitening filter may be multiplied by an arbitrary factor to arrange any desired normalization. A particularly appealing choice is to maintain the low-frequency normalization by using $[S(f=0)/S]^{1/2}$.

The possibility that the noise power spectrum goes to zero over some frequency interval must be contended with. The prewhitening filter may be regularized through the choice

$$\frac{S + \varepsilon}{S^2 + \varepsilon^2} \tag{14}$$

where ε is a small positive number. The ε in the numerator is needed to prevent the filter from going to zero as S goes to zero, which would preclude the extraction of possibly useful signal information. The maximum value this filter can attain is ε^{-1}. If S becomes smaller than ε, the spectrum of the prewhitened noise will not be flat, contrary to the assumption of the subsequent estimator. This is where the optimality of the prewhitening approach can break down.

A few precautions must be taken in applying the prewhitening filter. It is known that when the filtering operation is carried out by means of the discrete Fourier transform, the result is identical to a cyclic convolution. Thus, wraparound occurs in which the data at one end of the interval are strongly affected by those from the other end. This effect arises from the circulant approximation of the Toeplitz matrices and should be avoided by extending the interval of the data sufficiently. Extension of the interval also allows the use of fast Fourier transforms, whose speed is optimimum for interval lengths of powers of two. In the present work the interval length is chosen to be twice the next power of two greater than or equal to the length of the data segment. In extending the data interval it is necessary to fill in the mock data points with some specific values. If the mock data points are set to zero and the actual data at the ends of the valid interval are far from zero, severe edge effects may result. To ameliorate this unwanted behavior, in the present work the mock data points are set equal to a linear function that connects the first and last data points over the region of extension. Even with this precaution the discontinuity in slope that can occur at the ends of the data interval can produce undesirable effects when filtered. Therefore it is desirable to use a longer interval for the prewhitening filtering than is used for the fitting of the actual data. It is imperative that whatever procedure is followed in prewhitening the data, the same procedure be applied to the fitting function f in the minimum χ^2 calculation.

Goodness of fit

After a parametric model has been fit to a set of data, it is very desirable to know how closely the fitted function matches the data. An important reason for determining the goodness of fit is to judge the adequacy of the model. If it is determined that the observed residuals, the deviations between the fit and the data, are improbable, given the known statistical characteristics of the noise, then the model used to fit the data is suspect and the derived parameters may be meaningless. A commonly used test for the goodness of fit is to compare the χ^2 value to a table of the integrated χ^2 probability distribution for the number of degrees of freedom,[12] that is, the number of data points minus the number of parameters being estimated. If the indicated percentile value is improbable, for example if it is less than 5% or greater than 95%, then the fit is suspect. Of course, this test relies on the noise being uncorrelated and so can only be used for correlated noise after the prewhitening procedure described above. However, the χ^2 test is not very useful as it relies on accurate knowledge of the noise variance. For 100 degrees of freedom the rms value of the noise must be known to better than 10% for the test to be meaningful.

We propose to use an alternate measure for the goodness of fit that is based upon the observed serial correlation of the residuals. This approach has merit because it resembles the method employed by trained human observers to distinguish bad fits from good ones. Durbin and Watson[13,14,15] define the statistic

$$d = \frac{\sum_{i=1}^{n-1} (r_{i+1} - r_i)^2}{\sum_{i=1}^{n} r_i^2} \qquad (15)$$

where the r_i are the sequence of residuals $y_i - f_i$. Note that the denominator is an estimate of the variance in the residuals and that no estimate of the noise variance is required. Durbin and Watson have computed the cumulative probability distribution of d assuming that the residuals are those of linear regression fits to data that have uncorrelated, normally distributed noise. The specific value of d obtained for a given fit may be compared with their tables to obtain a percentile value as in the χ^2 test described above. If the percentile value is very low or very high, one may conclude that the residuals probably do not conform to the assumption of normally distributed, uncorrelated noise. As with the χ^2 test, the Durbin-Watson test is only useful when applied to fits of uncorrelated noise or correlated noise that has been prewhitened. For large numbers of degrees of freedom, the probability distribution for d is roughly Gaussian with a mean of about 2 and a variance of 4/(degrees of freedom). Thus, values of $d \leq 1.0$ are fairly improbable if the residuals conform to the assumptions. It is easy to show that the mean of d is closely related to the first moment of the power spectral density of the residuals. For noise that possesses short range, positive correlation, the mean of d will be less than 2.

Estimation of parameter uncertainties

It is usually desirable to know how accurately the parameters have been estimated. The statistical accuracy with which a parameter can be estimated clearly depends upon the rate of dependence of the likelihood function upon that parameter. In the vicinity of maximum likelihood, this is related to the curvature of the likelihood function, the logarithm of which is linearly related to χ^2. Thus, the accuracy in the estimation of a parameter is simply related to the second derivative of χ^2 with respect to that parameter.[16] Using Eq. (4) for uncorrelated noise, we find the standard deviation in position σ_Δ is given by

$$\sigma_\Delta^{-2} = \frac{1}{2} \frac{\partial^2 \chi^2}{\partial \Delta^2} = \sum_i \frac{1}{\sigma_i^2} \left[\frac{\partial f}{\partial x}\right]_i^2 \qquad (16)$$

The counterpart of this for uncorrelated Poisson noise was given in Ref. 17. A variation of the matched filter for this situation was also derived there. Equation (16) states that it is the derivative of a function that carries information about its position, which is eminently reasonable.

The optimal positional accuracy derived in Ref. 1 using the binary decision approach mentioned above, appropriate to 1-D signals, is given by

$$\sigma_{\Delta}^{-2} = 2\pi \int \frac{H^2}{S} F^2 u^2 \, du \tag{17}$$

where H is the contrast transfer function, F is the Fourier transform of the original unblurred function, S is the noise power spectrum, and u is the spatial frequency variable. Comparison of Eqs. 16 and 17 reveals that they are identical in the limit that the summation in the former becomes an integration. Then the two expressions are related by Parseval's theorem, Eq. (16) referring to the power in the spatial domain and Eq. (17) referring to the power of the identical quantity represented in the frequency domain. Accounting for the factors in Eq. (17) in terms of their inverse transforms, we see HF corresponds to the measured function f, $S^{-1/2}$ arises from the prewhitening filter that is necessary to derive Eq. (16) and is related to σ_i and u corresponds to the derivative with respect to x. In practice the evaluation of these two expressions may yield different results, because the discrete samples of f_i may be unequally spaced or not close enough or may not span the complete function f.

If an estimation procedure that is optimum for white noise is used to analyze data possessing correlated noise (without prewhitening), the resulting uncertainty in position is given by

$$\sigma_{\Delta}^{-2} = 2\pi \frac{\left[\int H^2 F^2 u^2 \, du \right]^2}{\int S H^2 F^2 u^2 \, du} \tag{18}$$

This result is obtained by applying the same approach in Ref. 1 to the formula derived by Wagner[6] for this non-optimum observer. It can be shown that the uncertainty derived from Eq. (18) is greater than or equal to that derived from Eq. (17).

Examples

It is possible to test the prewhitening and fitting procedures described above by generating numerous data sets with known noise characteristics and comparing the response of the algorithms against that predicted. To that end these procedures were implemented on a CDC7600 computer. Data were generated by adding Gaussian-distributed, random noise that had been filtered by the square root of a known noise power spectrum S to object functions filtered by a known contrast transfer function H. The two sets of spectra used in these studies are shown in Fig. 1a and 1b, hereafter designated as spectra A and B, respectively. In spectra A, S (designated NPS in the Figure) represents a rather unlikely condition in that it falls by a factor of 10^4 before it levels off. The situation for spectra B is more likely to occur in actual imaging systems and, indeed, closely resembles those obtained for Hi-Plus/XRP screen/film at diagnostic energies.[18,19] The rms values of the noise generated were 0.103 and 0.144, respectively. Two object functions were used. The first was a rectangle with a width of 1.0 mm and an amplitude of 1.0. The second was a Gaussian with a width of 0.2 mm FWHM and an amplitude of 2.83. The detection sensitivity index (d') for the ideal observer is about 10 for all four combinations of spectra with objects. One sample of the data sets generated for the rectangle combined with spectra A is shown in Fig. 2a. The 60 data points shown represent a subsample from 90 data points that are filtered on the basis of 256 points suitably extended with a taper function. It is observed that the smoothness in the object, centered on 3 mm, closely resembles the smoothness in the noise as suggested by the spectra in Fig. 1a.

The result of applying a conventional nonlinear fitting procedure, based on Eq. (4), to the central 40 data points from Fig. 2a is shown in Fig. 2b. Only the object's position is allowed to vary in this fit. The object's amplitude and width and the zero value of the background are assumed to be known. Fig. 2c shows the same data from Fig. 2a after prewhitening. The 60 data points are extended by a linear taper to a length of 128 for the filtering operation. It is seen that the end points respond somewhat violently to the filtering operation as it amounts to an edge enhancement and the tapered extension creates a discontinuity in slope at the end. For this reason, the central section consisting of 40 data points is used for fitting and the end points are discarded. The fit with respect to position of the prewhitened data is shown in Fig. 2d. The ringing in the predicted function arises from the abrupt frequency cutoff of the effective filter after prewhitening $H S^{-1/2}$,

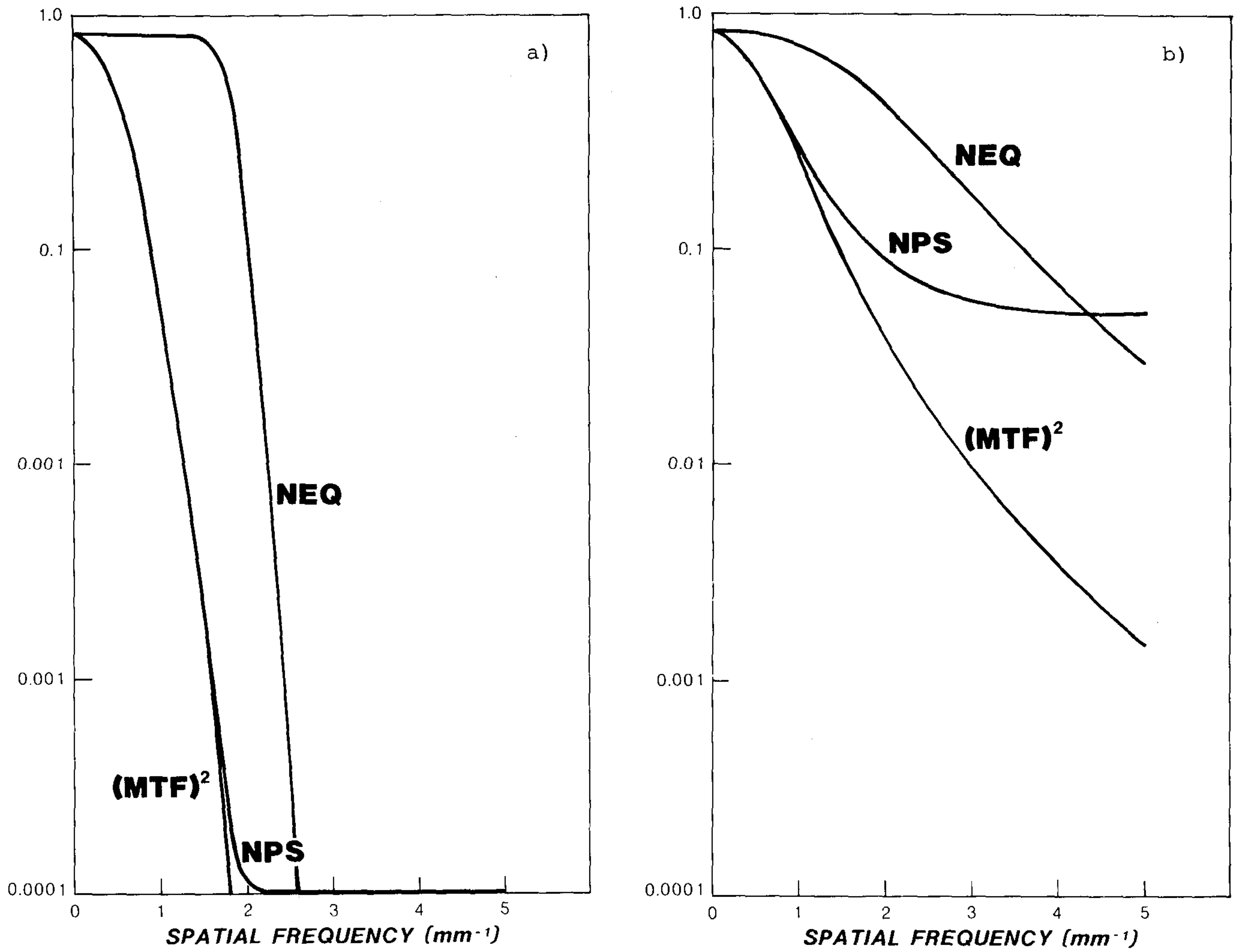

Figure 1. The square of the modulation transfer function (MTF), the noise power spectrum (NPS) and the noise equivalent quantum (NEQ) spectrum used to generate data for testing the algorithms. The designations are a) spectra A, and b) spectra B.

which is the same as $NEQ^{1/2}$. It should be clear that the fluctuations of the data about the fitted curve are strongly correlated in Fig. 2b whereas they appear uncorrelated in Fig. 2d. The prewhitening works!

A test series of 200 data sets were generated for each of the four combinations of object functions and spectra. Each data set was fit by the known function allowing only its position to vary. In all cases the mean of the estimated position was the same as that used to generate the data to very high accuracy. Thus, both the prewhitened and non-prewhitened procedures yield unbiased estimates. Each series started with the same seed for the pseudorandom number generator, so the results are not statistically independent. Table I summarizes the predicted and observed accuracies obtained with and without prewhitening. The accuracies are substantially improved by the prewhitening procedure for spectra A and modestly improved for spectra B. Since the correct rms noise is used in each fitting process, the average χ^2 value in each case is about 39, the same as the number of degrees of freedom. However, as can be seen from the Table, the result is an underestimate of the position uncertainty when Eq. (16) is used and the data have not been prewhitened. This is a consequence of using Eq. (4) to calculate χ^2 rather than Eq. (3). The mean value of the Durbin-Watson parameter d is 1.96 for the fits to the prewhitened data, which is close to the expected value of 2. But the mean d value is 0.089 for spectra A and 0.76 for spectra B without prewhitening, indicating that the residuals in these fits are correlated. This may be taken to imply that the predicted uncertainties in the parameter estimation are incorrect.

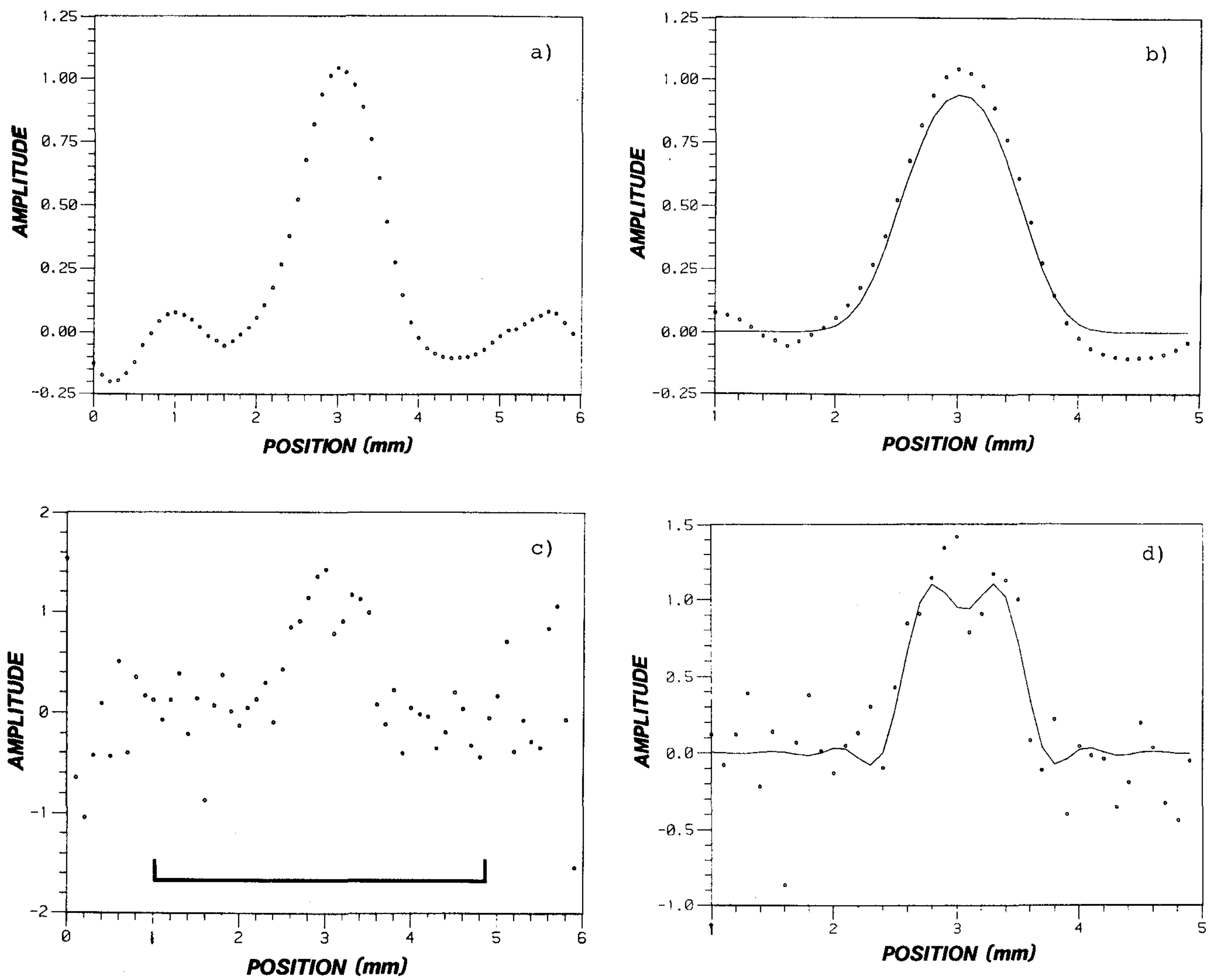

Figure 2. a) A data sample generated using spectra A and a rectangular object of width 1 mm and unity amplitude, b) the minimum χ^2 fit with respect to object position to a subsample of the original data, c) the same data after prewhitening, and d) a fit of the prewhitened object function to the subsample of (c) indicated by the bracket. The prewhitening procedure reduces the rms deviation in the estimated position by 34%.

TABLE I

Summary of predicted and observed rms deviations in object localization based on 200 sets of trial data. Spectra A corresponds to Fig. 1a and B to Fig. 1b.

			σ_x (mm)		
Object	Spectra	Prewhiten	Eq. 17 or 18	Eq. 16	Actual
Rectangle	A	Yes	.038	.038	.041
		No	.054	.023	.055
Point	A	Yes	.017	.018	.017
		No	.040	.021	.041
Rectangle	B	Yes	.039	.038	.045
		No	.047	.027	.049
Point	B	Yes	.016	.017	.015
		No	.021	.017	.023

The robustness of the prewhitening technique may be tested by assuming the wrong noise power spectra for the prewhitening filter. It is found that the assumption of an NPS that is ∓20% wider than the NPS used to generate the data yields only moderately worse fitting accuracies. For spectra B the positional uncertainties increase by less than 2% and the $\bar{d}$ values are 2.03 and 1.82. For spectra A, the uncertainty increases by as much as 44% and the $\bar{d}$ values are 2.31 and 1.25. It is worse to assume a narrower NPS than is correct for the data. While the value of d obtained helps indicate when the wrong prewhitening filter has been used, it is not as sensitive as one would like. In the face of no information about the NPS of a given data sample, it is probably safest to assume white noise, that is, use no prewhitening. When the NPS is known moderately well, it is best to overestimate the width of the NPS for the prewhitening calculation.

Discussion

We have shown that it is possible to construct an optimal estimator that can take into account known noise correlation using the prewhitening approach. Substantially improved performance can be achieved over that obtained by conventional estimators, which are optimal only for uncorrelated noise. The incorporation of noise correlation into the analysis also allows a reliable measure of the goodness of fit to be defined and provides an accurate estimation of the statistical uncertainties in the derived parameters. In the future it would be desirable to compare the prewhitening method to one based upon the minimization of the complete χ^2 expression given in Eq. (3).

Acknowledgments

The author is grateful to Robert F. Wagner, Arthur E. Burgess, Katherine Campbell, and George W. Wecksung for many helpful conversations. This work was supported by the U. S. Department of Energy under contract number W-7405-ENG-36.

References

1. Hanson, K. M., "Variations in Task and the Ideal Observer," Proc. SPIE, Vol. 419, pp. 60-67. 1983.
2. Van Trees, H. L., Detection, Estimation and Modulation Theory, Part I, pp. 52-73, John Wiley and Sons, New York, 1968.
3. ibid., p. 96f.
4. Hildebrand, F. B., Introduction to Numerical Analysis, p. 443ff, McGraw-Hill, New York, 1956.
5. Bevington, P. R., Data Reduction and Error Analysis for the Physical Sciences, pp. 222-231, McGraw-Hill, New York, 1969.
6. Wagner, R. F. , "Decision Theory and the Detail Signal-to-Noise Ratio of Otto Schade," Photogr. Sci. Eng., Vol. 22, pp. 41-46. 1978.
7. Hanson, K. M., "Detectability in Computed Tomographic Images," Med. Phys., Vol. 6, pp. 441-451. 1979.
8. Hanson, K. M., "On the Optimality of the Filtered Backprojection Algorithm," J. Comput. Assist. Tomo., Vol. 4, pp. 361-363. 1980.
9. Van Trees, H. L., op.cit., pp. 290-297.
10. Golub, G. H., and Reinsch, C., "Singular Value Decomposition and Least Squares Solutions," Numer. Math., Vol. 14, pp. 403-420. 1970.
11. Andrews, H. C., and Hunt, B. R., Digital Image Restoration, pp. 211-224, Prentice-Hall, Englewood Cliffs, 1977.
12. Bevington, P. R., op. cit., p. 187ff.
13. Durbin, J., and Watson, G. S., "Testing for Serial Correlation in Least Squares Regression, I," Biometrika, Vol. 37, pp. 409-428. 1950.
14. Durbin, J., and Watson, G. S., "Testing for Serial Correlation in Least Squares Regression, II," Biometrika, Vol. 38, pp. 159-178. 1951.
15. Durbin, J., and Watson, G. S., "Testing for Serial Correlation in Least Squares Regression, III," Biometrika, Vol. 58, pp. 1-19. 1974.
16. Bevington, P. R., op. cit., p. 242ff.
17. Hanson, K. M., Bradbury, J. N., Cannon, T. M., Hutson, R. L., Laubacher, D. B., Macek, R. J., Paciotti, M. A., and Taylor, C. A., "Computed Tomography Using Proton Energy Loss," Phys. Med. Biol., Vol. 26, pp.965-983. 1981.
18. Doi, K., Holje, G., Loo, L. N., Chan, H. P., Sandrik, J. M., Jennings, R. J., and Wagner, R. F., MTF's and Wiener Spectra of Radiographic Screen-film Systems, HHS Publication(FDA) 82-8187, U.S.GPO, Washington, DC, 1982.
19. Sandrik, J. M., Wagner, R. F., and Hanson, K. M., "Radiographic Screen-film Noise Power Spectrum: Calibration and Intercomparison," Appl. Opt., Vol. 21, pp. 3597-3601. 1982.