Commission III
Presented Paper

SPECTRAL ANALYSIS OF INTERPOLATION

Dr. V. Kratky
Photogrammetric Research
National Research Council of Canada
Ottawa, Ontario, Canada K1A OR6

ABSTRACT

Basic concepts of spectral analysis are applied to interpret interpolation, smoothing and parametric transformation based on uniform sampling, as different types of discrete convolution. Spectral properties of interpolation are then discussed with main emphasis on its linear least squares version. This method is shown to be a close approximation to the interpolation with a low-pass filtering effect. An operator derived in a functional form performs direct least squares interpolation avoiding the usual inversion of the covariance matrix. Finally, advantages of spectral analysis are demonstrated by examining the accuracy and stability of the procedure.

INTRODUCTION

With the current expansion of digital terrain modelling and automated mapping it becomes increasingly important that the methods used in the associated data compression, smoothing, resampling and transformation be as efficient as possible. The time consumption and reliability are of primary interest, but one should also ensure that the operations really perform their expected functions with no significant loss of information and with a minimum distortion of data quality. In order to understand the meaning of some functions it is useful to analyze their spectral properties.

In general, spectral analysis can be applied both to continuous functions represented, e.g., by analog signals, and to sampled functions given by a recorded stream of digital data collected in discrete time or space intervals. This form of recording is typical of most photogrammetric applications. Geometric information is collected by a regular one-dimensional sampling, such as an automated coordinate recording of profiles, or by sampling in regular two-dimensional rasters typical for more advanced photogrammetric systems, such as the Gestalt Photo Mapper or analytical plotters.

This present paper is an attempt to show spectral properties of elementary interpolations and of the method of linear least squares interpolation (Kraus, 1972, Moritz, 1973). The scope of the analysis is limited to the version preserving a fit in regularly distributed data points, and the noise filtering is disregarded. The findings are applicable also to more general configurations as long as an average point spacing is ensured.

BASIC CONCEPTS

A few basic concepts from the theory of Fourier transformation are needed for the analysis and will be briefly reviewed here, with reference to

details available in any of numerous suitable textbooks on the subject, e.g. in Bracewell (1965).

In the context of digital terrain modelling a function $f(x)$ representing a spatial distribution of elevations by positions x, corresponds to a function $F(s)$ representing amplitudes of related spatial frequencies s. The relationship $f(x) \leftrightarrow F(s)$ is defined in both directions by a pair of *Fourier transforms*.

Convolution of two continuous functions is defined by $h = c*f$ or

$$h(x) = c(x)*f(x) = \int_{-\infty}^{\infty} c(x-u)f(u) \, du \quad , \tag{1}$$

where $f(x)$ is a primary function, $c(x)$ is a weighting or convolving function and $h(x)$ is a resulting function. Accordingly, each value $h(x)$ is an integral functional of all values from $c(u)$ and $f(u)$. Obviously, $c(x)$ as a weighting function has a smoothing effect on $f(x)$, which is proportional to the width of $c(x)$.

Convolution theorem states that a convolution of two functions in one domain corresponds to a simple product of their transforms in the other domain, i.e.

$$h = c*f \leftrightarrow H = CF \quad \text{or} \quad h = cf \leftrightarrow H = C*F \quad . \tag{2}$$

The theorem has a great practical importance, because it enables one to avoid difficulties in certain computations or derivations by choosing a possibly easier solution of equivalent relations in the other domain, provided that the Fourier transforms exist. In terms of symbols, if it holds

$$a \leftrightarrow A \quad \text{and} \quad b \leftrightarrow B \quad ,$$

then an indirect, but technically simpler path

$$a \rightarrow A \rightarrow B \rightarrow b \tag{3}$$

is a substitute for a direct path $a \rightarrow b$.

If $f(x)$ is a band-limited function such that its transform $F(s)$ is zero for all absolute frequencies exceeding a limit of $s = 0.5$, the Fourier transform of function f sampled at unit intervals ($\text{III}f$) is expressed by

$$\text{III}f \leftrightarrow \text{III}*F \quad . \tag{4}$$

Sampling function III is basically a set of impulses δ spaced at unit intervals and convolution $\text{III}*F$ represents function $F(s)$ which is regularly replicated at unit intervals. Consequently, *replication is a Fourier equivalent to sampling*.

If function $f(x)$ is not band-limited or the above bandwidth is exceeded the replication results in overlaps of individual transforms F and in false summations of their higher-frequency tails. This effect is known as *aliasing* which distorts the spectral characteristic of the original function.

DISCRETE CONVOLUTION

The convolution of a sampled function yields a resulting function h and its spectrum H which are continuous

$$h = c*\text{III}f \leftrightarrow H = C(\text{III}*F) \quad . \tag{5}$$

The analysis of the effect of convolution in the raster of resampled points
$\text{III}h = \text{III}(c * \text{III}f)$, as presented in the Appendix, leads to important relations

$$\text{III}h = (\text{III}c) * (\text{III}f) \leftrightarrow \text{III} * H = (\text{III} * C)(\text{III} * F) \quad . \tag{6}$$

A resampling process can be described by convolution of two sampled
functions. The spectrum of a resampled function $\text{III}h$ is a product of repli-
cated spectra for functions c and f. This convolution deals only with dis-
crete values at given rasters in all three functions involved. Thus, we may
term it *discrete convolution*. When considering all three functions as
finite, the same set of numerical operations in the spatial domain can also
be expressed in matrix notation by

$$h = C \, f \tag{7}$$

with a graphical representation given in Fig. 1. Individual row vectors in
the band-limited matrix C are identical except for their mutual shift. The
width of the matrix is limited by the given length of vector f. The resul-
ting vector h stretches out beyond the length of f, thus adding non-zero
tails to both its ends which taper off rather rapidly. This fact reflects
the well-known boundary or edge effect typical for finite interpolations.
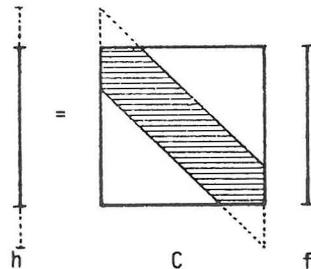
Fig. 1 Matrix representation of discrete convolution

INTERPOLATION

Convolution performs as a general interpolation if it yields a fit
$\text{III}h = \text{III}f$ expressed by

$$\text{III}f = (\text{III}c) * (\text{III}f) \leftrightarrow \text{III} * F = (\text{III} * C)(\text{III} * F)$$

from which one can derive conditions

$$\text{III}c = \delta \leftrightarrow \text{III} * C = 1 \tag{8}$$

to specify general interpolation by considering the well-known property of
the δ-function $(\delta \leftrightarrow 1)$. Any function with unity value in its origin and
zeros for any integer x, acts as a general interpolator regardless of its
form within the integer intervals. Any other convolving function c intro-
duces a smoothing effect and changes values also in the given raster of
sampled points.

Interpolation which provides full agreement, h = f, exists only if
function f is band-limited. Considering unit sampling interval the frequen-
cies in function f should not exceed the limit of $|s| = 0.5$. In other words,
the replicated spectrum F must be free of aliasing. The central unit sec-
tion of the replicated spectrum is then identical with the original spectrum
defined in the range s(-0.5 , 0.5). Since it follows from Equation (5)

$$f = c * \text{III}f \leftrightarrow F = C(\text{III} * F)$$

the only function C which preserves values of F up to $|s| = 0.5$ and sup-
presses everything beyond, is the rectangle function $\text{II}(s)$. Its equivalent
in the spatial domain is the sinc-function

$$c(x) = \text{sinc}(x) = \sin(\pi x)/\pi x \leftrightarrow C(s) = \Pi(s) \quad . \tag{9}$$

LEAST SQUARES INTERPOLATION

The least squares interpolation (LSI) represents a linear prediction with minimum variance. Its basic formulation in matrix notation is very simple

$$h = c^T Q^{-1} l. \tag{10}$$

Here, C is a covariance vector, Q is a covariance matrix and l is a vector of observations. In this analysis the filtering capability of LSI is disregarded. Covariances are computed from a Gaussian function which is either determined from an autocorrelation analysis of given data, or a priori prescribed. Equation (10) can be interpreted in two different ways and accordingly treated as a
  • parametric transformation

$$h = c^T g \quad , \quad g = Q^{-1} l \quad , \tag{11}$$

  where g is a vector of auxiliary parameters computed from a system of linear equations, or as a
  • linear estimation or direct interpolation

$$h = a^T l \quad , \quad a^T = c^T Q^- \tag{12}$$

  where a is an interpolator.

Values of vectors l and g are referred to the same positions and are functionally related. Both forms of LSI are equivalent and yield identical results.

Spectral Characteristic of LSI

The parametric form (11) of LSI represents a convolution of deconvolved observations. In this paper, the same covariance function is supposed to be used in both phases. The direct interpolation form (12) operates with original observations and the interpolator a must be numerically computed for each individual point from the inverse of the covariance matrix. The analytical form of the function a(x) is unknown. We will show here that a suitable spectral analysis helps to approximate this function. If this is successful the practical interpolation may proceed as direct convolution of sampled observations with a derived effective interpolator a(x), thus avoiding the need to invert the covariance matrix Q.

In accordance with Equations (6) and (7) the system of linear equations to be solved for parameters g

$$Qg = l \tag{13a}$$

has its equivalent form of a discrete colvolution which can be monitored in both the spatial and spectral domains

$$\text{III}c * \text{III}g = \text{III}l \leftrightarrow (\text{III}*C)(\text{III}*G) = \text{III}*L \quad . \tag{13b}$$

The process of solution for vector g from Equation (13a) has equivalent operations in Equation (13b) and these are extremely simple in the frequency domain

$$\text{III}*G = (\text{III}*L)/(\text{III}*C) \quad . \tag{14}$$

A complicated solution of linear equations is represented by a simple division of two replicated spectra.

Suppose that we have enough information on the spectral meaning of parameters g. The use of g or l in computing a new value

$$h = c^T g = a^T l \tag{15a}$$

has their convolution equivalents in functions

$$h = c * \mathrm{III}g = a * \mathrm{III}\ell \leftrightarrow H = C(\mathrm{III} * G) = A(\mathrm{III} * L) \tag{15b}$$

and with substitution for $\mathrm{III} * G$ from Equation (14) we obtain

$$H = C(\mathrm{III} * L)/(\mathrm{III} * C) \quad .$$

Hence we derive

$$\boxed{A = C/(\mathrm{III} * C)} \quad , \tag{16}$$

which is the spectral characteristic or transfer function to the yet unknown effective interpolation function a(x). It is evident that A meets the criterion from Equation (8) to test the interpolation capability, since $\mathrm{III} * A = (\mathrm{III} * C)/(\mathrm{III} * C) = 1$.
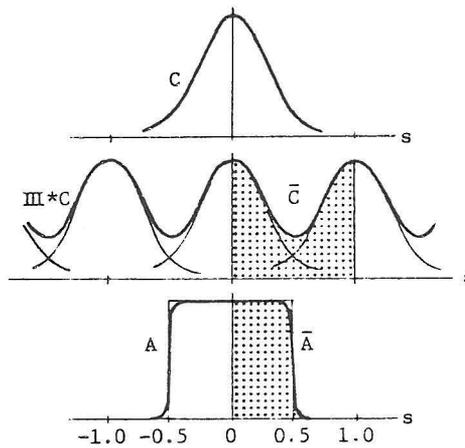
Fig. 2 Spectral characteristic of LSI

The form of the transfer function A depends on the definition of the Gaussian covariance function c and of its transform C

$$c(x) = \exp(-dx^2)/w \leftrightarrow C(s) = \exp(-Ds^2) \quad , \tag{17}$$

where $dD = \pi^2$ and $w^2 = D/\pi = \pi/D$. Figure 2 gives a graphical representation of C, of its replication $\mathrm{III} * C$ and of the resulting transfer function A. Assuming that the value of C is sufficiently small at s = 1, the aliasing effect can be approximated from the summation of two shifted Gaussians in the range of s(0,1). The corresponding part of the replicated spectrum is denoted as $\overline{C}$ and analytically expressed by

$$\overline{C} = \exp(-Ds^2) + \exp(-D(s-1)^2) =$$
$$= \exp(-Ds^2)(1 + \exp(D(2s-1))) \quad .$$

In accordance with Equation (16) we define A in the limited range of s by

$$\overline{A} = C/\overline{C}$$

and with substitutions for C and $\overline{C}$ we arrive at

$$\overline{A}(s) = (1 + \exp(D(2s-1)))^{-1} \quad . \tag{18}$$

A quick analysis yields values $\overline{A}(0) \approx 1$, $\overline{A}(0.5) = 0.5$ and $\overline{A}(1) \approx 0$ which together with Fig. 2 show that the transfer function A is numerically very close to the rectangular function Π which is the spectral characteristic of the true interpolation, if the covariance function c is fairly wide and its transform

is then narrow with respect to the unit interval.

The least squares interpolation yields results which are very close to or practically identical with those from the true interpolation, if $D \gg \pi$ or $d \ll \pi$.

Accuracy Limits of LSI

Although there is a certain freedom in choosing the form of the covariance function c without affecting the interpolation results in an appreciable way, one should be careful not to exceed limits beyond which the interpolation becomes either less accurate or reliable. Common sense would prevent one from choosing function c too narrow with respect to the sampling interval, because this would generate artificial depressions between given points. Experience with LSI would also warn against extending the spread of the c-function too far, since in this instance the solution suffers from numerical ill-conditioning of the covariance matrix to be inverted, even though it was reported by Lauer (1976) and Schut (1976) that up to a certain degree a loss of numerical stability was accompanied by a gain in interpolation accuracy. Spectral analysis can provide explanations to this phenomenon.

A variable spread of the covariance function is reflected in the frequency domain by different levels of spectral aliasing. A narrow spread of $c(x)$ causes high aliasing, for $C(s)$ is then too wide. For a narrow $c(x)$ the convolution form $h = c*\text{III}g$ is obviously limited to a serial product containing a low number of terms and, consequently is not as accurate as a convolution of continuous functions expressed by integrals. The error can be estimated from the difference between the integral and the serial product. If c as a weighting function is area-normalized then it holds true

$$\int_{-\infty}^{\infty} c(x)\,dx = 1 \qquad \text{and} \qquad C(0) = 1 \quad ,$$

because the area of a function is always equal to the value of its transform at $s = 0$. A corresponding serial product for $\text{III}c$ is distorted by a value ε which represents the relative inaccuracy of the serial product according to

$$\sum_n c(n) = 1 + \varepsilon \qquad \text{and} \qquad \overline{C}(0) = 1 + \varepsilon \ .$$

If $C(s) = \exp(-Ds^2)$ as defined by Equation (17), then with reference to Fig. 3

$$\varepsilon = 2C(1) = 2\exp(-D) \quad .$$

To ensure an n-digit accuracy we set a condition

$$\varepsilon \leqq 0.5 \times 10^{-n}$$

and hence

$$D > n\ln 10 + \ln 4 \simeq 2.3n + 1.4 \qquad \text{or} \qquad d < 9.87/D \quad .$$

The following table gives a few examples for values D and d meeting the requirement of the n-digit accuracy of interpolation:

n =	2	3	8
D >	6	8	20
d <	1.6	1.2	0.5
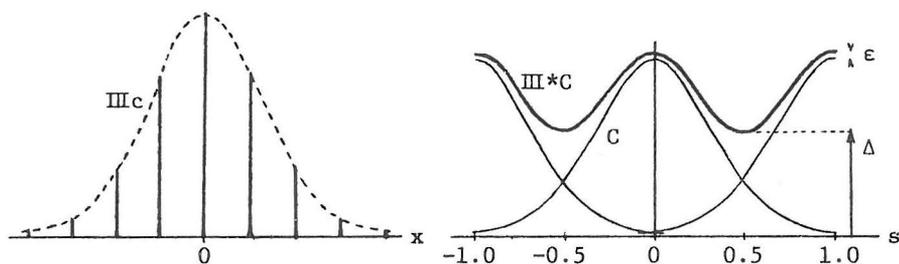
Fig. 3 Sampled covariance function and its aliased spectrum

The problem of numerical stability of the covariance matrix is reflected in the spectral domain by Equation (14) which contains a critical division by III*C or, within the range of s(0,1) by $\overline{C}$. With reference to Fig. 3 and to Equation (17) the numerical values in $1/\overline{C}$ range between 1 and $1/\Delta = 2\exp(D/4)$. The corresponding range of decimal exponents extends from 0 to $\log(1/\Delta) \simeq D/9.2 = 1.1/d$. Because of the inherent reversal of qualities in a function and its Fourier equivalent, a tall and narrow function has a flat and wide transform. An excessive range of numerical values within III*C in the frequency domain signifies a monotonous character of related values within IIIc in the spatial domain and, consequently, a rather ill-conditioned matrix Q to be inverted. Should one still succeed with the inversion, the lack of aliasing makes the resulting interpolation very accurate and practically identical with the ideal sinc-interpolation.

Direct Least Squares Interpolator

Equations (16) and (18) show the general character and form of the transfer function A typical for LSI. However, this form is not suitable to derive its transform a(x), which could be used as a functional operator for a direct interpolation expressed by Equations (12) and (15b). We should attempt to find an alternate formulation for A whose transform can be derived more easily.

Function A in Fig. 2 has an appearance of a blurred function II as if the ideal II were convolved with an auxiliary, strongly peaked and narrow Gaussian E which is area-normalized

$$A = E*II \quad . \tag{19}$$

This is illustrated in Fig. 4. As a Gaussian, function E conforms to the definition of a probability density function ϕ, i.e. $E = \phi$, and the left part of II is, of course, coincident with a unit step function H, as shown in Fig. 5. The shape of the transfer function as expressed in its left half A_L by a modification of Equation (18)

$$A_L(s) = (1 + \exp(-D(2s+1)))^{-1} \tag{18a}$$

can be modelled by a cumulative distribution function Φ obtained from a progressive summation of ϕ, i.e. by a convolution of ϕ and H

$$A_L \simeq \Phi = \phi*H \quad . \tag{20}$$

In turn, a derivative of Φ inherently defines the value of ϕ

$$\phi(s) = \Phi'(s) \quad .$$
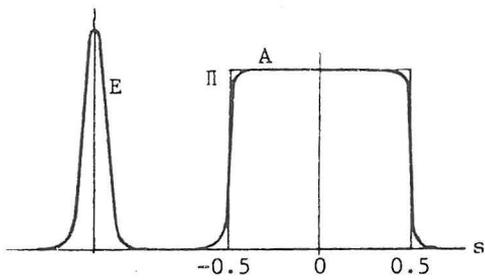
Fig. 4 Blurred function Π
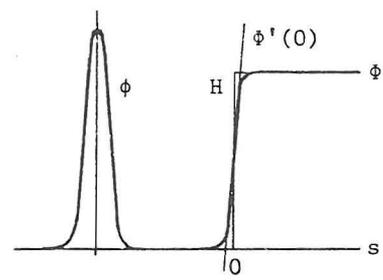
Fig. 5 Distribution function Φ

To achieve a close Φ-to-A_L fit we specify that Φ assumes the slope of A_L in its centre point. By definition, this determines the peak value for φ

$$\phi(0) = \Phi'(0) = A_L'(-0.5) \quad .$$

From Equation (18a) we develop

$$A_L'(s) = 2D \, \exp(-D(2s+1))(1+\exp(-D(2s+1)))^{-2}$$

and derive

$$A_L'(-0.5) = 2D(1+1)^{-2} = D/2 \quad .$$

Since φ is area-normalized and we know also its peak value $\phi(0) = D/2$, it is possible to derive its analytical form in terms of the covariance spectrum C. The same form also represents the auxiliary blurring function E

$$\phi(s) = E(s) = \tfrac{1}{2}D \, \exp(-\tfrac{1}{4}\pi D^2 s^2) \quad . \tag{21}$$

From Equation (19) we are finally able to determine the interpolation operator a as an analytical function

$$\boxed{a(x) = e(x)\mathrm{sinc}(x) \leftrightarrow A(s) = E(s)*\Pi(s)} \quad , \tag{22}$$

where

$$\boxed{e(x) = \exp(-4\pi x^2/D^2)} \quad .$$

Following our well established path (3)

$$c \longrightarrow e \equiv c \longrightarrow C \longrightarrow E \longrightarrow e$$

we also review the results of the derivations using value d from the original covariance function as an independent parameter

$$c(x) = \sqrt{d/\pi} \, \exp(-dx^2) \leftrightarrow C(s) = \exp(-\pi^2 s^2/d) \quad ,$$
$$\phi(0) = \pi^2/2d \quad , \tag{23}$$
$$e(x) = \exp(-4d^2 x^2/\pi^3) \leftrightarrow E(s) = (\pi^2/2d)\exp(-\pi^5 s^2/4d^2) \quad .$$

The effective LS interpolator a is derived by an auxiliary Gaussian attenuation of the sinc-function. Its spectral response A closely approximates the rectangle function Π. The degree of attenuation and smoothing is affected by the choice of the original covariance function c or of its transform C. This way, one can control the performance of LSI as suited to particular conditions of the application. More information on this and on numerical aspects of the direct LSI was presented by Kratky (1978).

CONCLUSIONS

Spectral analysis helps to understand the function and to estimate the effectiveness of various interpolation and smoothing methods. For regular sampling patterns, very common in photogrammetry, the formulation of discrete convolution is an analytical equivalent to linear transformations and

systems of linear equations to be solved for parameters, as typical in LSI and other methods. The method of LSI examined with the use of Fourier transforms shows a spectral response which can also be achieved by means of a direct interpolator expressed by an analytical function. Aspects of accuracy and stability of LSI are easy to understand and interpret from the analysis of spectral and spatial relations.

REFERENCES

Bracewell, R.N. (1965), "The Fourier Transform and Its Applications", McGraw-Hill, New York.

Kratky, V. (1978), "DTM Interpolation with Gliding Vectors", Proceedings of the DTM Symposium, ASP, St. Louis.

Kraus, K., E.M. Mikhail (1972), "Linear Least-Squares Interpolation", Photogrammetric Engineering, 38:6.

Lauer, S. (1976), "Automatische Höhenlinienzeichnung mit Darstellung von Geländeunstetigkeiten und digitales Geländemodell", Vermessungstechnische Rundschau, 38:1.

Moritz, H. (1973), "Neuere Ausgleichungs- und Prädiktionsverfahren", Zeitschrift für Vermessungswesen, 98:4.

Schut, G.H. (1976), "Review of Interpolation Methods for Digital Terrain Models", The Canadian Surveyor, 30:5.

APPENDIX

Resampling by Discrete Convolution

If $\overline{F}$ denotes the central unit section of an aliased spectrum $\text{III} * F$ then obviously

$$\text{III} * F = \text{III} * \overline{F} = S(\overline{F}) \quad , \tag{A}$$

where $S(\overline{F})$ represents a set of adjacent identical sections $\overline{F}$ composing the full replicated spectrum $\text{III} * F$. Similarly, the original spectrum F can also be partitioned into a set of unit sections F_i which are not identical

$$F = S(F_i) \quad , \qquad i = \ldots -2, -1, 0, 1, 2 \ldots \tag{B}$$

These sections F_i, shifted to a single unit interval obviously define section $\overline{F}$ by summation

$$\overline{F} = \sum_i F_i \quad . \tag{C}$$

Spectrum $H = C(\text{III} * F)$ can be modified by using the above S-definitions yielding $S(H_i) = S(C_i)S(\overline{F}) = S(C_i\overline{F})$ and hence

$$H_i = C_i\overline{F} \quad . \tag{D}$$

To derive a replicated spectrum $\text{III} * H$ one should consider the definition from Equation (C) first and then substitute for H_i from Equation (D) so that

$$\overline{H} = \sum_i H_i = \sum_i (C_i\overline{F}) = \left(\sum_i C_i\right)\overline{F} = \overline{C}\overline{F} \quad .$$

Since $\text{III} * H = S(\overline{H}) = S(\overline{C})S(\overline{F})$, one arrives at

$$\boxed{\text{III} * H = (\text{III} * C)(\text{III} * F)} \quad . \tag{E}$$