High speed correspondence for object recognition and tracking

D. D. A. P. Ariyawansa and T. A. Clarke

Department of Electrical, Electronic, and Information Engineering,
City University, Northampton Square, London EC1V 0HB, U.K.

ABSTRACT

Real-time measurement using multi-camera 3-D measuring systems requires three major components to operate at high speed: image data processing; correspondence; and least squares estimation. This paper is based upon a system developed at City University which uses high speed solutions for the first and last elements, and describes recent work to provide a high speed solution to the correspondence problem. Correspondence has traditionally been solved in photogrammetry by using human stereo fusion of two views of an object providing an immediate solution. Computer vision researchers and photogrammetrists have applied image processing techniques and computers to the same configuration and have developed numerous matching algorithms with considerable success. Where research is still required, and the published work is not so plentiful, is in the area of multi-camera correspondence. The most commonly used methods utilise the epipolar geometry to establish the correspondences. While this method is adequate for some simple situations, extensions to more than just a few cameras are required which are reliable and efficient. In this paper the early stages of research into reliable and efficient multi-camera correspondence method for high speed measurement tasks are reported.

Keywords: correspondence, matching, stereo, epipolar, close-range photogrammetry, real-time, high speed, multi-camera, 3-D measurement.

1. INTRODUCTION

Solutions to the correspondence problem for stereo imagery has been widely researched and applied. General solutions for *m* cameras have only been explored by fewer researchers for instance, Chen et al. (1995) and Faugeras and Mourrain, (1995). For real-time 3-D measurement where reliability and speed of measurement are of paramount importance this subject requires further research.

Computer vision users and developers have often used stereo approaches for acquiring 3-D data. This is largely because many of the tasks that they wish to apply their work to are analogous to human tasks (eg. robot vision) and also because high precision is not normally required. Extensions to trinocular vision have been developed chiefly as a means of reducing image matching errors and increasing the matching reliability (Faugeras and Robert, 1996, Dhond and Aggarwal, 1991). The constraints which are used are often different from those which are typical in convergent photogrammetry. For instance, patches may not be corrected for geometric differences between images because of the closeness of the two or three views.

Photogrammetrists have seldom addressed multi-camera correspondence as a single topic. Is this because the problems have been solved? Or is it because few systems use more than four cameras? It is true that four cameras are adequate for most measurement tasks and that little can be gained by using more cameras that cannot be gained by multiple observations, possibly with further camera roll diversity. However, future tasks are almost certainly likely to benefit by additional views because, in complex manufacturing situations the presence of many occluding objects will require them. It is therefore assumed that a *m* convergent camera correspondence solution is required and that issues of the proven reliability, certification of correspondence, and speed will become important as more real-time photogrammetric systems are implemented.

2. A REVIEW OF CORRESPONDENCE METHODS AND TECHNIQUES

2.1 Overview

Issues concerned with correspondence have often been addressed while solving other problems. Correspondence solving is a means to an end so a thorough search of the literature was necessary to assess the current state of research in this area. The areas where it has been necessary to research the correspondence problem are: 1. remote sensing where stereo images of the surface of the earth are used to construct digital terrain models and maps (Li et al. 1995); 2. close range photogrammetry using film based methods for measurement or mapping of archaeological (Koch, 1994), or industrial artefacts (Petran et al., 1996; Huang et al. 1996); 3. computer vision researchers who require 3-D information concerning a scene for the purposes of robot navigation, image understanding or tracking of moving objects (Venkateswar and Chellappa, 1995); 4. high precision measurement using targeted objects (Baltsavias, 1991); and 5. high speed measurement using dedicated hardware (Clarke et al. 1997). This section attempts to put correspondence research into perspective for those concerned with high precision close-range photogrammetry. The algorithms and methods discussed here are those which are appropriate to multiple-camera correspondence of discrete points of interest (targets) on the surfaces of opaque objects. For the purpose of discussion, the review of correspondence algorithms are separated into 2-D, 3-D, and hybrid methods. Complementary techniques such as use of unique markers, tracking, special constraints, and optimisations are discussed.

2.2. 2-D methods

2.2.1. Parallel optical axes

This is the conventional situation for stereo (Fig. 1). The optical axes of two cameras (V1, V2) are parallel and the image sensors ideally lie in a single plane. The y axes of the local camera co-ordinate systems are parallel. The images of a 3-D point P(X,Y,Z) on V1 and V2 are p(x,y) and p\`(x\`,y\`) respectively. In this configuration the epipolar geometry is simple. The image of the optical centre O2 of V2 on V1 (the epipole of V1) lies at infinity and the same applies for O1 on V2. Hence, the epipolar line *l*1 (the line joining the epipole of V1 and p) and the epipolar line *l*2, (the line joining the epipole of V2 and p\`) are same scan lines. Therefore searching for corresponding points of interest for matching can be performed conveniently along the same scan line in both images. The other advantage is that the computation of the 3-D points is simple. A common situation with stereo is that for a particular point on *l*1 there can be a number of potential matching points on *l*2. In such a situation global matching is considered using techniques such as neighbourhood relationships, disparity gradient limits, relaxation labelling, area based correlation, or other grey scale based methods to solve the ambiguities (Dhond and Aggarwal, 1991). Dhond and Aggarwal (1989) provide a comprehensive review of stereo research particularly the work related to correspondence. Haralick, and Shapiro, 1993 provides detailed description of gray scale matching methods.

This configuration has disadvantages in close-range photogrammetry unless the cameras are located close together or the object of interest is far away (e.g. remote sensing), both cameras may not see all the points of interest in 3-D space. Furthermore, while the X and Y accuracies are good the Z accuracy is poor and there is no redundancy. In most computer vision applications images are usually obtained with only a small base length *b* so that they are radiometrically and geometrically similar except at positions where sharp discontinuities exist in the object space. The computational effort in solving the correspondences is proportional to number of targets under the simplest situation without any ambiguities. Multiple cameras are rarely used in this configuration.

2.2.2. Non-parallel optical axis

It is often not possible or convenient to have stereo pairs with parallel optical axes and it may also be undesirable because of the disadvantage of the configuration with respect to the object coverage, hence convergent views are used. Under these conditions, the solution to the correspondence problem and the computation of 3-D world co-ordinates are not as simple since the orientation of both cameras needs to be included in the solution. There are number of algorithms that have been tested and used to solve the correspondences.

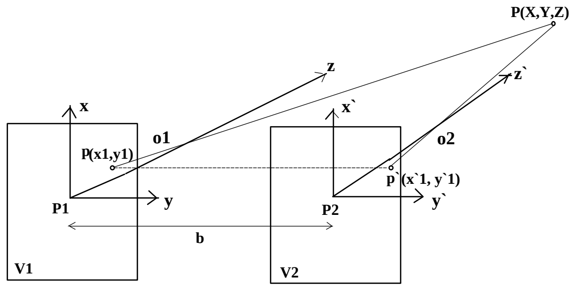

Figure 1. Parallel optical axis stereo pair

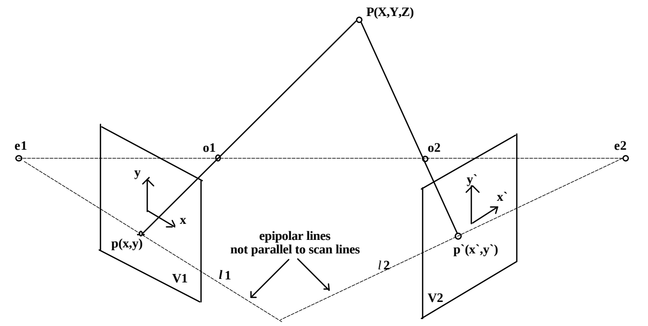

Figure 2. Non-parallel optical axis stereo pair

(a) Standard epipolar line. With non-parallel optical axes (Fig. 2), the epipolar lines are not the scan lines and the epipoles e1 and e2 lie at a finite distance away from cameras. The images of a 3-D point P(X,Y,Z) on cameras V1 and V2 are p and p` respectively. Using the epipolar geometry the equation of the epipolar line *l*2 can be obtained on which p` should ideally lie. In practice, due to the lens distortion point p` would not lie on straight line *l*2. Therefore search is normally performed along a band. This method requires epipolar line equations and a two dimensional search for matching candidates, hence it is computationally expensive. Further developments to this method have been reported. An alternative way of establishing the correspondences by means of a 3x3 matrix known as fundamental matrix has been introduced (Luong et al. 1993). The fundamental matrix contains all the geometric information that is required for establishing the correspondences between two images (Zhang, 1996a). The theory of epipolar geometry has been extended to more than two cameras and Faugeras et al. (1996) reported work for arbitrary number of cameras. Zhang, 1996b reported that the straight epipolar line considered in an ideal situation is in reality replaced by an epipolar curve.

(b) Epipolar line with rectification of images. If the interior and exterior parameters of the cameras are known then it is possible to rectify the images to provide images which have the same benefits as the parallel axis stereo pair for feature matching and 3-D point computation. The rectification process (Hallert, 1960) replaces the original image pair V1 and V2 with images V3 and V4 that lie in a single plane which is parallel to O1O2 (Fig. 3). The pixels or points of interest in the original images are mapped on to the new images (e.g. p(x,y) in V1 to p(u,v) in V3 and p`(x`,y`) in V2 to p`(u`,v`) in V4). The transformation parameters that are necessary can be calculated using some known points in the new images. This method has the advantage of avoiding the epipolar line equation calculations and (*n* x *n*) comparisons for the cost of the transformation of each point in the image where *n* is the number of targets in each image.

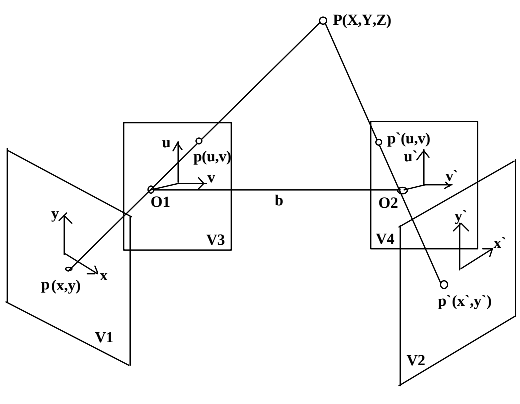

Figure 3. Rectification of convergent images to normal case stereo pair.

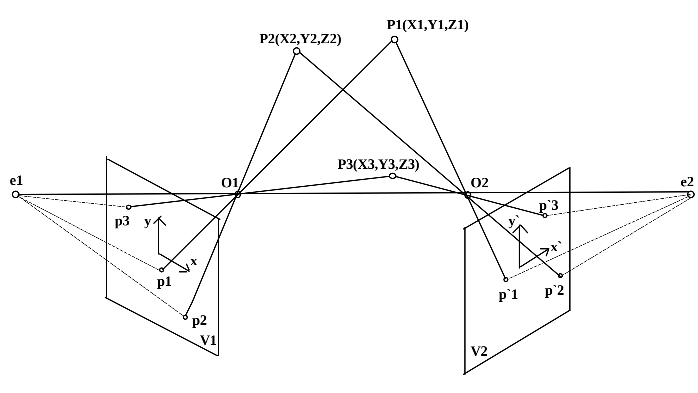

Figure 4. Epipolar line slope comparison.

(c) Epipolar line slope comparison. Another method was reported by Sabel et al. (1993) where a comparison of slopes of epipolar lines between image pairs was used for establishing the correspondences. The slope of an epipolar line is a single quantity which is unique to a particular point in an image with respect to the epipole. If the camera orientation parameters are known, the local co-ordinates of epipoles e1 and e2 of cameras V1 and V2 respectively (Fig. 4) can be determined. Hence, the slopes of epipolar lines e1p1, e1p2, e1p3 of camera V1 and those of e2p1`, e2p2`, e2p3` of camera V2 can be calculated. Using a second order polynomial slopes of lines in V2 can be converted to find the slopes of the corresponding lines in V1. By comparing the converted slopes of image V2 with the slopes of lines in image V1, correspondence can be established. This method requires an initial calibration process to estimate the unknown parameters of the polynomial. It can be extended for more than two cameras. A slope related to each point in V1 is compared with the slopes in V2, hence this method requires (*n* x *n*) comparisons.

2.3. 3-D method

Correspondences can be established by looking at the intersection between rays projected into object space from features or points of interest in each view (Chen et al. 1993; Chen et al. 1995; Clarke et al. 1995). Ideally, a ray projected into space from a point in one image will intersect with a ray projected from the corresponding point in another image. But in reality these rays intersect with a small distance, d (i.e. perpendicular distance) between them due to the distortion introduced by the lens and other errors (Fig 5). The rays from corresponding target images intersect with the minimum distance hence, can be distinguished. Algebraic (Chen et al. 1993) as well as vector (Sabel et al. 1993) based methods are available for obtaining the minimum distance. The ambiguous situations that could arise due to the same minimum distance produced by more than one point in V2 with a single point in V1 can be solved by using additional cameras or other 2-D techniques. It was reported by Chen et al. (1993) that this method has been successfully tested for more than 20 views. This method requires (*n* x *n*) calculations of minimum distances and comparisons.

2.4. Hybrid method

Back projection can be used as a means of propagating correspondences from a pair of corresponded views onto a third, fourth or more views (Fig 6). Initially a pair of cameras are selected and the correspondences are established. Any of the 2-D or 3-D methods described previously may be used. Then an estimate of the location of points in 3-D space can be calculated using either vector or DLT (direct linear transformation) based methods. While this 3-D estimation may not be very good it is likely to be sufficient for the purposes of correspondence solving provided that the object point density in 3-D space is low. The next step is to back project from the 3-D points into other views. By performing a 2-D image space search around the 2-D points computed by the back-projection, corresponding points can be found. This is inherently a method for multi-camera correspondence solving and has been successfully used in many close-range photogrammetric applications (Fraser, 1997). However, the disadvantage of this method is that if a 3-D point of interest is not in the field of view of all cameras or is partially occluded, then the correspondence may not be established reliably.

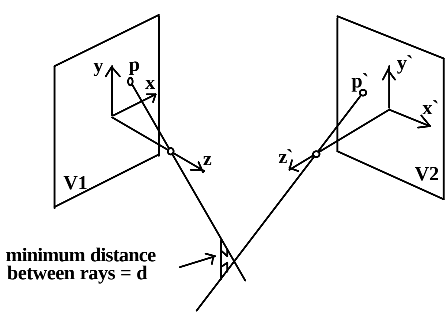

Figure 5. 3-D intersection

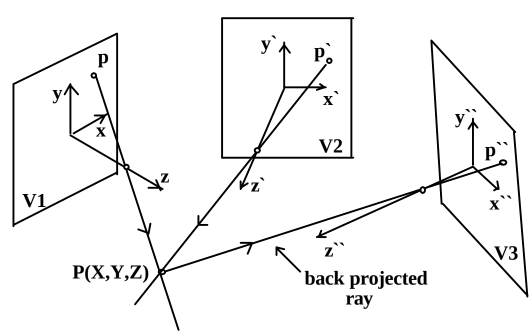

Figure 6. Back projection

2.5 Complementary Techniques

(a) Direct solution using unique markers

This is a solution which may be used in various situations to solve the correspondence problem. Coded targets or other unique markers (e.g. colour) are used to uniquely identify the features. These targets have taken many forms, vertical bar codes (Wiley and Wong, 1992), circular bar codes (Van der Heuvel et al. 1992), or square features (Sharp, 1997). The target is identified to each camera by use of a unique code which is observable either in close proximity to, or surrounding the target used for location purposes. Each camera recognises each target such that after collection the target co-ordinates can be used directly. A disadvantage of this method is the size of the feature that needs be recognised in the image. Such target images will also take longer to process and may have a detrimental effect on the usable angle of view due to perspective distortion. The image processing time can be traded against the computational time for other correspondence solving methods and will just be proportional to number of targets.

(b) Tracking

Tracking has been extensively researched those working in the motion analysis area. Researchers have been using tracking to solve correspondences between features of interest in a sequence of images. Research reported have mostly used a single camera and tracked moving objects (Robert and Charnley, 1993). Wang and Duncan (1996) reported the use of image sequences from a stereo pair for recovering the 3-D motion of multiple moving objects. In a multi-camera environment tracking is useful for maintaining the correspondences among cameras as solving correspondences for each set of images is computationally expensive. Hence only for appearing and disappearing targets correspondences need to be considered. Tracking can be carried out in variety of ways. The Kalman filter is an often used technique. Other methods based on velocity/direction predictors have been reported (Robert and Charnley, 1993). All such methods rely on the fact that the time between images is short enough to allow for tracking. It is likely that in all cases it will be beneficial to use additional evidence such as would be available if a group of targets were fixed to a solid object. So that for correspondence solving and tracking purposes a group of targets may be considered, hence increasing the efficiency. Grouping would become more efficient when more than one moving object is present in the scene.

(c) Constraints

CAD model to constrain search. Streilein (1996) developed the concept of using approximate CAD models to semi-automate the process of building up a full CAD model of architectural and industrial object. The approximate CAD model is input by the operator and, using the constraints thus imposed, the software is able to use computer vision and photogrammetric techniques to produce a refined model of the object. Such methods can be useful in that the initial correspondence is solved by the operator using stereo fusion or human intelligence and then features such as edges are used to create CAD elements. The bundle adjustment may then be used to provide a rigorous and redundant solution.

Use of object features such as: patches, edge segments, shared edges, tangents. Huang et al. (1996) used the length of lines or diameters of ellipses to compare between models. Graph matching methods were then used to find pairs between models and measured object features. Finally two further constraints were considered: orientation relationships using surface normals which had to satisfy angular requirements; and proximity relations, e.g. the distance between surfaces had to again satisfy the requirements of the model for a correct match. Brenner and Hahn (1996) discuss the use of active exploration using a range of sensors for object recognition and gauging in a manufacturing process. They postulated that the object recognition could be accomplished using global properties such as volume, roundness, or higher order moments to form a vector of parameters so that matching is performed in parameter space.

Smoothness of velocity fields. Maas (1995). discussed the determination of velocity fields in flow tomography sequences by 3-D least squares matching. A powerful laser sheet was scanned through a small volume of liquid which was seeded with fluorescein. The images were collected of the scanned sheets thus building up a 3-D picture of the turbulent flow that was being investigated. This work is of significance here not so much because or the direct 3-D correspondence methods but because of the constraints that were used to maintain tracking of voxels. Temporal constraints of smoothness of the velocity fields it was assumed that sequential data sets were correlated and that at least three data sets could be used. Intensity constraints, it was assumed that a voxel would maintain the same global level of intensity between adjacent voxel sets. Many other forms of constraint such as figural continuity, disparity gradient limit, cross-channel activity, left to right and right to left connections have been used. However, these are more applicable to image matching techniques after the correspondence constraint has been used so are not discussed not further in this paper.

(d) Optimisations

Tree searching. Chen et al. (1995) discussed how tree searching methods could be used to improve matching speed and efficiency. Chen used a 3-D method of minimum distance calculation to establish correspondences. The usual technique is to consider each point in one image with all the points in the other image to find the corresponding point. But with more than two images this method is inefficient. Chen optimised the matching by considering groups of points rather than individual points by using tree searching.

Hierarchical methods. Venkateswar and Chellappa (1995) discussed feature based hierarchical stereo matching. The hierarchy consisted of lines, vertices, edges, and surfaces. The matching started at the highest level of the hierarchy (i.e. surfaces) and proceeded down to the lowest level (i.e. lines). Matching becomes easier at the highest level as numbers to be matched are small. Also it constrains the matches at the lower levels hence making the matching process more efficient.

Use of look up tables. Thieling and Ameling (1993) described an active triangulation method for fast 3-D industrial measurement using a projector and a single CCD camera. A table was constructed in which 3-D co-ordinates were computed for every image location. This table was limited in size by the use of interpolation. The process of measurement occurred in two stages: off line and on line. The off line stage proceeded by calibration of the cameras, computation of light projection lines, calculation of epipolar lines, followed by calculation of table entries. The on line stage produced the image locations and via the look up table calculated the 3-D co-ordinates.

Iterative improvement of camera parameters. Chen et al. (1993) iteratively solved the correspondence at the same time as improving the estimation of the camera parameters. Forlani et al. (1996) discussed a semi-automated means of searching for homologous points. Multiple adjustments were used to refine camera parameters until the maximum number of targets were found for each set of three images. At this point another three images were used to find further correspondences and off-line computation of once-only parameters.

2.6 Conclusion

Various methods and complementary techniques for correspondence solving have been discussed. The reliability of the matching is an important issue and has been tested on real and calibrated scenes using binocular and trinocular images (Dhond and Aggarwal, 1989). For matching points the disparity error, that is the difference between y-disparity obtained from a pair of images and the true disparity using known depth, was used to judge the goodness of the matches. Fraser, 1997 reveals that for close-range photogrammetry, there is still a need for a comprehensive solution and few ,if any, comparisons of the efficiency of close-range correspondence methods have been reported. The next section considers the relative computational efficiencies of selected methods .

3. ANALYSIS OF THE COMPUTATIONAL COMPLEXITY OF ALGORITHMS USED FOR SOLVING THE CORRESPONDENCE PROBLEM

The previous section was devoted to the discussion of the algorithms used for solving the correspondence problem. This section presents an analysis of the time taken by each algorithm using a range of target and camera numbers under pre-defined circumstances. The computational effort required is a major characteristic of these algorithms. In general, as the number of cameras and/or targets increases the computational time increases. If occlusions and ambiguities are present then further computations are required to solve correspondences. The computational analysis was based on the number and the type of computations (additions, subtractions, multiplications, divisions, square root, & comparison) required in each algorithm. The time taken by each of these operation was estimated using a Sun Sparc Classic workstation and each operation was given a weighted time. Time consumed by tasks that are common to all the algorithms such as the computation of the rotation matrix parameters of each camera and accessing the data structures for obtaining 2-D target co-ordinates (e.g. tree searching) were not included in the time equation. It was assumed that the camera parameters both internal and external are accurately known and that the number of targets (*n*) seen by each camera was the same.

Five methods of correspondence solving were analysed: **(1) 3-D space intersection** method used the algebraic means of obtaining the minimum distances between two optical rays. It required rotation matrix parameters and 2-D co-ordinates of target image centroids. **(2) Epipolar line** method used a simple way of obtaining the equation of epipolar line in one view corresponding to a point in another view. This required 2-D co-ordinates of the epipole and another arbitrary point on the epipolar line. In order to obtain this arbitrary point an arbitrary Z value in 3-D space was used and obtained the corresponding X,Y using co-linearity equations. Then, using co-linearity equations, a point on the epipolar line was obtained which corresponded to the arbitrary 3-D point. **(3) Rectified Image** method used a transformation to map the target co-ordinates from original to new images. **(4) Epipolar Line Slope** method used the procedure as discussed in the previous section. **(5) Back-Projection** method used 3-D intersection based correspondence solving and vector based 3-D point estimation. The co-linearity equations were then used to obtain the 2-D co-ordinates on other views.

A set of curves were obtained for the illustration of the results of this analysis.

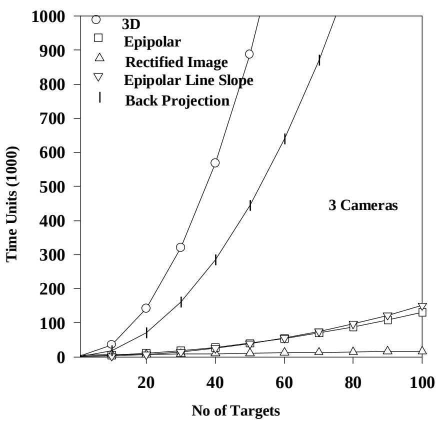

Figure 3.1 Time each algorithm takes for different number of targets with 3 cameras.

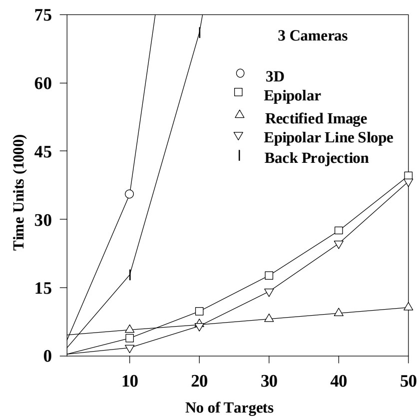

Figure 3.2 Highlight of the time taken by each algorithm at the beginning of the process.

Fig. 3.1 illustrates the weighted time units taken by each method where the number of cameras used was 3. The 3-D intersection algorithm consumes more time than any other method and increases sharply as the number of targets in each view increases. The reason for this is that each light ray projected into object space from a particular view is considered with all the rays projected from another view and the shortest distances between rays are calculated. It requires (*n* x *n*) shortest distance calculations hence, the time function is proportional to (*n* x *n*). The back projection algorithm that is commonly used by photogrammetrists falls in second place. For each 2-D point estimated by a back projected ray from a 3-D point in space, (*n* x *n*) 2-D comparisons need to be completed. Furthermore, obtaining the correspondences between the initial pair of views is also proportional to (*n* x *n*). Hence this method is computationally more efficient than the 3-D method. The Epipolar Line and the Epipolar Line Slope methods follow similar curves, and are more efficient than the 3-D and the Back Projection methods. Finally, the Rectified Image method shows different characteristics from all other methods. As the number of targets increases, the time consumed rises linearly where as the other methods have a non linear rise. The reason is because after a relatively expensive process of computing parameters for the rectification transformation, the time for target location transformation and matching are proportional to *n* as opposed to (*n* x *n*) in other methods.

Fig. 3.2 highlights the differences in computational effort required by each method at the beginning of the correspondence solving process. The Rectified Image and the 3-D methods have the highest and the lowest efforts respectively.

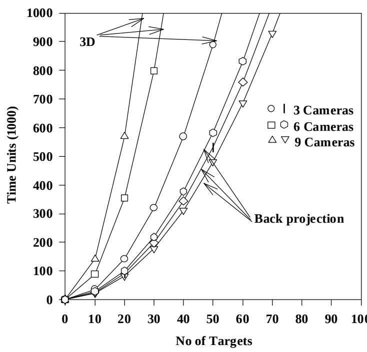

Figure 3.3 Comparison of time taken by the 3D and the Back Projection methods with 3, 6, and 9 cameras.

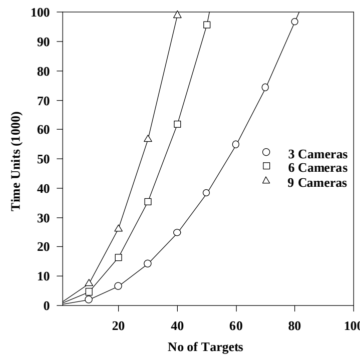

Figure 3.4 Comparison of time taken by the Epipolar line with 3, 6, and 9 cameras.

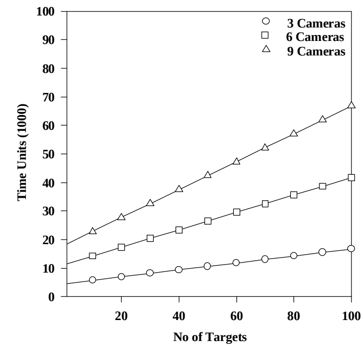

Figure 3.5 Comparison of time taken by the Rectified Image method with 3, 6, and 9 cameras.

The figures 3.3-3.5 illustrate the increase in computational time for each method when the number of cameras is increased. The 3D, Back Projection, and Epipolar Line methods do not exhibit a large change in the initial computational effort when the number of cameras is increased. On the contrary, the Rectified Image method shows a considerable change in this figure as expected due to the initial computational effort. Furthermore, the Back Projection method ,in general, shows a smaller variation in the computational time for different number of cameras compared to other methods.

In conclusion, this analysis illustrates important details for selecting a suitable algorithm for correspondence solving in a multi-camera environment where speed is critical. These initial results appear to indicate several important conclusions. First the 3-D method is inherently expensive and has the undesirable characteristics of a non-linear method. Second, the Back Projection used is in addition to another method and is expensive by comparison with Epipolar Line methods. Only in the case where Back Projection is used in conjunction with searching target images in the image itself could this method be considered as efficient since this removes (*n* x *n*) searching. Third, the Epipolar method used with or without the comparison of slopes is an efficient method but suffers because of the (*n* x *n*) co-ordinate searching problem. Finally, the Rectified Image method is the most efficient and has the best characteristics of all the methods because non of the components in the time equation are proportional to (*n* x *n*).

For cases where more than 20 targets must be corresponded it appears that the rectification method is the most efficient. If rectification parameters can be carried forward then this method will be efficient for fewer targets too. If 1-20 targets require correspondence then the Epipolar Line method may be the most efficient. However given that the time taken for either method is low for 20 targets it may not make a significant difference which method is used. In many situations it is required to track moving targeted objects. The problem in this situation is that some targets may move into, and others may move out of the field of view. Hence, a fast correspondence solution is required for appearing and disappearing targets to maintain continuous tracking. The Rectified Image method would be the most suitable for such a situation provided that the cameras are stable so that the same transformation parameters can be carried forward.

4. CONCLUSIONS

This paper has reviewed both computer vision and photogrammetric research and has considered how the research of the past can be best utilised to provide efficient correspondence solution for real-time 3-D measurement. To achieve this the computational complexities of the available methods have been analysed and using this information research for a correspondence algorithm has been started which it is hoped will provide fast and highly reliable correspondences for *m* camera real-time measurement situations.

5. ACKNOWLEDGEMENTS

Mr. D.D.A.P. Ariyawansa would like to acknowledge the support of The Worshipful Company of Scientific Instrument Makers in providing a PhD study bursary.

6. REFERENCES

Baltsavias, E. P. 1991. Multiphoto geometrically constrained matching. PhD Thesis, pp. 1-40.

Brenner, C. and Hahn, M. 1996. Object recognition using multi-sensor fusion and active exploration. ISPRS Vol. XXXI, Part B5. pp. 59-64.

Chen, J. Clarke, T.A. & Robson. S., 1993. An alternative to the epipolar method for automatic target matching in multiple images for 3-D measurement. Optical 3-D measurements techniques II, Pub. Wichmann, Karlsruhe: 197-204.

Chen, J. Clarke, T.A. Cooper, M.A.R. Grattan, K.T.V.G. 1995. An optimised target matching based on a 3-D space intersection and a constrained search for multiple camera views, Videometrics III. SPIE Vol. 2350. Boston. pp. 324 - 335.

Clarke, T.A. Cooper, M.A.R. Chen, J. and Robson, S. 1995. Automated three dimensional measurement using multiple CCD camera views. Photogrammetric record, 15(85):27-42.

Clarke, T.A. Gooch, R.M. Ariyawansa, D.D.A.P. and Wang, X. 1997. 3D-Net - the development of a new real-time photogrammetric system. Videometrics V. SPIE Vol. 3174. San Dieago.

Dhond, U.R. and Aggarwal, J.K. 1989. Structure from stereo - A review. IEEE Trans. on Systems, man and Cybernetics, 19(6): 1489-1510.

Dhond, U.R. and Aggarwal, J.K. 1991. A cost-benefit analysis of a third camera for stereo correspondence. Int. Journal of Computer Vision. 6:1, pp. 39-58.

Faugeras, O. and Mourrain, B. 1995. On the geometry and algebra of the point and line correspondences between N images. Research Report No-2665, INRIA Sophia-Antipolis, BP 93 Sophia-Antipolis, Cedex, France.

Faugeras, O. and Robert, L. 1996. What can two images tell us about a third one? International Journal of Computer Vision, 18. 5-19.

Forlani, G. Giussani, A. Scaioni, M. and Vassena, G. 1996. Target detection and epipolar geometry for image orientation in close-range photogrammetry. ISPRS Vol. XXXI, Part B5. pp. 518-523.

Fraser, C.S. 1997. Automation in digital close-range photogrammetry. First transaction. Tasman surveyors conference. Published on Internet.

Hallert, B. 1960. Photogrammetry: Basic principles and general survey. Published by McGraw-Hill civil engineering series. pp. 104-113.

Haralick, R.M, and Shapiro, L.G. 1993. Computer and Robot Vision Volume II. Published by Addison-Wesley Publishing Company Inc. USA. 630 pages.

Huang, Y. Trinder, J.C. and Donnelly, B.E., 1996. Object recognition for a flexible manufacturing system. ISPRS Vol. XXXI, Part B5. pp. 253-258.

Koch, R, 1994. Model-based 3-D scene analysis from stereoscopic image sequences. ISPRS Journal of Photogrammetry and Remote Sensing, 49(5):23-40.

Li, D. Wang, S. Li, R. 1995. Automatic quality diagnosis in DTM generation by digital image matching techniques. Geomatica, 49(2):65-73.

Luong, Q. Deriche, R. Faugeras, O. Papadopoulo, T. 1993. On determining the fundamental matrix: Analysis of different methods and experimental results. Research Report No - 1894, INRIA Sophia Antipolis, BP 93 - Sophia-Antipolis, Cedex, France.

Maas, 1995. Determination of velocity fields in flow tomography sequences by 3-D least squares matching. Optical 3-D measurements techniques II, Pub. Wichmann, Karlsruhe: pp. 366-376.

Petran, F. Krzystek, P and Bonitx, P. 1996. CAD-based reverse engineering with digital photogrammetry. ISPRS Vol. XXXI, Part B5. pp. 475-480.

Robert, J. and Charnley, D. 1993. Attentive visual tracking. BMVC Vol. II, pp. 459-468.

Sabel, J.C. van Veenendaal, H.L.J. and Furnee, 1993. PRIMAS, a real time 3D motion analysis system. Optical 3-D Measurement Techniques II. Published by Wichmann, Karlsruhe, pp. 530-537.

Sharp, K., 1997. Fixed-mount scanners. IO systems. European edition, May. pp. 20-30.

Streilein, A.1996. Utilisation of CAD models for the object oriented measurement of industrial and architectural objects. ISPRS Vol. XXXI, Part B5. pp. 548-553.

Thieling, L. and Ameling, W. 1993. An active triangulation based method for fast 3-D industrial measurement. Optical 3-D measurements techniques II, Pub. Wichmann, Karlsruhe, pp. 394-401.

Van den Heuval F.A., Kroon, R.J.G.A., and Re Poole, R.S., 1992. Digital close-range photogrammetry using artificial targets. ISPRS Vol. 29, Part B5, pp. 222-229.

Venkateswar, V. and Chellappa, R., 1995. Hierarchical stereo and motion correspondence using feature groupings. International Journal of Computer Vision, 15. pp. 245-269.

Wang, W. and Duncan, J.H. 1996. Recovering the three-dimensional motion and structure of multiple moving objects from binocular image flows. CVIU Vol. III No. 3, May 1996, pp. 430-446.

Webb, J.A. Warfel, T. Bing Kang, S. 1994 .A scalable video rate camera interface. 18 pages, published on the Internet as publicaton CMU-CS-94-192.

Wiley, A.G. and Wong, K.W. 1992. Geometric calibration of zoom lenses for computer vision metrology. ISPRS Vol. 29, Part B5, pp. 587-593.

Zhang, Z. 1996a. Determining the epipolar geometry and its uncertainty: A review. Research Report No - 2927, INRIA Sophia-Antipolis, BP 93 Sophia-Antipolis, Cedex, France.

Zhang, Z. 1996b. On the epipolar geometry between two images with lens distortion. ICPR Vol. I, pp. 407-411.

PAPER REFERENCE

Ariyawansa, D.D.A.P, and Clarke, T.A. 1997. High-speed correspondence for object recognition and tracking. SPIE Vol 3174, Videometrics V, pp. 70-79.