Dr. R. L. Hardy*
Iowa State University
Ames, IA 50011

Least Squares Prediction

Multiquadric and covariance functions are compared, and applications in topographic interpolation, gravity anomaly prediction, and image processing are presented.

Introduction

Least squares prediction with covariance functions has been applied to photogrammetry during recent years (Kraus, 1972; Kraus and Mikhail, 1972; Schut, 1974). This theory, originally developed from time series analysis, is now most rigorously handled as a part of the theory of stationary random functions. The expression "covariance function" was popularized in geodesy by Heiskanen and Moritz (1967). From the point of view of communications engineering, the more appropriate term is "autocorrelation function" (Blackman and Tukey, 1959). The autocorrelation function is the normalized autocovariance

Abstract: *The basic principles of least squares prediction using both multiquadric functions and covariance functions are covered briefly. Similarities and dissimilarities of the multiquadric and covariance methods are discussed. Multiquadric kernels are based on geometric and physical considerations rather than stochastic processes as is the case of covariance kernels. Thus the procedure of determining and fitting empirical covariances to select an analytical covariance function is unnecessary in multiquadric analysis. Comparative test results involving topographic features in Hawaii are given. A similar type test involving gravity anomalies in Iowa is also discussed. In general multiquadric kernels were found to be superior to covariance kernels in accuracy and computational efficiency for these applications. Topography, gravity anomalies, and other phenomena are not stationary in the sense of stationary random functions, which is at the heart of the justification for least squares prediction with covariance functions. This is probably true of many of the phenomena of interest in photogrammetry and remote sensing. Least squares prediction with multiquadric functions is applied in a demonstration of its capability with respect to image processing and analysis.*

function (i.e., normalized to variance equals one at the origin). In a time series, the autocovariance function expresses the covariance between $X(t)$ and $X(t+\tau)$, where t is time and τ is a "lag" on the time axis. Yaglom (1962) uses the more general term "correlation function", in lieu of autocorrelation or covariance. These different terminologies are mentioned to aid the reader who may wish to study these functions in greater depth than can be given here. The terms "covariance", "correlation" and "autocorrelation" will be used more or less synonymously throughout this paper.

Another type of interpolation or prediction method, called "multiquadric functions" has

* Currently on leave from Iowa State University as a Senior Scientist in Geodesy, National Research Council; assigned to the National Geodetic Survey.

also been applied in photogrammetry (Hardy, 1972b). This method first appeared in the geophysics literature in 1971 (Hardy, 1971). It has been compared with the bicubic spline function in hydrological applications by Shaw and Lynn (1972) who reported several advantageous features of multiquadric analysis. Brown (1973) and Trotter (1975) have also reported favorably upon the use of multiquadric functions for applications in the use of satellite altimetry to determine the geoid. Thus multiquadric theory and applications have been considered by other investigators from time to time as a topic worthy of examination. Meanwhile, new developments in multiquadric equations have been studied by Hardy and his colleagues (Hardy, 1972a, 1975b; Hardy and Göpfert, 1975a).

In 1974 several authors in the photogrammetric field took note of multiquadric equations as an incidental aspect of studies involving covariance functions. In one case, multiquadric equations were included in the class of covariance functions without qualification (Rauhala, 1974). In another case multiquadric equations were said to involve the formalism of covariance theory, but an ambiguity in correlation at the origin of coordinates was noted, leading to a conclusion that multiquadric theory is a non-rigorous application of covariance theory (Schut, 1974). In the third case, multiquadric analysis was said to be identical with least squares interpolation (without filtering), if the hyperboloid is chosen as a covariance function. On the other hand, it was noted that some functions (such as the hyperboloid) do not seem to have a statistical basis. Thus it was concluded that these functions for which the expression "covariance function" is really unsuitable, should be included as special cases in a more broadly defined general interpolation model, which would still be called the covariance function (Assmus and Kraus, 1974). Needless to say, the preceding references have stimulated a more thorough study of covariance functions on this author's part.

Consequently, the situation will be clarified at this time (1976) with respect to the similarities and dissimilarities of multiquadric and covariance methods. During this presentation it will be shown that some multiquadric functions cannot possibly be covariance functions, based on strictly theoretical definitions applied to covariance theory. Moreover it will be shown that comparative studies conducted at Iowa State University indicate superior interpolation or computational characteristics for multiquadric functions over covariance functions in applications involving topography and gravity anomalies. A possible reason for this is that variations of topography and gravity are not necessarily stationary random functions, which is at the heart of the justification for the use of covariance functions. With respect to gravity in particular, the point of non-stationarity has been made previously by Williamson and Gaposchkin (1973). This may be true for many other phenomena, including those encountered in photogrammetry and remote sensing. If so, research beyond that reported in this presentation will be needed to develop an optimal prediction theory for non-stationary random functions. Meanwhile, multiquadric functions should be considered as an experimental alternative to the use of covariance functions for phenomena that are not strictly stationary.

Results of an experiment in image analysis using multiquadric functions are also included, not involving a comparison with covariance functions.

Covariance Kernel Functions

The bare essentials of covariance function theory will be presented, modeled to a great extent on the text book presentation by Heiskanen and Moritz (1967). The symbol Z will be used for generality to represent the ordinate of any phenomenon, in lieu of Δg for gravity anomalies which was the text book application.

The computation of one-sided covariance functions with respect to a profile, rather than on a two-dimensional plane or other surface, will be used. One-sided covariance functions can generate symmetric covariance surfaces by revolving them around the Z axis. The arbitrary function profile in Figure 1(a) has been pre-adjusted so that the first moments of the ordinates are zero, i.e.,

$$\sum_{i=0}^{n} Z_i = 0 \qquad i = 0,1,2,\ldots,16. \tag{1}$$

This produces a centered arbitrary function for use in determining the covariance function. For discrete sequences a common formula for the covariance is

$$C(s) = \frac{1}{N+1} \sum_{x=0}^{N} \xi(x+s)\,\xi(x) \qquad s = 0,1,2,\ldots \tag{2}$$

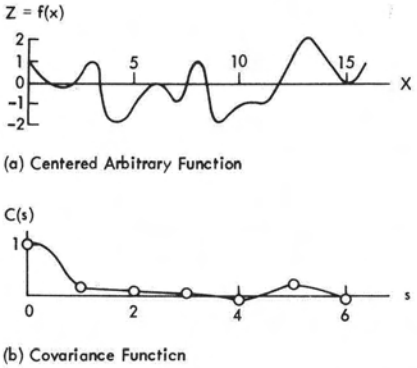

FIG. 1. Arbitrary and covariance functions.

in which the symbols x and s replace t and τ respectively. The symbol ξ represents a random variable.

The empirical covariance $C(s)$ for a computation for $s = 1,2,\ldots,6$ is shown in Figure 1(b) which is sufficient in this case. After $s = 3, C(s)$ oscillates near zero and theoretically should remain near zero. There are precautions one should consider in developing the empirical covariance function. For $s = 0, N = 16$, i.e., $N + 1 = 17$ which is the number of ordinates that can be squared at zero distance, meaned, and normalized to variance $C(0) = 1$. With $s = 1, N = 15$; with $s = 2, N = 14$; etc. Thus the number N steadily decreases and the mean becomes less reliable as s increases. According to Blackman and Tukey (1959) less than 10 percent of the available record in time series analysis is normally used to compute "apparent" covariances in communications engineering.

According to Yaglom (1962) the analytical definition of a correlation function produces the following characteristics

$$C(0) > 0 \qquad \text{(a)}$$
$$C(-s) = C(+s) \quad \text{(b)}$$
$$|C(s)| \leq C(0) \quad \text{(c)}$$

Specification (a) requires that correlation at the origin of coordinates be positive. Specification (b) requires the correlation function to be an even function, i.e., symmetric with respect to the origin. Specification (c) requires that the correlation at any distance s cannot be greater than the correlation at the origin. In other words the correlation at the origin is at least one of the maxima, if not the only maximum of the function. Several possible correlation functions are illustrated in Figure 2.

Multiquadric Kernel Functions

The first multiquadric kernels were standard quadric surfaces, of which the cone and hyperboloid are the principal examples (Hardy, 1971). In a profile mode these reduce to the absolute value function and hyperbola as illustrated in Figure 3. Although they are even functions as in (b) above they obviously do not satisfy the important specifications (a) and (c) for covariance functions. Moreover, the "bounded" cone and "bounded" hyperboloid fail to have non-negative Fourier transforms. Non-negative Fourier transforms are an indirect requirement of the famous Wiener-Khintchine relations for correlation functions and power spectra. These relations state that the covariance function of a function is the Fourier transform of the spectral density of that function. Spectral densities must be non-negative. Therefore a negative, or partially negative Fourier transform of an assumed covariance function, is proof that the assumed covariance function is incorrect (Figure 4(a)). Yaglom (1962) used this criteria over and over again in exam-

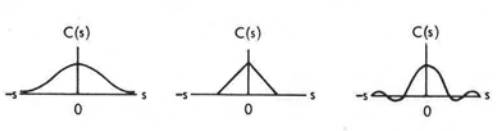

FIG. 2. Shapes of possible covariance kernels.

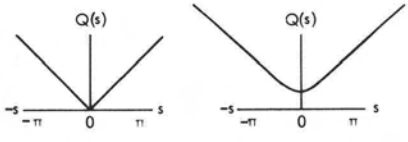

FIG. 3. Shapes of possible multiquadric kernels.

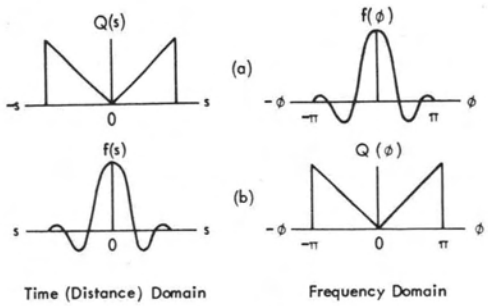

FIG. 4. Fourier transform relations for the bounded absolute value function.

ple problems to determine whether or not a specified analytical function qualified as a covariance function.

Thus there is no direct theoretical relationship of the quadric surfaces mentioned above to true covariance functions.

LEAST SQUARES PREDICTION WITH COVARIANCE FUNCTIONS

First it is assumed that a predicted $\widetilde{Z}_p$ anywhere is related linearly to measured or known values of Z at various positions in the region of consideration, i.e.,

$$\widetilde{Z}_p = \alpha_1 Z_1 + \alpha_2 Z_2 + \ldots \alpha_n Z_n = \sum_{i=1}^{n} \alpha_i Z_i \tag{3}$$

Any prediction method should be capable of predicting its own data ordinates exactly as in collocation polynomial theory (Scheid, 1968), or very closely in the more general sense of least squares. Thus, if a set of "predicted" $\widetilde{Z}_p$'s are exactly coincident with the Z_i's at each X position of the record, i.e., $X_p = X_i$, $i = 0,1,\ldots,16$, as in Figure 1(a) then we should have perfect correlation. For this case of zero "lag", i.e., distance $s_{pi} = 0$,
we have

$$C(s_{pi}) = C(0) = M\{Z_p \cdot Z_p\} = C_o \tag{4}$$

which is the mean value of the set of products Z_p^2 in this case, in fact the variance rather than covariance.

Now consider a set of Z_p's at regularly spaced intervals along the X axis, but not coincident with the Z_i's, i.e., they all have a "lag" or distance $s_{pi} \neq 0$. Then, in terms of covariance computations with ordinates Z_p and Z_i we have

$$C(s_{pi}) = M\{Z_p Z_i\} = C_{pi} \tag{5}$$

Also we may assign a pair of subscripts to the same sequence of data points in the same order. Then the covariance at a distance s_{ik} is

$$C(s_{ik}) = M\{Z_i Z_k\} = C_{ik} \tag{6}$$

Now we let the error of a predicted $\widetilde{Z}_p$ with respect to the true Z_p be expressed as

$$\epsilon_p = \widetilde{Z}_p - Z_p \quad .$$

Then from Equation 3, and squaring ϵ_p

$$\epsilon_p^2 = \left(\sum_{i=1}^{n} \alpha_i Z_i - Z_p\right)^2$$

$$= \left(\sum_{i=1}^{n} \alpha_i Z_i\right)^2 - 2Z_p \sum_{i=1}^{n} \alpha_i Z_i + Z_p^2 .$$

Rearranged, including the use of i and k subscripts to distinguish products at different positions on the X axis, we have

$$\epsilon_p^2 = \left(\sum_{i=1}^{n} \alpha_i Z_i\right) \left(\sum_{k=1}^{n} \alpha_k Z_k\right) - 2\sum_{i=1}^{n} \alpha_i Z_p Z_i + Z_p^2 .$$

Noting that

$$\left(\sum_{i=1}^{n} \alpha_i Z_i\right) \left(\sum_{k=1}^{n} \alpha_k Z_k\right) = \sum_{i=1}^{n} \sum_{k=1}^{n} \alpha_i \alpha_k Z_i Z_k$$

and substituting the symbols for covariance from Equations 4, 5, and 6 for the various ordinate products, we have

$$\epsilon_p^2 = \sum_{i=1}^{n} \sum_{k=1}^{n} \alpha_i \alpha_k C_{ik} - 2\sum_{i=1}^{n} \alpha_i C_{pi} + C_0 . \tag{7}$$

We can minimize the error of prediction by partially differentiating (ϵ_p^2) with respect to α_i and setting the new function equal to zero. Thus

$$\frac{\partial\left(\epsilon_p^2\right)}{\partial\, \alpha_i} = 2\sum_{k=1}^{n} \alpha_k C_{ik} - 2C_{pi} = 0 \quad i = 1,2, \ldots, n. \tag{8}$$

The result in Equation 8 may be viewed as a system of n linear equations with n unknowns

$$\sum_{k=1}^{n} \alpha_k C_{ik} = C_{pi} \qquad i = 1,2, \ldots, n . \tag{9}$$

Then the solution for the α_k's is

$$\alpha_k = \sum_{i=1}^{n} C_{ik}^{-1} C_{pi} \tag{10}$$

where C_{ik}^{-1} indicates the elements of an inverse of a matrix with elements C_{ik}. Substituting the now known α_k's into Equation (3) we have

$$\tilde{Z}_p = \sum_{k=1}^{n} \alpha_k Z_k = \sum_{i=1}^{n} \sum_{k=1}^{n} C_{ik}^{-1} C_{pi} Z_k \tag{11}$$

or in matrix notation

$$\tilde{\tilde{Z}}_p = \begin{bmatrix} C_{p1} C_{p2} \ldots . C_{pn} \end{bmatrix} \begin{bmatrix} C_{11} & C_{12} & \ldots . & C_{1n} \\ C_{21} & C_{22} & \ldots . & C_{2n} \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ C_{n1} & C_{n2} & \ldots . & C_{nn} \end{bmatrix}^{-1} \begin{bmatrix} Z_1 \\ Z_2 \\ . \\ . \\ . \\ . \\ Z_n \end{bmatrix} \tag{12}$$

In this form the covariance prediction method does not handle the general case of more equations than unknowns as indicated later for multiquadric prediction. Nevertheless Heiskanen and Moritz (1967) successfully applied covariance methods for determining the error covariance function and the standard error of least squares prediction. Moritz has subsequently improved and generalized this method in many ways, including filtering. However these details are not essential in a fundamental comparison involving pure prediction only. The most important point to note here is that the system of equations in Equation (9) corresponds to the normal equations in a least squares adjustment, thereby implying correctly that the more general case can be developed as in least squares adjustment theory.

Least Squares Prediction with Multiquadric Functions

Least squares prediction with MQ (multiquadric) functions is conceptually different from least squares prediction using covariance functions. Nevertheless, because the principle of least

squares is involved, the matrix manipulation can be made identical for both cases. This aspect of similarity probably accounts for the existing confusion about the covariance and multiquadric functions.

MQ kernels were based originally (Hardy, 1971) on exclusively geometric or physical considerations rather than stochastic processes. In fact the paper by Hardy (1972a) was virtually a polemic against what was considered to be an over-emphasis on statistical methods in some cases, without adequate consideration of geometry and physics in the "prediction" of topographic surfaces.

The basic hypothesis of multiquadric analysis is that any smooth mathematical surface, and also any smooth arbitrary surface (mathematically undefined), may be approximated to any desired degree of exactness by the summation of a wide variety of regular, mathematically defined surfaces, particularly quadric forms. The favoring of quadric forms came about through trial and error procedures with actual problems in topography, and more recently in geodesy (Hardy, 1974, 1975a). This reinforces an original conclusion that quadric forms were not only the simplest, but also the most efficient in converging on irregular surfaces.

It should be noted that the original approach to MQ analysis did not, and still does not exclude any kernel function from being considered as having an optimum role in some application. To make the innumerable possibilities more clear we now consider the concept of "any workable kernel". Let $W(X,Y,X_j,Y_j)$ be the workable kernels in a linear prediction function of the Cartesian form

$$f(X,Y) = Z = \alpha_1 W(X,Y,X_1,Y_1) + \alpha_2 W(X,Y,X_2,Y_2) + \ldots + \alpha_n W(X,Y,X_n,Y_n) = \sum_{j=1}^{n} \alpha_j W(X,Y,X_j,Y_j) \tag{13}$$

The basic requirement for a workable kernel is that with n or more data points (X_i,Y_i,Z_i), the system of equations formed from Equation (13) will lead to a non-singular coefficient matrix of the normal equations. As a result the solution for the coefficients is optimum in an ordinary least squares sense, regardless of the deterministic, stochastic, or other interpretation that can be placed on $f(X,Y)$ by other means.

From a purely geometric point of view the kernel functions in Equation 13 are interpretable as continuous single valued surfaces extending over the entire region of consideration. In a broad sense these kernels are not necessarily symmetric or even in the same class, e.g., quadric, cubic, or other. But for simplicity let us assume symmetric kernels. Then the coordinates X_j,Y_j represent the translation of the center of symmetry of each individual surface from the origin at $X,Y = 0,0$ to its respective X_j,Y_j, in general not 0,0. Thus $f(X,Y)$ in Equation 13 is represented by a series which applies the principle of superposition to an array of n separate geometric surfaces, each translated to different X_jY_j coordinates. By summation a single multiquadric, multicubic, or other combined surface is produced (Hardy, 1971).

If the workable surface functions in Equation 13 are not covariance functions, it is possible to interpret $f(X,Y)$ as representing at least one class of a broad field of non-stationary random functions. Yaglom (1962) indirectly permits this type of generalization. On his page 10 we see "There exist other methods of specifying a random function. Thus it is often convenient to define a random function by an analytic formula, containing parameters which are random variables. For example one may consider polynomials (ordinary or trigonometric) with random coefficients, We shall occasionally use this method of specifying a random function." Yaglom obviously recognized that in a general sense this method will produce non-stationary as well as stationary random functions. Since Yaglom was only interested in stationary random functions he invariably applied the Wiener-Khintchine theorem when he used this method, to determine if the random function was stationary.

Considering the broad implications above, we now substitute a quadric kernel function for the more general $W(X,Y,X_j,Y_j)$ in Equation 13. This will illustrate least squares prediction and error analysis with multiquadric functions in particular, using a quadric kernel that is known not to be a covariance kernel. Thus we let

$$W(X,Y,X_j,Y_j) = Q(X,Y,X_j,Y_j) = ((X - X_j)^2 + (Y - Y_j)^2 + \delta)^{1/2} \tag{14}$$

which is a hyperboloid. Arbitrarily letting $\delta = 0$ we can form a system of m linear equations with n unknown coefficients C_j, involving the cone as the kernel

$$\sum_{j=1}^{n} C_j ((X_k - X_j)^2 + (Y_k - Y_j)^2)^{1/2} = Z_k \quad k = 1,2, \ldots, m, \; m \geqslant n \tag{15}$$

Thus the problem involves m data points (or data and test points), of which any number n may be selected as nodal points for the functions centered at X_jY_j. Now we will go to matrix notation, using Q_{kj} as an abbreviation for the cones $((X_k - X_j)^2 + (Y_k - Y_j)^2)^{1/2}$.

Then the observation equations are

$$\begin{bmatrix} Q_{11} & Q_{12} & \cdots & Q_{1n} \\ Q_{21} & Q_{22} & \cdots & Q_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ Q_{m1} & Q_{m2} & \cdots & Q_{mn} \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ C_n \end{bmatrix} - \begin{bmatrix} Z_1 \\ Z_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ Z_m \end{bmatrix} = \begin{bmatrix} V_1 \\ V_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ V_m \end{bmatrix} \tag{16}$$

or more simply

$$\boldsymbol{QC} - \mathbf{Z} = \mathbf{V} \tag{17}$$

From this the unit weighted least squares solution for the column vector of coefficients is

$$\boldsymbol{C} = (\boldsymbol{Q}^T\boldsymbol{Q})^{-1} (\boldsymbol{Q}^T\mathbf{Z}) \tag{18}$$

With the known coefficients C_j, the summation formula for the prediction of a column list of $\widetilde{Z}_p$'s at the coordinates X_p,Y_p, is

$$\widetilde{Z}_p = \sum_{j=1}^{n} C_j((X_p - X_j)^2 + (Y_p - Y_j)^2)^{1/2} \quad p = 1,2, \ldots . \tag{19}$$

or in matrix notation, a column vector of predicted $\widetilde{Z}_p$'s is

$$[\widetilde{Z}_p] = [Q_{pj}][Q_{kj}{}^T\, Q_{kj}]^{-1} [Q_{kj}{}^T\, Z_k] \ . \tag{20}$$

For an analysis of the error of prediction we can determine $\Sigma\, \boldsymbol{V}^2$ with

$$\boldsymbol{V}^T\boldsymbol{V} = (\boldsymbol{Q}(\boldsymbol{Q}^T\boldsymbol{Q})^{-1} (\boldsymbol{Q}^T\mathbf{Z}) - \mathbf{Z})^T (\boldsymbol{Q}(\boldsymbol{Q}^T\boldsymbol{Q})^{-1} (\boldsymbol{Q}^T\mathbf{Z}) - \mathbf{Z}) \tag{21}$$

For a comparison with the covariance function development by Heiskanen and Moritz we take $n = m$ in Equation 15, then Equation 18 reduces to

$$\boldsymbol{C} = \boldsymbol{Q}^{-1}\, \mathbf{Z} \ . \tag{22}$$

Now if we predict only one $\widetilde{Z}_p$ instead of a vector as in Equation 20 we have

$$\widetilde{Z}_p = [Q_{pj}]\ [Q_{kj}]^{-1}\ [Z_k] \ . \tag{23}$$

Expanding from Equation 23 for comparison with Equation 12 we have

$$\widetilde{Z}_p = \begin{bmatrix} Q_{p1} & Q_{p2} & \cdots & Q_{pn} \end{bmatrix} \begin{bmatrix} Q_{11} & Q_{12} & \cdots & Q_{1n} \\ Q_{21} & Q_{22} & \cdots & Q_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ Q_{n1} & Q_{n2} & \cdots & Q_{nn} \end{bmatrix}^{-1} \begin{bmatrix} Z_1 \\ Z_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ Z_n \end{bmatrix} \tag{24}$$

Thus the solution for coefficients (unknowns) and the determination of the error of least squares prediction (adjustment) for any function, multiquadric, covariance or other "workable" kernel function can be found by the routine procedures of least squares adjustment. Consequently, least squares prediction is a matter of evaluating any analytical function whose coefficients can be solved for and adjusted by least squares. Optimality beyond that provided by least squares is a question involving the degree of correspondence between the intrinsic nature of a phenomenon, and the intrinsic nature of the kernel function used to replicate the phenomenon.

Similarities and Dissimilarities

The similarity of the covariance and multiquadric methods of analysis is represented almost entirely in the identity of Equation 12 and Equation 24, provided $C = Q$, i.e., a covariance

kernel is identical with a quadric kernel. Since this has been shown to be incorrect for specific cases, we are led to significant generalizations concerning the inequalities

$$Q \neq C, \text{ necessarily}$$

or even more generally that

$$W \neq C, \text{ necessarily}$$

where W is an arbitrary workable kernel. Careful consideration of the derivations of Equation 12 and Equation 24 show that covariance kernels are merely one of the many explicit workable kernels that are possible. A multiquadric kernel is another explicit form, with apparent advantages in some applications involving non-stationary or marginally stationary processes.

As will be suggested and verified to some extent in the section on applications, the competition among a variety of "workable" kernel functions (including covariance functions) that can be used in least squares prediction is likely to be decided by certain engineering considerations. Among these are economy of effort, efficiency, and the suitability of the results in actual application.

An important difference between the multiquadric and covariance function approach is that the choice of a covariance kernel is often (but not always) based on the computation of a discrete covariance sequence variously termed "empirical" or "apparent" covariance. After an apparent covariance is computed, an analytical function resembling the "apparent" covariance is usually chosen as the kernel. An exact fit of an "apparent" covariance with a single analytical kernel is seldom, if ever, possible. This, in itself, leads to a preliminary least squares problem. In other words an analytical covariance kernel function is frequently selected by means of a least squares fit to the apparent covariance. In the multiquadric approach this preliminary procedure is not relevant.

Can Some Multiquadric Kernels Be Spectral Density Functions?

The stimulation of the papers by Rauhala (1974), Assmus and Kraus (1974), and Schut (1974) led to a determination that the cone and hyperboloid are not covariance functions, according to any currently existing theory. On the other hand it may be highly significant that a large part of the mathematical formulation for the least squares prediction is identical in the two cases. This may be true of relevant Fourier and inverse Fourier transform relationships also. Therefore the author has speculated somewhat on how "bounded" multiquadric kernels and spectral density functions may be related. The interest in spectral density functions is due to the Wiener-Khintchine Fourier transform relation, already mentioned in connection with proof that the cone and hyperboloid cannot be covariance functions.

Since the spectral density is required to be non-negative by definition, the bounded absolute value function in Figure 4a fails to quality as a correlation function because its Fourier transform in the frequency domain is partially negative. On the other hand, if we view the bounded absolute value function as a spectral density function in the frequency domain, we appear to have no inconsistency (Figure 4b). The absolute value function is non-negative which enables it to satisfy the definition of spectral density. Geometrically the Fourier transform of the bounded absolute value function is the same as that in Figure 4a, but now it is in the time (distance) domain. Note that it resembles the sinc function in Figure 2(c) which is also a covariance function in the time (distance) domain. There is no restriction by definition or otherwise that the correlation of a centered arbitrary function is non-negative. Superficially at least, inconsistency has been exchanged for consistency by reversing the role of the Fourier transform pair in Figure 4a to that shown in Figure 4b. It is easy to show that this new consistency applies to the cone (generated by rotating the absolute value function around the Z axis), and for the same reason also applies to the hyperbola and hyperboloid.

It is also easy to show that the principle of least squares prediction with multiquadric functions is not affected by using an appropriate "bounding" of the otherwise indefinitely increasing multiquadric kernels (cone and hyperboloid). Assume that the region of availability of data and of interest for prediction is between $\pm \pi$ on the frequency axis ϕ in Figure 4b. Then let

$$Z = Q(\phi) = \sum_{j=1}^{n} C_j \left| \phi - \phi_j \right|, \text{ but } Z = 0, \text{ if } \left| \phi \right| \geq \left| \pm \pi \right|$$

Thus each kernel is 2π in width and each covers the entire region of interest. Consequently, there is no effect whatsoever on the least squares prediction in the region.

In conclusion, for this section, it appears that whereas an arbitrary function may be decomposed into covariance kernels for least squares prediction in the time (distance) domain, the same arbitrary function may possibly be decomposed into spectral density kernels for least squares prediction in the frequency domain. If so, the role of multiquadric functions for prediction purposes may be better understood in the future.

What has been said in this section is somewhat speculative, and studies in this area are continuing; however, it is hoped that others will also consider the possibility of some quadric kernels being spectral density functions, rather than considering them incorrectly to be covariance functions. Evidence that the cone and hyperboloid kernels are not covariance functions is overwhelming. Whether or not a satisfactory Fourier transform relationship relating multiquadric kernels to covariance kernels can be developed is not certain.

A Comparison of Covariance and Multiquadric Methods for Predicting Real Topography in Hawaii

A 1000 by 1000 foot square area with 240 feet of relief in the Haleiwa Quadrangle, Oahu Island, State of Hawaii was selected for a detailed comparative test. A topographic map of this area, greatly enlarged over the original 1:24,000 scale, is shown in Figure 5. Sample elevations Z and distances s in the area were used to compute the normalized apparent covariance, discretely as shown in Figure 6. Distances were scaled to unity for the length of the diagonal across the test area. Then these three analytical covariance kernels

$$C_1(s) = e^{-a^2 (s - s_j)^2}$$

$$C_2(s) = \sum_{k=0}^{3} b_k (s - s_j)^k$$

$$C_3(s) = \sum_{k=0}^{6} a_k (s - s_j)^k$$

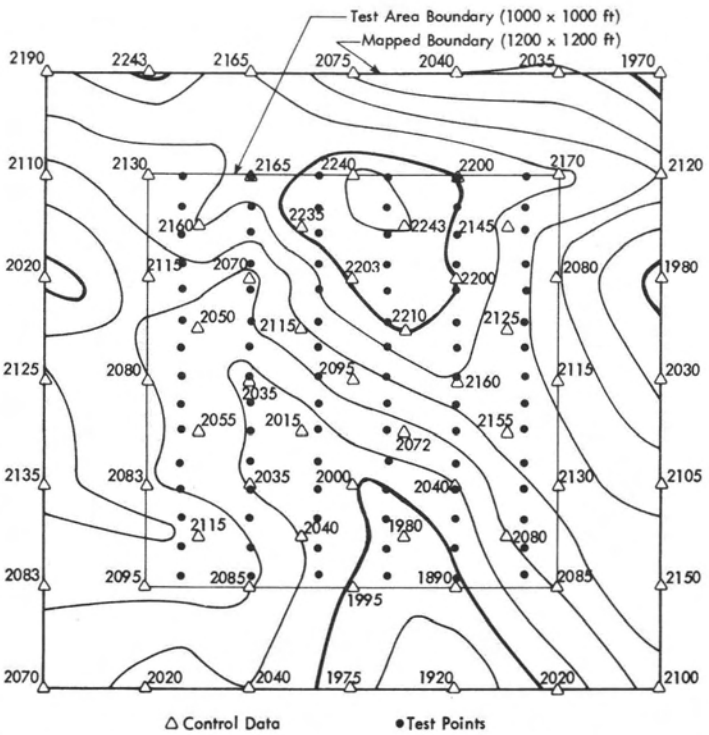

FIG. 5. A Portion of Haleiwa quadrangle, Oahu Island, State of Hawaii.

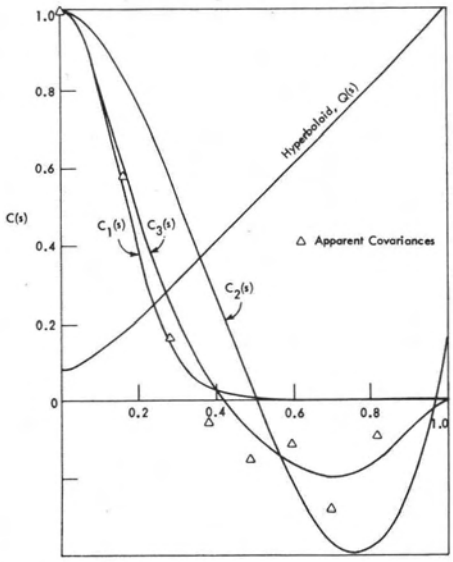

FIG. 6. Apparent covariances and selected covariance kernels, including an overlay of the selected multiquadric kernel.

were fitted (with $s_j = 0$) to the apparent covariance by least squares with constraints. The resulting curves are also shown in Figure 6. It is easy to recognize the analytical covariance functions as being the Gaussian distribution curve, a third degree polynomial, and sixth degree polynomial respectively.

The fourth curve in Figure 6 is a multiquadric kernel, specifically a hyperboloid

$$Q(s) = ((s - s_j)^2 + \delta_o)^{1/2}$$

There is obviously no relation of the hyperboloid kernel to the apparent covariance, and pre-fitting by least squares to correlation was irrelevant in its selection. A numerical value for δ_o was determined, using a relationship which indicates that an approximate optimum δ_o can be determined from

$$\delta_o = 0.665D^2 \tag{27}$$

where D is the grid spacing of the nodes. Using any positive δ, no matter how small, provides continuity of slope at data points which is not true of the absolute value function $|s| = (s^2)^{1/2}$. Sometimes a larger δ is desired for greater smoothing but if δ is too large some difficulties in "warping" may be experienced. Equation 27 provides a good working value for δ_o in all cases encountered to date.

The layout of map control data made available for the various approximations of the terrain in the test area were 41 elevations at specified horizontal locations on and inside the boundary, plus 24 elevations outside the area as shown by the spot elevations in Figure 5. Inside the test area no map control point was closer than 141 feet on the ground to another control point. Outside the test area no points were closer than 200 feet.

After a collocated fit of the multiquadric equations to the 65 map control points the multiquadric function was used to "predict" 90 elevations in a 6 by 15 grid array as shown also in Figure 5. The 90 predicted values were compared point by point with evaluations determined independently, by manual interpolation, in the same grid array. In other words, the original quadrangle map was used as a source of map control data for determining coefficients of the prediction function, and also as a source of data for testing the prediction function, the control and test sets being generally displaced relative to each other. A few points were nearly coincident by accident.

One hundred percent of the multiquadric predictions at the 90 test points were within one contour interval (40 feet), and 91 percent were within one half contour interval. This was without using the 0.02″ allowable shift at publication scale (1:24,000), or requiring the test

elevations to be at significant topographic or cultural features, both of which are permissible under National Map Accuracy Standards. Thus the technical standards with respect to contouring were more than satisfied. The mean difference in Z between the 90 predicted and the 90 manually interpolated values was +2 ft. The standard deviation was ±13 ft.

As shown in Table 1, the covariance predictions were compared with the multiquadric predictions and also with each other. The differences between the multiquadric predictions for the 90 test points and the cubic and sixth degree polynomial covariance kernels are negligibly small. On the other hand there are significant differences when the Gaussian prediction result is compared with the multiquadric, cubic, and sixth degree polynomial predictions. The Gaussian prediction should be regarded as quite marginal in this case for meeting National Map Accuracy Standards.

In summary one can say that the multiquadric function performed as well or better than two covariance functions and certainly better than the third covariance function for prediction purposes.

A Comparison of Covariance and Multiquadric Methods for Predicting Real Gravity Anomalies in Iowa

After the comparative test of covariance and multiquadric functions for predicting topography as described above, it was decided to perform a similar test in Iowa involving predicted gravity anomalies.

Seven free air gravity anomaly maps of Iowa were manually interpolated at a 10 milligal contour interval by six students and staff at Iowa State University, using 43 free air anomalies distributed throughout the state. (Anomalies outside the state boundary were not used, and the anomaly maps were not extrapolated to the state boundary.) One hundred non-data points were interpolated in each map, at identical locations in each map, and the mean was determined at each point. The standard deviation of a single point was found to range from ±1.2 to ±2.3 mgal. Later, based on a sample of 20 of the 100 points, the average standard deviation of a single point was found to be ±1.6 mgal. Figure 7 is an example of the manually

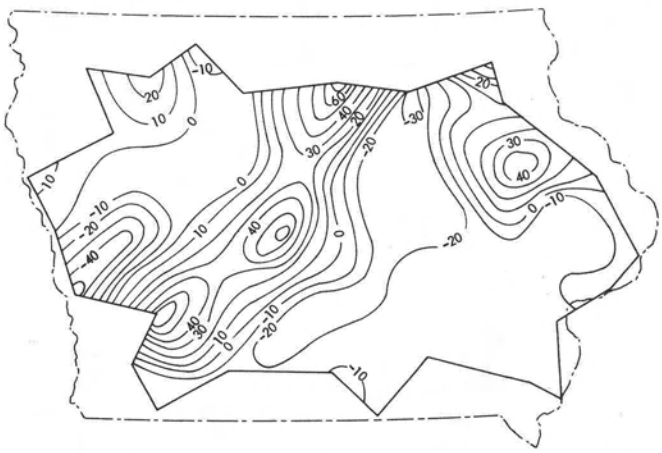

Fig. 7. Manually interpolated free air anomaly map of Iowa.

Table 1. Model Comparisons

Model Differences	Mean Differences in feet for 90 Test Points	Standard Deviation of a Single Difference
$Q(s) - C_2(s)$	0.2	±0.8
$Q(s) - C_3(s)$	0.1	±0.8
$C_2(s) - C_3(s)$	−0.1	±0.8
$Q(s) - C_1(s)$	−1.8	±6.6
$C_2(s) - C_1(s)$	−2.0	±7.0
$C_3(s) - C_1(s)$	−1.9	±6.8

interpolated maps. This one was found on the whole, to be the closest to the mean values of the 100 points. The mean values of these points were used later to judge the quality of predictions by the covariance and multiquadric methods.

Two covariance functions and two multiquadric functions were used for prediction purposes. A Gaussian and also a sixth degree polynomial were used as kernels in the covariance functions. A hyperboloid was used as the kernel in a non-harmonic form of the multiquadric function and the reciprocal distance was used as a kernel in a harmonic form of the multiquadric function. The graphical results of the prediction of the gravity anomalies is given for these four functions in Figure 8 through 11. These contour maps were determined by computer evaluation of the respective functions.

In each case the analytical predictions from the respective functions were compared with the mean manual interpolation of each of the 100 points. The maximum deviation, the mean deviation, and the standard deviation are given below for each function.

(a) Gaussian

Maximum deviation	+29.8 mgal
Mean deviation	+ 1.42 mgal
Standard deviation	±11.13 mgal

(b) Sixth Degree Polynomial

Maximum deviation	+29.6 mgal
Mean deviation	+ 2.88 mgal
Standard deviation	±10.46 mgal

(c) Non-harmonic Multiquadric (hyperboloid)

Maximum deviation	+ 7.60 mgal
Mean deviation	+ 0.11 mgal
Standard deviation	±2.78 mgal

(d) Multiquadric Harmonic (reciprocal distance)

Maximum deviation	+ 8.80 mgal
Mean deviation	− 0.26 mgal
Standard deviation	± 3.31 mgal

As to computational efficiency, the non-harmonic multiquadric requires no precomputation except for optimum δ_0 which is very easy. It is the most efficient of the four functions. The multiquadric harmonic requires an *a priori* determination of the best radius of point masses with the best r-formula (Hardy and Göpfert, 1975) but this is easier than the *a priori* determination of the empirical covariance and least squares fit of an analytical kernel function in the covariance method. The evaluation or prediction aspect, after solving the function, is also computationally easier with the multiquadric kernels than with the covariance kernels.

Applications of Multiquadric Functions in Photogrammetry and Remote Sensing

The applications of multiquadric functions to photogrammetry and remote sensing certainly encompass all the possibilities that have already been mentioned in the literature for covariance functions, and possibly for other computational methods as well. Among these possibilities are

(1) lens distortion corrections
(2) film deformation corrections
(3) prediction of corrections to pass points in strip and block triangulation
(4) digital terrain model (DTM) contouring and profiling
(5) camera and reseau calibration
(6) geometric correction of radar imagery
(7) geometric correction of panoramic camera imagery
(8) image function processing and analysis

Multiquadric equations have been applied to a limited extent at Iowa State for items (1), (3), (4), and (8). At the present time item (8) is receiving the most attention, and consequently will be the particular application presented here.

Image processing and analysis, has in itself, the potential for a breakdown into many sub-applications. Among these are

(1) pattern recognition
(2) boundary location (or more generally, gradient analysis)
(3) density slicing, color coding, and B & W to color conversion

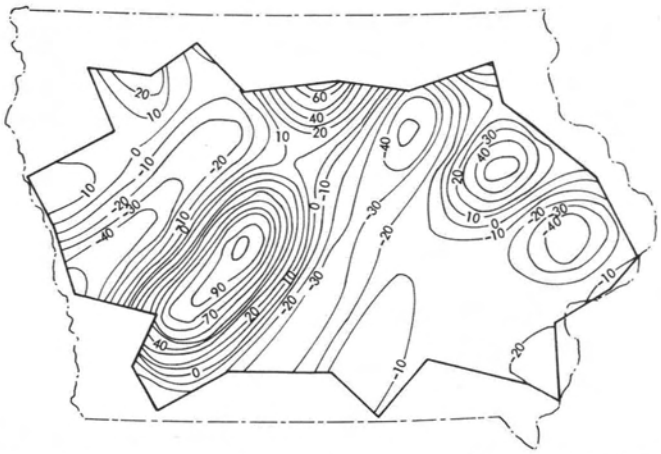

FIG. 8. Free air anomaly map of Iowa using a Gaussian kernel in a covariance function.

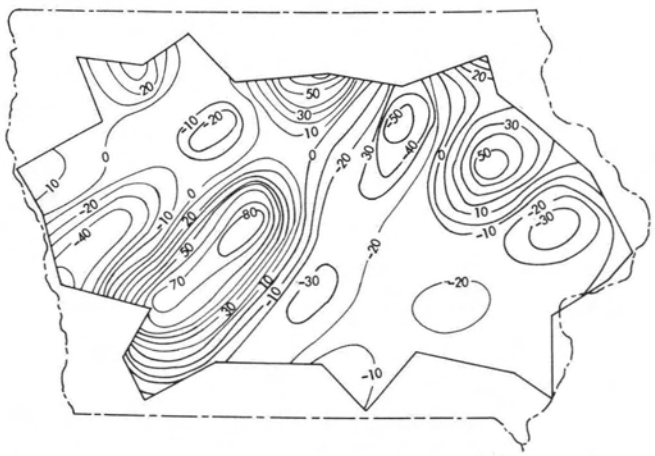

FIG. 9. Free air anomaly map of Iowa using a sixth degree polynomial as the kernel in a covariance function.

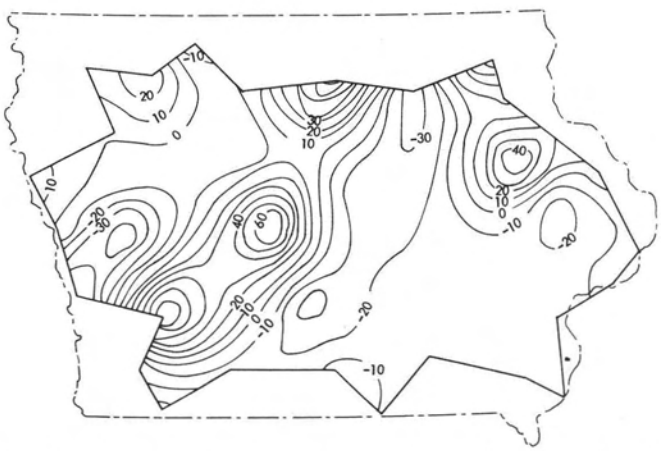

FIG. 10. Free air anomaly map of Iowa using a hyperboloid as the kernel in a multiquadric function.

FIG. 11. Free air anomaly map of Iowa using the reciprocal distance as a kernel in a multiquadric harmonic function (depth of 43 point masses, 30 km).

(4) image function compression
(5) image function reconstruction (expansion)
(6) correlation of perspective image functions
(7) orthoprojection of perspective image functions

Of these, experiments with items (4) and (5) will be reported in this paper. In a sense, items (4) and (5) involve an inverse or transform relationship with respect to each other. It is possible to describe an application and the corresponding process by starting with either (4) and (5) as the problem and ending up with (5) or (4) respectively as being essential to the solution. This is probably why Fourier, Hadamard, Walsh, and other transforms have received so much attention in communications engineering for the efficient transmission and reconstruction of images. It is possible for multiquadric equations to accomplish somewhat the same result in a conceptually more direct and more understandable fashion. However, it remains to be seen whether multiquadric analysis could be more computationally efficient than fast Fourier transform methods, for example. Meanwhile, the fundamentals of the multiquadric method will be presented.

The picture elements in Figure 12(b) may be viewed as containing the coarsely sampled and quantized information from Lincoln's portrait in Figure 12(a). Each pixel contains a uniform distribution of intensity (gray level) within its respective boundaries. It is more or less obvious that the gray level in each pixel of 12(b) is the mean value of the variable gray level in a corresponding, but unbounded "block area" of the original picture 12(a). From a close inspection point of view the original picture appears to have been degraded in Figure 12(b) to the point of extinction by this process. On the other hand, if both pictures are viewed from a distance of fifteen to twenty feet without magnification they are both recognizable and mutually indistinguishable. Another analog solution for restoring a substantial part of the "lost information" is to project a slide of Figure 12 (b) with deliberate defocusing. Generally there are analytical solutions which correspond to analog solutions and this case is no exception. One of the solutions involves Fourier transform methods as mentioned above. Figures 12(a) and 12(b) have been reprinted from the cover of Science, 15 June 1973 with the permission of the magazine (copyright 1973 American Association for the Advancement of Science), and of authors Harmon and Julesz (1973). As indicated in their article, Figure 12(a) is a reconstruction of a high resolution Lincoln photograph (not shown) by Fourier processing of the coarsely sampled and quantized information in Figure 12(b). The reader is referred to their article and references for a more detailed understanding of the Fourier processing. What we have done recently at Iowa State is to partially repeat this experiment, using multiquadric functions.

First we may view a photograph as being a picture function (Rosenfeld, 1969) or, more generally for remote sensing, an image function. The gray levels of a black and white photograph correspond to Z ordinates of a function $Z = f(X,Y)$ in which X and Y are plane coordinates for locating the Z information. For a picture quantized in pixels it is easily understandable that we can represent an image function in a general sense with the multiquadric function

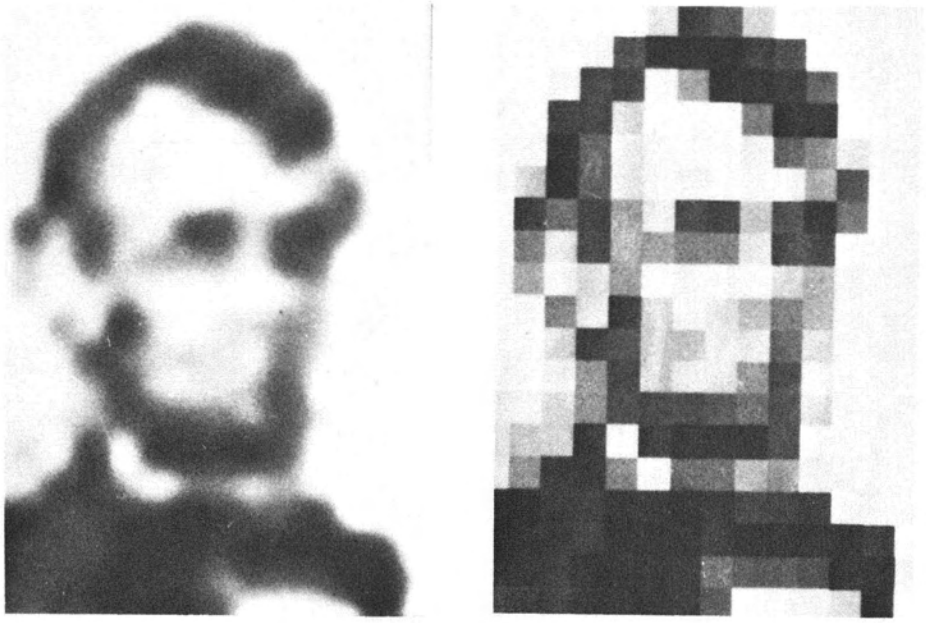

FIG. 12. Two images of Lincoln. (Reprinted by permission from the cover of the 15 June, 1973 issue of *Science* magazine. Copyright 1973 by the American Association for the Advancement of Science.)

$$Z = \sum_{j=1}^{n} C_j \, Q(X,Y,X_j,Y_j) \tag{28}$$

Now let us represent a particular image (say that in Figure 13 which is very similar to Figure 12(b) using the cone as a quadric kernel and a linear system of equations as follows

$$Z_i = \sum_{j=1}^{n} C_j \, ((X_i - X_j)^2 + (Y_i - Y_j)^2)^{1/2} \quad i = 1,2, \ldots, n \tag{29}$$

If we solve the coefficients for the whole image simultaneously, then $n = 266$, since the pixel array is 19×14. For large images this could become a formidable computational problem, but fortunately there are at least two approaches to computational simplification. One method is to use the one-time matrix inversion property of multiquadric functions for fixed or formatted arrays of input data. The inverse of the coefficient matrix of the normal equations is always the same in such cases since it is unaffected by a new vector of Z ordinates. Thus the reconstruction of any number of images in a specified X,Y format involves matrix multiplication only, using a once-inverted and stored inverse. Another feature of multiquadric analysis, could be called the "osculating sub-picture principle". In other words we can choose to reconstruct a subpicture of the total picture with a small matrix inverse and repeat the process until the entire picture is reconstructed. Moreover the subpictures can be forced to fit each other at each discrete element on the boundaries, using an osculating mode (Hardy, 1971; 1975). Details will not be given here.

A reconstruction of the coarse picture in Figure 13 is given in Figure 14. It is far from a "perfect reconstruction" because it involved only a 4 by 4 subpicture matrix covering the same area as a single pixel in the coarse picture, and I did not use the osculating principle to force fits at the edge of the subpicture. The multiquadric equation of the subpicture was then evaluated in 16 subpixels, thereby replacing an original large pixel with a set of 4 by 4 smaller ones. The boundaries of some of the 266 subpictures are visible in the reconstruction, particularly in the light areas. However the result is striking, considering that the entire computation was done on a Wang 600-14 programmable calculator, and the plot with an on-line Wang 602 typewriter plotter.

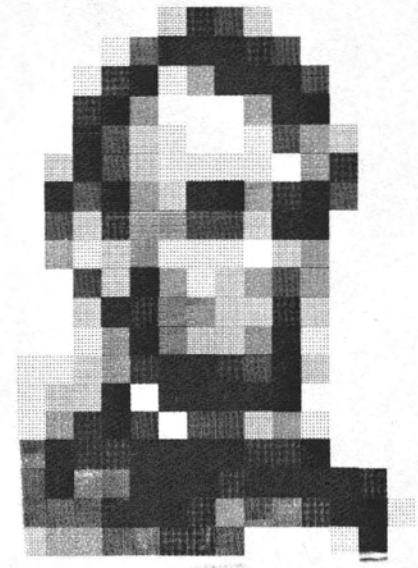

FIG. 13. Another coarse sample of a Lincoln image.

FIG. 14. A multiquadric reconstruction of the Lincoln image with a one-time 4 × 4 coefficient matrix inversion.

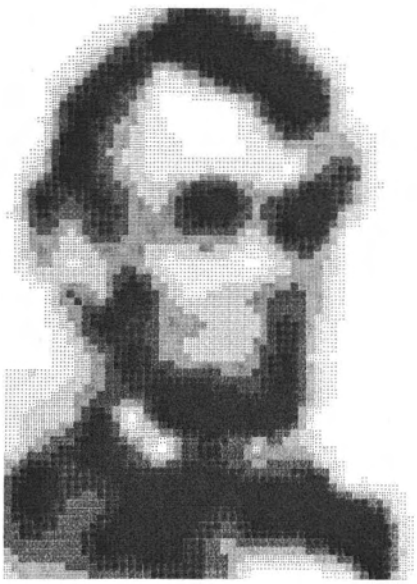

FIG. 15. A multiquadric reconstruction of the Lincoln image with a one-time 16 × 16 coefficient matrix inversion.

FIG. 16. Photograph of off-line TV reconstruction of the Lincoln image with a one-time 4 × 4 coefficient matrix inversion.

By contrast the reconstruction in Figure 12(a) was accomplished with a Facsimile transmitter/receiver in conjunction with a Honeywell DDP-224 computer for photographic-digital conversion, plus a Honeywell 6078 computer using a conventional two dimensional fast-Fourier-transform program to process the digitized information.

Another reconstruction of the coarse picture in Figure 13 is shown in Figure 15. It is also a Wang 600-14 Calculator—Wang 602 Plotter output. In this case, a one time inversion of a 16 by

16 subpicture matrix covering the same area as nine pixels in the coarse picture (Figure 13) was used. The evaluation or "prediction" mode was the same as for Figure 14. Each original pixel was evaluated in 16 sub-pixels. In this case the boundaries of the 266 coarse pixels have virtually disappeared, even though the more rigorous "osculating mode" was not applied.

The improved appearance of the reconstruction of the coarse picture in Figure 13 as shown in Figure 16, using the basic Wang generated "predictions", was accomplished at the Remote Sensing Institute of South Dakota State University. Digital tape data involving a "stretch" from the original 8 gray levels of Figure 14 to 32 gray levels was input to a video monitor and photographed. Although the overall appearance is improved, this procedure introduced geometric distortions and edge effects that were not in the data.

This experiment has suggested a new application, an extension of the image function reconstruction concept. The output of any data collection system involving pixels can be enlarged until the information content "appears" to be virtually destroyed as in Figures 12(b) and 13. Some computational processes, in this case multiquadric equations, are capable of restoring relative continuity to such an enlarged step function in a logical manner, thus bringing out "predicted" details of continuity that were never directly recorded in the original image.

Acknowledgments

This research has been sponsored by the Engineering Research Institute, Iowa State University. The studies involving gravity anomalies, as reported herein, have been supported in part by NSF Grant GK 40287.

References

1. Assmus, E., and K. Kraus, 1974, *Die Interpolation nach Kleinsten Quadraten Prädiktionswerte simulierte Beispiele und ihre Genauigkeiten*, Reihe A, Heft Nr 76, Deutsche Geodätische Kommission, Muenchen.
2. Blackman, R. B., and J. W. Tukey, 1959, *The Measurement of Power Spectra*, Dover Publications, Inc., New York.
3. Brown, Duane C., 1973, *Investigation of the Feasibility of a Short Arc Reduction of Satellite Altimetry for Determination of the Oceanic Geoid.* AFCRL Report No. 73-0520. Bedford, Massachusetts: Air Force Cambridge Research Laboratories.
4. Hardy, R. L., 1971, "Multiquadric equations of topography and other irregular surfaces," *Journal of Geophysical Research*, 76:8, pp. 1905-1915.
5. Hardy, R. L., 1972a, "The analytical geometry of topographic surfaces," *Proceedings*, American Congress on Surveying and Mapping, 32, pp. 163-181.
6. Hardy, R. L., 1972b, "Analytical topographic surfaces by spatial intersection," *Photogrammetric Engineering*, 38:5, pp. 452-458.
7. Hardy, R. L., 1974, "Geodetic application of multiquadric harmonic equations," Abstract, *EOS*, Transactions, American Geophysical Union, 55:7, p. 673.
8. Hardy, R. L., and N. M. Gopfert, 1975a, "Least squares prediction of gravity anomalies, geoidal undulations, and deflections of the vertical with multiquadric harmonic equations," *Geophysical Research Letters*, 2:10, pp. 423-426.
9. Hardy, R. L., 1975b, "Research results in the application of multiquadric equations to surveying and mapping problems," *Surveying and Mapping*, 35:4, pp. 321-332.
10. Harmon, L. D., and B. Julesz, 1973, "Masking in visual recognition: effects of two-dimensional filtered noise," *Science*, 180:4091, pp. 1194-1196.
11. Heiskanen, W. A., and H. Moritz, 1967, *Physical Geodesy*, W. H. Freeman, San Francisco, Calif.
12. Kraus, K., 1972, "Film deformation correction with least squares interpolation," *Photogrammetric Engineering*, 38:5, pp. 487-493.
13. Kraus, K., and E. M. Mikhail, 1972, "Linear least squares interpolation," *Photogrammetric Engineering*, 38:10, pp. 1016-1029.
14. Rauhala, V. A., 1974, *Array Algebra with Applications in Photogrammetry and Geodesy* (Doctoral Dissertation, Royal Institute of Technology, Stockholm), Fotogrammetriska Meddelanden, 6.
15. Rosenfeld, A., 1969, *Picture Processing by Computer*, Academic Press, New York.
16. Scheid, F., 1968, *Theory and Problems of Numerical Analysis*, McGraw Hill, New York.
17. Schut, G. H., 1974, "Two interpolation methods," *Photogrammetric Engineering*, 40:12, pp. 1447-1453.
18. Shaw, Elizabeth M. and P. P. Lynn, 1972, "Areal rainfall evaluation using two surface fitting techniques," *Bull. Int. Assoc. Hydrological Sc.*, 27, 419-433.

19. Trotter, Jerry E., 1975, *Skylab Altimetry Data Reductions and Development of Programs for Pre and Post Processing of Altimetry Data for Sarra Reductions.* AFCRL-TR-75-0359. Bedford, Massachussetts: Air Force Cambridge Research Laboratories.
20. Williamson, M. R. and E. M. Gaposchkin, 1973, *Estimate of Gravity Anomalies, 1973 Smithsonian Standard Earth III,* SAO Report 353, Smithsonian Astrophysical Observatory, Boston.
21. Yaglom, A. M., 1962, *An Introduction to the Theory of Stationary Random Functions,* Prentice Hall, Inc., Englewood Cliffs, New Jersey.

Ninth Surveying Teachers' Conference
University of New Brunswick, June 19-23, 1977

The Department of Surveying Engineering at the University of New Brunswick, Fredericton, Canada will be hosting the ninth Surveying Teachers' Conference from June 19th through 23rd, 1977. Surveying teachers' conferences are held every three years and provide a forum for the exchange of ideas about the nature and function of surveying education. It is expected that more than 150 surveying teachers from universities and colleges throughout Canada, the United States, and other countries will participate in this year's meeting. The theme of the conference is "resources for more effective teaching." Further information may be obtained from Professor Angus Hamilton at U.N.B.