TARGET DETECTION AND EPIPOLAR GEOMETRY FOR IMAGE ORIENTATION IN CLOSE-RANGE PHOTOGRAMMETRY

Gianfranco Forlani, Alberto Giussani, Marco Scaioni
Politecnico di Milano, Dip. I.I.A.R. sez. Rilevamento
p.za Leonardo da Vinci 32, 20133 Milano (ITALY)
Giorgio Vassena
Università degli Studi di Brescia, Dip. di Ingegneria Civile
v. Branze 38, 25123 Brescia (ITALY)

Commission V, Working Group 2

KEY WORDS: Target Detection, Close-Range Photogrammetry, Epipolar Geometry, images Orientation

ABSTRACT

The survey of complex three dimensional objects still require a considerable amount of human interaction in performing purely geometric tasks even in top digital photogrammetric systems. We present here an improvement of the automation of a subtask, namely the tie point transfer and the computation of an approximate exterior orientation by bundle block adjustment. We assume a close range environment, where object targeting plays a key role in the method. To speed up the workflow, we automated the localization and numbering of the targets, while the operator will still recognize the control points. In addition, after for a group of three images correspondencies are given, the use of the epipolar constraint make it possible to compute automatically the exterior orientation of any other image with a suitable number of points in common with the set of three images.

1. INTRODUCTION

The survey of complex three dimensional objects, though clear in the concept, still suffer from the difficulties of optimizing the photogrammetric network design (both in terms of number of images and of their position and attitude); apart from accuracy considerations, also practical but time consuming issues, such as to ensure a complete object coverage and the organization of the data acquisition for the digital surface model (D.S.M.) should deserve attention. Therefore even top digital photogrammetric systems still require a considerable amount of human interaction in performing purely geometric tasks (such as D.S.M. generation), to arrange the sequence of images, number the control points, provide approximate values if necessary and select (stereo)pairs for mass point determination by l.s. matching or other techniques.

To increase the degree of automation or just to improve the performance of such systems, several issues should be addressed, some perhaps specific to the survey of a particular class of objects, some more general. In the following we report about the experience gained in the survey of a sculpture, in the framework of a project aiming at assessing the feasibility of reproducing a faithful copy of the original by a numerical control milling machine. Details about the survey, performed by acquiring more than 100 images with two microscanning cameras JenScan coupled with a pattern generator and processed by the InduScan system of Zeiss (Kludas, 1995), can be found in (Vassena et al., 1996).

Here we focus on a subtask, namely the improvement of the tie point transfer and the computation of an approximate exterior orientation by bundle block adjustment. To this aim, the procedure can be split in three different stages:

- *target localization* (selection of all and only the targets);
- *target identification* (identifying the control points);
- *target transfer* (search for correspondencies in all images).

The first stage is the easiest to automate, if well defined targets are used; the second is feasible, provided that the c.p. are distinguished by individual marks. The last stage depends on informations about image geometry, target density and object shape. We assumed as starting point an operational environment where the operator, to provide data for a bundle block adjustment, works on image pairs and selects the homologous pairs; besides he labels them and identifies the control points. To speed up the workflow, we automated the localization of the targets, by using several levels of screening; the operator will still recognize the control points. In addition, after for a group of three images correspondencies are given, the use of the epipolar constraint make it possible to compute automatically the exterior orientation of any other image with a suitable number of points in common with the triplet of images.

2. IMAGE ORIENTATION

The determination of the exterior orientation of the images when the object is really 3D and the pictures are taken disregarding the normal case, is often a complex task. In fact it may be difficult to provide good approximations for the orientation parameters and, if some degree of automation is involved, also the selection of tie points may prove to be hard. The proper design of a survey might well include such problems in its framework. Particular attention should be paid to whether targeting may be worth or not in the case at hand. Retroreflective targets or even less expensive home-made targets can help substantially, simplifying both the ground control and the search for tie points.

As far as the ground control is concerned, there may be cases where, even if the object coordinate system is arbitrary, so that only a scale is to be fixed, some control points may be necessary to strenghten the block: in such circumstances only targets (no natural features of the object) may be determined with the necessary accuracy.
As far as the automatic or semi-automatic search for tie points is concerned, the procedure may well take advantage of well identifiable features in the object. Highly textured objects, rich in details, make life easy to humans or interest operators; to the contrary, dealing with poor or no texture, artificial patterns may be projected on the object. This will ensure dense and reliable measurements in the mass point determination stage. In a semi automatic procedure, nevertheless, where the operator's role is bounded to roughly select tie points, he would feel uneasy with such patterns, while targets might effectively support him and speeding up the procedure. A step further, in this respect, might be automatically recognizing targets in images, leaving the operator just with the task of assigning their correct labels (point numbers). This trade off can still be highly productive in a close-range environment, where the block structure and the object shape may prevent the algorithms designed for aerial images to connect models or photographs.
Dealing with a truly 3D object like a statue, the number of images necessary to cover the whole object may be considerably high; moreover, without a rough but adequate 3D model, no automatic design will provide with reliable results, since hidden parts of the object won't be covered: the expertise and the check of an operator will probably be still due for a long time to go, but methods for providing such a description simply and fast will be welcome.

3. AUTOMATIC TARGET DETECTION

Target selection and localization in the images may be performed by a combination of several criteria, in order to enhance the reliability of the whole process. If retroreflective targets are used, as in our case, they will usually appear as very brigth spots, often with g.v. close to the extreme of the brightness range: this characteristic can obviously be exploited in the localization process. Therefore we applied the following criteria:

- first selection by Foerstner's interest operator;
- ranking according to interest value and mean of the gray values within a mask;
- template l.s. matching;
- object coordinates consistency in case of redundant observations.

High contrast features like the targets used are easily located by *Förstner's interest operator* (Förstner, 1986; Förstner, 1991), provided that an appropriate mask size is chosen (either not too small nor too large, compare to target size). To reduce the computational load, the mask is moved in steps related to its size and, after local maxima suppression, the bias correction is applied computing the gravity center of the feature. Since not only targets will be selected (but no one will be missing, except perhaps for cases where collimation would be uneasy also for an human operator) false candidates should be discarded; due to the pattern generated by the projector, in our case the candidates were still in the order of thousands.

To discriminate the targets, the circularity index proved not to be good, while the interest value, though not in absolute terms, that is only on a single image base, was always in the top of the rank list. Given the maximum number v of targets visible in an image, a first criterium was therefore to rank the candidates and discard all points exceeding the v-th position. This has been complemented by the computation of the mean of the gray values within a mask around the candidate centre, removing those out of a threshold empirically fixed. This second criterium proved to be less sensitive to changes from image to image, compare to the interest value.

3.1 Template least squares matching

To improve the accuracy of the target location and to further discard false candidates a *template l.s. matching* is performed (Lemmens, 1988; Baltsavias et al., 1990; Baltsavias, 1991), where the template is a synthetic image representing the target. The matching program works in a X-Windows interface under Unix environment and has been written in C and Fortran 77.
The geometric model assumed for resampling within the search window is an affine transformation, where any of the parameter may be constrained, adjusting the model to a simpler one (shift only, conformal transformation, etc) prior to the processing (no automatic testing of significance and updating is performed in the iteration process). To stop the iterations, the relative change of sigma naught and of the correlation coefficient are used; in addition, if a correlation threshold (0.5) is not reached within the first 10 iterations or if the convergence rate is too low, the matching is stopped. If at the end of the process a second threshold, fixed depending on the quality (distinctness, image contrast, noise level, etc.) of the points to be matched is reached, the result is accepted. Since l.s. matching is sensitive to outliers in the gray values, due e.g. to dust or scratches or to inhomogeneous target background, the solution may converge to a local minimum, while still enjoying a degree of correlation sufficient to let the point to be accepted as a good one. To contrast this bias, the matching can be robustified by using a different adjustment principle or introducing some outlier rejection procedure. As a very basic but sometimes very effective idea is to remove from the solution the contribution of the pixels with g.v. outside a given range; this works well in a controlled enviroment. Rather than a robust adjustment method, a more conventional data snooping procedure has been used, rejecting all standardized residual larger than 4 at each iteration of the non linear process. The outlier rejection take place only after a certain degree of convergence is reached, since at the beginning radiometric differences between template and slave are likely to depend on the approximate registration of the images rather than on image noise. Experiences have shown that the procedure may cause some instability (oscillations of the parameters in subsequent iterations) but that it is effective in reducing the bias in many cases.
To improve the fitting between template and search window, a radiometric transformation is introduced, which should partly compensate for differences in illumination, reflectivity and so on. The parameter of the transformation are not included in the l.s. adjustment, but are computed prior to the matching: we used the *Wallis filter* (Wallis, 1976) in a particular case, forcing the

equalization of the mean and the variance of the g.v. between the two windows:

$$g_p^c = (g_p - \bar{g}_T) \cdot \frac{\sigma_T}{\sigma_p} + \bar{g}_T \qquad (1)$$

g_p^c, g_p: g.v. of a patch pixel after and before the filter correction;
$\bar{g}_p, \bar{g}_T$: mean values of g.v. of the patch and the template;
σ_p, σ_T: st.d. of g.v. of the patch and the template.

In the first iterations a constraint is put on the amplification of the contrast:

$$\frac{\sigma_T}{\sigma_p} < q_{max} \qquad (2)$$

q_{max}: threshold value for contrast amplification.

Since it has been noticed in previous experiences with poorly textured surfaces that, when the initial values are not very good, it may happen that the g.v. variance in the search window is very small, because just a small part of the target shows up in the window. Therefore the amplification factor becomes very high, affecting the convergence of the geometric transformation to the correct solution (basically, the scale factor may be misjudged); after a sufficient degree of convergence is reached, this effect disappears, so the equalization can improve the radiometric fitting without any bias of the solution.
Two types of targets were used on the sculpture: a set for control points, provided with point number, and a second one for tie points; both contain a white central disk surrounded by a black ring (see Fig. 2). To build a template good for both kinds of targets, the image of a target with minimal perspective deformation has been cut out. A synthetic copy of the target has been generated by selecting half of a profile across the centre, interpolating it by polynomials and rotating it around a vertical axis, to give raise to a symmetric distribution of the g.v. with respect to the window centre (see Fig. 1).

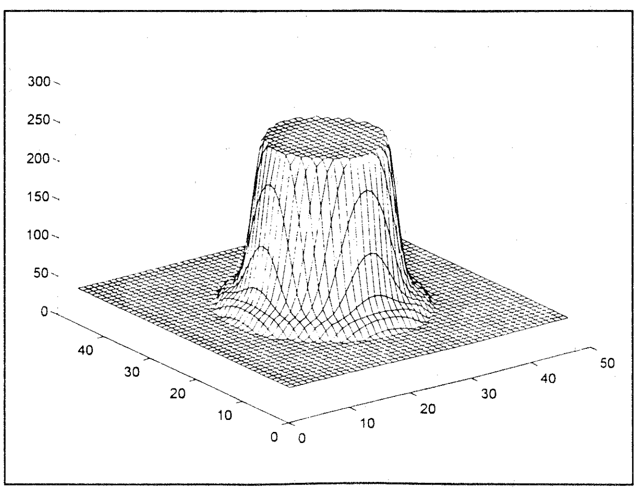

Figure 1 - 3D view of the template

The theoretical accuracy on the image, based on the (usually optimistic) estimate given by the inverse matrix by l.s.m., is in the order of 0.15 mm; the R.M.S. of the differences with respect to the same targets, measured by the InduScan system, are nevertheless in the same range.

4. THE SEARCH FOR HOMOLOGOUS POINTS

Once targets have been located in the images, correspondencies must be established between image points; moreover, control points need to be identified. If we could approximate the object by a plane, at least in significantly large areas, we may compute a 8-parameter transformation, identifying manually 4 points in each image and then transferring automatically all targets from a reference image. In aerial blocks the flight plan and approximate informations on the elevations may provide with a range along an approximate epipolar line (Liang and Heipke, 1994). The same idea may be used with 3D objects, but the approach may fail, because it is difficult to find good approximate orientation values for the images: it may be necessary to extend the search for candidates to a larger region along the epipolar line, increasing the risk of false matchings, particularly when the candidates all look the same, like retroreflective targets. To avoid inconsistencies, the search area should contain only one candidate or, in other words, epipolar geometry should be precisely known for at least three images. To automate the process, any available a priori information on the imaging geometry should be considered, but still in many cases the amount of information may not be enough.
We opted for an intermediate solution, where human intervention is still necessary, but limited to tasks easy for humans and more difficult to algorithms. In practice, we want to improve a workflow where the operator will roughly point to the targets on the screen, then will label them in all images, either assigning new numbers or reading the control point numbers (see Fig. 2). This is still much work when there are many images and many targets, like in our case; therefore we set up a semi-automatic procedure, where the operator's job will be restricted to restarting the procedure whenever it stops and to identifying the control points only. The other targets will be numbered and recognized automatically.

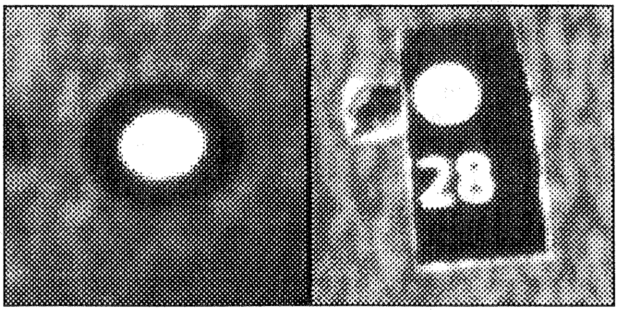

Figure 2 - The two kinds of targets used

4.1 Description of the procedure

The input data for the procedure are the target image coordinates and the object coordinates of the control points; the operator looks for pairs of images containing at least the same 6 targets, 3 of which being control points, and begins by pairing all these image points. The (asymmetric) relative orientation for each pair is computed, starting from approximate values of the rotations and base components. In our case, since the images were taken in pairs with two cameras mounted on a rod, the orientation of the images of a pair is easy, more difficult otherwise. Rather than using the Cardan sequence for the rotations, we express their approximate values in terms of a Euler sequence, which are easier to provide, and solve from the rotation matrix with respect to the angles ω, ϕ, κ. The models are thereafter absolutely oriented: since the Anblock hypotheses are not satisfied, we compute the transformation parameters by expressing the rotation matrix by the *Hamilton's quaternions* (Sansò, 1973) which do not require initial values. Finally, a bundle adjustment provides the exterior orientation of the images. Since the distribution of the targets in the images may not be ideal, the adjustment will be repeated at a later stage, when more image points are made available. If the initial configuration is too poor, nevertheless, the procedure may not continue correctly, due to a lack of targets in the area.

We can now take advantage of the known epipolar geometry selecting from the pairs sets of three images, to find automatically more corresponding targets (Ayache, 1991; Maas, 1992).

Let A be the reference image, B and C the other two images. For each target on A, the corresponding epipolar lines e_B and e_C are drawn on B and C respectively; targets P_{Bj} and P_{Ck} falling in a band along the two lines are considered (see Fig. 3).

Starting from a point P_{Bj} in B its epipolar is drawn on C and its intersection with e_C is computed. If a candidate P_{Ck} is found in a window defined by the two bands overlap around this point, then a candidate set is found; more candidates P_{Ck} may anyway exist in the same window. Once all possible sets are found, ambiguities must be solved through a consistency check of the intersecting rays. This is done by computing, for each set, the three distances between the rays and taking their mean value. The set enjoying the minimum distance d_{min} is taken. If d_{min} is smaller than a threshold, the ground coordinates of the point are computed. Once all candidates from A have been checked, the exterior orientation is improved by running a second bundle adjustment including the new points.

After this first stage we have (one or more) groups of three oriented images, some known object points and some image points still to be numbered on the images. To go further the operator looks for adjacent images sharing at least three known object points and three more targets with two of the oriented images. He just assign the correspondencies, labelling the new points on the new image. Based on the common set of points, the above procedure is repeated and the new image is incorporated in the block, while the number of targets determined in object space increases. The procedure continues adding new images but may eventually stop if there are no images satisfying the above conditions. If this is the case, the operator must find another group of three images to restart with.

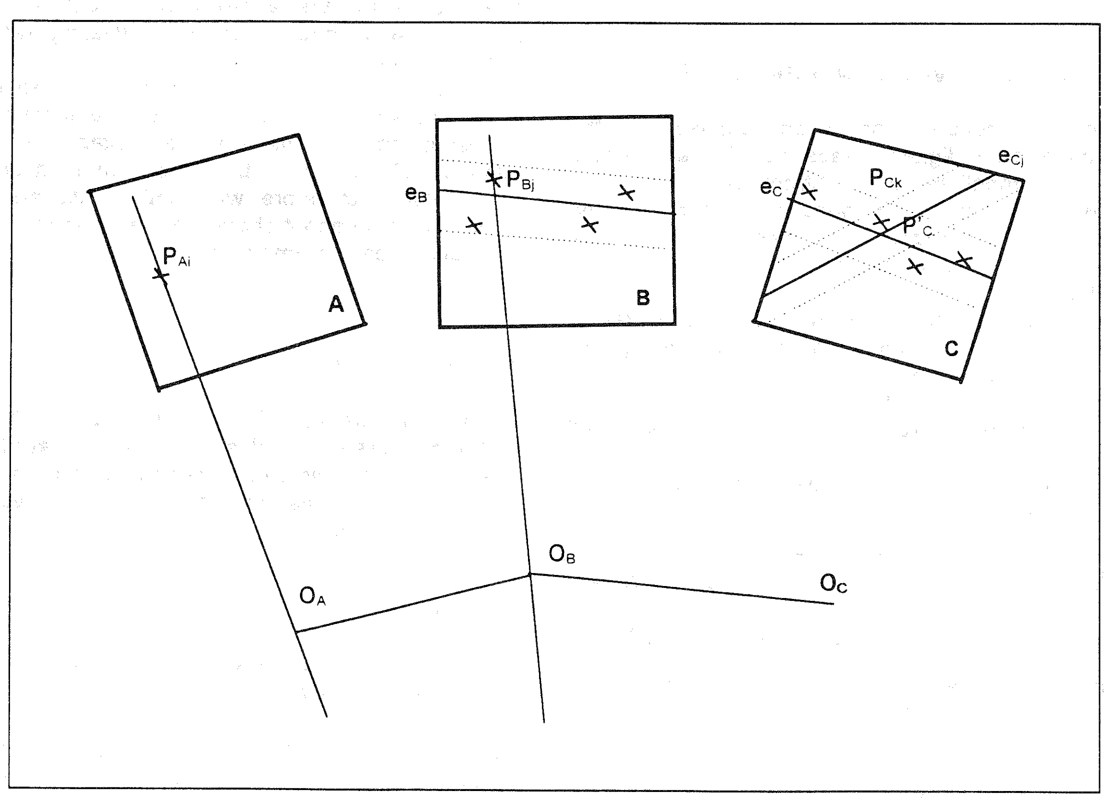

Figure 3 - The identification of the corresponding image points

Figure 4 - The statue surveyed

5. THE SURVEY OF A SCULPTURE

The above procedure has been applied to the determination of the digital surface model of a sculpture (see Fig. 4), as mentioned in the introduction.
For the whole survey, 108 images have been acquired and 11 of the 150 targets used were determined by theodolite survey with an accuracy of 1/10 mm on all coordinates.
The functioning of the proposed method has been tested with only 10 images, each containing on average 14 targets, covering the sculpture's head.
Control points were available only in the front part of the body. The 10 images were successfully tied together. All targets were automatically localized and measured on the images, while the control points were labeled by the operator (see Table 5).

image n.	1	2	3	4	5	6	7	8	9	10
targets	17	16	14	14	14	12	15	16	16	9
c. points	6	7	6	6	4	4	6	6	4	3

Table 5 - Number of targets and control points in the test images.

Three pairs of images (group 1) were selected out of the whole group, which contained at least 3 c.p. and 3 tie points; a bundle adjustment has been computed, providing the approximate exterior orientation.

Looking in all possible sets of three images, thanks to the epipolar constraints, the coordinates of additional tie points were determined (see Table 6).

image n.	1	2	3	4	5	6	7	8	9	10
targets	17	16	14	14	14	12	15	16	16	9
numbered	10	11	11	11	4	4	13	12	4	3

Table 6 - Results of the target location procedure for group 1 (in dark cells)

Each of the four remaining images has been linked to one of the three sets of group 1, by manual identification of the minimum number of common targets. Correspondencies between the targets on the new images and the images of group 1, were automatically found (see Table 7). Including these new observations, a new bundle adjustment have been run, improving the exterior orientation.

image n.	1	2	3	4	5	6	7	8	9	10
targets	17	16	14	14	14	12	15	16	16	9
numbered	11	14	13	13	13	12	13	13	9	8

Table 7 -Results of target location on the whole set

6. FINAL REMARKS AND PERSPECTIVES

The results shown above confirm that the implemented procedure is effective in speeding up the measurement of tie points for image orientation. Many developments can be still introduced either robustifying the algorithms or improving the user-interface.
The search for the approximate orientation may be simplified and made capable of dealing with strong convergent cases.
At present the procedure runs over all possible sets of three images to search for additional target correspondencies. This is still acceptable with a small number of images, but the combinations grow exponentially: therefore we need to discard all sets whose images cannot share, on the basis of the exterior orientation, any common target.

Aknowledgements

The authors are grateful to Carl Zeiss GmbH and to Dipl.-Ing. Torsten Kludas for making available the images used in this work. Many thanks go also to Dr. Sergio Grassi for his support in program developments under X-Windows.

References:

Ayache, N., 1991. Artificial vision for Mobile Robots. MIT Press, Boston, U.S.A..

Baltsavias, E.P., 1991. Geometrically Constrained Multiphoto Matching. Mitteilungen N. 49, Institute of Geodesy and Photogrammetry, ETH, Zurich, Switzerland.

Baltsavias, E.P., Beyer, H.A., Fritsch, D., Lenz, R.K., 1990. Tutorial on: Fundamentals of Real Time Photogrammetry. Institute of Geodesy and Photogrammetry, ETH, Zurich, Switzerland.

Förstner, W., 1986. A feature based correspondence algorithm for image mathching. In: IAPRS, Rovaniemi, Finland, Vol. 26, Part 3/3, pp. 150-166.

Förstner, W., 1991. Object Extraction from Digital Images. In: Proceedings of the 43rd Photogrammetric Week, Institut für Photogrammetrie, Universität Stuttgart, Germany, pp. 1-15.

Kludas, T., 1995. Three dimensional surface reconstruction with the Zeiss Photogrammetric Industrial Measurement System InduSURF. Edited by Carl Zeiss Jena GmbH, Jena, Germany.

Lemmens, M., 1988. A Survey on Stereo Matching Techniques. In: IAPRS, Kyoto, Japan, Vol. 27, Part B8/III, pp. 11-23.

Liang, T., Heipke, C., 1996. Automatic Relative Orientation of Aerial Images. PE&RS, Vol. 62, n. 1, München, Germany, pp. 47-55.

Maas, H.-G., 1992. Complexity analysis for the determination of dense spatial fields. In: Robust Computer Vision, Wichmann, Karlsruhe, Germany, pp. 243-256.

Sansò, F., 1973. An exact solution of the roto-translation problem. Photogrammetria, Vol. 29, pp. 203-216.

Vassena, G., Scaioni, M., Kludas, T., Pfeil, J.-U., 1996. Automatic D.E.M. generation using digital system InduSCAN: an application to the artworks of Milano Cathedral finalized to realize physical marble copies. In: IAPRS, Vienna, Austria.

Wallis, R., 1976. An approach to the space variant restoration and enhancement of images. In: Proceedings of Symposium on Current Mathematical Problems in Image Science, Naval Postgraduate School, Monterey, U.S.A.