Automatic Point Transfer: a practical application of optimal digital image matching based on local invariant properties

X. Xu, ZG. Tan*
Graz City Council, Dept. 10/6

Commission III, WG III/2

Key Words: invariant properties, local descriptor, optimal matching, automatic point transfer

Abstract

One practical example in digital photogrammetric production is that if one point in the new image is chosen, the same position in old digital image should be found and its geodesic coordinates be transferred to the new one automatically. Another well known example is to automatic produce digital terrain model from stereo pair. One of the key steps of automatic point transfer or automatic DTM production from aerial photos is introducing a robust matching algorithm. In our case, robust means even in tatteried area it should give out satisfied results. In this paper a feature-based matching method which depends on locally invariant properties is presented. The characteristics of this algorithm are the first, the matching inputs are the radiometric and geometric noise invariant properties of image patches; the second, the locally matching inputs are optimally used for extrinsic matching procedure. Some practical results are presented.

1. INTRODUCTION

1.1 The goals

Generally after every aerial photography it is necessary to follow the aerotriangulation procedure so that the new photos can be used in the production. Our experience told us the aerotriangulation was a time- energy consuming job, even with optimal working procedures and excellent adjustment program /Ganster1993/, /Ganster1994/. For a city like Graz, usually in every 3 ~ 4 years one new aerial photographic flight is to be carried out. So in the point of view of economic aspects, it will be very interesting if the aerotriangulation can be removed from the production chain. Now slowly digital image matching techniques show their practical marvelous sides. If one position in new image is defined by operator, the same position in old image(s) which were taken several years ago and absolute orientation parameters are known should be found and the geodesic coordinates can be transferred to the new one. Let's name this operation as *automatic point transfer*. On the base of one high accurate aerotriangulation results of the old image sets and available digital photos (the both sets of digital photos will be used for Orthophoto production), automatic point transfer is exact the right tool for this purpose.

It is true that in some areas under special conditions the DTM can be driven from digital stereo pair by computation. But in the typical european city Graz, it is almost unimaginable at now to get DTM (urgent demand for orthophoto production) full automatically. What we try now is that let feature-based matching results (or combined with other matching methods) make the measurement of DTM much more easier for human operators, especially in the tattered areas like quarry field or down town. The direct application is to use the matching results as prepositions for DTM measurement.

1.2 Overview of different methods

Before this matching method was developed, lots of existed methods were studied carefully /Xu1994/. Here let's have a simple overview. Due to the complexity and severity of different matching tasks, a great mounts of matching algorithm have been developed during last several decades. Each kind of them deals with special aspect of the problems. They can be generally sorted into following four types:

1.) signal based matching /Hannah1988/, /Dowman1977/, /Ackermann1983/, /Benard1986/; 2.) low-level feature based matching /Moravec1977/, /Foerstner1987/; 3.) high-level feature based matching /Shapiro1980/, /Cheng1985/; 4.) hybrid matching /Jordan1991/, /Hsieh1992/, /Xu1993/

Signal based matching methods (They are also named as area-based matching) can produce results with very high accuracy on the basis of rather precise initial values and on the cost of computation time. Low-level feature based matching is sometime used as first step of whole matching procedure to supply the initial values for signal based matching. For this aim, the methods should be qualitative robust, i.e. matching results with the same quality should be returned even with different image contents. In application of aerial photos in european city areas there are few methods which can produce the satisfied initial values for signal based matching, e.g. least square matching. High-level feature based matching has the difference with low-level one in the way that it takes the relations between

ZG. Tan* : Institute for Applied Mathematics, Tsinghua University, PR. China

the features into account in various forms. The difficulty of relation definition and time consuming computation make this kind of methods unpopular in the field of photogrammetry. The hybrid matching methods have almost the same disadvantages.

1.3 General idea of our matching method

We developed a robust method that belongs to the second sort and is suitable to be used in application of aerial photographs in tattered areas. The following strategies are engaged in order to guarantee the robustness:

- the epipolar geometry reduces the searching area from two dimensions to one;
- canny operator is used to extract the edges;
- the matching positions are located on where the edges run through epipolar lines almost perpendicularly;
- centred at such a crossing point, geometrical and radiometric noise invariant properties are extracted out either along the edge patch or around the point;
- these properties are used as inputs for the bipartite weighted matching for each pairs of epipolar lines /Tan1996/.

1.4 The experimental environment

The experimental photos were taken with WILD RC10 and RMK-TOP from Zeiss, and the photoscale is about 1:3500 to 1:3700. The digital images were obtained with *Microtek Scanmaker 600zs* of resolution of 600 dpi and Vexel scanner with 20 μm pixel size. The main computer for the whole processing is *Silicon Graphics Indigo²* (CPU: MIPS R4400, OS: IRIX5.3). The working language is IDL4.0 (Interactive Data Language, Version 4.0).

2. Matching based on local invariant properties

2.1 Locating of searching positions

The first step of the whole matching procedures is to locate the possible matching positions, i.e. searching positions. This is also a common step for all feature-based matching methods. For automatic point transfer, this step is somehow easier. The target area is defined by operator, and generally must be with unequivocal features. And the searching positions in other image(s) can be selected with identical features. In our second case, together with Canny edge detection algorithm, the advantage of application of epipolar geometry is profited. By mean of geometrical transformation, the scanned digital images are transformed into normal case, in which the epipolar lines are re-arranged parallel to each other. The homologous epipolar lines can be put in the same position in y-direction. In this way the matching will be carried out in x-direction. The detailed descriptions for these operations can be found in /Brandstaetter1992/, /Xu1994/.

It is advantageous to apply the epipolar geometry in the matching procedure. It reduces the searching areas from 2D to 1D. But at same time, the features which available for matching is also reduced from areas to lines if only the information along epipolar lines is used. One logical solution is to define small window along the epipolar lines. On the other hand due to the different camera orientations and different flight heights, it is very difficult to define the form and size of the window in the searching area when the target window is defined with certain size and form. When entire information inside the window be used, only these differences could mislead the matching procedures later. The solution in our case is that only the crossing points of epipolar lines and edges that run through the epipolar lines semi-perpendicularly will be used as matching positions. Around these positions the features either along edges or along epipolar lines will be extracted out and be applied for matching.

After comparison and analysis of different edge detection methods, the canny edge detection algorithm is applied due to its outstanding anti-bias property.

2.2. Extraction of local invariant properties

After determination of the matching candidate positions, the next problem is to setup the matching inputs from the matching candidates. And this problem is considered to be difficult as the matching procedure itself. What we want using for the matching inputs are those properties that are satisfied the following conditions:

- they should be insensible to either geometrical or radiometric noises;
- they should be invariant to the rotation, transformation, and scaling as much as possible;
- they should be able to distinguish the differences among the neighbouring edge segments;
- and at last but not least, they should be local properties.

2.2.1 Geometrical properties The edge polarity of the matching position is the first element to be used. The edges can be divided into two types, namely, positive and negative one. The edge which gray values change along positive x-direction from higher to lower is defined as negative edge, otherwise it is positive. With zero-crossing theory (LoG operator) /Marr1982/ it is easy to get that. In this way the matching candidates are grouped into two types, i.e. the probability of correct matching will be risen to double high.

Another geometrical noise invariant property is either local Fourier descriptor or the curve signature of the edge segment. One of them gives out the rotation and transformation invariant property in frequency domain while the other describes such a property in space domain. The power spectrum of one close edge segment $s(t)=x(t)+jy(t)$ on the interval [0, k] can be written as

$$p(u) = |a(u)|^2 = \left| \frac{1}{k} \sum_{t=0}^{k-1} s(t) \cdot \exp\left(-\frac{j2\pi t}{k} u \right) \right|^2 \qquad (2.1)$$

For an opened edge segment, if we trace once forward and then retrace backward, a closed edge is driven. The local Fourier descriptor (power spectrum) can be obtained with formula (2.1) in this way.

From the theory of differential geometry /Stoker1969/, a smooth curve can be uniquely reconstructed within a rigid motion using three geometrical invariances, i.e. the arc length ζ, curvature $k(\zeta,)$, and torsion $\tau(\zeta,)$. As we know, these values are too much influenced by the discreteness of the digitization. Specially for the higher order derivatives, their numerical stability is very doubtful /Rosenfeld1982/, /Gonzalez1992/. In order to use them in digital image processing procedure, some stabilization strategies may be applied. In our case we avoid to compute the second derivative directly, but use the differences of neighboring tangent vectors on the curve to get out the approximated curvature. And we use it as curve signature:

$$k = d\theta \approx dl = \sqrt{\left(\frac{a_1}{\sqrt{a_1^2+b_1^2}} - \frac{a_2}{\sqrt{a_2^2+b_2^2}}\right)^2 + \left(\frac{b_1}{\sqrt{a_1^2+b_1^2}} - \frac{b_2}{\sqrt{a_2^2+b_2^2}}\right)^2} \tag{2.2}$$

where, $\mathbf{V}_1=a_1 \cdot \mathbf{i} + b_1 \cdot \mathbf{j}$ and $\mathbf{V}_2=a_2 \cdot \mathbf{i} + b_2 \cdot \mathbf{j}$ are two neighboring tangent vectors on the curve.

2.2.2 Radiometric aspects The following two radiometric noises invariant properties can be used together or individually. They reflect the radiometric properties of the edge segment in x and y direction respectively.

The one in y-direction can be described as following model: The gray-value of the edge segment in the window is $g(i+p, j+q)$ where (i, j) are the coordinates of the crossing point. At the same time they are also the coordinates of the window center, and $\{p, q \in \text{window}\}$. The gray-value of the neighboring pixels of this edge segment are $g(i+p+l, j+q)$, where l is the edge step. First step is to eliminate the image gray-value independent part of noise, for example for edge with two pixel step:

$$\Delta g(p, q) = g(i+p+2, j+q) - g(i+p, j+q) \quad \{p, q \in \text{window}\} \tag{2.3}$$

The second step is to delete the image gray-value proportional noise,

$$R_y(i,j) = \Delta g(p',q') / \Delta g(u,v) \qquad \{(u,v) \in (p,q),(p',q') \notin (u,v)\} \tag{2.4}$$

$R_y(i, j)$ reflects almost the nude gray-value property of the cross point at (i, j) along the edge segment inside the window. Due to the fact that only those edges are to be used when they run through the horizontal epipolar line semi-perpendicularly, we say that $R_y(i, j)$ is the radiometric noise invariant property of the point (i, j) in y-direction. Together with LoG (Laplacian of Gaussian) operator, we can define the edge steps for different edge types, for example step, ramp, roof, spike edge etc.

In direction of epipolar line, we have to introduce weight index for the gray-value ratio map. If the position of the crossing point, say at (i, j), there are two neighboring points along the same epipolar line outside the edge range. Their coordinates are (p, j) and (q, j) respectively, where $0 \le p < i$ and $i < q \le x$-dimension. But in our case we select these two points mostly inside the window, that means $p_w \le p < i$ and $i < q \le q_w$, $\{p_w, q_w \in \text{window}\}$

$$R_x(i,j) = \left(\frac{q-i}{q-p}\right) \cdot \left(\frac{g(i,j)-g(p,j)}{i-p}\right) + \left(\frac{p-i}{q-p}\right) \cdot \left(\frac{g(q,j)-g(i,j)}{q-i}\right) \tag{2.5}$$

where $g(i, j)$ is the gray-value at the point (i, j).

2.3 Management of the different properties

Along two epipolar lines we find out sets of searching positions. At every position, there are several geometrical and radiometric noises invariant properties. Together we name them as local descriptor. The first element of local descriptor is edge polarity. With this zero-one element, we select the matching candidates into two parts. Inside each part, we do the following operations:

First of all one matching table based on one element set of the local descriptor will be setup. The other element sets will be used as support-index in form of weight matrix added to this ground matching table. For example, the weight matrix for R_x can be derived like following:

$$w_{uv}(j) = \min\left\{\frac{R_x(u,j)}{R_x(v,j)}, \frac{R_x(v,j)}{R_x(u,j)}\right\} \tag{2.6}$$

at j position in y direction, u is x-position of one line and v is the one the other line.

2.4. Optimal extrinsic matching

Now the matching procedure itself should satisfy the following two conditions:

1. based on the local invariant properties, the matching process should be a global one;
2. it should be processed optimally in order to avoid the rest local ambiguous parameter.

Along this direction, the bipartite weighted matching with conserved order is applied in our procedure. Detailed properties of this matching method is presented by /Tan1996/. Here we describe its mathematics model.

There are two ordered data sets, denoted by

$\boldsymbol{L} = \{1, 2, 3, ..., m\}$
$\boldsymbol{R} = \{1, 2, 3, ..., n\}$.

Let the weight of matching between element i in the set $\boldsymbol{L}$ and element j in the set $\boldsymbol{R}$ be w_{ij}, and the weight matrix $\boldsymbol{W}$ be given.

Let

$$I = \{i_1, i_2, \cdots, i_{k_1}\} \subseteq L \cup \Theta$$
$$J = \{j_1, j_2, \cdots, j_{k_2}\} \subseteq R \cup \Theta$$

(Θ means empty set) be two ordered subsets, i.e. there exist the relations $1 \le k_1 \le m$ and $1 \le k_2 \le n$. A possible matching set between $\boldsymbol{L}$ and $\boldsymbol{R}$ is denoted as

$$\Phi = \{\phi(\mathrm{I},\mathrm{J}) : \forall \mathrm{I} \subseteq \mathrm{L} \cup \Theta, \mathrm{J} \subseteq \mathrm{F} \cup \Theta\} \tag{2.7}$$

where:

$$\phi(\mathrm{I},\mathrm{J}) = \begin{pmatrix} i_1 & i_2 & \cdots & i_p & \cdots & i_{k(\phi)} \\ j_1 & j_2 & \cdots & j_p & \cdots & j_{k(\phi)} \end{pmatrix} \tag{2.8}$$

(i_p, j_p) is one correspondent pair, and the objective function at $\phi(I,J)$ can be written in (2.9):

$$Obj(\phi) = \sum_{p=1}^{k(\phi)} w_{i_p j_p} \tag{2.9}$$

The mathematical model of the bipartite weighted matching algorithm can be expressed as: We find out the optimal matching

$$\phi^*(\mathrm{I}^*, \mathrm{J}^*) = \begin{pmatrix} i_1^* & i_2^* & \cdots & i_{k^*}^* \\ j_1^* & j_2^* & \cdots & j_{k^*}^* \end{pmatrix} \quad (2.10)$$

such that

$$Max\left\{ \sum_{p=1}^{k(\phi)} w_{i_p j_p} : \phi(\mathrm{I},\mathrm{J}) \in \Phi \right\} = \sum_{p=1}^{k^*} w_{i_p^* j_p^*} \quad (2.11)$$

Now this model is used by pairs of scan lines. The application in the whole image at one time should be possible.

3. Experimental results

3.1 Results of automatic point transfer

The old image set was taken in 1989 with Wild camera RC10. The focal length is 213.91 mm. Photoscal is 1:3700 and scanned by Vexel scanner with resolution of 20 μm. The new image set was photographied in 1992 by Zeiss RMK-TOP 30/23 camera at about 1000 m high. The focal length is 305.412 mm. The images were also digitized by Vexel scanner with 1200 dpi.

After necessary preparation (automatic interior orientation and image resampling to epipolar geometry), the operator should choose the positions which has more features in the image. Anyway if there are great differences between the two image sets, the results will be very poor even no back position. Fig.3.1.1 gives out one example

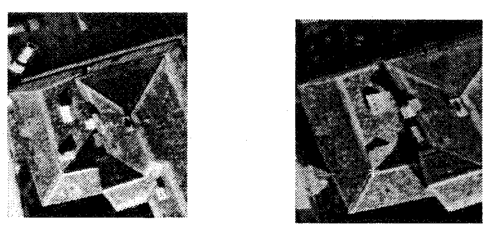

Fig.3.1.1 The two image sets with found point '*'

3.2 Results of DTM production

In order to check the robust ability of the of the algorithm, some tattered areas were chosen. Fig.3.2.1 and Fig.3.2.2 show some area in quarry field and its matching results. The matched points are marked with white points.

Unlike automatic point transfer, the returned matching point will be displayed. The operator can check them one by one with stereo viewing. By DTM computation, in order to have a qualitative checking, the digital terrain model of the same area was measured manually. Or like that in our case some amount of well-defined positions are chosen and checked on analytical plotter. In this example thirty well-defined feature points were picked out from both models for the comparison. The table 3.2.1 shows the checking results.

Table 3.2.1

	Error < 1 pixel	Error < 2 Pixels	Error > 2 pixels
Point Number	25	3	2
Percent	83%	10%	7%

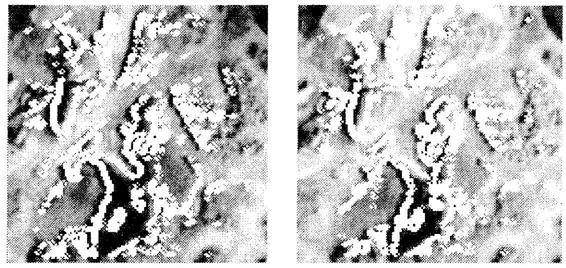

Fig. 3.2.1 Matching results in quarry field

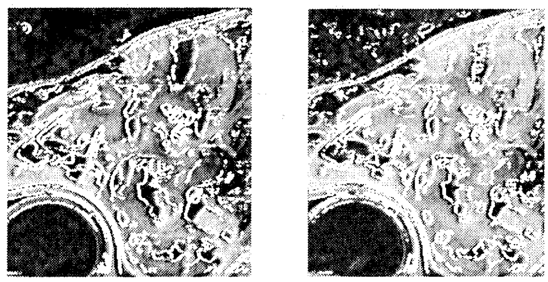

Fig.3.2.2 Matching results in mixed area

4. Discussion

4.1 About automatic point transfer

At presents there are some limitations about this procedure. First of all the positions of project centers of the same image areas between two flights should not be very far away. We have that for our example about 200 ~ 300 m. For the new flights it should not be far away of ± 80 m from each other. Beside this the flight height, the flight date and time, the camera of the two aerial photographies should be similar as much as possible. The resolutions of digital images can be the same, can also be differently.

Such a technique has a very bright future for daily photogrammetric production. As in our case, it can save aerotriangulation in every 3 ~ 4 years without any accuracy loss.

4.2 About DTM production

The matching method described in this paper is suitable for the different types of aerial photos, different image contents except that image with no features. Due to the fact that only those local geometrical and radiometric noise independent properties have been used as matching representations, even though the stereo pair are quite different, the global bipartite weighted matching can also get satisfied results. Here we have to point out that the noise means not only the real one but also those image differences within some degrees. The robustness of the method is tested in the region of quarry field and with lots of tattered natural features.

The direct application of this method is to use the matching results guiding operator for the DTM measurement. Together with signal based matching (in development) one robust automatic 3D model generation algorithm should be there which able to meet the requirements of daily production.

/Gonzalez1992/: R.C.Gonzalez, R.E.Woods, "Digital image processing", Addison-Wesley publishing company, 1992
/Hannah1988/: M.J.Hannah, "Digital stereo image matching techniques", International archives of Photogrammetry and Remote Sensing, Vol.27, No.3, 1988
/Hsieh1992/: YC.Hsieh, et al, "Performance evaluation of scene registration and stereo matching for cartographic feature extraction", PAMI-14, No.2, 1992

The following image pair show the matching results, where the white points are the matched positions.

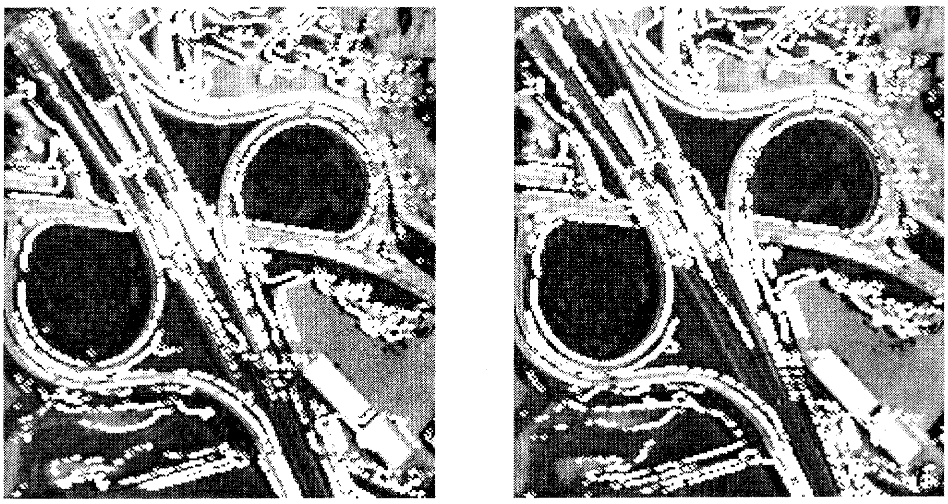

Fig.3.3 Aerial image with matching results (white points)

5. References:

/Ackermann1983/: F.Ackermann, "High precision digital image correlation", proceeding of 39. photogrammetric week, Stuttgart, 1983
/Benard1986/: M.Benard, et al, "Automatic stereo photogrammetry: implementation and comparison of classical methods and dynamic programming based techniques" International Archives of Photogrammetry and Remote Sensing (ISPRS), Vol.26, No.3/3, 1986
/Brandstaetter1992/: G.Brandstaetter, "Note on the direct projective transformation of general pairs into the rigorous normal case by image correlation", XVII ISPRS Congress. Comm.III, Washington D.C, 1992
/Cheng1985/: YC.Cheng, SY.Lu, "Waveform correlation by tree matching", PAMI-7, No.3, 1985
/Dowman1977/: I.Dowman, A.Haggag, "Digital image correlation along epipolar lines", Proc. International Symposium on image processing, Graz, 1977
/Foerstner1987/: W.Foerstner, E.Guelch, "A fast operator for detection and precise location of distinct points, corners, and centres of circular features", ISPRS, Intercomm. Workshop, Interlaken, 1987
/Ganster1993/: W. Ganster, X. Xu, “Aerotriangulation 1. Hauptblock”, Technischer Bericht des Magistrates Graz, Abt.10/6, 1993
/Ganster1994/: W. Ganster, X. Xu, “Aerotriangulation Graz”, Technischer Endbericht des Magistrates Graz, Abt.10/6, 1994
/Marr1982/: D.Marr: "Vision", Freeman, 1982
/Moravec1977/: HP.Moravec, "Towards automatic visual obstacle avoidance", Proc. of 5th IJCAI, 1977
/Rosenfeld1982/: A. Rosenfeld, A.C. Kak, "Digital picture processing", Academic press, New York, 1982
/Shapiro1980/: LG.Shapiro, "A structural model of shape", PAMI-2, No.2, 1980
/Stoker1969/: J.J.Stoker, "Differential geometry", Wiley interscience, 1969
/Tan1994/: ZG.Tan, G.Brandstaetter, X.Xu, "A method for solving the inverse problem of photogrammetry", ISPRS, Intercomm. Conference, COMM.III, Munich, 1994
/Tan1996/: ZG.Tan, X.Xu, "The bipartite weighted matching with conserved order and its application in image processing", to be published, 1996
/Xu1993/: X.Xu, G.Brandstaetter, "A hybrid stereo algorithm for photogrammetric image matching", <<Optical 3-D Measurement Techniques>> by Gruen/Kahman, Zurich,1993
/Xu1994/: X.Xu, "Digital image matching procedure based on local descriptor", Ph.D thesis, Graz University of Technology, Graz, 1994
/Xu1995/: X.Xu, G.Brandstaetter, "Optimal digital stereo image matching based on local descriptor", 40. Internationales Wissenschaftliches Kolloquium, Technische Universität Ilmenau, 1995, Ilmenau