7/3

TRANSFORMATION and INTERPOLATION

of

INCONSISTENT COORDINATES

INTERNATIONAL INSTITUTE FOR AERIAL SURVEY AND EARTH SCIENCES ENSCHEDE

TRANSFORMATION and INTERPOLATION of INCONSISTENT COORDINATES

by

F. Leberl

ITC - Enschede

March 1972

0. INTRODUCTION

Transformation and interpolation are considered to be independent mathematical tools, to be applied to different problems. This is only true to a limited extent. Number of applications are feasible, in which transformation and interpolation are closely related and can successfully be combined. Furthermore, the methods applied for each of the two are very similar.

The applications under discussion concern the transformation and interpola= tion of "inconsistent" cooordinates. The coordinates of two sets of points are called "inconsistent", if they are of different origin. There may be differences due to measurements or due to the method of computation or due to a combination of both. Thus there is a stochastic component involved.

In the following, the relation between transformation and interpolation is explained in some detail. Then, a number of mathematical concepts, referring to interpolation mainly, are introduced. Finally a number of commonly applicable methods for transformation and interpolation of in= consistent coordinates is surveyed.

1. THE PROBLEM

Given are two sets of points by their coordinates in n-dimensional space, respectively. Required is the establishment of a functional relationship among the two groups of points, such that each point in set A corresponds to one point in set B (see figure 1). Using this functional relationship, it is trivial then to create for any additional point P_A of set A the corresponding point in set B, denoted by P_B. In mathematical terms, set A is "mapped" onto set B. This is the basic idea of "transformation".

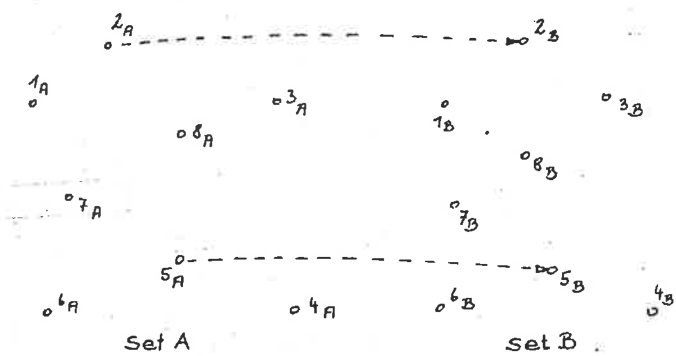

Fig.1 Two sets of points among which a functional relationship should be defined

If the coordinates of set A and B are inconsistent in the sense of the above definition, then the functional relationship for the transformation is not strict. It is reasonable to establish it by overdetermination, so that there will residuals be left. Furthermore, the establishment of the functional relation is subject to incertainities and consequently to interpretation and arbitrary decisions.

In photogrammetry and surveying, the problem of transforming inconsistent coordinates occurs, e.g. whenever one set of points is measured in different ways, so that different coordinates result, such as from ground survey and photogrammetry, or from an old geodetic triangulation and a new one etc. There are then usually points P_A, which are known in one of the sets only. The problem to be solved by transformation is then to create the corresponding point P_B in the other set.

The residual discrepancies left after transformation of inconsistent coordinates are to be compensated by an interpolation. In these cases, transformation is very strongly related to interpolation. It even is possible to define the transformation as an interpolation and vice-versa. One may define "interpolation"[1] as a procedure, by which a required unknown value, belonging to a given point P, is estimated using the known values at other given points 1, 2, ...

It is obvious now, that the difference between transformation and inter= polation can be defined by consideration of their inputs: the thransformation requires two sets of points to be given to map one into the other. An interpolation requires one point set to be given, and for a certain number of points in addition one or more quantities, which have to be estimated at other points of the set.

---

[1] The problem of interpolation sometimes is referred to as "approximation" (see e. g. J.R. Rice: "The Approximation of Functions", Addison-Wesley Publ. Co., Reading, Massachusetts, USA, 1964)

From the above, the relation between transformation and interpolation is clear. A transformation of set A into set B is carried out by overdetermination. Re= siduals will be left after the transformation (see fig. 2). Interpolation can now be applied to estimate the residuals at those points which are only measured in set A, e.g. point P′ in figure 2. The interpolation makes use of the given residuals and the transformed point P_B.

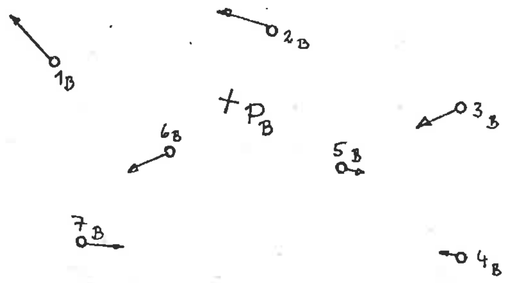

Fig.2 Residuals in point set B after transformation of set A into set B
....error vector

A complicated transformation, followed by a simple interpolation can be replaced by a simple transformation and more complicated interpolation. In the extreme case, no transformation is done anymore, but the residuals for the interpolation are derived as simple differences of coordinates in set A and B.

2. TERMINOLOGY AND MATHEMATICAL CONCEPTS

The functional relationship between set A and B in figure 1 may theoretically be of any kind, thus consist of polynomial, trigonometric, exponential, rational or other functions. In the case, however, where the transformation is improved by a subsequent interpolation, it is advisable to use simple functions, preferably polynomials of low order. The main purpose is to create the situation of figure 2, namely residuals.

The input for interpolation may be directly measured (e.g. digital terrain models) or consist of the output of a transformation (e.g. model deformation after absolute orientation). Points, in which the entity to be interpolated is known, are called "reference-points". The space of the reference points - for applic= ation in mapping - can be:

one-dimensional	:	e.g. for lens distortion curves
two-dimensional	:	e.g. for film shrinkage
three-dimensional	:	e.g. for model deformation

The entities to be interpolated ("residuals") can concern

one dimension : e.g. height deformation in a model

two dimensions : e.g. planimetric deformations in a model

three dimensions : e.g. planimetric and height deformation in a model

It is easy to see now the possible combinations to define interpolation tasks: a n-dimensional "field" is to be interpolated on a m-dimensional reference space, where n are the dimensions of the residuals, and m the dimensions of the reference points.

So, for example, height residuals in a spatial model are a one-dimensional field to be interpolated in a three dimensional reference space. Or, for example, Δx-, Δy-, Δz- residuals in the map plane do represent a three dimensional field to be interpolated in the two dimensional reference space.

The easiest case to visually realize is the one-dimensional field on a two-dimensional reference-plane: the residuals can be interpreted here as "heights" in the reference points (see figure 3). The problem of interpolation exists in the definition of a "surface" passing through the reference points.

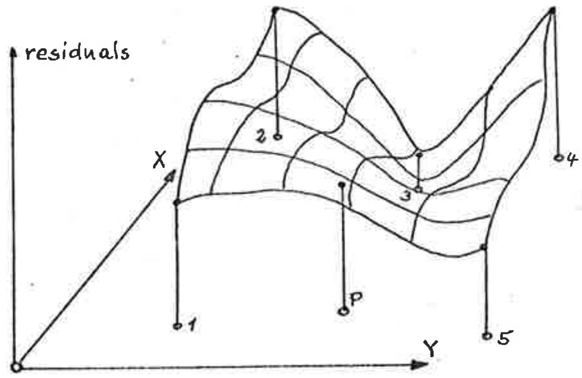

Fig. 3 : Visualization of a one-dimensional field on a two-dimensional reference space (x y- plane)

The residuals are the result of random causes, such as e.g. measuring errors, and therefore the field is called "stochastic field". The mathematical theory of random functions, an extension of classical statistics, considers stochastic fields in full detail.

Residuals in neighbouring points obviously must be correlated, if interpolation should be useful. The correlation, or better the co-variance, can be estimated from the given residuals. As a result, one obtains a covariance function. This function gives the covariance cov among two residuals $r_1(x_1, y_1, \ldots z_1)$,

r_2 (x_2 , y_2 ,... z_2) in function of their position:

$$\text{cov}\ (r_1 , r_2) = f\ (x_1 , y_2 , \ldots z_1 ; x_2 , y_2 , \ldots z_2)$$

There are cases, where the covariance does not depend on the position of the residuals, but only on the distance and direction between them or, in the simplest case, only on the distance between the residuals in consideration. In the last case, the covariance would be translation- and rotation-invariant or, as in the terminology of mathematics, it would be "isotropic". Thus the covariance for a field would be described by a function in two-dimensional space , as shown in figure 4.

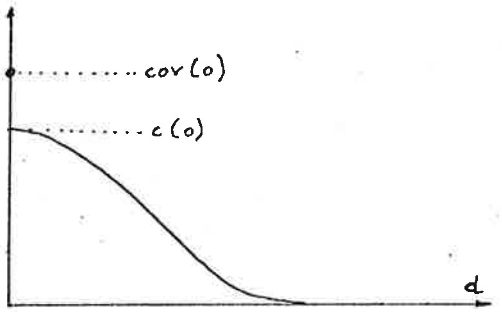

Fig. 4: Typical covariance func= tion.

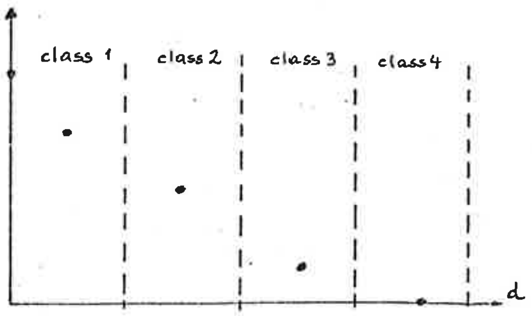

Fig. 5: Sample values for covari= ance function.

The computation of the covariance function of an isotropic field is as follows: A number m of classes of distances d among reference points is chosen. Then, all distances among reference points are grouped in these classes. For each class then, the covariance is computed using:

$$\text{cov}\,(d_i) = \frac{\sum_{j=1}^{k} c_j}{k}$$

i .... 0, 1, 2, ... n ;

k .... number of distances in class i ;

c_j .... product of residuals at end points of distance j.

For d_i = 0 , the variance cov(0) is obtained. For every class of distances, a sample value of the covariance function is computed. Thus, the covariance function is given as a number of discrete points (see figure 5). To derive a continuous function, one usually puts a two-parametric function through the discrete points, e.g.

$$\mathrm{cov}\ (d) = C(0) \cdot e^{-\lambda \cdot d}$$

or

$$\mathrm{cov}\ (d) = C(0) / (1 + (\frac{d}{\lambda})^2)$$

who's parameters $C(0)$ and λ are determined in a least squares algorithm. To do this, the sample value $\mathrm{cov}(0)$, thus the variance, is <u>not</u> used. As a result, the function of figure 4 may result:

There will be a covariance function, which, at $d = 0$, has two values, namely $\mathrm{cov}(0)$ and $C(0)$. What is the significance of this?

One may assume, that the residuals r are composed of a correlated and an uncorrelated component, c and u, respectively. The covariance-function of the correlated part, c, is

$$\mathrm{cov}_c\ (d) = C(0) \cdot e^{-\lambda . d}$$

for example. The covariance function of the uncorrelated part, u, is:

$$\mathrm{cov}_u\ (d) = \mathrm{cov}\ (0) - C(0) \qquad \text{if } d = 0$$

$$= 0 \qquad \text{if } d \neq 0$$

One usually is interpreting the uncorrelated part, u, as a measuring error, if the "residuals" are directly observed. This measuring error, as a hypothesis, is thus considered to be uncorrelated. Summarizing the above, one can formulate

$$r = c + u$$

$$\mathrm{cov}_r\ (0) = \mathrm{cov}_c\ (d) + \mathrm{cov}_u\ (d)$$

3. <u>TRANSFORMATIONS</u>

Transformations can be classified according to:

- linear or non-linear transformations
- simple, piece-wise and pointwise transformations,
- the dimension of the space to be transformed .

Furthermore, transformations are called "aphyllactic", if angles as well as surfaces are deformed. They are called "equivalent", if surfaces remain unchanged. They are "conformal", if angles are not deformed. For this last feature, the transformation must fulfil the conformity condition

$$\frac{\partial x_A}{\partial x_B} = \frac{\partial y_A}{\partial y_B} \qquad\qquad (1)$$
$$\frac{\partial x_A}{\partial y_B} = \frac{\partial y_A}{\partial x_B}$$

3.1 TRANSFORMATION WITH SIMPLE FUNCTIONS

3.1.1 Linear Transformation

If the functional relationship between set A and B is linear in the variables, then the transformation is called "linear".

Linear conformal transformation:

This is the most important linear transformation. Denoting the coordinates in set A by small, in set B by bold letters, its formula states:

$$X = a_1 . x + a_2 . y + a_3$$

$$Y = -a_2 . x + a_1 . y + a_4$$

or:

$$X = \lambda (\cos \alpha . x + \sin \alpha . y) + c_x$$

$$Y = \lambda (-\sin \alpha . x + \cos \alpha . y) + c_y$$

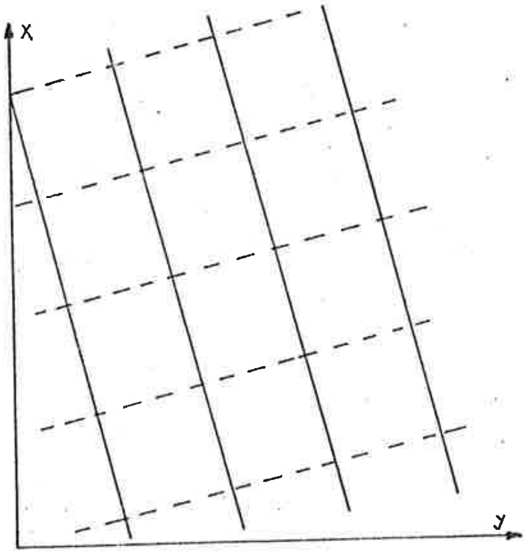

Fig. 6 : Two orthogonal pairs of parallel equidistant lines after linear, conformal transformation

The transformation consists of two shifts (c_x , c_y), a rotation α and a uniform scale change, λ. Figure 6 shows, how a network of equidistant, parallel straight lines with x = const, y = const, is transformed into the new set.

In matrix notation, the transformation is:

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \lambda \cdot R \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} c_x \\ c_y \end{bmatrix}$$

R is an orthogonal rotation matrix. The transformation has four unknowns, which can be determined from two points.

The linear conformal transformation in three dimensions states:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \lambda \cdot R \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} c_x \\ c_y \\ c_z \end{bmatrix}$$

Here, R obviously is a 3 x 3 orthogonal matrix. The transformation has 6 unknowns, which again can be determined from two given identical points in sets A and B.

Affine transformation

This is a general linear transformation, who's formula is for two dimensions:

$$X = a_1 x + a_2 y + c_x$$

$$Y = b_1 x + b_2 y + c_y$$

The affine transformation consists basically of 2 shifts (c_x, c_y), a general rotation plus "shear deform= ation", i.e. angles are deformed, and a general scale change as well as "differential scale change", i.e. scale change is different in different directions. The extrema occur in the "affinity axes". Figure 7 illustrates, how lines of x = const and y = const. are transformed into the set (X,Y):

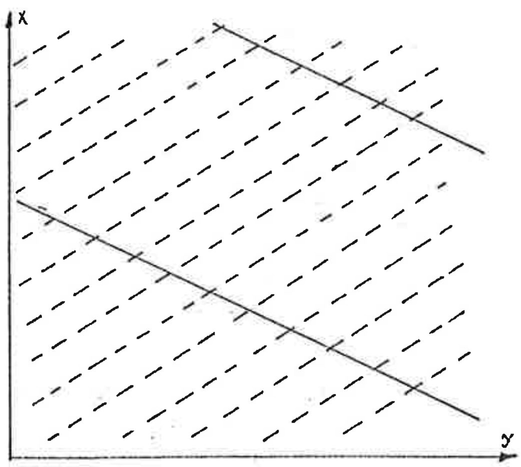

Fig.7 : Two orthogonal pairs of parallel, equidistant lines after affine transformation

the lines are not perpendicular anymore ("shear deformation"), and the distances among lines of $x = \text{const.}$ are different from distances among lines of $y = \text{const.}$ ("differential scale change"). There are 6 unknowns in the transformation formula. They can be computed as long as 3 control points are given.

Generalization to three dimensions gives:

$$X = a_1 x + a_2 y + a_3 z + c_x$$

$$Y = b_1 x + b_2 y + b_3 z + c_y$$

$$Z = c_1 x + c_2 y + c_3 z + c_z$$

3.1.2 Non-Linear Transformation

Projective transformation

The transformation equation for this case consists of a ratio of two linear polynomials

$$X = \frac{a_1 x + a_2 y + a_3}{c_1 x + c_2 y + 1}$$

$$Y = \frac{b_1 x + b_2 y + b_3}{c_1 x + c_2 y + 1}$$

The transformation has 8 unknowns which can be found from four control points. Figure 8 shows the typical projective deformation of the network of parallel, equi=distant lines of $x = \text{const.}$, $y = \text{const.}$ They do, after being transformed, intersect in one point, the so-called "vanishing point".

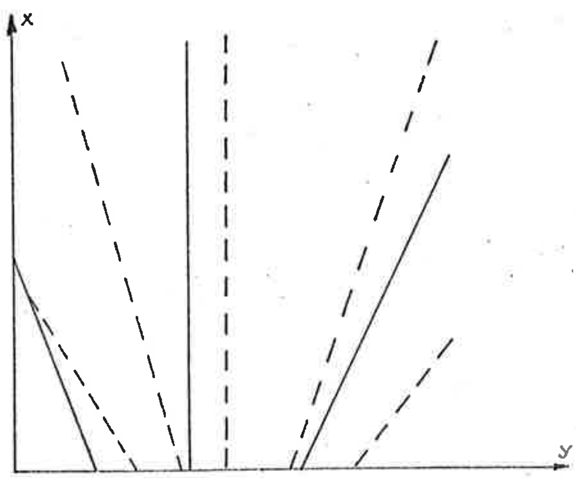

Fig.8 : Two orthogonal pairs of parallel, equidistant straight lines after projective transformation

Non-linear conformal transformation

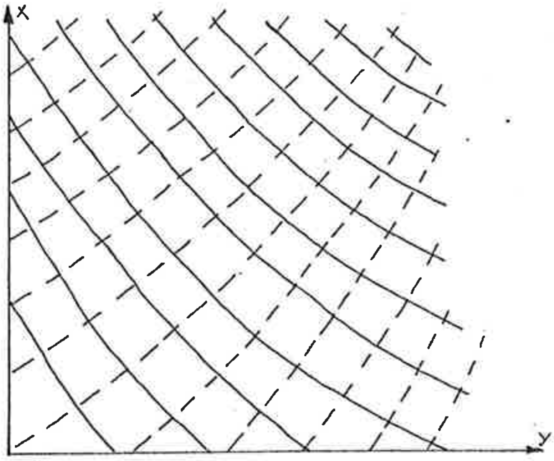

Fig.9 : Two orthogonal pairs of parallel, equidistant straight lin after 2nd order, conformal trans= formation

The transformation equation is:

$$x = c_x + a_1 x - a_2 y + a_3 (x^2+y^2) - 2a \; .x.y + \ldots$$

$$y = c_y + a_1 y + a_2 x + 2a_3 \; x.y + a_4 (x^2-y^2) + ..$$

Conformal transformations are feasible of any order. The equation represents a second order conformal transformation. Conformity can be checked by the condition (1). It may be easier for some to find the higher order conformal polynomials from a formulation with complex numbers:

$$X + i\,Y = (c_x + i\,cy) + (a_1 + i.a_2).(x + i.y) + (a_3 + i.a_4)\;\;(x + i.y)^2 + ..$$

Figure 9 shows, how a second order conformal transformation works. Obvious is the conformity : the lines intersect at angles of 90^o, such as before transformation.

General non-linear transformation with polynomials:

Here

$$X = \sum_{i=0}^{n} \sum_{j=0}^{i} a_{i,j} \cdot x^{i-j} \cdot y^{j}$$

$$Y = \sum_{i=0}^{n} \sum_{j=0}^{i} b_{i,j} \cdot x^{i-j} \cdot y^{j}$$

or

$$X = a_{oo} + a_{10}\, x + a_{11} y + a_{20}\, x^2 + a_{21}\, xy + a_{22} y^2 + ..$$

$$Y = b_{oo} + b_{10}\, X + b_{11}\, y + b_{20} x^2 + b_{21}\, xy + b_{22} y^2 + ..$$

There are, in the general case of a polynomial of order n, (n+1).(n+2)/2 unknowns. The number of coefficients becomes rather large for higher order polynomials.

Therefore, one does often not use the full polynomials, but just parts of it, setting few of the coefficients equal to zero, e.g.:

$$X = a_{oo} + a_{10}x + a_{11}y + a_{21}\ x.y$$

$$Y = b_{oo} + b_{10}x + b_{11}y + b_{21}\ x.y$$

The use of higher order polynomials certainly must be done with care, since surprising results are possible.

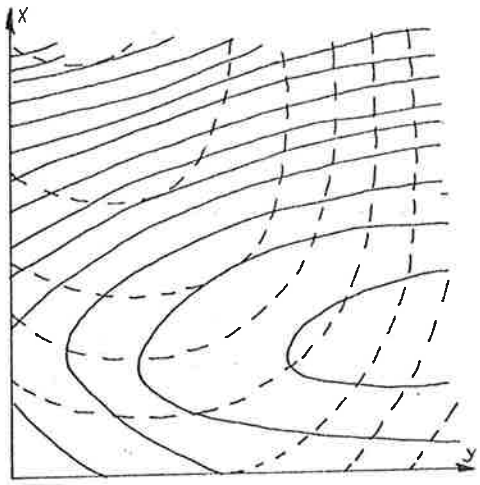

Fig.10 : Two orthogonal pairs of parallel, equidistant straight line after transformation with 2nd order polynomial

Figure 10, showing graphically the effect of a second order polynomial trans= formation, makes clear that straight lines are transformed into curves of the order of the polynomial in use.

Other types of non-linear transformation

Complicate non-linear transformations are in use in the field of geodesy, especially to image the earth's surface onto a plane. For these transform= ations, various kinds of functions are used, such as trigonometric, exponential, logarithmic and others. Since discussion of these transformations is not relevant in the present context, reference is made to the geodetic literature on map-projections.

3.2 PIECE-WISE TRANSFORMATION

Up to here it was silently assumed that for each set A of points only one function is used to transform it into set B. But it is possible to use various functions for various subsets of A, so that actually separate transformations are performed. See figure 11 for visualization of this concept.

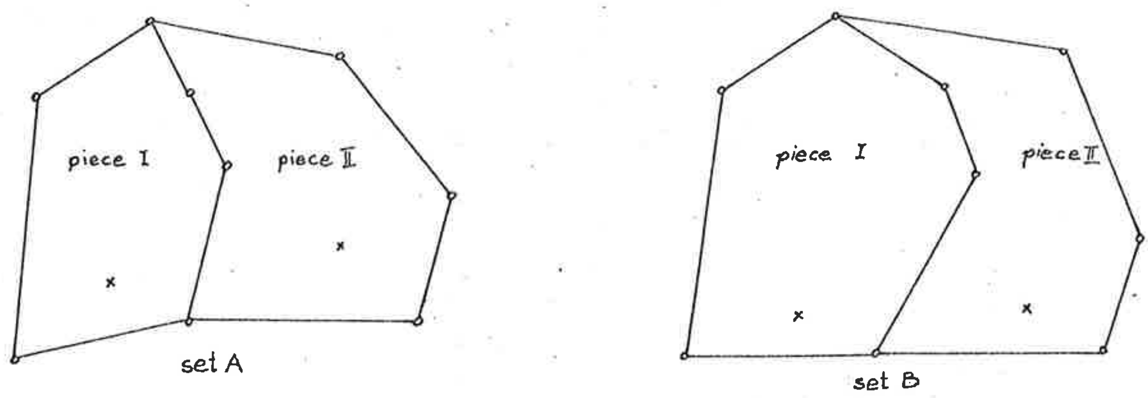

Fig. 11 : Concept of piecewise transformation, illustrated by 2 pieces

The advantage of this concept is, that relatively simple functions are used in each piece, but still a great flexibility is possible through subdivision in many pieces.

A disadvantage can be, that sudden changes of the transformation occur along the boundary lines of neighbouring pieces. These changes mostly are illogical. Therefore one introduces, in applying the concept, joining conditions for adjacent pieces: e.g. one can ask for the same value and even first and higher order derivatives along boundary lines.

3.2.1 Piece-wise Affine Transformation

The network of given points in set A and B is used to build corresponding triangles. Each triangle then defines exactly an affine transformation through the three corner points. Each new point of set A is then transformed into B by using the triangle he falls in (see fig. 12).

It can be easily seen, that along boundary-lines, the transformation formulae do give the same values, so that for a point on the boundary line the trans=formation will produce the identical result with either of the two triangles. This is a very attractive feature of the method.

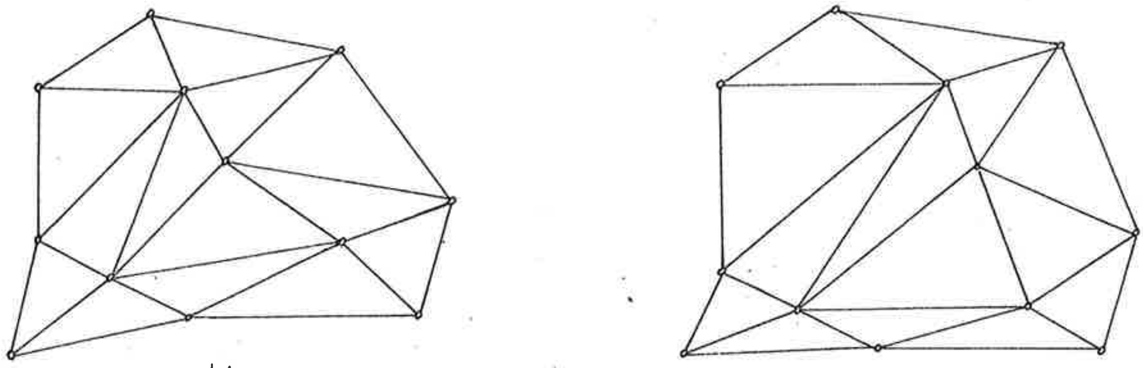

figure 12 : Illustration of the principle of piecewise affine transformation

3.2.2 Other Piece-wise Transformations

In general, the network of points present in set A and B can be separated to larger than just triangular pieces, as was shown in fig.11. In each piece, a certain transformation can be defined. The main problem is to avoid dis= continuities of the piece-wise transformation function along boundary-lines. The most effective way of defining boundaries is as parallel lines, eventually in the direction of coordinate axes. Then the number of conditions to avoid discontinuities along boundaries is kept small.

3.3 POINT-WISE TRANSFORMATION

The method of point-wise transformation is very logical: every point is trans= formed with another transformation function. Numerically this can be very simple.

3.3.1 Principle of Arithmetic Mean

The coordinates in set B of a transformed point are found from the formula:

$$X = x + \frac{p_1 \cdot X_1 + p_2 \cdot X_2 + \cdots + p_n \cdot X_n}{p_1 + p_2 + \cdots + p_n} - \frac{p_1 x_1 + p_2 \cdot x_2 + \cdots + p_n \cdot x_n}{p_1 + p_2 + \cdots + p_n}$$

$$Y = y + \frac{p_1 Y_1 + p_2 \cdot Y_2 + \cdots + p_n \cdot Y_n}{p_1 + p_2 + \cdots + p_n} - \frac{p_1 \cdot y_1 + p_2 \cdot y_2 + \cdots + p_n \cdot y_n}{p_1 + p_2 + \cdots + p_n}$$

The weight factors p_1 , p_2, ...p_n are found from set A :

$$p_i = 1/(x - x_i)^2 + (y - y_i)^2)^{\frac{k}{2}}$$

k has to be chosen in a proper way.

It is sufficient to use only the points in the vicinity of the one to be transformed for the computation of the weighted mean, since the weight of points farther away is very small.

3.3.1 Moving Average

In this method the coordinates of each new point in set B are computed by a function, whose coefficients are repeatedly determined for every new point. For the example of the polynomial

$$X = a_0 + a_1 x + a_2 y + a_3 x^2 + ..$$

$$Y = b_0 + b_1 x + b_2 y + b_3 x^2 + ..$$

the unknown coefficients a_o , b_0 , ... are found from the given points. The advantage of the method consists in the fact, that only the points in the vicinity of the one to be transformed are used to find the coefficients of the transformation formula.

In figure 13, the principle of the Moving Average is illustrated.

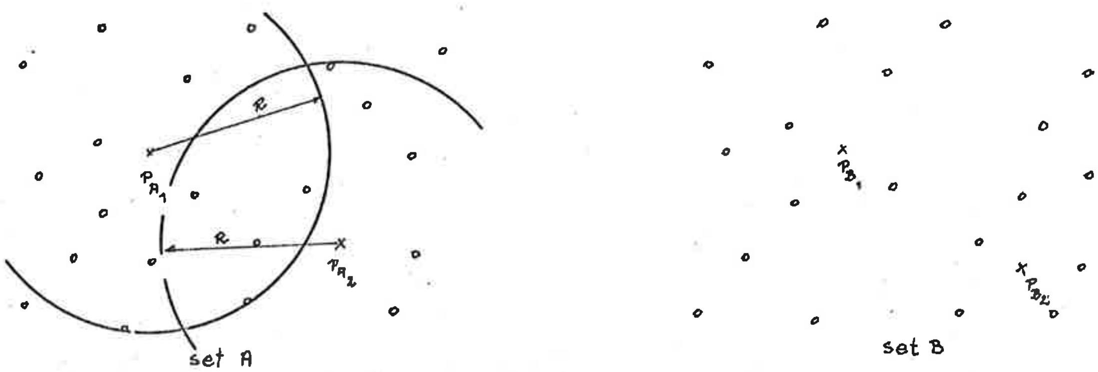

Fig.13 : Moving average : for every other new point P_A , another function is found from the given points in close vicinity of point P_A , e.g. selected according to a critical radius R.

The method is probably the most flexible of all discussed so far. The trans= formation function can be very simple, e.g. a polynomial of low order. A dis= advantage may only consist in the eventually large computing effort required with this method.

4. INTERPOLATION

Once a transformation is carried out and residuals are obtained at the reference points, or "residuals" are observed directly, an interpolation can be used to estimate residuals at the new points and thus obtain a better estimate of the true situation of the new point in set B.

4.1 INTERPOLATION WITH SIMPLE FUNCTIONS

The problem is to put one function through the residuals at the reference points. These residuals are considered as "height" on the reference plane, such as shown in fig. 3. The function represents a "surface".

Since it is usually a complicated "surface" of residuals, a higher order polynomial or other complicated function would have to be used. Basically the input consists of:

ΔX , ΔY, ... ΔZ , which are the residuals to the coordinates in set B at all reference points.

X_i , Y_i , ... Z_i , which are the coordinates of the reference points in set B

X_p , Y_p , ... Z_p , which are the coordinates of the new point in set B.

So the problem is to find the coefficients of the functions

$$\begin{aligned}
\Delta X &= X(X_i, Y_i, \dots Z_i) \\
\Delta Y &= Y(X_i, Y_i, \dots Z_i) \\
&\vdots \\
\Delta Z &= Z(X_i, Y_i, \dots Z_i)
\end{aligned}$$

and, once the coefficients are known, to compute ΔX_p, ΔY_p, ... ΔZ_p from the formula using X_p , Y_p , ... Z_p as the independent variables.

The choice of functions is as large as for transformations. It is advisable, however, not to use linear functions. These, obviously, would represent a plane surface through the residuals and this may mostly be a surface too simple to approximate the behaviour of the residuals.

For the description of photogrammetric strip deformations, traditionally various kinds of non-linear polynomials were used (see Adjustment of Aerial triangulation

4.2 INTERPOLATION WITH PIECEWISE FUNCTIONS

This is the method most in use at the present time.

4.2.1 Meshwise Bi-linear Polynomial

This method is very effective and represents the equivalent to the piecewise affine transformation. The reference points form triangles which exactly define a plane. As a result the "surface" through the reference points is composed of triangular plane pieces, as a polyeder. The interpolation formula is

$$\Delta X = a_o + a_1 X + a_2 Y$$

The coefficients a_o , a_1 , a_2 are found for each of the triangles from the given corner points.

4.2.2 Piecewise Non-linear Polynomials

In contrary to transformations, this concept can more easily be applied to interpolation, since it is much easier to define boundary-lines for the polynomial pieces. These boundaries can be chosen arbitrarily and do not have to pass through reference points. In this way a minimum number of straight lines is used for definition of each boundary. Along these few lines the number of joining conditions to be fulfilled by neighbouring function pieces is not too large (see figure 14).

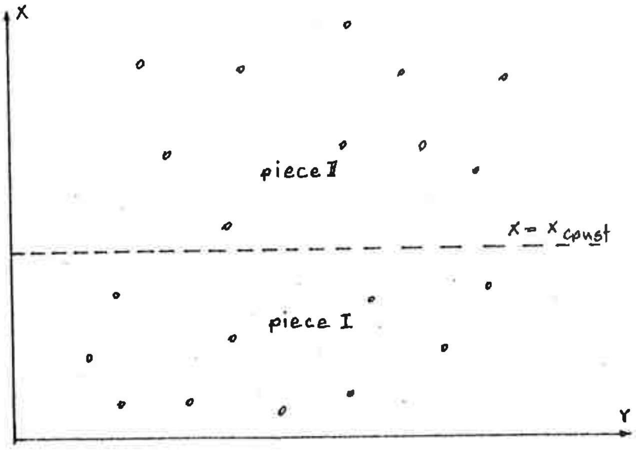

Fig. 14 : Concept of piecewise inter= polation, boundary line x = x const.

For the example of a second order polynomial in two pieces, one can give for every reference point the equation:

$$\Delta x = a_o + a_1 x + a_2 y + a_3 x^2 + a_4 y^2 + a_5 \cdot xy$$

if it is in piece 1, or

$$\Delta x = b_o + b_1 x + b_2 y + b_3 x^2 + b_4 y^2 + b_5\, xy,$$

if it is in piece 2.

The joining conditions to ensure the same function values for the two polynomials along the boundary line $x = x_{const}$ are:

$$a_o + a_1\, x_{const} + a_3\, x^2_{const} = b_o + b_1\, x_{const} + b_3\, x^2_{const}$$

$$a_2 + a_5\, x_{const} = b_2 + b_5\, x_{const}$$

$$a_4 = b_4$$

The concept can be generalized in a straight forward manner to other than polynomial functions and more than 2-dimensional reference-space.

4.3 Pointwise Interpolation

4.3.1 Arithmetic Mean

The formula for interpolating residuals is:

$$\Delta x = \frac{p_1 \cdot \Delta x_1 + p_2 \cdot \Delta x_2 + \ldots + p_n \cdot \Delta x_n}{p_1 + p_2 + \ldots + p_u}$$

$$\Delta y = \frac{p_1 \cdot \Delta y_1 + p_2 \cdot \Delta y_2 + \ldots + p_n \cdot \Delta y_n}{p_1 + p_2 + \ldots + p_n}$$

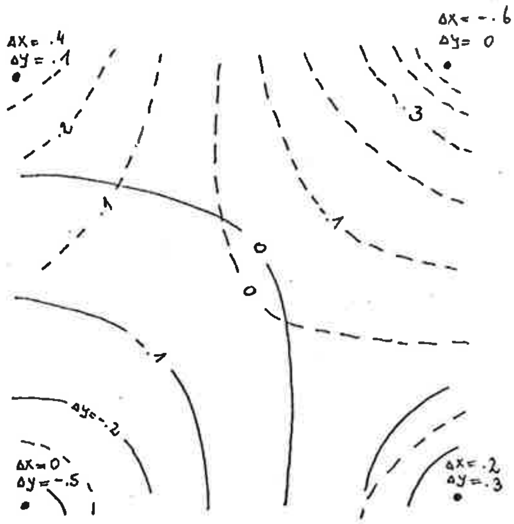

Fig. 15 : Isolines through interpolation with arithmetic mean • .....ref. pts.

The weight factors p_1 usually are chosen as:

(a) $P_i = \frac{1}{d^i}$, $d^2 = (x - x_i)^2 + (y - y_i)^2$

or as

(b) $P_i = 1 / (1 + d^k)$

In the first case (a), the interpolation procedure would in theory give the exact value x_i , if it would be interpolated. In the second case, however, the interpolated "surface" would not pass through the reference points. A "smoothing" would be performed of the surface.

The weight function p (d) should approximate the eventually computed covariance function.

4.3.2 Moving Average

In analogy to what was said about the method as applied to transformation, the ΔX , ΔY , ... ΔZ are found in a new point by putting a function (preferably simple polynomial) through the neighbouring reference points. The method is used in various Digital Terrain Models, e.g. in France, Finland and the CSSR. The name and concept has its origin in the mathematical theory of random functions.

To simplify the numerical procedure, one is reducing the coordinates of the reference points onto the new point. In this way it is only necessary to compute the polynomial coefficients a_o and b_o:

$$\Delta X_P = a_o$$

$$\Delta Y_P = b_o$$

since $X_P = Y_P = 0.$

4.3.3 Least-Squares Interpolation

The method of "interpolation with least squares", which has actually the name "linear prediction and filtering", comes from the theory of random functions. Provided that some theoretical conditions do apply, the method has very good performance.

The interpolation formula is:

$$\Delta x = a_1 \cdot \Delta x_1 + a_2 \cdot x_2 + \quad a_n \cdot \Delta x_n$$

with $(a_1 , a_2 , \ldots ,a_n) = \underline{a}$ defined as:

$$\begin{array}{lllll} a_1 = & q_1 \cdot Q_{11} & + q_2 \cdot Q_{21} + & \vdots & + q_n \cdot Q_{n1} \\ a_2 = & q_1 \cdot Q_{12} & + q_2 \cdot Q_{22} + & \vdots & + q_n \cdot Q_{n2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_n = & q_1 \cdot Q_{1n} & + q_2 \, Q_{2n} + & \vdots & + q_n \cdot Q_{nn} \end{array}$$

or, in matrix notation:

$$\underline{a} = \underline{q} \cdot Q$$

The vector $\underline{q}$ is:

$$\underline{q} = (\mathrm{Cov}\ (d_1)\ ,\ \mathrm{Cov}\ (d_2)\ ,\ \ldots\ldots\ \mathrm{Cov}\ (d_n)\ )$$

The elements $\mathrm{Cov}(d_i)$ are the covariances between the n given points and the new point.

The matrix Q is the inverse of the following:

$$Q^{-1} = \begin{bmatrix} \mathrm{Cov}\,(0) & \mathrm{Cov}\,(d_{12}) & \mathrm{Cov}\,(d_{1n}) \\ \mathrm{Cov}\,(d_{21}) & \mathrm{Cov}\,(0) & \mathrm{Cov}\,(d_{2n}) \\ & & \\ \mathrm{Cov}\,(d_{n1}) & \mathrm{Cov}\,(d_{n2}) & \mathrm{Cov}\,(0) \end{bmatrix}$$

The diagonal elements do contain the variance of the whole field, including thus the eventually present uncorrelated component (measuring error, see fig. 4). With this method, it is possible to obtain an interpolated value for the reference points too (as, by the way, also with the overdetermined simple and piecewise functions, the moving average and the arithmetic mean with weight (b)). In this case, if thus reference point j is interpolated, the vector $\underline{q}$ is then:

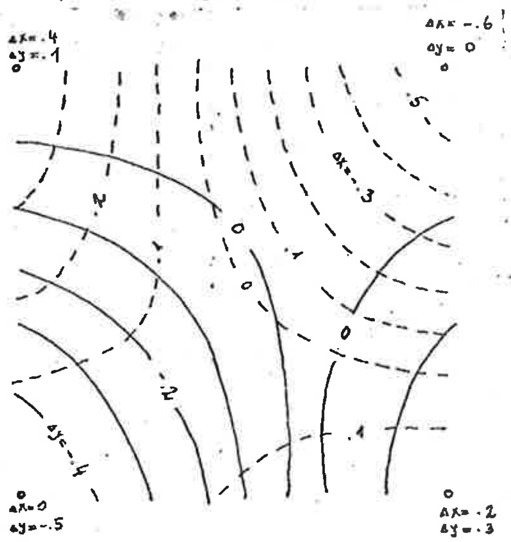

Fig.16 : Isolines as interpolated with "linear prediction" ..ref.pts.

$$\underline{q} = (\ \mathrm{Cov}(d_1)\ ,\ \mathrm{Cov}\,(d_2),\ \ldots\ \mathrm{Cov}\,(d_{j-1}), C(0), \mathrm{Cov}(d_{j+1}) .. \mathrm{Cov}(d_n))$$

Q remains unchanged.

So for $\underline{q}$, not Cov(0), but C(0) is taken. The uncorrelated component is thus not used in the interpolation, it is "filtered out". The purpose of this is obvious : first, interpolation of reference points will produce better estimates of their true values, since the measuring error is eliminated, and secondly the interpolation of new points is more accurate. If, however, the uncorrelated component cannot be interpreted as measuring error, then there is no point in filtering the reference points.

The method will only be successfully applicable, if a few conditions concerning the covariance function and the mean of the field are fulfilled:

- the mean must be zero
- the mean must be about the same when computed separately for various subsets of the field
- the covariance function must go to zero for $d \rightarrow \infty$
- the covariance function must be the same when computed separately for various subsets of the field

If these conditions are not fulfilled, then a preliminary interpolation has to be made by means of a simple function. So the method usually should only be applied after a reasonably flexible transformation or as refinement after an interpolation with a simple function.

4.4 A SPECIAL INTERPOLATION PROBLEM WITH PHOTOGRAMMETRIC MODEL DEFORMATIONS

Model deformations usually have to be interpolated from 4 reference-points located in the corners of the model. An appropriate method to do this would be the arithmetic mean and least squares interpolation.
A simple function, however, causes the problem, that there will be breaks along the edges of adjacent models. Only the piecewise bi-linear inter= polation, according to figure 18, could avoid cracks.

But the piecewise bi-linear method has the drawback of arbitrarily choosing a diagonal to obtain two triangles.Which diagonal then? The problem can be overcome by using the following incomplete third order polynomial.

$$\Delta z = a_0 + a_1 x + a_2 y + a_3 x^2 + a_4 y^2 + a_5\, x\, y + a_6\, x^2 y + a_7\, x\, y^2$$

There are, however, eight coefficients in this formula, but only four

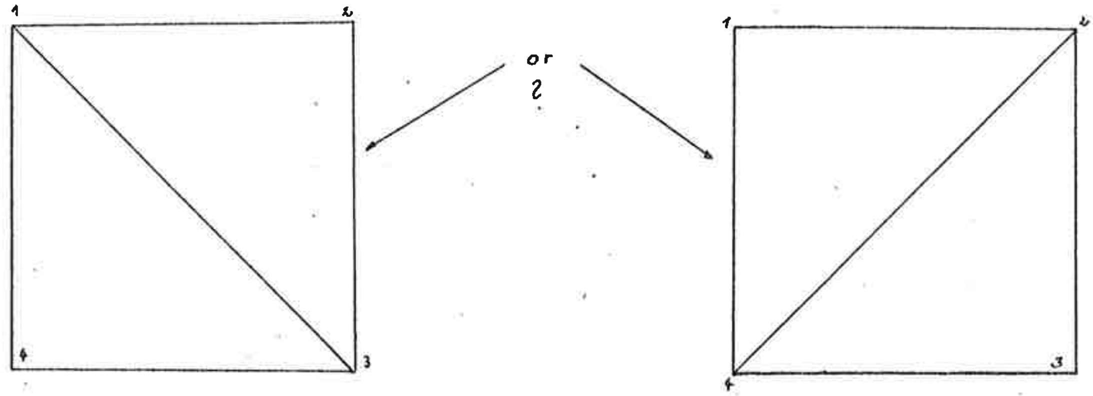

Fig. 17 : Piecewise bilinear method to interpolate model-deformations from four reference points would pose the problem of selecting one diagonal

reference points with known residuals. Therefore, one has to create four more values in the middle of the edges of the model by taking the arith= metic mean of the residuals in the endpoints of the sides (1,2), (2,3), (3,4), (4,1), respectively. These four values, together with the four given ones, can just define the surface.

The main advantage of this method consists in the fact, that there will be no cracks between the interpolating polynomials along the edges of adja= cent models.
The functions along the edge line do only depend on the residuals in the end -points.

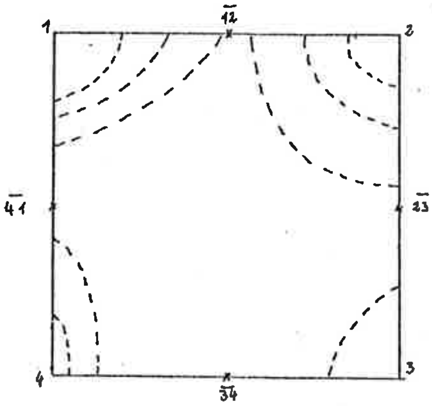

Fig. 18 : Isolines after special interpolation method

The interpolation method has been developed in 1964 in Germany for the special requirement of describing model deformations. At the present time, however, there are the methods of pointwise interpolation developed, so that the one just described is outdated. Nevertheless it is interesting to see, how a method was developed for a very peculiar application.

4.5 GRAPHICAL INTERPOLATION AFTER STRINZ

The residuals in the reference points can be considered as heights and contour lines be graphically interpolated. For every point P (x_p, y_p), the residual can easily be interpolated from the contour plot (see figure 19). Usually the results obtained with this method compare quite favourable with all other methods. The only disadvantage may consist in the fact, that the procedure is rather laborious. On the other hand, it is completely independent from calculating tools. This may be one of the main reasons, why the method has been so popular in photogrammetry, where it is used for many years as e.g. graphical strip-adjustment after Zarzizky.

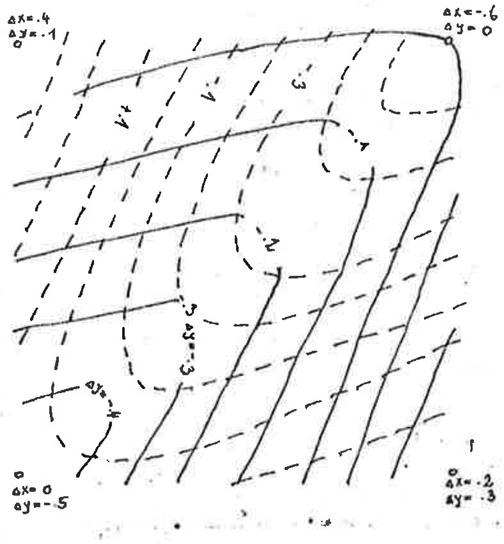

Fig. 19 : Isolines after Strinz

5. SUMMARY

A survey is given of transformation and interpolation methods relevant to applications in photogrammetry. To begin with, a number of general mathematical concepts was explained, such as "stochastic field" and "covariance function".

The transformation and interpolation - closely related and even interchangeable - were classified as linear and non-linear, simple, piecewise and pointwise methods. The various procedures were shortly explained, accompanied by graphical representations of the deformations to be achieved with the various methods.

Finally the description of the modern and rather powerful methods of interpolation with moving average and linear prediction do provide the reader with the new developments in interpolation theory.

LITERATURE

Jaglom, A. M. "Einführung in die Theorie stationärer Zufallsfunktio= nen", Akademie Verlag Berlin, 1959,(in German).

Yaglom, A. M. "Introduction to the Theory of Stationary Random Functions", Prentic-Hall Inc., 1962

Schatz, U. "Das Problem der optimalen Stützpunktdichte und der optimalen Maschengrösze bei der Transformation un= gleichartiger Koordinaten", Doctoral thesis, Univ. of Bonn, W. Germany, 1970, (in German).