논문 96-5-4-08

Motion Estimation of 3D Planar Objects using Multi-Sensor Data Fusion

Woo Suk Yang

센서 융합을 이용한 움직이는 물체의 동작예측에 관한 연구

梁 雨 錫

Abstract

Motion can be estimated continuously from each sensor through the analysis of the instantaneous states of an object. This paper is aimed to introduce a method to estimate the gleneral 3D motion of a planar object from the instantaneous states of an object using multi-sensor data fusion.

The instantaneous states of an object is estimated using the linear feedback estimation algorithm. The motion estimated from each sensor is fused to provide more accurate and reliable information about the motion of an unknown planar object. We present a fusion algorithm which combines averaging and deciding. With the assumption that the motion is smooth, the approach can handle the data sequences from multiple sensors with different sampling times. Simulation results show proposed algorithm is advantageous in terms of accuracy, speed, and versatility.

Key Word : Multisensor, Data Fusion, Dynamic Environment, Motion Estimation, 3D Object

1. Introduction

Recently, there has been a growing interest in upgrading robot intelligence through the integration of multiple sensors. Sensor data fusion is important in improving a complex robot's performance by providing hybrid information, by reducing sensor errors, and by maximizing sensor use. A comprehensive survey of multisensor data fusion is published by Luo and Kay[15]. Although many good individual ideas have been explored for the fusion of multiple sensors, several key issues still remain unresolved. One of the key issues is the data fusion under time evolving conditions.

The time evolving condition involves the prediction of the future states of moving objects. The estimation of the parameters of 3D motion makes it possible to predict the location and configuration of the moving objects at the next stage. However, It is not easy to estimate the motion parameters directly from sensor data, because the time evolving problem suffers from the lack of a priori knowledge about the underlying dynamics which govern the motion as well as the structure of the objects. With some

홍익대학교 전기공학과
(Dept. Electrical Engineering, Hong Ik Univ.)
※ 이 논문은 1994년도 한국학술재단의 공모과제 연구비에 의하여 연구되었음.
<접수일자 : 1996년 3월 30일>

constraints, the problem of motion estimation becomes easier to solve. Most of the constraints are made on the allowed motion, the underlying dynamics[5,11], and/or object structure[19]. Some researchers[5,11] require the known principal moments of inertia and the structures of objects. Weng et al.[22] consider the motion of an object in which at least two elements of a diagonalized inertia tensor are the same. Motion can also be analyzed without any constraints on the object's structure and/or dynamics from the time-sequential images using point correspondence or line correspondence[13,14]. An alternative approach is to estimate the motion parameters using optical flow[1], but this approach can not be used to estimate the general 3D motion.

Another difficulty in time evolving condition is a sensor synchronization problem. If the object properties being observed are varying with respect to time, the sensor measurements with different sensing times can not be used directly for fusion. To fuse the multiple sensor readings from different sensors in a dynamic environment, we need to consider different sampling time of each sensor. One of the approaches to exploring the dynamic data fusion is the use of a Kalman filter. If the environment can be described by a linear model and sensor error can be modeled as Gaussian, the Kalman filter provides unique statistically optimal estimates for the fused data. Quadratic moving curve fitting and weighted least square error estimation[18,19] are also discussed under time-evolving conditions. Another approach [16] is to model the time varying properties of an environment. If it is possible to provide a certain way to represent the time varying properties of an object as a function of time, the problem of data fusion for a moving object can be solved by applying well known static fusion algorithms to the motion parameters[15,16].

Our approach is based on the analysis of corresponding surface patches from the observation of consecutive images. Given data sets of a moving object taken at consecutive time instants, we estimate the instantaneous positions and orientations of an object. General object motion is modeled using the Taylor expansion technique. The motion parameters of an object model can be considered as constants during a small time intervals. The sequences of the estimated instantaneous positions and orientations are used to determine the motion parameters. These parameters are later to be fused using a general static fusion algorithm. In this paper, we focus on the estimation of general 3D motion of a planar object with no a priori knowledge of environment, and provide a method to specify an uncertainty of the estimated motion. We use a sequence of instantaneous data from multiple sensors with different sampling times. It is assumed that the correspondence between object surfaces is resolved, and that the parameters of object surfaces are already estimated with some well-defined procedure so that the computational error bound can be obtained.

2. object model

An object in the environment can be assumed to be a convex polyhedron (or simplified by convex polyhedron) in a three dimensional Cartesian space $\mathbf{R}^3$ such as

$$S = \{ x \mid A \cdot x \le c \} \qquad (1)$$

For an n-faced polyhedral object, $A = [\ a_1 : a_2 : \cdots : a_n\ ]^T$ is an $n \times 3$ matrix and $c = (\ c_1 : c_2 : \ . \ . \ . : c_n)^T$ is an $n \times 1$ vector. The i^{th} face is described by $a_i^T \cdot x = c_i$.

Equation (1) describes an object model which is used as a basis for estimating the object motion. The object parameters, A and c, can be obtained from the observation at any time instant

t_0. The object observed at time t, $t > t_0$, is described as

$$S(t) = \{x(t) \mid F(t) \cdot x(t) \le g(t)\} \quad (2)$$

where $F(t) = [\, f_1 : f_2 : \ldots : f_m \,]^T$ is an $m \times 3$ matrix, and $g(t) = [\, g_1 : g_2 : \ldots : g_m \,]^T$ is an $m \times 1$ vector. The integer m is the number of object faces observed. The i^{th} row vector of $F(t)$, f_i, and the i^{th} element of $g(t)$, g_i, specify the i^{th} face of an object. as

$$f_i^T(t) \cdot x(t) = g_i(t)$$

It is obvious that $n \ge m$. Without the loss of generality, let us discard the row terms corresponding to unobserved faces from A and c of equation (1), and let A and c be an $m \times 3$ reduced matrix and an $m \times 1$ reduced vector respectively such that a_i and c_i will be the parameters of the face of a reference object corresponding to the face of the observed object with parameters f_i and g_i. An object motion can be described by

$$x(t) = R(t)\, x(t_0) + d(t) \quad (3)$$

where $R(t)$ is a 3×3 orthogonal matrix which describes the rotation of an object. $d(t)$ is a 3×1 vector which describes the translation between two objects. It is clear that $R(t_0) = I$ and $d(t_0) = 0$. From the equation (3) and (2), we obtain

$$F(t)R(t)x(t_0) \le g(t) - F(t)d(t) \quad (4)$$

Comparing equations (1) and (4), we get

$$\{F(t) + \delta F(t)\} R(t) = H(t)\{A + \delta A\} \quad (5)$$

$$\{g(t) + \delta g(t)\} - \{F(t) + \delta F(t)\} d(t) = H(t)\{c + \delta c\} \quad (6)$$

where $\delta A = \delta F(t)$, $\delta c = \delta g(t_0)$. $H(t)$ is a $m \times m$ time varying diagonal matrix whose i^{th} diagonal element describes a dummy variable for corresponding the surface equation of the i^{th} object face observed at time t to that of the i^{th} face of the reference object. The $m \times 3$ matrix, $\delta F(t)$, and an $m \times 1$ vector, $\delta d(t)$, denote the observation errors introduced during the computation of parameters $F(t)$ and $g(t)$. An $m \times 3$ matrix, $\Delta_F(t)$, and an $m \times 1$ vector, $\varepsilon_g(t)$, can be specified to bound the computation errors $\delta F(t)$ and $\delta d(t)$ such that

$$|\,\delta F(t)\,| \le \Delta_F(t) \text{ and } |\,\delta g(t)\,| \le \varepsilon_g(t) \quad (7)$$

where $|\,P\,| \le Q$ is defined such that the absolute value of an element of P is less than or equal to the corresponding element of Q. The values of $\Delta_F(t)$ and $\varepsilon_g(t)$ do not represent the absolute values of possible maximum errors. They rather describe the maximum possible errors under which a certain uncertainty is guaranteed to the observed data.

3. Estimation of the Orientation of an object

There are many approaches to estimating the orientation of an object. If there are many samples, least square error estimation(LSEE) is popular. However, LSEE may not be suitable in our application because object faces are a rather small number of samples. Moreover, LSEE does not always guarantee the best estimate in some circumstances. Thus, instead of estimating the best orientation, we try to compute the orientation of an object ignoring the existence of $\delta F(t)$ and $\delta d(t)$, then, consider the worst case and estimate the maximum value of the possible error bounds from the solution. These error bounds are used as an uncertainty measure for the orientation computed.

3.1 Estimation of the Dummy Variable H

In order to estimate the dummy variable $H(t)$ in equation (5), We first ignore the error term δF and δd, and consider FR = HA. Let h_{ii} be the i^{th} diagonal element of H. h_{ii} is the multiplier for matching the surface equation of the i^{th} surface of an object at time t to the equation of the corresponding surface at time t_0. For the i^{th} surface, we have

$$f_i^T \cdot R = h_{ii} a_i^T \quad (8)$$

Since R is orthogonal, we have $\| f_i \|_2 = \| f_i^T \cdot R \|_2$ where $\| f_i \|_2$ represents the 2-norm of f_i such that $\| f_i \|_2 = (f_i^T f_i)^{1/2}$. The absolute value of h_{ii} is computed as follows.

$$h_{ii} = \left[\frac{f_i^T \cdot f_i}{a_i^T \cdot a_i} \right]^{1/2} \quad (9)$$

The sign of h_{ii} is determined by the sign of $f_i^T \cdot a_i$.

Let an $m \times m$ diagonal matrix δH represent the computation error of H due to the ignoring of δF, and δh_{ii} be the i^{th} diagonal element. Then,

$$(f_i + \delta f_i)^T \cdot R = (h_{ii} + \delta h_{ii})(a_i + \delta a_i)^T \quad (10)$$

where δf_i^T and δa_i^T represent the i^{th} row of F and A respectively. Since R is orthogonal,

$$(f_i + \delta f_i)^T \cdot (f_i + \delta f_i) = (h_{ii} + \delta h_{ii})^2 (a_i + \delta a_i)^T \cdot (a_i + \delta a_i) \quad (11)$$

Let us neglect the second order error terms. Since $a_i^T \cdot a_i$ can not be zero, the maximum error bound of $| \delta h_{ii} |$ is obtained as

$$| \delta h_{ii} | \leq \frac{1}{(a_i^T \cdot a_i)^{1/2} (f_i^T \cdot f_i)^{1/2}} \left\{ | f_i^T | \cdot (\Delta_F)_i^r + \frac{f_i^T \cdot f_i}{a_i^T \cdot a_i} | a_i^T | \cdot (\Delta_A)_i^r \right\}$$

$$= (\Delta_H)_{ii} \leq \frac{\| (\Delta_F)_i^r \|_2}{\| a_i \|_2} + \frac{f_i^T \cdot f_i}{a_i^T \cdot a_i} \frac{\| (\Delta_A)_i^r \|_2}{\| f_i \|_2} \quad (12)$$

Now, we have the following equation.

$$(F + \delta F) R = B + \delta B \quad (13)$$

where $B = HA$ and $\delta B = \delta HA + \delta AH$. Let Δ_H be a diagonal matrix whose i^{th} diagonal element is $(\Delta_H)_{ii}$. Then, the bounds Δ_B of error δB becomes

$$| \delta B | \leq \Delta_H | A | + \Delta_A | H | = \Delta_B \quad (14)$$

The error bounds Δ_F of $\boldsymbol{\delta F}$ is specified by equation (7).

3.2 Estimation of the Orientation of an Object

The orientation R of an object is computed from equation (13) with the consideration of equations (7) and (14). Equation (13) is always consistent mathematically. However, δF and δB, the observation errors, are not known. Moreover, the number of faces observed from a sensor may not be enough to guarantee the unique solution of R. We solve the following perturbed equation by first ignoring the observation errors $\boldsymbol{\delta F}$ and δB. We then estimate the error bounds of the result as its uncertainty measure. By ignoring δF and δB terms from equation (13), and by multiplying both sides with F^T, we have the following consistent associated system of a normal equation.

$$F^T F \hat{R} = F^T B \quad (15)$$

where F and B are $m \times 3$ matrices. Assuming that we can observe at least two faces in the worst case, the matrix F has a rank of at least two. The solution $\hat{R}$ has a unique decomposition as the sum of two terms such that

$$\hat{R} = R_1 + R_2 \qquad (16)$$

where the column vector of R_1 is in the range space of F^T, $R(F^T)$, and R_2 is in the null space of F, $N(F)$. The solution of LSEE represents the best estimate of the projection of x onto $R(F^T)$. That means it can only provide the R_1 information, but does not provide any information about R_2. By using the pseudoinverse theory, R_1 becomes

$$R_1 = F^+ B \qquad (17)$$

where F^+ is computed from

$$F^+ = Q^T L^{-1/2} P^T \qquad (18)$$

The matrices P and Q are obtained using the diagonalization of matrices FF^T and F^TF. Since FF^T and F^TF are symmetric positive semidefinite, all of their eigenvalues are non-negative real numbers, and the eigenvectors are orthogonal to each other. Let

$$FF^T = TD_1T^T \quad \text{and} \quad F^TF = S^TD_2S$$

D_1 and D_2 are $n \times n$ and $m \times m$ diagonal matrices whose first r diagonal terms, λ_i, $1 \le i \le r$, are nonzero. Let us express T and S as

$$T = [P : P_0], \quad \text{and} \quad S^T = [Q^T : Q_0^T] \qquad (19)$$

where the column vectors of P and Q^T are the eigenvectors of FF^T and F^TF respectively corresponding to the nonzero eigenvalues λ_i, and the column vectors of P_o and Q_0^T are the eigenvectors of FF^T and F^TF corresponding to the eigenvalues of zero value. Then,

$$L = diag(\lambda_1, \lambda_2, \ldots, \lambda_\gamma) \qquad (20)$$

$$Q = L^{-1/2} P^T F \qquad (21)$$

If the rank of F is three, Q_0 is a null matrix. For F of rank two, Q_0 becomes a single vector. Since $R(I - F^+F) = R(Q_0)$,

$$r_i^2 = \begin{cases} 0 & \text{for } F \text{ of rank } 3 \\ Q_0^T z & \text{for } F \text{ of rank } 2 \end{cases} \qquad (22)$$

where r_i^2 be the i^{th} column vectors of R_2, and z_i is a scalar number. Let r_i^1 and r_i^2 be the i^{th} column vectors of R_1 and R_2 respectively. And let $r_i = r_i^1 + r_i^2$ be the i^{th} column vector of the rotation matrix $\hat{R}$. r_i and r_j are orthogonal for $i \neq j$. r_i^1 and r_i^2 are also orthogonal for any i and j. Using these constraints of orthogonality, we have

$$z_i z_j + (r_i^1)^T \cdot r_j^1 = \delta_{ij} \qquad (23)$$

where δ_{ij} is the Kroneker delta. R_1 is a projection of R into a two dimensional range space $R(F^T)$. Hence, if R_1 is not so heavily disturbed, z_i can be computed from equation (23). There are two solutions with different sign. The determinant of $\hat{R}$ can be derived as

$$determ(\hat{R}) = determ(R_1) + trace\{adj(R_1)R_2\} \qquad (24)$$

where determ (X) denotes the determinant, trace {X} represents the trace, and adj (X) denotes the adjoint of a matrix X. If F is of rank 2, determ (R_1) is zero. Therefore, the determinant of R has a different sign based on the signs of z_i. The correct answer of z_i makes determ ($\hat{R}$) positive.

Next, let us estimate the error bounds for those solutions for R_1 and R_2. Norms seem to be the best measure for the bounds of matrices. The following equation is always true:

$$(F+\delta F)(R+\delta R)=B+\delta B \quad (25)$$

If $\| F^{+}\triangle_F \| \le 1$, we get the following error bound for R1.

$$\frac{\|\delta R\|}{\|R\|} \le \frac{\|F\|\,\|F^{+}\|}{1-\|F^{+}\triangle_F\|}\left\{\frac{\|B\|}{\|F\|\,\|R\|}\frac{\|\triangle_B\|}{\|B\|}+\frac{\|\triangle_F\|}{\|F\|}\right\}$$

$$\le \frac{\|F\|\,\|F^{+}\|}{1-\|F^{+}\triangle_F\|}\left\{\frac{\|\triangle_B\|}{\|B\|}+\frac{\|\triangle_F\|}{\|F\|}\right\} \quad (26)$$

This equation shows that the change of $\|\triangle_F\|$ has more influence on the error bound of R1 than the change of $\|\triangle_B\|$.

The error of R2 is related to the error of R1. Let r_i^1 and r_i^2 be the i^{th} column vectors of R1 and R2 respectively. And let δr_i^2 denote the error of our solution r_i^2. Since the column vectors of R are orthogonal to each other, we can obtain the following equation.

$$\{(r_i^1+\delta r_i^1)+(r_i^2+\delta r_i^2)\}^T\{(r_j^1+\delta r_j^1)+(r_j^2+\delta r_j^2)\}=0 \quad \text{for } i\neq j \quad (27)$$

If we neglect the second order errors,

$$z_j\delta z_i+z_i\delta z_j = (\delta_{ij}-r_i^Tr_j)-\{(r_i^1)^T\delta r_j^1+(r_j^1)^T\delta r_i^1\} \quad (28)$$

where δz_i denotes a possible error of scalar z_i. The maximum error of δz_i can be easily obtained from

$$|\delta z_i| \le \begin{cases}\dfrac{A_{ii}}{2|z_i|} & \text{for } z_i\neq 0\\[2mm] \dfrac{A_{ii}}{|z_i|} & \text{for } z_i=0 \text{ with some } z_j\neq 0\end{cases} \quad (29)$$

where

$$A_{ij}\le|\delta_{ij}-r_i^Tr_j|+|r_i^1|\,\|\delta r_j^1\|+|r_j^1|\,\|\delta r_i^1\|$$

and

$$\frac{\|\delta r_i^1\|}{\|r_i^1\|} \le \frac{\|F\|\,\|F^{+}\|}{1-\|F^{+}\triangle_F\|}\left\{\frac{\|(\triangle_B)_i\|}{\|b_i\|}+\frac{\|\triangle_F\|}{\|F\|}\right\}$$

where $(\triangle_B)_i$ represents the i^{th} column vector of $\triangle_B$. The 1-norm of δR_2 becomes $\|\delta R_2\|_1=|\delta z_i|$. It can be proved that $\|\delta R_2\|_2\le 3^{1/2}\|\delta R_2\|_1$. Therefore, we get the following equation for the upper bound of $\|\delta R\|$:

$$\|\delta R\| \le \|\delta R_2\|+\|\delta R_1\| \quad (30)$$

$\|\delta R\|$ measures how much perturbed at the worst case under the transformation by $\hat{R}$. As depicted in the following figure, this can be obtained by determining the maximum magnification of vectors on the unit sphere. In orther words, $\|\delta R\|$ is defined as follows.

$$\|\delta R\| = \underset{|x|=1}{\mathrm{MAX}}\frac{\|R_x-\hat{R}_x\|}{\|x\|} = \underset{|x|=1}{\mathrm{MAX}}\|\delta R_x\|$$

Until now, we have computed $\hat{R}$ with the maximum value of its possible error. However, the estimated $\hat{R}$ may not be orthogonal because δF and δB are not zero. In the presence of observation error, we use the optimized rotation matrix, R, which minimizes

$$E = trace\{(R\cdot\hat{R})^T(R\cdot\hat{R})\} \quad \text{subject to } R^TR=I \text{ and } RR^T=I \quad (31)$$

E is a square of the Euclidean norm of $(R-\hat{R})$. The matrix R has three eigenvalues: one is real, and the other two eigenvalues are complex. Considering a rotation around a fixed point, the real eigenvalue is 1 and the corresponding eigenvector describes the direction of the rotation axis. The complex eigenvalues, $e^{\pm j\varphi}$, whose absolute value is also 1, correspond to the isotropic lines in the plane perpendicular to l, and the displacement within the plane is planar rotation with the angle φ[3]. Hamilton[7] showed that the rotation is translated into a product of quaternions such that

$$Rx = q * x * \bar{q} \tag{32}$$

where $q = (\sin(\varphi/2)p^T, \cos(\varphi/2))^T$. The equation (31) becomes

$$E = \sum_{i=1}^{3} \| q * u_i - r_i * \bar{q} \|^2 \tag{33}$$

where u_j and r_j are the i^{th} column vector of an identity matrix and $\hat{R}$ respectively. Equation (33) can be modified by a simple calculus such that

$$E = q^T G q \tag{34}$$

Since G is symmetric positive semidefinite, the optimal value of q is the unit eigenvector of G corresponding to the smallest eigenvalue. The rotation matrix becomes

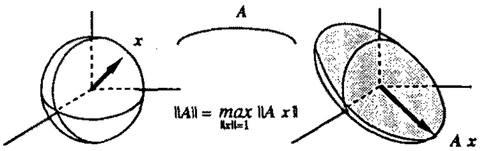

Fig. 1. Definition of $\| \delta R \|$

$$R = pp^T + \cos\varphi(I - pp^T) + \sin\varphi \begin{bmatrix} 0 & -p_3 & p_2 \\ p_3 & 0 & -p_1 \\ -p_2 & p_1 & 0 \end{bmatrix} \tag{35}$$

where p_i represents the i^{th} element of p which is the unit eigenvector corresponding to the real eigenvalue of R.

4. Estimation of the Object Translation

Estimation of object translation using point correspondence is straightforward. In the previous section, we obtained the following equation :

$$\begin{aligned}\{g(t) + \delta g(t)\} - \{F(t) + \delta F(t)\}d(t) \\ = \{H(t) + \delta H(t)\}\{c + \delta c\}\end{aligned} \tag{36}$$

Equation (36) is simplified as

$$(F + \delta F)d = e + \delta e \tag{37}$$

where $\boldsymbol{e}$ and $\delta \boldsymbol{e}$ are 3×1 vectors such that

$$e = g - Hc, \quad \delta e = \delta g - \delta H c - H \delta c \tag{38}$$

The bounds of $\delta \boldsymbol{e}$ becomes

$$|\delta e| \le \triangle_g + \triangle_H |c| + |H| \triangle_c = \triangle_e \tag{39}$$

The translation of an object can be obtained from equations (37) and (39). Thus,

$$\hat{d} = (F^T F)^{-1} e \tag{40}$$

$$\frac{\|\delta d\|}{\|\hat{d}\|} \le \frac{\|F\| \, \|(F^T F)^{-1} F^T\|}{1 - \|(F^T F)^{-1} F^T \triangle_F\|} \left\{ \frac{\|\triangle_e\|}{\|e\|} + \frac{\|\triangle_F\|}{\|F\|} \right\} \tag{41}$$

5. Estimation of a 3D Object Motion

Object motion can be specified by its translation and rotation changes. The displacement of an object is fully described by the changes of its translation vector. And the rotation of the motion is described by the changes of its rotation matrix, which is specified by the change of the rotation axis, i.e. the eigenvector p as discussed in previous section, and by the changes of rotation angle φ. Given p and φ, rotation matrix R can be found uniquely, and vice versa. However, it is not a good idea to specify continuous motion of an object using either p and φ or R directly. Although small perturbations in p and φ exert little influence on R, the reverse is not always true. As φ approaches to zero, even a small noise in R may cause large perturbations in p and φ. If the changes of R with respect to time are used to describe object motion, it is not easy to continuously keep the rotation matrix R orthogonal. For this reason, we use several virtual points to specify object motion[9,10]. For an object, we just define three arbitrary points, $P_1(t_0)$, $P_2(t_0)$, and $P_3(t_0)$ at time t_0 with constraints

$$\sum_{i=1}^{3} P_i(t_0) = 0 \tag{42}$$

As time goes, P_i is updated as equation (3) such that

$$P_i(t) = R(t)P_i(t_0) + d(t) \quad \text{for } t \geq t_0 \tag{43}$$

Clearly,

$$\frac{1}{3}\sum_{i=1}^{3} P_i(t) = d(t) \tag{44}$$

$R(t)$ and $d(t)$ are obtained from equations (34) and (40) respectively. Given $P_i(t)$ for i = 1, 2, and 3, $R(t)$ and $d(t)$ can be uniquely reconstructed. $P_i(t)$ is expanded as

$$P_i(t) = P_i(\tau) + \frac{1}{1!}P_i'(\tau)(t-\tau) + \frac{1}{2!}P_i'(\tau)(t-\tau)^2 + \cdot\cdot\cdot \tag{45}$$

If the time intervals between t and $d\tau$ are short, $P_i(t)$ can be approximated by the first k terms. Letting $q_{i,s+1} = (1/j!)P_i^{(s)}(\tau)$,

$$P_i(t) \cong \sum_{j=1}^{k} (t-\tau)^{j-1} q_{i,j} \tag{46}$$

Let us assume that the information about a motion is required at every time instant T_i, for i = 0, 1, 2, · · ·. Set $\triangle T = T_{i+1} - T_i$. Similarly, assume that the sampling data by the n^{th} sensor are available at time ${}^n t_j$, for j = 0, 1, 2, · · ·. Let ${}^n\triangle t = {}^n t_{j+1} - {}^n t_j$. Here, $T_j = {}^n t_j$ is not warranted. Let $\tau = T_j$ at time t, $T_{j+1} \geq t \geq T_j$. By given $q_{i,1}, q_{i,2}, \ldots, q_{i,k}$, we can obtain the estimated $P_i(T_{j+1})$ value, as $\hat{P}_i(T_{j+1})$.

$$\hat{P}_i(T_{j+1}) \cong \sum_{s=1}^{k} (\triangle T)^{s-1} q_{i,s} \tag{47}$$

Let us first consider a one-sensor case. Let $P_i(t_j)$ represent a location of a i^{th} virtual point at time t_j where $\tau \leq t \leq \tau + \triangle T$. We use error bounds discussed in the previous section as an uncertainty measure. Let $1/w_{i,j}$ be the uncertainty of $P_i(t_j)$ at time t_j. Set $\| P_i(t_0) \| = m_i$. Since

$$P_i(t) + \delta P_i(t) = \{R(t) + \delta R(t)\}P_i(t_0) + \{d(t) + \delta d(t)\}$$

we define

$$\frac{1}{w_{i,j}} = \max\{ \| \delta R(t_j) \| m_j + \| \delta d(t_j) \| \} \tag{48}$$

The uncertainty of a sensory output increases as the error bounds of the sensory output increase. The estimated motion $\hat{P}_i(t_j)$ at time t_j becomes

$$\hat{P}_i(t_j) = \sum_{s=1}^{k} (t_j - \tau)^{s-1} q_{i,s} = \sum_{s=1}^{k} (x_j)^{s-1} q_{i,s} \tag{49}$$

where $x_j = t_j - \tau$. Let $k \times 1$ vector X_j at time t_j be $X_j = [1, x_j, \ldots, (x_j)^{k-1}]^T$, and $3 \times k$ parameter matrix, Q_{ij}, at time t_j for $\hat{P}_i(t_j)$ be $Q_{i,j} = [q_{i,1} \vdots q_{i,2} \vdots \ldots \vdots q_{i,k}]$. Using $\hat{P}_i(t_j)$, $P_i(t_j)$ and $w_{i,j}$, we can estimate the motion parameter $Q_{i,j}$ at each time instant t_j which minimizes the following weighted least square error.

$$E_j = \sum_{i=1}^{3} E_{i,j} = \sum_{i=1}^{3} \left\{ \sum_{p=1}^{s} (w_{i,j-p}\ \varepsilon^T_{i,j-p}\ \varepsilon_{i,j-p}) \right\} \quad (50)$$

subject to

$$\frac{1}{3} \sum_{i=1}^{3} \hat{P}_i(t_j) = d(t_j)$$

where

$$\varepsilon_{i,j-p} = P_i(t_{j-p}) - \hat{P}_i(t_{j-p})$$

and S represents the number of the sensor data samples to be considered for continuous motion. $E_{i,j}$ represents the modeling error of the motion of the point P_i at time t_j. By using the Hamiltonian[7], we obtain the following linear equation:

$$Q_{i,j}\Big[\sum_{p=0}^{S-1} w_{i,j-p} X_{j-p} \vdots \sum_{p=0}^{S-1} w_{i,j-p} x_{j-p} X_{j-p} \vdots \dots \vdots \sum_{p=0}^{S-1} w_{i,j-p} (x_{j-p})^{k-1} X_{j-p} \Big]$$

$$= \Big[\sum_{p=1}^{S-1} w_{i,j-p} P(j-p) \vdots \sum_{p=0}^{S-1} w_{i,j-p} x_{j-p} P(t_{j-p}) \vdots \dots \vdots \sum_{p=0}^{S-1} w_{i,j-p} (x_{j-p})^{k-1} P(t_{j-p}) \Big] - \frac{1}{6} \lambda X_j^T \quad (51)$$

where 3×1 vector λ represents the Lagrange multiplier. Let $A_{i,j}$ and $B_{i,j}$ be $k \times k$ and $3 \times k$ matrices respectively such that

$$A_{i,j} = \Big[\sum_{p=0}^{S-1} w_{i,j-p} X_{j-p} \vdots \sum_{p=0}^{S-1} w_{i,j-p} x_{j-p} X_{j-p} \vdots \dots \vdots \sum_{p=0}^{S-1} w_{i,j} (x_{j-p})^{k-1} X_{j-p} \Big]$$

$$B_{i,j} = \Big[\sum_{p=0}^{S-1} w_{i,j-p} P(t_{j-p}) \vdots \sum_{p=0}^{S-1} P(t_{j-p}) x_{j-p} P(t_{j-p}) \vdots \dots \vdots \sum_{p=0}^{S-1} w_{i,j-p} (x_{j-p})^{k-1} P(t_{j-p}) \Big]$$

The Lagrange multiplier λ can be obtained from equation (51),

$$2\Big(\sum_{i=1}^{3} B_{i,j} A_{i,j}^{-1} \Big) X_j - 6d(t_j) = X_j^T \Big(\sum_{j=1}^{3} A_{i,j}^{-1} \Big) X_j \lambda \quad (52)$$

Equation (52) can be decomposed into three scalar linear equations. Once λ is obtained, motion parameter $Q_{i,j}$ of a virtual point P_i at time t_j is found using equation (51). The solution $Q_{i,j}$ is valid during $t_j \leq t \leq t_{j+1}$. The value E_i is used as a uncertainty measure for $Q_{i,j}$ at time t.

6. Motion Estimation using Multi-Sensor Fusion

When we try to incorporate information from different sensors and/or the same sensor in different positions, three approaches have been suggested[8,21]: averaging, deciding, and guiding. First, averaging is most appropriate when the observation is contaminated by considerable additive noise. Usually a weighted average is performed by giving more weight to the sensor data with the smaller noise variance[16]. Second, a decision can be made between measurements. This method is most suitable when the information provided by one sensor is vastly superior to the rest of the sensors[6]. Finally, information from some sensors can also be used to guide other sensors to produce a more accurate measurment[21]. It is hard to say which method is better than the others. We introduce an approach which combines the techniques of averaging and decision as illustrated in Fig. 2.

See Fig. 2. Using row data from each sensor, the object surface parameters F and g are computed with error bounds $\Delta_F(t)$ and $\varepsilon_g(t)$ respectively. These error bounds are not the absolute errors, but the maximum values of possible errors which guarantee a certain uncertainty to F and g. The instantaneous rotation information, R, and translation d are then estimated with maximum possible errors $\| \delta R \|$ and $\| \delta d \|$. In other words, $\| \delta R \|$ and $\| \delta d \|$ represent the maximum possible errors of R and d which guarantee the previously defined uncertainty of the observation. This rotation and translation information is used for estimating motion parameters if their possible maximum errors are small. If not, fedbacked motion parameters based on the previously estimated motion are used for estimating motion parameters. The weighted LSEE technique is applied as a local motion estimator. If all S consecutive displacements are chosen from the feedback loop for a certain sensor, so that this sensor has no contribution to the final decision, this particular sensor is considered to be erroneous.

The final decision is made by weighted averaging. Consider that we have N sensors, and let ${}^nQ_{i,j}$ denote the latest estimation of the motion parameter of a virtual point P_i for the n th sensory data at time, T_j. Let $1/{}^nU_{i,j}$ be the uncertainty corresponding to ${}^nQ_{i,j}$. ${}^nU_{i,j}$ is defined as

$$\frac{1}{{}^nU_{i,j}} = {}^nC_1\frac{{}^nE_{i,j}}{{}^nS-{}^nM_{i,j}} + \frac{{}^nC_2}{{}^nS}\sum_{t_{j-1}<t_l\le t_j}\frac{1}{{}^nw_{i,j}} + {}^nC_3\cdot({}^nM_{i,j}) \quad (53)$$

where ${}^nE_{i,j}$ represents $E_{i,j}$ in equation (50) for n^{th} sensor at time t_j, $t_j\le T_j\le t_{j+1}$, and ${}^nw_{i,l}$ represents $w_{i,j}$ in equation (48) for n^{th} sensor at time t_l. ${}^nS$ represents the number of data samples needed to estimate the motion parameters by the n^{th} sensor as in equation (52). ${}^nM_{i,j}$ represents how many times the fedbacked P_i's are used for the computation of ${}^nE_{i,j}$. And ${}^nC_1$, ${}^nC_2$, and ${}^nC_3$ is a weighting constant for n^{th} sensor. Equation (53) implies how well the instantaneous states of an object fit to the motion formulated, how accurate the instantaneous states are, and possible degradation of a certain sensor. The fused parameter $Q_{i,j}$ for a virtual point P_i at time T_j becomes

$$Q_{i,j} = \sum_{n=1}^{N}\frac{{}^nU_{i,j}}{\sum_{m=1}^{N}{}^mU_{i,j}}\,{}^nQ_{i,j} \quad (54)$$

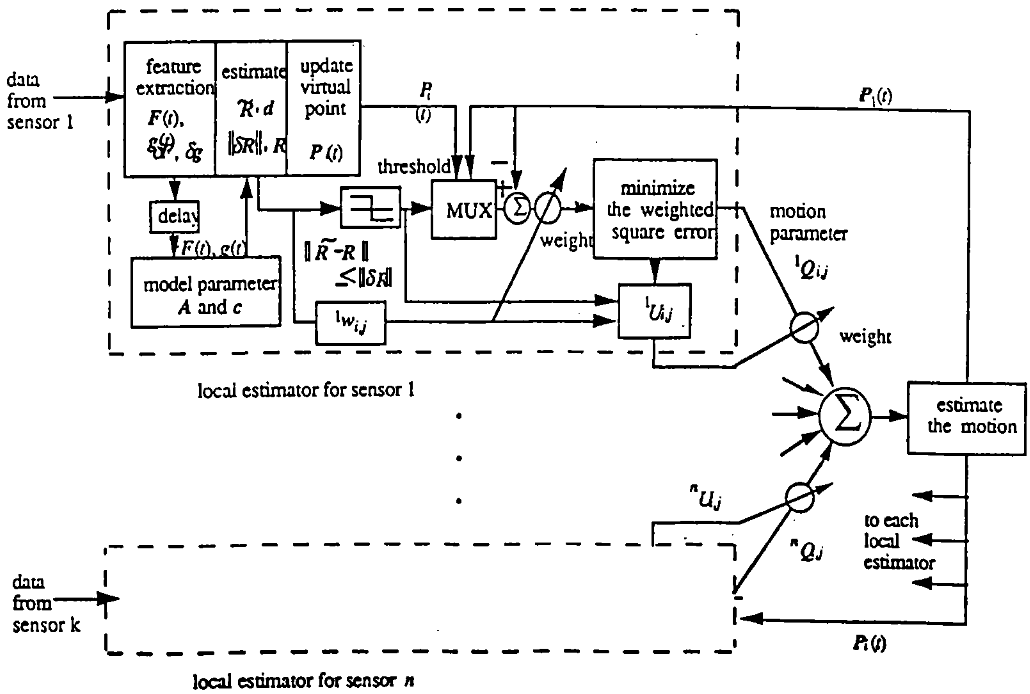

Fig. 2. Diagram of Overall Multi-Sensor Data Fusion

7. Simulation Results

For simulation a heptahedron is used as a polyhedral object. We generate the data based on this heptahedron. This object is assumed to be observed by different sensors with different observation uncertainties assuming that they have no a priori knowledge of the object. We consider that only the rotation portion $R(t)$ of this heptahedron is changed with respect to time because the translation portion $d(t)$ can be easily calculated using the same procedure of estimating the rotation parameter. The rotation axis p can be specified by two angles: ζ, the angle between p and xz-plane, and ξ, the angle between p and xy-plane.

We consider the rotation which consists of three types of time varying functions; linear, parabolic and sinusodial such that $\zeta(t) = 0.4t$, $\xi(t) = 0.05t(t-3)$, and $\varphi(t) = \sin(t/1.5)$. The combination of these three types of time varying functions can more generally describe the 3D motion of an object. Each sensor provides the parameters of visible surfaces of the heptahedron. These observed surface parameters contain observation errors which depend on the direction of the surface normal with respect to the view angle from a sensor. Because of the existence of different observation errors of the surfaces of the same object, we need to introduce a normalized observation error(NOE) in order to compare the observation errors of surface parameters from different surfaces. Let $ax+by+cz=d$ represent the equation of a certain object surface with no error and $\hat{a}x+\hat{b}y+\hat{c}z=\hat{d}$ represent the corresponding perturbed object surface equation observed. The NOE is defined as:

$$NOE(\%) = \frac{\sqrt{(a-\hat{a})^2+(b-\hat{b})^2+(c-\hat{c})^2+(d-\hat{d})^2}}{\sqrt{a^2+b^2+c^2+d^2}} \times 100\% \tag{55}$$

We assume that the NOE of an object surface increases proportionally to θ^2, where θ is defined as the angle between the view direction from a sensor and the surface normal. Thus, each object surface has its own surface parameters with different NOE, although several object surfaces are observed by the same sensor. This is a reasonable assumption because the error will increase as the object surface rotates towards a position parallel to the viewing direction. Each sensor is considered to have its inherent NOE bound which guarantees a certain uncertainty, and we assume only the surfaces whose θ are less than 80° can be observed within the NOE bound defined for each sensor.

In our simulation, we consider four sensors located 90° apart at the four different sides of an object, and assume that these sensors have different sampling times and different NOE bounds. For proof of concept purposes, we use two sensor with a small NOE bound(20 %) and two other sensors with larger NOE bounds(40%, and 80%). Sensor 1 is located on the positive x-axis and has a sampling time of 0.1 seconds.

The maximum NOE of sensor 1 is assumed to be 20 %. Sensor 4 is located on the negative x-axis with a sampling time of 0.2 seconds and the maximum NOE 80%. Sensors 2, and 3 are located on the positive and negative y-axes with a maximum NOE of 40% and 20%, respectively. Both sensors 2 and 3 are assumed to have the same sampling time of 0.15 seconds.

Fig. 3-6 illustrate the error portions of the estimated rotation. Average $\|\delta R\|$ denotes the average value of $\|\delta R\|$ for consecutive S virtual points used to estimate the object motion. Our approach to estimating object motion is a kind of curve fitting method. Thus, our method tends to smoothe the estimating errors especially when k is small. As k increases, the location of virtual points computed from the estimated motion approaches more closely to the instantaneous location of virtual points but more fluctuates. Clearly, Figures 3 and 5 with k=3 show less E than Figures 4 and 6 which have k of value 2, but tend to fluctuate. Usually, error is large with large E. However, the low value of modeling error, E, does not guarantee the small error especially when NOE is large. Average $\| \delta R \|$ seems to have more influence on the error. Thus, we can see that it is reasonable to specify an uncertainty of object motion using E and average $\| \delta R \|$ as in equation (55).

Fig. 7 shows the fused result. Table 1 illustrates the maximum and minimum errors of

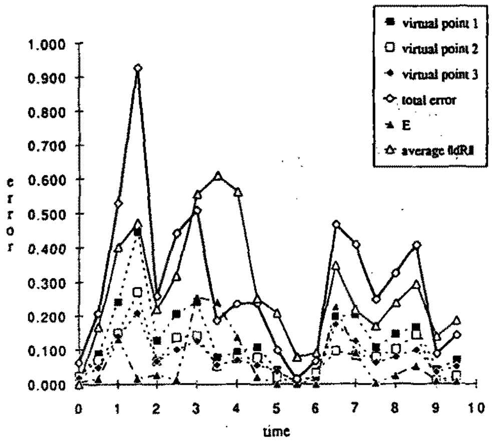

Fig. 3. Error of Estimated Locations of Virtual Points from Sensor 1

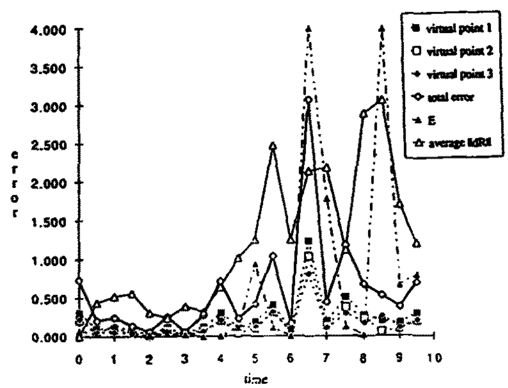

Fig. 4. Error of Estimated Locations of Virtual Points from Sensor 2

each sensor and fused result. We set C_1, C_2, and C_3 to be 1, 1, and 3 respectively. For demonstration purpose, we co

nsidered only four sensors with a wide range of NOE bounds in this particular simulation. Considering the number of sensors used and their wide range of uncertainties(i.e. from 20% to 80% NOE), the simulation results show that our approach for motion estimation is satisfactory.

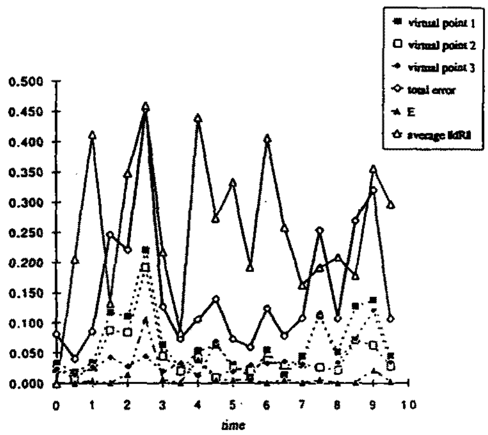

Fig. 5. Error of Estimated Locations of Virtual Points from Sensor 3

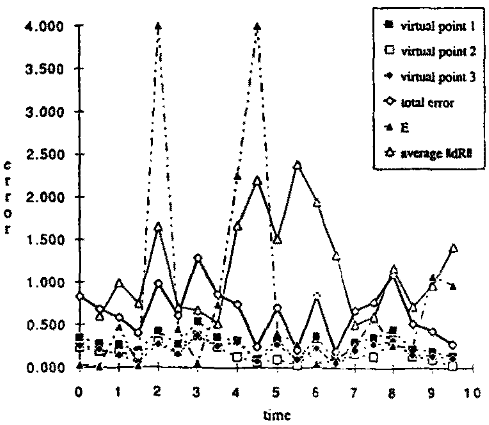

Fig. 6. Error of Estimated Locations of Virtual Points from Sensor 4

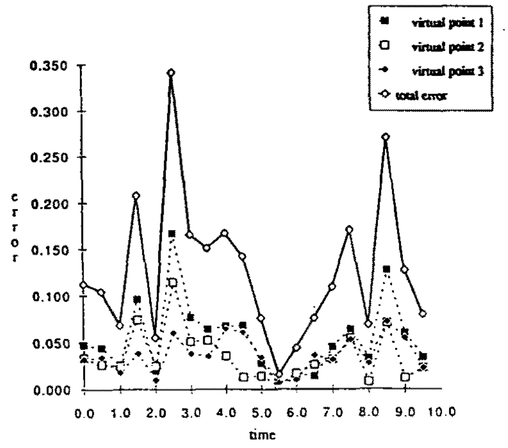

Fig. 7. Error of Fused Locations of Virtual Points

Table. 1 Maximum and Minimum Error of Each Sensor

Sensor / Error	Sensor 1	Sensor 2	Sensor 3	Sensor 4	Fused Result
Max. Error	1.062	1.781	0.512	2.232	0.341
Min. Error	0.041	0.076	0.014	0.317	0.016

8. Conclusion

In this paper a new approach is proposed to estimating the general 3D motion of polyhedral objects using multiple sensor data. Proposed method can provide answers to some of the difficult problems of dynamic sensor fusion: the lack of a priori knowledge of environment and the problem of sensor synchronization. Motion can be estimated continuously from each sensor through the analysis of the instantaneous state of an object. If at least two surfaces are observed, this approach can compute an object's orientation. For the data fusion, the uncertainty of each sensor data should be adequately modeled and specified. The means have been provided to adequately represent the error or uncertainty of the results for the estimation of object's orientation so that it can be reliably fused and integrated with other sensors. Using the instantaneous state of an object with its associated uncertainty, the object's motion is estimated from each sensor and fused to provide more accurate and reliable information about the motion. We have presented a general paradigm of multisensor data fusion which combines the techniques of averaging and decision. The property of decision makes it possible to detect malfunctioning sensors.

REFERENCES

[1] G. Adiv, "Determining Three-Dimensional Motion and Structure from Optical Flow Generated by Several Moving Objects", IEEE Trans. on Pattern Anal. Machine Intell., Vol. 7, pp. 348-401, 1985.

[2] A. Albert, Regression and the Moore-Penrose Pseudoinverse, Academic Press, NY. 1972.

[3] O. Bottema and B. Roth, Theoretical Kinematics, North-Holland Publishing Co., NY., 1979.

[4] H. H. Chen and T. S. Huang, "Maximal Matching of Two 3D Point Sets", in Proc. of the ICPR, pp. 27-31, Paris, France, Oct. 1986.

[5] B. A. Conway, J. E. Tulingliwski, and P. D. Webber, "Dynamics of Remote Orbital Capture", in Proc. AAS/AIAA Astrodynamics Specialist Conf., 1983.

[6] A. M. Flynn, "Combining Sonar and Infrared Sensors for Mobile Robot Navigation", Int. J. Robotics Research, pp. 5-14, 1988.

[7] W. R. Hamilton, Elements of Quaternions, Chelsea, N.Y. 1969.

[8] S. Y. Harmon, G. L. Bianchini, and van Aslst, "Sensor Data Fusion through a Distributed Blackboard", in Proc. of IEEE Int. Conf. on Robotics and Automation, pp. 1449-1454, 1986.

[9] T. S. Huang, Image Sequence Analysis, Springer-Verlag, Heidelberg, FRG, 1981.

[10] T. S. Huang, Image Sequence Processing and Dynamic Scene Analysis, Springer-Verlag, Heidelberg, FRG, 1983.

[11] M. H. Kaplan and A. A. Nadkarni, "Control and Stability Problems of Remote Orbital Capture", Mechanism and Machine Theory, vol. 12, pp. 57-64, 1977.

[12] Kirk, "Optimal Control Theory", Wiley, New York, 1978.

[13] Y. C. Liu and T. S. Huang, "Estimation of Rigid Body Motion Using Straight-Line Correspondences", in Proc. of the IEEE Workshop on Motion: Representation and Analysis, pp. 47-52, Kiawah Island, SC, May 1986.

[14] Y. C. Liu and T. S. Huang, "Estimation of Rigid Body Motion Using Straight-Line Correspondences", Computer Vision, Graphics, and Image Processing 43, pp.37-52, 1988.

[15] R. C. Luo and M. G. Kay, "Multisensor Integration and Fusion in Intelligent Systems", Pattern Anal. Machine Intell., Vol. 19, pp. 901-931, Sep. 1989.

[16] R. C. Luo and M. Lin, "Dynamic multi-sensor data fusion system for intelligent robots," in Proc. of IEEE Int. Conf. on Robotics and Automation, pp. 1076-1081, Philadelphia, PA, Apr. 1988.

[17] R. C. Luo and W. S. Yang, "Multi-Sensor Fusion for Estimation of a Moving Polyhedral Object", in Proc. of SPIE Sensor Fusion II: Human and Machine Strategies, Philadelphia, PA, Nov. 1989.

[18] R. C. Luo and W. S. Yang, "Estimation of a Moving Planar Object using Multi-Sensor Fusion", J. of Robotics System, 1990.

[19] Ren C. Luo, W. S. Yang and M, Lin, "Multi-Sensor Data Fusion under Time Evolving Environments", in Proc. of IEEE Int'l Conf. on Intelligent Control Systems, Arlington, VA, Aug. 1988.

[20] P. S. Maybeck, Stochastic Models, Estimation, and Control, Vols. 1 and 2. New York, NY. Academic Press, 1979 and 1982.

[21] S. W. Shaw, R. deFigueiredo, and K. Krishen, "Fusion of Radar and Optical Sensors for Space Robotic Vision", in Proc. of IEEE Int. Conf. on Robotics and Automation, Philadelphia, PA, Apr. 1988.

[22] J.Weng, T.S.Huang, and N. Ahuja, "3D Motion Estimation, Understanding, and Prediction from Noisy Image Sequences", IEEE Trans. on Pattern Anal. Machine Intell., Vol. 9, pp. 370-387, 1987.

著 者 紹 介

梁雨錫(Woo Suk Yang)
1979년 서울대학교 전기공학과 졸업 (공학사). 1990년 미국 North Carolina State Univ. Electrical and Computer Engineering 졸업 (공학박사). 1978년 - 1884년 (주) 대우. 1990년 - 1991년 금성사 책임연구원. 1991년 - 현재 홍익대학교 전기공학과 조교수. 주관심분야 : 로보틱스, 컴퓨터비젼, FA, Sensor Fusion