論文 98-23-9-20

Motion Estimation using Regions

Sanghoon Sull* *Regular Member*

ABSTRACT

We present a two step approach for estimating the motion and structure parameters from region correspondences in two frames. Given four or more region correspondences on the same planar surface, the motion and planar orientation parameters are first linearly estimated based on second-order approximation of the displacement field of the image plane. Then, using this linear estimate as an initial guess, a nonlinear estimate is obtained by iteratively minimizing an objective function using the exact expression of the displacement field. The objective function involves the centroids of corresponding regions and relationships among low-order moments. Through simulations, we show that the two-step region-based approach gives robust estimates. The performance of nonlinear region-based estimation is compared with that of linear region-based and point-based methods. Experimental results for two image pairs, one synthetic and one real, are presented to show the practical applicability of our approach.

I. Introduction

This paper is concerned with estimation of motion and structure parameters using image regions as features, given their correspondences in an image pair. Compared to point features which are commonly used [1], macro features such as lines and regions [2, 3, 4, 5, 7] if available are less sensitive to quantization noise and are easier to match than points. Further, it is desirable to use region-based motion estimation for efficient MPEG-4 shape coding [8].

Several region-based approaches have been proposed. Kanatani [3] developed a method for computing the motion parameters for scenes with known initial structure; his method is based on the values of certain integrals computed on one planar region under the assumption that the motion is infinitesimal. Wu et al. [4] proposed an iterative method based on optical flow on a contour. However, it is generally difficult to extract reliable region contours from a pair of noisy images. It was reported that the method usually converges if the maximal motion is smaller than a threshold related to the contour size. Some of the recent methods such as [7] minimize the displaced pixel difference with respect to 6 affine parameters, but these methods are not computationally efficient since they should scan the whole pixels inside the regions of interest when they compute the partial derivatives with respect to each parameter during the minimization.

In this paper, we present a two-step approach for estimating the motion and structure parameters for planar surfaces from region correspondences in two views. We use region correspondences instead of commonly used point features because regions are more robust features than points: being smaller in number, they are easier to match; they have more measurable properties than points (which are characterized only by their positions); and, in some images, it is difficult to extract good point features (e.g., see Figs. 5 and 6).

In the first step of our algorithm (which was partially presented in [5]), the motion and planar orientation parameters are linearly estimated by approximating the displacement field of the image plane corre-

* School of Electrical Engineering Korea University
論文番號 : 98088-0302
接受日字 : 1998年 3月 2日

sponding to each rigid planar patch by the second-order polynomials. In the second step, using the linear estimate as an initial guess, a nonlinear but accurate estimate is obtained by iteratively minimizing the objective function using the exact expression of the displacement field. The objective function is based on the differences between the observed region centroids and those predicted by the motion and structure parameters, and thus more robust to the noise-sensitive boundaries of the extracted regions.

II. Mathematical Formulation

2.1 Description of Displacement Field

Let a coordinate system (X, Y, Z) be fixed on a camera with the origin coinciding with the projection center of the camera. Assuming that the focal length is unity, the perspective projection (x, y) on the image of a point (X, Y, Z) is given by

$$x = X/Z,$$
$$y = X/Z.$$

Consider a point P on an object in 3D. Let $\vec{X} = [X, Y, Z]'$ be the 3D coordinate vector of P at time t_1 and let $\vec{X}' = [X', Y', Z']'$ be the corresponding vector at time t_2. Let $\vec{T}$ and $\mathbf{R}$ denote the translation vector and rotation about the unit axis $\vec{n}_\omega = [n_X, n_Y, n_Z]'$ by an angle ω, respectively. Then,

$$\vec{X}' = \mathbf{R}\vec{X} + \vec{T} \tag{1}$$

where

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \tag{2}$$

and

$$\vec{T} = [T_X \; T_Y \; T_Z]'.$$

Let (x, y) and (x', y') be the image coordinates corresponding to $\vec{X}$ and $\vec{X}'$, respectively. If the point P is on a plane $aX + bY + cZ = 1$ at time t_1 whose unit surface normal is denoted by $\vec{n}_S$, then

$$Z = \frac{1}{ax + by + c}. \tag{3}$$

Hence, from Eqs. (1), (2) and (3), we get

$$x' = \frac{X'}{Z'} = \frac{a_1 x + a_2 y + a_3}{a_7 x + a_8 y + a_9} \tag{4}$$

$$y' = \frac{Y'}{Z'} = \frac{a_4 x + a_5 y + a_6}{a_7 x + a_8 y + a_9} \tag{5}$$

where

$$\mathbf{A} \overset{\text{def}}{=} \begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{bmatrix} = \mathbf{R} + \vec{T}\,\vec{n}'_S. \tag{6}$$

2.2 A Two-Step Approach based on Region Correspondences

We first describe how to obtain a linear but approximate estimate using second-order approximation of the displacement field. Then, we present a nonlinear objective function to be iteratively minimized whose minimum should yield an exact estimate.

2.2.1 First Step: Linear Estimation

Using Eqs. (4) and (5), the displacement vector (D_x, D_y) is represented by

$$D_x = x' - x = \frac{a_3 + (a_1 - a_9)x + a_2 y - a_8 xy - a_7 x^2}{a_7 x + a_8 y + a_9} \tag{7}$$

$$D_y = y' - y = \frac{a_6 + a_4 x + (a_5 - a_9)y - a_7 xy - a_8 y^2}{a_7 x + a_8 y + a_9} \tag{8}$$

Then, by using Eqs. (3) and (6), the denominator in Eqs. (7) and (8) can be rewritten as follows:

$$a_7 x + a_8 y + a_9 = r_{33} + \frac{T_Z}{Z} + (r_{31} x + r_{32} y). \tag{9}$$

When the rotation angle ω is small, we have

$$r_{33} \approx 1, \quad r_{31} \approx -\omega_Y \quad \text{and} \quad r_{32} \approx \omega_X \tag{10}$$

where

$$\omega_X \stackrel{\text{def}}{=\!=} n_X \omega$$
$$\omega_Y \stackrel{\text{def}}{=\!=} n_Y \omega. \tag{11}$$

Then, the second term in Eq. (9) can be ignored if the translation in the Z-direction is small relative to the object distance from the camera, and the third term can be ignored if ω_X, ω_Y, and the image plane coordinates are small. Hence, if we assume that (1) $\frac{Tz}{Z} \ll 1$, (2) the field of view of the camera is small, and (3) the rotation about the X and Y axes is small, then

$$a_7 x + a_8 y + a_9 \approx 1. \tag{12}$$

These assumptions which are quite common in motion analysis are not very restrictive since the field of view of a camera is small in practice and the amount of motion is small if the time interval between two images is short. As in [6], we will approximate the displacement vector by the second-order polynomials in (x, y). Note that we are approximating the displace ments D_x and D_y instead of x' and y' since $|D_x|$ and $|D_y|$ are usually smaller than $|x'|$ and $|y'|$, respectively. Therefore, we have

$$D_x = a_3 + (a_1 - a_9)x + a_2 y - a_8 xy - a_7 x^2 \tag{13}$$
$$D_y = a_6 + a_4 x + (a_5 - a_9)y - a_7 xy - a_8 y^2. \tag{14}$$

It is well-known that we can also derive the second-order polynomials for (D_x, D_y) by using the instantaneous velocity formulation used for computing the optical flow.

Now, we would like to relate region properties to 3D motion and structure [5]. Equations (7) and (8) define a mapping $G: (x, y) \rightarrow (x', y')$. Hence,

$$G(x, y) \stackrel{\text{def}}{=\!=} (g_1(x, y), g_2(x, y)) = (x', y') = (x + D_x, y + D_y).$$

Let M and N be the corresponding regions at two time instants. For a smooth function f, we know

$$\int\int_N f(x, y)\, dxdy = \int\int_M f(x + D_x, y + D_y) J dxdy \tag{15}$$

where J is the Jacobian:

$$J = \left| \frac{\partial g_1}{\partial x} \frac{\partial g_2}{\partial y} - \frac{\partial g_1}{\partial y} \frac{\partial g_2}{\partial x} \right|. \tag{16}$$

Using Eqs. (4), (5), (6) and (16), we arrive at

$$J = \frac{|\, det(\mathbf{A})\,|}{|\,(a_7 x + a_8 y + a_9)^3\,|}, \tag{17}$$

where $det(\mathbf{A})$ represents the determinant of the matrix $\mathbf{A}$. The numerator of J in Eq. (17) is a constant, and the expression in the denominator can be rewritten as previously:

$$a_7 x + a_8 y + a_9 = r_{33} + \frac{Tz}{Z} + (r_{31} x + r_{32} y).$$

We assume that three conditions stated earlier are satisfied. Then, the second and third terms are small compared to 1. Further, variations in Z, x and y are not large since the integration is performed over one region M in the image plane which is small. Hence, the denominator of J can be approximated as a constant over M. Then, Eq. (15) becomes

$$\int\int_N f(x, y)\, dxdy \approx J \int\int_M f(x + \tilde{D}_x, y + \tilde{D}_y)\, dxdy. \tag{18}$$

Note that if $f = 1$, then the equation simply repre-

sents the area relationship:

$$N_{00} = J M_{00} \tag{19}$$

where

$$N_{ij} \stackrel{\text{def}}{=} \int\int_N x^i y^j \, dxdy$$
$$M_{ij} \stackrel{\text{def}}{=} \int\int_M x^i y^j \, dxdy. \tag{20}$$

Hence, given a pair of regions at two time instants, J can be treated as a constant which can be easily computed. In principle, we can use a sufficiently large number of functions f to find the unknowns. However, we use only two functions $f = x$ and $f = y$ since these choices make $f(x + \tilde{D}_x, y + \tilde{D}_y)$ linear in the coefficients $a_1, \ldots, a_9$. Hence, by Eqs. (13), (14), (18) and (19), we have two equations for each region correspondence:

$$\frac{N_{10}}{N_{00}} - \frac{M_{10}}{M_{00}} = a_3 + (a_1 - a_9)\frac{M_{10}}{M_{00}} + a_2\frac{M_{01}}{M_{00}} - a_8\frac{M_{11}}{M_{00}} - a_7\frac{M_{20}}{M_{00}} \tag{21}$$

$$\frac{N_{01}}{N_{00}} - \frac{M_{01}}{M_{00}} = a_6 + a_4\frac{M_{10}}{M_{00}} + (a_5 - a_9)\frac{M_{01}}{M_{00}} - a_7\frac{M_{11}}{M_{00}} - a_8\frac{M_{02}}{M_{00}} \tag{22}$$

where $N_{10}, \ldots, M_{01}$ are defined in Eq. (20). Therefore, we can linearly compute 8 coefficients $a_1, \ldots, a_8$ from four or more region correspondences with a_9 set to one since the coefficient a_9 can have any value in the above two equations. Then, using the algorithm presented in [1], we can noniteratively solve for the motion and plane normal.

2.2.2 Second Step: Nonlinear Estimation

Since the set of nine coefficients in Eqs. (4) and (5) are not independently determined, we define a new set of coefficients $\mathbf{C} = (c_1, \ldots, c_8)$ as follows:

$$c_i = \frac{a_i}{a_9} \quad \text{for} \quad i \le 8. \tag{23}$$

Our goal is to compute $\mathbf{C}$ without any approximation from region correspondences. Thus, the problem is nonlinear. Given $\mathbf{C}$, we can easily compute the motion and structure parameters using the algorithms in [1].

Let M_i and N_i be the corresponding regions at two time instants, and let M_i' be the predicted region for moving region M_i according to given $\mathbf{C}$. If X is a quantity computed from the region N_i, then let $\widehat{X}$ denote the value of X computed from the predicted region M_i.

The solution obtained by the method to be presented in this section minimizes a measure of inconsistency between the observed regions (i.e. N_i) at t_2 and those predicted by the estimates of $\mathbf{C}$ (i.e. M_i'). To obtain such measure for a given $\mathbf{C}$, we define two error terms for each region correspondence i:

$$E_{i,x}(\mathbf{C}) \stackrel{\text{def}}{=} \frac{N_{i,10}}{N_{i,00}} - \frac{\widehat{N_{i,10}}}{\widehat{N_{i,00}}} \tag{24}$$

$$E_{i,y}(\mathbf{C}) \stackrel{\text{def}}{=} \frac{N_{i,01}}{N_{i,00}} - \frac{\widehat{N_{i,01}}}{\widehat{N_{i,00}}} \tag{25}$$

where $N_{i,00}$, $N_{i,10}$ and $N_{i,01}$ are the moments of N_i defined in Eq. (20). Note that $E_{i,x}(\mathbf{C})$ and $E_{i,y}(\mathbf{C})$ represent the differences in x and y coordinates of the centroids for N_i and M_i', respectively.

The estimates of $\mathbf{C}$ are computed by minimizing the following objective function with respect to $\mathbf{C}$ using L region correspondences:

$$E(\mathbf{C}) \stackrel{\text{def}}{=} \sum_{i=1}^{L} (E_{i,x}^2 + E_{i,y}^2). \tag{26}$$

Here we have used only low-order moments of region

shape to avoid the noise sensitivity of higher moments. Also, note that our region-based method does not use region contour information since the exact contours are very difficult to extract reliably.

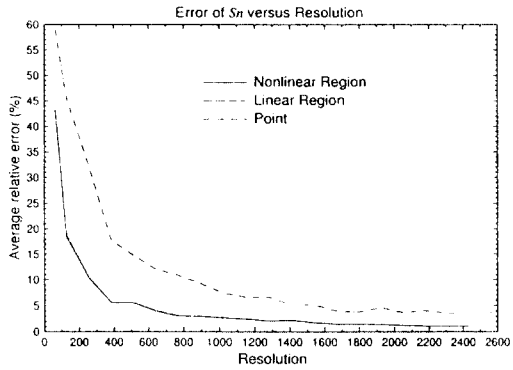
III. Performance Evaluation

Using known motion parameters and simulated images, performance of the approximate linear estimation method is compared with that of the exact iterative method. Performance of the point-based estimation [1] is also presented for comparison. The values of $\frac{N_{i,10}}{N_{i,00}}$ and $\frac{N_{i,01}}{N_{i,00}}$ are computed once and the lists of boundary points of each region M_i at t_1 are stored. To compute $\frac{\widehat{N_{i,10}}}{\widehat{N_{i,00}}}$ and $\frac{\widehat{N_{i,01}}}{\widehat{N_{i,00}}}$ for each pair of corresponding regions M_i and N_i at each iteration given C, the boundary points of M_i are predicted by using Eqs. (4) and (5). Then, $\widehat{N_{i,00}}$, $\widehat{N_{i,10}}$ and $\widehat{N_{i,01}}$ are computed by applying Green's Theorem to Eq. (20). To iteratively minimize Eq. (26), we use a modified Levenberg-Marquardt algorithm.

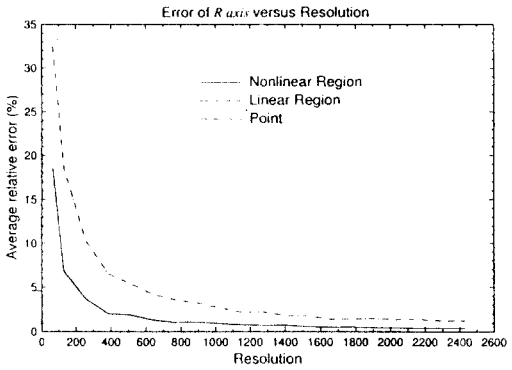
The size of the simulated image plane is 0.5×0.5. Twelve feature correspondences are used for points and regions. At each trial, a plane passing through the point at (0, 0, 10) is randomly generated. Then, 3D coordinates of point features on the given plane are randomly generated. A region boundary is generated by four random variables: two for the center point, one for the radius (from 0.01 to 0.08) and one for the number of points (from 16 to 64) on a region boundary. Only those features which are within visual fields in two frames are generated. The image coordinates of the points are quantized to the nearest integer for each resolution. The relative error of a vector is defined by the Euclidean norm of the error vector divided by the Euclidean norm of the correct vector. The surface normal $\vec{n}_{S,1}$ is scaled to the unit vector. All errors represent the average errors over 50 random trials.

We try three sets of motion parameters at various image resolutions. Figure 1 shows the average relative errors of surface normal $\vec{n}_{S,1}$, rotation axis $\vec{n}_\omega$, rotation angle ω and translation $\vec{T}$ for the first set of known motion parameters ($\vec{n}_\omega = [0, 0, 1]'$, $\omega = 10$ deg and $\vec{T} = [0.2, 0.2, 0]'$). We note that this ideal set of parameters perfectly satisfies three assumptions stated in Section 2.2.1. Figure 2 shows the average relative errors of $\vec{n}_{S,1}$, $\vec{n}_\omega$, ω and $\vec{T}$ for the second set of motion parameters $\vec{n}_\omega = [0.5774, 0.5774, 0.5774]'$, $\omega = 4$ deg and $\vec{T} = [0.2,0.2,0.2]'$). Finally, Fig. 3 shows the average relative errors for the third set of motion parameters $\vec{n}_\omega = [0.5774, 0.5774, 0.5774]'$, $\omega = 7$ deg and $\vec{T} = [0.2, 0.2, 2]'$). Note that this set of motion parameters has relatively large values of T_Z and rotation about the X and Y axes, which do not well satisfy the three assumptions.

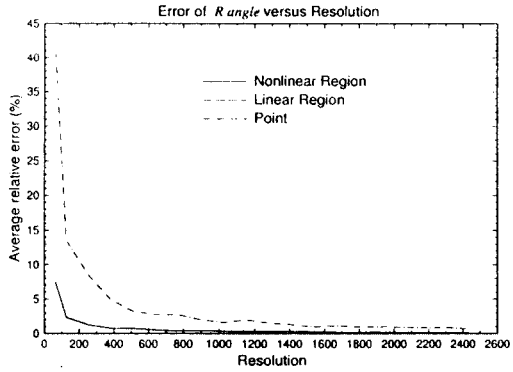
The following observations can be made from these simulations. First, as the resolution becomes higher, the linear point-based estimates and the nonlinear region-based estimates become more accurate whereas the accuracy of the linear region-based estimates does not improve beyond certain resolution as we can see from Figs. 2 and 3. This is expected since the approximations are made under three assumptions (See Section 2.2.1.) when the region-based equations for the linear estimation are derived. Note that both linear and nonlinear region-based estimations have the same performance in the ideal case such as Fig. 1. Second, for the type of motions where three assumptions are relatively well satisfied (for example, Figs. 1 and 2), the region-based estimations, both linear and nonlinear, yield more reliable result than point-based method for low to mid resolutions due to the robustness of the centroids of regions to the quantization errors of the region boundaries. However, when the assumptions are not well satisfied, (for example, Fig. 3), the linear point-based estimates are better than the linear region-based ones. Third, estimation of the translation is the most noise-sensitive.



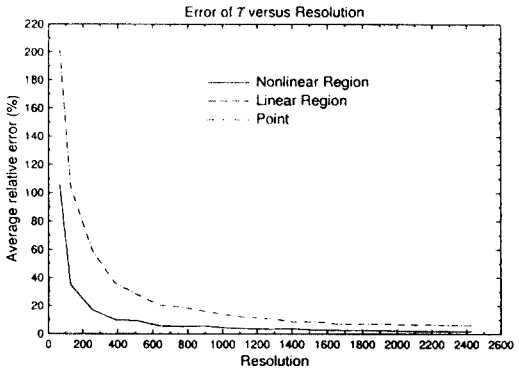
(a) Error of surface normal ($\vec{n}_{S,1}$).



(b) Error of rotation axis ($\vec{n}_{\omega}$).

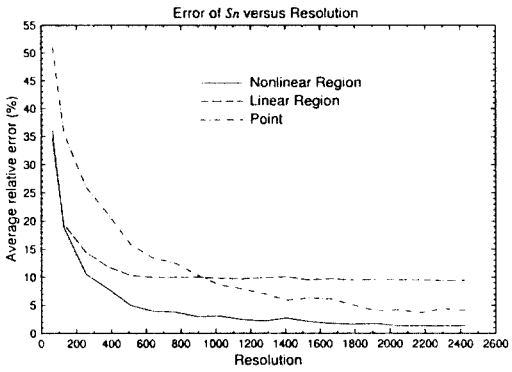


(c) Error of rotation angle (ω).

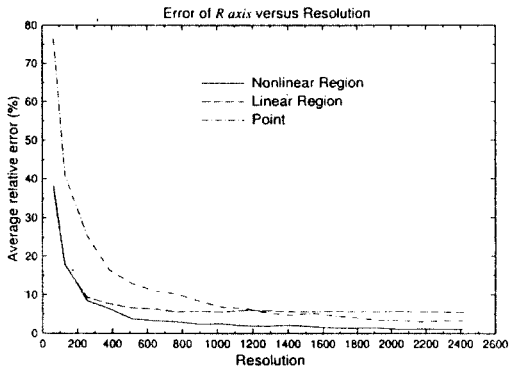


(d) Error of translation ($\vec{T}$).

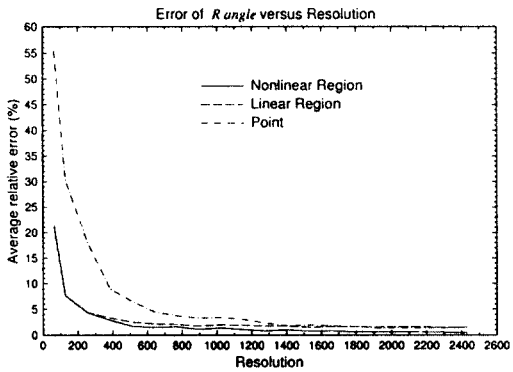
Fig. 1 Performance Comparison. Resolutions are varied from 64×64 to 2432×2432 for a set of actual parameters ($\vec{n}_{\omega}=[0,\ 0,\ 1]'$, $\omega=10$ deg and $\vec{T}=[0.2,\ 0.2,\ 0]'$).



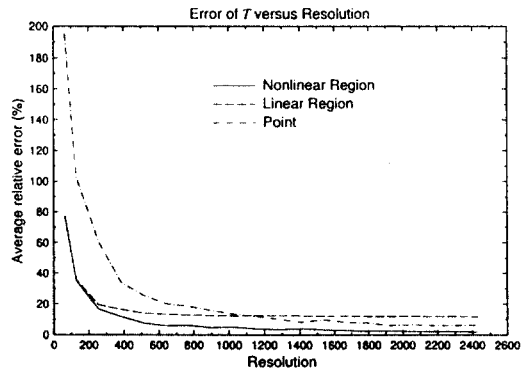
(a) Error of surface normal ($\vec{n}_{S,1}$).



(b) Error of rotation axis ($\vec{n}_{\omega}$).

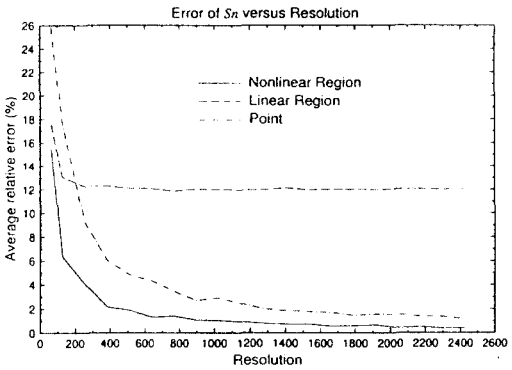


(c) Error of rotation angle (ω).

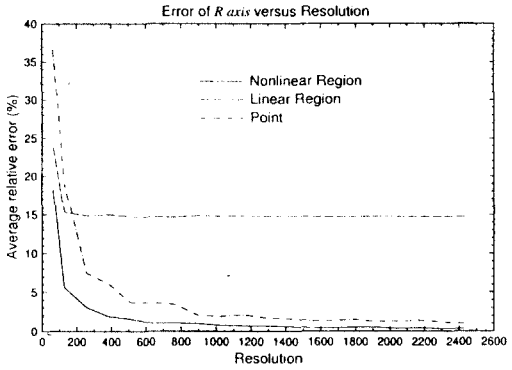


(d) Error of translation ($\vec{T}$).

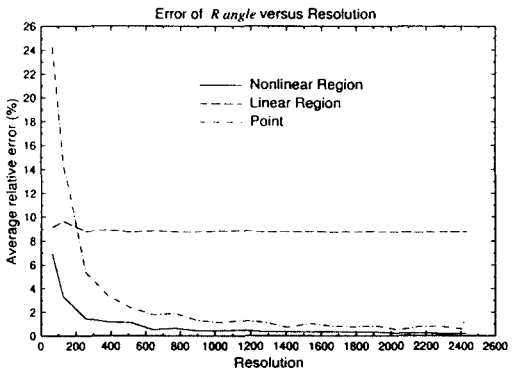
Fig. 2 Performance Comparison. Resolutions are varied from 64×64 to 2432×2432 for a set of actual parameters ($\vec{n}_{\omega}$ = [0.5774, 0.5774, 0.5774]′, ω = 4 deg and $\vec{T}$ = [0.2, 0.2, 0.2]′).



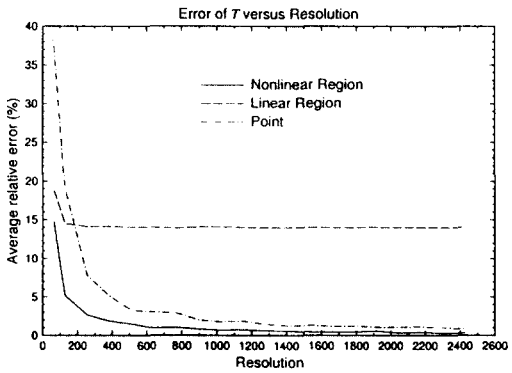
(a) Error of surface normal ($\vec{n}_{S,1}$).



(b) Error of rotation axis ($\vec{n}_{\omega}$).



(c) Error of rotation angle (ω).



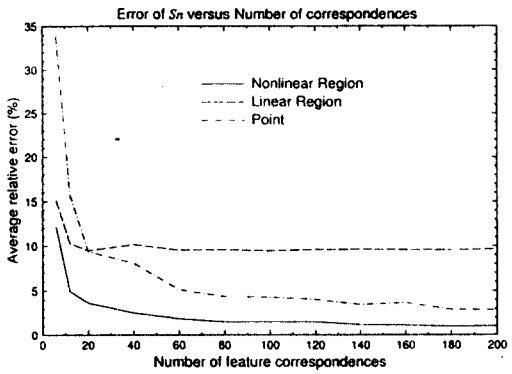
(d) Error of translation ($\vec{T}$).

Fig. 3 Performance Comparison. Resolutions are varied from 64×64 to 2432×2432 for a set of actual parameters ($\vec{n}_{\omega}=[0.5774,\ 0.5774,\ 0.5774]'$, $\omega=7$ deg and $\vec{T}=[0.2,\ 0.2,\ 2]'$).

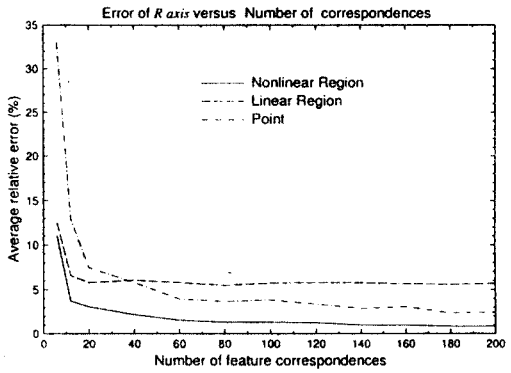
We also vary the number of feature correspondences used for estimation at a fixed resolution. Figure 4 shows the average relative errors of $\vec{n}_{S,1}$, $\vec{n}_{\omega}$, ω and $\vec{T}$ for the second set of motion parameters ($\vec{n}_{\omega}$ = [0.5774, 0.5774, 0.5774]′, ω = 4 deg and $\vec{T}$ = [0.2, 0.2, 0.2]′). We observe that increasing the number of features gives better estimates for the nonlinear region-based estimates and the linear point-based estimates. However, the accuracy of the linear region-based estimates does not improve beyond certain resolution due to the approximations.

From these simulations, we see that the two-step region-based approach gives

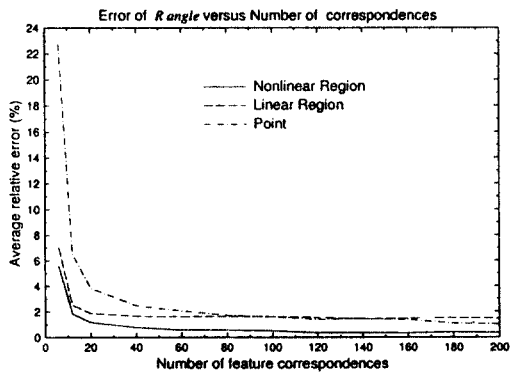
robust estimates at all resolutions for any type of motion of a planar surface. Note that these simulations are based on the assumption of perfect extraction and matching of features up to the quantization errors.



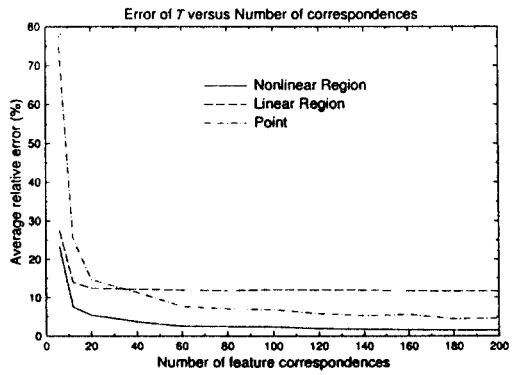
(a) Error of surface normal ($\vec{n}_{S,1}$).



(b) Error of rotation axis ($\vec{n}_{\omega}$).



(c) Error of rotation angle (ω).



(d) Error of translation ($\vec{T}$).

Fig. 4 Performance Comparison. The number of feature correspondences is varied from 6 to 200 at a fixed resolution of 512×512 for a set of actual parameters ($\vec{n}_{\omega}$ = [0.5774, 0.5774, 0.5774]′, ω = 4 deg and $\vec{T}$ = [0.2, 0.2, 0.2]′).

IV. Experimental Results with Image Pairs

4.1 Implementation Details

Regions are extracted using multiple intensity thresholds. For point feature detection, we use local maxima and minima of intensity values. Then, the algorithm described in [5] is applied to pairs of the images to match the points and regions and then to segment them into the locally planar surface patches. The approximate linear estimation method and the exact nonlinear method are applied to estimate the motion and structure parameters from region correspondences. The linear algorithm based on point correspondences presented in [1] is also applied for comparison.

4.2 Experimental Results for Two pairs of Images

We conducted the experiments with two pairs of images, one synthetic and one real. Although there are dual solutions for a planar surface, only the closer solution to the ground truth is shown.

4.2.1 Runway Images

We use two frames of a synthetic sequence. The focal length is 8 mm. The field of view of the camera is 40 by 34.4 (deg) corresponding to the image resolution of 560 by 480. The matching and segmentation algorithm yields 21 region correspondences and 57 point correspondences. The input images, the segment corresponding to the ground and matching result are shown in Fig. 5.

True values and results of nonlinear region-based estimation of the motion and structure parameters are shown in Table 1, where the estimates can be seen very good. In Table 2, relative errors of point-based estimates, linear region-based estimates and nonlinear region-based estimates are presented for comparison. Errors of the linear region-based estimates are close to those of the nonlinear region-based ones. Points do not give good results in this case since it is difficult to extract good point features.

(a)

(b)

Fig. 5 Runway Images: (a) First image, and its matched points and regions. (b) Second image, and its matched points and regions.

Table 1. Runway: True values and nonlinear region-based estimates of surface normal ($\vec{n}_{S,1}$) at t_1, rotation axis ($\vec{n}_\omega$), rotation angle (ω), and translation ($\vec{T}$).

	True Values	Nonlinear Sol.
$\vec{n}_{S,1}$	[−0.9150, 0.3624, 0.1772]	[−0.8934, 0.4109, 0.1815]
$\vec{n}_\omega$	[0.1222, −0.0368, 0.9918]	[0.1266, −0.0413, 0.9911]
ω	7.2000 deg	7.3108 deg
$\vec{T}$	[0.0460, −0.0160, 0.2300]	[0.0487, −0.0092, 0.2206]

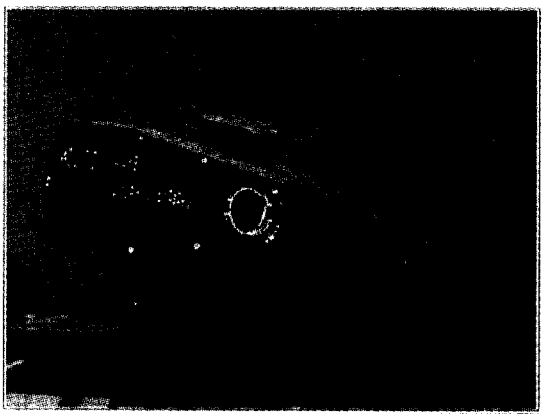
Table 2. Runway: Relative estimation errors.

Error in %	Point	Linear Reg.	Nonlinear Reg.
$\vec{n}_{S,1}$	71.5692	5.3520	5.3267
$\vec{n}_{\omega}$	3.7027	0.7257	0.63326
ω	15.0444	2.0319	1.5389
$\vec{T}$	25.1390	5.7663	5.0667

4.2.2 PUMA Image

Two real images of a PUMA robot are used. The focal length is 35 mm and the field of view of the camera is approximately 13 deg. The image size is 500 by 384. The matching and segmentation algorithm yields 6 region correspondences and 45 point correspondences for the segment corresponding the large arm of the robot. The input images, the segment corresponding to the large arm and matching result are shown in Fig. 6.

True values and results of nonlinear region-based estimation are shown in Table 3 where we see that the estimates are good. In Table 4, relative errors of point-based estimates, linear region-based estimates and nonlinear region-based estimates are presented for comparison. The linear region-based estimates are almost identical to the nonlinear region-based ones. Errors of the point-based estimates are large due to the difficulty of extracting reliable point features.



(a)

(b)

Fig. 6 PUMA Images: (a) First image, and its matched points and regions. (b) Second image, and its matched points and regions.

Table 3. PUMA : True values and nonlinear estimates.

	True Values	Nonlinear Sol.
$\vec{n}_{S,1}$	[−0.0214, 0.6180, 0.7860]	[0.1388, 0.6874, 0.7129]
$\vec{n}_{\omega}$	[−0.0214, 0.6180, 0.7860]	[−0.0897, 0.5442, 0.8342]
ω	−3.7529 deg	−3.6514 deg
$\vec{T}$	NA	[0.0543, 0.0070, 0.0086]

Table 4. PUMA : Relative estimation errors.

Error in %	Point	Linear Reg.	Nonlinear Reg.
$\vec{n}_{S,1}$	85.5166	18.9255	18.9255
$\vec{n}_{\omega}$	70.8301	11.15	11.15
ω	19.6808	2.7019	2.7046

V. Conclustion

We have presented a two-step approach for estimating the motion and structure parameters for piecewise planar surfaces from region correspondences in two views. We have shown that our approach gives robust estimates. We also showed that the linear, but approximate solution yields a performance similar to the nonlinear (exact) solution at moderate resolutions

if the commonly-assumed conditions are satisfied. Our algorithm is less sensitive to noise than the point-based methods because it uses regions as features which are more robust features than points. The problem with point-based methods mainly comes from the difficulty with extracting good point features despite its easier mathematical manipulation. For the comparison with the existing region-based approaches [7], our method is computationally efficient since it scans the boundary points of regions to compute the first moments (using Green's Theorem) during the minimization rather than scanning the whole pixels of the regions.

For the MPEG-4 shape coding, it is necessary to compute the motion of regions (or contour) and thus our approach for region-based estimation can be used.

References

1. R. Tsai, T. Huang and W. Zhu, "Estimating three dimensional motion parameters of a rigid planar patch, II: Singular value decomposition," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. ASSP-30, pp. 525-534, Aug. 1982.
2. Y. Liu and T. Huang, "Estimation of rigid body motion using straight line correspondences," *Comput. Vis., Graph., Image Process.*, vol. 43, pp. 37-52, Jul. 1988.
3. K. Kanatani, "Detecting the motion of a planar surface by line and surface integrals," *Comput. Vis., Graph., Image Process.*, vol. 29, pp. 13-22, 1985.
4. J. Wu, R. Brockett, and K. Wohn, "A contour-based recovery of image flow: Iterative method," in *Proc. IEEE Conf. Comput. Vision Patt. Recogn.*, (San Diego, CA), pp. 124-129, 1989.
5. S. Sull and N. Ahuja, "Integrated 3D analysis and analysis guided synthesis of flight image sequences," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 16, pp. 357--372, April 1994.
6. G. Adiv, "Determining three-dimensional motion and structure from optical flow generated by several moving objects," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. PAMI-7, pp. 384-401, July 1985.
7. V. Dang, A. Mansouri and J. Konrad, "Motion estimation for region-based video coding," in *Proc. IEEE Intern. Conf. Image Proc.*, (Washington, D.C.), vol II, pp. 189-192, Oct. 1995.
8. MPEG Group, "Overview of the MPEG-4 Version 1 standard," ISO/SC29/WG11, N2196, Mar. 1998.

설 상 훈(Sanghoon Sull) 정회원

1981년 : 서울대학교 전자공학과 학사

1983년 : 한국과학기술원 전기 및 전자공학과 석사

1993년 : University of Illinois, Urbana-Champaign, Dept. of Electrical and Computer Eng. Ph.D

1983년 2월 ~ 1986년 7월 : KBS 기술연구소

1993년 1월 ~ 1994년 2월 : Univ. of Illinois, Beckman Institute, Artificial Intelligence/Vision Group

1994년 3월 ~ 1996년 8월 : NASA Ames Research Center, Flight Deck Branch

1996년 8월 ~ 1997년 8월 : IBM Almaden Research Center, Visual Media Management Department

현재 : 고려대학교 전기전자전파공학부 부교수

※주관심분야 : 영상 이해 및 처리, 멀티미디어 데이터 검색 및 관리