

Multilevel Multiagent Based Team Decision Fusion for Autonomous Tracking System*

Tse Min Chen† and Ren C. Luo†

Abstract: Multilevel fusion is a key issue for developing the decision-making kernel of multilevel systems. This article presents a formulation of predictive decision-making algorithm for multilevel multiagent based team decision-making system with the I/O mode characterizations of feature in–decision out or data in–decision out methods. The sequential data fusion is conducted through a dynamic behavior modeling method capable of estimating the observed system parameters from the raw sensory measurements over period of time. The temporal estimated model is used to forward prediction of the observed system output for decision-making. A self-evaluation method to estimate the prediction quality is used to generate the individual decision confidence for final decision integration, which is conducted through a multi-layered fuzzy linguistic reasoning engine. The method is implemented for an autonomous tracking system that consists of a target tracking agent whose whose inputs are visual and ultrasonic range measurements and a collision avoidance agent whose inputs are speed data. The experimental results conducted by a mobile robot and intelligent electrical wheelchair will demonstrate the feasibility, accuracy, and robustness of the system based on the multisensor fusion method.

Keywords: Multisensor Fusion, Multiagent Team Decision, Target Tracking

1. Introduction

A SYSTEM only with single perceptual sensor has its inherent limit of capabilities. Due to the possible weaknesses of uncertainty, missing observation, and incompleteness of single sensor, there is a growing need to integrate and fuse multisensor information for advanced systems with high robustness and flexibility. The multimedia sensor fusion system for realistic application on automated videoconferencing and surveillance[1] is an excellent example of operational fusion systems. It integrates the stereo sound and color video information for improving the system ability and performance to detect and track human speaker. The accurate differential global positioning system[2] that integrates rate gyro, speedometer, and GPS receivers for the application on vehicle localization is another successful example of multiple sensor fusion systems.

The multisensor fusion and integration concept has been applied in a wide variety of application fields, such as military command, image processing, robotics, automation, and environmental monitoring. Using multisensor fusion techniques on a system may enhance the quality of system performance of reliability, robustness, confidence, efficiency, and resolution as Varshney’s discussion [3]. Luo and Kay suggest the taxonomy with four levels of sensor fusion: signal level, pixel level, feature level, and symbol level[4]. Time varying sensor measurements corrupted with noise can be fused at signal level where great degree of sensory registration is required. Pixel-level fusion can be used to combine multiple images for richer information content. The environmental abstraction extracted from raw sensor data can be combined by the feature-level fusion to increase the likelihood or obtaining additional composite features. Symbol-level fusion is used for making final decision through the symbolic reasoning processes that integrate multiple symbols derived from sensory aspect of environment.

A complete intelligent system may involve multiple-level hierarchical sensor fusion processes for final decision [1], [2]. For design and implementation it can be simplified to consider the following fusion processes based on I/O characterizations as Varshney’s classification [3], i.e. data in–data out, feature in–feature out, decision in–decision out, data in–feature out, feature in–decision out, and data in–decision out. Dasarathy’s classification [4] doesn’t consider the data in–decision out process, but add a temporal fusion that is used for data/feature/decision integration over a period of time. The latest two fusion types of Varshney’s classification are especially important in most intelligent motion control systems that required qualified performance of fast-response efficiency, high flexibility, and robustness, such as navigating a mobile robot [6] in uncertain environment, mobile tracking of various targets with nonlinear motion behavior [7], and the autonomous land vehicle road following[8].

To reach the qualified performance through low-cost processing units and multiple sensors of the system, it is needed to reduce the complexity of the multilevel fusion. The ALVINN project made such efforts on navigating control of a land vehicle on a highway using a neural network that inputs color image pixels and outputs motion decision [8]. The integrated multi-behavior system [6] by multi-layered fuzzy logic reasoning engine with multisensor input and motion decision output has also contributed in this area.

©1999 Cyber Scientific

---

* Received October 22, 1999; accepted December 18, 1999. This work was supported by National Science Council, Taiwan, R.O.C., under the reference number, NSC-87-2212-E-194-002.
† Department of Electrical Engineering, National Chung Cheng University, 160 Shang-Shang, Ming-Huang, Chia-Yi, Taiwan 621, R.O.C. E-mail: {tunchen, luo}@ia.ee.ccu.edu.tw

Paper No. 1345–269X/99/020063-07
Final Decision

Defuzzification

Linguistic Inference

Other Agents Local Decisions

Agent #2 Local Decision

Fuzzification

Fuzzification

Fuzzification

Fuzzification

Environmental Recognition

Self-evaluation

Signal/Pixel Level Fusion Modules

Feature Extraction

Self-evaluation

Adaptive Modeling

CCD Model

Linguistic Inference

Linguistic Inference

Linguistic Inference

Agent #1 Local Decision

Fig. 1 The conceptual architecture for sensor fusion across multiple levels

For the system uses multiple sensors with different resolutions or with uncertain noisy outputs of data for motion decision, it is difficult to employ a method for accomplishing sensor fusion across levels. There are two major reasons, one is the little common of sensory output and the other is the different acquisition frequency of the sensor modules. The two problems can be solved based on the concept of signal-level and symbol-level fusion individually. Signal-level fusion requires stochastic filter that needs high degree of sensory registration information in the fusion processes. The Kalman filter based method for global positioning system [2] is an example of signal-level fusion. The key point of the symbol-level fusion is to integrate multiple symbols of sensory abstraction by referring to the symbol evidential confidence. The use of D-S (Dempster-Shafer) theory to integrate four types of sensors for environmental recognition is an example [9] of symbol-level fusion.

This article presents a formulation of predictive decision-making algorithm for multilevel multiagent based systems with the I/O mode characterizations of feature in–decision out or data in–decision out. The proposed method involves a dynamic behavior estimator for data prediction over a period of time to solve the problem caused by various sensor-response time and multi-layered fuzzy reasoning engine to integrate the predicted data and confidence sets. The predictive fusion method can provide the system with faster response performance [7] and can compensate the sensory data loss caused by the potential detection error or time-delay. The multi-layered fuzzy reasoning engine can perform fast fusion for multi-dimensional input [6] at the signal or symbol levels. The method is implemented for an autonomous tracking system that consists of a target tracking agent whose inputs are visual and ultrasonic range measurements and a collision avoidance agent whose inputs are ultrasonic range measurements. The experimental results conducted by a mobile robot and intelligent electrical wheelchair will demonstrate the feasibility, accuracy, and robustness of the system based on the multisensor fusion method.

2. Conceptual Architecture for Multi-Level Fusion

Advanced multisensor based systems for some sets of goals or tasks always involve a team of local decision maker that works cooperatively to solve decision problems [10]. Each of local decision makers in the system can be treated as an agent who is an expert capable of lower level fusion to suggest recommendations for the global decision maker [11]. The function of the global decision maker is to fuse the local decisions from the agents to derive the team decision using symbol-level fusion. Therefore each agent needs to perform sensor fusion across levels and the global decision maker works in high-level symbolic fusion. The objective in this research is to develop a flexible multilevel fusion method for decision-making of local agents and the global decision maker.

The conceptual architecture for decision making from the multilevel fusion of the multiple time-varying data, features, and symbols is shown in Fig. 1 which is based on the four levels of Luo and Kay’s taxonomy [5]. In the lower-level fusion of time-sequential data fusion, we first assume the parameters of detected target model are unknown in a priori. The adaptive modeling modules are used to online estimate the temporal-change of dynamic parameters.
Based on the estimated model parameters it can perform the prediction for the coming sensory data/measurements for higher-level fusion. The look-ahead method has the advantage of fast error convergence rate for high performance systems, and can perform data extrapolation when data loss problem occurs, but we have to evaluate the confidence of the prediction to ensure the validity. For this purpose, the self-evaluation module calculates the confidence according to the modeling accuracy and timing parameters. The details of this issue are discussed in Section 3.

Fuzzy inference theory has been successfully applied in many data fusion applications. For higher-level fusion and decision-making we use the multi-layered fuzzy inference method. The inputs for fusion process are first translated to fuzzy linguistic representation by the membership functions, e.g. the triangular function shown in Eq. (1):

$$\mu_A(x) = \begin{cases} 
0 & \text{for } x < \alpha \\
\frac{x-\alpha}{\gamma-\alpha} & \text{for } \alpha \leq x < \gamma \\
\frac{\beta-x}{\beta-\gamma} & \text{for } \gamma \leq x < \beta \\
0 & \text{for } x \geq \beta
\end{cases}$$

(1)

where $\gamma = (\alpha + \beta)/2$. Through the membership functions, an input $x$ falls into the range of $\alpha, \beta$ can be represented by a set of such fuzzy linguistic $F_\gamma = [\mu_{a_1}, \mu_{a_2}, \ldots, \mu_{a_n}]$. The fuzzy linguistic interference process is a rule-based method, e.g. Mamdani method[12]. The final decision is accomplished through a defuzzification function.

Each of the linguistic inference blocks in the architecture may include multiple input lines, where the input number is dependent on the application. To consider the system efficiency and easy implementation, the linguistic inference block is designed by multi-layered fuzzy inference engine. Each inference process is called elementary engine that handles two inputs and outputs a fusion data of fuzzy linguistic representation as the example shown in Fig. 2.

The elementary engine performs fusion looking up tabular and minimum-maximum processes such as the Mamdani method[12]. For a linguistic inference block with $m$ inputs the total number of the elementary engines and needed fuzzy rule sets are $2m^2/2 - 1$ if $m$ is even; otherwise $2(m-1)^2/2$. Therefore, for a 5-level fuzzy linguistic representation the total rules needed in the $m$-input block are $25 \times (2^{m/2} - 1)$ if $m$ is even; otherwise $25 \times 2^{(m-1)/2}$. Compared to the single-layered inference engine that needs $5^m$ rules, the multi-layered method will result in a better efficiency for the system with large amount of inputs [1].

3. Adaptive Modeling of Sequential Sensory Data

The problem addressed in this section is to explore the dynamic modeling method of the time-varying sensory measurement $s(t)$. For continuously detecting an unknown dynamic target we can model the motion dynamics using nth-order ordinary differential equations as follows:

$$\frac{d^n z(t)}{dt^n} + a_1 \frac{d^{n-1} z(t)}{dt^{n-1}} + \cdots + a_n z(t) = b$$

(2)

where $z(k) = \sum_{i=1}^{k} s(t)$ is the accumulated generating operator (AGO). In grey theory, Eq. (2) is called “white descriptor” for modeling the system so that we can find its parameters $(a_1, a_2, \ldots, a_n, b)$ directly from the observed system outputs $s(t)$. To estimate the parameters of the partial-known system or called “grey system,” it is approximated by the following grey-differential equation [13]:

$$\frac{d^n z(t)}{dt^n} + a_1 \frac{d^{n-1} z(t)}{dt^{n-1}} + \cdots + a_n g(t) = b$$

(3)

where $g(t) = (z(t+1) + z(t))/2$. We can obtain the optimal parameters $(a_1, a_2, \ldots, a_n, b)$ by least square estimation algorithm by introducing the accumulated generating operation in a time interval. For simplification, the sampled time of past measurement is taken as a unit. The derivative terms from $m = 1$ to $n$ of Eq. (3) in a discrete system can be written as:

$$\frac{d^m z(t)}{dt^m} = \frac{d^{m-1} z(t)}{dt^{m-1}} - \frac{d^{m-2} z(t)}{dt^{m-2}} + \cdots + a_n g(t)$$

(4)

Substitute the sequential data of $X$ and $Z$ in the time interval $t = [1, 2, \ldots, \xi]$ into Eq. (3), we get the following matrix relation:

$$Y = \begin{bmatrix} 
\frac{d^n z(2)}{dt^n} & \frac{d^n z(3)}{dt^n} & \cdots & \frac{d^n z(\xi)}{dt^n}
\end{bmatrix} = [A B] \phi$$

(5)

where

$$A = \begin{bmatrix}
-\frac{d^{n-1} z(2)}{dt^{n-1}} & -\frac{d^{n-2} z(2)}{dt^{n-2}} & \cdots & -\frac{d^2 z(2)}{dt^2} \\
-\frac{d^{n-1} z(3)}{dt^{n-1}} & -\frac{d^{n-2} z(3)}{dt^{n-2}} & \cdots & -\frac{d^2 z(3)}{dt^2} \\
\vdots & \vdots & \ddots & \vdots \\
-\frac{d^{n-1} z(\xi)}{dt^{n-1}} & -\frac{d^{n-2} z(\xi)}{dt^{n-2}} & \cdots & -\frac{d^2 z(\xi)}{dt^2}
\end{bmatrix},$$

$$B = \begin{bmatrix}
-\frac{1}{2}(z(2) + z(1)) \\
-\frac{1}{2}(z(3) + z(2)) \\
\vdots \\
-\frac{1}{2}(z(\xi) + z(\xi - 1))
\end{bmatrix},$$

and

$$\phi = [a_1 a_2 \cdots a_n b]^T.$$
minimizing the least square error term using the matrix derivation with respect to \( \phi \), the optimal solution
\[
\hat{\phi} = [\hat{a}_1 \hspace{1em} \hat{a}_2 \hspace{1em} \cdots \hspace{1em} \hat{a}_n \hspace{1em} \hat{b}]^T
\]
can be obtained by the following equation:
\[
\hat{\phi} = \left( [A B]^T [A B] \right)^{-1} [A B]^T Y. \tag{6}
\]

Similarly, for the case \( \xi \leq n + 1 \), we can calculate the optimal parameters by minimum-norm method that minimizes the \( (\hat{\phi}^T \phi)/2 \) under the constraint of \( Y = [A B]\phi \). The corresponding solution for \( \phi \) is obtained by the following equation:
\[
\hat{\phi} = [A B]^T \left( [A B] [A B]^T \right)^{-1} Y. \tag{7}
\]

The estimated parameters can be brought into the response solution of the \( n \)th-order ordinary differential equation Eq. (2) for the prediction of the accumulated generating operator value at \( t > \xi \). In real applications, the higher order AGO model of Eq. (2) is more sensitive to the signal input and the solution of parameters is more difficult to be found than lower level model. For example, the solution of the first order AGO model is:
\[
\dot{z}(t + 1) = \left( z(1) - \frac{\hat{b}}{\hat{a}_1} \right) e^{-\hat{a}_1 t} + \frac{\hat{b}}{\hat{a}_1} \tag{8}
\]
and the solution of 2nd-order AGO model is:
\[
\dot{z}(t + 1) = c_1 e^{-\lambda_1 t} + c_2 e^{-\lambda_2 t} + \frac{\hat{b}}{\hat{a}_2} \tag{9}
\]
where \( \lambda_1 \) and \( \lambda_2 \) are the eigenvalues of the system. Finally, the prediction result \( s_p \) of the data at \( t = \xi + \Delta t \) can then be obtained by:
\[
s_p(\xi + \Delta t) = \dot{z}(\xi + \Delta t) - \dot{z}(\xi + \Delta t - 1). \tag{10}
\]

4. Self-Evaluation of the AGO Model

To evaluate the dynamic models we can bring the variable \( t \) (i.e. \( t = 1, 2, \cdots, \xi \)) into Eq. (10) and calculate the error between the estimated data and the measurements by the following equation:
\[
\varepsilon(t) = s_p(t) - s(t). \tag{11}
\]
The average of absolute errors can be obtained through the following equation:
\[
\bar{\varepsilon} = \frac{1}{\xi} \sum_{t=1}^{\xi} |\varepsilon(t)| \tag{12}
\]
and the error variance is calculated by:
\[
\sigma = \sqrt{\frac{1}{\xi} \sum_{t=1}^{\xi} (|\varepsilon(t)| - \bar{\varepsilon})^2}. \tag{13}
\]
The variance of the data can be described as:
\[
\rho = \sqrt{\frac{1}{\xi} \sum_{t=1}^{\xi} (s(t) - \bar{s})^2}, \tag{14}
\]
where \( \bar{s} = (1/\xi) \sum_{t=1}^{\xi} s(t) \). The quality index \( Q \) of the learned AGO model is defined by the ratio of the two variances as the following equation:
\[
Q = \frac{\sigma}{\rho}. \tag{15}
\]
According to the \( Q \) we can evaluate the confidence of the prediction, i.e. the confidence is proportional to \( Q^{-1} \). Beside, the confidence is also proportional to the inverse of time-delay \( \Delta t \). Therefore, we define the confidence \( C \) of current prediction based on the latest learned AGO model as Eq. (16):
\[
C(t) = \exp \left( -\frac{Q \Delta t}{\tau} \right) \tag{16}
\]
where impedance coefficient \( \tau > 0 \) can be set according to the sensor parameter of accuracy.

5. Illustrative Application

This section presents the experiment of autonomous target tracking system using the adaptive modeling method and the multi-layered fuzzy reasoning engine based on the conceptual architecture. The experimental setup consists of one autonomous mobile robot and a multisensor-based electrical wheelchair as shown in Fig. 3. The mobile robot “Chung Cheng-1” [6] is a three-wheel mobile platform equipped with a vertical sliding manipulating arm and other sensory modules. The experimental target is the multisensor-based electrical wheelchair named “Luoson-3” which was developed in our laboratory. The mobile platform is driven by the differential velocity from the two individually controlled rear wheels. The Chung Cheng-1 served as a tracker and it is desirable to keep a constant distance from the target Luoson-3 in the dynamic changed environment with unknown obstacles. When the tracker is tracking the target the man may walk through the interval casually between the target and tracker as shown in the Fig. 3.

5.1 System implementation

Two types of sensors for target tracking agent are equipped on our autonomous mobile robot “Chung-Cheng-1” which is used in the experiments, one is an ultrasonic range sensor and the other is a color vision system. The ultrasonic sensor is used directly to measure the distance
between the target and the robot. The color vision is based on the color-histogram backprojection method [14] to detect the marked tag of target. The ultrasonic range sensor has the advantage of high-frequency detection rate of 15 Hz, but its resolution is limited in 1 [in]. The visual detection is operating in 5 Hz and its resolution is limited in 0.1 [in]. One major drawback of the vision system based on the histogram backprojection method is that the measurements may contain noisy uncertainty due to the marked area size. The distance measurement from vision system is floating in different tests even both the target and tracker are static. Therefore, the fusion of these two sensing results will be meaningfully necessary for accurate motion decision-making.

**Figure 4** illustrates the implementation structure of the autonomous target tracking system. It contains two major agents for local decisions, one is the target-tracking agent whose inputs are the input position measurements from ultrasonic and vision sensors and the other is the collision-avoidance agent whose inputs are the surrounding range measurements from 16 ultrasonic sensors. The target-tracking agent is shown in the shadowed area in Fig. 4, where $u_x$ is the sequential measurements of distance between robot and target from the ultrasonic range sensor. Similarly, $v_x$ is the distance measurement from the visual detection and $s_x$ is the robot driving velocity. The local decision relative to target position is made by fusion of the two sensory data, the error of predicted distance and the desired distance 40 [in], and the error changes. It will output the local decision $D_1$ and the relative confidence $C_1$ of $D_1$, where $C_1$ is calculated from the fusion of the two prediction confidences and the difference of the two predictions. The local decision represents the relative velocity of target and tracker.

The collision avoidance agent is based on our previous developed method in [15], which integrates the ultrasonic range sensor array equipped on the mobile robot to detect obstacle and make motion decision. The confidence of the decision from collision avoidance module is inversely proportional to the distance from robot to the closest obstacle. The final decision calculated by fusing the two local decisions is the absolute driving velocity of mobile robot.

### 5.2 Experimental results

The target tracking results for the period of 45,000 [ms] are shown in **Fig. 5**. The robot follows the target by keeping a distance of 40 [in] from the target. It starts on tracking the target from time 0 [ms] and slows down to avoid the potential collision to the human at 7,500–12,000 [ms] as shown in the Fig. 5(a). After the human passes the interval between robot and target it continuously tracks the target and smoothly approaches the target at 12,000–35,500 [ms], and so on. Figures 5(a) and (b) are the measurement and prediction results from vision system and the ultrasonic sensor respectively. The measurements are plotted as dot points in the figures and the prediction results are plotted as marked lines. At time interval 7,500–12,000 [ms] the human walk across the interval between the robot and
target. This causes the vision system failure on detecting the target (i.e. the zero value of the plot) at time interval 8,000–12,000 [ms]. For the case, the system extrapolates the target position by prediction based on the previously obtained AGO model and the self-evaluated confidence is decreased as the time increased as shown in Fig. 5(c). The case at time interval 35,500–39,500 [ms] is similar. In the meanwhile, the measurements from ultrasonic sensor are the distance between target and obstacle. This results the target tracking module to make a rapid deceleration decision for the tracker robot as shown in Fig. 5(e). Due to the large difference of the two predictions and the low confidence of the vision prediction, the decision confidence of target tracking module becomes low (e.g. see the Fig. 5(d) \(t = 7,500–12,000 \text{ [ms]}\)). On the other hand, the collision avoidance module reports a high confidence of control decision at the periods. This leads the robot slowing down to avoid the potential collision according to the fusion of the two decisions.

Without obstacle interrupting (i.e. time 0–7,500, 12,000–35,500, and 39,500–45,000 [ms]) the tracking system operates as a fusion mechanism for integration of the redundant data under feature level. The obstacle avoidance module reports a maximum decision 70 for the robot but the confidence is zero (i.e. useless for final decision fusion). Therefore, the final decision in the case is fully

---

**Fig. 5** Experimental results of the intelligent target tracking system: (a) the vision measurement and prediction results, (b) the ultrasonic sensor measurement and prediction results, (c) confidences of target position prediction, (d) confidence of local decisions, and (e) local decisions and final decisions
dependent on the local decision from target tracking module. The overall result of this experiment shows the success of developing an intelligent target tracking system in real robot applications. It also demonstrates the capability of adaptive multilevel fusion and decision kernel for the robust systems.

6. Conclusions

This paper has presented a formulation of predictive decision-making algorithm for multilevel multiagent based team decision systems with the I/O mode characterization of feature in–decision out or data in–decision out. The illustrative experimental results of autonomous target-tracking application, i.e., a typical multilevel multiagent based team decision system, have shown the success of the method. In addition to the multilevel fusion characteristic, the generalized architecture and the adaptive modeling method provide the system designers with systematic and modularized means to design the decision kernels of the systems with multiple sensors across levels without the need of a priori model parameters of the observed data dynamics. In the future, we will focus on real-time embedded intelligent system based on this multilevel multiagent based team decision architecture.

Acknowledgments

The authors wish to express their sincere appreciation to Mr. C. Y. Hu for his enthusiastic support on constructing the second mobile robot “Luoson-1” in the laboratory. The authors thank for the support of this research by National Science Council, Taiwan, R.O.C.

References


Biographies

Tse Min Chen received B.S. degree in electrical engineering in 1993 from Feng Chia University and the M.S. degree in electrical engineering in 1995 from National Chung Cheng University. He is currently working toward the Ph.D. degree in electrical engineering. He is a research assistant in the Department of electrical engineering at National Chung Cheng University. His research interests include Internet based robotics, multiagent multiprocessor system, target tracking, and behavior fusion.

Ren C. Luo received the Diplom.-Ing. and Ph.D. degrees from Technische Universituet Berlin, Berlin, Germany, in 1979 and 1982, respectively. He is currently a Professor in the Department of Electrical Engineering and the Dean of the College of Engineering, National Chung Cheng University. From 1992 to 1993, he was the Toshiba Endowed Chair Professor in the Institute of Industrial Science, University of Tokyo. From 1990 to 1995 he was a Professor in the Department of Electrical and computer Engineering and the Director of Center for Robotics and Intelligent Machines, North Carolina State University, Raleigh, North Carolina. Dr. Luo is a fellow of IEEE and is currently the President of IEEE Industrial Electronics Society.