ENHANCED SINGLE CAMERA PHOTOGRAMMETRY ALGORITHMS FOR REAL-TIME CONTROL APPLICATIONS

R.C. Hughes
National Aeronautical Establishment
National Research Council of Canada
Ottawa, Ontario, CANADA, K1A 0R6

ABSTRACT

A methodology is presented for developing photogrammetry algorithms which provide high accuracy relative position and attitude information and which can run at real-time video camera update rates on microprocessor based systems. A discussion of the performance and operating bounds of a series of algorithms is presented.

INTRODUCTION

Single camera photogrammetry can establish the full six degrees-of-freedom relations (i.e., position and attitude) of an object with respect to a camera, as long as the object has suitable reference markings or control points in view. A typical situation is illustrated in Figure 1, where the "object" could be a satellite or payload for space applications or could be a part on an assembly line for industrial applications. The reference markings could be structural features, such as holes or corners of panels or alternatively could be high contrast target elements specifically applied for photogrammetric purposes.

For real-time control applications, the emphasis is on establishing the position and orientation of an object with respect to the camera in an on-line mode at high update rates. In a companion paper, Pinkney and Perratt (1986) describe a photogrammetry based guidance and control system which uses video camera technology and custom image processing hardware to provide complete position and attitude information at a 30 Hz update rate. In this paper we are primarily concerned with the single camera photogrammetry algorithms that go into such a system.

A number of single camera resection algortihms have already been developed. Approximate iterative solutions have been presented for a square target array (4 elements in a plane) by Pinkney (1978) and Kratky (1979). Kratky (1979) has also presented a linearized iterative least squares approach for use with arbitrary target arrays. More recently Fischler and Boles (1981), Tietz and Germann (1982) and Hung, Yeh and Harwood (1985) have presented direct (i.e. closed form) solutions for 3 and 4 elements-in-a-plane target arrays.

For a one time solution, direct or closed forms frequently take less computer time than an iterative approach for the same accuracy. However, for a real-time control problem where successive solutions are produced at a high update rate, iterative solutions can be very efficient as the previous solution provides a good starting point for the current iterations and usually only a single iteration is required at each update period.

Of the iterative solutions available, the approximate solutions provide minimal computer loading but only produce acceptable results over a narrow range of orientation angles due to the use of trignometric approximations and the neglecting of some cross-coupling terms. The standard Gauss-Newton iterative least squares approach offers high accuracy, large operating range and greater flexibility but is too computationally demanding for some applications.

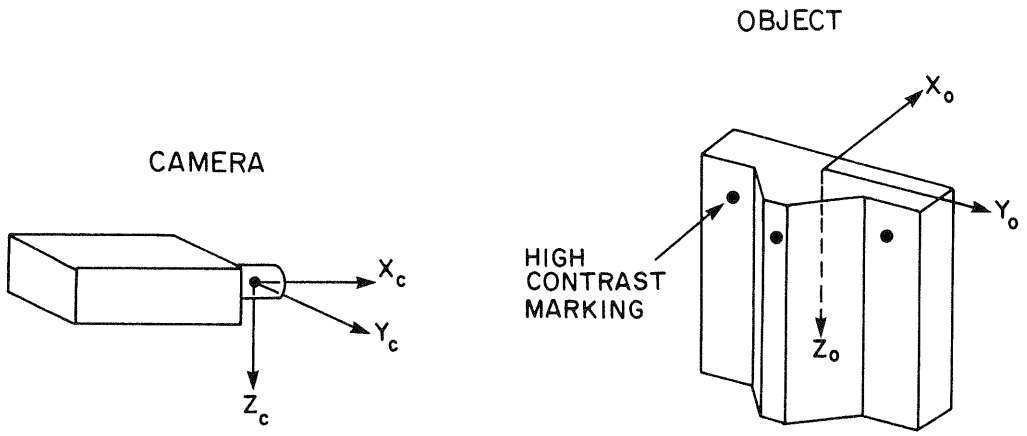

FIG. 1: SINGLE CAMERA PHOTOGRAMMETRY FOR CONTROL APPLICATIONS

The purpose of this paper is to introduce a new approach for generating exact (as opposed to approximate) iterative photogrammetry algorithms which offer improved accuracy and extended operating range over the approximate solutions but at the same time execute 3 to 5 times faster than the more general least squares approach. The term "exact" is used because these algorithms iterate to a solution which exactly satisfies the photogrammetric collinearity equations under noise free conditions, as does the least squares approach. At the same time an increase in speed over a Gauss-Newton least squares approach is achieved by a tailoring process in which the algorithm is optimized for speed about a particular operating point by judicious use of symmetries inherent in a particular choice of target array.

This tailoring methodology was originally developed by Pinkney (Pinkney and Perratt, 1986) and has been further refined and evaluated as described in this paper.

COLLINEARITY EQUATIONS

With aerospace applications in mind, we will work with a camera coordinate system (denoted by subscript C) with origin at the camera projection centre and with X_C axis along the camera optic axis as illustrated in Figure 2. We will assume that the object for which we desire position and attitude information has its own coordinate system (denoted by subscript O) which is used to specify the relative positions of target elements on the body.

Our overall objective in positioning and control applications is to establish, using photogrammetric techniques, the location in camera coordinates of the object coordinate origin (denoted by components X_{CO}, Y_{CO}, Z_{CO}) and the orientation of the object coordinate system with respect to the camera coordinate system, as expressed by a rotation matrix (denoted by A_{CO}) or by a set of Euler angles (denoted by yaw, ψ, pitch θ, and roll, ϕ, applied respectively about the camera Z_C, Y_C and X_C axes).

The position on the object of the jth target element (see Figure 2) is presumed to be known in the object coordinate system (e.g. $X_O(j)$, $Y_O(j)$, $Z_O(j)$). However, this position in space can also be expressed in the camera coordinate system as given by

$$\begin{bmatrix} X_C(j) \\ Y_C(j) \\ Z_C(j) \end{bmatrix} = \begin{bmatrix} X_{CO} \\ Y_{CO} \\ Z_{CO} \end{bmatrix} + A_{CO} \begin{bmatrix} X_O(j) \\ Y_O(j) \\ Z_O(j) \end{bmatrix} \quad (1)$$

where X_{CO}, Y_{CO}, Z_{CO} and matrix A_{CO} are at present unknown quantities which must be determined. The rotation matrix A_{CO} can be expressed in terms of both direction cosines α_{ij} or our choice of Euler angles ψ, θ, ϕ as given by

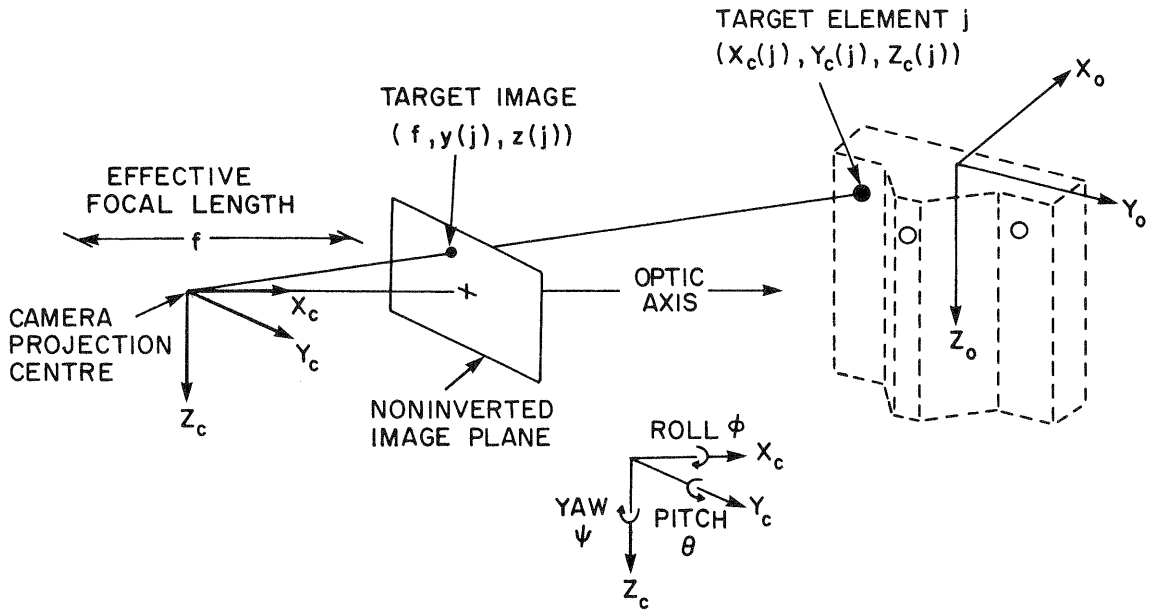

FIG. 2: CAMERA COORDINATE SYSTEM

$$A_{CO} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} \tag{2a}$$

$$= \begin{bmatrix} \cos\theta\cos\psi & -\cos\phi\sin\psi + \sin\phi\sin\theta\cos\psi & \sin\phi\sin\psi + \cos\phi\sin\theta\cos\psi \\ \cos\theta\sin\psi & \cos\phi\cos\psi + \sin\phi\sin\theta\sin\psi & -\sin\phi\cos\psi + \cos\phi\sin\theta\sin\phi \\ -\sin\theta & \sin\phi\cos\theta & \cos\phi\cos\theta \end{bmatrix} \tag{2b}$$

In actual practice, a video camera produces an inverted image of the object and target elements, on a photo-active element located behind the camera projection centre. However, for our purposes, it is easier to represent a non-inverted image plane at a distance f (the effective focal length) ahead of the projection centre as illustrated in Figure 2. In any event the point projection operation of an idealized camera implies that vectors from the projection centre to the target element and to the target element image are proportional, resulting in the standard collinearity equations given for the jth element by (using Equation 1)

$$y(j) = f\left[\frac{Y_{CO}+\alpha_{21}X_O(j)+\alpha_{22}Y_O(j)+\alpha_{23}Z_O(j)}{X_{CO}+\alpha_{11}X_O(j)+\alpha_{12}Y_O(j)+\alpha_{13}Z_O(j)}\right] \tag{3a}$$

$$z(j) = f\left[\frac{Z_{CO}+\alpha_{31}X_O(j)+\alpha_{32}Y_O(j)+\alpha_{33}Z_O(j)}{X_{CO}+\alpha_{11}X_O(j)+\alpha_{12}Y_O(j)+\alpha_{13}Z_O(j)}\right] \tag{3b}$$

where y(j) and z(j) are camera coordinate system image plane measurements.

ITERATIVE SOLUTIONS

For N target elements, the 2N collinearity equations of the form of Equation 3 can be expressed in a vector form as given by

$$\underline{u} = \underline{h}(\underline{t}) + \underline{v} \tag{4}$$

where $\underline{u}$ is a 2N length vector of photo-coordinate measurements, $\underline{v}$ is a corresponding vector of measurement errors, $\underline{t}$ is a 6 element unknown parameter vector (consisting of

X_{CO}, Y_{CO}, Z_{CO} and either ψ, θ, ϕ or a suitable set of 3 independent direction cosines) and finally $\underline{h}(\underline{t})$ is a nonlinear vector function of $\underline{t}$ which represents the right hand side of Equation 3. Since the problem is nonlinear and frequently over specified (i.e. $2N > 6$), it is not possible in general to directly invert Equation 4 in order to solve for $\underline{t}$ in terms of $\underline{u}$. However, iterative solutions to this class of problem can usually be found.

As an example of an iterative solution, the Gauss-Newton least squares approach first linearizes Equation 4 about a current best guess $\underline{t}_k$ as given by

$$\underline{u} \simeq \underline{h}(\underline{t}_k) + B_k(\underline{t}-\underline{t}_k) \tag{5a}$$

$$B_k = \frac{\partial \underline{h}}{\partial \underline{t}}\bigg|_{\underline{t}=\underline{t}_k} \tag{5b}$$

The unweighted least squares iterative solution of Equation 5 is then

$$\underline{t}_{k+1} = \underline{t}_k + [B_k^T B_k]^{-1} B_k^T [\underline{u}-\underline{h}(\underline{t}_k)] \tag{6}$$

where k is the iteration index.

A general iterative solution takes the form given by

$$\underline{t}_{k+1} = \underline{f}(\underline{u},\underline{t}_k) \tag{7}$$

where $\underline{f}$ is some possibly nonlinear function of the measurements $\underline{u}$ and the previous parameter estimate $\underline{t}_k$. The iteration process represented by Equation 7 will converge in a region for which the magnitude of the gradient of $\underline{f}$ with respect to $\underline{t}$ is less than unity. Although there is not the space to elaborate here, there are recursive filtering techniques which frequently can modify the function $\underline{f}$ in order to reduce the absolute gradient and thereby not only expand the range of convergence but also speed up the rate.

Our method of algorithm tailoring is to assemble iterative photogrammetry solutions of the form of Equation 7 by solving for the 6 independent position and attitude components from 6 independent linear combinations of the 2N simultaneous collinearity equations. In practice, a nominal operating point for the algorithm is first chosen and then the dominant effects on the camera image of the particular target array, due to various spatial translations and attitude changes about the operating point, are analyzed in order to choose suitable linear combinations of the collinearity equations. This process is best understood by following through an example in the next section.

PEG ARRAY ALGORITHM

We will now illustrate the tailoring process for a 3 element peg target array. The peg array consists of two base target elements in the plane of the target array and a centre target element (on a "peg") projecting either into or out of the target array plane. A horizontal peg array is illustrated as three target elements in Figure 1.

In this example we will use a vertically oriented peg target array with the three target elements (numbered 1, 2 and 3 respectively) located in the object coordinate system (i.e. $X_O(j)$, $Y_O(j)$, $Z_O(j)$) at the coordinates of (0,0,-a), (d,0,0) and (0,0,a) where d is the depth of the peg element out of the plane and 2a is the separation of the base elements. For these element locations the corresponding collinearity equations (see Equation 3) can be written as

$$y(1)[X_{CO}-a\alpha_{13}] = f[Y_{CO}-a\alpha_{23}] \tag{8a}$$

$$z(1)[X_{CO}-a\alpha_{13}] = f[Z_{CO}-a\alpha_{33}] \tag{8b}$$

$$y(2)[X_{CO}+d\alpha_{11}] = f[Y_{CO}+d\alpha_{21}] \tag{8c}$$

$$z(2)[X_{CO}+d\alpha_{11}] = f[Z_{CO}+d\alpha_{31}] \tag{8d}$$

$$y(3)[X_{CO}+a\alpha_{13}] = f[Y_{CO}+a\alpha_{23}] \tag{8e}$$

$$z(3)[X_{CO}+a\alpha_{13}] = f[Z_{CO}+a\alpha_{33}] \tag{8f}$$

In order to establish appropriate linear combinations of these equations we first turn to Figure 3 where the camera image of such a vertical peg array is illustrated for each of a yaw, pitch and roll of the target array with respect to the camera.

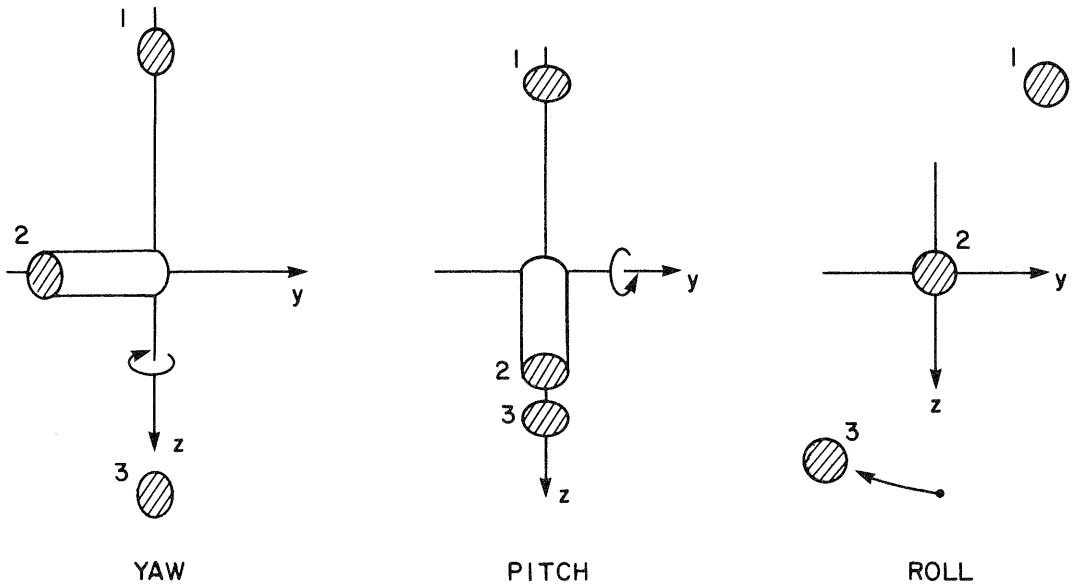

FIG. 3: EFFECT OF YAW, PITCH AND ROLL ON IMAGE OF A VERTICAL PEG TARGET ARRAY

With reference to Figure 3, we observe that: the image separation of the base elements is a good indication of the range (i.e. X_{CO}) of the target array; the mean image position of the base elements is a good indication of lateral displacement (i.e. Y_{CO} and Z_{CO}); the y and z displacement of the peg element image from the mean image base position are good indications of yaw and pitch respectively; and finally the difference in the y components of the two base element images is a good indication of roll. An approximate solution could use these listed dependencies to directly estimate (X,Y,Z,ψ,θ,ϕ). However, we only use them to guide the choice of linear combinations.

As an example, an iterative solution for X_{CO} can be found by subtracting Equations 8b from 8f and solving for X_{CO} to give

$$X'_{CO} = a[2f\alpha_{33}-\alpha_{13}(z(3)+z(1))]/[z(3)-z(1)] \tag{9}$$

where the primed quantity is the latest iterative update. In a similar manner, expressions can be derived for the other translational components Y_{CO} and Z_{CO}.

Rather than working directly with Euler angles, ψ,θ and ϕ, as originally used by Pinkney (1978), we now choose a group of three independent direction cosines to represent the object attitude. Any additional direction cosines appearing in the collinearity equations then become dependent parameters which are determined from the three independent direction cosines through the standard relations among the elements of an orthogonal rotation matrix. For the example of a vertical peg target array, the A_{CO} matrix component α_{21} is selected as the direction cosine most closely associated with yaw angles (see Equation 2). We can solve for α_{21} in an iterative form from suitable combinations of Equations 8a, 8c and 8e as given by

$$\alpha'_{21} = [2(X_{CO}+d\alpha_{11})y(2) - (X_{CO}-a\alpha_{13})y(1) - (X_{CO}+a\alpha_{13})y(3)]/(2df) \tag{10}$$

Iteration expressions are derived in a similar manner for the other two direction cosines most closely associated with pitch and roll. To complete the iteration cycle, the dependent direction cosines are evaluated from the selected independent triad and the most current parameter estimates (e.g. the left hand sides of Equations 9 and 10) become input parameters for the next iteration (e.g. the right hand sides of Equations 9 and 10). If yaw, pitch and roll Euler angles are needed as outputs from the algorithm they are derived from the direction cosines (e.g. using Equation 2).

At this point we would like to demonstrate the performance of our tailored iterative photogrammetry algorithms. As a suitable error metric on which to base comparisons between algorithms we have chosen to use a measure of the root-mean-square target element positioning error in 3 dimensional object space. In order to evaluate this metric we first back project the positions of the individual target elements based on the current solution (i.e. X_{CO}, Y_{CO}, Z_{CO}, A_{CO}) and then calculate the root-mean-square positioning error among the target elements. An advantage of this measure is that it includes both the positioning and attitude errors of the solution.

Our test example for the peg array consists of a vertical peg with a base separation 2a equal to 56 cm and a peg depth d equal to -20 cm (where the minus sign indicates a peg pointed towards rather than away from the camera) at a nominal range of 3 m from the camera. We assume that the camera has a focal length of 12.5 mm and an image plane width and height (y and z dimensions) of 12.9 by 9.7 mm which are divided respectively into 640×480 pixels. An ensemble error statistic was estimated on a root-mean-square basis over 300 separate trials. Each trial was simulated on a computer by feeding the algorithm "noise corrupted" image plane measurements. These were generated by first calculating the true target element image positions (from the collinearity equations) for the particular camera to target geometry and then adding in random error components drawn from a Gaussian distribution with a variance equivalent to 0.1 pixels in the image plane. For each trial, the algorithm under test used the results of the previous trial as a starting point and cycled through 3 iterations.

In Figure 4, the resulting position error statistic for 3 convergence cycles is plotted, versus the pitch angle of the target array with respect to the camera, for two

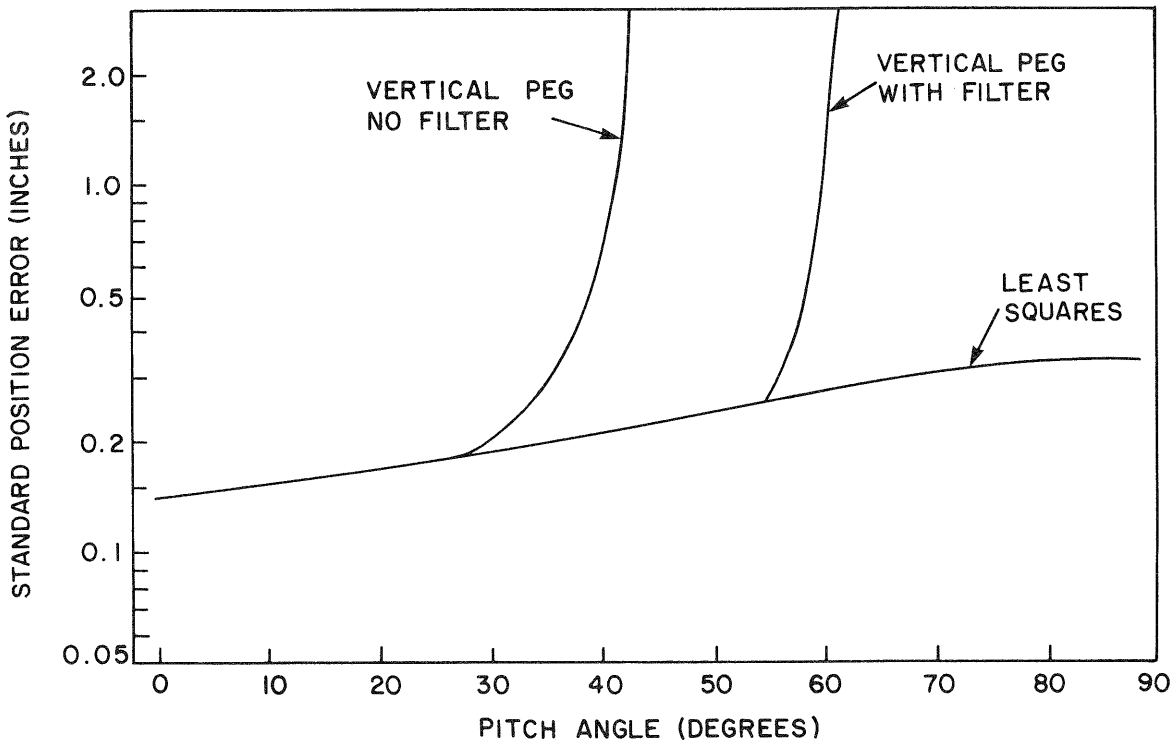

FIG. 4: PERFORMANCE OF VARIOUS ALGORITHMS FOR A 3-ELEMENT PEG TARGET ARRAY

versions of the vertical peg algorithm and for a general least squares algorithm. Referring to Figure 4, the standard (i.e. "no filter") vertical peg algorithm has a nearly identical performance to that of the least squares approach out to a pitch angle, θ, of roughly 30° at which point the error begins to climb due to a failure of the algorithm to converge in 3 cycles. The introduction of a recursive iteration filter can generally improve convergence properties and extend the operating range of the vertical peg algorithm out to a pitch of approximately 55° as illustrated by the "with filter" curve in Figure 4.

FOUR ELEMENT ARRAY ALGORITHMS

A second family of algorithms which we will now consider is that of an array of four target elements confined to a plane. For this situation we will examine two versions of the algorithm, each of which has a different nominal operating point. In order to appreciate the need for multiple operating points, consider Figure 5 in which the image of a four target element square array is sketched in both a non-pitched and a strongly pitched (60°) configuration. Referring to the square array images of Figure 5, the shortening of the vertical sides due to a pitch motion is a very dramatic effect at large pitch angles which is evident regardless of the range of the object (and target) from the camera but is a very weak effect at small pitch angles (since the image length of a side is roughly proportional to the cosine of the pitch angle). In contrast to this effect, the change in the ratio of the top length to bottom length is strongest at very small pitch angles but is only very significant when the target array occupies a large angular field of view of the camera.

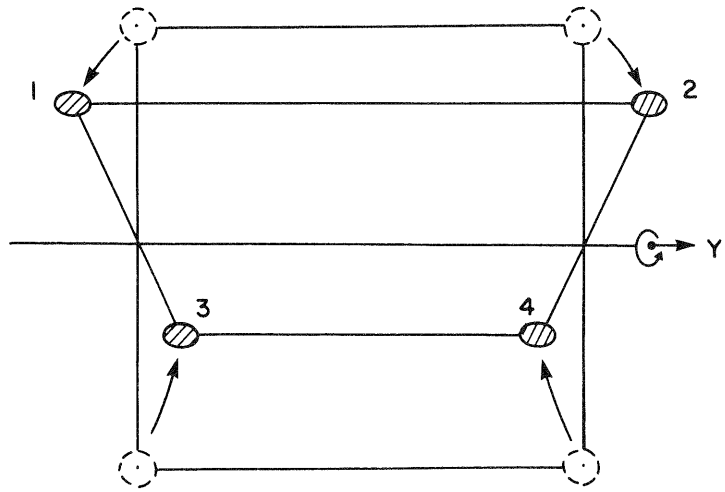

FIG. 5: EFFECT OF 60° OF PITCH ON IMAGE OF A SQUARE TARGET ARRAY

Our "Quadrilateral" algorithm determines both yaw and pitch angles from the effect that a side of a quadrilateral which is pitched or yawed towards the camera increases in length while one pitched or yawed away from the camera decreases in length. As a result this algorithm works best for small pitch and yaw angles and for close range photogrammetry. Our "Pitched Rectangle" algorithm, which handles rectangular target arrays that are strongly pitched with respect to the camera, is based on the general shortening of sides perpendicular to the pitch axis and is suitable for both close in and longer range photogrammetry but cannot be used in the neighbourhood of 0° of pitch.

In Figures 6 and 7, the performances of these two 4-element-in-a-plane algorithms are compared, for both a 86 cm and a 56 cm square target array respectively, with that of a least squares approach over a range of target pitch angle (where the target range, camera parameters and simulation conditions are the same as those used for Figure 4). Referring to Figures 6 and 7, the Quadrilateral algorithm can be seen to have close to

least squares performance in the neighbourhood of 0° of pitch (where this neighbourhood shrinks as the size of the target array gets smaller) while the pitched rectangle algorithm (having a nominal operating point of 45°) is seen to have a large operating range except near 0° and 90° of pitch.

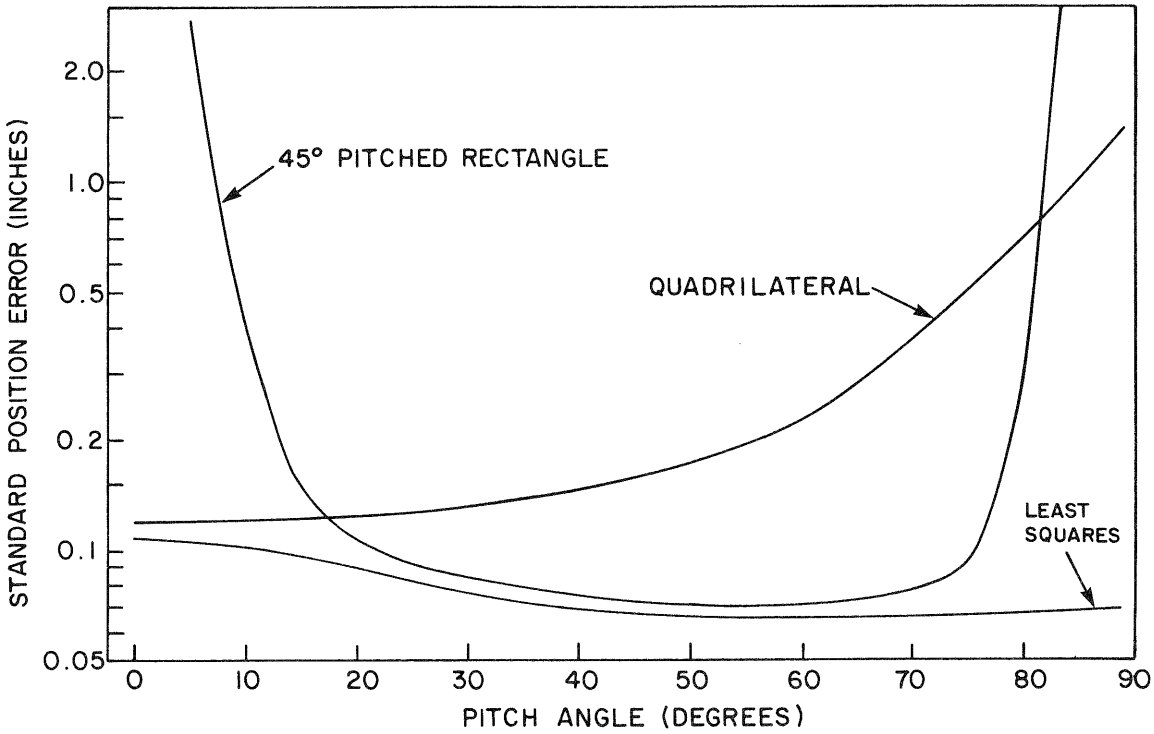

FIG. 6: PERFORMANCE OF VARIOUS ALGORITHMS FOR A 86 cm SQUARE TARGET ARRAY

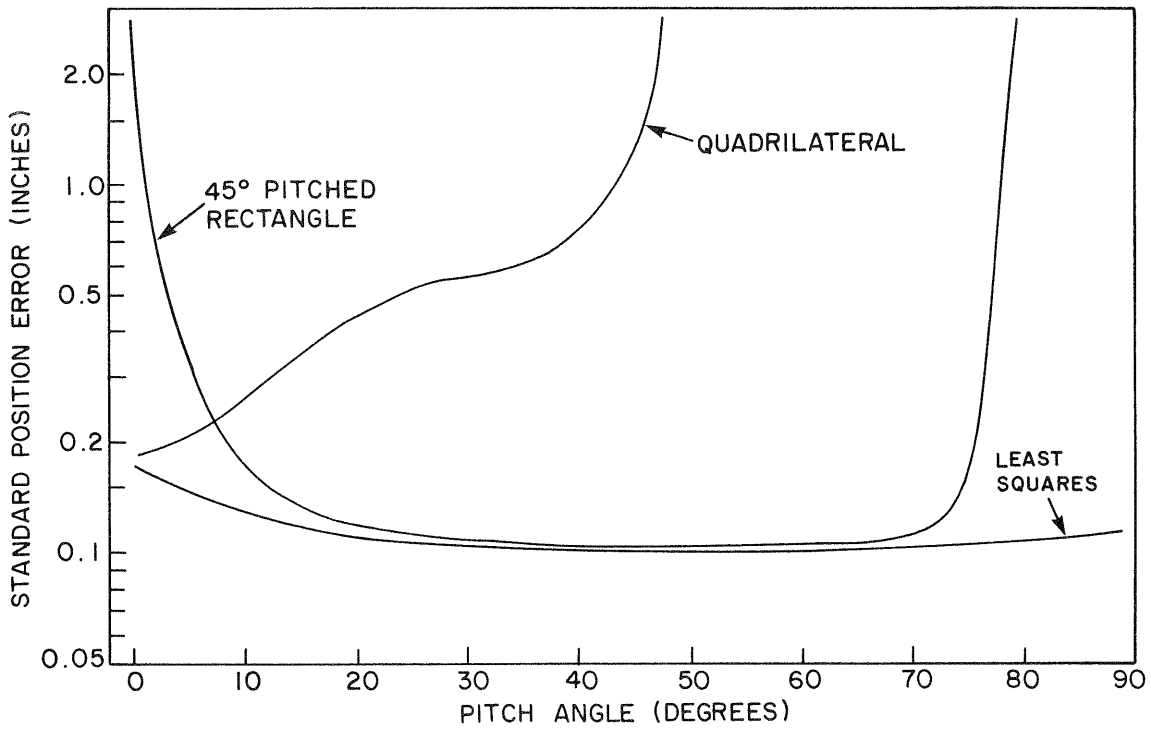

FIG. 7: PERFORMANCE OF VARIOUS ALGORITHMS FOR A 56 cm SQUARE TARGET ARRAY

CLOSING REMARKS

We have introduced an approach to developing iterative single camera photogrammetry algorithms which have low execution overheads. Within their nominal operating ranges these algorithms provide a level of accuracy close to that of least squares approaches.

ACKNOWLEDGEMENTS

The author would like to thank H.F.L. Pinkney, G.L. Basso and C.I. Perratt of the National Research Council of Canada and S.N. Prasad of SPAR Aerospace Ltd. for valuable discussion regarding the design and evaluation of photogrammetry algorithms.

REFERENCES

Fischler, M.A. and Bolles, R.C., 1981. "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography," Communcations of the ACM, Vol. 24, No. 6.

Hung, Y., Yeh, P. and Harwood, D., 1985. "Passive Ranging to Known Planar Point Sets," Proc. of IEEE Int. Conf. on Robotics and Automation, March 1985, St. Louis.

Kratky, V., 1979. "Real-Time Photogrammetric Support of Dynamic Three-Dimensional Control," Photogrammetric Eng. and Remote Sensing, Vol. 45, No. 9.

Pinkney, H.F.L., 1978. "Theory and Development of an On-Line 30 Hz Video Photogrammetry System for Real-Time 3-Dimensional Control," Proc. of ISP Symposium on Photogrammetry for Industry, Aug. 1978, Stockholm.

Pinkney, H.F.L. and Perratt, C.I., 1986. "A Flexible Machine Vision Guidance System for 3-Dimensional Control Tasks," Proc. of ISPRS Commission V Symposium, June 1986, Ottawa.

Tietz, J.C. and Germann, L.M., 1982. "Autonomous Rendezvous and Docking," Proc. of American Control Conference, June 1982, Arlington.