DETERMINATION OF MOTION PARAMETERS OF A MOVING OBJECT FROM MOVING CAMERA DATA

Sunil Fotedar
Tracking and Laser Systems Section,
Lockheed Engineering and Sciences Company (LESC),
Houston, Texas 77058

Rui J. P. deFigueiredo
Electrical and Computer Engineering Department,
Rice University,
Houston, Texas 77251-1892

Kumar Krishen
Mission Support Directorate,
Johnson Space Center, NASA,
Houston, Texas 77058

ABSTRACT

The determination of 3-D motion parameters of an object from its image-sequences is discussed for three types of motion analysis: (1) *monocular vision*, (2) *stereo vision*, and (3) *stereo motion*. These parameters enable one to obtain attitude, attitude rate, surface shape, identification/recognition, and track of the object. Under suitable conditions, these parameters can be estimated from 2-D image coordinates of a set of points on the object's surface in consecutive images, using the *Image Point Correspondence* (IPC) algorithm. In this paper, a *Generalized Image Point Correspondence* (GIPC) algorithm has been developed to enable the computation of motion parameters for a general configuration where both the object and the camera are moving. Furthermore, this algorithm was successfully tested on both simulated and video-acquired data.

1 INTRODUCTION

Research in motion-analysis has evolved over the years as a challenging field in the area of robotic vision. Its major contribution is in application to dynamic image-sequence analysis. The information contained in the image-sequences of a moving object aids in the computation of the motion parameters as well as the segmentation and shape analysis.

1.1 APPLICATIONS OF MOTION ANALYSIS

The potential applications of image-sequence analysis (Fig. 1) are listed below [1,2]:

(i) Robotics/Automation
(ii) Medicine and Biological Sciences
(iii) Military Applications
(iv) Meteorological Applications
(v) Traffic Monitoring
(vi) Segmentation and Scene Analysis

Robotics and automation find applications in industry and space. In industry, a camera mounted on the end-effector of a manipulator (robot arm) takes images of a moving object or target. The range and the orientation of the object relative to the end-effector are estimated from a knowledge of the displacements of the images of its features. This information, in turn, is used as a feedback to control the manipulator. For space applications, a robot's task can be, for example, to retrieve a defective satellite. After knowing the position, orientation, and the velocities in all six degrees of freedom of the satellite, the motion of the manipulator under computer control in matched to that of the satellite. It is done so that when the robot grasps the satellite, no excessive forces and torques are produced, which might otherwise damage the satellite or the manipulator. A future use of the robot is in its application to an autonomous vehicle. Such a robot-controlled vehicle could be extensively used for lunar/space exploration, using vision for its navigation. One of the functions of the vision system is to prevent collision with obstacles, such as rocks, while moving and taking pictures of different types of terrains

* With Tracking and Laser Systems Section, Lockheed Engineering and Sciences Company (LESC), Houston, Texas 77058
+ With Electrical and Computer Engineering Department, Rice University, Houston, Texas 77251-1892
# With Mission Support Directorate, Johnson Space Center, NASA, Houston, Texas 77058

in planetary exploration.

1.2 CURRENT RESEARCH THRUSTS IN MOTION ANALYSIS

The current research in motion analysis for perception is concentrated in the following areas (Fig. 1):

(i) Feature extraction and matching of consecutive images.
(ii) Efficient algorithms for motion parameter detection.
(iii) Hardware Implementations.
(iv) Applications developments.

Techniques are being developed that enable extraction and matching of features over consecutive frames. Features can also be estimated from images for identification purposes. Furthermore, several images can be matched using a set of features corresponding to a particular object or scene. Another approach to the identification of the object/target is the correlation of the acquired scenes with those stored in the computer. Feature extraction and image matching provide the needed inputs to the algorithms for motion parameter estimation. These algorithms perform unique operations on image sequences. Considerable effort is presently being expended in the area of parameter estimation for applications in relative motion between the camera and the object/scene. Applications of motion parameter estimation in various fields of engineering and science continue to be explored at an accelerated pace. The thrust of automation and robotics in industrial, space, and defense operations has created a need for the motion parameter estimation algorithms. Research efforts in the areas of algorithm development and applications are determining requirements for hardware development. Hardware subsystems are comprised of cameras, videos, pointing, tracking, and on-board data processing. The technology innovations for space applications include reduction of size and weight, increase in speed and reliability, and automatic operation. For the medical and ground-based systems, technology is being developed for data processing/recording and display subsystems. In the past two decades, a number of approaches have evolved in the development of efficient algorithms for motion perception from a sequence of images. These approaches are discussed in the following subsection.

1.3 APPROACHES TO THE SOLUTION FOR THE MOTION ANALYSIS PROBLEM

In order to develop efficient algorithms, one can use either of the two broadly classified approaches [1]:

1.3.1 INTENSITY-BASED APPROACH

Here, image processing techniques like subtraction, correlation, convolution, Fourier analysis, or differentiation, are applied to the object's image to estimate its motion parameters. The algorithms that fall in this category are explained below [3]:

(i) Reflectance Map Method: The dependence of surface reflection on the geometry of incident and reflected rays is given by the bidirectional reflectance distribution function. The reflectance that gives a relationship between the surface orientation and brightness can be derived from this function and the distribution of the light sources. The photometric stereo method for recovering the orientation of surface patches from a number of images taken under different lighting conditions has been developed. If a single image is available, the shape can also be recovered from the spatial variation of brightness called shading, since parts of the surface are oriented differently and are visible with different brightnesses.

(ii) Optical Flow Approach: *Optical flow* is a velocity field that defines motion in an image. A velocity vector is assigned to each point in the image. Brightness patterns in the image move as the object moves. It is the apparent motion of these brightness patterns that gives the optical flow. Movement through the environment maps information onto a pattern. From this pattern through inverse mapping, it is possible to derive information about the environment and the observer's motion. This approach requires iterative searches. The orthographic (parallel) projection is also assumed.

1.3.2 FEATURE-BASED APPROACH

In this approach, prominent features are found and then matched over consecutive time-varying frames. The features can be points, line segments, edge fragments, or moment invariants. The correspondence or the matching problem over a set of frames is assumed to be known *a priori.*

The algorithms that fall in this category are explained below:

(i) Moment Invariant/Attributed Graph Approach: 3-D objects are recognized, and the motion parameters are determined from their 2-D orthographic projections. Geometric transforms are used instead of the iterative matching techniques [4,5]. For identification purposes, a 3-D object is represented by an attributed graph where a node represents a face of the object. Associated with each node is a feature vector containing moment invariants of the face. A link between two nodes means that the two faces are adjoining. Associated with each edge is a scalar, which is the angle between two nodes. Any 2-D projection of the object can be similarly represented by a graph, which is a subgraph of the above graph. This is so because a part of the object is facing the camera. Thus the identification problem becomes a subgraph isomorphism between the observed image and the 3-D object. The moment invariants are found to stay unchanged under 3-D motion. Attitude parameters of the observed object relative to the 3-D object are determined from the knowledge of 2-D moments of its faces. Orthographic projection is assumed in this analysis. This approach considers the object with flat surfaces only.

(ii) Straight Line Correspondence Algorithm: 3-D motion/structure of a rigid body, containing straight line segments as features, can be determined if a sequence of three 2-D perspective views is given [6]. The projections of 3-D lines over the three consecutive image frames are assumed to be known. A seven line correspondence (LC) involves an iterative search without any constraints on the 3-D line. If the 3-D lines lie on parallel planes, and the orientation of the rotation axis is fixed over the three image frames, an eight LC results in a linear method. The surface of a unit sphere is used in place of the plane of the perspective projection. In this analysis, the projections of moving 3-D lines on this sphere over three frames are studied. A fairly accurate initial guess is required for convergence in the iterative search.

(iii) Image Point Correspondence Algorithm: Three different cases of motion analysis have been identified [7]. They are: (i) Two-view motion analysis (monocular vision), (ii) stereo vision, and (iii) stereo motion. A discussion of these cases appears in the following section where the approach used in the paper is described.

1.4 STATEMENT OF THE PROBLEM

In this paper, three cases of motion analysis based on vision, have been investigated. These cases are:

(1) The Two-view Motion Analysis or Monocular Vision Case (Fig. 2): Pictures of a moving target are taken by a stationary camera at different instants of time. Using the IPC algorithm, the motion parameters of the target are found from 2-D image coordinates of the target's features (e.g. surface points). The knowledge of the correspondence of these 2-D coordinates over consecutive frames is required. The surface of the target is also determined.

(2) The Stereo Vision or Binocular Vision Case (Fig. 3): Pictures of a stationary target are taken by two cameras stationed at different locations. The motion parameters that relate the two camera coordinate systems are found by using the IPC algorithm. The surface of the target can also be determined from a minimum of eight available data points.

(3) The Stereo Motion Case: This case is similar to the second case. However, instead of two stationary cameras, one moving camera is used to take the pictures of the stationary target from two different locations at different instants of time.

In all the cases mentioned, the image plane is assumed to be at the focal point of the camera with its X- and Y-axes parallel to those of the camera coordinate system, where z-axis is the line of sight.

In this context, the previous work done in developing the IPC algorithm in three different ways has been studied [7,8,9]. However, the IPC algorithm does not apply to the more general problem of motion analysis. The generalized version of motion analysis involves a situation where both the object and the camera are moving [10]. For example, industrial and space robots face this situation in locating and tracking of various objects/scenes. The space robot takes pictures of a satellite for motion analysis. Both the camera/video system and the object move asynchronously. An algorithm for motion parameter estimation is required for this general case of relative motion. In the GIPC algorithm, a generalization of the IPC algorithm, this problem of motion analysis is discussed. The three cases of motion analysis mentioned above are special cases of the present one.

2 THE GENERALIZED IMAGE POINT CORRESPONDENCE ALGORITHM

2.1 THE ALGORITHM

The general case of motion analysis is illustrated in Fig. 4. For simplicity in presentation, we consider in detail the equations that track a single point P on a moving object by a moving camera. F_i and F_j are the two frames with which the camera coordinate system coincides at two different instants of time τ_i and τ_j (where $\tau_j > \tau_i$) respectively. Point P moves from one position P_i to another position P_j due to rigid-body motion of the object. We assume (R_i, $\mathbf{T}_i$) and (R_j, $\mathbf{T}_j$) to be the transformation parameters (rotation and translation) that link the frames F_i and F_j respectively with the standard frame S. Also, let (R_{ij}, $\mathbf{T}_{ij}$) be the transformation parameters that link the frame F_i with the frame F_j. The object moves with the unknown motion parameters (R, $\mathbf{T}$).

The desired relationship between the coordinates of the initial and the final positions of point P recorded by the camera, with respect to the frames F_i and F_j respectively, is given by the following equations (see Appendix):

$$\mathbf{p}_{jj} = R_{ij}' \ \mathbf{p}_{ii} + \mathbf{T}_{ij}' \qquad (1a)$$

where

$\mathbf{p}_i = (x_i, y_i, z_i)^T$ is the vector of 3-D coordinates of P_i relative to S

$\mathbf{p}_j = (x_j, y_j, z_j)^T$ is the vector of 3-D coordinates of P_j relative to S

$\mathbf{p}_{\alpha\beta} = (x_{\alpha\beta}, y_{\alpha\beta}, z_{\alpha\beta})^T$ is the vector of 3-D coordinates of P_β relative to F_α at instant τ_β (α, β = i,j)

and

$$R_{ij}' = R_{ij}^T \ R_i \ R \ R_i^T = R_j \ R \ R_i^T \quad \text{and} \quad \mathbf{T}_{ij}' = - R_j \ R \ R_i^T \ \mathbf{T}_i + R_j \ \mathbf{T} + \mathbf{T}_j \qquad (1b,c)$$

Eq. (1a) gives the expression for the *generalized version of the motion-analysis equation*. Clearly, it does not matter whether the object or the camera is moved first. For the rest of the discussion, R and $\mathbf{T}$ are defined as:

$$R \equiv \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad \text{and} \quad \mathbf{T} \equiv \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} \qquad (1d,e)$$

where $r_{\alpha\beta}$ (α, β = 1,2,3) are the rotational elements and t_α (α = 1,2,3) the translations along x-, y-, and z-axes respectively.

Special Cases: If F_i coincides with S, the motion equation can be written as

$$R_{ij}' = R_{ij}^T \ R = R_j \ R \quad \text{and} \quad \mathbf{T}_{ij}' = R_j \mathbf{T} + \mathbf{T}_j \qquad (2a,b)$$

The IPC algorithm can be used to estimate the motion parameters (R_{ij}',$\mathbf{T}_{ij}'$) and hence (R, $\mathbf{T}$) of the moving object, assuming R_{ij} and $\mathbf{T}_{ij}$ are known.

(i) Monocular Vision: Eq. (1a) reduces to the case of the two-view motion-equation when the location of the camera, taking the pictures of the moving object, is fixed (Fig. 2). In that case,

$$R_i = R_j = R_{ij} = I \quad \text{and} \quad \mathbf{T}_{ij} = \mathbf{T}_j = \mathbf{O} \qquad (3a)$$

Therefore, the generalized motion-equation reduces to

$$\mathbf{p}_{jj} = R \ \mathbf{p}_{ii} + \mathbf{T} \qquad (3b)$$

or to the more familiar two-view motion equation

$$\mathbf{p}' = R \ \mathbf{p} + \mathbf{T} \qquad (3c)$$

as the frames F_i and F_j coincide with the frame S, such that $\mathbf{p}_{ii} = \mathbf{p}_i = \mathbf{p}$ and $\mathbf{p}_{jj} = \mathbf{p}_j = \mathbf{p}'$.

(ii) Stereo Vision/Stereo Motion: For a stereo vision/stereo motion case, the object is assumed to be stationary (Fig. 3). In that case,

$$R = I; \ R_{ij} = R_j \quad \text{and} \quad \mathbf{T} = \mathbf{O} \qquad (4a)$$

and the generalized motion equation reduces to

$$\mathbf{p}' = R_j\,\mathbf{p} + \mathbf{T}_j \tag{4b}$$

as should be the case.

In another special case of stereo vision, the object is moving and the cameras are rigidly fixed to each other in a manner that their optical axes are parallel, as shown in Fig. 5. In this case, F_i coincides with S so that

$$R_{ij} = R_j = I;\ \mathbf{T}_{ij} = \mathbf{T}_j = (-d, 0, 0)^T \tag{4c}$$

The three cases of motion analysis have been found to be equivalent [7]. The motion of point P_i to point P_j with a fixed frame F_j is similar to the motion of frame F_j to frame F_i with respect to a fixed point P.

2.2 DETERMINATION OF 3-D MOTION PARAMETERS

The motion parameters to be determined [7,8,9,10,11] are:

2.2.1 OBJECT SURFACE SHAPE

The shape of the object surface is found by a map of relative depths of 3-D object surface points by the method discussed as follows:

From the generalized motion equation for any two consecutive frames F_i and F_{i+l} , one finds that:

$$X_{i+1} = \frac{x_{i+1}}{z_{i+1}} = \frac{\mathbf{r}_1 \cdot (\mathbf{p}_i - \mathbf{T}')}{\mathbf{r}_3 \cdot (\mathbf{p}_i - \mathbf{T}')} = \frac{\mathbf{r}_1 \cdot (\mathbf{v}_i - \mathbf{T}'/z_i)}{\mathbf{r}_3 \cdot (\mathbf{v}_i - \mathbf{T}'/z_i)} \tag{5a}$$

where $\mathbf{r}_\alpha$ (α = 1,2,3) is the α^{th} row of the rotation matrix $R_{i,i+1}$' and $\mathbf{v}_i = \mathbf{p}_i / z_i$. Therefore, the 'z' coordinates (*depth* or *range*) of the features, from their 2-D image points, can be found from the following equation:

$$z_{\lambda,i} = \frac{(\mathbf{r}_1 - X_{i+1}\mathbf{r}_3) \cdot \mathbf{T}'}{(\mathbf{r}_1 - X_{i+1}\mathbf{r}_3) \cdot \mathbf{v}_i} = \lambda z_i \tag{5b}$$

Therefore,

$$x_{\lambda,i} = z_{\lambda,i} X_i = \lambda x_i \quad \text{and} \quad y_{\lambda,i} = z_{\lambda,i} Y_i = \lambda y_i \tag{5c}$$

Eqs. (5b, 5c) give the object surface shape at instant τ_i with respect to the frame F_i. The shape of the object surface is the set of 3-D coordinates of surface points relative to a frame. Therefore, the shape of an object can be determined before it moves. However, its size cannot be determined due to the scale factor λ involved. In order to determine the shape of the object surface after the object has moved with respect to the frame F_{i+1}, one has to compute 'X_{i+l}' and 'Y_{i+l}' from

$$x_{\lambda,i+1} = z_{\lambda,i+1} X_{i+1} \quad \text{and} \quad y_{\lambda,i+1} = z_{\lambda,i+1} Y_{i+1} \tag{5d}$$

For finding the shape of an object surface one needs to know the correct solution for the rotation matrix R. If the sign of the coordinates 'z' and 'z' ' of any point is the same, the corresponding rotation matrix R is considered. When the signs are opposite, the other rotation matrix R' is considered as the final result [8,9].

2.2.2 RANGE AND INTERFRAME RANGE RATE

The *range* of available features, which determines the depth of object surface up to the scale factor λ at instant τ_i, is defined as:

$$\Re_i = z_{\lambda,i} \tag{6a}$$

since the object is conventionally viewed along the z-axis of the camera coordinate system. The *interframe range rate* (up to the scale factor λ) between any two consecutive frames F_i and F_{i+l}

at instants τ_i and τ_{i+1}, is given by:

$$\Re_{i,i+1} = \frac{\Re_{i+1} - \Re_i}{\Delta\tau} = \frac{z_{\lambda,i+1} - z_{\lambda,i}}{\Delta\tau} \qquad (\Delta\tau = \tau_{i+1} - \tau_i) \qquad (6b)$$

2.2.3 INTERFRAME ATTITUDE AND ATTITUDE RATES

The rotation matrix R can be estimated between the two consecutive frames F_i and F_{i+1} from the following equations:

$$R = R_i^T \; R_{i,i+1} \; R_{i,i+1}' \; R_i = R_{i+1}^T \; R_{i,i+1}' \; R_i \qquad (7a)$$

assuming the relative transformation parameters between these frames are known. Once the rotation matrix R is estimated, the attitude parameters of the object are computed [7,8,9]. Two different representations of R have been used [8,10,11].

The *interframe attitude rates* between the two frames are given as:

angular velocity of rotation, $\omega_\theta = (\theta_{i+1} - \theta_i)/\Delta\tau$ (7b)

if the first representation of R is used [8,10]. If the second representation of R is used [10,11], then

roll rate, $\omega_\theta = (\theta_{i+1} - \theta_i)/\Delta\tau$ (7c)

yaw rate, $\omega_\phi = (\phi_{i+1} - \phi_i)/\Delta\tau$ (7d)

pitch rate, $\omega_\psi = (\psi_{i+1} - \psi_i)/\Delta\tau$ (7e)

2.2.4 OBJECT TRACKING AND RECOGNITION

The process of tracking can be divided into three tasks: *object/target acquisition*, *attitude determination*, and *prediction* [2]. In the acquisition task, the object is located in the scene and its attitude and attitude rate parameters are determined to aid in the process of recognition at a later time. This task is divided into three major portions: feature tracking, the stereo solution, and matching to the object model. Features are detected and tracked over several frames, corresponding to different instants of time. These features are matched between consecutive frames, and the motion parameters are estimated by the stereo solution using the GIPC algorithm. These motion parameters are used in the prediction stage. In this stage, the attitude and attitude rates are predicted so as to get an idea about the future locations of the object. This allows one to track the object successfully in near real-time, limited by processing time of the computer used for the analysis.

The process of recognition is carried out after the object surface is reconstructed, and hence identified, from a knowledge of 3-D object space coordinates of the features that reside on the surface of the object. Since we find 3-D coordinates up to the scale factor λ, the object can be matched and recognized with a reference object in the main computer memory.

3 EXPERIMENTAL RESULTS

In this section, we present experimental results for the IPC and the GIPC algorithm tested successfully on real data. Separate experiments, corresponding to 2-D images of the two positions of an *octbox* (Fig. 6(a)), were conducted. The octbox has two parallel octagonal faces opposite to each other and eight rectangular faces. In the first experiment (Fig. 6(b)), the octbox was rotated around the x-axis by 15^o. In the second experiment (Fig. 6(c)), the octbox was rotated around the x-axis by 105^o. In both experiment, rotation around the x-axis means that the angle of rotation, by definition, is *roll*. Therefore, direction cosines of the arbitrary axis, around which the octbox rotates, are given by $\upsilon_1 = 1.0$; $\upsilon_2 = 0.0$; $\upsilon_3 = 0.0$. The translation along the three axes is 1 unit each. Data for these two cases of octbox rotations are shown in Fig. 7(a) and Fig. 7(b) respectively, where 3-D coordinates of the vertices of the octbox before the motion and their 2-D coordinates after the motion are given.

The GIPC algorithm has been applied to the first and second experiments, and results are shown in Figs. 8(a) and 8(b). In both the experiments, the camera was rotated through 10^o around its x-axis. With the same set of data for these experiments with the octbox rotation (Figs. 6(b), 6(c)), the directional cosines are found to be same. The angles of rotation are -5^o and 95^o respectively, which means that the angles of rotation of the oct-box are subtracted by the amount of rotation of the camera, and that indeed should be the case. These two experiments show the success of GIPC algorithm with real data.

4 CONCLUSIONS

In this paper, an extension to the IPC algorithm called the Generalized Image Point Correspondence (GIPC) was developed. A generalized expression for the motion analysis equation was derived, and three cases of motion analysis were found to be the special cases of this case. Experimental verification for this development was also provided.

5 ACKNOWLEDGMENTS

This research was supported by the NASA Contract NAS 9-17145, and NASA/JSC (RICIS) Grant NCC 9-16. Thanks are due to Mr. Olin Graham, Mr. D. Rhoades, Mr. K. Land, and Mr. K. Vorhaben (Supervisor). Special thanks to Mrs. L. Fotedar for technical assistance.

6 BIBLIOGRAPHY

[1] H. Shariat," The Motion Problem: How to use more than Two Frames," Institute for Robotics and Intelligent Systems, School of Engineering, University of Southern California, L. A., California 90089-0273, Oct. 6, 1986. *Ph. D. Dissertation*.

[2] D. B. Gennery," Stereo Vision for the Acquisition and Tracking of Moving Three-dimensional Objects," *Techniques for 3-D Machine Perception*, Elsevier Science Publishers B.V. (North-Holland), 1986, pp. 53-74.

[3] B. K. P. Horn, *Robot Vision*, MIT Press, McGraw Hill, 1986.

[4] B. B. Bamieh and R. J. P. deFigueiredo," A General Moment Invariants/Attitude-Graph Method for Three-Dimensional Object Recognition from a Single Image," *IEEE Journal of Robotics and Automation*, vol. RA-2, No. 1, Mar. 1986, pp. 31-41.

[5] R. J. P. deFigueiredo, B. A. Bamieh, S. Fotedar, E. Hack, K. Trahan and C. Wu," Determination of a Methodology for Machine Recognition and Attitude Determination of a 3-D Object from a Single TV Picture Frame," *Technical Report EE8520*, Dept. of Electrical and Computer Engineering, Rice University, Houston, TX 77251-1892, Dec. 12, 1985.

[6] B. L. Yen and T. S. Huang," Determining 3-D Motion/Structure of a Rigid Body over 2 Frames using Correspondences of Straight Lines lying on Parallel Planes," *Proc. ICASSP*, vol. 2, Mar. 19-21, 1984, San Diego, California, pp. 23.5.1-23.5.4.

[7] X. Zhuang and R. M. Haralick," Two-View Motion Analysis, Stereo Vision and a Moving Camera's Positioning: Their Equivalence and a New Solution Procedure," *IEEE Int. Conf. on Robotics and Automation*, Mar. 25-28, 1985, pp. 6-10.

[8] R. Y. Tsai and T. S. Huang," Uniqueness and Estimation of Three-Dimensional Motion Parameters of Rigid Objects with Curved Surfaces," *IEEE Trans. Pattern Anal. & Machine Intel.*, vol. PAMI-6, No. 1, Jan. 1984, pp. 13-27.

[9] H. C. Longuet-Higgins," A Computer Algorithm for Reconstructing a Scene from Two Projections," *Nature*, vol. 293, pp. 133-135, Sept. 1981.

[10] S. Fotedar, and R. J. P. deFigueiredo," Determination of Motion Parameters of a Moving Object from Moving Camera Data," *Technical Report EE8708*, Dept. of Electrical and Computer Engineering, Rice University, Houston, TX 77251-1892, Apr. 29, 1987.

[11] S. Ganapathy," Decomposition of Transformation Matrices for Robot Vision," *Proc. Int. Conf. on Robotics*, Mar. 13-15, 1984.

APPENDIX

To carry out a detailed analysis of the problem sketched in Fig. 4, one needs to establish the relationship between the three sets of frames in terms of the coordinates of point P. For this purpose, the problem is split into two steps - one corresponding to the *initial* position P_i and the other to the *final* position P_j. These two steps are considered separately.

A.1.1 INITIAL POSITION OF THE OBJECT

The transformation from the frame F_i to the frame F_j [7] is:

$$\Phi_i \mathbf{p}_{ii} = \Phi_i R_{ij} \mathbf{p}_{ji} + \mathbf{T}_{ij} \qquad \text{(A.1)}$$

where Φ_i is the matrix of the basis vectors of F_i .

The transformation from F_i to the standard frame S, and F_j to S are special cases of transformation of Eq. (A.1) and are expressed respectively as:

$$\mathbf{p}_{ii} = R_i \mathbf{p}_i + \mathbf{T}_i \text{ and } \mathbf{p}_{ji} = R_j \mathbf{p}_i + \mathbf{T}_j \qquad \text{(A.2a,b)}$$

A.1.2 FINAL POSITION OF THE OBJECT

The motion of the object relates $\mathbf{p}_i$ and $\mathbf{p}_j$ as follows:

$$\mathbf{p}_j = R \mathbf{p}_i + \mathbf{T} \qquad \text{(A.3a)}$$

As in Eq. (A.1), the transformation from F_i to F_j is expressed as:

$$\Phi_i \mathbf{p}_{ij} = \Phi_i R_{ij} \mathbf{p}_{jj} + \mathbf{T}_{ij} \qquad \text{(A.3b)}$$

The transformation from F_i to S, and F_j to S, are special cases of Eq. (A.3b) and are expressed as:

$$\mathbf{p}_{ij} = R_i \mathbf{p}_j + \mathbf{T}_i \quad \text{and} \quad \mathbf{p}_{jj} = R_j \mathbf{p}_j + \mathbf{T}_j \qquad \text{(A.3c,d)}$$

This results in two sets of Eqs. (A.1) and (A.3a, A.3b), corresponding to two positions of the object. From either sets of equations, we find the relationship between (R_{ij}, $\mathbf{T}_{ij}$), (R_i, $\mathbf{T}_i$) and (R_j, $\mathbf{T}_j$) as:

$$R_{ij} = R_i R_j^T \text{ and } \quad \mathbf{T}_{ij} = \Phi_i (\mathbf{T}_i - R_i R_j^T \mathbf{T}_j) \qquad \text{(A.3e,f)}$$

The pictures of the moving object are taken by a moving camera at discrete instants of time. From this general motion problem, we have two different cases. In the first case, the initial picture of the moving object is taken by the camera at some location. Then the camera is moved to another location so that the object is in the field of view when the second picture is taken. In the second case, the camera can be moved first to another location in order to take the second picture of the object. These two cases are found to be identical in terms of the motion equation.

A.2.1 FIRST CASE

The motion of the object preceding the motion of the camera is assumed in this case. The relationship between the initial and final positions of the point P is expressed by the following set of motion equations:

$$\mathbf{p}_j = R \mathbf{p}_i + \mathbf{T} \qquad \text{(A.4a)}$$

for the motion of the object, and

$$\Phi_i \mathbf{p}_{ij} = \Phi_i R_{ij} \mathbf{p}_{jj} + \mathbf{T}_{ij} \qquad \text{(A.4b)}$$

for the motion of the camera from Eq. (A.3b). For transformation of F_i to S, let F_j coincide with S. In that case,

$R_{ij} = R_i$; $\mathbf{T}_{ij} = \mathbf{T}_i$; $\mathbf{p}_{ji} = \mathbf{p}_i$; and $R_j = I = \Phi_j^T$

Substituting the above in Eq. (A.1) gives Eq. (A.2a). Similarly, let F_i be S, where

$R_{ij} = R_j^T$;$\mathbf{T}_{ij} = - R_j^T \mathbf{T}_j$; $\mathbf{p}_{jj} = \mathbf{p}_i$; and $\Phi_i^T = I = R_i$

Substituting the above in Eq. (A.1) results in Eq. (A.2b). Eliminating $\mathbf{p}_i$ from Eq. (A.2a) and Eq. (A.2b), we have

$$\Phi_i \mathbf{p}_{ii} = \Phi_i R_i R_j^T \mathbf{p}_{ji} + \Phi_i (\mathbf{T}_i - R_i R_j^T \mathbf{T}_j) \qquad \text{(A.4c)}$$

Comparison of Eq. (A.4c) with Eq. (A.1) gives Eqs. (A.3e,A.3f).

The derivation of Eq. (1a) is as follows:

2-D image coordinates of $\mathbf{p}_{ij}$ and $\mathbf{p}_{ji}$ are not observed on the image image plane, because of the manner in which the pictures are taken. So, eliminating $\mathbf{p}_{ij}$ from Eqs. (A.3b,A.3c) yields

$$\Phi_i (R_i \mathbf{p}_j + \mathbf{T}_i) = \Phi_i R_{ij} \mathbf{p}_{jj} + \mathbf{T}_{ij} \qquad \text{(A.5a)}$$

Elimination of $\mathbf{p}_j$ from Eqs. (A.3a,A.5a) gives

$$\Phi_i (R_i R \mathbf{p}_i + R_i \mathbf{T} + \mathbf{T}_i) = \Phi_i R_{ij} \mathbf{p}_{jj} + \mathbf{T}_{ij} \qquad \text{(A.5b)}$$

Furthermore, elimination of $\mathbf{p}_i$ from Eqs. (A.2a,A.5b) yields the desired relationship between $\mathbf{p}_{jj}$ and $\mathbf{p}_{ii}$ as shown in Eq. (1a).

A.2.2 SECOND CASE

In this case, the motion of the camera preceding that of the object is assumed. The analysis follows:

The camera moves first in this case, so that

$$\Phi_i \mathbf{p}_{ii} = \Phi_i R_{ij} \mathbf{p}_{ji} + \mathbf{T}_{ij} \qquad \text{(A.6a)}$$

Then the object moves, so that

$$\mathbf{p}_j = R \mathbf{p}_i + \mathbf{T} \qquad \text{(A.6b)}$$

From Eqs. (A.6a,A.6b,A.2b,A.2c, and A.3d), the relationship between $\mathbf{p}_{jj}$ and $\mathbf{p}_{ii}$ can be found to be the one expressed in Eqs. (1a,1b,1c).

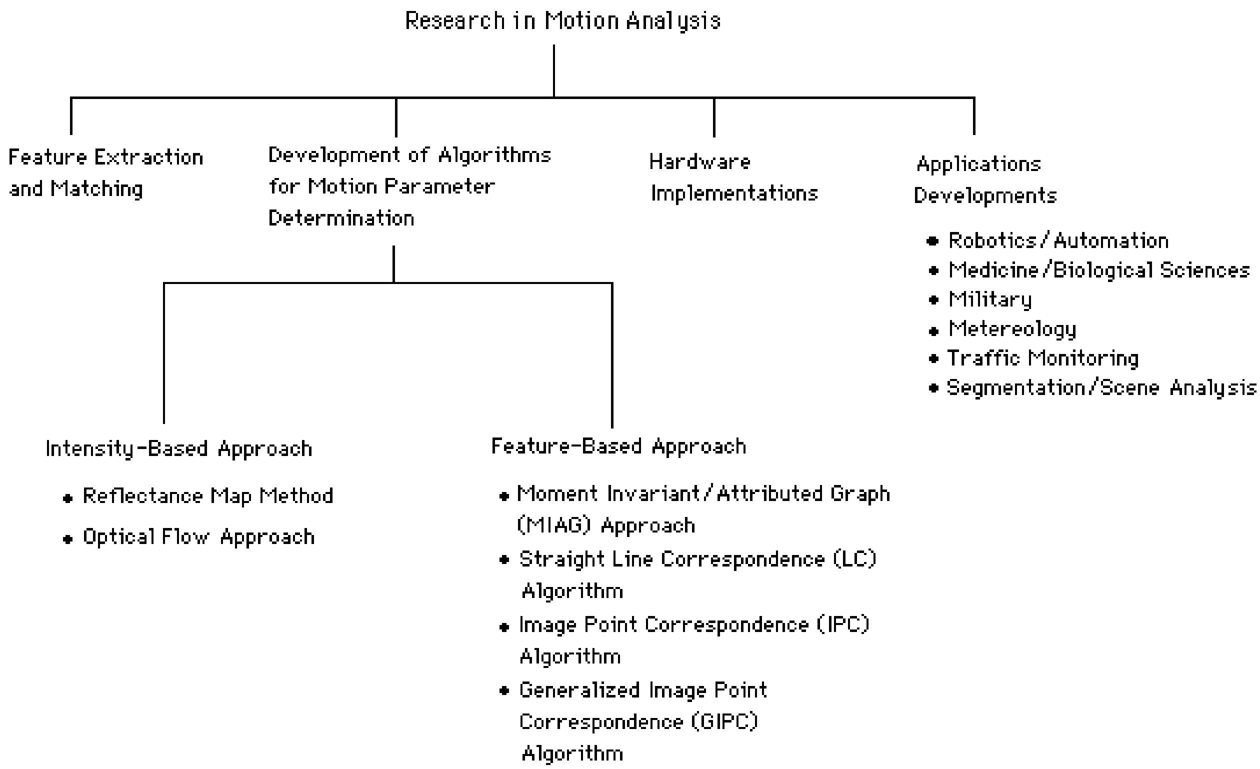

Fig. 1. Current research thrusts in motion analysis.

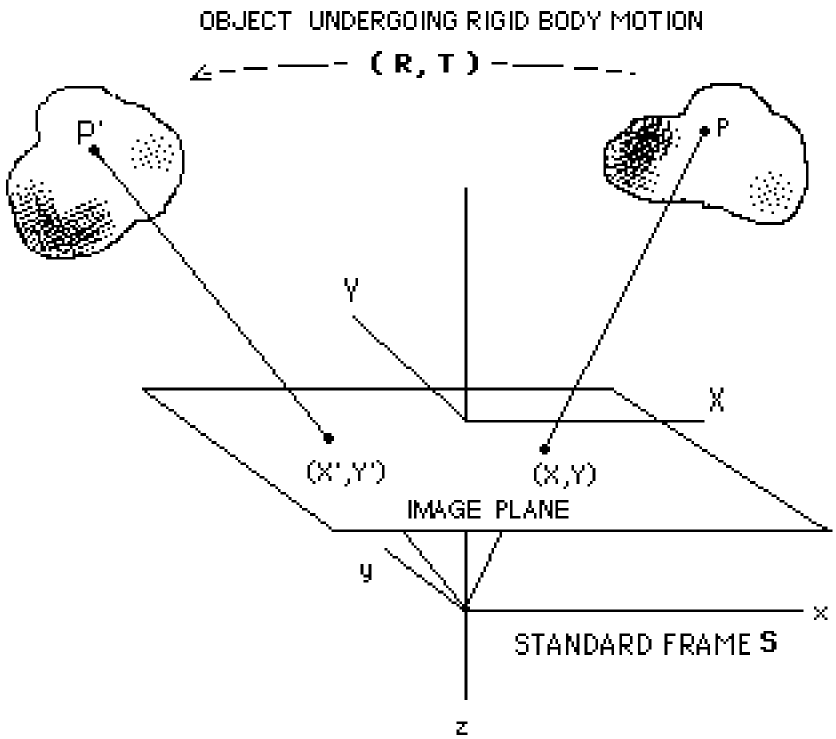

Fig. 2. Two-view motion analysis case.

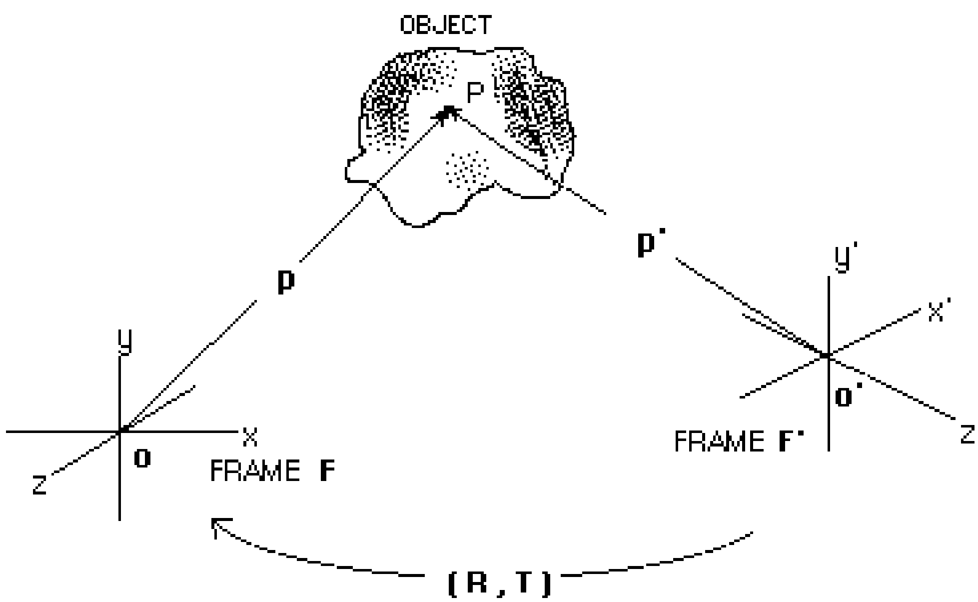

Fig. 3.   Stereo vision case.

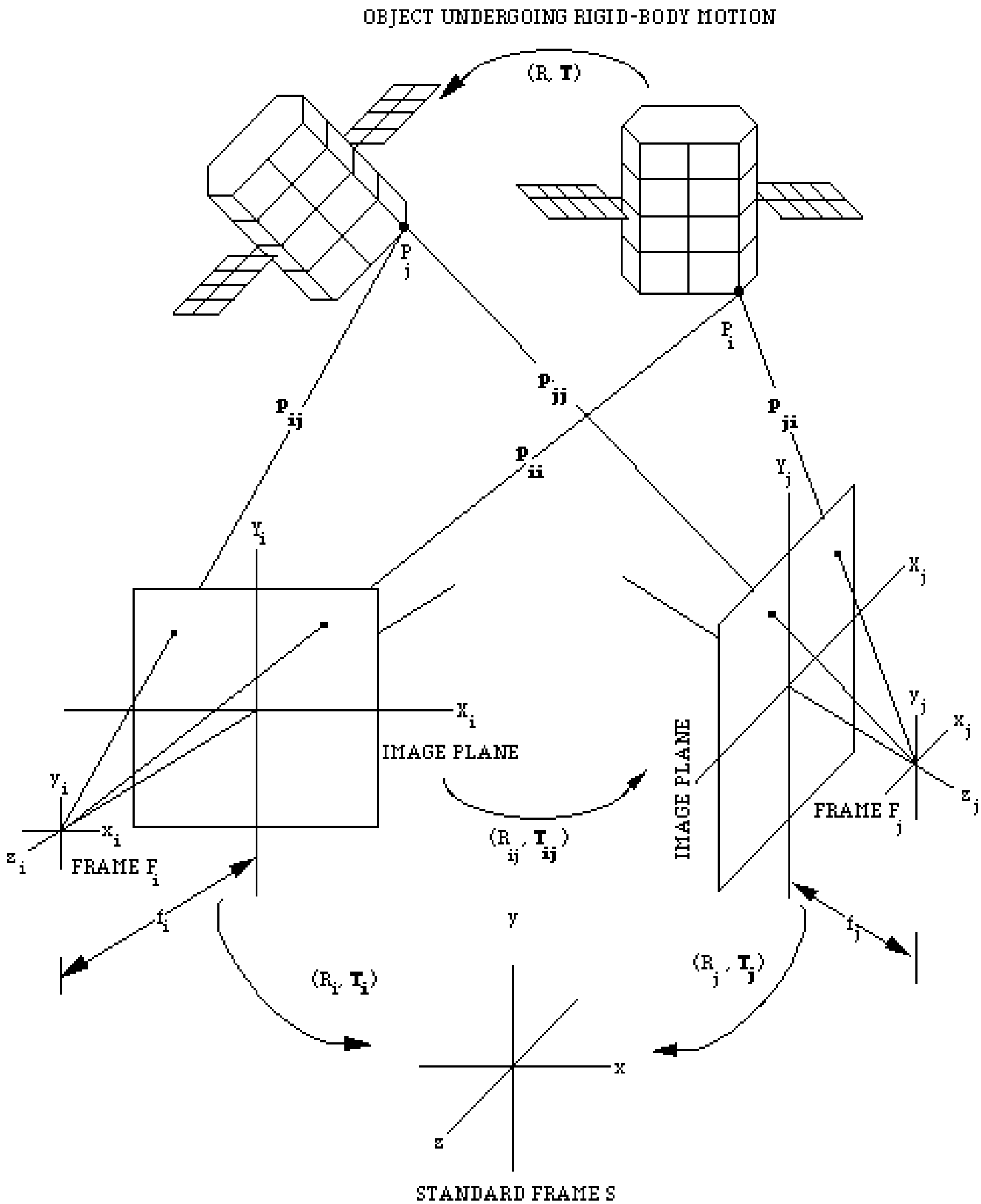

Fig. 4. Geometry illustrating GIPC Algorithm.

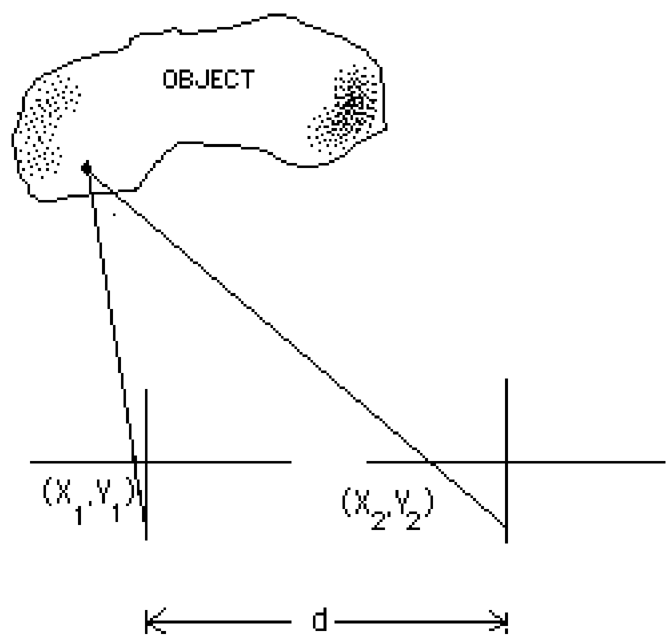

Fig. 5. Special case of stereo vision where the optical axes of two cameras are parallel.

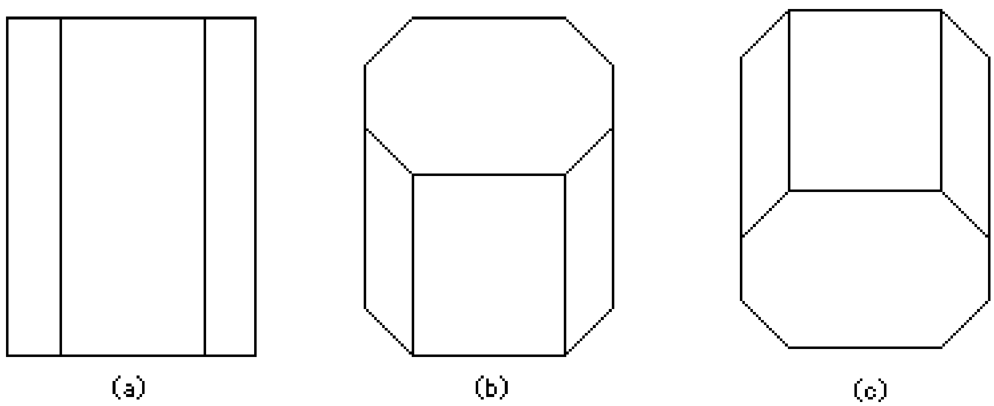

Fig. 6. Experiments with real-data.

(a) *Octbox* in its initial position; (b) first case of motion where *Octbox* is rotated through 15^{o} around x-axis; and (c) second case of motion where *Octbox* is rotated through 105^{o} around x-axis.

x	y	z	X'	Y'
0	-1	-1	3.414214	7.595754
-1.5	-1	-0.5	-3.058409	10.654163
-2	1	1	-0.585786	-0.131652
-1.5	1	-0.5	-0.238625	0.584223
-1.5	1	2.5	-0.379110	-1.268983
0	1	-1	0.449490	0.767327
1.5	1	-0.5	1.193126	0.584223
2	1	1	1.757359	-0.131652

(a)

x	y	z	X'	Y'
0	-1	-1	-4.44949	-1.303225
-1.5	-1	-0.5	-1.936348	0.633123
-2	1	1	-0.449490	0.767327
-1.5	1	-0.5	-0.644449	2.700675
-1.5	1	2.5	-0.136105	0.359012
0	1	-1	3.414214	7.595754
1.5	1	-0.5	3.222247	2.700675
2	1	1	1.348469	0.767327

(b)

Fig. 7. Real-data for motion of *Octbox*.

(a) Data for the first case of motion, and
(b) data for the second case of motion.

Estimated Translational Vector (up to a scale factor) is:

$T = [\ 3.863706,\ 3.863707,\ 3.863717\ ]^T$

Two possible solutions of Rotation Matrix are:

$$R = \begin{bmatrix} -0.333334 & 0.816496 & 0.471405 \\ 0.772304 & -0.050319 & 0.633257 \\ 0.540773 & 0.575154 & -0.613810 \end{bmatrix} \quad \text{and} \quad R' = \begin{bmatrix} 1.000000 & 0.000000 & -0.000000 \\ -0.000000 & 0.996195 & -0.087156 \\ 0.000000 & 0.087156 & 0.996195 \end{bmatrix}$$

The directional cosines of the axis and the angle of rotation about the axis (corresponding to R and R') are respectively:

$\nu_1 = 0.576984$; $\nu_2 = 0.688845$; $\nu_3 = 0.438843$; $\theta = 182.886128$

$\nu_1' = 1.000000$; $\nu_2' = 0.000006$; $\nu_3' = 0.000001$; $\theta' = 354.999983$

Conclusion: Choose R' and its associated parameters as the final solution.

(a)

Estimated Translational Vector (up to a scale factor) is:

$T = [\ 1.035277,\ 1.035273,\ 1.035283\ ]^T$

Two possible solutions of Rotation Matrix are:

$$R = \begin{bmatrix} 1.000000 & -0.000002 & 0.000003 \\ 0.000003 & -0.087156 & -0.996195 \\ 0.000002 & 0.996195 & -0.087156 \end{bmatrix} \quad \text{and} \quad R' = \begin{bmatrix} -0.333332 & 0.471405 & -0.816496 \\ 0.772304 & 0.633257 & 0.050321 \\ 0.540773 & -0.613810 & -0.575153 \end{bmatrix}$$

The directional cosines of the axis and the angle of rotation about the axis (corresponding to R and R') are respectively:

$\nu_1 = 1.000000$; $\nu_2 = 0.000000$; $\nu_3 = 0.000002$; $\theta = 95.000000$

$\nu_1' = 0.431055$; $\nu_2' = 0.860937$; $\nu_3' = -0.195299$; $\theta' = 230.385837$

Conclusion: Choose R and its associated parameters as the final solution.

(b)

Fig. 8. Demonstration of GIPC algorithm using real-data.

(a) For the first case of *Octbox* rotation, and
(b) for the second case of *Octbox* rotation.

BIOGRAPHIES

Sunil Fotedar is presently senior associate engineer in the Tacking and Laser Systems Section at Lockheed Engineering and Sciences Company (LESC), Houston, Texas. He received B.E. degree in Electrical Engineering from Regional Engineering College, Srinagar, India; and M.S. degree in Electrical and Computer Engineering from Rice University, Houston, Texas.

Mr. Fotedar is currently involved with vision perception for extra-vehicular activity (EVA) retriever, which is being designed at NASA/Johnson Space Center to retrieve drifting tools/astronauts in space during an EVA. His research areas include digital image processing and robotic vision.

Rui J. P. deFigueiredo received his B.S. and M.S. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, and his Ph.D. degree in applied mathematics from Harvard University, Cambridge, Mass.

Since 1965 he has been a member of the faculty of Rice University in Houston, Texas, where he holds the position of Professor of Electrical and Computer Engineering and of Mathematical Sciences. Other positions that he has held include Visiting Research Professor of mathematics at the Mathematics Research Center of the University of Wisconsin, Madison, from 1972 to 1973; Visiting Professor of electrical engineering and computer sciences at the University of California, Berkeley, in 1973; and Visiting Research Professor of telecommunications at the Swiss Federal Institute of Technology E.T.H., Zurich, Switzerland in the summer of 1981. He has authored or co-authored 150 research papers. In particular, he is a co-author of the chapter on Space Robots of the *Wiley Encyclopedia of Robotics*.

Dr. deFigueiredo is the technical editor for space applications of the *IEEE Journal of Robotics and Automation*, associate editor of *Circuits, Systems and Signal Processing*, and a member of the editorial board of *Electrosoft*. He is a member or chairman of a number of national committees, including the IEEE Fellow Award Committee, the Administrative Committee of the Ocean Engineering Society and the IEEE Robotics and Automation Council. He is a Fellow of IEEE, a member of the New York Academy of Sciences and listed in *Who is Who in Engineering*. In 1987, he received a NASA Certificate of Recognition for Creative Development and Technical Innovation, and in 1988, an NCR Faculty Excellence Award.

Kumar Krishen is presently Assistant to the Director for Technology and Advanced Projects within the Mission Support Directorate at NASA/Johnson Space Center (JSC). He is responsible for planning, advocating, coordinating, and reviewing the Directorate Technology efforts and advanced projects which support the strategic technology and exploration goals of the Directorate, Center, and the Agency. He received his degrees from Kansas State University (Ph.D. and M.S.), Calcutta University (M. Tech and B. Tech.), and Jammu and Kashmir University (B.A.) in Electronics, Electrical Engineering, Radio Physics, and Mathematics.

Dr. Krishen was an Assistant Professor in Electrical Engineering at Kansas State University before joining Lockheed in 1969 as Staff Scientist. In 1976, he joined NASA/JSC and has held various positions initiating, coordinating, managing, and directing specific Advanced Programs in Earth Observations, Science Payloads, Experiment Systems, and Tracking and Communications Division. He is the author or coauthor of more than 50 paper/reports.

Dr. Krishen is a senior member of the IEEE and is a recipient of many awards, commendations, medals, and honors from universities, Lockheed, and NASA, including NASA/JSC Certificate of Commendation and is listed in *Who is Who in the World*. He serves as adjunct faculty at Rice University where he is guiding research and teaching courses in the area of electromagnetic wave propagation applications to vision and sensing. He is also a NASA/National Research Council Doctoral/Post Doctoral advisor guiding research in the areas of communications, radar, tracking, and robotic vision.