MOTION AND STRUCTURE FROM POINT CORRESPONDENCES: A ROBUST ALGORITHM FOR PLANAR CASE WITH ERROR ESTIMATION

JUYANG WENG, NARENDRA AHUJA, THOMAS S. HUANG

Coordinated Science Laboratory
University of Illinois Urbana, IL 61801

Abstract

This paper deals with the problem of determining motion and structure for a planar surface and the error estimation. Since the motion of a planar patch is a degenerate case for linear algorithms (algorithms that consist of solving mainly linear equations and give closed-form solution) for general surfaces, the motion of such a planar surface should be considered separately. First, a new algorithm is presented that gives a closed-form solution to motion parameters using monocular perspective images of the points on a planar surface. The algorithm is simpler and more reliable, in the precence of noise, than the existing ones. There are generally two solutions given two image frames. From three image frames the solution is generally unique. An approach is proposed to test whether the points are coplanar.

Based on this algorithm, the errors in the motion parameters and surface structure can be estimated for each pair of images. Specifically, the standard deviation of the errors is calculated in terms of the variance of the errors in the image coordinates. The presented approach to estimating errors is applicable to least-squares, pseudo-inverse and eigenvalue eigenvector problems.

1. INTRODUCTION

If the object points are coplanar, *i.e.*, they all lie in a plane in 3D, the existing linear algorithms that gives closed-form solutions to motion and structure from two images fail to give unique solution. An algorithm that is devoted to coplanar points is required to solve the problem.

Tsai and Huang [Tsai83] develop a linear algorithm to solve the problem of motion of a planar patch using singular value decomposition. However their algorithm is primarily for the noise free images and solve for the exact solution. In the presence of noise, several problems have to be solved. First, A simple and stable algorithm is required in the presence of noise. Second, how can we check for the case of degeneracy or near-degeneracy? More generaly how can we assess the reliability of the solutions? Third, how can we test whether the points are coplanar, given the images of the points? In this paper, we address these problems.

2. A TWO-VIEW MOTION ALGORITHM FOR A PLANAR PATCH

Without loss of generality, we assume that the focal length is unity. Visible objects are always located in front of the image plane, *i.e.*, $z>1$. Geometry of the setup is shown in Fig. 1.

We first introduce some notations. Matrices are denoted by capital italics. Vectors are denoted by bold fonts, either capital or small. A vector is sometimes regarded as a column matrix. Vector operations such as cross product ($\times$) and matrix operations such as matrix multiplication are appled to three-dimensional vectors. Matrix operations precede vector operations. For a matrix $A=[a_{ij}]$, $\|A\|$ denotes the Euclidean norm of the matrix. *i.e.*, $\|[a_{ij}]\|=\sqrt{\sum_{ij} a_{ij}^2}$. We define a mapping $[\cdot]_\times$ from a three-dimensional vector to a 3 by 3 matrix:

$$[(x_1, x_2, x_3)]_\times = \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix}. \tag{2.1}$$

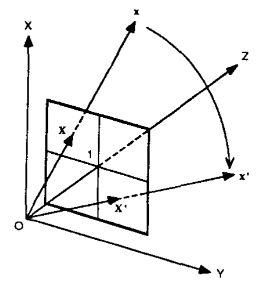

Fig.1. Geometry of the camera model

Consider a point P on the object which is visible at two time instants. The following notation is used for the spatial vectors and the image vectors (see Fig. 1).

$\mathbf{x} = (x, y, z)$ *spatial vector of P at time* t_1

$\mathbf{X} = (u, v, 1) = (\frac{x}{z}, \frac{y}{z}, 1)$ *image vector of P at time* t_1

where (u, v) are the image coordinates of of the point. The corresponding vectors at t_2 are primed. Let R and $\mathbf{T}$ be the rotation matrix and the translational vector, respectively. The spatial points at the two time instants are related by

$$\mathbf{x}' = R\mathbf{x}+\mathbf{T} \tag{2.4}$$

Assume the plane that the points lie in at time t_1 is represented by

$$\mathbf{N}^t\mathbf{x}=1 \tag{2.8}$$

where $\mathbf{N}$ is the normal vector of the plane and the z component of $\mathbf{N}$ is always positive, since the plane is located at the space $z>0$. By this representation, we exclude the cases where the plane goes through the origin, since in that case the plane is invisible to the camera (the image of the plane is a straight line).

We want to determine the relative depth $\frac{z}{\|\mathbf{T}\|}$ and $\frac{z'}{\|\mathbf{T}\|}$. Equivalently we determine the relative normals instead

$$\tilde{\mathbf{N}}=\|\mathbf{T}\|\mathbf{N}^t, \tag{2.11}$$

From the relative normal we can determine the relative depth.

The magnitude of the translational vector ($\|\mathbf{T}\|$) can not be determined by monocular vision and the depths of the object points (z_i and z_i') can only determined up to a scale factor $\|\mathbf{T}\|^{-1}$.

Due to space limitation we shall just state the algorithm and results without presenting proofs.

Algorithm

(i) *Solve for intermediate parameter matrix* $E = R + \mathbf{T}\mathbf{N}^t$

Let $\mathbf{X}_i = (u_i, v_i, 1)$, $\mathbf{X}_i' = (u_i', v_i', 1)$, $i = 1, 2, \cdots, n$, be the corresponding image vectors of n ($n \geq 4$) points. Let A be a $2n$ by 9 matrix such that

$$A = \begin{bmatrix} \mathbf{X}_1^t & \mathbf{0} & -u_1'\mathbf{X}_1^t \\ \mathbf{0} & \mathbf{X}_1^t & -v_1'\mathbf{X}_1^t \\ \mathbf{X}_2^t & \mathbf{0} & -u_2'\mathbf{X}_2^t \\ \mathbf{0} & \mathbf{X}_2^t & -v_2'\mathbf{X}_2^t \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \mathbf{X}_n^t & \mathbf{0} & -u_n'\mathbf{X}_n^t \\ \mathbf{0} & \mathbf{X}_n^t & -v_n'\mathbf{X}_n^t \end{bmatrix} \tag{Ag.1}$$

and $\mathbf{h}$ be a 9-dimensional vector

$$\mathbf{h} = (h_1, h_2, h_3, h_4, h_5, h_6, h_7, h_8, h_9). \tag{Ag.2}$$

We solve for unit vector $\mathbf{h}$ such that

$$\| A\mathbf{h} \| = min \tag{Ag.3}$$

The solution of $\mathbf{h}$ is the unit eigenvector of A^tA associated with the smallest eigenvalue. Then E_s is determined by

$$E_s = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix} \tag{Ag.4}$$

Let $H = [\mathbf{h}_1\ \mathbf{h}_2\ \mathbf{h}_3]$ be a 3 by 3 orthonormal matrix such that

$$H^t E_s^t E_s H = diag(\gamma_1, \gamma_2, \gamma_3) \tag{Ag.5}$$

(where $diag(\gamma_1, \gamma_2, \gamma_3)$ denotes the conventional diagonal matrix with the corresponding diagonal entries) with $\gamma_1 \leq \gamma_2 \leq \gamma_3$. Then

$$E = \frac{1}{\sqrt{\gamma_2}} E_s. \tag{Ag.6}$$

If

$$\sum_i \mathbf{X}_i' \cdot E\mathbf{X}_i < 0 \tag{Ag.7}$$

$E \leftarrow -E$. The summation in (Ag.7) is over several values of i's to reduce the sensitivity to noise (usually three or four values of i will suffice).

(ii) *Solve for* R, $\hat{\mathbf{T}}$ *and* $\tilde{\mathbf{N}}$ *from* E

We have

$$H^t E^t E H = diag(\gamma_1/\gamma_2, 1, \gamma_3/\gamma_2) \triangleq diag(\lambda_1, \lambda_2, \lambda_3) \tag{Ag.8}$$

Case (1). $\lambda_1 < 1 < \lambda_3$ (iff $\mathbf{T} \times (R\mathbf{N}) \neq \mathbf{0}$):

There exist two set of solutions that give the same images. Let

$$\alpha = \sqrt{\frac{\lambda_3 - 1}{\lambda_3 - \lambda_1}}, \qquad \beta = \sqrt{\frac{1 - \lambda_1}{\lambda_3 - \lambda_1}}. \tag{Ag.9}$$

The first set of solutions: Let

$$\mathbf{V}_1 = \alpha\mathbf{h}_1 + \beta\mathbf{h}_3, \qquad \mathbf{V}_2 = \mathbf{h}_2. \tag{Ag.10}$$

Solve for R such that

$$\sum_{i=1}^{2} \| R\mathbf{V}_i - E\mathbf{V}_i \|^2 = min, \ subject\ to: R\ is\ a\ rotation\ matrix. \tag{Ag.11}$$

The solution of (Ag.11) is as follows:

Let $\mathbf{W}_i = E\mathbf{V}_i$, $i = 1, 2$. Define a 4 by 4 matrix B by

$$B = \sum_{i=1}^{2} B_i^t B_i \tag{Ag.12}$$

where

$$B_i = \begin{bmatrix} 0 & (\mathbf{V}_i - \mathbf{W}_i)^t \\ \mathbf{W}_i - \mathbf{V}_i & [\mathbf{W}_i + \mathbf{V}_i]_\times \end{bmatrix}. \tag{Ag.13}$$

Let $\mathbf{q} = (q_0, q_1, q_2, q_3)$ be the unit eigenvector of B associated with the smallest eigenvalue. The solution of rotation matrix R in (Ag.11) is

$$R = \begin{bmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 - q_0q_3) & 2(q_1q_3 + q_0q_2) \\ 2(q_2q_1 + q_0q_3) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_2q_3 - q_0q_1) \\ 2(q_3q_1 - q_0q_2) & 2(q_3q_2 + q_0q_1) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{bmatrix}$$

Let

$$\hat{\mathbf{N}} = \mathbf{V}_1 \times \mathbf{V}_2. \tag{Ag.15}$$

If the third components of $\hat{\mathbf{N}}$ is negative, $\hat{\mathbf{N}} \leftarrow -\hat{\mathbf{N}}$. Let

$$\tilde{\mathbf{T}} = E\hat{\mathbf{N}} - R\hat{\mathbf{N}} \tag{Ag.16}$$

$$\hat{\mathbf{T}} = \frac{\tilde{\mathbf{T}}}{\| \tilde{\mathbf{T}} \|} \tag{Ag.17}$$

$$\tilde{\mathbf{N}} = \| \tilde{\mathbf{T}} \| \hat{\mathbf{N}}. \tag{Ag.18}$$

The second set of solutions: Change the sign of the β we got in (Ag.9) (*i.e.* $\beta \leftarrow -\beta$) and keep the α unchanged. (Ag.10)-(Ag.18) give the second set of solutions.

Case (2). $\lambda_1 = 1 < \lambda_3$ (iff $\mathbf{T} // R\mathbf{N}$ and $2R\mathbf{N} \cdot \mathbf{T} > -\| \mathbf{T} \|^2 \| \mathbf{N} \|^2$) :

(Ag.9)-(Ag.18) give the unique solution. In this case $\alpha = 1$ and $\beta = 0$.

Case (3). $\lambda_1 < 1 = \lambda_3$ (iff $\mathbf{T} // R\mathbf{N}$ and $2R\mathbf{N} \cdot \mathbf{T} < -\| \mathbf{T} \|^2 \| \mathbf{N} \|^2$) :

(Ag.9)-(Ag.18) give the unique solution. In this case $\alpha = 0$ and $\beta = 1$.

Case (4). $\lambda_1 = 1 = \lambda_3$ (iff $\mathbf{T} // R\mathbf{N}$ and $2R\mathbf{N} \cdot \mathbf{T} = -\| \mathbf{T} \|^2 \| \mathbf{N} \|^2$) :

If $\det(E) > 0$, report $\mathbf{T} = 0$. $R = E$. $\tilde{\mathbf{N}}$ can not be determined.

$\det(E) < 0$ happens only if the back side of the plane faces the camera after motion. This can not happen for a opaque plane in reality. If the plane is transparent and the points on the plane are visible on both sides, this case can happen. If so, the solutions are infinitely many. For any unit $\hat{\mathbf{N}}$, the following is a solution

$$R = E(I_3 - 2\hat{\mathbf{N}}\hat{\mathbf{N}}^t) \tag{Ag.19}$$

$$\tilde{\mathbf{T}} = -2R\hat{\mathbf{N}}. \tag{Ag.20}$$

(Ag.17) and (Ag.18) give $\hat{\mathbf{T}}$ and $\tilde{\mathbf{N}}$, respectively.

Note: The necessary and sufficient conditions for each case in step (ii) to occur are for the noise-free images. With noise, generaly only Case (1) could happen. following way. Without noise, if the rank of A in (Ag.1) is more than 8 the points are not coplanar. In the presence of noise, if

$$\lambda_1 > \varepsilon, \tag{Ag.22}$$

where epsilon is a threshold based on the error analysis of λ_1 discussed in Section 3, the points are not coplanar.

We now present some uniqueness results.

We can determine $\mathbf{h}$ (and consequently E) up to a scale factor from (Ag. 3) if and only if $rank(A) = 8$.

Theorem 1. $rank(A) = 8$ if and only if there exist a set of four points such that the images of any three points in this set do not lie in a straight line either at time t_1 or t_2. □

Corollary. $rank(A) = 8$ if and only if

1) there exist a set of four points in the object plane such that any three points in this set do not lie in any straight line in the object plane and

2) if the object plane is extended, it does not go through the focal

point of the camera either at time t_1 or time t_2. □

Theorem 2. Let the eigenvalues of E^tE be $\lambda_1, \lambda_2, \lambda_3$, with $\lambda_1 \le \lambda_2 \le \lambda_3$. Then

1) E^tE has multiple eigenvalues iff $\mathbf{T} // R\mathbf{N}$
2) $\lambda_1 = 1 < \lambda_3$ iff $\mathbf{T} // R\mathbf{N}$ and $2R\mathbf{N}\cdot\mathbf{T} > -\|\mathbf{T}\|^2\|\mathbf{N}\|^2$
3) $\lambda_1 < 1 = \lambda_3$ iff $\mathbf{T} // R\mathbf{N}$ and $2R\mathbf{N}\cdot\mathbf{T} < -\|\mathbf{T}\|^2\|\mathbf{N}\|^2$
4) $\lambda_1 = 1 = \lambda_3$ iff $\mathbf{T} // R\mathbf{N}$ and $2R\mathbf{N}\cdot\mathbf{T} = -\|\mathbf{T}\|^2\|\mathbf{N}\|^2$. □

Theorem 3. Assume the eigenvalues of E^tE are distinct, there are exactly two sets of solutions for R, $\hat{\mathbf{T}}$ and $\hat{\mathbf{N}}$ in the equation

$$R + \hat{\mathbf{T}}\hat{\mathbf{N}}^t = E \tag{2.40}$$

with the constraints that R is a rotation matrix and $\hat{\mathbf{N}}$ is a unit vector. □

In fact, generally there exist two planes (under different motions) that give the same images at the two time instants.

Theorem 4. If

$$R_a + \hat{\mathbf{T}}_a\tilde{\mathbf{N}}_a^t = R_b + \hat{\mathbf{T}}_b\tilde{\mathbf{N}}_b^t$$

and $\tilde{\mathbf{N}}_a \cdot \mathbf{X} > 0$, $\tilde{\mathbf{N}}_b \mathbf{X} > 0$ hold for all the image vectors $\mathbf{X}$, then there exist two planes with $\tilde{\mathbf{N}}_a$ and $\tilde{\mathbf{N}}_b$ as normals, respectively, and there exist points in the plane with positive z components at time t_1 such that if they undergo the two motions corresponding to R_a, $\hat{\mathbf{T}}_a$ and R_b, $\hat{\mathbf{T}}_b$, respectively, they produce the same images at the two time instants. □

Fortunately, if we have three image frames taken at three different time instants, in general we can uniquely determine the motion parameters and the relative normal of the object plane. Assume we have three image frames taken at time t_1, t_2 and t_3, respectively. Consider the two motions: one from t_2 to t_1 and the other from t_2 to t_3. The true solution sets for the two motions should have the same answer for the unit normal $\hat{\mathbf{N}}$ of the object plane at time t_2. By the following theorem, the false solutions generally do not have the same answer for the unit normal of the object plane.

Theorem 5. Let E and E' correspond to two motions and assume there exist rotation matrices R_a and R_b, unit vectors $\hat{\mathbf{N}}_a$ and $\hat{\mathbf{N}}_b$ such that

$$E = R_a + \tilde{\mathbf{T}}_a\hat{\mathbf{N}}_a^t = R_b + \tilde{\mathbf{T}}_b\hat{\mathbf{N}}_b^t$$

Let H be an orthonormal matrix with $\det(H)=1$ such that

$$H^tE^tEH = diag(\lambda_1, 1, \lambda_3)$$

with $\lambda_1 < 1 < \lambda_3$.

Then, the necessary and sufficient condition for the two solutions of the unit normal are the same for E and E':

$$E' = R_c' + \tilde{\mathbf{T}}_c'(\hat{\mathbf{N}}_a')^t = R_d' + \tilde{\mathbf{T}}_d'(\hat{\mathbf{N}}_b')^t$$

is that the ambiguity condition is satisfied, *i.e.*, there exist an orthonormal matrix Q and positive number k such that

$$E' = Q\ diag(\sqrt{1-k(1-\lambda_1)},\ 1,\ \sqrt{1+k(\lambda_3-1)})H^t$$

with $1-k(1-\lambda_1) \ge 0$ or

$$E' = Q\ diag(\sqrt{1+k(1-\lambda_1)},\ 1,\ \sqrt{1-k(\lambda_3-1)})H^t$$

with $1-k(\lambda_3-1) \ge 0$. □

For checking the condition that all the object points are coplanar, we have the following theorem.

Theorem 6. The necessary and sufficient condition for the rank of A in (Ag.1) to be less than 9 is that there exists a 3 by 3 nonzero matrix F such that all the points lie in the intersection of two quadratic surfaces before motion:

$$(\mathbf{x}-\mathbf{O}')\times F\mathbf{x}$$

where $\mathbf{O}' = -R^t\mathbf{T}$. □

In summary, we have presented close-form solutions of the problem. Given 4 or more point correspondences, the algorithm first solves for the intermediate parameters E. Then the motion parameters and the relative normal of the plane are obtained from E. The algorithm uses weighted least-squares to combat noise.

3. ERROR ESTIMATION

Formally, let the image coordinates of all the points be represented by I, and the errors in the image coordinates of these points be represented by a random variable ε. The error e in the estimated motion parameters is a function of I and ε. Denoting this function by f, informally we can write:

$$e = f(I, \varepsilon) \tag{3.1}$$

Our goal is to estimate the error e given the images I. However we don't know ε. If we can estimate the standard deviation of e (with ε as a random variable) given the noise-corrupted image I, we can use it to estimate the errors of the estimates. The images I corresponding to a degenerate or nearly degenerate spatial configuration should give large estimates of e and that corresponding to a stable configuration should give small estimates.

We assume the noises in the image coordinates have zero mean, known variance and are pairwise uncorrelated.

For the sake of conciseness, we use the following notation: The perturbation matrix of A is denoted by Δ_A. The noise-corrupted version of A is denoted by $A(\varepsilon)$. Thus we have

$$A(\varepsilon) = A + \Delta_A. \tag{3.2}$$

Similarly for vectors, we use δ with corresponding subscript to denote the noise vectors:

$$\mathbf{X}(\varepsilon) = \mathbf{X} + \boldsymbol{\delta}_{\mathbf{x}}. \tag{3.3}$$

Γ with the corresponding subscript is used to denote the auto covariance matrix of the noise vector (if only the first order errors are considered, the means of the errors are zero):

$$\Gamma_{\mathbf{x}} = \mathrm{E}(\delta_{\mathbf{x}}\delta_{\mathbf{x}}^t) \tag{3.4}$$

where E denotes expectation. A matrix $A = [\mathbf{A}_1\ \mathbf{A}_2\ \cdots\ \mathbf{A}_n]$ is associated with a corresponding vector $\mathbf{A}$ with

$$\mathbf{A} = (\mathbf{A}_1^t, \mathbf{A}_2^t, \cdots \mathbf{A}_n^t)^t \tag{3.5}$$

Similarly Γ_A denotes the corresponding covariance matrix of the vector $\mathbf{A}$ associated with matrix A. δ_A denotes the perturbation vector associated with the perturbation matrix Δ_A.

Assuming two variables a and b with small errors:

$$a(\varepsilon) = a + \delta_a, \quad b(\varepsilon) = b + \delta_b, \tag{3.6}$$

we have

$$a(\varepsilon)b(\varepsilon) = ab + \delta_a b + a\delta_b + \delta_a\delta_b \triangleq ab + \delta_{ab}. \tag{3.7}$$

The errors in $a(\varepsilon)b(\varepsilon)$ is

$$\delta_{ab} = \delta_a b + a\delta_b + \delta_a\delta_b \cong \delta_a b + a\delta_b. \tag{3.8}$$

In the last approximation we keep the linear terms (first order perturbation) of the error and ignore the higher order terms. Later in this paper we use the sign $\cong$ for the equations that are equal in the linear terms ($\approx$ for the approximate equality in the usual sense).

The algorithm presented involves computation of the eigenvectors of a symmetrical matrix. With small perturbation in the matrix, we need to known the corresponding perturbation in its eigenvalues and eigenvectors. We have the following theorem.

Theorem 7. Let $A = [a_{ij}]$ be an n by n symmetrical matrix and H be an orthonormal matrix such that

$$H^{-1}AH = diag(\lambda_1, \lambda_2, \cdots, \lambda_n) \tag{3.9}$$

Let the eigenvalues be ordered according to increasing magnitudes. Without loss of generality, consider the eigenvalue λ_1. Assuming λ_1 is a simple eigenvalue.

$$\lambda_1 < \lambda_2 \le \lambda_3 \le \cdots \le \lambda_n . \tag{3.10}$$

Let

$$H = [\mathbf{h}_1\ \mathbf{h}_2 \cdots \mathbf{h}_n] \tag{3.11}$$

and $\mathbf{X}$ be an eigenvector of A associated with λ_1. $\mathbf{X}$ is then a vector in $span(\mathbf{h}_1)$ (the linear space spanned by $\mathbf{h}_1$). Let $\mathbf{X}(\varepsilon)$ be the eigenvector of the perturbed matrix $A(\varepsilon)=A+\Delta_A$ associated with the perturbed eigenvalue $\lambda_1(\varepsilon)$. $\mathbf{X}(\varepsilon)$ can be written as

$$\mathbf{X}(\varepsilon) = \mathbf{X}+\boldsymbol{\delta}_{\mathbf{x}} \tag{3.12}$$

with $\boldsymbol{\delta}_{\mathbf{x}} \in span(\mathbf{h}_2, \mathbf{h}_3, \cdots, \mathbf{h}_n)$. Letting ε be the maximum absolute value of the entries in $\Delta_A=[\delta_{a_{ij}}]$, we have

$$\Delta_A = \varepsilon B \tag{3.13}$$

where $B=[b_{ij}]$, with $b_{ij}=\delta_{a_{ij}}/\varepsilon$. Therefore $|b_{ij}| \le 1$, $1 \le i \le n, 1 \le j \le n$. Then for sufficiently small ε, the perturbation of λ_1 can be expressed by a convergent series in ε:

$$\delta_{\lambda_1} \triangleq \lambda_1(\varepsilon) - \lambda_1 = p_1\varepsilon + p_2\varepsilon^2 + p_3\varepsilon^3 + \cdots \tag{3.14}$$

and the perturbation vector $\boldsymbol{\delta}_{\mathbf{x}}$ can be expressed by a convergent vector series in the space $span(\mathbf{h}_2, \mathbf{h}_3, \cdots, \mathbf{h}_n)$. In other words, letting $H_2=[\mathbf{h}_2, \mathbf{h}_3, \cdots, \mathbf{h}_n]$, then for sufficiently small positive ε, there exist (n-1)-dimensional vectors $\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3, \cdots$ such that

$$\delta_{\mathbf{x}} = \varepsilon H_2\mathbf{g}_1 + \varepsilon^2 H_2\mathbf{g}_2 + \varepsilon^3 H_2\mathbf{g}_3 + \cdots \tag{3.15}$$

The liner term (in ε) in (3.14) is given by

$$p_1\varepsilon = \mathbf{h}_1^t \Delta_A \mathbf{h}_1 \tag{3.16}$$

The linear term (in ε) in (3.15) is given by

$$\varepsilon H_2\mathbf{g}_1 = H\Delta H^t \Delta_A \mathbf{X} \tag{3.17}$$

where

$$\Delta = diag(0, (\lambda_1-\lambda_2)^{-1}, \cdots, (\lambda_1-\lambda_n)^{-1}) \tag{3.18}$$

That is, suppressing the second and higher order terms (*i.e.*, considering first order perturbation), for the eigenvalue we have

$$\delta_{\lambda_1} \cong \mathbf{h}_1^t \Delta_A \mathbf{h}_1 \tag{3.19}$$

and for the eigenvector:

$$\boldsymbol{\delta}_{\mathbf{x}} \cong H\Delta H^t \Delta_A \mathbf{X} \tag{3.20}$$

□

A similar result holds for other simple eigenvalues and eigenvectors.

From the theorem if the perturbation matrix Δ_A can be estimated, the corresponding perturbation in the eigenvalue and the eigenvectors of A can be estimated (by first order perturbation). The steps (i), and (ii) in the algorithm are to find the eigenvalues and the eigenvectors of the corresponding matrices. The problem now is to estimate the perturbation of the corresponding matrices from the perturbation in the image coordinates. Again we use the first order approximation to estimate these perturbations in the matrices. We will not include detailed derivations here because of space limitation. Based on (3.8) and the algorithm, it is not difficult to derive the first order perturbation of A^tA: Δ_{A^tA}. From Theorem 7, we have the first order perturbation of $\mathbf{h}$:

$$\boldsymbol{\delta}_{\mathbf{h}} \cong H\Delta H^t \Delta_{A^tA}\mathbf{h}$$

$$= H\Delta H^t [h_1 I_9\ h_2 I_9\ \cdots\ h_9 I_9]\boldsymbol{\delta}_{A^tA} \triangleq G_{\mathbf{h}}\boldsymbol{\delta}_{A^tA} . \tag{3.28}$$

We can relate $\boldsymbol{\delta}_{A^tA}$ in (3.28) to $\boldsymbol{\delta}_{A^t}$. Finally we get the perturbation vector of E, $\boldsymbol{\delta}_E$. For the perturbation vectors of R, $\hat{\mathbf{T}}$ and $\mathbf{N}$ we will get the linear expression in terms of $\boldsymbol{\delta}_E$. The corresponding covariance matrix then is obtained. For example, if we get $D_{\hat{T}}$ such that $\boldsymbol{\delta}_{\hat{T}} \cong D_{\hat{T}}\boldsymbol{\delta}_E$, we have then the covariance matrix $\Gamma_{\hat{T}} \cong D_{\hat{T}}\Gamma_E D_{\hat{T}}^t$. Similarly, the covariance matrix of $\mathbf{q}$ then R, $\mathbf{N}$ are derived using the results of Theorem 7.

We can estimate the Euclidean norm of the perturbation vector and the perturbation matrix by

$$\| \Delta_R \| = \| \boldsymbol{\delta}_R \| \approx \sqrt{trace(\Gamma_R)}$$

$$\| \boldsymbol{\delta}_{\hat{T}} \| \approx \sqrt{trace(\Gamma_{\hat{T}})}.$$

Similarly we can estimate perturbations of relative normal $\tilde{\mathbf{N}}$.

4. SIMULATIONS

In the simulation the object feature points are generated randomly. The image coordinates of the points are quantized according to the resolution of the camera. These quantization errors result in the errors in the motion parameters and the relative depths calculated by the algorithm. Other additional random errors such as corner detection errors can be simulated by a reduced image resolution. All the errors of the estimated R, $\hat{\mathbf{T}}$ and $\tilde{\mathbf{N}}$ shown in the figures are the relative errors defined by the Euclidean norm of the error vector (or matrix) divided by the Euclidean norm of the original vector (or matrix). Since no confusion may arise, we may call relative errors simply errors.

Different motion parameters with different image resolutions are simulated. The error is reduced roughly by a factor of two when the image resolution is doubled. Fig.2 shows the results of a typical sequence of trials with 5 point correspondences. The image resolution is 512 by 512. The results in here use an object normal (0, 0, 1) at time t_1, which is a typical unstable case (not discussed here), to show the performance of the motion estimation algorithm. Rotation is about axis (1, 1, 1) by 2.86°. The translation is (−0.176, 0.176, −1.995). In Fig.2, 20 random trials (randomly generated sets of points on the plane) are shown in the order of their generation. As can be seen from the figure, the estimated errors (dashed lines) are strongly correlated with the actual errors (solid lines). The estimated errors are especially important to detect a relatively unreliable configuration (trial No. 7 in Fig. 2(a)).

The average performance of the error estimation as well as that of the motion estimation algorithm is presented in Fig. 3. The solid lines show the relative errors observed over 20 random trials (randomly generated sets of points, with the same motion). The dot-dashed lines indicate the mean deviation (average of the absolute difference between the estimated error and the actual error). The dashed lines give the bias (difference between the mean of the estimated errors and the mean of the actual errors) over these 20 trials.

5. CONCLUSIONS

A new algorithm has been presented which gives a closed-form solution for the motion parameters and the relative normal of the plane. It uses least-squares techniques to give a stable solution in the presence of noise. For more robust estimation, a two step approach proposed by Weng, Huang and Ahuja [Weng88] can be used. Namely, the solution of this algorithm can be used as an initial guess for the iterative method that improves the solution based on optimal estimation. All the cases of the solution have been investigated and uniqueness has been proved. An approach is presented that test whether the points are coplanar. The necessary and sufficient condition for such a test to fail has been derived.

The error estimation algorithm is based on first order perturbation. The simulation results show a strong correlation between the estimated and the actual errors.

ACKNOWLEDGEMENTS

This research was supported by the National Science Foundation under grants ECS-83-52408 and IRI-86-05400. The authors would like to thank X. Zhuang for some illuminating discussions.

REFERENCES

[Tsai83] R. Y. Tsai and T. S. Huang, Estimating three-dimensional motion parameters of a rigid planar patch, II: singular value decomposition, Acoustics, Speech, and Signal Process.,, vol. ASSP-30, No. 4, pp. 525-534, 1983.

[Weng88] Closed-form solution + maximum likelihood: a robust approach to motion and structure estimation, in Proc. *IEEE Conf. Computer Vision and Pattern Recognition*, Ann Arbor, Michigan, June 5-9, 1988, pp. 381-386.

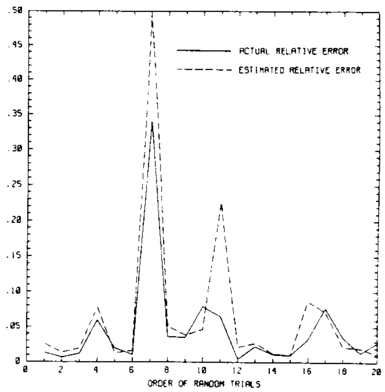

Fig. 2(a)

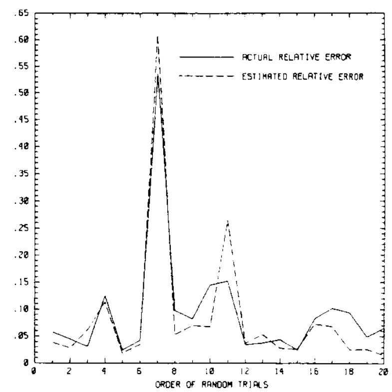

Fig. 2(b)

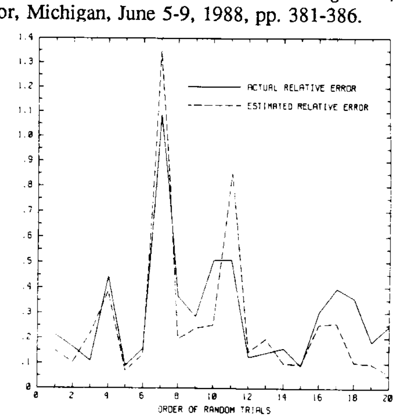

Fig. 2(c)

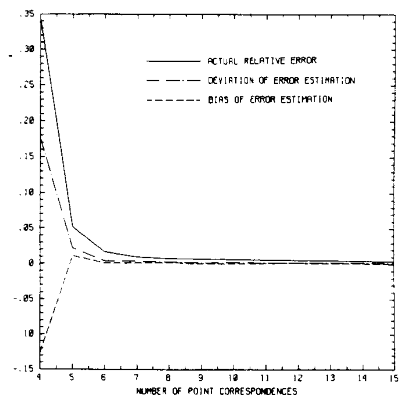

Fig. 3(a)

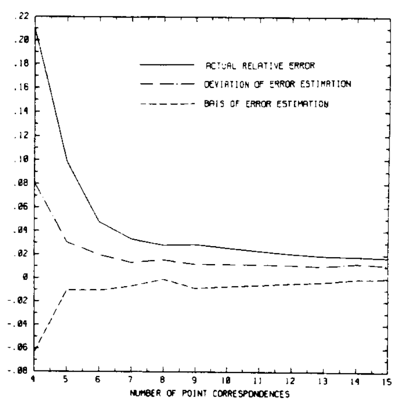

Fig. 3(b)

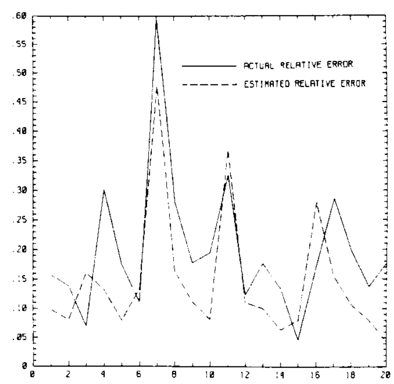

Fig. 2(d)

Fig.2. Actual Relative Error (solid lines) and Estimated Relative Error (dashed lines) of (a): E; (b): R; (c): $\mathbf{T}$ (d): $\mathbf{N}$. The horizontal index is the order of trials

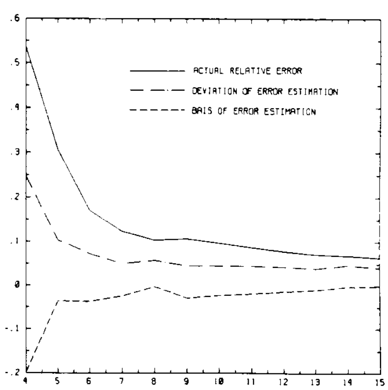

Fig. 3(c)

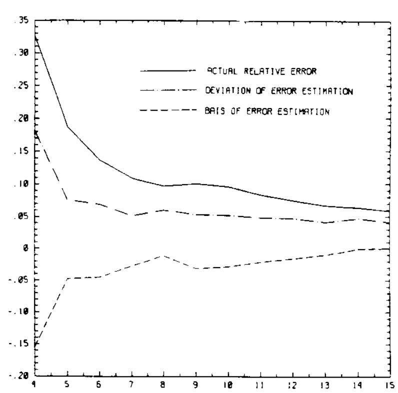

Fig. 3(d)

Fig.3. Mean Actual Relative Errors (solid lines), Deviation of Error Estimation (dot-dashed lines) and Bias of Error Estimation (dashed lines) for 20 trials versus Number of Point Correspondences. (a): E, (b): R, (c): $\mathbf{T}$, (d): $\mathbf{N}$