Adaptive Fusion by Reinforcement Learning for Distributed Detection Systems

NIRWAN ANSARI, Senior Member, IEEE
EDWIN S. H. HOU, Member, IEEE
BIN-OU ZHU
JIANG-GUO CHEN
New Jersey Institute of Technology

Chair and Varshney have derived an optimal rule for fusing decisions based on the Bayesian criterion. To implement the rule, the probability of detection $P_D$ and the probability of false alarm $P_F$ for each detector must be known, but this information is not always available in practice. An adaptive fusion model which estimates the $P_D$ and $P_F$ adaptively by a simple counting process is presented. Since reference signals are not given, the decision of a local detector is arbitrated by the fused decision of all the other local detectors. Furthermore, the fused results of the other local decisions are classified as “reliable” and “unreliable.” Only reliable decisions are used to develop the rule. Analysis on classifying the fused decisions in terms of reducing the estimation error is given, and simulation results which conform to our analysis are presented.

I. INTRODUCTION

Distributed detection systems with data fusion have been investigated widely in recent years [1–6]. The problem of decision fusion in a binary hypothesis system was considered by Chair and Varshney [1], and Thomopoulos et al. [6]. Chair and Varshney [1] developed an optimal decision rule by using the minimum probability of error criterion. Thomopoulos et al. [6] proposed the optimal decision rule based on the Neyman-Pearson test. They showed that the optimal fusion rule is obtained by a weighted sum of local decisions through a hard limiter. The weight associated with each local detector indicates the degree of reliability of the detector. Each weight is a function of the probability of detection $P_D$ and the probability of false alarm $P_F$ of the detector. The $P_D$ and $P_F$ can be obtained when either the distribution of the observations at each detector is given, or when some reference signals are provided to estimate the $P_D$ and $P_F$ by an empirical method. However, in practice, neither $P_D$ nor $P_F$ is known. Furthermore, since the sensors are usually exposed to a changing environment, the performance of each individual detector may not always be the same, i.e., the $P_D$ and $P_F$ may vary with time. We propose an adaptive system to estimate the $P_D$ and $P_F$. Without knowledge of the performance of each detector, the proposed system is capable of approximately estimating the $P_D$ and $P_F$ of the detector in the course of performing the decision fusion.

Consider a binary hypothesis testing system consisting of $n$ local detectors with the probabilities of two hypotheses $H_0, H_1$ denoted as $P(H_0) = P_0$ and $P(H_1) = P_1$, respectively. Assume that under each hypothesis, the observations at each detector are statistically independent. Let $u_i$ and $u$ denote the decisions made by the $i$th detector and the fusion center, respectively. When the $i$th local detector favors the $H_1$ hypothesis, $u_i = +1$; otherwise $u_i = -1$. The output $u$ is similarly defined. We let $P_{Di}$ and $P_{Fi}$ denote the probability of detection and the probability of false alarm of the $i$th detector, respectively.

Chair and Varshney [1] showed that the optimal fusion rule for the minimum probability of error criterion is

$$ u = \begin{cases} +1, & \text{if } a_0 + \sum_{i=1}^{n} a_i u_i > 0 \\ -1, & \text{otherwise} \end{cases} $$

(1)

where

$$ a_0 = \log \frac{P_1}{P_0} $$

(2)

$$ a_i = \begin{cases} \log \frac{P_{Di}}{P_{Fi}}, & \text{if } u_i = +1 \\ \log \frac{1-P_{Fi}}{1-P_{Di}}, & \text{if } u_i = -1 \end{cases} $$

(3)
For the case \( P_0 = P_1 \) and the probability of false alarm \( P_{Fj} \) is equal to the probability of miss \( P_{Mj} \), \( a_0 = 0 \) and the optimal fusion rule can be simplified to

\[
u = \begin{cases} 
+1, & \text{if } \sum_{j=1}^{n} w_j u_j > 0 \\
-1, & \text{otherwise}
\end{cases}
\] (4)

where

\[
w_j = \log \frac{P_{Dj}}{P_{Fj}}, \quad \text{for each } j.
\] (5)

The system structure is shown in Fig. 1, where

\[
y = \sum_{j=1}^{n} w_j u_j.
\] (6)

The structure shown in Fig. 1 is similar to a single neuron system, in particular, the perceptron [7]. If reference signals are given, they can be used as a “reference” to train the system such that the weights converge to the optimal values defined by (5). However, in practice, such a reference is not readily available and at the same time, the \( P_D \) and \( P_F \) of a detector may vary with time. Since the fused decisions are usually better than local decisions, they can be considered as the reference. When the \( i \)th local decision \( u_i \) is equal to the fused decision \( u \), then \( u_i \) is considered to be correct; otherwise, \( u_i \) is considered to be incorrect. Since \( u = \sgn(y) = \sgn(\sum_{j=1}^{n} w_j u_j) \), the fused decision \( u \) has already taken into account the decision of the \( i \)th detector, \( u_i \). If \( u \) is used as a reference for \( u_i \), a bias is established for \( u_i \). Thus, in the proposed system, the decision of the \( i \)th local detector \( u_i \) is arbitrated by the fused decision of all the other \( (n - 1) \) local detectors. Denote the fused decision as \( \bar{u}_i \), and define

\[
y_i = \sum_{j \neq i} w_j u_j
\] (7)

i.e., \( y_i \) is the weighted sum of all local decisions except \( u_i \), then

\[
\bar{u}_i = \sgn(y_i).
\] (8)

Note that \( \bar{u}_i \) and \( u_i \) are conditionally independent given \( H_j \), \( j = 0, 1 \). The “reference” \( \bar{u}_i \) may not always be correct. To reduce the possibility of using incorrect references, the decisions \( \bar{u}_i \) are further classified. The decision \( \bar{u}_i \) is considered unreliable when the weighted sum defined by (7) is close to the decision threshold 0. Our strategy is to determine an “unreliable range” around the decision threshold such that when the weighted sum \( y_i \) falls in this range, the fused decision \( \bar{u}_i \) is considered unreliable and will not be used for training the system. The selection of this unreliable range is discussed next.

II. ADAPTIVE MODEL ANALYSIS

Consider the structure shown in Fig. 1. From (6) and (7), we have

\[
y_i = y - w_i u_i.
\] (9)

Under the assumptions that \( P_0 = P_1 \) and \( P_{Fj} = P_{Mj} \), the conditional probability mass functions \( f_i(y_i|H_1) \) and \( f_i(y_i|H_0) \) are symmetric with each other, i.e., \( f_i(y_i|H_1) = f_i(-y_i|H_0) \), as shown in Fig. 2.

We establish the above relationship as follows:

\[
y_i = \sum_{j \neq i} w_j u_j = \sum_{j \in S^+} w_j - \sum_{j \in S^-} w_j
\] (10)

where \( S^+_i = \{j : j \neq i \text{ and } u_j = 1\} \) and \( S^-_i = \{j : j \neq i \text{ and } u_j = -1\} \). By the earlier assumption of independent observations,

\[
P(y_i = \xi|H_1) = \prod_{S^+_i} P(u_j = 1/H_1) \prod_{S^-} P(u_j = -1/H_1)
\] (11)

where \( S_i = \{S^+_i, S^-_i\} \); combinations of \( S^+_i \) and \( S^-_i \) such that \( \sum_{S^+_i} w_j - \sum_{S^-_i} w_j = \xi \). Since we have assumed that \( P_{Mj} = P_{Fj} \), i.e.,

\[
P(u_j = -1/H_1) = P(u_j = 1/H_0) = P_{Fj}
\] (12)

which also implies that

\[
P(u_j = 1/H_1) = P(u_j = -1/H_0) = P_{Dj};
\] (13)

we have

\[
P(y_i = \xi|H_1) = \prod_{S^+_i} P_{Dj} \prod_{S^-} P_{Fj}.
\] (14)

Now, assume that the local detectors, except the \( i \)th detector, make opposite decisions as compared with (11) such that \( y_i \) becomes \(-\xi\). That is, \( S^+_i \) and \( S^-_i \) remain the same, but the decisions are reversed. Thus,
\[
P(y_i = -\xi / H_1) = \sum_{s_i} \prod_{s_i'} P(u_j = -1 / H_1) \prod_{s_i} P(u_j = 1 / H_1)
= \sum_{s_i} \prod_{s_i'} P_{D_i} \prod_{s_i} P_{D_j}
= \sum_{s_i} \prod_{s_i'} P_{F_i} \prod_{s_i} P_{D_j}.
\]

We next evaluate \( P(y_i = \xi / H_0) \),
\[
P(y_i = \xi / H_0) = \sum_{s_i} \prod_{s_i'} P(u_j = 1 / H_0) \prod_{s_i} P(u_j = -1 / H_0)
= \sum_{s_i} \prod_{s_i'} P_{D_i} \prod_{s_i} P_{D_j}.
\]

Since (15) and (16) are the same,
\[
f_i(y_i / H_0) = f_i(-y_i / H_0).
\]

Because of the symmetry of the conditional probability mass function (see (17)),
\[
P(u_i^* = 1 / H_0) = P(u_i^* = -1 / H_1)
\]
\[
P(u_i^* = 1 / H_1) = P(u_i^* = -1 / H_0).
\]

Thus,
\[
\hat{P}_{D_i} = P_{D_i} P(u_i^* = 1 / H_1) + P_{F_i} P(u_i^* = 1 / H_0).
\]

By the same reasoning, we have
\[
\hat{P}_{F_i} = P_{D_i} P(u_i^* = 1 / H_0) + P_{F_i} P(u_i^* = 1).
\]

Let \( \xi_1 < \xi_2 < \cdots < \xi_N \), where \( \xi_1 = \text{max} \{ y_i \} \), be the set of values that \( y_i \) can attain for the \( i \)th local detector. Without loss of generality, let \( \xi_1 < \tau < \xi_N \), and \( k \in \{ 1, 2, \ldots, N \} \) be the smallest integer such that \( \xi_j > \tau \), \( \forall j \leq k \). Define
\[
A = P(\bar{u}_i^* = 1 / H_1) = \sum_{j=1}^{N} P(y_j = \xi_j / H_1)
\]
\[
B = P(\bar{u}_i^* = 1 / H_0) = \sum_{j=1}^{N} P(y_j = \xi_j / H_0).
\]

Then, (21) and (23) can be written as
\[
\hat{P}_{D_i} = P_{D_i} A + P_{F_i} B
\]
and
\[
\hat{P}_{F_i} = P_{D_i} B + P_{F_i} A.
\]

Let \( r_i = \hat{P}_{D_i} / \hat{P}_{F_i} \), \( \hat{r}_i = \hat{P}_{D_i} / \hat{P}_{F_i} \), then, \( \log r_i = w_i \) is the weight of the \( i \)th detector defined by (5) and \(
\hat{r}_i = \hat{w}_i \) is the estimate for \( w_i \),
\[
\hat{r}_i = \frac{\hat{P}_{D_i}}{\hat{P}_{F_i}} = \frac{P_{D_i} A + P_{F_i} B}{P_{D_i} B + P_{F_i} A} = \frac{1 + B}{1 + r_i A} \hat{r}_i A = w_i + \hat{\varepsilon}_i.
\]

As seen in (28), the estimate for the weight is equal to the correct weight plus an error term \( \hat{\varepsilon}_i \), where,
\[
\hat{\varepsilon}_i = \log \frac{1 + B}{1 + r_i A} 
\]
\[
\hat{\varepsilon}_i = \log \frac{r_i A}{r_i B}.
\]

Since \( r_i \) is fixed, \( \varepsilon_i \) will approach 0 as \( B / A \) is approaching 0. We prove that increasing the reliability threshold \( \tau \) will reduce the fraction \( B / A \), and thus the error.
For notational convenience, let \( p_i = P_{Di}, q_i = P_{Fi} \). Since \( P(y_i = \xi_i | H_0) = \frac{P(y_i = \xi_i | H_1)}{P(y_i = \xi_i | H_0)} \) (see (17)),

\[
P(y_i = \xi_i | H_1) = \sum_{S_i} \prod_{S_i^{+}} p_i \prod_{S_i^{-}} q_i
\]
\[
P(y_i = \xi_i | H_0) = \sum_{S_i} \prod_{S_i^{+}} q_i \prod_{S_i^{-}} p_i
\]

From (10), we have

\[
\exp(y_i) = \frac{\exp(\sum_{S_i^{+}} w_j)}{\exp(\sum_{S_i^{-}} w_j)} = \frac{\prod_{S_i^{+}} \exp(w_j)}{\prod_{S_i^{-}} \exp(w_j)}.
\]

Applying (5) to (31) yields

\[
\exp(y_i) = \prod_{S_i^{+}} p_i \prod_{S_i^{-}} q_i
\]
\[
= \frac{\prod_{S_i^{+}} p_i \prod_{S_i^{-}} q_i}{\prod_{S_i^{+}} \prod_{S_i^{-}} q_i}
\]

The above equation holds for any combination of \( S_i^{+} \) and \( S_i^{-} \) such that

\[
y_i = \sum_{j \in S_i^{+}} W_j - \sum_{j \in S_i^{-}} W_j = \xi_i.
\]

Thus, using the following equality

\[
a + c = \frac{b + d}{b} = \frac{a + c}{b},
\]

Equation (30) becomes

\[
P(y_i = \xi_i | H_1) = \frac{\prod_{S_i^{+}} p_i \prod_{S_i^{-}} q_i}{\prod_{S_i^{+}} q_i \prod_{S_i^{-}} p_i}
\]
\[
= \frac{\prod_{S_i^{+}} p_i \prod_{S_i^{-}} q_i}{\prod_{S_i^{+}} q_i \prod_{S_i^{-}} p_i} = \exp(\xi_i)
\]

and

\[
P(y_i = \xi_i | H_0) = \exp(-\xi_i).
\]

Thus far, we have proved that for each \( y_i = \xi_i \), (36) holds. Using this equation and induction, we shall prove that \( B/A \) is monotonically decreasing with respect to \( \tau \).

As assumed earlier, \( \xi_1 < \xi_2 < \cdots < \xi_N \). From (36), we have

\[
P(y_i = \xi_i | H_0) > \frac{P(y_i = \xi_i | H_0)}{P(y_i = \xi_i | H_1)} > \cdots
\]
\[
> \frac{P(y_i = \xi_{N-1} | H_0)}{P(y_i = \xi_{N-1} | H_1)}
\]
\[
> \frac{P(y_i = \xi_N | H_0)}{P(y_i = \xi_N | H_1)}.
\]

Repeatedly applying the inequality,

\[
\frac{X}{Y} > \frac{a}{b} \Rightarrow \frac{X}{Y} > \frac{X + a}{Y + b} > \frac{a}{b}
\]
to (37), and using the definition of \( A \) and \( B \) in (23) and (24), it is clear that \( B/A \) is monotonically decreasing with respect to \( k \), and thus it is also monotonically decreasing with respect to \( \tau \). This is consistent with our intuitive reasoning. However, \( \tau \) cannot go to infinity; the maximum value of \( \tau \) is \( (y_i)_{\max} \). When \( \tau \) attains its maximum, \( B/A \) reaches its minimum value. According to the definition of \( A \) and \( B \), the minimum of \( B/A \) is

\[
\begin{align*}
\left( \frac{B}{A} \right)_{\min} &= \frac{P(y_i = (y_i)_{\max} | H_0)}{P(y_i = (y_i)_{\max} | H_1)} \\
&= \exp(-\xi_i) \\
&= \exp \left( -\sum_{j=1, j \neq i}^{n} \log \frac{P_{Di}}{P_{Fj}} \right) \\
&= \prod_{j=1, j \neq i}^{n} \frac{P_{Di}}{P_{Fj}}.
\end{align*}
\]

Thus, \( \left( \frac{B}{A} \right)_{\min} = \exp(-\xi_i) \)

\[
= \exp \left( -\sum_{j=1, j \neq i}^{n} \log \frac{P_{Di}}{P_{Fj}} \right) = \prod_{j=1, j \neq i}^{n} \frac{P_{Di}}{P_{Fj}}.
\]

According to (29), the minimum error that can be achieved at steady state for a fixed \( i \) is

\[
\varepsilon_i = \log \left( \frac{\prod_{j=1}^{n} \frac{P_{Di}}{P_{Fj}}}{1 + \prod_{j=1}^{n} \frac{P_{Di}}{P_{Fj}}} \right).
\]

Note that \( (y_i)_{\max} \) varies from sensor to sensor. In order to let every sensor adjust its weight and achieve the least error, the maximum value of \( \tau \) is chosen to be the minimum of all \( (y_i)_{\max} \):

\[
\tau_{\max} = \min \{ (y_1)_{\max}, (y_2)_{\max}, \ldots, (y_N)_{\max} \}.
\]

III. REINFORCEMENT LEARNING ALGORITHM

We assume that the distributed decision system has no knowledge of the probability mass functions of the observations. However, the probabilities of detection and false alarm for the \( i \)th detector \( \hat{P}_{Di} \) and \( \hat{P}_{Fi} \) can be approximated by relative frequencies. That is, in contrast to (18) and (22),

\[
\frac{\hat{P}_{Di}}{\hat{P}_{Fi}} \approx \frac{m_i}{n_i},
\]

where \( m_i \) and \( n_i \) are, respectively, the number of decisions made by the \( i \)th detector that agree and disagree with the reliable fused decisions. Both \( m_i \) and \( n_i \) are simply obtained by counting in the simulations. We next develop the updating rule for the fusion
As the number of iterations increase, $m_i$ approaches infinity. In this case,

$$ \frac{1}{m_i} P(u_i = \bar{u}_i^*) - \frac{1}{m_i} \exp(\hat{\theta}_i^*) P(u_i \neq \bar{u}_i^*) = 0 \quad (51) $$

and from the definition of $\hat{P}_{DI}$ and $\hat{P}_{FI}$,

$$ \hat{P}_{DI} - \hat{P}_{FI} \exp(\hat{\theta}_i^*) = 0. \quad (52) $$

Thus,

$$ \hat{\theta}_i^* = \hat{\theta}_i. \quad (53) $$

Hence, $\hat{\theta}_i^* \rightarrow \hat{\theta}_i$, for $i = 0, 1, \ldots$.

IV. SIMULATION RESULTS

In this section, we present some computer simulation results to demonstrate the validity of our proposed adaptive scheme. Fig. 3 shows the simulation set-up. Here, equally likely binary signals $\{-1, 1\}$ are randomly generated as source signals. Additionally, $N_1, N_2, \ldots, N_n$ are assumed to be independent identically distributed (IID) zero mean additive Gaussian random processes. Having selected the random noise process, the theoretical probabilities of detection and false alarm for each detector can be readily evaluated. For the Gaussian case, they can be determined by the standard deviation. They can be calculated according to:

$$ P_{FI} = Q \left( \frac{1}{\sigma_i} \right) = \int_{1/\sigma_i}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt \quad (54) $$

where $\sigma_i$ is the standard deviation of the Gaussian noise fed into the $i$th sensor,

$$ P_{DI} = 1 - P_{FI}. \quad (55) $$

Note that these theoretical probabilities and weights are calculated for comparison purposes only, and they are not readily available in practice. They are not used in the proposed adaptive fusion system. In the experiment, all the weights are first set to an initial value of 1, and then updated according to (49). The
steady state values are obtained after convergence (≈ 1000 iterations).

Figs. 4 and 5 and Table I show the results for two different cases. The first case assumes that each local detector is identical. Here, $P_{D_i} = 0.8413$, and $P_{F_i} = 0.1587$, for all $i = 1, 2, \ldots, 8$, where $w_i = \log P_{D_i}/P_{F_i} = 1.6679$. Fig. 4 shows the mean error among 8 sensors between the estimate $\hat{w}_i$ and the actual weight $w_i = 1.6679$ for different values of $\tau$, the reliability threshold. The figure conforms to our analytical results. That is, the larger the $\tau$, the smaller the error. On the other hand, larger training time is needed to reach the steady state for a larger $\tau$.

In the second case, the eight local detectors are assumed different, i.e., $P_{D_i} = 0.9234$ and $P_{F_i} = 0.0766$, for $i = 1, 2, 3, 5, 6, 7$; $P_{D4} = 0.8667$ and $P_{F4} = 0.1333$, and $P_{D8} = 0.9772$ and $P_{F8} = 0.0228$. Fig. 5 shows how the estimated weights approach the theoretical values.

In the figure, $w_1 = 2.4895, w_2 = 1.8721, w_8 = 3.7579$. Only three of the eight weights are shown. However, other weights also follow the same trend. Table I summaries the results for this experiment. It is readily seen that the simulation results conform closely to the theoretical results.

Though it has been shown that $\hat{w}_i^\tau$ converges to $\hat{w}_i$, it does not converge to $w_i$. The error, (42), depends on the number of sensors, the $P_{D_i}$ and $P_{F_i}$.

In the Gaussian noise environment, $P_{D_i}$ and $P_{F_i}$ are determined by the Signal-to-Noise ratio (SNR) of the $i$th sensor. Thus, the error $\epsilon_i$ is totally determined by the number and the SNRs of sensors. Fig. 6 shows, for the case of identical sensors, the error, $\epsilon_i$ versus $n$ (the number of sensors) for various SNRs. In this case, according to (54) and (55), the error can be simplified to

$$\epsilon_i = \log \frac{1 + \left(\frac{Q}{1-Q}\right)^n}{1 + \left(\frac{Q}{1-Q}\right)^{n-2}}$$

where $Q$ is the $Q$-function defined in (54) with the same standard deviation, $\sigma_i$ for all sensors. Note that the error is the same for every sensor.

When the SNR is different from sensor to sensor, the error can be written as

$$\epsilon_i = \log \frac{1 + T}{1 + \left(\frac{1-Q_i}{Q_i}\right)^2 T}$$

where

$$T = \prod_{j=1}^{n} \frac{Q_i}{1-Q_j}$$

and

$$Q_j = Q\left(\frac{1}{\sigma_j}\right).$$

From (57), it is readily seen that the larger the SNR, the smaller the $\epsilon_i$.

<table>
<thead>
<tr>
<th>Sensor</th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
<th>5th</th>
<th>6th</th>
<th>7th</th>
<th>8th</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noise Variance $\sigma^2$</td>
<td>0.49</td>
<td>0.49</td>
<td>0.49</td>
<td>0.81</td>
<td>0.49</td>
<td>0.49</td>
<td>0.49</td>
<td>0.25</td>
</tr>
<tr>
<td>SNR (dB)</td>
<td>3.098</td>
<td>3.098</td>
<td>3.098</td>
<td>3.098</td>
<td>3.098</td>
<td>3.098</td>
<td>3.098</td>
<td>6.02</td>
</tr>
<tr>
<td>Weights</td>
<td>Theoretical</td>
<td>2.4895</td>
<td>2.4895</td>
<td>2.4895</td>
<td>1.8721</td>
<td>2.4895</td>
<td>2.4895</td>
<td>2.4895</td>
</tr>
<tr>
<td>Steady state</td>
<td>2.4895</td>
<td>2.4916</td>
<td>2.4889</td>
<td>1.8721</td>
<td>2.4939</td>
<td>2.4971</td>
<td>2.4927</td>
<td>3.7941</td>
</tr>
</tbody>
</table>

ANSARI ET AL.: ADAPTIVE FUSION BY REINFORCEMENT LEARNING FOR DISTRIBUTED DETECTION SYSTEMS 529
V. CONCLUSIONS

In a real-world environment, the probability mass functions of the observations at local detectors may not be known and the performance of the local detectors may not be stationary. Under such circumstances, it is desirable to have a system which can adapt itself during the decision making process. This paper proposes such an adaptive system with the assumption that \( P_0 = P_1 \) and \( P_{D_1} = P_{F_1} \). The major advantage of the system is that a priori knowledge of the probability mass functions of the observations is not required. The system can acquire the knowledge about the reliability of the local detectors by itself—*it can learn by doing*. A reinforcement learning rule is proposed and adopted, and its convergence is analytically proven. The simulation results conform to our theoretical analysis.

If the **reliability threshold** \( \tau \) can be adjusted adaptively during the process of data fusing, the system may converge faster. Future efforts will focus on adaptively adjusting the **reliability threshold**, and developing a model for unequiprobable sources.

REFERENCES


Nirwan Ansari (M'88—SM'94) received the B.S.E.E. (summa cum laude) from New Jersey Institute of Technology (NJIT), Newark, NJ, in 1982, the M.S.E.E. from the University of Michigan, Ann Arbor, in 1983, and the Ph.D. degree from Purdue University, West Lafayette, IN, in 1988.

In 1988, he joined the Electrical and Computer Engineering Department of NJIT, where he is an Associate Professor and the Assistant Chair for Graduate Studies. His research interests include neural computing, pattern recognition, data fusion, nonlinear signal processing, ATM networks and adaptive CDMA detection.

Dr. Ansari serves as a frequent referee, a session chair, a session organizer, and a technical representative for various major journals, conferences, and federal agencies. He publishes regularly in his areas of research. He co-edited a book with B. Yuhua, *Neural Networks in Telecommunications*, published by Kluwer Academic Publishers, 1994.

Edwin S. H. Hou (S'83—M'89) received two B.S. degrees (magna cum laude) in electrical engineering and computer engineering from the University of Michigan, Ann Arbor, in 1982, the M.S. degree in computer science from Stanford University, Stanford, CA, in 1984, and the Ph.D. degree in electrical engineering from Purdue University, West Lafayette, IN, in 1989.

He is an Assistant Professor in the Electrical and Computer Engineering Department and the Assistant Director in the Electronic Imaging Center at New Jersey Institute of Technology, Newark. His research interests include infrared imaging, genetic algorithms, scheduling, and neural networks.

Bin Ou Zhu received the B.S. and M.S. degree from Shanghai Jiao Tong University in 1983 and 1986, respectively.

She came to New Jersey Institute of Technology, Newark, NJ, to pursue a doctoral degree in 1990, and is currently a technical staff member at OpenCom Systems, Inc. Her research interests include communication networking, data fusion and neural networks.

Jiang-Guo Chen received the B.S. degree in 1985, and the M.S. degree in 1988, both in electrical engineering, from Xidian University, Xi'an, Shaanxi, P.R. China. He is currently enrolled in the Ph.D. program in electrical engineering at the New Jersey Institute of Technology, Newark, NJ.

Between 1988 and 1993, he was a lecturer at the Computer Center, Taiyuan University of Technology, Taiyuan, Shaanxi, P.R. China. His research interests include data fusion, neural networks and image processing.