

MATCHING OF DIGITAL SCANNER DATA  
BY FEATURE EXTRACTION AND CROSS CORRELATION

Mathias J.P.M. Lemmens,<sup>1)</sup> Wim J. Looyen<sup>2)</sup>

1) Delft University of Technology, Faculty of Geodesy,  
Institute of Photogrammetry and Remote Sensing,  
Thijsseweg 11, 2629 JA Delft, The Netherlands

2) National Aerospace Laboratory NLR,  
Space division, Remote Sensing Department,  
Postbus 153, 8300 AD Emmeloord

Comm. VII/1

**Abstract:** A procedure is presented to trace automatically corresponding points (matching) in CAESAR forward and downward looking images in the land mode. With an interest operator points suitable for matching are indicated in one image. By cross correlation the corresponding points in the other image are traced. Test results are given and improvements of the procedure discussed. The method is probably also suited to determine tie points in the overlap of satellite images for mozaicing.

## 1 Introduction

Many remote sensing data, both from aircraft and satellite sources, are at present in digital format. Since a lot of restoration, enhancement and feature extraction operations are much easier carried out in digital domain than in analogue domain, the potential applications of digital image processing techniques in remote sensing are many. This paper is devoted to the tracing of corresponding points in conjugate image pairs, using digital signal matching techniques -in particular cross correlation- in combination with an interest operator to find suitable correlation points. Suitable points are (sub)pixels which show high grey value variances in all directions, e.g. corners, blobs and spots. The method is adopted from Moravec (1977), where it originally is developed for the guidance of autonomous robots by vision.

Corresponding points are mappings of the same object point in different images. Matching is a general problem and occurs in many vision applications. In photogrammetry, e.g., it is necessary to obtain a surface recovery from stereo pairs of images and to perform aerotriangulation. The last years much effort is spend and many techniques are developed to obtain digital elevation models automatically by digital matching methods. Lemmens (1988) gives a survey.

Very often, e.g. in the case of satellite imagery and flat areas, the imaged object space may be considered as two-dimensional (2-D); matching is necessary to obtain corresponding points in the overlap of the images for mozaicing and for the overlay of multitemporal images. In digital remote sensing usually points are manually indicated with a cursor. Just pixel level accuracy is achieved. In the present investigation tracing of corresponding points is necessary to overlay the downward with the forward looking images of the CAESAR airborne multispectral scanner. The connection must be performed very accurate, so two requirement have to be fulfilled:

- the network of corresponding points should be dense and regular;
- corresponding points in the two images have to be identified at subpixel level.

Manual identification isn't appropriate to accomplish the two tasks, so the assistance of digital matching is called in.

Matching consists of the following three stages:

- 1 Distinction check: selection of distinct points, i.e. points which differ significantly from the neighbouring points; we call them characteristic points;
- 2 Similarity check: selection of candidate corresponding points in the image sequences, using some similarity measure;
- 3 Consistency check: testing of the correspondence assignments using some global model.

### Remarks:

- In feature matching, characteristic points are traced in both images. According to the approach of Moravec (1977), we just search for characteristic points in one image;

- Correspondence analysis is performed by cross correlation, shifting a target area over search windows in the other image;
- The similarity check just yields local consistency, but no global consistency, i.e. because of imperfections in operators and similarity measures, the result will be ambiguous. In stereo imagery the global consistency check is executed by introducing an object model: very often just a smoothness constraint. The present problem shows much resemblance with stereo matching. Because of irregular platform movements, scanner data show, contrary to photographs, non-rigid image geometry. So, the mapping geometry is unknown and can just be approximated by some geometric transformation model (GTM), e.g. a second or third order polynomial. The parameters are adjusted by, for instance, flight path registration. The consistency check is executed by assuming, according to the smooth surface model in stereo matching, that the imaging geometry is varying smoothly.

Before treatment of the procedure some brief notes are devoted to the CAESAR scanner.

## 2 The CAESAR scanner

CAESAR is an airborne multispectral scanner. The acronym reads: CCD Airborne Experimental Scanner for Applications in Remote Sensing. The scanner is jointly built by the National Aerospace Laboratory NLR, and TNO Institute of Applied Physics, TPD Delft, as a part of the Dutch national remote sensing programme (Bunnik et al. 1986). The scanner is composed of commercially available silicium semiconductors of the same type as used in the SPOT HRV's. The total length of the CCD linear array is 1728 elements or 22.46 mm, which leads, together with the focal length of 50 mm, to a total field of view of  $25.7^\circ$ . Since the size of each element is  $13 \times 13$  micrometer, the instantaneous field of view (IFOV) is  $0.013/50 = 0.26$  mrad, which results in a ground resolution of 0.26 H, with the ground resolution in meter and H, the flying height, in km. CAESAR is developed for both land and sea applications. The spectral resolutions for land and sea observations are defined by user requirements. They are listed in table 1.

Land observation		Sea observation	
Band	spectral range (micrometer)	Band	spectral range (micrometer)
1	0.535 - 0.565	1	0.400 - 0.420
2	0.655 - 0.685	2	0.435 - 0.455
3	0.845 - 0.895	3	0.510 - 0.530
		4	0.555 - 0.575
		5	0.620 - 0.640
		6	0.675 - 0.695
		7	0.770 - 0.800
		8	0.990 - 1.050

Table 1

CAESAR consists of a cluster of four camera's. Three camera's point downward and one forward with an off-nadir view angle of  $52^\circ$  (see fig. 1<sup>a</sup>) The forward looking camera results from studies on the non-Lambertian behaviour of vegetation canopies. Vegetation types can be better distinguished by simultaneous observation from different viewpoints. Each camera comprises three CCD arrays in the focal plane (see fig. 1<sup>b</sup>). The center array cuts the optical axis, i.e. points nadir. The other two are positioned at each side of the center array, such that their viewing angles are  $-11.5^\circ$  and  $11.5^\circ$ , respectively. For land applications just the center array's are used, employing exchangeable filters to attain the required spectral bands. For sea applications all 9 arrays of the downward modules are employed; band 8 is recorded twice. The three arrays of the forward looking module point  $45^\circ$ ,  $52^\circ$  and  $59^\circ$  off-nadir.

The forward looking and downward looking images have to be compared, so they have to be brought into the same reference system. Positional data, consisting of flight velocity and platform attitude, is auxiliary recorded during flight; the platform attitude by an inertial navigation system. The flight data determine the image corrections, which are carried out by the software package OPTIPARES. The images are resampled and rescaled by interpolation. Although the positional data is, in principle sufficient to link forward and downward looking data, the accuracy is just enough to reach a first order approximation. To achieve higher accuracies, tie points together with a geometric transformation model (GTM) have to be introduced. To arrive at this aim and to reduce manual interference we investigated the capabilities of digital matching, subject of the next section.

### 3 The Digital Matching Procedure

As sketched in the introduction matching comprises 3 stages: (1) distinction check, (2) similarity check and (3) consistency check. Distinct or characteristic points are searched by the interest operator described by Moravec (1977). The similarity check is performed by cross correlation. Subpixel level is achieved by approximating the discrete correlation function by a quadratic expression. The consistency check is performed by the B-method of testing (Baarda, 1968). For the present procedure, the assumptions are:

- the object space is a plane, (i.e 2-D)
- the geometric differences between the two images are, after correction with OPTIPARES, smooth;
- the objects are rigid, i.e. they don't change form during successive recordings.

In separate subsections the three steps will be described in more detail.

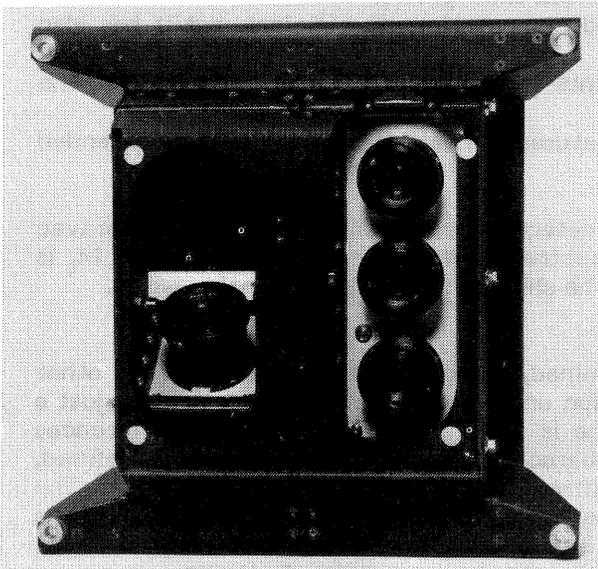


Fig. 1<sup>a</sup> CAESAR camera configuration: three camera's point nadir and one looks forward.

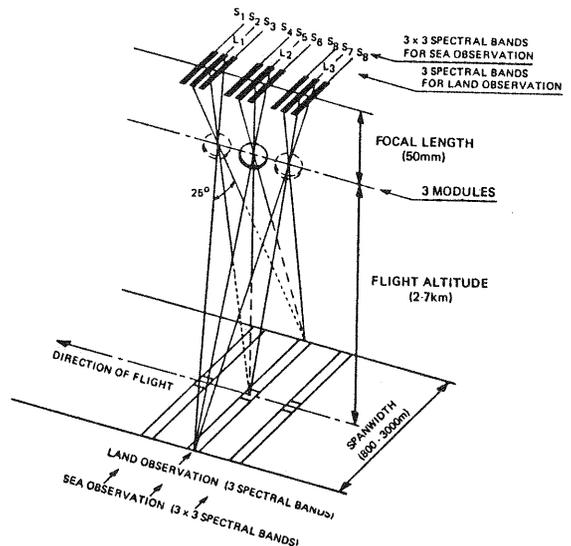


Fig. 1<sup>b</sup> Sketch of the downward looking part. In each focal plane 3 CCD linear array's are mounted.

#### 3.1 Distinction Check

Characteristic points in digital images are automatically traced by interest operators. In the course of time, several operators are developed. Their design depend on the way an image is looked at. Hannah (1974) considers an images as a 2-D signal and defines characteristic points in terms of autocorrelation, characteristic points show a steep autocorrelation in all direction. Moravec (1977) and Förstner (1986) consider the grey values in a statistical manner; characteristic points show high variances in all directions when moving away from the point. Dreschler (1981) looks at a grey value distribution as a continuous surface in a 3-D space. Characteristic points are defined in terms of principal curvatures  $k_1$  and  $k_2$  and the Gaussian curvature:  $K = k_1 \times k_2$ . We adopted the approach of Moravec (1977).

Grey value variances in some directions are computed, in particular the four main directions are chosen: rows, columns and the two diagonals. The variance in each direction is computed as the sum of squared grey value differences between adjacent pixels. The original Moravec operator is defined in a 5 x 5 window:

$$\begin{aligned}
 M_1(i,j) &= \frac{1}{20} \sum_{k=-2}^2 \sum_{l=-2}^1 (g(i+k,j+l) - g(i+k,j+1))^2 \\
 M_2(i,j) &= \frac{1}{20} \sum_{k=-2}^1 \sum_{l=-2}^2 (g(i+k,j+l) - g(i+k+1,j+l))^2 \\
 M_3(i,j) &= \frac{1}{16} \sum_{k=-2}^1 \sum_{l=-2}^1 ((g(i+k,j+l) - g(i+k+1,j+1+l))^2 \\
 M_4(i,j) &= \frac{1}{16} \sum_{k=-2}^1 \sum_{l=-2}^1 (g(i+k,j+1+l) - g(i+k+1,j+l))^2
 \end{aligned}
 \tag{3.1}$$

The operator response  $M$  is defined by:  $M = \min(M_i), i = 1, \dots, 4$ .

If  $M$  exceeds a certain threshold  $M_t$ , the pixel position  $(i, j)$  is a characteristic point. Since characteristic points will cause also high responses in a neighbourhood around the point, non-maximum suppression has to be carried out. For multispectral data the single grey value operator can be modified in two manners (Dreschler-Fischer, 1987):

- the operator is applied to the  $n$  distinct spectral bands, leading to  $M_k, k=1, \dots, n$ . The multispectral response is the maximum of the responses in the distinct bands:  $M_C = \max(M_k)$ ;
- The directional variances are computed for the spectral band vectors. The squared differences of the vectors are calculated. The response should exceed a threshold.

For the CAESAR scanner, the multispectral approach is useful when more equally directed modules are used, i.e. when the terrain is recorded in several spectral bands from the same viewpoint. Also when the matching procedure is applied for mozaicing of digital satellite images, e.g. Landsat MSS and TM, and Spot, the modification will show profit.

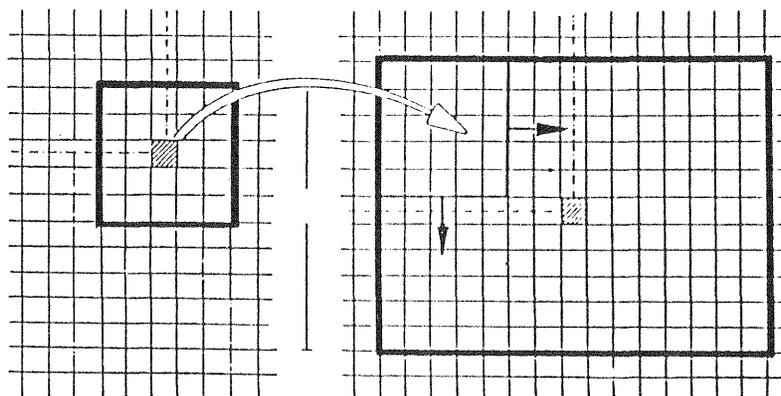
The Moravec operator is easy to implement and computationally efficient, but it has also some disadvantages (Dreschler, 1981):

- not the real positions of the characteristic points are found, but a shift is introduced, the larger the window size, the larger the shift;
- the operator is sensitive to low resolution features, i.e. small points cause an extended operator response;
- the operator is non-rotational invariant.

Fig. 3<sup>a</sup> shows the characteristic points (in white) detected by the single grey value Moravec operator in a CAESAR downward looking image, spectral band  $0.445 \mu\text{m}$ . The threshold  $M_t$  is 2500, which number is experimentally determined. The shift of the points is clearly visible.

### 3.2 Similarity Check

Once characteristic points in one image are determined, the corresponding points in the other image have to be traced. We employ cross correlation under the assumption that there is just a shift between the two images, i.e. the object space is a plane and the geometric differences between the two images are smooth. Around the characteristic point a target area is defined, for instance a  $5 \times 5$  window. The task is to find the corresponding window in the other image. The search space is confined by the positional data using the plane and smoothness assumption. The target area is shifted over the search area (see fig. 2). The cross correlation  $R$  is assigned to the mid-pixel of the target area and the mid-pixel of the concerning window in the search space. This leads to a discrete 2-D cross correlation function. The simplest method is to choose the mid-pixel of the most resembling window as corresponding point. However, just correspondence at pixel level is achieved. To arrive at subpixel level, the  $R$ 's are considered as discrete samples, taken from a continuous cross correlation function. The function is approximated by a second order polynomial. The maximum of the function defines the position which resembles the most the mid pixel position in the target area.



**Fig. 2** cross-correlation: a  $5 \times 5$  target area (left area, the hatched pixel is mid pixel) is shifted over a search space (right part) with increments of 1 pixel. In this manner a discrete correlation function is determined. The maximum of the function indicates the corresponding point (here indicated as hatched pixel).

When  $g_i^t$  denotes the greyvalues of the pixels in the target area (for a  $5 \times 5$  window  $i=25$ ), and  $g_i^s$  those of a search area in search space, the cross correlation  $R_{ts}$  between the target and the search area is defined by:

$$R_{ts} = \frac{\sum_i (g_i^t - \bar{g}^t)(g_i^s - \bar{g}^s)}{\sqrt{\sum_i (g_i^t - \bar{g}^t)^2 \cdot \sum_i (g_i^s - \bar{g}^s)^2}} \quad (3.2)$$

To achieve a computationally more efficient expression (3.2.) is rearranged to:

$$R_{ts} = \frac{\sum_i g_i^t \cdot g_i^s - (\sum_i g_i^t \cdot \sum_i g_i^s) / n}{\sqrt{(\sum_i (g_i^t)^2 - (\sum_i g_i^t)^2 / n) \cdot (\sum_i (g_i^s)^2 - (\sum_i g_i^s)^2 / n)}} \quad (3.3)$$

It is possible to take account of smooth relief variations by introducing a weight function  $w_i$ , which favours the central parts of the window at the cost of the border parts:

$$R_{ts} = \frac{\sum_i w_i \cdot (g_i^t - \bar{g}^t)(g_i^s - \bar{g}^s)}{\sqrt{\sum_i w_i (g_i^t - \bar{g}^t)^2 \cdot \sum_i w_i (g_i^s - \bar{g}^s)^2}} \quad (3.4)$$

The shape of the weighting function  $w_i$  depends on the expected kind of relief and can be, e.g., of linear or Gaussian form. The discrete  $R(i,j)$ 's are approximated by a continuous function  $R(x,y)$ . The maximum  $R_{max}$  at location  $(I_{max}, J_{max})$  and its neighbouring  $R$ 's are used to adjust the parameters of a second order polynomial:

$$\begin{aligned} R(x,y) &= a_0 + a_1x + a_2x^2 \\ R(x,y) &= b_0 + b_1y + b_2y^2 \end{aligned} \quad (3.5)$$

At a local maximum the derivatives to respectively  $x$  and  $y$ :  $dR(x,y)/dx$  and  $dR(x,y)/dy$ , will be zero, which leads to the most likely corresponding point:  $(X_{max}, Y_{max})$ :

$$X_{max} = \frac{-a_1}{2a_2} ; \quad Y_{max} = \frac{-b_1}{2b_2} \quad (3.6)$$

In the present investigation the polynomial is adjusted by  $R_{max}$  and four neighbours, two at each side. The five observations give rise to a least squares adjustment to determine the three unknown parameters  $a_i$  or  $b_i$ ,  $i = 0, 1, 2$ .  $(I_{max}, J_{max})$ , the pixel belonging to  $R_{max}$ , is taken as local origin  $(0,0)$ . So, for  $X_{max}$  computation, the least squares adjustment reads:

$$\begin{pmatrix} R(-2,0) \\ R(-1,0) \\ R(0,0) \\ R(1,0) \\ R(2,0) \end{pmatrix} = \begin{pmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} \quad (3.7)$$

With  $R(0,0) = R_{max}$ . For  $Y_{max}$  computation a similar expression is achieved. (3.7) expressed as:  $x = A y$ , the least squares solution reads:

$$y = (A^t A)^{-1} A^t x$$

Since for all points the same configuration is chosen, the design matrix  $A$  is the same for all points, so  $(A^t A)^{-1} A^t$  can be precomputed to save computation time. Further, for  $(X_{max}, Y_{max})$  determination, also the computation of  $a_0$  is superfluous.

So, applying precomputation,  $(X_{\max}, Y_{\max})$  becomes:

$$X_{\max} = -0.7 \frac{-2 R(-2,0) - R(-1,0) + R(1,0) + 2 R(2,0)}{2 R(-2,0) - R(-1,0) - 2 R(0,0) - R(1,0) + 2 R(2,0)} + I_{\max} \quad (3.8)$$

$$Y_{\max} = -0.7 \frac{-2 R(0,-2) - R(0,-1) + R(0,1) + 2 R(0,2)}{2 R(0,-2) - R(0,-1) - 2 R(0,0) - R(0,1) + 2 R(0,2)} + J_{\max}$$

One of the disadvantages of applying an interest operator on just one image becomes apparently clear when the second image contains substantially less local structure (texture) than the first image, in which the characteristic points are traced. In that case the correlation function is flat and consequently the location of the corresponding point is very insecure. So, as a part of the similarity check, besides the magnitude of the cross correlation coefficients, also the narrowness of the correlation function should be checked. A measure for narrowness is the second derivative, in case of the above quadratic function:  $2 a_2$  and  $2 b_2$ . A large number indicates a steep function. Since  $a_2$  and  $b_2$  are determined separately, a combined measure, such as  $(a_2^2 + b_2^2)^{\frac{1}{2}}$  can be applied. Steepness check is not yet implemented in the present investigation.

### 3.3 Consistency Check

The Moravec operator combined with the similarity check by cross correlation will lead to an initial set of possible matches. Due to, primarily, repetitive structures, the initial set of matches will contain ambiguities, i.e. matches which are equally likely. A global consistency check has to be carried out, using some model. Common methods in stereo vision to perform the consistency check are:

- dynamic programming;
- relaxation;
- robust statistics;

No surface recovery is necessary in our case, just a determination of the geometric transformation parameters. Therefore we suffice with classical hypothesis testing, using the B-method (Baarda, 1968).

The aim of testing is to check the validity of a certain null-hypothesis,  $H_0$ .  $H_0$  states:

- there are no gross errors in the observations;
- the applied GTM is a realistical one;
- the chosen variance-covariance matrix fits the stochastic behaviour of the observations.

Observations errors are often responsible for  $H_0$  rejection. Mismatches are inevitable because of repetitive structures. Also the GTM is a critical point, since the common polynomial models are just a rough approximation of the real imaging process. An insight into the stochastic behaviour of observations can be gained by evaluation of sequences of observations.

To check the validity of the above null-hypothesis a multi-dimensional overall test is carried out. If  $H_0$  is accepted the above assumptions are right with a certain probability and we are ready. If  $H_0$  is rejected first of all the observations are subject of testing. Detected gross errors are removed and to avoid disturbance of the design, new measurements should be added. Next,  $H_0$  is tested again. If  $H_0$  is accepted testing is finished, if not, the GTM is tested. When the GTM appears to be okay, but  $H_0$  is still rejected, the stochastic model has to be checked.

The common GTM's, e.g. affine and second order polynomials are linear expressions. So, without prefatory linearization and without introducing approximate values, the GTM is expressed as a linear system:  $x = A y$ , with,  $x$  the observations,  $y$  the unknowns and  $A$  the design matrix. However, matrix  $A$  will contain observations: coordinates  $(X, Y)$ , determined in image 2 (see also (3.10)). So, for error propagation computation and for carrying out the testing procedure, the observations in  $A$  have to be split up in a stochastic part  $(\Delta X, \Delta Y)$  and a non-stochastic part  $(X^0, Y^0)$ . For this purpose the system should be linearized. But, since the actual GTM will just differ slightly from a shift, after linearization the terms of the linear system can be rearranged such that each  $\Delta X$  and  $\Delta Y$  is additive to the corresponding coordinates in the observation vector. For  $(X^0, Y^0)$  the actual observations  $(X, Y)$  are used. Therefore, it isn't necessary to linearize the expression; the stochastic parts of  $(X, Y)$  are just added to accompanying coordinates in the observation vector. Since it is assumed that the stochastic behaviour of the coordinates, determined in both images, is the same, in the testing procedure the variance factor  $\sigma_0^2$  has just to be doubled.

To improve precision and reliability, more observations are taken than there are unknowns. So, a least squares adjustment can be carried out. For an affine transformation:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} + \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} \quad (3.9)$$

with:  $x, y$ : the coordinates in image 1  
 $X, Y$ : the coordinates in image 2  
 $a_i, b_i, i=0,1,2$ : transformation parameters

the linear system reads for  $n$  observations and  $m$  unknowns ( $m = 6$ ):

$$\begin{pmatrix} x_1 \\ y_1 \\ \vdots \\ \vdots \\ x_n \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & X_1 & Y_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & X_1 & Y_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & X_n & Y_n & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & X_n & Y_n \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ b_0 \\ b_1 \\ b_2 \end{pmatrix} \quad (3.10)$$

(Don't confuse the coordinates  $(x, y)$  and  $(X, Y)$  with the symbols  $x$  and  $y$  for observations and unknowns in the linear system.)

To the observations  $x$  a variance-covariance matrix  $\Sigma_x = Q_x \sigma_0^2$  is assigned, with  $\sigma_0^2$  the variance factor. The weight coefficient matrix  $Q_x^{-1}$  is assumed to be known. When all observations are uncorrelated and equally weighted, i.e. they have the same precision,  $Q_x$  is the identity matrix, and can actually be neglected, leading to the standard least squares solution:  $\hat{y} = (A^t A)^{-1} A^t x$  with  $\hat{y}$  denoting estimate of  $y$ . For hypothesis testing the variance factor  $\sigma_0^2$  is assumed to be known. The observations may have any arbitrary statistical distribution, but for the sake of testing they are assumed to be Gaussian distributed. The least squares solution of the overdetermined system leads to an estimate  $\hat{y}$  of the unknowns  $y$ :

$$\hat{y} = (A^t Q_x^{-1} A)^{-1} A^t Q_x^{-1} x \quad (3.11)$$

$\hat{y}$  is the best linear unbiased estimate (BLUE) for  $y$ . Once  $\hat{y}$  is calculated, the adjusted observations  $\hat{x}$  become:

$$\hat{x} = A \hat{y}; \quad \hat{x} = x + \hat{v} \quad (3.12)$$

The differences  $\hat{v} = \hat{x} - x$  are corrections to the observed values (residuals). From  $\hat{v}$  an estimate  $\hat{\sigma}_0^2$  of the variance factor  $\sigma_0^2$  is computed:

$$\hat{\sigma}_0^2 = \frac{\hat{v}^t Q_x^{-1} \hat{v}}{r} \quad (3.13)$$

with  $r$  the redundancy,  $r = n - m$ ;  $n$  the number of observations and  $m$  the number of unknowns,  $n > m$ .

The precision of the estimates  $\hat{v}$ ,  $\hat{x}$  and  $\hat{y}$  are determined by error propagation. For testing, just the variance-covariance matrix:

$$\Sigma_{\hat{x}} = \sigma_0^2 Q_{\hat{x}} \quad (3.14)$$

and the variance-covariance matrix of the corrections:

$$\Sigma_{\hat{v}} = \sigma_0^2 (Q_x - Q_{\hat{x}}) \quad (3.15)$$

with:  $Q_{\hat{x}} = A (A^t Q_x^{-1} A)^{-1} A^t$

are of importance.

To check the validity of the null-hypothesis,  $H_0$ , a multi-dimensional overall test is performed:

$$\frac{\hat{\sigma}_0^2}{\sigma_0^2} = F_{1-\alpha; r; \infty} \quad (3.16)$$

with  $F$  the critical value of the Fischer distribution and  $\alpha$  the probability that  $H_0$  is rejected wrongly. The other terms are already treated in previous parts.

When  $H_0$  is rejected, obviously one or more  $H_0$  assumptions are erroneous. The most likely source are gross errors in the observations. Therefore, first the observations are tested, using data snooping. Gross error testing is carried out by introducing the conventional alternative hypothesis  $H_{ai}$ , which states that in the entire observation set  $x$ , just one observation  $x^i$  has a gross error. To simplify the procedure further, the observations are supposed to be uncorrelated. The assumption isn't correct, but the approach has shown to be suited for photogrammetric purposes. The test variate  $w_i$  becomes now:

$$w_i = \frac{-v^i}{\sigma_{v^i}} \quad (3.17)$$

with  $v^i$  the correction placed on the observation  $x^i$  and  $\sigma_{v^i}$  the standard deviation of the correction.  $w_i$  has a Fischer distribution and is tested as follows:

if  $(w_i)^2 \leq F_{1-\alpha; 1; \infty}$   
 reject  $H_{ai}$  (i.e. no gross error in  $x^i$ )

else

accept  $H_{ai}$  (i.e. the observation set contains a gross error, remove the observation with the largest  $w_i$ )

endif

When, after gross error removal,  $H_0$  is still rejected, the GTM has to be checked. For instance, the applied affine transformation is replaced by a second order polynomial, and the testing procedure is executed again. If still  $H_0$  is rejected and no gross errors are detected anymore, the stochastic model is probably erroneous. For instance, the weight matrix is not the assumed identity matrix because the observations are correlated or not measured with equal precision.

In the B-method of testing, the choice of the testing parameters  $\alpha$  and  $\beta$  is such that a certain error is detected with the same probability by the overall  $H_0$ -test and the  $w$ -test. So, if  $H_0$  is accepted, no further testing is actually necessary. In practice, however, very often data snooping is performed additionally, regardless whether  $H_0$  is rejected or not.

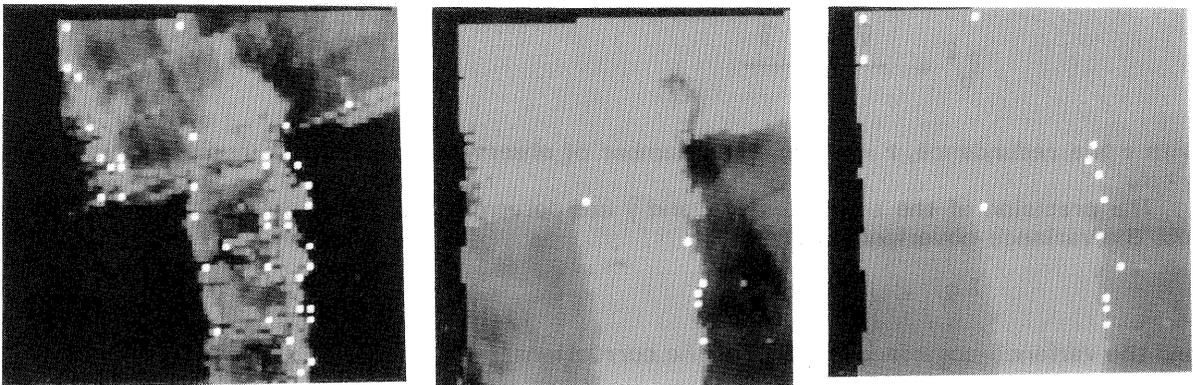


Image 1

Image 2

Image 3

**Fig. 3:** Three downward looking images. Characteristic points are traced in image 1 and searched in the other two images by cross correlation. The white dots in image 1 indicate the characteristic points, the white dots in image 2 and 3 indicate the corresponding points found by correlation.

#### 4 Results

No forward looking images were available during the test period. So, to be able to carry out some tests, three central array images of the downward looking mode are used; spectral bands: 0.445  $\mu\text{m}$  (image 1), 0.630  $\mu\text{m}$  (image 2) and 1.020  $\mu\text{m}$  (image 3). Image  $i$ ,  $i = 1, 2, 3$ , are shown in fig. 3. The imaged area is a lake and its environment ('Loosdrechtse Plassen'). Just image 1 contains moderate contrast; the other images show little local structure.

Since the three images are simultaneously recorded, they should show the same geometry. Deviations are due to erroneous location of the corresponding points. To attain an insight into the deviations, as GTM, a shift  $(t_x, t_y)$  is introduced. Characteristic points are traced in all three images to obtain an indication of the influence of local structure. For each image the corresponding points in the other two images are searched by cross correlation.

The thresholds  $M_t$  on the response of the Moravec operator are chosen such that the number of characteristic points are approximately the same in all three images ( $\pm 40$ ). The thresholds are experimentally determined; for image 1  $M_t = 2500$ , for image 2 and 3  $M_t = 100$ . These numbers indicate already the large difference in local structure. The target area, i.e. the window around the characteristic point, is set to  $5 \times 5$  pixels and the search space in the other image is set to  $13 \times 13$  pixels. The cross correlation coefficients have to exceed the threshold  $R_t = 0.8$  for all images. At present no constraints are imposed on the steepness of the correlation function. The results of the experiments are shown in table 2.

Image 1	$M_t = 2500$	$N_p = 40$	Image 2	$M_t = 100$	$N_p = 49$	Image 3	$M_t = 100$	$N_p = 34$
	Image 2	Image 3		Image 1	Image 3		Image 1	Image 2
$R_t$	0.8	0.8	$R_t$	0.8	0.8	$R_t$	0.8	0.8
$N_c$	6	13	$N_c$	19	16	$N_c$	25	18
$t_x$	3.16	- 0.34	$t_x$	- 1.18	- 1.23	$t_x$	- 0.18	- 5.06
$t_y$	0.53	2.01	$t_y$	- 0.70	- 2.49	$t_y$	- 0.13	2.35

- with:  $M_t$ : threshold on the response of the Moravec operator  
 $N_p$ : number of characteristic points found by the Moravec operator  
 $R_t$ : minimum value of cross correlation coefficient  
 $N_c$ : number of corresponding points found in second image  
 $t_x$ : shift in x-direction  
 $t_y$ : shift in y-direction

Table 2

The sizes of the shifts  $(t_x, t_y)$  indicate that in the present images the correspondence assignment isn't unproblematic. To attain a better insight, the standard deviations  $\sigma_{ij}$ ;  $i = 1, 2, 3$ ;  $j = 1, 2, 3$ ;  $i \neq j$ , of the correspondence assignments are computed;  $i$  refers to the image in which the characteristic points are traced and  $j$  to the other image. The results are listed in table 3.

$\sigma_{ij}$	1	2	3
1	-	7.15	7.67
2	4.55	-	5.83
3	3.40	7.68	-

Table 3

When the characteristic points are detected in image 1, which has the highest local structure,  $\sigma$  is high. Image 3 contains the least contrast and  $\sigma_{31} = 3.40$ , which is the lowest value. So, a first conclusion may be that the characteristic points should be traced in the image with the lowest contrast. But, the correlation threshold (0.8) and even more the negligence of the steepness check of the correlation function, may affect this conclusion.

## 5 Discussion

The procedure is developed for plane regions and land applications only. Cross correlation does only yield a correct match in the case of shift between signals. For regions the relief can't be neglected, more advanced matching techniques must be developed, such as least squares signal matching and feature matching.

In section 3 some drawbacks of the Moravec operator are listed. The Dreschler and Förstner operator doesn't show these disadvantages. Although computationally more costly, their suitability for automatic tie point tracing have to be investigated.

For application at sea, wave movements cause that the object rigidity assumption isn't valid. Moreover there is little texture and there are many repetitive structures which will cause many matching problems. In computer vision, methods are under current development to follow non-rigid objects in time varying image sequences. From a fixed viewpoint a 3-D object space is observed. The objects are moving around. With the CAESAR scanner, a 2-D non-rigid object space is observed from varying sensor positions. Because of the low grey value differences in the images, the first ideas are to employ methods which are suited to handle low textured images. Such methods are based on simulated annealing (Barnard, 1986, 1987). Simulated annealing is founded on the physical phenomenon of annealing a system of molecules to its ground state by reducing temperature. Two conditions have to be fulfilled:

- some function of the greyvalues in image  $G_1$  and  $G_2$  has to be optimized, e.g.:

$$\sum_i \sum_j (G_1(i, j) - G_2(i + k_i, j + l_j))^2 = \min \quad (5.1)$$

with  $k_i$  and  $l_j$  the disparity (=shift) between pixel  $(i, j)$  in  $G_1$  and the corresponding pixel in  $G_2$ .

- the disparities vary smooth, i.e. some smoothing measure of the disparities  $S(k_i, l_j)$  has to be optimized, e.g.:

$$\sum_i \sum_j (S(k_i, l_j))^2 = \min \quad (5.2)$$

with  $S(k_i, l_j)$ , for instance, the sum of the absolute differences between the disparities of a pixel and the eight neighbouring pixels. Both conditions can be incorporated into one decision function  $E$ , which has to be minimized:

$$E = \sum_i \sum_j (G_1(i, j) - G_2(i + k_i, j + l_j))^2 + \alpha (S(k_i, l_j))^2 \quad (5.3)$$

with  $\alpha$  a weighting factor. To save computation time  $E$  can be modified by replacing the quadrates by absolute values. Following the procedure of simulated annealing, a random state is chosen, i.e. to each pixel in  $G_1$  a pixel in  $G_2$  is arbitrarily assigned.  $E$  is computed. A new random state is chosen. States with lower  $E$  are accepted as better solutions. After many iterations the minimum state is reached, defining  $k_i$  and  $l_j$ .

To reduce search space, the epipolar geometry is applied and the maximum amount of possible disparity is limited. The method reveals a dense map of corresponding points, but is computationally costly. It can be refined, saving computation time, by a coarse to fine approach in an hierarchical way. Matches at a coarse resolution, guide matching at fine resolutions. Whether simulated annealing or one of its modifications is fruitful for matching CAESAR sea images, has to be subject of thoroughly investigation.

In summary the further research items are:

- suitability test of other interest operators;
- extension of the single grey value Moravec operator to the multispectral case;
- determination of the most appropriate threshold for cross correlation;
- determination of the lower bound on steepness of the correlation function;
- the suitability of other, more simple correlation measures, like the sum of the absolute values of the differences, to reduce computation time;
- implementation of a testing procedure based on the B-methode of testing, to detect mismatches and to check to assumed geometric transformation model;
- determination of a suitable geometric transformation model for CAESAR images;
- determination of the matching precision and its adequate measure;
- optimal distribution of tie points;
- extension of the procedure for sea applications.

## 6 Conclusions

A matching procedure to trace automatically corresponding points in CAESAR images is presented. Characteristic points are traced by the Moravec operator in one image. With cross correlation the corresponding points in the other image are traced. The search space is size limited by an approximate known exterior orientation and assumption of flatness of terrain. Mismatches are detected by the B-method of testing. The procedure is confined to the land mode of CAESAR and assumes that the area is approximately a plane.

The presented procedure is under current development. A part of it is already implemented on the NLR RESEDA system. The procedure has to be further extended and improved. The method is probably also suited for the determination of tie points in overlaying image parts of satellite images of areas with rare ground control points for mozaicing. For applications, for which the procedure isn't suited, because, e.g. the assumptions don't hold, other, more sophisticated procedures have to be exercised, such as least squares signal matching and feature matching.

## References

**Baarda, W., 1968**, A testing procedure for use in geodetic networks, Neth. Geod. Comm. publ. on geodesy, new series, vol. 2., nr. 5., Delft.

**Barnard, S.T., 1986**, A stochastic approach to stereo vision, SRI International, Technical note 373.

**Barnard, S.T., 1987**, Stereo matching by hierarchical microcanonical annealing, SRI international, Technical note 414.

**Bunnik, N.J.J., Pouwels, H., Smorenborg, C., Valkenburg A.L.G. van, 1986**, CAESAR: CCD Airborne experimental scanner for applications in remote sensing, ISPRS symp. comm. VII, Enschede, pp. 207-210.

**Dreschler, L., 1981**, Ermittlung markanter Punkte auf der Bildern bewegter Objekte und Berechnung einer 3-D Beschreibung auf dieser Grundlage, Ph.D. Thesis, University of Hamburg.

**Dreschler-Fischer, L.S., 1987**, A blackboard system for dynamic stereo matching, in: Hertzberger, L.O., Groen, F.C.A. (edt.) Intelligent Autonomous Systems, Elsevier Science Publishers B.V., Amsterdam, pp. 189-202.

**Förstner, W., 1986**, a Feature based correspondence algorithm for image matching, Int. Arch. of Photogr. vol. 26-III, Rovaniemi, pp. 1-17.

**Hannah, M., 1974**, Computer matching of areas in stereo images, Ph.D. Thesis, University of Stanford.

**Lemmens, M.J.P.M., 1988**, A survey on stereo matching techniques, 16th ISPRS Congress, Comm. V/6, Kyoto, Japan.

**Moravec, H.P., 1977**, Towards automatic visual obstacle avoidance, Proc. 5th Joint Conf. Art. Intell., Cambridge, p. 584.