Motion Estimation of a Block-shaped Rigid Object for Robot Vision

Mutshiro TERAUCHI†, Koji ITÔ†, Takuo JOKO‡ and Toshio TSUJI†
†Faculty of Engineering, Hiroshima University
1-4-1 Kagamiyama, Higashi-Hiroshima, 724 Japan
‡Atsugi Development Center, Minolta Camera Corporation
231-1 Kamiochiai, Atsugi, 243 Japan

Abstract: We propose a new conventional method to reconstruct 3D motion of rigid polyhedral objects from a sequence of monocular images. In general the problem is ill-posed, therefore additional information is required to recover depth. In this paper we utilize line correspondence between sequential images and existence of parallel line segments in the scene. The relation between the coordinates of points can be described only by a rotation matrix, if it is formulated relatively. Then if there exists a pair of parallel line segments, the matrix can be solved by using linear equations. After that the translation vector is computed. It is necessary that there exists at least one pair of parallel line segments in the scene in order to obtain motion parameters. We also propose the method to extract pairs of parallel line segments in the image. Finally some experimental results for simulated data sets are demonstrated.

1. Introduction

The reconstruction of 3D motion or structure from a monocular image is ill-posed problem, because the information on depth would be reduced through the projection onto the image. It follows that it is necessary for the reconstruction to introduce some constraints about characteristics of motion or the structure of the objects. In the problem of "Motion from Image Sequence" there often assumed that the motion is sufficiently smooth (Marr, 1982; Kitahashi, 1989; Aggarwal and Nandhakumar, 1988).

Ullman studied motion estimation based on points correspondence between sequential images by using the assumption of rigidity of the objects. He showed it is possible to obtain motion and structure of an object from points correspondence of 4 points on 3 frames of the sequential images under the parallel projection, and 5 points on 3 frames under the perspective projection (Ullman, 1979). Tsai and Huang derived linear equations by using decomposition of the singular point of the matrix which contains medial parameters obtained from motion information [Tsai and Huang, 1984]. The equation consists of 8 variables and is linear, and it enables us to obtain solutions from 8 points correspondence on 2 frames under the perspective projection. They realized to get linear solution, while it had been analysed using nonlinear equations. It is, however impossible to solve the linear equations, if all the 8 points lie on 2 planes where one of two planes intersects the origin of 3D coordinates (at least 5 of 8 points satisfy the condition), or on the surface of the cone which intersects the origin (at least 6 of 8 points satisfy the condition).

On the other hand we proposed a method combining optical flow and correspondence of line segments in which we derived linear equations by using an assumption of existence of two pair of parallel line segments [Joko et al, 1990]. However it was limited to motion with small displacement between points on two different frames, because the velocity components were approximated using differential equations.

In this paper we propose a linear algorithm to estimate motion of the rigid object utilizing the relative expression of coordinates by assuming the existence of a pair parallel line segments. This method does not adopt any approximation so that it yields error-free results for all types of motion. It is also possible to examine if a chosen pair of line segments is parallel or not before estimating the motion parameters. The present method can be applied to a number of industrial applications, because it needs only a pair of parallel line segments, i.e. 4 points correspondence on 2 frames.

2. Formulation of Motion

In this paper we focus on the motion reconstruction from two sequentialized images onto which one rigid object in the scene is projected by the perspective projection. In this section we formulate the motion of one end-point of a line segment and the relative motion of another end-point. Motion of a line segment on the image is shown in Fig.1. We define \((x, y)\) as one end-point of a line segment and \((r, s)\) as the relative position of another one, and the positions after motion are defined as \((x', y')\) and \((r', s')\) respectively as well as ones before motion.

2.1 Motion of End Points of a Line Segment

We show the geometrical relationship between the motion of point \(P\) in the scene and the image plane in Fig.2. The origin of the world coordinate system \(OXYZ\) is set at the lens center of camera, and the \(Z\) axis of the coordinate system is placed along the optical axis. Then the 3D motion of point \(P\) can be generally represented by the rotational component \(\Theta\) and the translational
component $\Gamma$ as follows,
\[
\begin{bmatrix}
X' \\
Y' \\
Z'
\end{bmatrix} = \Phi \begin{bmatrix}
X \\
Y \\
Z
\end{bmatrix} + \Gamma ,
\]
\quad (1)
where $\Phi$ and $\Gamma$ are represented as follows,
\[
\Phi = \begin{bmatrix}
\phi_1 & \phi_2 & \phi_3 \\
\phi_4 & \phi_5 & \phi_6 \\
\phi_7 & \phi_8 & \phi_9
\end{bmatrix}, \quad \Gamma = \begin{bmatrix}
\Delta X \\
\Delta Y \\
\Delta Z
\end{bmatrix} ,
\]
\quad (2)
where the rotation matrix $\Phi$ can be described as follows,
\[
\Phi = \begin{bmatrix}
\sigma_1^2 + (1 - \sigma_2^2) \cos \theta & \sigma_1 \sigma_3 (1 - \cos \theta) - \sigma_2 \sin \theta & \sigma_1 \sigma_2 (1 - \cos \theta) + \sigma_3 \sin \theta \\
\sigma_1 \sigma_2 (1 - \cos \theta) + \sigma_3 \sin \theta & \sigma_2^2 + (1 - \sigma_1^2) \cos \theta & \sigma_2 \sigma_3 (1 - \cos \theta) - \sigma_1 \sin \theta \\
\sigma_1 \sigma_3 (1 - \cos \theta) - \sigma_2 \sin \theta & \sigma_2 \sigma_3 (1 - \cos \theta) + \sigma_1 \sin \theta & \sigma_3^2 + (1 - \sigma_2^2) \cos \theta
\end{bmatrix} ,
\]
\quad (3)
where $\theta$ and $(\sigma_1, \sigma_2, \sigma_3)$ denote the rotation angle and the rotation axis respectively. Then the unit cosine vector can be written as
\[
\sigma_1^2 + \sigma_2^2 + \sigma_3^2 = 1 .
\]
The relations between $\phi_i$ (i = 1 to 9) and their parameters $\theta, \sigma_1, \sigma_2, \sigma_3$ yields a nonlinear function. It is, however, possible to solve a pair of solution sets of $\theta, \sigma_1, \sigma_2, \sigma_3$ [Tsai and Huang, 1984]. The two solutions are dual and the difference of them is based on whether the parameters are described in $cw$ or $ccw$. Hence as the difference can not be discriminated physically, they can be assumed as the same. Here the relations between two points in the 3D world $P(X, Y, Z)$ and $P'(X', Y', Z')$, and their projection onto image plane $p(x, y)$ and $p'(x', y')$ are given as follows,
\[
x = \frac{X}{Z} , \quad y = \frac{Y}{Z} , \quad x' = \frac{X'}{Z'} , \quad y' = \frac{Y'}{Z'} .
\]
\quad (4)

2.2 Relative Motion between Feature Points of an Object

Here we introduce a new Cartesian coordinate system $\hat{O}O'\hat{Z}'$ whose origin $\hat{O}$ is set at the end-point of line segment $P$ as shown in Fig.3. Then another end-point can be represented as $M(m_1, m_2, m_3)$ in this coordinate system. The corresponding point after motion can be represented as $M'(m_1', m_2', m_3')$, where $M$ and $M'$ are the vectors along the line-segments. The relations between the vectors $M, M'$ and their projection onto the image plane $(r, s), (r', s')$ are given as follows,
\[
r = \frac{m_1 - m_3}{Z + m_3} x , \quad s = \frac{m_2 - m_3}{Z + m_3} y ,
\]
\[
r' = \frac{m_1 - m_3}{Z + m_3} x' , \quad s' = \frac{m_2 - m_3}{Z + m_3} y' .
\]
\quad (5)
Through the motion between $M$ and $M'$ is represented only by the rotation matrix $\Phi$ as follows,
$M' = \Phi M$.

Here we introduce parameters $t$ and $t'$ defined as follows,

$$t = \frac{m_1}{Z + m_3}, \quad t' = \frac{m_1'}{Z' + m_3},$$

we obtain the following equations,

$$M = (Z + m_3)N, \quad N = \begin{bmatrix} r + t \xi \cr s + t \eta \cr t \end{bmatrix},$$

$$M' = (Z' + m_3')N', \quad N' = \begin{bmatrix} r' + t' \xi' \cr s' + t' \eta' \cr t' \end{bmatrix}.$$  \hspace{1cm} (8)

For each line segment before and after motion we define unit vectors along the line segments as $I$ and $I'$ respectively,

$$I = \frac{N}{|N|}, \quad I' = \frac{N'}{|N'|}.$$  \hspace{1cm} (9)

Substituting Eq.(9) into Eq.(8), we get the following equations,

$$M = (Z + m_3)N \mid I,$$

$$M' = (Z' + m_3')N' \mid I'.$$  \hspace{1cm} (10)

$M$ and $M'$ has the same length, because they are relatively represented and denote the same line segment. As the $Z$ components of the coordinates of the end-points $Z+m_3$ and $Z'+m_3'$ are positive, we can define the following value $K$ from Eq.(10),

$$K = \frac{Z' + m_3'}{Z + m_3} = \frac{|N|}{|N'|}.$$ \hspace{1cm} (11)

Here it is possible to compute the value $K$, if $t$ and $t'$ defined in Eq.(7) could be obtained.

After Eq.(6) representing the rotation in the 3D space can be described by using the coordinates on the image as follows,

$$KN' = \Phi N.$$ \hspace{1cm} (12)

In this equation the unknown parameters are $\Phi, t$ and $t'$.

2.3 Relationship between Pre and Post-Motion Vector Products

The relative representation of motion of two line segments is shown in Fig.4. Here we define the relative vectors from one end-point to another one of the line segments as $M_1$ and $M_2$, and those after motion as $M'_1$ and $M'_2$. Then relationship between vector products of $M_1$ and $M_2$, and of $M'_1$ and $M'_2$ are given by using just the rotation matrix $\Phi$ as follows,

$$M'_1 \times M'_2 = \Phi (M_1 \times M_2).$$ \hspace{1cm} (13)

From Eq.(8) and Eq.(11) we get,

$$K_1 K_2 N'_1 \times N'_2 = \Phi (N_1 \times N_2).$$ \hspace{1cm} (14)

In this equation the unknown parameters are $\Phi, t_1, t'_1, t_2$ and $t'_2$.

3. Estimation of Motion Parameters

3.1 Relationship between Two Line Segments

When the correspondence between two line segments in the 3D space is found from an image, the parameters in Eq.(12) can be obtained, and so forth the parameters in Eq.(14) from the vector products of the line segments. In order to cover those relationships we introduce matrices $H$ and $H'$ defined below,

$$H = \begin{bmatrix} N_1, & N_2, & N_1 \times N_2 \end{bmatrix},$$

$$H' = \begin{bmatrix} K_1 N'_1, & K_2 N'_2, & K_1 K_2 \left(N'_1 \times N'_2\right) \end{bmatrix}.$$ \hspace{1cm} (15)

From Eq.(12) and Eq.(14) we get,

$$H' = \Phi H,$$ \hspace{1cm} (16)

where $\Phi$ is a unitary matrix, then we obtain the following equation,

$$H'^T H' = H^T \Phi^T \Phi H = H^T H.$$ \hspace{1cm} (17)

The above can be rewritten as the equations of components as follows,

$$|K_1 N'_1|^2 = |N_1|^2,$$

$$|K_2 N'_2|^2 = |N_2|^2,$$

$$K_1 K_2 N'_1 \times N'_2 = N_1 \times N_2.$$ \hspace{1cm} (18) \hspace{1cm} (19) \hspace{1cm} (20)

3.2 Utilization of Parallelism of Line Segments

Eq.(16) states that if the relative lengths $t$ and $t'$ could be computed, it is possible to obtain the rotation matrix $\Phi$ by using the relations of two line segments. Therefore we assume there exists a pair of parallel line segments on a rigid object in 3D space. This assumption is obviously the constraint for
obtaining a unique solution for the parameters \( t \) and \( t' \).

When two line segments \( l_1 \) and \( l_2 \) are parallel, then the directions of two line segments are the same, hence we get the following relation from Eq.(8),

\[
\begin{bmatrix}
  r_1 + t_1 x_1 \\
  s_1 + t_1 y_1 \\
\end{bmatrix}
\begin{bmatrix}
  r_2 + t_2 x_2 \\
  s_2 + t_2 y_2 \\
\end{bmatrix}
= \alpha
\begin{bmatrix}
  r_2 + t' x_2 \\
  s_2 + t' y_2 \\
\end{bmatrix}
\]

(21)

We eliminate \( \alpha \) in Eq.(21) and solve in terms of \( t \) and \( t' \). Then we get,

\[
\begin{align*}
t_1 &= \frac{r_2 s_1' - r_1 s_2'}{s_2 x_1' - s_2 x_2' + r_2 y_2' - r_1 y_1'} \\
t_2 &= \frac{r_2 s_1 - r_1 s_2 + r_2 y_2' - r_1 y_1'}{s_2 x_1 - s_2 x_2 + r_2 y_2 - r_1 y_1} \\
\end{align*}
\]

(22)

Thus we can get the relative lengths \( t \) and \( t' \) from the assumption that there exists a pair of parallel line segments.

3.3 Motion Parameters from a pair of Parallel Line Segments

As mentioned above, the parameters \( t \) and \( t' \) can be obtained by using parallelism. It is, however, impossible to compute motion parameters, i.e. the rotation matrix, by using just the relationship between parallel two line segments, because the equations derived from them does not get independent. Therefore we also introduce virtual lines which are the line segments obtained by connecting each end-point of two line segments as they don't make diagonals. The virtual lines are shown in Fig.5a. Next we show how to obtain the relative motion in terms of \( K \) defined in Eq.(11) by following processes.

1. In Fig.5a the line segments \( l_1 \) and \( l_2 \) denote the real line, and \( l_1' \) and \( l_2' \) the virtual line. At first the parameters \( t_1 \), \( t_1' \), \( t_2 \) and \( t_2' \) can be obtained from parallelism of line segments \( l_1 \) and \( l_2 \).

2. From the parameters \( t_1 \), \( t_1' \), \( t_2 \) and \( t_2' \), we can get the vectors \( N_1, N_1', N_2 \) and \( N_2' \) which are functions of \( t_1, t_1', t_2 \) and \( t_2' \). Substituting these results into Eq.(11) we have the ratios of vector lengths \( K_1 \) and \( K_2 \), which are for the real line segments.

3. Next step we now get the parameter \( K \) for the virtual line segments by using the computed parameters \( K_1 \) and \( K_2 \), which are for the real line segments.

4. Eq.(8) which defines the parameter \( t \) can be rewritten as follows,

\[
Z + m_j = \frac{Z_m}{l - t} \\
\]

(8')

We define the absolute coordinates along Z axis of each base-point (the origin of local coordinates system) of the line segments as \( Z_1, Z_2, Z_3 \) and \( Z_4 \), and the relative ones of the end-points as \( \mu_1, \mu_2, \mu_3 \) and \( \mu_4 \) paying attention to \( Z \) (depth) component of all points. Then the absolute coordinates of each point is shown in Fig.5b.
(19') \[ |K_3 N_j \|^2 = |N_j |^2, \]

(20') \[ K_1 K_3 N_j^T N_j = N_j^T N_j', \]

where $N_j$ and $K_2$ are given from Eq.(8), Eq.(1) etc. As $K_3 = \alpha K_1$ from Eq.(29), Eq.(19') and Eq.(20') have unknown parameters $N_j$ and $N_j'$. The unknowns $N_j$ and $N_j'$ can be expressed by $t_3$ and $t'_3$ using Eq.(8), so forth Eq.(19) is regarded as a second order equation with respect to $t$ and $t'$, and Eq.(20) as a first order equation.

Then we get two solutions for the above simultaneous equations. Furthermore we can select the unique solution using coefficients of variation, if the image data contains less error or noises. The selecting method will be discussed in section 4.

On the other hand if we introduce the assumption that there exist two pairs of parallel line segments, we can independently solve each $t$ and $t'$ from the parallelism and Eq.(22). Therefore the solutions are robust against added error or noise. The results of this will be also discussed in section 4.

3.4 Detection Method of Parallel Line Segments

In this section we discuss how to find pairs of parallel line segments. The outline of the algorithm is shown in Fig.6.

At first, one or two pair of line segments are chosen arbitrarily on the image plane, and are assumed parallel in 3D space. Next, from Eq.(22) we get $t$ and estimate $\Phi$ by using Eq.(16). By using Eq.(1) we compute translational vector $T_j$ for both the real and the virtual line segments. If the chosen line segments are parallel one another, computed vector $T_j$ are all the same. Therefore after the investigation of all combinations of the line segments, we get which ones are parallel. It is, however necessary in practice to take into account the fact there exist noise and/or quantization errors.

The experimental verification is performed as follows. Note that there are three types of geometric relationship between a pair of line segments in 3D space, which are described as follows,

(i) parallel,
(ii) on the same plane,
(iii) twisted location.

The results for three types of a pair of line segments by using the algorithms shown in Fig.6 are listed in Table1. In the table (1), (2) and (3) denote the types of geometric relationship between a pair of line segments. Each pair of line segments is assumed to be parallel, and the motion parameters are estimated, and then the translational components are computed. The values of $\Delta x/\Delta z$, $\Delta y/\Delta z$ are computed, and then the coefficients of variance, S.D.(standard deviation) / $\bar{x}$ (mean value), are computed. There is a large difference of the coefficients between the type (1) and the types (2) and (3). Hence it is effective to adopt the coefficients of the translational components in order to validate the parallelism of a pair of line segments. It is, however difficult to validate the parallelism, if the added noise increases too much because of the fact that the coefficients of type (1) also increase as the noise increases.

Next, the results for two pairs of line segments are listed in Table2, where the added noise is 0.1% of each coordinates value of all points. The table shows that it is possible to adopt the coefficients of variance in order to validate the parallelism of line segments for all combinations of the geometric relationship (type a, "*, f).

4. Experiment

4.1 Estimation of Motion Parameters

The estimation of the motion parameters are performed for one or two pairs of parallel line segments on the same plane (See Table 3). The set of terminal points of four line segments as the input data is one of the examples which Huang's method could not solve [Tsai and Huang, 1984]. The experiment is performed as follows.

At first, the terminal points of the line segments are given as $(X_{ij}, Y_{ij}, Z_{ij}), i=1, \ldots, 4, j=1,2$. Next, the rotational and translational components are given, and then we get the coordinates of the terminal points after the motion as $(X'_{ij}, Y'_{ij}, Z'_{ij}), i=1, \ldots, 4, j=1,2$. All the points are projected onto the image plane by using Eq.(4) and are added random noises according to their values. Then values $r_1 = x_1 - x'_{11}$ and $r_2 = x_0 - y'_{11}$ are computed. The reason why we added the random noise is that we must take into account the error of quantization and miss-correspondence of the line segments between images.

Fig.6 The flow of the processing.
In error free case, the estimated motion parameters are listed in the second column of Table 3, and coincide with the true values. This implies that this method is capable even when there exist more than six points on the same plane, on the other hand it is not by Huang’s method (Tsuai and Huang, 1984). In addition, our method can get the solution with the same accuracy, even if the translational components along $X$ or $Y$ axis get zero or if the rotational component is too small.

The next results are for a pair of line segments with 0.1% random noise (See line segment 1,2 of Table 3 below). In this case we have two solutions from Eq. (19). However we can select one plausible solution according to the coefficient of variance. In the Table 3 the coefficient of the first solution (shown in third column) is much smaller than that of the second one (shown in fourth column). Hence we can select the first one as plausible solution. This means that the coefficients can be used to select a plausible solution, even when the added random noise gets large.

4.2 Error effected by adding Random Noise

In order to examine the effects of the quantization error to the solution, we have experiments for the terminal points of line segments with 0.1%, 0.5%, 1%, and 2% errors. The results for two pairs of parallel line segments are shown in Table 4. From the results our method is sufficiently robust for about 0.1% noise addition. Especially the rotational components are much accurate, even when random noise is added.

5. Concluding Remarks

In this paper we proposed a new linear algorithm to estimate 3D motion parameters of a rigid object from two sequential images by using parallelisms of a pair of line segments. We formulate the motion in 3D space by means of the relative coordinate systems paying attention to the line segments on the surface of the object. Then the motion can be described only by the rotational components, and hence we get the nine nonlinear equations. We derived that if there exist a pair of parallel line segments, the nonlinear equations can be resolved into the linear ones. We also give a method to extract pairs of parallel line segments.

The present method seems especially effective for a parallel-piped object, because there must exist parallel line segments on the surface of the object. When the motion trajectory is smooth, it is not difficult to catch up with the parallel line segments. In addition, as the method achieves the linear solutions, the algorithm is very simple and runs very fast.

References


![Table 1: Detection of a pair of parallel line.](image)

![Table 2: Detection of two pairs of parallel lines](image)

![Table 3: Simulation result by our method](image)

![Table 4: Simulation results using the perturbed data.](image)