**Linear estimation of object pose
from local fits to segments**

**G.L. Bilbro
and
W.E. Snyder**

**Center for Communications and Signal Processing
Department of Electrical and Computer Engineering
North Carolina State University**

November 1986

CCSP-TR-86/24

Linear estimation of object pose from local fits to segments

G. L. Bilbro and W. E. Snyder

Center for Communications and Signal Processing
Department of Electrical and Computer Engineering
North Carolina State University
Raleigh, NC 27695-7914

ABSTRACT: A planar³ or quadric⁵ surface can be fit to a segment of range data in a locally optimal sense by selecting the minimum eigensolution of a scatter matrix for that segment. We obtain a globally optimal fit by perturbing the local eigensystems with constraints reflecting relations among the corresponding primitives of a model. These pairwise relations define a view-invariant description of the model. For segments containing a few hundred pixels, the resulting perturbation is small enough to justify a linear treatment of the coupled system. From this globally optimal fit, we determine the pose of the object algebraically.

1. Introduction

In structural image understanding, a rigid object is imagined to be constructed of primitives rigidly related to each other. The edges of wireframe models of polyhedra or the faces of quadric polyhedra^{*} are examples. These primitives are chosen so that instances of them can be identified in an image by low level, data driven, bottom up image processing techniques. These instances are called segments. After such a segmentation, a correspondence is developed between the segments and the primitives of a model. If the correspondence preserves certain properties and relations of the primitives satisfactorily, the object is said to be identified.

The pose of an imaged object may be thought of as the rigid motion which carries the model of the object from its standard displacement and orientation to the observed

^{*} We define a quadric polyhedron to be a bounded solid object whose surface can be regarded as a finite set of quadric sections.

values of position. Because of noise and because of local symmetries, the pose of an imaged object is a "global" property. It depends on all the segments comprising the imaged object, on the correspondences between segments and model primitives, and on the properties of those primitives and the relations among them. On the other extreme, a "local" value or decision³ is one that is based only on the fit parameters of one segment, independent of all other segments. All planes, for example, are locally indistinguishable (except perhaps for boundary shape, which is not invariant under occlusion), as are all spheres of identical radius. These two extremes suggest the definition of "intermediate" properties or decisions as those that depend on some proper subset of all the segments.

Faugeras, Ayache, and Faverjon⁶ used intermediate estimates of the pose (which they call "partial positioning") to constrain rigid object recognition in range imagery. In the search for correspondences between segments and primitives, they realized enormous computational savings by immediately rejecting any hypothesized correspondence that was inconsistent with the object pose estimated from currently unrejected hypotheses. Barnich and De Figueiredo¹ made similar observations in the analysis of luminance images of polyhedra. Without the use of pose or some other constraint such as adjacency, the determination of correspondences requires a worst case time that is factorial in the number of locally indistinguishable segments. Adjacency provides a means of distinguishing such congruent segments on the basis of some neighborhood of that segment. Pose provides a (stronger) means of distinguishing congruent segments on the basis of any previously identified segment of the rigid object from anywhere in the image.

As we showed² in an earlier theory paper, intermediate estimates of the pose can do more than evaluate hypothetical correspondences. A current estimate of the pose can be used to predict consistent correspondences: When a set of correspondences is sufficient to uniquely specify the pose (even if it is slightly incorrect due to noise), the location of some currently unmatched primitive can be mapped from model space back into image space. If the currently unrejected correspondences are correct, then the segment corresponding to the primitive in question must be located at the predicted position on the image plane or be occluded by some interposed surface. In the absence of congruencies among the segments within the model, the ability to predict position can accelerate a recognition program to a worst case time that is linear in the number of segments. In principle, recognition (even in the case of polyhedra) requires only the time to find the first three correct correspondences (since the pose of three non-parallel planes completely determines the pose of a rigid body) plus the time to check that all other predicted correspondences are either valid or explained by occlusion, in the sense of Cowan, Bolles, Hannah, and Herson.⁴ This suggests that recognition is a $O(n^4)$ problem in range imagery, if the segmentation is correct, if noise can be neglected, and if the computation of intermediate pose does not dominate the problem. In this paper, we present a fast, noise tolerant technique for computing pose.

We are building a procedural recognition environment to facilitate the application of a variety of *a priori* information toward the robust reduction of that initial estimate of the image pose. In that environment, a procedural model will be able to evaluate a partial list of candidate correspondences (bindings) between segments of the image and prim-

itives of a static geometric version of the same target model. The evaluation routine returns both the global likelihood that the current list (and model) is correct and a global estimate of the pose of the imaged object. The procedural model can then use this estimate of the pose to navigate in model 3-space and to map significant model points back into the image plane, thereby predicting the location of other primitives.

2. Related work

Recently, Bolle and Cooper ³ reported a technique for the maximum likelihood estimation of global pose from (1) the maximum likelihood fits of a collection of segments, (2) the correspondences between those segments and the primitives of a model, and (3) the absolute geometry of the model primitives in the model coordinate system. They based this technique on the observation that in the neighborhood of each local best fit, the parametric dependence of the fitting error is quadratic. This allowed the expression of the likelihood of an arbitrary point in parameter space (near the minimum) as the exponential of a sum of quadratic terms. By defining the parameter sets of all the segments in terms of a single rigid motion of the model, one can determine a global maximum likelihood estimate of that motion.

The technique is marvelously general. It can be used to optimally combine any local measurements, consistent with any certain global constraints. Any global transformation (and the measurements it embodies) can be optimally defined by relating all model/data pairs simultaneously and minimizing the parametric dependence of the resulting combination of local errors.

In Bolle and Cooper's formulation, the model is defined absolutely in its own coordinate system, so that the unknown transformation (a rigid motion) enters the problem explicitly and in a nonlinear way. Furthermore, the rotation part of the rigid motion itself (and all approximations to it in the nonlinear optimization search) must be prevented from stretching, shearing, or otherwise distorting its argument. This introduces into the calculation an additional nonlinearity, either with the six constraints of orthogonality for a 3 by 3 rotation matrix or with the Lie-group-theoretic exponential matrix representation⁷ for the rotations described by Faugeras, Ayache, and Faverjon.

Faugeras, Ayache, and Faverjon recently⁷ devised a related nonlinear optimization technique for autonomous navigation. They developed a complete signal processing approach for the real time estimation of the current pose of a mobile robot in a static environment. This approach optimally combines three sources of noisy data: (1) estimates of the angular locations of imaged features such as sharp edges which can be tracked in a video sequence; (2) estimates of the rigid motion of the mobile robot between image frames; and (3) an initial estimate of the pose of the robot.

In this formulation also, constraints are imposed on the local observations by relating them to the model in its own coordinate system. In this case, however, the model specification is exquisitely simple: only that the model (the environment) may not change in its own frame. This single condition can be used to design a Kalman filter to optimally model the noise in the data sequences. As its state, the filter maintains the current rigid motion that relates the pose of the robot to the scene, which is regarded as a fixed arrangement of observed features. This is ideal for navigation in an unknown static

scene, and the authors state that they can adapt the technique to detect (presumably small) changes in the scene.

In that work, Faugeras, Ayache, and Faverjon⁷ do not specify *a priori* knowledge about the scene. The correspondences used are between identical features in successive frames of a video sequence. Complex geometrical relations between different parts of the model (scene) are not readily incorporated into that approach. That solution to the navigation problem is therefore not suitable to object recognition applications, in which an object is identified by the relative geometry of its primitives. Furthermore, that system is designed to update the estimate of the robot's pose along a trajectory of small changes. It is not intended to determine a completely unknown pose without an initial estimate. The analytic representation of rotations introduced in that work could be applied to such a determination, though it would also require nonlinear optimization.

3. Two subproblems: pose and fit

We separate the problem of global pose determination into (1) determining the globally optimal fit of the model to the data, and (2) determining the rigid motion that carries the model from its standard frame of specification to that best fit in the image frame.

In sections 4 and 5, we solve the first subproblem for the case of planes. It is the solution of a linear system that defines the minimum of a sum of quadratic expressions. The resulting globally optimal fit is an adjustment for the locally optimal fit of each segment. These adjustments to the original minima are of first order in the constraints imposed by the model. This is the same order as the adjustments to the locally optimal fits implicit in

the global optimization of Bolle and Cooper, which also minimizes a similar sum of quadratic variations of the error around the local optima. In this sense, our calculation is identical to theirs.

Rigid motion manipulations enter our computation only in the second subproblem. In the first step, the local segment fits were adjusted to conform with the model to first order. And to that same order of accuracy, any conveniently chosen set of adjusted segments determines the global pose. In section 6, we calculate the pose of planar polyhedra by simple linear algebra without the imposition of orthogonality constraints on the rotation part of the transformation.

In section 7, we show that the analysis of quadric polyhedra is very similar to the analysis of planar polyhedra, and we present only the modifications.

We regard our approach as a reformulation of Bolle and Cooper's work. However ours is performed entirely in the image coordinate system, where it is valid to treat the small errors in the local fits to first order. We also express the model constraints in a coordinate invariant form as pairwise relations between primitives. This difference allows us to avoid the complexities of rigid motion manipulations while we are fitting the image. Other differences are mentioned in section 8.

4. Locally optimal fit of polyhedra

The eigensystem description of noninteracting planar fits to segments of data has a familiar intuitive interpretation. The eigenvectors are the normals of certain planes and the eigenvalues are the associated sum-squared errors of the fit. We therefore introduce

the perturbation theoretic calculation of pose in terms of polyhedral objects. The application to quadric polyhedra, while more interesting, is formally so similar that we may refer to the same equations in the subsequent analysis of quadric polyhedra.

Consider one planar segment of data from a range image. It is (or defines) a collection of N 3-vectors,

$$\{ x_i, y_i, z_i \}_{i=1}^N,$$

and the 4-vector function

$$v_i = (1, x_i, y_i, z_i)^T, \quad (4.1)$$

and the coefficient vector

$$\alpha = (a_d, b_x, b_y, b_z)^T, \quad (4.2)$$

which we have partitioned into the sub vector a (a scalar for planes which will specify a distance from the origin) and the sub vector b (which defines the normal of the plane).

We take the b part of α to have unit length and use the constraint matrix

$$K = \text{diag}(0, 1, 1, 1) ,$$

to write that condition in terms of α

$$\alpha^T K \alpha = b^T b = 1 .$$

Now α represents a plane and the data vector v_i is on that plane if

$$v_i^T \alpha = 0 . \quad (4.4)$$

In general, there will exist no plane passing through all the points of the segment. We define a sum square error for a plane α as the sum of the squares of the normal distances from the data points to a plane

$$\xi = \alpha^T S \alpha , \quad (4.5)$$

where the scatter matrix

$$S = \sum_{i=1}^N v_i v_i^T, \quad (4.6)$$

completely characterizes the segments in the subspace of planes. We can define the locally optimal plane as the one that minimizes this sum square error by finding the extreme values of

$$\alpha^T S \alpha - \lambda \alpha^T K \alpha, \quad (4.7)$$

where we have introduced the constraint with the Lagrange multiplier λ , which will ultimately assume the value of the sum squared error in the usual way.

The extreme values of E occur when

$$\begin{bmatrix} S_{aa} & S_{ab} \\ S_{ba} & S_{bb} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \lambda \begin{bmatrix} 0 \\ K_{bb} b \end{bmatrix}, \quad (4.8)$$

where we have written the a and b parts of α explicitly and partitioned S into parts associated with those parts: S_{aa} is the upper left 1×1 part of S and S_{ab} is the upper right 1×3 part, and so forth. For planar fits, K_{bb} is simply the 3×3 identity matrix and could be suppressed. However the subsequent analysis for quadric primitives, in section 7, differs in this point, so we have written K_{bb} here to facilitate that presentation. We have omitted the term containing the first 1×1 part of the constraint matrix K because it vanishes.

We can solve the preceding partitioned equation with

$$a = -S_{aa}^{-1} S_{ab} b, \quad (4.9)$$

where the scalar reciprocal of S_{aa} will become a matrix inverse for the case of quadrics.

The unknown b satisfies the eigensystem equation

$$Hb = \lambda b \quad , \quad (4.10)$$

where the covariance H is

$$H = S_{bb} - S_{ba}S_{aa}^{-1}S_{ab} \quad . \quad (4.11)$$

For planes, this eigensystem comprises three eigenvalues λ_e and their associated eigenvectors b_e . This is useful ³ since the *smallest* eigenvalue (which we will refer to as λ_1) identifies the best local fit to the segment.

The sum squared fit error for an arbitrary plane

$$\alpha = \begin{bmatrix} A \\ B \end{bmatrix} \quad , \quad (4.12)$$

is

$$E = A^T S_{aa} A + 2A^T S_{ab} B + B^T S_{bb} B \quad , \quad (4.13)$$

where we have used capital A and B to indicate the a and b parts of an arbitrary α . This expression can be simplified by using equation 4.11 to write S_{bb} in terms of H and by writing the arbitrary A as a deviation, δA , from the special value defined by B and equation 4.9

$$A = S_{aa}^{-1}S_{ab}B + \delta A \quad . \quad (4.14)$$

With these substitutions, the sum squared fit error becomes

$$E = \delta A^T S_{aa} \delta A + B^T H B \quad . \quad (4.15)$$

The second term is the error resulting from a rotation of the plane from the optimal orientation. The first term is the error resulting from a displacement of the plane from the value that is optimal for the given B .

We can resolve

$$B = \sum_{e=1}^3 C_e b_e , \quad (4.16)$$

in terms of the fitting eigensystem of the segment, with the coefficients

$$C_e = b_e^T B , \quad e=1,2,3 , \quad (4.17)$$

and rewrite the error

$$E = \delta A^T S_{aa} \delta A + \sum_{e=1}^3 \lambda_e C_e^2 . \quad (4.18)$$

The total error incurred by fitting M arbitrary planes to M segments is the sum of M such contributions. The best zeroth order fit to such a collection is the list of local optima, the $e=1$ terms. In the neighborhood of this zeroth order fit, the additional error incurred derives from the other two values of e . Bolle and Cooper³ observed this, and the reader is referred to their excellent account for planar fits. However, Bolle and Cooper did not address in their report the possible application of eigensystem analysis to general quadric fits. Nor did they note that the constraints imposed on the local optima by the model could be expressed in terms of those local eigensystems. When expressed in terms of eigensystems, the constraints are not only view-invariant, but are also small in the sense that they only weakly perturb the local optima. The advantage gained is the possibility of obtaining a good solution for global pose from a linearization of the problem.

5. Globally optimal fit

To find the total fit error over all segments for a collection of arbitrary planes, we add a Lagrange term constraining the angles between pairs of fitting planes to equal the angles between corresponding planes in the model. For simplicity, we consider only the angular constraints imposed by the model on the local optima; only the orientations of

the local fits are affected. With no distance constraints, the globally optimal planar fits to the segments are not translated from their locally optimal distance from the origin. By construction, the first term of equation 4.18 vanishes for such an optimal distance, because $\delta A = 0$. This allows us to write the local error for segment k as:

$$E_k = \sum_{c=1}^3 \lambda_{kc} C_{kc}^2 . \quad (5.1)$$

The incorporation of constraints on the relative displacements between planes is similar to angular constraints except in one respect, which we discuss in section 8.

5.1. The model planes

We now introduce a model polyhedron comprised of $L \geq M$ plane primitives of the form of equation 4.2 in some convenient model coordinate system. We assume that this is the correct model and that the M segments together are an instance of this model. We assume that the correspondence between segments and primitives is given and simply label the primitives with the corresponding segment number k , so that

$$\alpha_k^T v = 0 \quad , \quad k=1, \dots, M < L \quad (5.2)$$

where $v = (1, x, y, z)^T$ is any point belonging to the k^{th} model plane and α_k is the parameter vector defining the primitive plane in the model frame

$$\alpha_k = \begin{bmatrix} d_k \\ n_k \end{bmatrix} , \quad (5.3)$$

and $n_k^T = (n_{kx}, n_{ky}, n_{kz})$ is a vector normal to the primitive. We assume α_k is normalized (just as were the coefficient vectors defining the segment fits of section 4), so that their b parts have unit length

$$\alpha_k^\top K \alpha_k = n_k^\top n_k = 1, \quad k=1, \dots, M. \quad (5.4)$$

This form will apply to quadric models; only the values and dimensions will differ.

5.2. The constraint imposed by the model planes on the fitting planes

With this normalization, the inner product $n_k^\top n_{k'}$ is a view-invariant (or frame-invariant or scalar) property of the model. For planes, it is the cosine of the angle between the two planes. There are $(M-1)^2$ such scalars (half that many distinct ones) that together determine every angle in the model: the angle between every distinct pair of model planes. We require the globally optimal fit to the data to obey

$$B_k^\top B_{k'} = n_k^\top n_{k'}, \quad k, k'=1, 2, \dots, M, \quad k \neq k', \quad (5.5)$$

where B_k defines the arbitrary plane fit to the k^{th} segment so that $B_k^\top B_{k'}$ is the cosine of the angle between fit k and k' . We therefore seek the M fitting planes defined by the $3M$ coefficients C_{ke} that minimize the total error and that obey

$$\sum_{e=1}^3 C_{ke} C_{k'e} b_{ke}^\top b_{k'e} = n_k^\top n_{k'}, \quad k, k'=1, 2, \dots, M, \quad k \neq k', \quad (5.6)$$

where $b_{ke}^\top b_{k'e}$ is the dot product of the e^{th} eigenvectors of the k^{th} segment with the k'^{th} segment, just as $n_k^\top n_{k'}$ is the dot product of the normals of the k^{th} and k'^{th} primitive planes of the model.

5.3. The exact system

We define an augmented total error F' by adding the constraint term to the sum of local errors.

$$F' = \sum_k E_k - \sum_{k, k'} \sum_{e, e'=1}^3 \mu_{kk'} b_{ke}^\top b_{k'e} C_{ke} C_{k'e}, \quad (5.7)$$

The constraints are the usual $(M-1)^2$ Lagrange multipliers $\mu_{kk'}$. Now we would like to

solve the $3M$ equations

$$\frac{\partial F}{\partial C_{ke}} = 0 \quad , \quad k=1, \dots, M \quad , \quad e=1,2,3 \quad . \quad (5.8)$$

Unfortunately, since both the μ 's and the C 's are unknown, the Lagrange terms are cubic, and the simultaneous system, equation 5.8, is quadratic.

5.4. Linearization of the system

When the locally optimal fits are good and the segments are well represented by the first eigenvectors of their eigensystems, then the set

$$C_{k1}=1 \quad , \quad C_{k2}=C_{k3}=0 \quad , \quad k=1,2, \dots, M \quad , \quad (5.9)$$

is close to the global optimal set.

We therefore write the globally optimal B_k as the locally optimal B_k with small corrections: $C_{k1}=1$ and C_{k2}, C_{k3} small. Note that this preserves the normalization of B_k to first order since the correction is orthogonal to the local optimum. The equations of constraint are, to first order,

$$\sum_{e'=2,3} b_{k1}^T b_{k'e'} C_{k'e'} + \sum_{e=2,3} b_{ke'}^T b_{k'1} C_{ke} = n_k^T n_{k'} - b_{k1} b_{k'1} \quad , \quad (5.10)$$

where the b 's and n 's are known.

5.5. The linear solution

To the first order, the solution is

$$C_{ke} = \frac{1}{\lambda_{ke}} \sum_{k' \neq k} \mu_{kk'} b_{ke}^T b_{k'1} \quad k=1, \dots, M \quad , \quad e=2,3 \quad , \quad (5.11)$$

where the $\mu_{kk'}$ are determined by substituting this result into the $(M-1)^2$ equations constraining the C_{ke} . Since the constraints are symmetric, we may take $\mu_{kk'} = \mu_{k'k}$, which

leaves us with a system of

$$\binom{M}{2} = \frac{M(M-1)}{2}$$

linear equations with the same number of unknowns. The solution is unique and defines the global fit. From this solution the total fit error can easily be calculated. A small fit error verifies a set of correspondences between segments and primitives. A large error means that either the correspondences are incorrect or the model does not identify the imaged object. The pose of the object can also be calculated, as section 6 illustrates.

5.6. Summary of concept

At this point, we summarize for the readers the two fundamental principles underlying this theory.

Principle 1:

The surfaces of the image are constrained to have the same pairwise angular relationship as the corresponding surfaces in the model. The constraint imposed in equation 5.5 guarantees this.

Principle 2:

There exists a rigid motion which carries each surface into almost exact coincidence with the corresponding image surface. Therefore, the minimizing eigenvector of the fit will align very closely with the normal of the corresponding model surface. In equation 5.9, we use this assumption to find a linear solution to the problem. In section 6, we determine the rigid motion.

6. Determination of the global pose from the adjusted fit

The pose can be calculated to first order by determining first the rotation that takes three orthonormal vectors in the model frame to three corresponding unit vectors in the image frame. We construct the first set from three convenient model primitives, e.g. the first, n_1 , the projection of n_2 orthogonal to n_1 , and the cross product of those two

$$f_M = n_1 \quad , \quad (6.1)$$

$$g_M = \frac{n_2 - \frac{n_1 n_1^T n_2}{n_1^T n_2}}{1 - \frac{(n_1^T n_2)^2}{n_1^T n_2}} \quad , \quad (6.2)$$

$$h_M = f_M \times g_M \quad . \quad (6.3)$$

From the C_{ke} , the globally adjusted segment fits can be calculated from corresponding segments

$$f_I = B_1 \quad , \quad (6.4)$$

$$g_I = \frac{B_2 - \frac{B_1 B_1^T B_2}{B_1^T B_2}}{1 - \frac{(B_1^T B_2)^2}{B_1^T B_2}} \quad , \quad (6.5)$$

$$h_I = f_I \times g_I \quad , \quad (6.6)$$

where B_k is the globally adjusted fit for the k^{th} segment

$$B_k = \sum_{e=1}^3 C_{ke} b_{ke} \quad . \quad (6.7)$$

Then the linear transformation

$$R = f_I f_M^T + g_I g_M^T + h_I h_M^T \quad , \quad (6.8)$$

is the desired rotation which carries any linear combination of the model bases, f_M, g_M, h_M into the same combination of corresponding image bases, f_I, g_I, h_I .

The remaining translation can be found by identifying a point in the model frame view-invariantly, say as the intersection of the first three model planes, which is the

solution of the simple system

$$d_k + n_{kx}x_M + n_{ky}y_M + n_{kz}z_M = 0 \quad , \quad k=1,2,3 \quad , \quad (6.9)$$

and the point in the image frame defined by the corresponding system

$$A_k + B_{kx}x_I + B_{ky}y_I + B_{kz}z_I = 0 \quad , \quad k=1,2,3 \quad , \quad (6.10)$$

where the a part of the adjusted fit is

$$A_k = \sum_{e=1}^3 C_{ke} a_{ke} \quad , \quad (6.11)$$

and a_{ke} is the a part of the fit associated with the e^{th} eigenvector of the k^{th} segment by equation 4.9 with the S submatrices for the k^{th} segment. If we choose the convention that the translation T is to be applied after the rotation, then the correct T must carry every model point to the corresponding image point. In particular, it must satisfy

$$(x_I, y_I, z_I)^T = R (x_M, y_M, z_M)^T + T \quad ,$$

so that

$$T = (x_I, y_I, z_I)^T - R (x_M, y_M, z_M)^T \quad .$$

The choice of segments is not critical, since all have been adjusted to conform exactly with the model to first order. Intuitively, those segments that are the least adjusted by the global optimization should be somewhat better. Segments with very poor local fits should be avoided in the determination of R and T . In addition, no two of the three segments used to fix a point for the calculation of T should be even nearly parallel. Otherwise there appears to be little justification for a more sophisticated determination of R and T .

7. Quadric polyhedra

The formal analysis of quadric primitives and segments is almost identical to the preceding presentation. For the 3-vector in equation 4.1 associated with the i^{th} point of a segment we now use the 10-vector function

$$v_i = (1, x_i, y_i, z_i, x_i^2, y_i^2, z_i^2, y_i z_i, z_i x_i, x_i y_i)^T$$

in the definition of S in equation 4.6. The parameter vector is still partitioned in the same way, but has now a 4×1 a part,

$$(a_1, a_x, a_y, a_z) , \quad (7.1)$$

which specifies the constant and linear terms of a quadric function instead of simply the constant distance for planes, and a 6×1 b part

$$(b_{x^2}, b_{y^2}, b_{z^2}, b_{xy}, b_{yz}, b_{yz}) , \quad (7.2)$$

which specifies the pure and mixed quadratic terms. Similarly, S_{aa} of the partitioned segment scatter matrix in equation 4.8 becomes a 4×4 submatrix, S_{ab} becomes a 4×6 submatrix, and so forth. The final difference is the constraint matrix K , which is still diagonal, but partitions into a vanishing 4×4 upper left submatrix and a 6×6 submatrix,

$$K_{bb} = \text{diag}(1, 1, 1, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}) , \quad (7.3)$$

in the lower right. This particular view-invariant constraint on the quadric fit was suggested by Faugeras, et al.⁵ and examined in more detail⁸ by Groshong. Some such constraint is necessary to exclude the trivial solution of a null parameter vector during the minimization of the error for the segment. This particular constraint has the added benefit of excluding simple planar fits (degenerate quadrics) also. We have exercised⁸ this formalism thoroughly and successfully with synthetic images of spheres, cones,

cylinders, and general ellipsoids. The most serious reservation is that the segment must contain at least 100 points and must have a diameter that is on the order of the important lengths of the quadric section. We refer the reader to these reports ^{8,5} for further information.

The chief formal consequence of this new K_{bb} is that equation 4.10 must be transformed into

$$H' = \lambda b'$$

where

$$H' = J^T H J, \quad b' = J b, \quad (7.5)$$

and the 6×6 diagonal matrix

$$J = K_{bb}^{-1/2} = \text{diag}(1, 1, 1, \sqrt{2}, \sqrt{2}, \sqrt{2}) \quad (7.6)$$

so that $K_{bb} = J^{-2}$. This recovers the form of a simple symmetric eigenvalue problem. This transformation must also be applied to the representation of the primitives, but with this change, the remaining treatment of planes applies to quadric fitting and quadric constraints. The primitives are represented by 10-vectors and are normalized as before to have unit b parts (in the transformed space, which amounts to $\alpha^T K \alpha = 1$ in the original space.) The b part of the k^{th} model primitive is n_k and is now a 6-vector. The dot products between pairs of primitives are now sums over all six components. They no longer have interpretations so simple as the cosine of the angle between planes, but they still represent the invariant properties of the model and are used the same way. Of course, the upper limit on all other sums over the eigensystem index, e , is also increased from $e = 3$ to $e = 6$.

Finally, the determination of the pose from the adjusted fit will be substantially different for quadrics, but it is still algebraic and elementary. Some of this calculation can be found in Hemler.⁹

8. Discussion

In addition to the exploitation of the entire eigensystem of the segment fits and the expression of the model in a view-invariant form, there are several other differences between our approach and that of Bolle and Cooper.³ We use general quadrics instead of restricting the form of the fitting functions to cylinders and spheres. Their additional restriction gives tighter fits to segments that are of fixed "optimal" size. We assume a problem domain with primitives large enough for a connected component analysis to produce segments typically containing a few hundred pixels. In this domain, our general quadric fits to synthetic data yield axis directions and semimajor axis lengths within a few percent of the correct values. This is certainly within the power of a first order analysis. For quadrics, the pairwise constraints correct small errors in the relative magnitudes of the semimajor axis lengths just as they correct small errors in axis orientation (which was the only effect in the case of planes).

An important omission from this presentation is the treatment of distance constraints. Bolle and Cooper do include displacement constraints, which are formally similar to rotational constraints in their treatment. In our treatment of planes, distance constraints involve a four segment interaction instead of the two segment interaction of the preceding section, but are analogous otherwise. Quadrics can also be treated with a four segment interaction, but may yield to simpler constraints. We are currently studying this

problem.

Finally, Bolle and Cooper³ arrive at their technique by considering the likelihood of the fit, given the data. Ultimately, they minimize the logarithm of the probability of the data. This quadratic log probability is the sum square error that we minimize in this paper. We prefer to consider the problem in terms of sum square error, but clearly each view affords its own useful insight.

References

1. B. Bamieh and R. J. P. De Figueiredo, "A General Moment-Invariants/Attributed-Graph Method for Three-Dimensional Object Recognition from a Single Image," *IEEE Journal of Robotics and Automation* **RA-2**(1)(March, 1986).
2. G. L. Bilbro, W. E. Snyder, and J.E. Franke, "A Linear Time Theory for Recognizing Surfaces in 3-D," (*Submitted to*) *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (April 1986).
3. R.M. Bolle and D.B. Cooper, "On Optimally Combining Pieces of Information, with Application to Estimating 3-D Complex-Object Position from Range Data," *IEEE Trans. on Pattern Analysis and Machine Intelligence* **PAMI-8**, 5(September 1986).
4. C.K. Cowan, R. C. Bolles, M.J. Hannah, and J.A. Herson, "Edge Chain Analysis for Object Verification," *Proc. IEEE International Conf. on Robotics and Automation*, pp. 937-942 (April, 1986).
5. O. D. Faugeras, M. Hebert, and E. Pauchor, "Segmentation of Range Data Into Planar and Quadric Patches," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8-13 (1983).
6. O. D. Faugeras, N. Ayache, and B. Faverjon, "A Geometric Matcher for Recognizing and Positioning 3-D Rigid Objects," *The First Conference on Artificial Intelligence Applications*, (December, 1984).
7. O.D. Faugeras, N. Ayache, and B. Faverjon, "Building visual maps by combining noisy stereo measurements," *Proc. IEEE International Conf. on Robotics and Automation*, pp. 1433-1438 (April, 1986).
8. B. R. Groshong, *Segmentation of Range Images Containing Curved Surfaces*, Master's Thesis, North Carolina State University (June 1986).
9. Paul F. Hemler, "Active Model Matching in Range Images," (*submitted to*) *Proc. IEEE International Conf. on Robotics and Automation*, (April 1987).