International Journal of Navigation and Port Research
Vol.27, No.3 pp. 273~282, 2003 (ISSN-1598-5725)

AGV Navigation Using a Space and Time Sensor Fusion of an Active Camera

Tae-Seok Jin · Bong-Ki Lee* · Jang-Myung Lee***

**, **Dept. of Electronics Engineering, Pusan Nat'l Univ., Pusan, 609-735, korea*

Abstract : *This paper proposes a sensor-fusion technique where the data sets for the previous moments are properly transformed and fused into the current data sets to enable accurate measurement, such as, distance to an obstacle and location of the service robot itself. In the conventional fusion schemes, the measurement is dependent only on the current data sets. As the results, more of sensors are required to measure a certain physical parameter or to improve the accuracy of the measurement. However, in this approach, instead of adding more sensors to the system, the temporal sequence of the data sets are stored and utilized for the measurement improvement. Theoretical basis is illustrated by examples and the effectiveness is proved through the simulations. Finally, the new space and time sensor fusion (STSF) scheme is applied to the control of a mobile robot in the indoor environment and the performance was demonstrated by the real experiments.*

Key words : *Navigation, Sensor Fusion, Active Camera, Obstacle Avoidance, Image Processing*

1. Introduction

There have appeared many research studies on the spatial fusion technique. That is, multiple sensor data are utilized for the purpose of providing complementary or redundant data to measure physical parameters. That is, many current data from the various sensors are integrated and fused to obtain a useful combination(R.C. Luo, Chih-Chen Yih, and Kuo Lan Su, 2002), (R. C. Smith, P. Cheeseman, 1986), (R. C. Luo and K. L. Su, 1999). In recent years interests have been growing in the synergistic use of multiple sensors to increase the capabilities of intelligent machines and systems. For these systems to use multiple sensors effectively, a strategy to integrate the information provided by these sensors for the operations of the system is necessary. While in many multi-sensor systems the information from each sensor serves as a separate input to the system, the actual combination or fusion of information prior to its use in the system has been a particularly active area of research. Typical applications that can benefit from the use of multiple sensors are industrial tasks like an assembly, military command and control for battlefield management, mobile robot navigation, multi-target tracking, and aircraft navigation(Jang M. Lee, B. H. Kim, M. H. Lee, M. C. Lee, J. W. Choi, and S. H. Han, 1999), (P. Weckesser and R. Dillman, 1997). In all of these applications, the system needs to intelligently interact with the human and operates in an unstructured environment with the human operator's involvement.

Fusing temporal information recursively is crucial to many applications, such as navigation, robotics, target identification, and multi-target tracking. There are algorithms that can be used to integrate temporal information(R.C. Luo, Chih-Chen Yih, and Kuo Lan Su, 2002). Among them, the distributed Kalman filtering(L. Hong, A. Lynch, 1993), (Ren C. Luo and Kuo L. Su, 1999) and the Bayesian approach(S. Thrun, 1996) are popular for many applications. However, these algorithms require substantial prior information, such as initial values and initial covariance matrices for the distributed Kalman filtering and prior probabilities for the Bayesian approach. In many cases, the prior information is either not available or not known precisely. Theoretically, some estimated values could be used as the prior information if the models are correct and consistent measurements are provided. However, it takes some time for the systems to converge to the correct values. On the contrary, the Dempster-Shafer technique has the strong capability of handling information uncertainty, at the cost of more expensive computation(A. P. Dempster, N. M. Laird, and D. B. Rubin, 1977), (W. Pieczynski, 2000), (J. Llinas and D. L. Hall, 1994). Note that the technique has been employed mostly for spatial information fusion(J. Llinas and D. L. Hall, 1994) and for temporal information fusion only when the data structure is disjoint(L.A. Zadeh, 1973)

In this paper, as a general approach of sensor fusion, a STSF (Space and Time Sensor Fusion) scheme is proposed for either joint or disjoint data structure and applied to the landmark identification for mobile robot navigation. This

** Corresponding Author : Jang-Myung Lee, jmlee@pusan.ac.kr, 051-510-2378
* ybkybk2000@pusan.ac.kr, 051-510-1696
* jints@pusan.ac.kr, 051-510-1696

newly proposed STSF is necessary for the complementary data case where unless there is the sensor fusion, the measurement cannot be completed. Therefore the utilization method will be determined by the sensor data structure. However for the redundant data case where some sensor data are not essential for the measurement, it is required to define the method to fuse the previous data set to the current data set. The least-squares solution is generally adopted for the redundant data fusion without considering the error variance in the measurement.

2. Space and Time Sensor Fusion

Multi sensor fusion refers to any stage in the integration process where there is an actual combination (or fusion) of different sources of sensory information into one representational format.

2.1. A General Pattern of Sensor Fusion

Fig. 1 means to represent a general pattern of multi sensor integration and fusion in a system. In this figure, n sensors are integrated to provide information to the system.

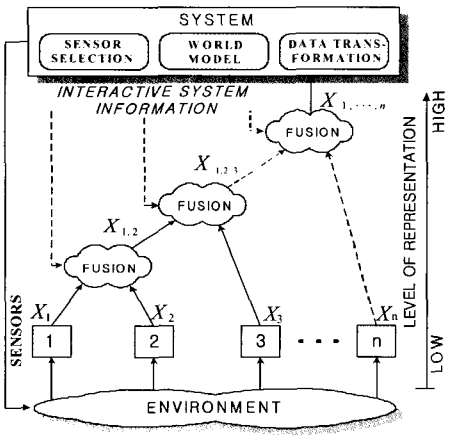


Fig. 1 General pattern of multi sensor integration and fusion.

The output x_1 and x_2 from the first two sensors are fused at the lower left hand node into a new representation $x_{1,2}$. The output x_3 from the third sensor could then be fused with $x_{1,2}$ at the next node, resulting in the representation $x_{1,2,3}$, which might then be fused at nodes higher in the structure.

2.2 Sensor Fusion Transformation

Let us define the k th moment data set provided by i th sensor as, $\mathbf{z_i(k)}$, and the k th measurement vector as $\mathbf{x(k)}$. Then the conventional sensor fusion technique provides the measurement as

$$\hat{x}(k) = \sum_{i=1}^{n} W_i x_i(k) \tag{1}$$

where $x_i(k) = H_i z_i(k) \in R^m$,

H_i represents transformation from the sensory data to the measurement vector, and $W_i \in R^{m \times m}$ represents the weighting value for i th sensor.

Note that in the measurement of $\mathbf{z_i(k)}$, the low level fusion might be applied with multiple sets of data with known statistics[9]. The determination of H_i is purely dependent on the sensory information and the decision of W_i can be done through the sensor fusion process. Later this measured data are provided to the linear model of the control/measurement system as current state vector, $\mathbf{x(k)}$. In this approach, we propose a multi sensor data fusion using sensory data, $Tz_i(j)$, as

$$\hat{x}(k) = \sum_{i=1}^{n} W_i \{ \sum_{j=1}^{k} P_j Tz_i(j) \} \tag{2}$$

where $\sum_{j=1}^{k} P_j = 1$

Note that when each of sensor information can provide the measurement vector, that is, the redundant case, $Tz_i(j)$ can be expanded as

$$Tz_i(j) = T_j + H_i z_i(j) \tag{3}$$

where T_j represents the homogeneous transformation from the location of the j th to the k th measurements.

However, when the multi sensors are utilized in the complementary mode, the transformation relationship cannot be defined uniquely; instead it will be defined depending on the data constructing algorithm from the measurements. For example, a single image frame captured by a camera on an AGV cannot provide the distance to an object until the corresponding object image is provided again from a different location. This algorithm will be described in detail in the section 3.1.

Fig. 2 illustrates the concept of this multi sensor temporal data fusion. Estimation of parameter may provide the measurement vector at each sampling moment. The verification of significance and adjustment of weight steps are pre processing stages for the sensor fusion. After these steps, the previous data set will be fused with the current data set, which provides a reliable and accurate

data set as the result of multi sensor temporal fusion. Significance implies that how much the previous data set is related to the current data. An arbitrary value of significance may cause the problem to be complex. Therefore, here we may consider whether it corresponds to the same data or not, that is, 1 or 0. When the significance is 0, the weight can be adjusted simply to be 0. However, when the significance equals 1, the adjustment of weight should be properly performed to provide reliable and accurate data. Here in the following sub section, we introduce a simple methodology for the weight adjustment.

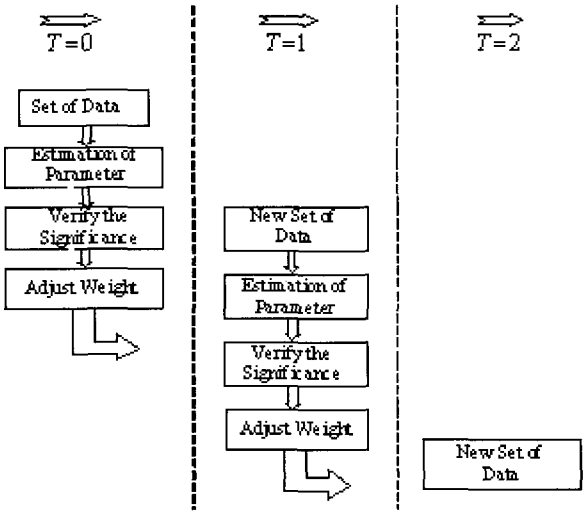


Fig. 2 Concept of Space and Time Sensor Fusion.

2.3 Auto-correlation for Estimation Techniques

Each previous data set is transformed to the k th (current) sampling location, and represented by the measurement vector, $Tz_i(j)$. Now how can we fuse the k data sets into a reliable and accurate data set? In the Eq. (2), W_i can be determined by the geometrical relationship among sensors, in other words, by the spatial sensor fusion.

While the estimating sensor is tracking a feature, it generates a stream of measurements. When there is relative motion between the feature and the sensor, the processes then cease to be stationary. As an illustration of gathering the model information from the sensor, we shall only consider the stationary case, in order words, there is no motion between the sensor and the feature being tracked. Our interest in the proceeding analysis lies only in determining whether the process noise is white or not.

Random processes are defined in terms of their ensemble averages and these can be estimated. Our model shall be in terms of such averages. In practice, we require to estimate these averages from finite sequences. We consider a process y_k as realized (estimated) by the finite sequence $y(k)$, for $0 \le k \le N-1$. That $y(k)$ is an estimate of the random process y_k is made plausible by a consideration of ergodic processes. From $y(k)$, we can, therefore, estimate the averages for the process, the mean is estimated by $\hat{\mu}_x = \frac{1}{N}\sum_{k=0}^{N-1} x(k)$ and the variance by $\hat{\sigma}_x^2 = \frac{1}{N}\sum_{k=0}^{N-1}(x(k)-\mu_x)^2$,

A biased estimate of the autocorrelation is given by

$$\tilde{\phi}_{xx}(m) = \frac{1}{N}\sum_{k=0}^{N-|m|-1} x(k)\cdot x(k+m) \tag{4}$$

with a variance of $\tilde{\sigma}_{\tilde{x}} = \frac{1}{\sqrt{N}}\tilde{\sigma}_x^2$, (5)

where $|m| < N$ Similarly, an unbiased autocorrelation is estimated by

$$\hat{\phi}_{xx}(m) = \frac{1}{N-|m|}\sum_{k=0}^{N-|m|-1} x(k)\cdot x(k+m) \tag{6}$$

and the variance of the unbiased autocorrelation is given by

$$\hat{\sigma}_{\hat{\phi}} = \frac{\sqrt{N}}{N-|m|}\hat{\sigma}_x^2 \tag{7}$$

for the biased autocorrelation and similarly for the unbiased one. Therefore, determination of P_j is the final step for the temporal sensor fusion. Note that this expands the dimension of sensor fusion from one to two.

As one of solid candidate, we propose here to use the auto correlation as an index for the weight adjustment and have the form,

$$\Psi_j = \sum_{k=-\infty}^{+\infty} x_i(k)x_i(j-k). \tag{8}$$

Depending on the correlation, P_j will be determined as

$$P_j = \frac{\Psi_j}{\sum_{j=1}^{k}\Psi_j}\quad \text{g} \tag{9}$$

3. Applications to an AGV

3.1 Complementary Usage for 3D Vision

If the image for an object is well matched to one model in the database, the position of the object can be obtained directly. In a well structured environment, it may be a

usual case. However, when the AGV is navigating in an unstructured environment, it needs to recognize the position/orientation of an object located in the middle of its path, which is not known to the AGV a priori.

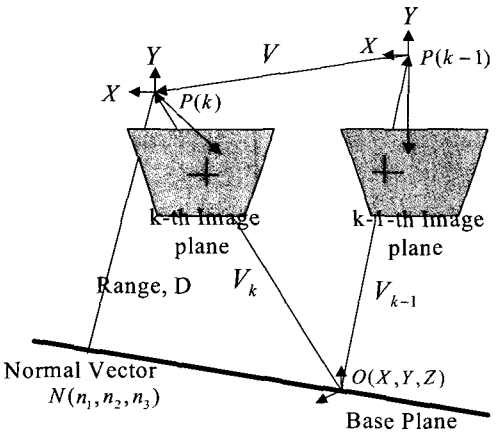


Fig. 3 Transformation of camera coordinates.

As a typical geometrical model for camera, a pinhole model is widely used in vision application fields as shown in Fig. 3. At the k th sampling moment, a scene point O(X,Y,Z) is captured by a camera on the AGV. The vectors from the scene point to the k th and $(k-1)$th camera perspective center are represented by V_k and V_{k-1}, respectively. The motion of AGV from $(k-1)$th moment to k th moment is represented by $\boldsymbol{V}$. Now we can write the vector relationship as

$$V_{k-1} = V_k - V \tag{10}$$

This can be represented as a matrix form,

$$\alpha \begin{bmatrix} x_{k-1} \\ y_{k-1} \\ -f \end{bmatrix} = \beta \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} x_k \\ y_k \\ -f \end{bmatrix} - \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \tag{11}$$

where (x_k, y_k, f) and (x_k, y_{k-1}, f) represent the projection of the scene point onto the camera image planes; V(v_1, v_2, v_3) represents the translational motion of the AGV; r_{ij} is an element of the rotation matrix, R represents the relative rotation between the two camera frames; α and β are constants and f is the focal length of a camera .

Now consider the reference base plane passing through the scene point P with a direction vector N(n_1, n_2, n_3); then the range value, D, can be represented as

$$D = V_k \cdot N \tag{12}$$

This can be represented again as

$$D = \beta (n_1 x_k + n_2 y_k - n_3 f) \tag{13}$$

Now, Eq. (11) is reformulated as

$$\alpha \begin{bmatrix} x_{k-1} \\ y_{k-1} \\ -f \end{bmatrix} = \beta \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} x_k \\ y_k \\ -f \end{bmatrix} - \frac{\beta}{D} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \begin{bmatrix} n_1 & n_2 & n_3 \end{bmatrix} \begin{bmatrix} x_k \\ y_k \\ -f \end{bmatrix} \tag{14}$$

$$(\alpha / \beta) \begin{bmatrix} x_{k-1} \\ y_{k-1} \\ -f \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_k \\ y_k \\ -f \end{bmatrix} \tag{15}$$

where $a_{ij} = r_{ij} - (v_i \cdot n_j / D)$.

Expanding the matrices and dividing rows one and two by row three gives

$$D(R_3 x_{k-1} + R_1 f) = C_3 x_{k-1} + C_1 f \tag{16}$$

$$D(R_3 y_{k-1} + R_2 f) = C_3 y_{k-1} + C_2 f \tag{17}$$

where $R_i = r_{i1} x_k + r_{i2} y_k - r_{i3} f$ and $C_i = v_i (n_1 x_k + n_2 y_k - n_3 f)$

In matrix form, these equations can be expressed as

$$AD = B \tag{18}$$

where $A^T = [a \quad b]$, $B^T = [c \quad d]$, $a = R_3 x_{k-1} + R_1 f$, $b = R_3 y_{k-1} + R_2 f$, $c = C_3 x_{k-1} + C_1 f$, and $d = C_3 y_{k-1} + C_2 f$.

Use of the pseudo inverse matrix enables computation of the range value, D which is associated with image point (x_k, y_k), and is written as,

$$D = (A^T A)^{-1} A^T B \quad \text{or} \tag{19}$$

$$D = \frac{(ac + bd)}{a^2 + b^2} \tag{20}$$

So far, we have shown that using the consecutive two image frames, the distance information of the scene point can be obtained as using the stereo images at a certain moment.

3.2 Space and Time Fusion Filter

We represent a single observation of an object as a two dimensional Gaussian distribution in Fig. 4. The center, or mean of the distribution is the estimated location of the object and the standard deviations along the maximum and

minimum axes of the distribution correspond to the estimates of the uncertainty (or noise) in the observation along the corresponding axis. The value of the distribution at any point corresponds to the probability that the object is actually in that location.

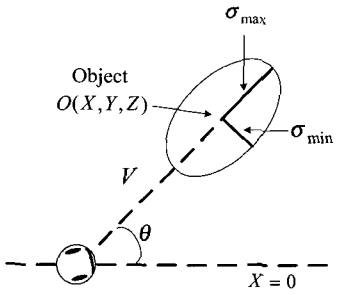


Fig. 4 Gaussian distribution parameter definition: mean (X, Y), standard deviations along maximum and minimum axes σ_{max} and σ_{min}, and distance to the object, V.

For the given observations, we need to determine the mean, standard deviations, and angle of the merged distribution to estimate object position and characterize the quality of the estimate. We compute the mean, standard deviations, and angle of measurement distributions from sensor readings (mean and angle) and models of sensor error (deviations). Thus, we require a method of determining combined parameters from those of individual distributions. The matrix form of Kalman filtering adopted by Smith and Cheeseman makes this computation relatively simple(R.C. Smith, 1986). Because the mean, standard deviations, and orientation of the major axis are independent of scaling, they can be extracted from the merged covariance matrices without considering scaling factors.

The canonical form of the two-dimensional Gaussian distribution depends on standard deviation, σ, a covariance matrix, C, and the mean(R.C. Smith, 1986). The covariance matrix of an observation, C, is initially determined from the major and minor axis standard deviations as

$$\underline{C} = \begin{bmatrix} \sigma_{max}^2 & 0 \\ 0 & \sigma_{min}^2 \end{bmatrix}. \quad (21)$$

Since the observation may be oriented arbitrarily with respect to the global coordinate frame, it must first be rotated to align with this frame:

$$C = R(-\theta)^T \underline{C} R(-\theta) \quad (22)$$

where θ is the angle of the distribution's principal axis with respect to the global x-axis. This rotation accomplishes the transformation from observation parameters to the canonical form. Once the observations are in the canonical form, we continue to merge the observations into one.

The covariance matrices of two distributions, C1 and C2, can be combined into a single covariance matrix, C, as

$$C = C_1 - C_1[C_1 + C_2]^{-1} C_1 \quad . \quad (23)$$

Now the mean of the resulting merged distribution, X, is computed from the individual distribution means and covariance matrices as follows:

$$\hat{X} = \hat{X}_1 + C_1[C_1 + C_2]^{-1}(\hat{X}_2 - \hat{X}_1). \quad (24)$$

The angle of the resulting principal axis is obtained from the merged covariance matrix:

$$\theta = \frac{1}{2} \tan^{-1}\left(\frac{2B}{A-C}\right) \quad (25)$$

where A, B, and C represent the components of the covariance matrix (upper left, upper right/lower left, and lower right, respectively). Lastly, the resulting major and minor axis standard deviations are extracted by rotating the covariance matrix to align with those axes:

$$C = R(\theta)^T C R(\theta) \quad (26)$$

and the resulting major and minor axis standard deviations can be directly computed from the covariance matrix by reversing Eq. (21)

4. AGV Type in Experiments Setup

The AGV used in the experiments is an *IRL 2001* developed in the *IRL*, PNU which is designed for an intelligent service vehicle.

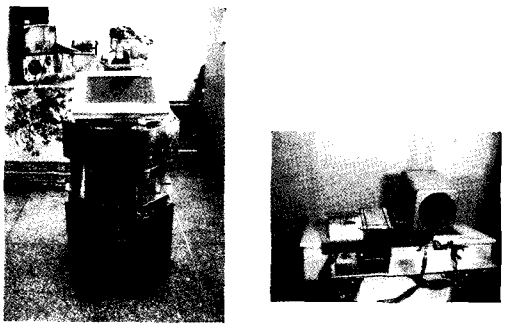


Fig. 5 *IRL* 2001 AGV and CCD camera system.

This vehicle is shown in Fig. 5 along with some of its sensory components. Its main controller is made on system clock 600 MHz, Pentium III Processor. The sensors, 16 ultrasonic and a robust odometry system are installed on the AGV. Ultrasonic sensors and infrared sensors in eight sides(25°) sense obstacles of close range, and the main controller processes this information.

For visual information, a CCD camera is mounted on the top of the AGV in order to sense obstacles or landmarks of the side and the rear of AGV. And DC servomotors are used for steering and driving of *IRL* 2001 AGV.

5. Experimental Results

5.1 Vehicle Localization used Landmark pattern recognition

The AGV robot(*IRL* 2001) is commanded to follow the environment as shown from (1) to (6) of Fig. 7. We performed the experiment for two cases.

To begin with, the 2 D landmark used by *IRL 2001* is shown in Fig. 6. The primary pattern of landmark is a 10cm black square block on white background and a 5cm square block. The major reasons for choosing the square blocks are

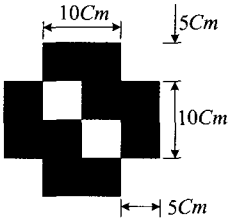


Fig. 6 The landmark pattern and size used by *IRL 2001.*

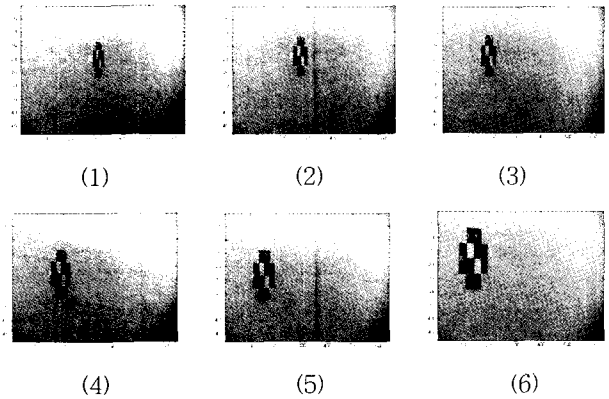


Fig. 7 A landmark locations detected by camera.

The projection of a square block in the image plane can always be approximated by an ellipse, which is easy to recognize using the elliptical Hough transformation technique(V. Ayala, 2000).

A square pattern is more robust to noise and occlusion than circular, polygonal patterns during template matching process, even though all these patterns can be detected by using Hough transformation technique(V. Ayala, 2000).

The image corners are then automatically extracted by camera parameters, and displayed on Fig. 7 and the blue squares around the corner points show the limits of the corner finder window. The corners are extracted to an accuracy of about 0.1 pixel.

The extrinsic parameters, relative positions of the landmark with respect to the camera, are then shown in a form of a 3D plot as Fig. 8. And In Fig. 9, every camera position and orientation are represented by red pyramid, therefore we can see the location and the orientation of an AGV in the indoor environment.

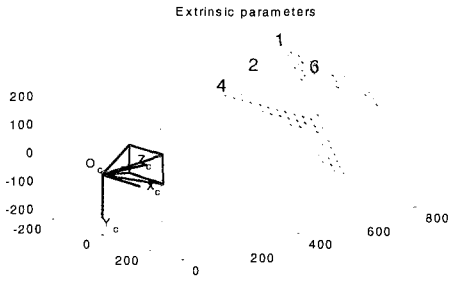


Fig. 8 Relative positions of the landmark w.r.t the camera.

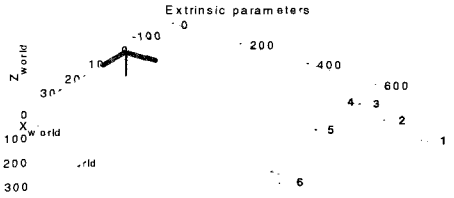


Fig. 9 AGV position and orientation.

To measure the relative distance of the landmark from the AGV, we first measure the distance of image from the fixed position in *IRL* lab corridor. The predefined values of the landmark defined in this section are given as follows the origin of coordinates is equal to the origin of AGV, a Y axis is fit to the front of AGV and an X axis is perpendicular with Y axis.

Table 1 shows the data measured in open-corridor. The Left direction marks negative. From table 1, we find the maximum and the minimum error on distance is 0.32 m and 0.13m, respectively.

It shows that the distance error becomes less and less by frames, which composes the environment map. And so,

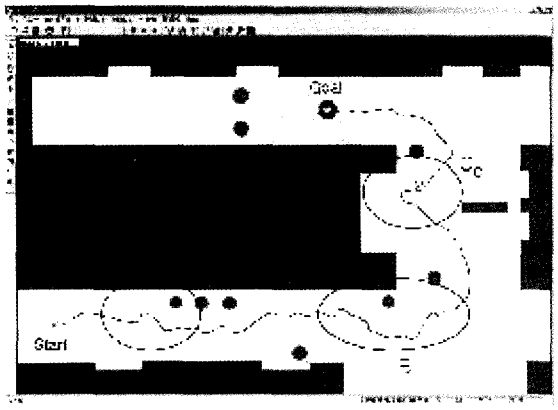
we can use it to measure the relative distance of the AGV.

Table 1 The result of relative distance (Dim.:m).

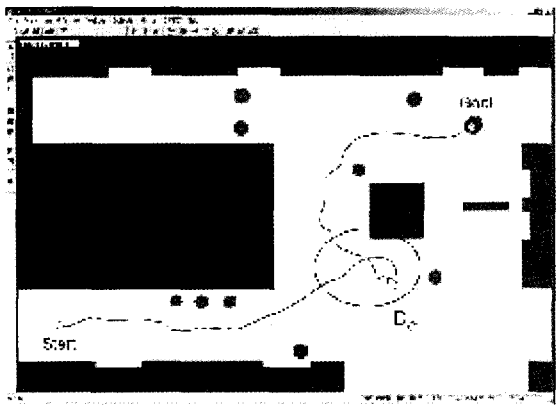
Frame Number	World Coordinate Distance	Image Coordinate Distance	Error
1	7.81	8.13	0.32
2	7.02	7.30	0.28
3	6.28	6.53	0.25
4	5.06	4.89	0.17
5	5.52	5.39	0.13
6	6.32	6.46	0.14

5.2 AGV Navigation

Conventional fusion and STSF(a space and time sensor fusion) have first been tested with simulation to show the usefulness of STSF in two environments respectively. Starting at (0.3m, 5m, 0 degree), a virtual AGV was driven around a virtual square corridor one time. The walls in the artificial environment are denoted by the real map, open-corridor of PNU.



(a) Experiment in a corridor.



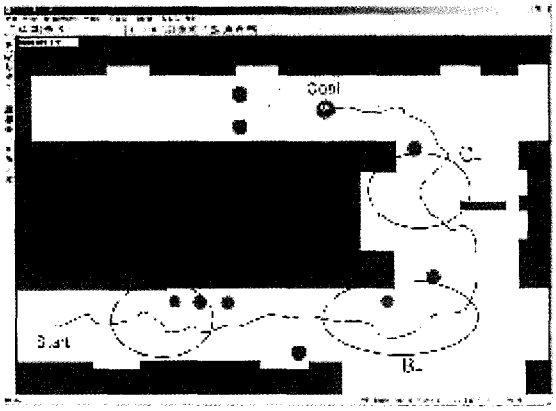
(b) Experiment in a corridor with wide space.

Fig. 10 Simulation for pointing vector based upon current readings.

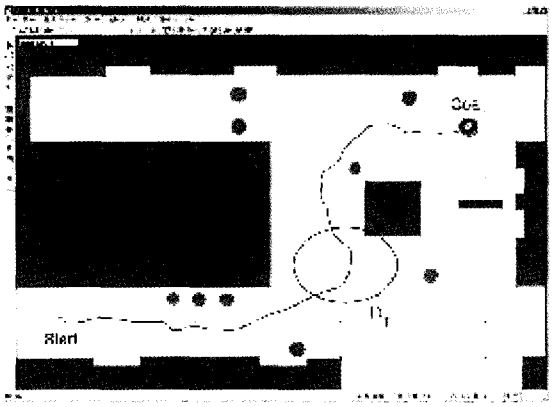
In each round, the AGV stops a total of 12 times to rescan the environment. The size of given map is 12m X 8m, the total distance traveled is 12 + 8=20 meters, and the total number of scanning points is 38. The comparison of simulation position and direction at all stops is shown in Fig. 10 and Fig. 11.

Fig. 10 shows determination of the pointing vector based upon only current readings used conventional sensor fusion, i.e. spatial fusion. This AGV was made to move randomly within the confines of the above setup and at the region, *Cc*. There are a little of difference between conventional fusion and the new STSF.

But at the region, *Ac*, the AGV moves not keeping the distance between AGV and wall constant and have some difficult local minimum trap problems at some places.



(a) Experiment in a corridor.



(b) Experiment in a corridor with wide space.

Fig. 11 Simulation used a STSF scheme.

Fig. 11 shows multi sensor STSF scheme is applied for the measurement. And the results are compared to show the superiority of the proposed scheme. The AGV was allowed to move keeping the distance between AGV and obstacles constant at the region, A_T and B_T. The region B_T, shows the improvement in steering at corner. And the

simulation experiments show that an AGV, utilizing our scheme, can avoid obstacles and reach a given goal position in the workspace of a wide range of geometrical complexity. Experiments results using new STSF, show the AGV can avoid obstacles (boxes and trash can) and follow the wall. Fig. 10 through Fig. 11 demonstrate one of many successful experiments. The algorithm is very effective in escaping local minima encountered in laboratory environments.

The AGV navigates along a corridor with 3m width and with some obstacles as shown in Fig. 10 and Fig. 11. It demonstrates that the AGV avoids the obstacles intelligently and follows the corridor to the goal.

Also notice that especially at the region, A_T, the errors of the AGV position converse to zero as the same reason, referring to the simulation result and experimental result in Fig.. 10 (a) and 10 (b) respectively, Fig. 11 (a) and 11 (b) represent the reference of AGV direction produced by the proposed STSF. Finally, the AGV is tested to follow the whole trajectory from start position to final position as shown in Fig. 10 and Fig. 11.

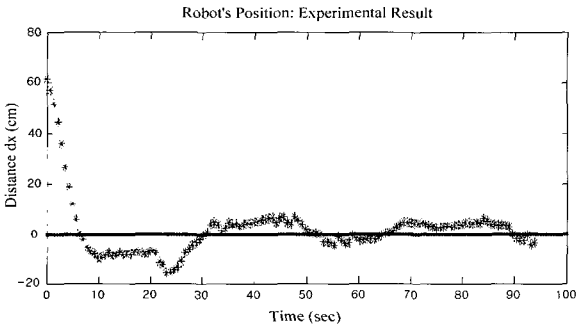


Fig. 12 AGV's position experiment results.

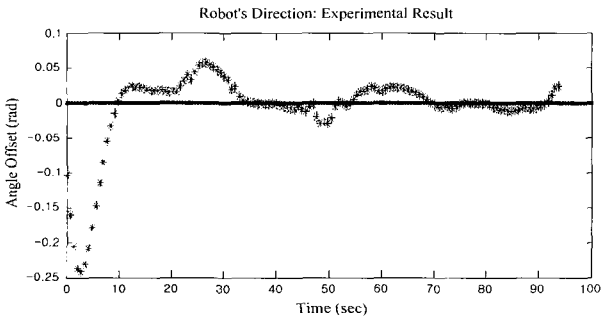
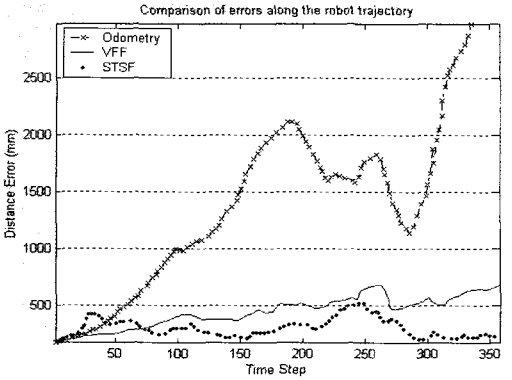


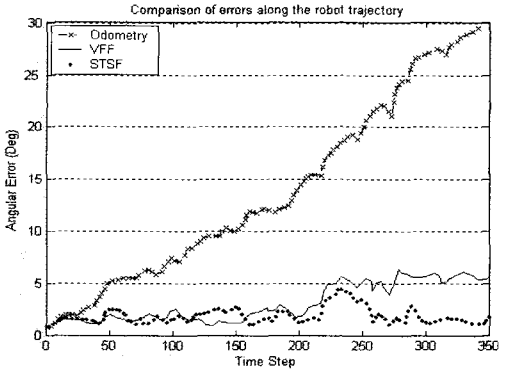
Fig. 13 AGV's direction experiment results.

The experimental results of the AGV status for distance between AGV and corridor wall under such control strategy are given in Fig. 12 and 13.

5.3. Experimental Comparison



(a) distance error.



(b) angular error.

Fig. 14 Comparison of errors along the AGV trajectory.

Table 2 Summary of the solution obtained by the different approaches.

Maximum errors $(x_{max}, y_{max}, \text{and } \phi_{max})$, Distance a_{rel} and Angular ϕ_{rel} Error growing rates results.

	Odometry	VFF	STSF
x_{max}(mm)	1622.22	426.44	283.02
y_{max}(mm)	2144.70	355.62	202.65
ϕ_{max}(deg)	27.20	4.44	3.93
d_{rel}(mm/m)	47.4	10.8	7.8
ϕ_{rel}(deg/m)	0.432	0.144	0.114

Table 2 compares the largest errors obtained by the STSF method with the approach neglecting correlations. Clearly, the STSF method obtained estimations for the AGV localization with an upper bound of around 5.9 mm/m of the total trajectory length for the distance error and 0.04°/m for the orientation error, which represented an order of magnitude below dead reckoning errors.

Therefore, pairings between local observations at time k and previously stored knowledge up to time k 1 were found at each point along the AGV trajectory even when the vehicle returned to previously learned places of the navigation area. An average of 74% of the number of available observations were matched with previous known features. Fig. 14 compares the real errors along the AGV path, obtained by each approach. The highest errors correspond to odometry based navigation, while the smallest correspond to the STSF method. From the Fig. 14 note how the location uncertainty of the AGV decreases when the vehicle revisits previously learned places (i.e., from trajectory point 40 onwards).

6. Conclusions

In this paper, a new sensor fusion concept, STSF (space and time sensor fusion), was introduced. The effectiveness of STSF was demonstrated through the examples, simulations and experiments. To generate complete navigation trajectories without *a prior* information on the environment, not only the data from the sensors located at different places but also the previous sensor data are inevitably utilized. Although we have tried using the sonar system for map building and navigation in indoor environment, the result from the above experiments clearly shows that by utilizing both systems and applying active sensing to adapt to differing situation, a high level of competent collision avoidance behavior by STSF can be achieved.

Sonar system and visual systems are cooperatively utilized for collision avoidance based upon STSF such that an AGV was successfully navigated in an unstructured environment as well as in a structured environment. Although we have used an active CCD camera system for landmark recognition and navigation in outdoor environment, the results from the experiments clearly suggest that by utilizing the AGV and by adapting the active sensing to different situations, a high level of competent collision avoidance behavior can be achieved with the STSF method.

Based on these results, further experiments to be carried out will aim at applying the STSF method to the control of an AGV in an unstructured environment with various sensors.

Acknowledgement

This work was partially supported by the KOSEF through the CIIPMS at Dong A University and by the RICIC of Pusan National University.

References

[1] A. P. Dempster, N. M. Laird, and D. B. Rubin (1977), "Maximum likelihood from incomplete data via the EM algorithm," J. R. Statist. Soc., Vol. 39, pp. 138.

[2] Jang M. Lee, B. H. Kim, M. H. Lee, M. C. Lee, J. W. Choi, and S. H. Han (1999), "Fine Active Calibration of Camera Position/Orientation through Pattern Recognition," Proc. of IEEE Int'l. Symp. on Industrial Electronics, pp. 100-105.

[3] J. Llinas and D. L. Hall (1994), "A challenge for the data fusion community II: Infrastructure imperatives," in Proc. 7th Nat'l. Symp. on Sensor Fusion, Albuquerque, pp. 134-138.

[4] L.A. Zadeh (1973), "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes," IEEE Transactions on Systems, Man and Cybernetics, 3(1): pp. 28-44.

[5] L. Hong, A. Lynch (1993), "Recursive temporal-spatial information fusion with applications to target identification," Aerospace and Electronic Systems, IEEE Transactions on, Vol. 29, Issue. 2, pp. 435-445.

[6] M. Rombaut, and D. Meizel (2001), "Dynamic data temporal multisensor fusion in the Prometheus Prolab2 demonstrator," IEEE Int. Conference on Robotics and Automation, San Diego, pp. 36-76.

[7] P. Weckesser and R. Dillman (1997), "Navigating a Mobile Service-Robot in a Natural Environment Using Sensor-Fusion Techniques," Proc. of IROS, pp. 1423-1428.

[8] R.C. Luo, Chih-Chen Yih, and Kuo Lan Su (2002), "Multisensor fusion and integration: approaches, applications, and future research directions," Sensors Journal, IEEE, Vol. 2, Issue. 2, pp. 107-119.

[9] R. C. Luo and K. L. Su (1999), "A Review of High level Multisensor Fusion: Approaches and applications," Proc. of IEEE Int'l. Conf. on Multisensor Fusion and Integration for Intelligent Systems, pp. 25-31..

[10] R. C. Smith, P. Cheeseman (1986), "On the Representation and Estimation of Spatial Uncertainty." in the International Journal of Robotics Research, Vol. 5, no. 4, pp. 56-68.

[11] Ren C. Luo and Kuo L. Su (1999), "A Review of High-level Multisensor Fusion: Approaches and applications," Proc. of IEEE Int'l. Conf. on Multisensor Fusion and Integration for Intelligent Systems, pp. 25-31.

[12] S. Thrun (1996), "A Bayesian approach to landmark discovery and active perception for mobile robot

navigation," (Tech. rep. CMU-CS-96-122). Pittsburgh, PA: Carnegie Mellon University, Department of Computer Science.

[13] V. Ayala, J.B. Hayet, F. Lerasle, and M. Devy (2000), "Visual localization of a mobile robot in indoor environments using planar landmarks," Intelligent Robots and Systems, (IROS 2000). Proceedings. IEEE/ RSJ International Conference on, Vol. 1, pp. 275-280.

[14] W. Pieczynski (2000), "Unsupervised DempsterShafer fusion of ependent sensors," in Proc. IEEE Southwest Symp. Image Analysis and Interpretation (SSIAI'2000), pp. 247-251.

Received 11 March 2003
Accepted 14 August 2003