An Alternative to the Epipolar Line Method for Automatic Target Matching in Multiple Images for 3-D Measurement

J.Chen[1],T.A.Clarke[2] & S.Robson[3]

Department of Electrical, Electronic, and Information Engineering[1&2]

Department of Civil Engineering[3]

City University

Northampton Square

LONDON EC1V 0HB

U.K.

Abstract

In digital photogrammetry, difficulties are often experienced when automatically matching differing viewpoints of similar targets, such as retro-reflective targets. In this paper, an alternative target matching technique to the epipolar line method is developed. The method, which is combined with a bundle adjustment process, is based on a 3-D intersection and "epipolar plane", as opposed to the 2-D intersection of the epipolar line method. The theory of the approach is described and discussed. Comparisons are made between the epipolar line method and the 3-D matching method. Results from both simulation and practical testing are given and critical evaluation is made of the technique.

1. Introduction

In digital photogrammetric 3-D engineering measurement many problems have been successfully solved using various target matching techniques [Gruen,1988]. The main matching algorithms applied are area [Heikpe,1992] and feature [Blake,1990] based methods. Unfortunately, targets viewed from differing angles and distances seldom provide enough unique information to allow successful matching. This is especially true of circular retro-reflective targets which are often used for high accuracy measurement. The popular epipolar line method [Ayache,1991], founded on the projection of straight lines into the image plane, is often used to find corresponding targets. The method requires good estimations of the camera orientation parameters, also the straight epipolar lines cannot easily include system distortions such as lens distortion [Maas,1992].

Limitations encountered with the epipolar line method and poor network geometry can be overcome by combining bundle adjustment techniques with target matching. The bundle adjustment can be used to iteratively improve the camera exterior and interior orientation parameters during the matching procedure. The method can initially use the strongest network elements to obtain good target and camera orientation parameters before any remaining target image measurements are gradually added. Such a method can provide advantages, for example in the estimation of

additional parameters, over the epipolar line method for target correspondence matching.

In this paper a "3D space intersection method" is developed whereby target labelling is carried out iteratively within the bundle adjustment procedure and successful matches are identified by 3-D object-space intersection discrepancies. A master image is selected, its target image locations identified to sub-pixel accuracy (Clarke et al,1993) and reprojected into the object space. Given reasonable initial camera parameters, corresponding target images projected through a number of viewpoints will intersect in space within some pre-determined 3D tolerance. Such a 3D tolerance can be usefully obtained from the computed RMS standard deviations of the estimated target object coordinates. Target images which do not originate from the same point in space will produce a large intersection error, and can be rejected. The accepted candidates are used in a bundle adjustment where the initial camera parameters are refined. On successful adjustment any unused targets image locations can be retested with the refined parameters to ascertain whether they now fall within the space intersection tolerance.

The 3D method is formalised, tested and discussed in this paper. Consideration is also given to overcoming ambiguities, the selection of a space tolerance value, and solving target occlusion problems.

2. The 3-D space intersection matching algorithm

A 2-D image is formed by targets placed on an object in the 3-D world when light rays from the targets are projected through the camera lens onto the sensor. The co-ordinates x_i and y_i of the various target images can be identified and measured automatically. These measured image positions are not likely to be exactly correct, but will contain errors caused by image intensity variations, lens distortion, and any systematic or random errors in the system used. To account for these variations some additional parameters must be included in the mathematical model used to compute the 3D co-ordinates of each target.

If additional parameters are considered, the well known collinear equations can be expressed as [Methley, 1986]:

$$\frac{(Xa-Xo)}{(m11(x-x_o+\Delta x)+m21(y-y_o+\Delta y)+m31(-c))}=\frac{(Za-Zo)}{(m13(x-x_o+\Delta x)+m23(y-y_o+\Delta y)+m33(-c))} \quad (1a)$$

$$\frac{(Ya-Yo)}{(m12(x-x_o+\Delta x)+m22(y-y_o+\Delta y)+m32(-c))}=\frac{(Za-Zo)}{(m13(x-x_o+\Delta x)+m23(y-y_o+\Delta y)+m33(-c))} \quad (1b)$$

where Δx and Δy are the corrections for lens distortion.

In such a form the equations satisfy the space line equation:

$$\frac{Xa-Xo}{p}=\frac{Ya-Yo}{q}=\frac{Za-Zo}{r} \quad (2)$$

The intersection of two such lines determines a point in space. The line equation represents a ray projected from the target through the perspective centre of the lens. If each of the target images from another viewpoint are projected through its perspective centre, then the corresponding target projection ray will come close to the

single projected 3D line. If additional viewpoints are considered in the same way, further target image candidates will be found. The lines will rarely intersect exactly due to geometric distortions and target image location errors. The principle of the 3D space matching method is illustrated in Figure 1 using two camera viewpoints. It can be shown that whilst the two projected rays of a matching target are not likely to intersect at a single point in space because of the errors mentioned previously, their separation D can be calculated directly from equation 2:

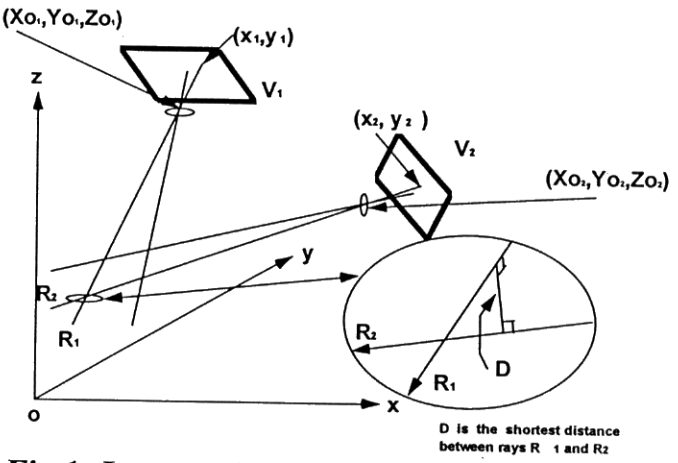

Fig.1: Intersection of two projected rays

$$D = \begin{vmatrix} Xo_2-Xo_1 & Yo_2-Yo_1 & Zo_2-Zo_1 \\ p_1 & q_1 & r_1 \\ p_2 & q_2 & r_2 \end{vmatrix} \Big/ \left\{ \begin{vmatrix} p_1 & q_1 \\ p_2 & q_2 \end{vmatrix}^2 + \begin{vmatrix} q_1 & r_1 \\ q_2 & r_2 \end{vmatrix}^2 + \begin{vmatrix} r_1 & p_1 \\ r_2 & p_2 \end{vmatrix}^2 \right\}^{-1/2} \quad (3)$$

where: p_1,q_1,r_1 and p_2,q_2,r_2 are the vectors of equation 2 for each of the two viewpoints.

If targets in the master image are projected, then matching targets from other images can be identified. For all correctly corresponding targets, D should be small and lie within some pre-determined tolerance δ. For all other non-corresponding targets, the value of D will depend on the spatial distribution of the targets, and will generally be much larger than that for the correctly corresponding targets. The tolerance band needs to be selected so that it excludes all incorrect targets, but encompasses the correct ones. The value of the tolerance band can be allowed to change, becoming successively smaller as the solution is strengthened and refined. All matched 3D target co-ordinate starting values can be estimated by 3D space intersection before the bundle adjustment. Since the 3D tolerance band is chosen using target object co-ordinate standard deviations computed from the bundle adjustment, it includes a complete reflection of the network properties at each bundle adjustment iteration. Additionally the 3D tolerance will reduce as the network improves.

There are two distinct differences between the 3D space method and the epipolar line method: the epipolar line method is a 2D projection of a space ray, which passes close to a point in another view, whilst the 3D-space method is the space intersection of rays through image points from two or more views; the threshold values for judging correspondence between image views are different, the 3D-space method uses the computed standard deviations of the target coordinates in 3D space, whilst the epipolar method relies on image co-ordinate residuals.

During the matching procedure, attention must be paid to target ambiguities and occlusions. Two stages are used to improve the correspondence: a **global uniqueness constraint** in multiple views to match all targets except where an occlusion or

ambiguity occurs (stage one); and a **local uniqueness constraint** to overcome such problems by selecting a subset of the viewpoints (stage two).

2.1 Global uniqueness constraints by matching groups of targets from multiple views

An epipolar plane is defined as a space plane $\mathbf{P_{11}T_1P_{21}}$ defined by two space rays, along which target $\mathbf{T_1}$ projects onto two image planes at $\mathbf{P_{11}}$ and $\mathbf{P_{21}}$ (Figure 2). The epipolar plane will intersect each image plane along an epipolar line. Ambiguities will occur if a target lies on or close to any other target's epipolar plane, for example target $\mathbf{T_2}$.

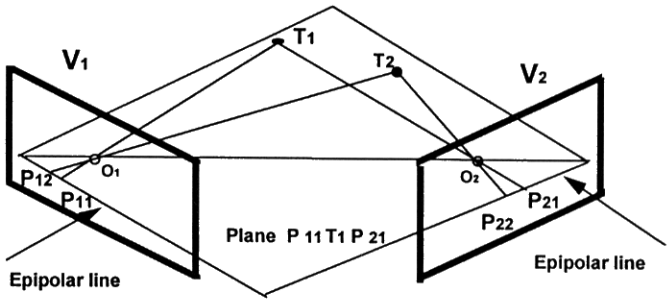

Fig.2: Epipolar plane generation

If **n** targets in the object space are projected onto **m** cameras, there will be **m** sets of points where: $V_i = \{P_{ij}, (j=1,..,m_i\ (m_i<n))\ (i=1,..m\ )\}$.

The global uniqueness constraint algorithm can be described as follows: Given a group of image points G(P), which includes all possible reprojections of image points with respect to a ray from a single image point in a master image:

$$P = \{P_{1j},P_{2j},...P_{mj}\}\ (P_{1j} \in V_1, P_{2j} \in V_2,... P_{mj} \in V_m)$$

where: $V_1,V_2,...V_m$ are complete sets of points in each view.

There will be a subset of points: G(M), where the minimum distance between the reprojected rays which satisfy the 3D object space tolerance:

$$M = \{P_{1j},P_{2j},...P_{mj}\}\ (P_{1j} \in V1, P_{2j} \in V_2,...P_{mj} \in V_m), \text{ satisfy } D_{ik} < \delta\ (i=0,..m,\ k=0,..m,\ i \neq k)$$

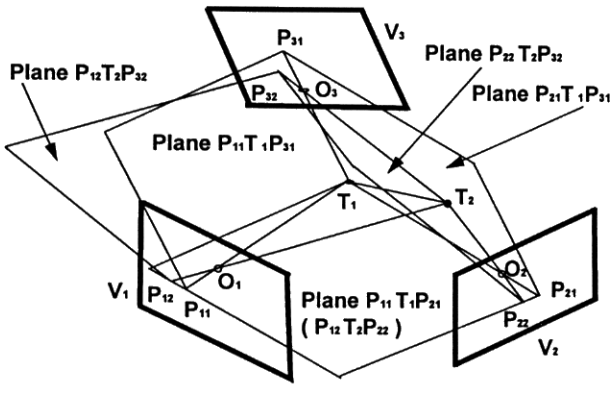

Fig.3a: Two targets on the epipolar plane of V1,V2

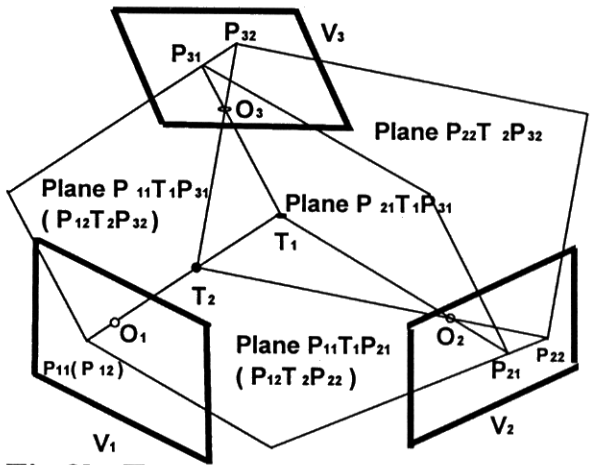

Fig.3b: Two targets on the epipolar plane of V1,V2 and V2,V3

where: D_{ik} is the space distance between point P_{ij} and P_{kj}, and δ is the tolerance value.

$G(M) \in G(P)$. when M is composed of $\{P_{1j},P_{2j},...P_{mj}\}$, this set is termed an initial matching group. If another group $M' = \{P_{1j}',P_{2j}',...P_{mj}'\}$, based on a different image point, exists and contains at least one common image with group M, then at least two targets lie on the same epipolar plane. In this situation, both of the image points in group M and M' must be rejected as non-matching target image points. If the object point in M does not appear in any other whole set, M is a matching group and labelled.

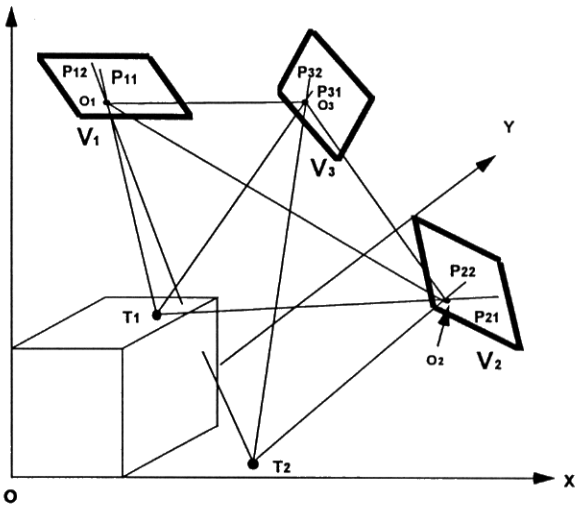

Fig.4: An object point occluded in V1

The matching process is carried out to find G(M) from any set G(P). Each element in G(M) must be tested and the group rejected if two targets are found on the same epipolar plane. Targets on the same epipolar plane can be matched during this stage. In Figure 3a, targets T_1 and T_2 lie on the epipolar plane of viewpoints V_1,V_2, so that T_1 and T_2 are irresolvable between V_1 and V_2. However if the global uniqueness constraint is applied to viewpoints V1, V2 and V3, then: $M_1 = \{P_{11},P_{21},P_{31}\}$ and $M_2 = \{P_{12},P_{22},P_{32}\}$ are unique matching groups consequently, both sets can be accepted. In Figure 3b, targets T_1 and T_2 lie on the epipolar plane of viewpoints V_1,V_2 and V_1,V_3. Such a situation cannot be solved with a global uniqueness constraint since the two groups, $M_1 = \{P_{11},P_{21},P_{31}\}$ and $M_2 = \{P_{11},P_{22},P_{32}\}$, both contain target T_1 such that no unique matching group exists. Targets in occlusion, (Figure 4) cannot be matched with the global uniqueness constraint because no group G(M) can be found.

2.2 Local uniqueness constraint for overcoming ambiguity and occlusion

This stage of the matching process is designed to remove any remaining ambiguities and occlusions remaining from stage one. The process is carried out between pairs of images. On completion of stage one, there will generally be a large number of matched targets. These targets are included in the bundle adjustment network to enhance the geometry and to strengthen the network. Consequently most of the estimated parameters including any camera inner orientations will be close to their true values. Targets which have not been matched in stage one can be explained by: occlusion in at least one image; or ambiguities; or single projections. If following stage one the object space RMS standard deviation is small, a small tolerance can be chosen. Consequently combinations of two viewpoints can be used to match the remaining targets.

Initially a two point group M is found, where:

$M = \{P_{ik},P_{jl}\}$ $(i != j,\ i=1,2...,m,\ j=1,2,...,m,\ k=1,2,...,m_i,\ l=1,2...,m_j)$

If another group M' can be found, where $M' = \{P_{ik}',P_{jl}'\}$ and in which $P_{ik}'=P_{ik}$ or

$P_{j1'}=P_{j1}$, both groups M and M' must be rejected as non-matching points. If only one G(M) exists M is a matching group and labelled. For ambiguous targets such as T_1 and T_2 (Figure 3b), which lie on the epipolar plane of V_1,V_2 and V_1,V_3 and remain from stage one, it is possible to obtain correspondences with V_2 and V_3. Target T_2 occluded in viewpoint V_1 (Figure 4) can be matched by isolating viewpoint V_1 and using V_2 and V_3. However matching reliability will decrease in such cases where only two viewpoints can be used.

Theoretically, after this stage, all targets will be matched except for any which are single projections since these are indeterminate. However under practical conditions, with very dense target arrays, problems of multiple ambiguity for a single master image point can arise (Maas, 1992).

3. Evaluation of the method by simulation and experiment

To test the method, several simulations and practical experiments were carried out under a variety of conditions. In each case the matching method produced good results.

3.1 Evaluation by simulation

A computer simulation model was constructed (Figure 5). Two object planes were defined, one on the XOY plane and another parallel to it but displaced in Z. On each plane a number of targets were randomly distributed. Five viewpoints were generated around the object such that the simulation included ambiguities, occlusions and single projections. The 3D matching method was successful in the simulated trials, all targets were matched, with the exception of the indeterminate single projections.

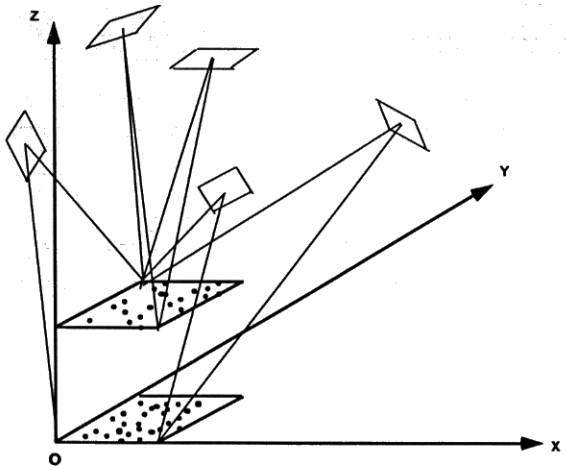

Fig.5: The simulation network

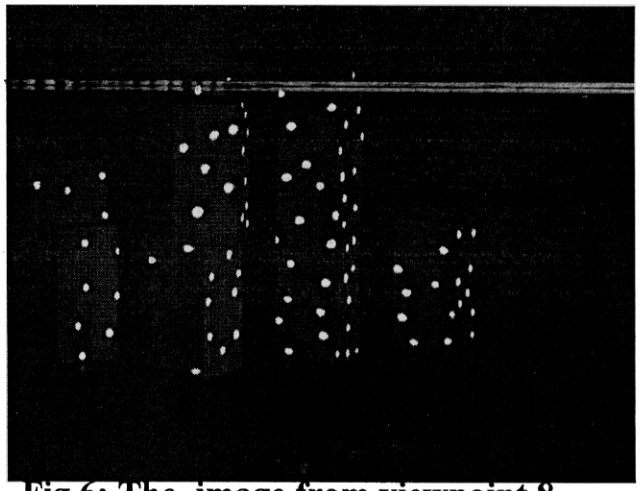

Fig.6: The image from viewpoint 8

3.2 Test by experiment

To test the method further a target field was constructed using a number of toy bricks arranged on a surface. A piece of opaque acetate, in which a number of pin holes had been randomly made, was positioned in a projector to produce target blobs on the surface of the bricks (Figure 6). A black cover was positioned to minimise ambient

illumination. Nine images were collected from viewpoints located in front of the test field. Final results after the two-stage matching process are shown in Table 1. Some remained unmatched because of target images which were falsely identified as targets in some of the viewpoints, for example those at the base of viewpoint 8.

	Iteration No.	V_1	V_2	V_3	V_4	V_5	V_6	V_7	V_8	V_9
Image Points	-	62	65	67	102	112	111	75	87	84
Stage 1	1	51	51	51	81	81	81	56	56	56
Stage 1	2	3	3	3	8	8	8	6	6	6
Stage 2	-	6	8	8	13	23	22	8	16	16
Remd.		2	3	5	0	0	0	5	9	6

Table 1: Number of correct matches per viewpoint.

3.3 Experimental test using a car gearbox

A car gearbox was chosen as a significantly complex object (Figure 7a). About 250 circular retro-reflective targets were placed on the gearbox surface. Four camera viewpoints were used, each at a distance of approximately 1 metre from the gearbox. A weighted centroid algorithm was used on the grey scale image to compute a sub-pixel location of the target image. A filling algorithm (Clarke et al. 1993) was used during this process to trace and automatically set the window size in which to locate the target. Results from the 3D matching process, which took about 4 minutes on a 33Mhz 80486, are shown in Table 2. Figure 7b shows a primitive reconstruction produced by downloading the data into a CAD system.

Fig.7a: The car gearbox from viewpoint 3

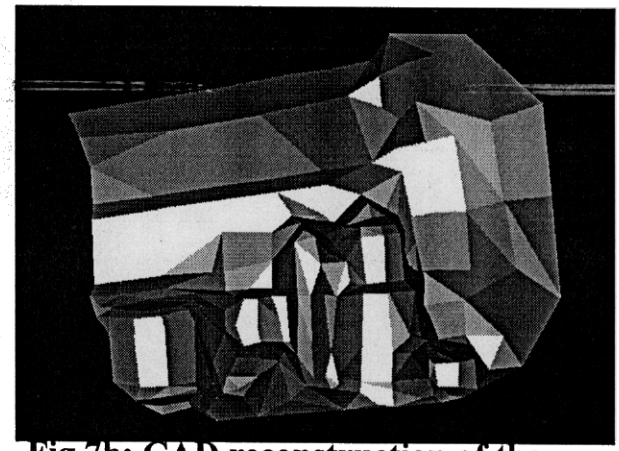

Fig.7b: CAD reconstruction of the gearbox

	V_1	V_2	V_3	V_4
Total targets	179	221	237	220
Stage 1	42	42	42	42
Stage 1	101	101	101	101
Stage 1	9	9	9	9
Stage 2	26	69	81	67
Final result	178	221	233	119

Table 2: Successful matches from the gearbox application.

4. Conclusions.

A target matching method for multi-viewpoints has been successfully developed whereby a two stage iterative process is used. During the procedure the photogrammetric network is iteratively refined and strengthened as matching targets are gradually identified and introduced. Multi-view constraints and constrained self calibration can be applied during the target correspondence process to increase matching reliability. Solutions for target image groups which include ambiguities such as occlusion can be obtained automatically. An epipolar plane is defined such that the cause of such ambiguities can be explained in 3D space and removed by automatically restricting the viewpoints used in the matching process. The 3D target RMS standard deviations computed and refined during the adjustment procedure are chosen as space intersection tolerances. The method has been found to be valid in the general case, and appears particularly suited to accurate industrial measurement where highly complex objects requiring multiple views must be measured reliably.

5. References

Ayache, N., 1991, Artificial Vision for Mobile Robots: Stereo Vision and Multisensory Perception, Massachusetts Institute of Technology, The MIT Press, 342pp.

Blake, A., McCowen, D., Lo., H.R., Konash, D., 1990, Epipolar Geometry for Trinocular Active Range-sensors, British Machine Vision Conference, BMVC90, pp19-24.

Clarke, T.A., Cooper, M., Fryer, J., 1993, An Estimator for the Random Error in Sub-pixel Target Location and its Use in the Bundle Adjustment. In Proc. Optical 3-D Measurement Techniques, Zurich, In press.

Gruen, A.W., 1988. Geometrically Constrained Multiphoto Matching, Photogrammetric Engineering & Remote Sensing, Vol.54, No.5, pp.633- 641

Heipke, C., 1992. A Global Approach for least-Squares Image Matching and Surface Recognition in Object Space, Photogrammetric Engineering & Remote Sensing, Vol.58, No.3, pp.317-323

Maas, H.G., 1992a. Complexity Analysis for the Establishment of Image Correspondences of Dense Spatial Target Fields, Int. Archives of Photogrammetry and Remote Sensing, Vol.29(5), pp.102-107

Methley, B.D.F., 1986, Computational Models in Surveying and Photogrammetry, Thomson Press Limited, 346pp.