LEAST SQUARES MATCHING BY SEARCH

Tapani Sarjakoski and Jussi Lammi
ESPA Systems Ltd.
Tekniikantie 12, FIN-02150 Espoo
Finland

ISPRS Working Group III/2

KEY WORDS: Digital Aerial Imagery, Least Squares Matching, Line Matching, Area Matching, Search Techniques

ABSTRACT

The paper studies a method for solving the least squares matching by search. The basic principle of the method is straightforward: 1) the object function (the minimum sum of squared residuals) is expressed in terms of original, non-linearized equations, 2) search space is defined in terms of unknown parameters and 3) the object function is solved by means of numerical analysis using search.

The paper begins with a general formulation of the least squares object function in terms of object-space gray value matching and non-linear observation equations. Bilinear interpolation of the gray level values is used as a method for achieving subpixel accuracy. A search method based on a predetermined search space and complete search is described in its generic form. Usage of hierarchical search methods and case-dependent knowledge are introduced for making the search efficient. Supersampling of gray level values with bilinear interpolation is used as a preprocessing procedure for making the computations simple in the innermost loop.

The method as such is independent of the application area. The treatment taken in this paper is based on cases where aerial images are used. The case-dependent heuristic search methods are introduced for image to multi-image matching when the image orientations are known either precisely or approximately. The cases for height determination and template matching for signalized points in aerial triangulation are also covered. It is shown that the effect due to the terrain tilt can be handled rigorously using only two parameters. This is essential for keeping the size of the search space as small as possible.

It is concluded that the applicability of the least squares matching by search is as wide as the applicability of least squares matching in general, ranging from aerial triangulation to digital elevation computation. It clearly has the benefit of being less sensitive of approximate values. Its straightforward formulation yields a straightforward implementation. It is also straightforward to implement other optimality criteria, like the use of L_1-norm. Its computation cost is tolerable in practice, even when used as a part of interactive measurement methods. In addition, the search-based method is directly suitable for multithreaded programming techniques to utilize multiprocessor workstations in a highly efficient way.

1. INTRODUCTION

Least squares matching or correlation is known as one of the best methods for image matching on subpixel accuracy. It also provides a unified optimization criterion for matching on multiple images. There are several variants of least squares matching that are based on conventional least squares adjustment, including use of approximate values and solving a linearized equation system. For convergence to a correct minimum, these methods are critically dependent on approximate values in the range of few pixels, expressed in image scale.

This paper studies a method for solving the least squares matching by search. The basic principle of the method is straightforward: 1) the object function (the minimum sum of squared residuals) is expressed in terms of original, non-linearized equations, 2) search space is defined in terms of unknown parameters and 3) the minimum of the object function is found by means of numerical analysis using search. Bilinear interpolation of the gray level values is used as a method for achieving subpixel accuracy. Although the field of numerical analysis identifies several methods for solving systems of non-linear equations by iteration or search, the techniques used in this study are mainly originated from heuristic search techniques used in artificial intelligence.

2. BASIC FORMULAS

The basic formulation for least squares matching has been introduced simultaneously by Förstner (1982) and by Thurgood and Mikhail (1982), further refined by Ackermann (1984) and adapted for multi-image matching by Grün (1985). Those formulations are based on linearized equations and conventional techniques for solving and otherwise treating overdetermined linearized equations systems by means of least squares adjustment. Probability theory for linear or linearized equation systems can therefore be applied directly.

Least squares matching can also be expressed in terms of non-linear functions. For matching two continuous and two dimensional gray level functions we have the following formula

$$\int_{x_1}^{x_2}\int_{y_1}^{y_2}\left[g_1(x+\Delta x, y+\Delta y)-g_0(x,y)\right]^2 dydx = \min! \quad (1)$$

where

$g_0(x,y)$ a template or a reference image

$g_1(x,y)$ the image to be matched

x_1, x_2, y_1, y_2 lower and upper bounds of the window in the reference image

$\Delta x, \Delta y$ the **unknown** shifts

For multi-image matching, when one of the images is kept as a reference image, the basic formulation becomes

$$\sum_{k=1}^{l}\int_{x_1}^{x_2}\int_{y_1}^{y_2}\left[g_k(x+\Delta x, y+\Delta y)-g_0(x,y)\right]^2 dydx = \min! \quad (2)$$

where

$g_0(x,y)$ a reference image

l number of images in the match set (the total number of images is $l+1$)

Formulas (1) and (2) are expressed in a simplified form so that there are no free parameters for radiometric or geometric corrections. For modelling the geometric differences of the images rigorously, it is most straightforward to use object space formulation (Wrobel, 1987; Ebner and Heipke, 1988; Helava 1988; Heipke 1992):

$$\sum_{k=0}^{l}\int_{X_1}^{X_2}\int_{Y_1}^{Y_2}\left[g_k\left(S\left(X,Y,Z(X,Y,\mathbf{p}),\mathbf{r}_k\right)\right)-G(X,Y)\right]^2 dydx = \min! \quad (3)$$

where

X,Y,Z object coordinate system

$S(X,Y,Z,\mathbf{r}_k)$ geometric transformation function from the object coordinate system (X,Y,Z) to the image coordinate system $(x,y)_k$

$\mathbf{r}_k$ an (approximately) known vector of the orientation parameters for each image k

X_1, X_2, Y_1, Y_2 lower and upper bounds of the match window (patch) in the object space

$Z(X,Y,\mathbf{p})$ an elevation function for the patch in the object space

$\mathbf{p}$ an **unknown** vector of the parameters for the elevation function

$G(X,Y)$ an **unknown** fictitious gray level function of the patch in the object space

Formula (3) is the corner stone for this paper. It defines the object function and free functions $Z(X,Y,\mathbf{p})$ and $G(X,Y)$. The type of these functions have to be defined to model the physical reality with sufficient fidelity. Thereafter the numerical parameters related to the model can be estimated by means of numerical analysis. Throughout of this paper we assume that orientation parameters are known for each image, at least approximately.

For the purposes of further elaboration, a modified version of formula (3) is:

$$\sum_{k=0}^{l}\sum_{i\in A}\left[g_k\left(S\left(X_i,Y_i,Z(X_i,Y_i,\mathbf{p}),\mathbf{r}_k\right)\right)-G_i\right]^2 = \min! \quad (4)$$

where

A a set of groundels used for match

X_i, Y_i the planimetric object coordinates of the centre point of the groundel i

G_i the **unknown** fictitious gray level value of the groundel i

Here a rectangular match window in the object space is replaced with a finite set of groundels that usually form a topologically connected area The gray level function has been replaced with simple variables.

In computer implementations, function $g_i(x,y)$ is presented with an integer valued function, $I_k(line, column)$, where *line* and *column* are also integer variables. Using this as input, the numerical values for $g_i(x,y)$ are available by means of bilinear interpolation using the neighboring pixels. The use of bilinear interpolation is essential if we want to stay in accordance with the original least squares matching (Förstner, 1982). This can by justified by noticing that both methods use gray value differences of neighboring pixels for approximating the gray level function locally. The technique is used for reaching sub-pixel accuracy.

3. LEAST SQUARES MATCHING BY SEARCH

Search methods are based on the idea that a solution to a problem can be found by searching a finite number of states and that a criterion is given for evaluating the 'goodness' of each state. This criterion defines the goal state, i.e. the state to stop the search. When applied to numerical analysis and especially to least squares adjustment this means that the search space is defined in terms of the unknown parameters, whereas the object function (here the minimum sum of squared residuals) defines directly the goal state.

Search-based techniques rely on integer programming and therefore the unknown parameters have to been discretized. Regarding formula (4), a finite search space consists of the discretized values of $\mathbf{p}$ and $G_i, i\in A$. This assumes that realistic lower and upper bounds are available for each scalar component. The goal state (where the sum of the squared residuals is minimum) would always be found by complete search. However, for computer implementation further development is necessary to keep the computational load tolerable.

4. MAKING SEARCH EFFICIENT

The efficiency of any search-based method is dependent on three factors: the total size of the search space, the efficiency of pruning techniques for reducing the search space, and the effort required to evaluate a single state in the search space.

4.1 Size of the search space

The total size of the search space can be reduced by keeping a) the number of unknown parameters as small as possible, b) the value ranges of the unknown parameters as tight as possible, and c) the final step size with respect to the unknown parameters as coarse as possible. An overall treatment of these issues is given below and we will return to them throughout the rest of this paper.

In our problem it is possible to reduce the number of unknown parameters by taking the fictitious gray level variables , $G_i, i \in A$, outside the search space. Instead, they can be evaluated within each state of the search simply by computing the average individually for each variable. This is possible if the normalization of the gray levels is not necessary. In this case, the approach is mathematically rigorous because the variables are mutually independent. The size of the parameter vector $\mathbf{p}$ in $Z(X,Y,\mathbf{p})$ should be kept minimal to reduce the total size of the search space.

The condition b) requires some case-dependent knowledge of the realistic ranges of unknown variables. Condition c) is closely related with the measuring accuracy of the matching algorithm. It could also be called resolution. It should be sufficiently high for not reducing the accuracy of the matching. If the expected accuracy of a variable is for example 100 mm, then it is not usually necessary to use a smaller than a 50 mm step size and in no circumstances smaller than a 10 mm step size.

4.2 Pruning techniques

Most of the pruning techniques used to reduce a search space are heuristic and do not always guarantee that the optimal state is found but rather a sub-optimal state (Pearl, 1984; Sarjakoski, 1988). For our application a hierarchical search is an obvious choice. It is based on the principle that the search is first completed using a wide range and a coarse step size, whereafter the search is repeated using a tighter range and a smaller step size.

Hierarchical matching techniques have become a standard method in digital photogrammetry, especially in the production of digital elevation models (Grün and Baltsavias, 1988; Helava, 1988; Ackermann and Krzystek, 1991; Heipke, 1992). In principle they are not identical with hierarchical search methods because also the input gray level functions is different, due to the use of image pyramids. The motivation for using hierarchical matching is also different: they are used primarily to avoid mismatches on high frequency features, whereas the reduction of computation is only of secondary importance.

4.3 Evaluating a single state

The effort required to evaluate a single candidate is the third criterion for making a search-based technique efficient. In our problem the number of the groundels involved has a direct influence on the computing time and therefore this number should be considered carefully. The issue is strongly dependent on the application. The use of hierarchical approaches is motivated also by this efficiency factor.

Regarding the processing of a single groundel, the computations involve the evaluation of the elevation function $Z(X,Y,\mathbf{p})$ and the spatial transformation function $S(X,Y,Z,\mathbf{r}_k)$. A continuous elevation function $Z(X,Y,\mathbf{p})$ can be expressed as a linear or as a piece-wise linear function in X and Y. This makes it possible to approximate the mapping from the 3-D object space to the 2-D image space with a bilinear interpolation. In this technique the mapping is computed rigorously in the corner points of a patch and by bilinear interpolation within the patch. The technique is well-known as an anchor-point method in the rectification of digital images.

The bilinear interpolation of the gray-level values can be replaced by nearest-neighborhood interpolation in supersampled input images. The approach is fully rigorous if the supersampling is carried out using bilinear interpolation with a sufficient enlargement factor. The image pyramid is, so to say, extended some levels below the ground by means of supersampling with bilinear interpolation. These levels are computed only locally and stored temporarily, due to their high storage requirements. With this technique the repeated resampling will be avoided in the search.

5. APPLICATION CASES

The principle of least squares matching by search has been described above in its generic form. In this section we specialize the technique for some frequently encountered tasks in digital photogrammetry, especially when aerial images are used. It must be emphasized that the list is not meant to be complete.

5.1 Height determination of a point

In height determination for a point with a predetermined planimetric position, the search space is one-dlmensional, defined by the unknown height. This makes the search highly efficient. If unknown terrain inclinations are also treated, two variables must be added, one for X-directional and one for Y-directional inclination. Because the inclinations are included only for compensating image deformations, even the final step size for these variables can be rather coarse. In practice it is rather questionable wheather inclinations can be considered at all, because the match area is usually small.

5.2 Matching a point

With the title 'matching a point' we mean here that a point feature is identified in one image (by human or by algorithm) whereafter the homologous points must be found from the other images. The task is only slightly different from the previous one: instead of moving vertically, we now move along the ray determined by the image point on the reference image. The search is as efficient as in the previous task. The technique resembles the geometrically constrained multiphoto matching by Grün and Baltsavias (1988). The crucial difference is that the final least squares matching for subpixel accuracy is made by search without forming linearized observation equations.

5.3 Matching a point when orientation is imprecise

We have assumed throughout this paper that the orientation parameters are known for each image so that epipolar search can be used. This assumption is not strictly valid for example in digital aerial triangulation, especially in its early stages when external orientation is known only approximately. For this purpose our search method must be extended so that in each state of the main search a subsearch is made to compensate the imprecise epipolar condition. The subsearch is made for each image in two dimensions. The dimensionality of the search space for this subsearch is $2l$, thus making it computationally rather demanding. Careful design is necessary for making a well-balanced choice of the search ranges and step sizes.

5.4 Template matching of a point

Matching a point feature to a template is a necessary operation in digital aerial triangulation using signalized control points. This can be accomplished either simultaneously or in two phases so that the images are first matched mutually and thereafter the resulting 'groundel image' with a known template. Based on the good results with a resembling method by Lammi (1994), the two-phase approach is favored here. It is likely to be more robust because the mutual matching of the images is not disturbed by the template matching. For cross targets the matching between the groundel image and a template image can also be made with least squares matching by search in three dimensions (two translocations and a rotation in $90°$ range). For point symmetric targets the search space is obviously two-dimensional.

5.5 Matching a line

When a line feature has been identified in one of the images, it can be matched from the other images by keeping the Z-coordinates of the end points of the line as unknown, thus creating a two-dimensional search space. The end points move along the rays as described in Section 5.2. The set of groundels to be matched is now defined by a narrow linear band. The optimizition criterion must be modified slightly so that root mean square error (RMSE) is used. This is necessary for comptensating the virtually varying number of groundels involved. For matching polylines the method can be applied sequentially so that the search for one line section is made only with the next vertex, the height of the previous vertex being already determined.

5.6 Matching a planar surface

When a planar surface is delineated by a quadrangle or any polygon in one image, the search for match can be made in three dimensions as in Section 5.2, the vertical translocation and the inclinations of a plane being the unknowns. The XY-position of each vertex is determined by spatial intersection with respect to this plane. The set of groundels to be matched is now defined by the planar surface projected to this place. For some applications it may be useful to extend this region slightly by buffering, to guarantee that texture on the edges support the matching procedure.

6. PRACTICAL EXPERIENCES

Until now the search method has been implemented for line features (Lammi,1996). The results are as expected. For well-defined lines the method works well but in low contrast areas we have encounter typical problems of an area-based matching. The earlier experiences with two-image cross-correlation on supersampled images (Lammi, 1994) let us assume that the method should work very well for point matching. The theoretical similarity of cross-correlation and least squares matching has been emphasized by Helava (1976).

Least squares matching by search has the advantage that the pull-in range is totally controlled by the case-specific range settings for the unknown variables. Overly wide ranges will, of coarse, increase the computing time.

7. QUALITY INDICATORS

Neither observation equations nor normal equations are formed in the least squares matching by search. Therefore the variance-covariance matrix of unknown parameters is not directly available for accuracy estimation. This drawback of the method could be circumvented by forming the normal equations explicitly in the goal state. Direct analysis of the texture in the match windows would be an alternative for quality analysis. The issue requires further studies.

8. OTHER CRITERIA FOR OPTIMUM

Least squares matching by search uses, as defined here, the sum of squared residuals as the object function. This criterion could be replaced by any other function of the residuals. Use of the sum of absolute values of residuals (L_1-norm) could be suitable. It would be less sensitive on large residuals which can be interpreted as blunders in the gray level values. The robust estimation methods and weight reduction methods used for blunder detection could also be implemented in a very efficient way.

9. ON COMPUTER IMPLEMENTATIONS

The numerical methods for making the search efficient where treated in Section 4. Regarding computer implementation of many numerical algorithms, the innermost loop is the most critical for speed. All 'tricks' for making it efficient should be regarded if the speed is a bottle neck or an obstacle for use. In our problem the innermost loop deals with the resampling of the gray level values and some rather simple computations on them. Use of methods like integer arithmetics instead of floating points arithmetics, multithreaded programming, or even assembly level programming could be justified here.

10. CONCLUSIONS

Least squares matching by search is based on the well established theory on object space least squares matching. The implementation is rigorous and straightforward because the linearization of the observation equations is not necessary. Use of case-dependent knowledge on geometry makes it possible to keep dimensionality of the search space moderate. Further reduction of the search space is possible by reducing the range of unknown variables hierarchically when the corresponding step size is decreased. The principle for reaching subpixel accuracy is similar compared to the original formulation of least squares matching, although the numerical realization is very different. Further study is proposed for developing other quality criteria to replace the use of the variance-covariance matrix of unknown parameters. The effect of using other norms in the object function should be investigated, to be able to compensate radiometric disturbances on the gray level values.

REFERENCES

Ackermann, F., 1984. Digital Image Correlation: Performance and Potential Application in Photogrammetry. The Photogrammetric Record, 11(64), pp. 429-439.

Ackermann, F. and P. Krzystek, 1991: MATCH-T: Automatic Mensuration of Digital Elevation Models. Technical Seminar of the Sociedad Espanola de Cartografia Fotogrammetria y Teledeteccion, Barcelona, Spain, pp. 67-73.

Ebner, H. and C. Heipke, 1988. Integration of Digital Image Matching and Object Surface Reconstruction. In: International Archives of Photogrammetry and Remote Sensing, Kyoto, Japan, Vol. 27., Part B11, pp. 534-545.

Förstner, W., 1982. On the Geometric Precision of Digital Correlation. International Archives of Photogrammetry and Remote Sensing, Helsinki, Finland, Vol. 24-III, pp. 176-189.

Grün, A., 1985. Adaptive Least Squares Correlation: A Powerful Image Matching Technique. South African Journal of Photogrammetry, Remote Sensing and Cartography, 14(3), pp. 175-187.

Grün, A. and E. Baltsavias, 1988. Geometrically Constrained Multiphoto Matching. Photogrammetric Engineering and Remote Sensing, 54(5), pp. 633-641.

Heipke, C., 1992. A Global Approach for Least Squares Image Matching and Surface Reconstruction in Object Space. Photogrammetric Engineering and Remote Sensing, 58(3), pp. 317-323.

Helava, U., 1976. Digital Correlation in Photogrammetric Instruments. A paper presented at the XIII International Congress for Photogrammetry, Helsinki, Finland, 26 p.

Lammi, J., 1994. User-Guided Measuring in Digital Aerial Triangulation. In: International Archives of Photogrammetry and Remote Sensing, Munich, Germany, Vol. 30, Part 3/1, pp. 467-474.

Lammi, J., 1996. 3-D Modelling of Buildings from Digital Aerial Imagery. A paper presented at the XVIII ISPRS Congress, Viena, Austria, Intercommission Working Group II/III.

Pearl, J., 1984. Heuristics: Intelligent Search Strategies for Computer Problem Solving. Addison-Wesley Publishing Company, 382 pages.

Sarjakoski, T., 1988. Artificial Intelligence in Photogrammetry. Photogrammetria. 42, pp. 245-270.

Thurgood, J. and E. Mikhail, 1982. Photogrammetric Analysis of Digital Images. In International Archives of Photogrammetry and Remote Sensing, Vol. 24-III, Helsinki, pp. 576-590.

Wrobel, B. 1987. Facets stereo vision (FAST vision) - A New Approach to Computer Stereo Vision and to Digital Photogrammetry, Interlaken, Switzerland, Proceedings of a Conference on Fast Processing of Photogrammetric Data, pp. 231-258.