

Fig. 10. Average speedup values with respect to full resolution classification. Two different SRNN confidence thresholds are used.

#### IV. CONCLUSIONS

The SRNN image classification technique, which is based on human fixation behavior is proposed for fast classification of images. Starting from the lowest possible resolution, SRNN classifier sequentially increases the resolution on the image segment to be classified, as long as the embedded  $k$ -nearest neighbors classifier gives the "no decision" answer. In our texture discrimination experiments, SRNN obtained up to two orders of magnitude speedup with respect to full resolution classification without reducing the classification accuracy.

The implementation strategy for the "no decision" class is found to be very important for SRNN classification accuracy and speed. It is possible to make a compromise between accuracy and speed by choosing the confidence threshold of SRNN appropriately. If the "no decision" class appears frequently (high confidence threshold), then the multiresolution performance curves of SRNN iterations spread apart, the classification accuracy increases but the overall speedup decreases. Conversely, using low confidence thresholds, one can obtain very high speedup in comparison to full resolution classification, but a drop in the classification accuracy is also observed.

#### REFERENCES

- [1] C. A. Curcio, "Human photoreceptor topography," *J. Compar. Neurol.*, vol. 292, pp. 497–523, 1990.
- [2] C. A. Curcio and K. A. Allen, "Topography of ganglion cells across human retina," *J. Compar. Neurol.*, vol. 300, pp. 5–25, 1990.
- [3] J. L. Croner and E. Kaplan, "Receptive fields of P and M ganglion cells across the primate retina," *Vision Res.*, vol. 35, no. 1, pp. 7–24, 1995.
- [4] W. S. Geisler and M. S. Banks, "Visual performance," in *Handbook of Optics*. New York: McGraw-Hill, 1995.
- [5] T. Kuyel, W. S. Geisler, and J. Ghosh, "A nonparametric statistical analysis of texture segmentation performance using a foveated image preprocessing similar to the human retina," in *Proc. IEEE SSIAI'96*, 1996, pp. 207–212.
- [6] T. Kuyel and J. Ghosh, "Multiresolution nearest neighbor classifier: A fast algorithm for classification of images," in *Proc. SIP'96*, 1996, pp. 122–126.
- [7] I. Tomek, "A generalization of the  $k$ -NN rule," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-6, no. 2, pp. 121–126, 1976.
- [8] S. A. Dudani, "Distance-weighted  $k$ -nearest-neighbor rule," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-6, no. 2, pp. 121–126, 1976.
- [9] J. Macleod and J. A. Givens, "A re-examination of the distance-weighted  $k$ -NN classification rule," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-15, no. 4, pp. 689–696, 1987.
- [10] C. K. Chow, "On optimum recognition error and rejection tradeoff," *IEEE Trans. Inform. Theory*, vol. 16, pp. 41–46, 1970.
- [11] J. Shurmann, *Pattern Classification: A Unified View of Statistical and Neural Approaches*. New York: Wiley, 1996.

- [12] A. Sharkey, "On combining artificial neural networks," *Connect. Sci.*, vol. 8, nos. 3/4, pp. 299–314, 1996.
- [13] C. Thianhorng, "Texture analysis and classification with tree structured wavelet transform," *IEEE Trans. Image Process.*, vol. 2, no. 4, pp. 429–442, 1993.
- [14] M. Unser and M. Eden, "Multiresolution feature extraction and selection for texture segmentation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-11, pp. 717–728, 1989.
- [15] S. J. Roan, J. K. Aggarwal, and W. N. Martin, "Multiple resolution imagery and texture analysis," *Pattern Recognit.*, vol. 20, pp. 17–31, 1987.
- [16] T. Kuyel, "Foveated models for compression, texture discrimination and classification," Ph.D. dissertation, Univ. Texas, Austin, 1997.

#### Detecting 3-D Motion Field from Range Image Sequences

Krishnendu Chaudhury, Rajiv Mehrotra, and Cid Srinivasan

**Abstract**—An algorithm for computing three-dimensional (3-D) velocity field and motion parameters from a range image sequence is presented. It is based on a new integral 3-D rigid motion invariant feature—the trace of a  $3 \times 3$  "feature matrix" related to the moment of inertia tensor. This trace can be computed at every point of a moving surface and provides a quantitative measure of the local shape of the surface. Based on the feature's conservation along the trajectory of every moving point, a 3-D Flow Constraint Equation is formulated and solved for the velocity field. The robustness of the feature in presence of noise and discontinuity is analyzed.

**Index Terms**—Motion analysis, optical flow, rigid motion invariants, stereo, 3-D flow, 3-D motion detection.

#### I. INTRODUCTION

Detecting motion from a sequence of images is an important problem in computer vision. Intuitively, it seems obvious that detection of three-dimensional (3-D) motion will be easier if we are given all the three coordinates of the position vectors of image points (as in range images) rather than two coordinates (as in intensity images). However, because of the unreliability associated with range image detection techniques, the most widely used form of image in motion detection so far has been the intensity image. An inherent problem here is the recovery of the 3-D motion field from the detected two-dimensional (2-D) velocity field (optical flow).

Recent advances in range imaging technology make it realistic and necessary to address the problem of detecting motion from a sequence of range images. Most of the existing techniques for range images are feature-based. Consequently, they depend on the detection of reliable range image features and establishment of interframe correspondence among them. A variety of techniques have been suggested for recovering the 3-D motion parameters, once correspondence of features is

Manuscript received January 15, 1995; revised June 17, 1996 and August 6, 1997. This paper was recommended by Associate Editor K. Pattipati.

K. Chaudhury is with the Department of Advanced Development, Adobe Systems, Inc., San Jose, CA 95110-2702 USA (e-mail: kchaudhu@adobe.com).

R. Mehrotra is with the Kodak Research and Development Laboratories, Imaging Science Division, Imaging Research and Advanced Development, Eastman Kodak Company, Rochester, NY 14650-18161 USA.

C. Srinivasan is with the Department of Statistics, University of Kentucky, Lexington, KY 40506 USA.

Publisher Item Identifier S 1083-4419(99)00909-7.

established [1], [4], [9]. The real difficulty with feature-based motion detection from image sequences lies in establishing correspondence of features. Clearly, a “correspondence-less” technique for detecting motion from range images is highly desirable. Recently Horn and Harris [5] reported a pioneering effort in this direction. Given a range image in Cartesian or spherical coordinate system, they convert it to a Cartesian Elevation Map (CEM), in which the height (depth)  $Z$  is expressed as a function of  $X$  and  $Y$ , displacements in the horizontal plane. A time varying CEM can be expressed as a function of the form  $Z(X, Y, t)$ , where  $t$  denotes time. It should be noted that a set of 3-D points obtained by transforming from spherical to Cartesian coordinates will not, in general, be placed regularly on a rectangular grid. Horn and Harris used an interpolation scheme, using the elastic membrane model to obtain a smooth, dense CEM in which the points are regularly placed on a grid. The following identity (called the *elevation rate constraint equation* by the authors) forms the basis of the approach:

$$\frac{dZ}{dt} = \frac{\partial Z}{\partial x} \frac{dX}{dt} + \frac{\partial Z}{\partial Y} \frac{dY}{dt} + \frac{\partial Z}{\partial t}. \quad (1)$$

The instantaneous position vector of a point  $\vec{r} = [X \ Y \ Z]^T$ , the instantaneous translational velocity as  $\vec{T} = [T_1 \ T_2 \ T_3]^T$ , and the instantaneous angular velocity vector as  $\vec{\omega} = [A \ B \ C]^T$  are related as  $(d\vec{r}/dt) = -\vec{T} - \vec{\omega} \times \vec{r}$ . Rewriting the above vector equation in coordinate form, we obtain three equations, expressing  $(dX/dt)$ ,  $(dY/dt)$ ,  $(dZ/dt)$  in terms of  $T_1, T_2, T_3, A, B$ , and  $C$ . Substituting in (1), a constraint equation is obtained, with six unknowns,  $T_1, T_2, T_3, A, B$ , and  $C$ . One such equation is obtained for each point in the CEM. The resulting large overdetermined system is solved by pseudo-inverse technique to obtain the unknowns.

Though commendable as a pioneering effort, the above technique suffers from a number of serious drawbacks. It assumes the presence of a single rigid motion, which imposes the rather restrictive assumption of passive motion, or the absence of motion boundaries in the scene. Error involved in estimation of the partial derivatives of  $Z$  at points of discontinuity can introduce significant errors in the overall motion parameters, as the least-square overdetermined technique is known to be sensitive to large errors in one or more equations. The situation becomes worse when we consider the effect of noise in the  $Z$  values. Three-dimensional range data is known to be noise prone. The partial derivatives enhance the effect of any noise present in that data.

In this paper, a new “correspondence-less” approach has been proposed for estimating 3-D motion (velocity field and motion parameters) from a range image sequence. The proposed technique overcomes all the above-mentioned shortcomings of the Horn–Harris approach. Our approach utilizes an integral feature that is *local* in nature and is *invariant to rigid 3-D motion* to compute the 3-D flow field (or velocity vectors). The computed 3-D flow field is used to recover rigid motion parameters. In the following section, the invariant feature used in the proposed technique is described. The methods for the estimation of 3-D flow field and the rigid motion recovery are presented in Section III. The influence of noise and image discontinuities on the proposed invariant feature are analyzed in Sections IV and V. Some experimental results are presented in Section VI and conclusions are offered in Section VII.

## II. THE INVARIANT FEATURE

In this section, we define a “feature matrix” which can be computed for any set of 3-D points and whose trace does not change if the set of points undergo 3-D rigid motion transformation.

Let  $S = \{\vec{r}_j = [x_j \ y_j \ z_j]^T; j = 1 \cdots N\}$  denote the 3-D position vectors of a set of points. The center of mass of this set (assuming

equal mass for all points) is given by

$$\vec{\mu} = \frac{1}{N} \sum_{j=1}^N \vec{r}_j. \quad (2)$$

Consider the  $3 \times 3$  matrices  $C_j = (\vec{r}_j - \vec{\mu})(\vec{r}_j - \vec{\mu})^T$  and  $C = (1/N) \sum_{j=1}^N C_j$ . Henceforth, the matrix  $C$  defined above will be referred to as the feature matrix<sup>1</sup> corresponding to  $S$ . Let  $R$  be a proper rotation matrix satisfying  $R^T R = R R^T = I$  ( $I$  stands for the  $3 \times 3$  identity matrix) and  $\det(R) = 1$ , and  $\vec{b}$  be a translation vector. Let  $S' = \{\vec{r}'_j; j = 1 \cdots n\}$  denote the set of points obtained by applying the rigid transformation  $(R, \vec{b})$  to the points in  $S$ , such that  $\vec{r}'_j = R\vec{r}_j + \vec{b}$ . It can be easily shown that, the new center of mass of the set of points  $S'$  is given by  $\vec{\mu}' = R\vec{\mu} + \vec{b}$  and the new feature matrix  $C' = R C R^T$ . Thus the matrices  $C$  and  $C'$  are related by orthogonal similarity transformations, ( $R$  being a proper rotation matrix). Hence, they have the same eigenvalues [10], trace, and determinant. Expressed in terms of components, the trace of the feature matrix is

$$\text{Tr}(C) = \frac{1}{N} \sum_{j=1}^N ((x_j - \mu^x)^2 + (y_j - \mu^y)^2 + (z_j - \mu^z)^2). \quad (3)$$

The above results are equally true for a continuous set of points (the summations become integrals). In particular, if  $S = \{\vec{r}(x, y) = [x \ y \ z(x, y)]\}$  denotes a continuous surface patch, and  $\|S\|$  denotes area of the surface, the corresponding trace of the feature matrix can be expressed in terms of its components as

$$\text{Tr}(C) = \frac{1}{\|S\|} \int_S ((x - \mu^x)^2 + (y - \mu^y)^2 + (z - \mu^z)^2) dS \quad (4)$$

and is invariant to 3-D rigid motion.

Thus, if  $S$  denotes the set of points belonging to a surface patch, the corresponding  $\text{Tr}(C)$  and  $\det(C)$  will be rigid motion invariant features of the surface patch. Due to poor noise and discontinuity properties of the determinant, in this paper, only the trace is used as an invariant feature.

Examining (3), we find that  $\text{Tr}(C)$  measures the “spread” of the given set of points around the center of mass  $\vec{\mu}$ . In other words, the trace of the feature matrix  $C$  corresponding to the surface patch  $S$  is a quantitative measure for the “geometrical shape” of the surface patch. It is intuitively obvious that such a quantity will be invariant to rotation and/or translation of the given set of points  $S$ . The feature matrix  $C$  is also related to the moment of inertia matrix  $M$  for a system of particles (the moment of inertia matrix for a system of particles  $M$  relates the angular momentum  $\vec{L}$  and the angular velocity  $\vec{\omega}$  as  $\vec{L} = M\vec{\omega}$  and thus plays a role in rotational motion analogous to that of mass in linear motion [2]) as  $C = ((1/N) \sum_{j=1}^N r_j^2)I - M$ , where  $I$  is the identity matrix. It is a well known fact in classical mechanics that the trace of the moment of inertia matrix is invariant to rotation of the system of particles. Also, the quantity  $((1/N) \sum_{j=1}^N r_j^2)I$  is invariant to rotation of the system of particles. This corroborates our conclusion that trace of the feature matrix is rigid motion invariant.

## III. MOTION ESTIMATION

The trace feature introduced in Section II can be computed *locally* at every point of a 3-D surface from a neighborhood around this point. This neighborhood should also be invariant to rigid motion. For that

<sup>1</sup>The expression for the feature matrix is similar to that for the covariance matrix in multivariate statistics. However, the points  $\vec{r}_j$  here are not random vectors. They are deterministic points belonging to some surface patch.

reason, we have used a “spherical neighborhood” around each point on the surface<sup>2</sup>. Given a simple surface in the form of a Monge patch [3],  $S = \{\vec{r}(x, y) = [x \ y \ z(x, y)]^T: (x, y) \in D\}$  ( $D$  is some domain) the value of the trace feature at a point  $p = [x_p \ y_p \ z(x_p, y_p)]^T$ , is given by (4) where the surface of integration  $S$  is the spherical neighborhood  $S_p$  of point  $p$  (see Section II). We denote the trace feature as  $T$ . Since  $T_p$  can be computed at every point  $p$  on the surface, corresponding to any surface  $S$ , we can have a “feature surface”  $F = \{[x \ y \ T(x, y)]^T: (x, y) \in D\}$ .

Now, if the surface  $S$  undergoes rigid motion with time, the moving surface can be represented as  $S(t) = \{\vec{r}(x, y, t) = [x(t) \ y(t) \ z(x(t), y(t), t)]^T: (x, y, t) \in \mathcal{V}\}$  where  $\mathcal{V}$  denotes the 3-D spatio-temporal volume. Corresponding to the moving surface  $S(t)$ , we will have a moving feature surface  $F(t) = \{[x(t) \ y(t) \ T(x(t), y(t), t)]^T: (x, y, t) \in \mathcal{V}\}$ .

Since the feature  $T$  is invariant to 3-D motion, it remains conserved along the trajectory of any specific moving point (i.e.,  $T$  for a moving point does not change with time). Hence,  $(dT/dt) = (\partial T/\partial x)(dx/dt) + (\partial T/\partial y)(dy/dt) + (\partial T/\partial t) = 0$ . Denoting partial derivatives with subscripts and  $u = (dx/dt)$  and  $v = (dy/dt)$ , we have

$$T_x u + T_y v + T_t = 0. \quad (5)$$

For every point of the spatio-temporal volume, we have an equation of the above form.

It should be noted that in case of a moving surface, the spherical neighborhood comprises of *points in the same time frame only*. The spherical nature of the neighborhood makes it invariant to 3-D rigid motion and hence ensures that corresponding sets of points from each frame are used for feature computation.

It can be seen that (5) is the exact 3-D analog of the optical flow constraint equation of Horn and Schunk [6]. While the latter expresses the conservation of optical intensity along the trajectory of a moving point, the former expresses the conservation of the trace of the feature matrix along the same. Accordingly, we will refer to (5) as the *3-D flow constraint equation*.

In practice, we will be given a sequence of *discrete* range images. In that case, our algorithm consists of the following steps.

*Step 0:* Obtain the 3-D Cartesian coordinates corresponding to every point in every image in the sequence.

*Step 1:* Compute the value of trace feature at every point, thus generating the feature surface  $T(x, y)$  for every image in the spatio-temporal (ST) volume.

There is a tradeoff involved in selecting the radius of the spherical neighborhood in feature computation. While a large radius is better from the viewpoint of noise and discontinuity elimination, too large a neighborhood will “oversmooth” the image, i.e., the local surface patches around neighboring points will become indistinguishable, so that the algorithm will estimate zero values of  $u, v$  there. In our experiments, the radius of spherical neighborhood in feature computation is 10 (the units of measurement for  $a$  and  $h$  are same as that for the coordinate system).

*Step 2:* Compute partial derivatives  $T_x, T_y$ , and  $T_t$  at every point in ST volume. Given a discrete surface, the planar neighborhood of any point  $p: [x_i \ y_i \ z_i]^T$  on it is the set of points  $\{[x_j \ y_j \ z_j]^T: (x_j - x_i)^2 + (y_j - y_i)^2 \leq c\}$  for some constant  $c$ . Corresponding to the planar neighborhood, we have a local discrete feature surface  $N = \{(x_i, y_i, T(x_i, y_i)); i \text{ varying over all the neighborhood points}\}$ .

<sup>2</sup>At any point, say  $p$ , on the surface, a sphere of specific radius, say  $a$ , is drawn, with  $p$  as center. The points on the surface that lie within the volume of this bounding sphere constitute the spherical neighborhood of  $p$ , denoted  $S_p$ . Thus,  $S_p$  is a portion of the surface on which any point  $q$  satisfies  $\text{dist}(p, q) \leq a$ .

We assume that the continuous feature surface corresponding to  $N$  is bicubic of the form  $T(x, y) = k_1 + k_2 x + k_3 y + k_4 x^2 + k_5 xy + k_6 y^2 + k_7 x^3 + k_8 x^2 y + k_9 xy^2 + k_{10} y^3$ . The value of the coefficients  $k_1$  through  $k_{10}$  are computed by least squares data fitting. From these values, the partial derivatives  $T_x$  and  $T_y$  can be easily computed. The partial derivative  $T_t$  is computed by the first difference between the  $T$  values at the same  $(x, y)$  location in the successive frames.

*Step 3:* Thus, at any point of the ST volume, we have a pair of equations (3-D flow constraint equation and elevation rate constraint equation)

$$T_x u + T_y v + 0w + T_t = 0 \quad (6)$$

$$Z_x u + Z_y v - w + Z_t = 0. \quad (7)$$

$u, v$ , and  $w$  denote respectively the  $x, y$ , and  $z$  components of 3-D velocity. Now consider a small spherical neighborhood  $\Omega$  of this point  $p$  (the radius of this spherical neighborhood is usually smaller than that used in Step 1). Since  $\Omega$  is a small neighborhood, we assume a uniform velocity over the neighborhood. This, by the way, enforces a measure of smoothness on the velocity field [7]. Thus every point in the neighborhood yields two equations in three unknowns. Assuming  $m$  points in the neighborhood, we have  $2m$  equations in three unknowns, which forms an overdetermined system.

The pair of equations contributed by a point  $q$  (i.e., the position vector  $\vec{r}_q \in \Omega$ ) is weighted by a “window function”  $W(\vec{r}_q)$ , which varies inversely as the distance of that point from the center of the neighborhood,  $p$  (position vector  $\vec{r}_p$ ), i.e.,  $W(\vec{r}_q) = (1/c_1 \|\vec{r}_p - \vec{r}_q\|)$ . Thus equations contributed by points further away from the center are weighed less compared to the closer ones. In our experiments,  $c_1$  was taken to be two in all cases.

Moreover, as pointed out earlier, 3-D flow constraint equations are far more robust with respect to noise and discontinuity compared to elevation rate constraint equations. Hence, the former are further weighted by a factor  $w_1$  (in all our experiments,  $w_1 = 1000$ ). Thus we have an overdetermined system of weighted linear equations over the neighborhood  $\Omega$ , from which  $u, v$ , and  $w$  are solved by singular value decomposition techniques [8].

#### A. Recovering Rigid Motion Parameters

The instantaneous angular velocity  $\vec{\omega} = [\omega_1 \ \omega_2 \ \omega_3]^T$  and the instantaneous translational velocity  $\vec{T} = [T_1 \ T_2 \ T_3]^T$  can be recovered from the calculated 3-D flow field. If  $\vec{v} = [u \ v \ w]^T$  be the computed velocity of a 3-D scene point  $\vec{r} = [X \ Y \ Z]^T$ , then  $\vec{v} = \vec{T} + \vec{\omega} \times \vec{r}$ . Writing the above equation in terms of its components

$$\begin{aligned} u &= T_1 + \omega_2 Z - \omega_3 Y & v &= T_2 + \omega_3 X - \omega_1 Z \\ w &= T_3 + \omega_1 Y - \omega_2 X. \end{aligned} \quad (8)$$

This set of equations has six unknowns,  $T_1, T_2, T_3, \omega_1, \omega_2, \omega_3$ . Each point  $(X, Y, Z)$  for which a reliable estimate for  $u, v$ , and  $w$  exists, will yield three equations like (8). All these equations together form a heavily overdetermined system, from which the six unknowns can be solved. The (un)reliability of the  $u, v, w$  estimate obtained at a point  $\vec{r}$  is given by  $w_1(T_x u + T_y v + T_t)^2 + (Z_x u + Z_y v + Z_t - w)^2$ , where  $w_1$  is the weight associated with the 3-D flow constraint equations as described in the last section. Thus, while computing  $\vec{\omega}$  and  $\vec{T}$ , we discard any point whose unreliability is larger than a threshold.

#### IV. NOISE ANALYSIS

In this section we will analyze the noise properties of the trace feature denoted in (3). We assume that the  $x, y$ , and  $z$  components of the noise vector are uncorrelated. The expression is symmetric with respect to the  $x, y$ , and  $z$  axes, i.e., the three terms  $(x_i - \mu^x)^2$ ,

$(y_i - \mu^y)^2$ ,  $(z_i - \mu^z)^2$  are identical in form. Accordingly, we will carry out our analysis on a generic term of the same form. Let  $F = \{f_i: i = 1 \cdots N\}$  denote a set of discrete values. Let  $n_i$  denote the (zero-mean white) noise function, added to the true value  $f_i$  to produce the (noisy) observed quantity  $\tilde{f}_i = f_i + n_i$ . Let the variance of the noise,  $\text{var}(n_i)$  be  $\sigma_n^2$ . We will analyze the effect of noise on the quantity  $\gamma = (1/N) \sum_{i=1}^N (f_i - \mu_f)^2$ .

Now,  $\tilde{f}_i$  itself is another Random Variable. Since the noise is zero-mean (i.e.,  $E(n_i) = 0$ ), its expected value  $E(\tilde{f}_i)$  is equal to  $f_i$  and variance  $\text{var}(\tilde{f}_i) = \sigma_n^2$ .

Consider the center of mass  $\mu_f = (1/N) \sum_{i=1}^N f_i$  and the corresponding observed quantity  $\tilde{\mu}_f = (1/N) \sum_{i=1}^N \tilde{f}_i$ . Using well known properties of linear combinations of random variables, we have the expected value of the center of mass,  $E(\tilde{\mu}_f) = \mu_f$ . Similarly, we can compute the variance  $\text{var}(\tilde{\mu}_f) = \text{var}((1/N) \sum_{i=1}^N \tilde{f}_i) = (1/N^2) \sum_{i=1}^N \text{var}(\tilde{f}_i) = (1/N^2) \sum_{i=1}^N \sigma_n^2 = (\sigma_n^2/N)$ . Thus,  $\text{var}(\tilde{\mu}_f) \rightarrow 0$  as  $N \rightarrow \infty$ . Hence, by Tchebycheff inequality, the value of the random variable  $\tilde{\mu}_f$  approaches a deterministic value equal to its expected value  $\mu_f$  as  $N$  becomes large. In other words, the effect of noise is negligible on  $\mu_f$  for large values of  $N$ .

Now, let us consider the quantity  $\gamma$  and its observed value  $\tilde{\gamma} = (1/N) \sum_{i=1}^N (\tilde{f}_i - \tilde{\mu}_f)^2$ . For large  $N$ ,  $\tilde{\mu}_f = \mu_f$ , and  $\tilde{\gamma} = (1/N) \sum_{i=1}^N (f_i + n_i - \mu_f)^2$ . It can be easily shown that the expected value,  $E(\tilde{\gamma}) = \gamma + \sigma_n^2$  and the variance  $\text{var}(\tilde{\gamma}) = (4/N^2) \sum_{i=1}^N (f_i - \mu_f)^2 \sigma_n^2 + (2\sigma_n^4/N)$ . Assuming,  $(1/N) \sum_{i=1}^N (f_i - \mu_f)^2$  remains finite, (i.e., the set  $F$  is bounded)  $\text{var}(\tilde{\gamma}) \rightarrow 0$  as  $N \rightarrow \infty$ . Thus, using Tchebycheff's inequality, for large  $N$ , the random quantity  $\tilde{\gamma}$  approaches a deterministic value equal to its expected value  $\gamma + \sigma_n^2$ , i.e., its true value plus a constant equal to the variance of the noise.

Since the trace feature is a sum of three terms of the form of  $\gamma$ , the cumulative effect of noise on it is the addition of a constant value. Since in this algorithm, we always use the derivative of the trace feature, rather than the feature itself, this constant term contributed by noise cancels out. In other words, (5) is expected to be quite robust with respect to noise if the number of points in the neighborhood used to compute the trace feature is large.

## V. DISCONTINUITY ANALYSIS

In this section, we analyze the behavior of the trace feature in the presence of a step discontinuity in the moving surface. Fig. 1 shows a surface which has such a discontinuity. Without loss of generality, we can assume that the coordinate axes are aligned as shown in Fig. 1 (the constant  $Z$  lines are parallel to the  $Y$  axis). Since there is no variation along the  $Y$  axis, we will drop it from further analysis. This will allow us to carry out our analyses in 2-D. The corresponding 2-D situation is shown in Fig. 2.

The equation of the 2-D curve with step discontinuity (Fig. 2) can be written as

$$Z(x) = \begin{cases} -\frac{h}{2} + \alpha x & x \leq 0 \\ \frac{h}{2} + \alpha x & x > 0. \end{cases}$$

Thus,  $\lim_{x \rightarrow 0^-} Z \neq \lim_{x \rightarrow 0^+} Z$  and the derivative of  $Z$  with respect to  $x$  at  $x = 0$  is unbounded. Let us analyze the nature of the proposed trace feature in the vicinity of the discontinuity. In order to compute the trace feature, we need to consider a spherical neighborhood which, in the 2-D case reduces to a circular neighborhood. In the following discussion,  $a$  denotes the radius of the bounding circle and  $h$  denotes the size (jump) of the discontinuity. Let  $\vec{p}(x) = (x, Z(x))$  be the point at which we are evaluating our trace feature. The corresponding domain of integration is  $A(x)$ , which stands for the set of points  $(s, Z(s))$  such that  $\text{dist}(\vec{p}(x), \vec{p}(s)) \leq a$ ,

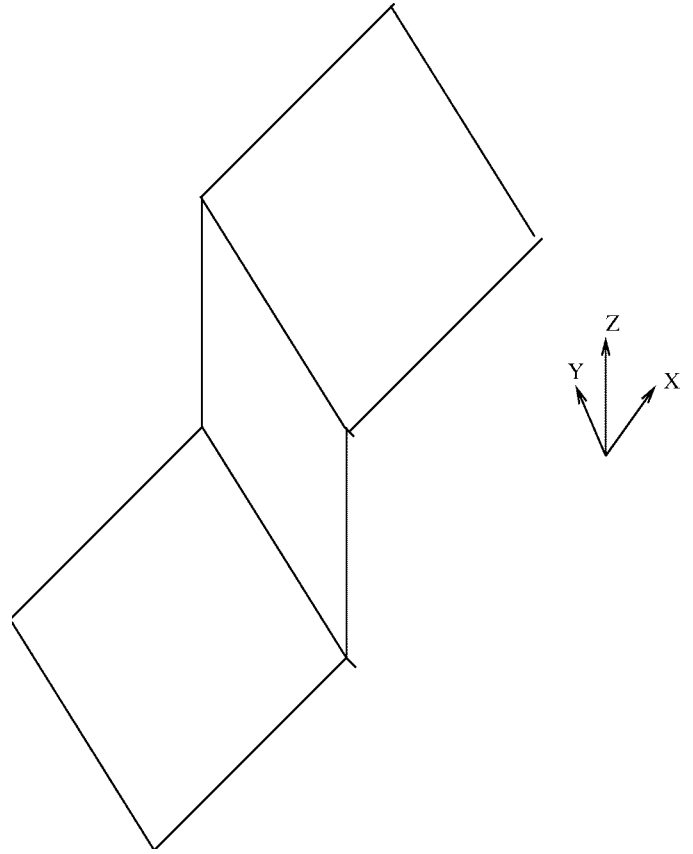


Fig. 1. Generalized step discontinuity in 3-D.

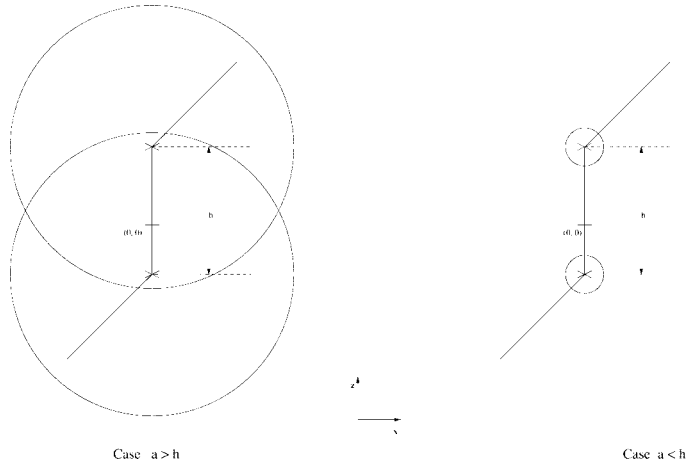


Fig. 2. Simplified 2-D diagram.

i.e.,

$$A(x) = \{s: (s - x)^2 + (Z(s) - Z(x))^2 \leq a^2\}.$$

Denoting  $\|A(x)\| = \int_{A(x)} ds$ , we have the center of mass

$$\begin{aligned} \vec{\mu}(x) &= [\mu_1(x) \mu_2(x)]^T \\ &= \frac{1}{\|A(x)\|} \left[ \int_{A(x)} s ds \int_{A(x)} Z(s) ds \right]^T \end{aligned} \quad (9)$$

$$f(x) = \frac{1}{\|A(x)\|} \int_{A(x)} ((s - \mu_1(x))^2 + (Z(s) - Z(x))^2) ds. \quad (10)$$

Two cases may arise, case 1 is  $a > h$  and case 2 is  $a \leq h$ .

Consider case 1 first. We will evaluate the trace feature  $f(x)$  at two points lying on two sides of the discontinuity but arbitrarily close to it, i.e., we will evaluate  $f(0-)$  and  $f(0+)$  and show them to be equal. Since  $\lim_{x \rightarrow 0-} Z(x) = -(h/2)$ , we have  $\lim_{x \rightarrow 0-} A(x) = A(0-) = \{s: (s)^2 + (Z(s) + (h/2))^2 \leq a^2\}$ . Breaking this domain into two portions corresponding to  $s \leq 0$  and  $s > 0$  we get

$$A(0-) = \begin{cases} \left\{ s: s^2 + \left(-\frac{h}{2} + \alpha s + \frac{h}{2}\right)^2 - a^2 \leq 0 \right\} & s \leq 0 \\ \left\{ s: s^2 + \left(\frac{h}{2} + \alpha s + \frac{h}{2}\right)^2 - a^2 \leq 0 \right\} & s > 0. \end{cases}$$

The first domain ( $s \leq 0$ ) yields  $\{-p \leq s\}$  where  $p = (a/\sqrt{1+\alpha^2})$ . Factorizing the quadratic term, the inequality on the second domain can be written as

$$\left( s + \frac{\sqrt{\alpha^2 h^2 + (a^2 - h^2)(1 + \alpha^2)} + \alpha h}{1 + \alpha^2} \right) \left( s - \frac{\sqrt{\alpha^2 h^2 + (a^2 - h^2)(1 + \alpha^2)} - \alpha h}{1 + \alpha^2} \right) \leq 0.$$

The first factor is  $>0$  for  $s > 0$ . Hence for the product to be  $\leq 0$ , the second factor has to  $\leq 0$ . Hence the second domain becomes  $s \leq q$  where  $q = (\sqrt{\alpha^2 h^2 + (a^2 - h^2)(1 + \alpha^2)} - \alpha h / 1 + \alpha^2)$ . It should be noted that for  $a > h$  the term under the root sign as well as  $q$  are nonnegative quantities. Summarizing, we have  $A(0-) = \{s: -p \leq s \leq q\}$  where  $p$  and  $q$  are shown above. Hence,  $\|A(0-)\| = p + q$ . Substituting these values in (9) and (10) we get

$$\mu(0-) = \left[ \frac{q-p}{2} \left( \frac{h}{2} - \frac{ph}{q+p} + \frac{\alpha}{2}(q-p) \right) \right]^T$$

$$f(0-) = (1 + \alpha^2) \frac{(q+p)^2}{12} + \frac{\alpha p q h}{q+p} + \frac{p q h^2}{(p+q)^2}.$$

Proceeding in a similar fashion, it can be shown that  $A(0+) = \{s: -q \leq s \leq p\}$  and  $\|A(0+)\| = p + q$ . Substituting in (9) and (10) we get

$$\mu(0+) = \left[ \frac{(p-q)}{2} \left( \frac{h}{2} - \frac{qh}{q+p} + \frac{\alpha}{2}(p-q) \right) \right]^T$$

$$f(0+) = (1 + \alpha^2) \frac{(q+p)^2}{12} + \frac{\alpha p q h}{q+p} + \frac{p q h^2}{(p+q)^2}.$$

Thus, we have shown that  $f(0+) = f(0-)$ , meaning  $f(x)$  is continuous at  $x = 0$ , even though  $Z(x)$  is not.

In order to provide a feel for the behavior of the trace feature near the point of discontinuity, the expressions for the above quantities as general functions of  $x$  (in the special case of  $\alpha = 0$ ) are as follows:

$$\mu(x-) = \left[ \left( x - \frac{1}{2}(a - \sqrt{a^2 - h^2}) \right) \cdot \left( \frac{h}{2} + \frac{h(x-a)}{a + \sqrt{a^2 - h^2}} \right) \right]^T$$

$$f(x-) = \frac{1}{12}(a + \sqrt{a^2 - h^2})^2 - \frac{h^2(x-a)(x + \sqrt{a^2 - h^2})}{(a + \sqrt{a^2 - h^2})^2}$$

$$\mu(x+) = \left[ \left( x + \frac{1}{2}(a - \sqrt{a^2 - h^2}) \right) \cdot \left( \frac{h}{2} + \frac{h(x - \sqrt{a^2 - h^2})}{a + \sqrt{a^2 - h^2}} \right) \right]^T$$

$$f(x+) = \frac{1}{12}(a + \sqrt{a^2 - h^2})^2 - \frac{h^2(x+a)(x - \sqrt{a^2 - h^2})}{(a + \sqrt{a^2 - h^2})^2}.$$

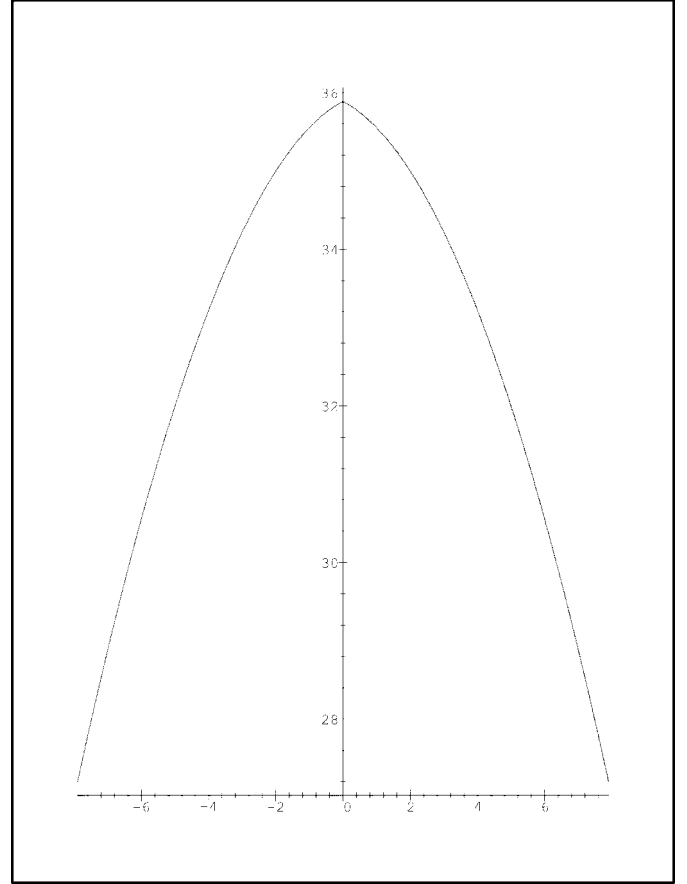


Fig. 3. Trace feature  $f(x)$  with  $Z(x)$  discontinuous at  $x = 0, \alpha = 0, a = 10, h = 6$ .

A plot of  $f(x)$  on two sides of the discontinuity is shown in Fig. 3 for  $a = 10$  and  $h = 6$ . It should be noted that as  $x \rightarrow 0$ , more points are included from the other side of the discontinuity. Since, these points are further away from the center of mass, their contributions to the trace feature are larger. Hence, the overall value of the trace feature is higher near the point of discontinuity.

In case 2 ( $a < h$ ), the domain of integration includes only one side of the discontinuity. In other words, points from only one side of the discontinuity are involved in computation of the center of mass and the trace feature. Consequently, the discontinuity is not smoothed out. Thus it has been shown that the proposed technique smooths out the symmetric discontinuities whose size is smaller than the radius of the bounding sphere.

## VI. IMPLEMENTATION AND RESULTS

The proposed technique has been implemented on a Sun Sparcstation 2 in C language. In order to study the effect of noise on the conservation of the trace feature, we took a set of 3-D points,  $S_1 = \{\vec{r}_{i1}: i \in [1, N]\}$  and imparted a known rigid motion  $(\mathbf{R}, T)$  to obtain the transformed set  $S_2 = \{\vec{r}_{i2}: i \in [1, N]\}$ , where  $\vec{r}_{i2} = \mathbf{R}\vec{r}_{i1} + T$ . White Gaussian noise was synthetically added to the points in  $S_2$ . The trace feature was computed at every point in  $S_1$  and  $S_2$  using a radius of spherical neighborhood  $r$ . Thus, two sets of trace features were obtained,  $\mathcal{T} = \{T_j: j \in [1, N]\}$  and  $\mathcal{T}' = \{T'_j: j \in [1, N]\}$ . Ideally,  $T_j$  should equal  $T'_j$ , but that would not be true in the presence of noise. The corresponding relative error measure  $(1/N)\sum_{j=1}^N ((T_j - T'_j)^2 / T_j^2)$  is denoted as the mean square conservation error (MSCE). The MSCE was computed at

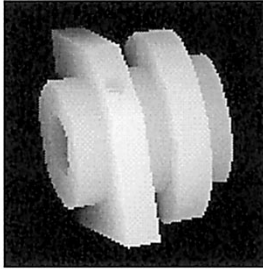


Fig. 4. The industrial part image.

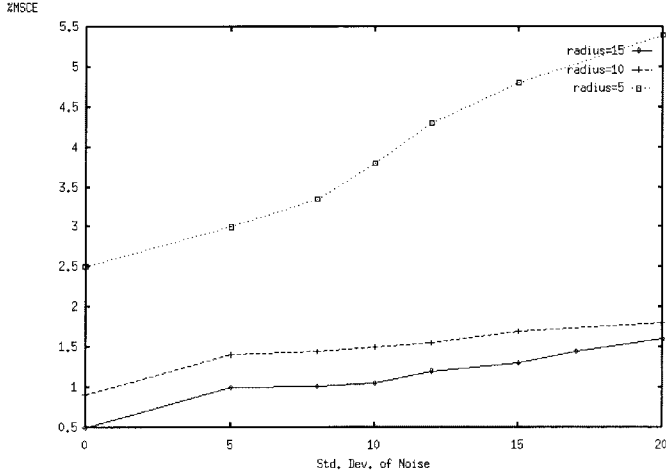


Fig. 5. Percentage mean square conservation error.

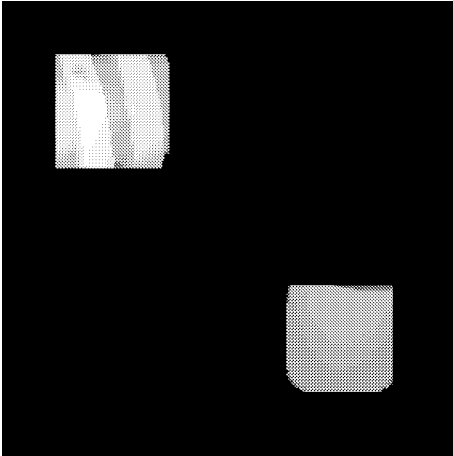


Fig. 6. The multiple moving object image.

different levels (i.e., standard deviations) of noise and different values of  $r$ . The results are presented graphically in Fig. 5. As expected, the error increases with the standard deviation of the noise. For a given standard deviation, the MSCE decreases with increasing  $r$ , corroborating the analysis in Section IV. For all these experiments,  $\|S_1\| = \|S_2\| = 512 \times 512$ .

In order to benchmark the proposed approach against that of Horn and Harris, numerous experiments were conducted with known ground truths, at different levels of synthetic noise. In each case, the angular and translation velocities were estimated using the method outlined in Section III-A. In one of these experiments, a complex

TABLE I  
EXPERIMENT SET 1

Std. Dev. of Noise	Proposed Approach % error $\ \vec{\omega}\ $ , $\ \vec{T}\ $	Horn and Harris Approach % error $\ \vec{\omega}\ $ , $\ \vec{T}\ $
0	5.04, 1.49	8.04, 1.94
5	6.06, 2.08	10.41, 4.59
8	7.19, 3.31	11.47, 6.32
10	8.45, 4.48	12.38, 7.64

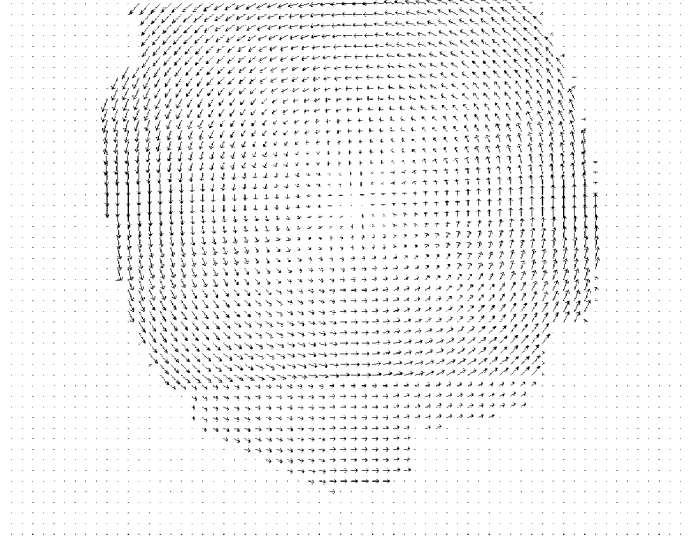


Fig. 7. X,Y components of the estimated velocity field for the industrial part image sequence.

motion (0.02 rad about an axis parallel to the  $Z$  axis passing through the center with a translation of 0.5 units along  $Z$  axis) was imparted to the industrial parts image shown in Fig. 4.<sup>3</sup> The *percentage* errors in angular and translational velocities for the proposed as well as Horn and Harris approach has been provided in tabular form in Table I. The accuracy of the proposed approach was greater in each case. In order to provide a visual feeling for the performance of the proposed approach, the vector needle diagram corresponding to  $X,Y$  components of the estimated velocity field (0 noise case) is also provided in Fig. 7.

The second experiment demonstrates the capability of the proposed algorithm to handle multiple motions. The input image was prepared by cutting two square patches from different images, and pasting them at opposite corners of a  $256 \times 256$  black background. The top patch was given a rotation of 0.1 rad about an axis parallel to the  $Z$  axis, passing through middle of the right border, along with a translation of 0.6 units along  $Z$  axis. The bottom patch was given a rotation of  $-0.1$  rad about an axis parallel to  $Z$  axis, passing through the middle of the left border, along with a translation of  $-0.6$  units along  $Z$  axis. The needle diagrams corresponding to  $X,Y$  components of the estimated velocity field (0 noise case) are shown in Fig. 8.

## VII. CONCLUSIONS

A technique for computing the 3-D velocity field from a sequence of range images is presented. The technique utilizes a new integral 3-D rigid motion invariant feature, whose conservation along the

<sup>3</sup>These range images were obtained from University of Michigan Artificial Intelligence Laboratory, the original source being the Environmental Research Institute of Michigan.

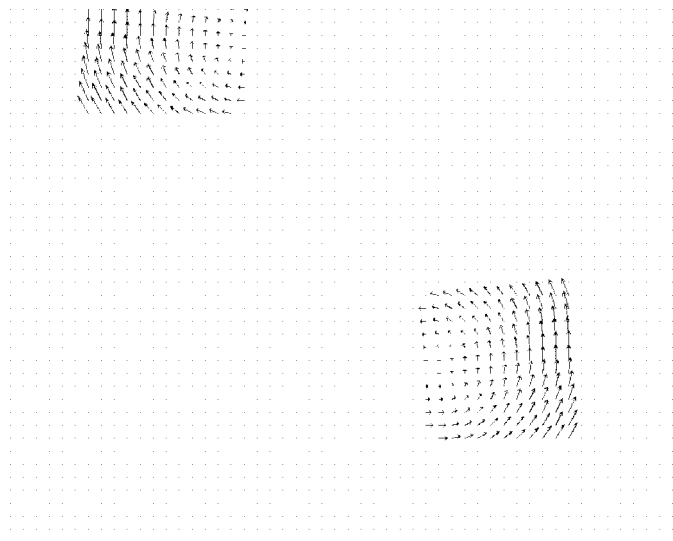


Fig. 8.  $X, Y$  components of the estimated velocity field for the multiple object image sequence.

trajectory of a moving point forms the basis of a 3-D flow constraint equation. This equation is the 3-D analog of the well known optical flow constraint equation. The velocity components at every point is computed by solving an overdetermined system of equations obtained from the neighborhood of the point. Theoretical analyses are presented to establish the noise and discontinuity robustness of the proposed 3-D flow constraint equation. Results of the experimental performance evaluation of the proposed technique support the theoretical analysis results.

The authors are currently investigating the application of the proposed invariant to correspondence based 3-D motion detection. The proposed invariant has potential applications in range image segmentation and analysis. Further research is necessary.

#### REFERENCES

- [1] K. Arun, T. Huang, and S. Blostein, "Least square fitting of two 3-D point sets," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-9, pp. 698–700, 1987.
- [2] H. Goldstein, *Classical Mechanics*. Reading, MA: Addison-Wesley, 1990.
- [3] R. S. Millman and G. D. Parker, *Elements of Differential Geometry*. Englewood Cliffs, NJ: Prentice-Hall, 1977.
- [4] B. K. P. Horn, E. Hildreth, and S. Negahdaripour, "Closed-form solution of absolute orientation using orthonormal matrices," *J. Opt. Soc. Amer.*, vol. 5, pp. 1127–1135, 1988.
- [5] B. K. P. Horn and J. Harris, "Rigid body motion from range image sequences," *CVGIP: Image Understanding*, vol. 53, no. 1, pp. 1–13, 1991.
- [6] B. K. P. Horn and B. G. Schunk, "Determining optical flow," *Artif. Intell.*, vol. 17, pp. 185–203, 1981.
- [7] B. Lucas, Ph.D. dissertation, Dept. Comput. Sci., Carnegie Mellon University, Pittsburgh, PA, 1984.
- [8] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C*. Cambridge, U.K.: Cambridge Univ. Press, 1988.
- [9] B. Sabata and J. K. Agarwal, "Estimation of motion from a pair of range images," *CVGIP: Image Understanding*, vol. 54, no. 3, pp. 309–324, 1991.
- [10] G. Strang, *Linear Algebra and Its Applications*. New York: Academic, 1980.

## An Intelligent Mobile Vehicle Navigator Based on Fuzzy Logic and Reinforcement Learning

Nelson H. C. Yung and Cang Ye

**Abstract**—In this paper, an alternative training approach to the EEM-based training method is presented and a fuzzy reactive navigation architecture is described. The new training method is 270 times faster in learning speed; and is only 4% of the learning cost of the EEM method. It also has very reliable convergence of learning; very high number of learned rules (98.8%); and high adaptability. Using the rule base learned from the new method, the proposed fuzzy reactive navigator fuses the obstacle avoidance behavior and goal seeking behavior to determine its control actions, where adaptability is achieved with the aid of an environment evaluator. A comparison of this navigator using the rule bases obtained from the new training method and the EEM method, shows that the new navigator guarantees a solution and its solution is more acceptable.

**Index Terms**—Behavior fusion, fuzzy logic, goal seeking, neural network, obstacle avoidance, reinforcement learning, vehicle navigation.

#### I. INTRODUCTION

The navigation of a mobile vehicle can be considered as a task of determining a *collision free path* that enables the vehicle to travel through an obstacle course from an initial configuration to a goal configuration. The process of finding such path is also known as the *path planning problem* which could be classified into: global path planning and local path planning. *Global path planning* methods are usually conducted off-line in a completely known environment. Many attempts at solving this problem have been tried [1], where an *exact environment model* has been used for planning the path. Although these approaches have an exact solution, their time complexity grows with the geometry complexity and grows exponentially with the number of degrees of freedom in the vehicle's motion [2]. Thus they are only practical when the environment model is simple and the number of degrees of freedom is reasonably low. However, real environments are never simple enough. This fact has led to the emergence of numerous *heuristic approaches*, which rely on either calculating the *potential fields* [3] or performing a search through a *state space model* [4]. In general, these methods trade reliability for speed, in which they do not guarantee a solution even if there exists one. Worse still, all these approaches fail when the environment is not fully known.

On the other hand, the *local path planning* techniques, also known as the obstacle avoidance methods, are potentially more efficient in vehicle navigation when the environment is unknown or only partially known. It utilizes the on-line information provided by sensors such as the ultrasonic sensor, laser range finder, radar, and vision sensor, to tackle the uncertainty. An efficient local path planning method is the potential field method, which was first proposed by Khatib [3] and has been widely used in obstacle avoidance cases [5]. In spite of its simplicity and elegance, this method has three problems: First, local minimum could occur and cause the vehicle to be stuck; second, it

Manuscript received March 4, 1997; revised January 15, 1998. This work was supported by the CRCG of The University of Hong Kong under Grant 337/062/0016. This paper was recommended by Associate Editor A. Kandel.

The authors are with the Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong (e-mail: nyung@eee.hku.hk). Publisher Item Identifier S 1083-4419(99)00910-3.