algorithm always converges to consistent unambiguous labelings [1], [5].

IV. CONCLUSION

In this correspondence, we analyzed the automata algorithm for relaxation labeling [1] for the case of symmetric compatibility functions. It is proved that starting with any initial label probabilities, the algorithm always converges to a consistent labeling. Further, all consistent unambiguous labelings are locally asymptotically stable. The algorithm analyzed in this correspondence has been employed successfully in computer vision problems such as stereopsis and object recognition [5].

REFERENCES

Problem 1: Given two sets of vectors \( \{ m_\alpha \} \) and \( \{ m'_\alpha \} \), \( \alpha = 1, \cdots, N \), compute a rotation \( R \) such that
\[
\sum_{\alpha=1}^{N} W_\alpha \| m_\alpha - Rm'_\alpha \|^2 \rightarrow \min.
\] (1)

Here, \( W_\alpha \) is a positive weight for the \( \alpha \)-th datum. The weights should be determined so that reliable data are given large weights whereas unreliable data are given small weights (we will discuss this later).

We use parentheses \((\cdot, \cdot)\) for vector inner product, superscript \( T \) for vector and matrix transpose, and \( \text{tr}(\cdot) \) for matrix trace. The left-hand side of (1) is expanded in the form
\[
\sum_{\alpha=1}^{N} W_\alpha \| m_\alpha \|^2 - 2\text{tr} \left( R^T \sum_{\alpha=1}^{N} W_\alpha m_\alpha m_\alpha^T \right) + \sum_{\alpha=1}^{N} W_\alpha \| m'_\alpha \|^2.
\] (2)

Hence, if the correlation matrix \( K \) between \( \{ m_\alpha \} \) and \( \{ m'_\alpha \} \), \( \alpha = 1, \cdots, N \), is defined by
\[
K_{\alpha\beta} = \sum_{n=1}^{N} W_n \frac{m_\alpha m_\beta^T}{\| m_\alpha \|^2},
\]
Problem 1 is restated as follows:

Problem 2: Given a correlation matrix \( K \), compute a rotation \( R \) such that
\[
\text{tr} (R^T K) \rightarrow \max.
\] (4)

The following two lemmas are fundamental:

**Lemma 1:** If \( S \) is a semipositive definite symmetric matrix, \( \text{tr}(RS) \) is maximized over all rotations by \( R = I \). The solution is unique if \( \text{rank} S > 1 \).

**Proof:** If \( S \) is semipositive definite symmetric, it has nonnegative eigenvalues \( \sigma_1 \geq \sigma_2 \geq \sigma_3 \geq 0 \), and the corresponding eigenvectors \( u_1, u_2, \) and \( u_3 \) can be chosen to be mutually orthogonal unit vectors. Hence, \( S \) is expressed in the following form (the spectral decomposition [9]):
\[
S = \sum_{i=1}^{3} \sigma_i u_i u_i^T.
\] (5)

Then,
\[
\text{tr}(RS) = \sum_{i=1}^{3} \sigma_i \text{tr}(Ru_i u_i^T) = \sum_{i=1}^{3} \sigma_i \text{tr}(Ru_i u_i).
\] (6)

By the Schwartz inequality \( (Ru_i u_i) \leq \|Ru_i\| \cdot \|u_i\| = 1 \), we see that
\[
\text{tr}(RS) \leq \sum_{i=1}^{3} \sigma_i = \text{tr} S.
\] (7)

If \( S \) is nonsingular (i.e., \( \text{rank} S = 3 \)), all eigenvalues are positive, so the equality holds if and only if \( Ru_i = u_i, i = 1, 2, 3 \). Since \( \{ u_i \} \) is an orthonormal system, this is true if and only if \( R = I \). If \( \text{rank} S = 2 \), then \( \sigma_1 \geq \sigma_2 > 0 \) and \( \sigma_3 = 0 \). Hence, the equality holds if and only if \( Ru_i = u_i, i = 1, 2 \). Since \( \{ u_i \} \) is an orthonormal system and \( R \) is a rotation, \( Ru_3 \) is necessarily \( u_3 \). Thus, \( R = I \). If \( \text{rank} S = 1 \), then \( \sigma_1 > 0 \) and \( \sigma_2 = \sigma_3 = 0 \), so the equality holds as long as \( Ru_1 = u_1 \). In other words, any rotation around \( u_1 \) gives a solution. If \( \text{rank} S = 0 \), then \( S = 0 \); any rotation is a solution.

The above proof is the same as that given by Arun et al. [1]. The following lemma is equivalent to the result of Umeyama [13].

**Lemma 2:** If \( S \) is a semipositive definite symmetric matrix, \( \text{tr}(RS) \) is maximized over all improper rotations for
\[
R' = I - 2u_i u_i^T,
\] (8)
where \( u_i \) is the unit eigenvector of \( S \) for the smallest eigenvalue. The solution is unique if \( \text{rank} S > 1 \) and the smallest eigenvalue of \( S \) is a simple root.

**Proof:** As in Lemma 1, let \( \sigma_1 \geq \sigma_2 \geq \sigma_3 \) be the eigenvalues of \( S \), and \( \{ u_1, u_2, u_3 \} \) an orthonormal system of the corresponding eigenvectors. If \( \sigma_3 = 0 \), the proof of Lemma 1 shows that \( \text{tr}(RS) \) is maximized by a (proper or improper) rotation \( R' \) such that \( R'u_1 = u_1 \) and \( R'u_2 = u_2 \). Since \( R' \) is improper, we automatically have \( R' u_3 = -u_3 \). This means that \( R' \) is diagonal with diagonal elements \( \{ 1, 1, -1 \} \) with respect to basis \( \{ u_1, u_2, u_3 \} \). Since \( u_1 u_1^T + u_2 u_2^T + u_3 u_3^T = I \) (the unit matrix), we obtain
\[
R' = u_1 u_1^T + u_2 u_2^T - u_3 u_3^T = I - 2u_3 u_3^T.
\]
So, assume that \( S \) is nonsingular. If \( R' \) attains the maximum of \( \text{tr}(RS) \), superimposition on \( R' \) of an infinitesimally small rotation in the form of \( (1+\epsilon W + O(\epsilon^2))R' \) yields zero perturbation to a first approximation in \( \epsilon \), where \( W \) is an arbitrary antisymmetric matrix.

Since
\[
\text{tr}(1+\epsilon W + O(\epsilon^2))R'S) = \text{tr}(R'S) + \epsilon \text{tr}(WR'S) + O(\epsilon^2),
\]
the term \( \text{tr}(W(R'S)) \) must vanish for any antisymmetric matrix \( W \). This occurs if and only if \( R'S \) is a symmetric matrix, namely \( R'S = (R'S)^T \). Hence,
\[
\text{tr}(R'S) = \text{tr}(R'(RS)) = \text{tr}(R'SR') = \text{tr}(R'^{-1}R'S) = \text{tr} S.
\]
(11)

Since \( S \) is nonsingular and \( R' \) is proper, Lemma 1 implies that this holds if and only if \( R'^{2} = I \), i.e., \( R' = R^{-1} \). Hence, \( R'S = SR'^{-1} = SR' \), i.e., \( R' \) *commutes* with \( S \). This means that \( R' \) and \( S \) are diagonalized at the same time; the orthonormal system \( \{ u_i \} \) of eigenvectors of \( S \) can be chosen to be the eigenvectors of \( R' \) as well. Since \( R' \) is orthogonal, its real eigenvalues are \( \pm 1 \). Hence, \( R'u_i = \pm u_i, i = 1, 2, 3 \), but \( R' \) is improper, so the three eigenvalues cannot be all \( 1 \). Since \( \sigma_1 \geq \sigma_2 \geq \sigma_3 > 0 \), (6) implies that the maximum of \( \text{tr}(RS) \) is attained when \( (R'u_i, u_i) = (R'u_2, u_1) = 1 \) and \( (R'u_3, u_1) = -1 \). Hence, \( R'u_1 = u_1, R'u_2 = u_2, \) and \( R'u_3 = -u_3 \). Thus, \( R' \) is diagonal with diagonal elements \( \{ 1, 1, -1 \} \) with respect to basis \( \{ u_1, u_2, u_3 \} \) and hence is given by (9). The uniqueness condition is obtained in the say way as in Lemma 1.

The first procedure for solving Problem 2 is to decompose the correlation matrix \( K \) into the form
\[
K = VAU^T,
\]
where \( V \) and \( U \) are orthogonal matrixes. This decomposition is called the singular value decomposition, and \( \sigma_1, \sigma_2, \) and \( \sigma_3 \) the singular values. The number of nonzero singular values is the rank of \( K \), which is equal to the number of linearly independent columns (or rows) of \( K \). The following theorem is mathematically equivalent to Umeyama's extension [13] of the method of Arun et al. [1].

**Theorem 1:** If \( K = VAU^T \) is the singular value decomposition, \( \text{tr}(R^T K) \) is maximized over all rotations by
\[
R = V \left( \begin{array}{c}
1 \\
\begin{array}{cc}
1 & \det(VU^T) \end{array}
\end{array} \right) U^T.
\]
(13)
The solution is unique if \( \text{rank} K > 1 \) and \( \det(VU^T) = 1 \), or if \( \text{rank} K > 1 \) and the minimum singular value is a simple root.
Proof: If \( K = V A U^T \), we have
\[ \text{tr}(R^T K) = \text{tr}(R^T V A U^T) = \text{tr}((U^T R^T V) A). \]  
(14)
Since \( V \) and \( U^T \) are orthogonal matrices, \( \det(VU^T) \) is either 1 or -1. If it is 1, the matrix \( U^T R^T V \) ranges over all proper rotations as \( R \) ranges over all rotations. Since \( A \) is a semipositive definite symmetric matrix, Lemma 1 implies that (14) attains its maximum when \( U^T R^T V = I \) or \( R = VU^T \). If \( \det(VU^T) = -1 \), the matrix \( U^T R^T V \) ranges over all improper rotations as \( R \) ranges over all rotations. By Lemma 2, the maximum is attained when \( U^T R^T V = A' \) or \( R = V A' U^T \), where \( A' \) is the diagonal matrix with diagonal elements \( \{1, 1, -1\} \) in this order. The uniqueness condition also follows from Lemmas 1 and 2.

The second procedure for solving Problem 2 is to decompose the correlation matrix \( K \) into the form
\[ K = VS = S'V, \]  
(15)
where \( V \) is an orthogonal matrix, while \( S \) and \( S' \) are semipositive definite symmetric matrices. This decomposition is known as the polar decomposition. The following theorem is an extension of the method of Horn et al. [3].

Theorem 2: If \( K = VS = S'V \) is the polar decomposition, \( \text{tr}(R^T K) \) is maximized over all rotations by
\[ R = V(I+(\det V-1)u_m v_m^T) = (I+(\det V-1)v_m v_m^T)V, \]  
(16)
where \( u_m \) and \( v_m \) are the unit eigenvectors of \( S \) and \( S' \), respectively, for the smallest eigenvalue. The solution is unique if \( \text{rank}(K) > 1 \) and \( \det V = 1 \). If \( \det V \neq 1 \), or \( \text{rank}(K) > 1 \) and the smallest eigenvalue of \( S \) (and of \( S' \)) is a simple root.

Proof: If \( K = VS \), we have
\[ \text{tr}(R^T K) = \text{tr}((R^T V)S), \]  
(17)
Since \( V \) is an orthogonal matrix, \( \det V = \pm 1 \). If \( \det V = 1 \), the matrix \( R^T V \) ranges over all proper rotations as \( R \) ranges over all rotations. By Lemma 1, (17) attains its maximum when \( R^T V = I \) or \( R = V \). If \( \det V = -1 \), the matrix \( R^T V \) ranges over all improper rotations as \( R \) ranges over all rotations. By Lemma 2, the maximum is attained when \( R^T V = I - 2u_m u_m^T \) or \( R = V(I-2u_m u_m^T) \). The same argument holds for \( K = S'V \). The uniqueness condition also follows from Lemmas 1 and 2.

The third method is based on the well-known fact that for any rotation matrix \( R \), there exist four numbers \( q_0, q_1, q_2, \) and \( q_3 \) such that \( q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1 \), and \( (q_0, q_1, q_2, q_3) \) is a four-dimensional unit vector. Hence, \( \text{tr}(R^T K) \) is maximized by the unit eigenvector \( \hat{q} \) of \( K \) for the largest eigenvalue.

Proof: Recall the relationship \( \text{tr}(R^T K) = \sum_{i=1}^N W_i(m_{0i} Rm_{0i}), \) for \( K = \sum_{i=1}^N W_i m_{0i} m_{0i}^T, \) from the quaternion representation (18), the \( i \)-th component of vector \( Rm_{0i} \) is written as
\[ (Rm_{0i})_i = m_{0i}q_0^2 + 2 \sum_{j=1}^3 \epsilon_{ijk} m_{0j} q_j / q_0 \]
\[ + 2q_0 \sum_{j=1}^3 m_{0j} q_j - m_{0i} \sum_{j=1}^3 q_j^2, \]  
(20)
where \( m_{0i} \) is the \( i \)-th component of vector \( m_{0i} \), and the symbol \( \epsilon_{ijk} \) denotes the Eddington epsilon, taking the value 1 if \( (ijk) \) is an even permutation of \( (123) \), -1 if \( (ijk) \) is an odd permutation of \( (123) \), and 0 otherwise. Then, it is easy to see that
\[ \sum_{i=1}^N W_i (m_{0i}, Rm_{0i} ) = q_0^2 \sum_{i=1}^3 K_{ii} + 2q_0 \sum_{i=1}^3 \left( \sum_{j=1}^3 \epsilon_{ijk} K_{kj} \right) \]
\[ + 3 \sum_{i,j=1}^3 q_i q_j K_{ij} - \delta_{123} \sum_{k=1}^3 K_{kk}, \]  
(21)
If we define a four-dimensional symmetric matrix \( \hat{K} \) by (19), (20) is rewritten in the form \( \text{tr}(\hat{R}^T \hat{K}) = \langle \hat{q}, K\hat{q} \rangle \), where \( \hat{q} = (q_0, q_1, q_2, q_3)^T \) is a four-dimensional unit vector. Hence, \( \text{tr}(R^T K) \) is maximized by the unit eigenvector \( \hat{q} \) of \( K \) for the largest eigenvalue.

III. ORTHOGONALITY RECONSTRUCTION FROM PROJECTION

The orthogonal projection of vector \( a \) onto the plane with unit surface normal \( h \) is given by \( P_h a \), where
\[ P_h = I - hh^T. \]  
(22)
(See. Fig. 1(a).) Let us consider the following problem (Fig. 1(b)): Problem 3: Given three coplanar vectors \( \{t_1, t_2, t_3\} \), compute a right-handed orthonormal system \( \{r_1, r_2, r_3\} \) such that
\[ t_i = P_h r_i, \]  
(23)
where \( h \) is the unit normal to the surface on which \( \{t_i\} \) lie.

If we define matrices \( R = (r_1, r_2, r_3) \) and \( T = (t_1, t_2, t_3) \), the requirement \( t_i = P_h r_i, i = 1, 2, 3 \), is equivalent to \( T = P_h R \). Let us call a matrix \( T \) a degenerate rotation if there exists a rotation \( R \) and a unit vector \( h \) such that \( T = P_h R \). We call the unit vector \( h \) the axis of \( T \).

Theorem 3: Given correlation matrix \( K \), define a four-dimensional symmetric matrix as shown in (19) at the bottom of the page. Let \( \hat{q} \) be the four-dimensional unit eigenvector of \( K \) for the largest eigenvalue. Then, \( \text{tr}(R^T K) \) is maximized by the rotation represented by \( \hat{q} \). The solution is unique if the largest eigenvalue of \( K \) is a simple root.

Proof: Recall the relationship \( \text{tr}(R^T K) = \sum_{i=1}^N W_i (m_{0i}, Rm_{0i} ) \) for \( K = \sum_{i=1}^N W_i m_{0i} m_{0i}^T, \) From the quaternion representation (18), the \( i \)-th component of vector \( Rm_{0i} \) is written as
\[ (Rm_{0i})_i = m_{0i}q_0^2 + 2 \sum_{j=1}^3 \epsilon_{ijk} m_{0j} q_j / q_0 \]
\[ + 2q_0 \sum_{j=1}^3 m_{0j} q_j - m_{0i} \sum_{j=1}^3 q_j^2, \]  
(20)
where \( m_{0i} \) is the \( i \)-th component of vector \( m_{0i} \), and the symbol \( \epsilon_{ijk} \) denotes the Eddington epsilon, taking the value 1 if \( (ijk) \) is an even permutation of \( (123) \), -1 if \( (ijk) \) is an odd permutation of \( (123) \), and 0 otherwise. Then, it is easy to see that
\[ \sum_{i=1}^N W_i (m_{0i}, Rm_{0i} ) = q_0^2 \sum_{i=1}^3 K_{ii} + 2q_0 \sum_{i=1}^3 \left( \sum_{j=1}^3 \epsilon_{ijk} K_{kj} \right) \]
\[ + 3 \sum_{i,j=1}^3 q_i q_j K_{ij} - \delta_{123} \sum_{k=1}^3 K_{kk}, \]  
(21)
If we define a four-dimensional symmetric matrix \( \hat{K} \) by (19), (20) is rewritten in the form \( \text{tr}(\hat{R}^T \hat{K}) = \langle \hat{q}, K\hat{q} \rangle \), where \( \hat{q} = (q_0, q_1, q_2, q_3)^T \) is a four-dimensional unit vector. Hence, \( \text{tr}(R^T K) \) is maximized by the unit eigenvector \( \hat{q} \) of \( K \) for the largest eigenvalue.
The singular values of $-T$ are also the singular values of $T$, which are 1, 1, and 0. Hence, $-T$ is also a degenerate rotation. Changing the sign of $T = (t_1, t_2, t_3)$ means changing the sign of each $t_i$. Since $\{t_i\}$ are coplanar, changing their signs means rotating them by angle $\pi$ in the plane on which they lie. Hence, the vectors $r_i$ to be reconstructed are also rotated by angle $\pi$ around the unit normal $h$ to the plane. It is easy to see that $I_h$ is a half-rotation about $h$, and the resulting rotation is $(I_h r_1, I_h r_2, I_h r_3) = I_h (r_1, r_2, r_3)$.

Let us call a pair $\{R, I_h R\}$ of rotations a twisted pair. It is easy to test if a given matrix $T$ is a degenerate rotation because of Proposition 1.

**Proposition 1:** A matrix $T$ is a degenerate rotation if and only if

$$\det T = 0, ||T|| = ||TT^T|| = \sqrt{2}. \quad (28)$$

**Proof:** Let $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$ be the eigenvalues of symmetric matrix $TT^T$. Then,

$$\det(TT^T) = \lambda_1 \lambda_2 \lambda_3, \quad \text{tr}(TT^T) = \lambda_1 + \lambda_2 + \lambda_3,$$

$$\text{tr}((TT^T)^2) = \lambda_1^2 + \lambda_2^2 + \lambda_3^2. \quad (29)$$

If $T$ has singular value 1, 1, and 0, then $\lambda_1 = \lambda_2 = 1^2$ and $\lambda_3 = 0^2$. Hence,

$$\det T = \sqrt{\det(TT^T)} = 0, ||T|| = \sqrt{\text{tr}(TT^T)} = \sqrt{2}. \quad (30)$$

Conversely, suppose (28) holds. Since $\det(TT^T)^2 = \det(TT^T) = 0$, it follow from (29) that $\lambda_3 = 0$, and the remaining eigenvalues of $TT^T$ satisfy

$$\lambda_1 + \lambda_2 = \lambda_1^2 + \lambda_2^2 = 2. \quad (31)$$

This means that $\lambda_1 = \lambda_2 = 1$. Hence, matrix $T$ has singular values $\sqrt{1}, \sqrt{1},$ and $\sqrt{0}$. $\Box$.

In order to obtain a robust method of solving Problem 3, we replace it by

**Problem 4:** Given three arbitrary vectors $\{t_1, t_2, t_3\}$, find an axis $h$ such that

$$\sum_{i=1}^{3} (t_i, h)^2 \rightarrow \min, \quad (32)$$

and then find a right-handed orthonormal system $\{r_1, r_2, r_3\}$ such that

$$\sum_{i=1}^{3} ||P_h r_i - t_i||^2 \rightarrow \min. \quad (33)$$

Let us call the above procedure the *optimal resolution* of matrix $T = (t_1, t_2, t_3)$ into an axis $h$ and a rotation $R = (r_1, r_2, r_3)$. It appears that we must first compute the axis $h$ by (32) and then compute (33) from the computed axis $h$. However, these two steps can be carried out independently as follows: optimal-resolution($T$)

1) Compute the unit eigenvector $h$ of matrix $TT^T$ for the smallest eigenvalue.
2) Compute the rotation matrix determined by
\[
\text{tr}(R^T T) \rightarrow \max .
\]

*Derivation:* The left-hand side of (32) is written as
\[
3 \sum_{i=1}^{3} (t_i, h)^2 = (h, \sum_{i=1}^{3} t_i i) h = (h, T T^T h) \rightarrow \min .
\]

Hence, the axis \(h\) is given by the unit eigenvector (up to sign) of matrix \(TT^T\) for the smallest eigenvalue. Take unit eigenvectors \(v_1\) and \(v_2\) for the remaining eigenvalues so that \(\{v_1, v_2, h\}\) form an orthonormal system. Now,
\[
3 \sum_{i=1}^{3} ||P_h t_i - t_i||^2 = \text{tr}(P_h^TT) - 2\text{tr}(P_h R) + \text{tr}(T T^T).
\]

Since \(\text{tr}(P_h R) = \text{tr}(R^T P_h^T T) = \text{tr}(R^T P_h T)\), the minimization (33) is equivalent to finding a rotation \(R\) such that
\[
\text{tr}(R^T P_h T) \rightarrow \max .
\]

The singular value decomposition of \(T\) in vector form is
\[
T = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \sigma_3 u_3 v_3^T,
\]
where \(\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq 0\) are the singular values of \(T\), and \(\{u_1, u_2, u_3\}\) are an orthonormal system of the unit eigenvectors of \(T^T T\). Since \(v_1\) and \(v_2\) are both orthogonal to \(h\), we see that
\[
P_h T = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + 0 h u_3^T.
\]

This means that \(T\) and \(P_h T\) have the same singular value decomposition except for their singular values. The method of singular value decomposition is applied to (37), the singular values themselves do not affect the solution (in fact, the solution is \(R = v_1 u_1^T + v_2 u_2^T \pm h u_3^T\), where an appropriate sign is chosen so that \(\det R = 1\) (cf. Theorem 1). Hence, (37) can be replaced by (34). \(\square\)

IV. 3-D MOTION ESTIMATION

Suppose the camera is rotated by \(R\) around the center of the lens and translated by \(h\). If two images of the same scene before and after the motion are given and point-to-point correspondences of multiple feature points are detected over the two images, we can compute the motion parameters \(\{R, h\}\) and the depth of each of the feature points up to scale [7], [10], [12], [14]. In order to remove the scale indeterminacy, it is customary to scale the translation \(h\) to a unit vector. The core of the problem is summarized as follows:

Let \(r_1, r_2, \) and \(r_3\) be the first, second, and third columns of \(R\), respectively. Define the matrix
\[
G = (h \times r_1, h \times r_2, h \times r_3).
\]

We abbreviate this matrix as \(h \times R\). This matrix is called the essential matrix and is directly determined from at least eight point-to-point correspondences over the two images. Hence, the problem reduces to the following form:

**Problem 5:** For a given matrix \(G\), compute a unit vector \(h\) and a rotation \(R\) such that
\[
G = h \times R.
\]

Let \(J_h\) be a quarter-rotation (rotation by angle \(\pi/2\)) about unit vector \(h\). We observe the following fact:

**Proposition 2:**
\[
G = P_h(J_h R).
\]

V. STATISTICS OF ROTATION FITTING

Let \(R = (r_1, r_2, r_3)\) be a rotation matrix. The three columns \(\{r_1, r_2, r_3\}\) form an orthonormal system. If \(R\) is computed from image data, it may be perturbed into \(R' = (r_1', r_2', r_3')\). However, the three columns \(\{r_1', r_2', r_3'\}\) form an orthonormal system, so the error in each element cannot be independent. Since the transformation from \(R\) to \(R'\) is a rotation, the error is also a rotation of some angle \(\Delta\theta\)
around some axis $l$. As is well known (e.g., [5]), if we define $\Delta l = \Delta l_l$, we have

$$r'_i = r_i + \Delta l \times r_i + O(\Delta l)^2, \quad i = 1, 2, 3,$$

(48)

where $O(\Delta l)^2$ denotes a term of order 2 or higher in the components of $\Delta l$. In matrix form,

$$R' = R + \Delta l \times R + O(\Delta l)^2.$$  

(49)

Regarding $\Delta l$ as a vector random variable, we define the covariance matrix $V[R]$ by

$$V[R] = E[\Delta l \Delta l']$$

(50)

where $E[\cdot]$ denotes expectation. The eigenvector of $V[R]$ for the largest eigenvalue indicates the axis of the error rotation that is most likely to occur, and the corresponding eigenvalue indicates the mean square of the angle of rotation around it.

Consider the problem of fitting an orthonormal frame $\{r_1, r_2, r_3\}$ to three unit vectors $\{m_1, m_2, m_3\}$ by the least-squares criterion

$$J = \sum_{i=1}^{3} W_i \|r_i - m_i\|^2 \rightarrow \text{min}.$$  

(51)

Here, $W_i$ is a weight for the $i$th datum. If we define matrix $R = \begin{pmatrix} r_1 & r_2 & r_3 \end{pmatrix}$, the condition that $\{r_i\}$, $i = 1, 2, 3$, be a right-handed orthonormal system is equivalent to $R$ being a rotation matrix. Also define matrix $M = \begin{pmatrix} W_1 m_1 & W_2 m_2 & W_3 m_3 \end{pmatrix}$, which can be viewed as the correlation matrix $\sum_{i=1}^3 W_i m_i e_i$, where $e_i$ is the $i$th coordinate basis vector. Then, it is easy to see that (51) is rewritten as

$$\text{tr}(R^T M) \rightarrow \text{max}.$$  

(52)

Let us call the rotation fitted in this way the best fitting rotation. If $\{m_1, m_2, m_3\}$ is an orthonormal system from the beginning, we obviously obtain $r_i = m_i$, $i = 1, 2, 3$. In the presence of noise, each $m_i$ is perturbed by $\Delta m_i$ (see Fig. 2(b)). Regarding $\Delta m_i$ as a vector random variable, we define the covariance matrix $\{8\}$ of unit vector $m_i$ by

$$V[m_i] = E[\Delta m_i \Delta m_i'].$$  

(53)

The weight $W_i$ should be large for reliable data and small for unreliable data. A reasonable choice is to weigh the term $\|r_i - m_i\|$ by the inverse of the root-mean-square error $\sqrt{E[\|\Delta m_i\|^2]}$ of possible error $\Delta m_i$. This is equivalent to choosing

$$W_i = \frac{\text{constant}}{\text{tr}V[m_i]}.$$  

(54)

We assume that no two of $W_1$, $W_2$, and $W_3$ are simultaneously 0 and adjust the constant so that $\sum_{i=1}^3 W_i = 1$.

If each $m_i$ is perturbed by noise into $m'_i = m_i + \Delta m_i$ independently, the fitted vectors $r_i$ also change into $r'_i = r_i + \Delta r_i$, however, $\{m'_i\}$ are not necessarily an orthonormal system, and hence each $\Delta r_i$ is not necessarily equal to each $\Delta m_i$ (see Fig. 2(b)). Let $[a, b, c] = (a \times b, c) = (b \times c, a) = (c \times a, b)$ be the scalar triple product of vectors $a$, $b$, and $c$. The perturbation of $R$ is determined as follows:

**Lemma 3:** A perturbation $\Delta m_i$ of $m_i$, $i = 1, 2, 3$, causes the best fitting rotation $R$ to undergo a perturbation by

$$\Delta l = \sum_{i=1}^3 \left( \frac{\sum_{j=1}^3 W_j r_j r_j m_{ij}}{1 - W_i} \right) r_i.$$  

(55)

**Proof:** According to (51), $\Delta r_i$ are determined so that

$$J = \sum_{i=1}^3 W_i \|\Delta r_i - \Delta m_i\|^2 \rightarrow \text{min}.$$  

(56)

From (48), we can write $J$ in terms of $\Delta l$ as

$$J = \sum_{i=1}^3 W_i \left( \|\Delta l \times r_i\|^2 - 2\|\Delta l\| \|r_i\| + \|\Delta m_i\|^2 \right) \rightarrow \text{min}.$$  

(57)

If $\Delta l$ attains the minimum of $J$, an arbitrary perturbation $\Delta l = \Delta l + \delta \Delta l$ causes zero first variation of $J$ in $\delta \Delta l$. To a first approximation in $\delta \Delta l$,

$$\delta J = 2(\delta \Delta l) \sum_{i=1}^3 W_i (\Delta l - (r_i, \Delta l)r_i - r_i \times \Delta m_i),$$  

(58)

which must vanish for arbitrary $\delta \Delta l$. Hence,

$$\sum_{i=1}^3 W_i (\Delta l - (r_i, \Delta l)r_i) = \sum_{i=1}^3 W_i r_i \times \Delta m_i.$$  

(59)

If we define matrix

$$L = \sum_{i=1}^3 W_i (I - r_i r_i^T),$$  

(60)
the left-hand side of (59) is written as $L \Delta l$. Since $\sum_{j=1}^{3} r_{xj} r_{yj} = I$ and $\sum_{i=1}^{3} W_i = 1$, we can write

$$L = \sum_{i=1}^{3} W_i \sum_{j=1}^{3} r_{xj} r_{yj} - \sum_{i=1}^{3} W_i r_i = \sum_{i=1}^{3} (1 - W_i) r_i r_i^T. \quad (61)$$

Its inverse is given by

$$L^{-1} = \sum_{i=1}^{3} \frac{r_i r_i^T}{1 - W_i}. \quad (62)$$

Since no two of $\{W_i\}$ are simultaneously 0, we have $W_i \neq 1$. From (59), we obtain

$$\Delta l = \sum_{i=1}^{3} \frac{W_i r_i (r_i^T r_i \Delta m_i)}{1 - W_i}, \quad (63)$$

which is rewritten as (55).

From this, it is easy to see that the covariance matrix $V[R] = E[\Delta l \Delta l^T]$ is given as follows:

**Theorem 5:** If each $m_i$ is independent and has covariance matrix $V[m_i]$, the covariance matrix $V[R]$ of the best fitting rotation $R$ to $\{m_i\}$, $i = 1, 2, 3$, is given by

$$V[R] = \sum_{i=1}^{3} \sum_{j=1}^{3} W_i W_j r_i (r_i^T r_i) r_j r_j^T \quad (64)$$

**Example:** Fig. 3(a) is a real image (270 x 300 pixels) of a rectangular box, and Fig. 3(b) shows detected edges. The focal length is estimated to be $f = 1750$ (pixels). The unit vectors that point to the three vanishing points are estimated by least squares as follows:

$$m_1 = \begin{pmatrix} 0.244 \\ -0.792 \\ 0.599 \end{pmatrix}, \quad m_2 = \begin{pmatrix} 0.300 \\ 0.636 \\ 0.711 \end{pmatrix}, \quad m_3 = \begin{pmatrix} 0.030 \\ 0.045 \\ -0.914 \end{pmatrix}. \quad (65)$$

(The $X$-axis extends upward, the $Y$-axis rightward, and the $Z$-axis away from the viewer.) These indicate the three 3-D orientations of the edges [6]. If the image resolution $n$ is assumed to be unity, their covariance matrices can be evaluated theoretically as follows [8]:

$$V[m_1] = 10^{-6} \begin{pmatrix} 0.402 & -1.114 & 1.755 \\ -1.114 & 3.790 & -5.857 \\ 1.755 & -5.857 & 9.070 \end{pmatrix},$$

$$V[m_2] = 10^{-5} \begin{pmatrix} 0.887 & 1.681 & -1.878 \\ 1.681 & 3.505 & -3.845 \\ -1.878 & -3.845 & 4.231 \end{pmatrix},$$

$$V[m_3] = 10^{-5} \begin{pmatrix} 1.241 & 0.019 & 2.794 \\ 0.019 & 0.023 & 0.041 \\ 2.794 & 0.041 & 6.292 \end{pmatrix}. \quad (66)$$

The best fitting rotation matrix is

$$R = \begin{pmatrix} 0.239 & 0.320 & -0.917 \\ -0.789 & 0.626 & 0.015 \\ 0.578 & 0.712 & 0.399 \end{pmatrix}. \quad (67)$$

The discrepancies of $m_1$, $m_2$, and $m_3$ from the corresponding orientations are 1.35°, 1.25°, and 0.97°, respectively. By evaluating the covariance matrix $V[R]$ given by (64), we can see that the root-mean-square error $\Delta \Omega$ of the angle of error rotation (from the true frame, which we do not know) is 0.49°.

**VI. CONCLUDING REMARKS**

In this paper, we first recapitulated methods of fitting a 3-D rotation to 3-D data in a refined form as optimization over proper rotations, extending three existing methods—the method of singular value decomposition, the method of polar decomposition, and the method of quaternions representation. As an application of these three methods, we formulated the problem of optimal resolution of a degenerate rotation and showed how this solves the problem of 3-D motion estimation from two images in a succinct way. Finally, we defined the covariance matrix of rotation fitting and analyzed the statistical behavior of error of the fit in terms of the covariance matrices of the data.

REFERENCES