Adaptive Fusion of Correlated Local Decisions

Jian-Guo Chen and Nirwan Ansari

***Abstract*—In this paper, an adaptive fusion algorithm is proposed for an environment where the observations and local decisions are dependent from one sensor to another. An optimal decision rule, based on the maximum posterior probability (MAP) detection criterion for such an environment, is derived and compared to the adaptive approach. In the algorithm, the log-likelihood ratio function can be expressed as a linear combination of ratios of conditional probabilities and local decisions. The estimations of the conditional probabilities are adapted by reinforcement learning. The error probability at steady state is analyzed theoretically and, in some cases, found to be equal to the error probability obtained by the optimal fusion rule. The effect of the number of sensors and correlation coefficients on error probability in Gaussian noise is also investigated. Simulation results that conform to the theoretical analysis are presented at the end of the paper.**

***Index Terms*—Correlated decisions, distributed detection, probability of detection, probability of false alarm, reinforcement learning.**

I. Introduction

Sensor fusion, the study of optimal information processing in distributed multisensor environments through intelligent integration of the multisensor data, has gained popularity in recent years [1], [2]. Such a technique is expected to increase the reliability of detection, to be fairly immune to noise interference and sensor failures, and to decrease the bandwidth requirement. Among papers dealing with fusing decisions from a number of sensors [3]–[9], logical AND, OR, K out of N, majority, Bayes, and Neyman–Pearson (N–P)-based fusion rules have been introduced. When the observations and sensor decisions are independent, two different kinds of Bayes-based optimal fusion rules have been developed [7], [8]. Chair and Varshney [8] established an optimal fusion rule with the assumption that each local sensor made a predetermined decision and each observation was independent. Reibman and Nolte [7] found the global optimal solution by combining both the sensors and the fusion processor. The fusion of correlated decisions has also been studied [10]–[12]. Drakopoulos and Lee [12] have developed an optimum fusion rule for correlated decisions based on N–P criteria. The major drawback of all of the aforementioned optimal decisions is the requirement of the knowledge of *a priori* probabilities and the probabilities of a miss and detection of each local sensor that are not readily available in practice. The lacking of *a priori* probability information makes these algorithms impractical for actual applications. To circumvent this difficulty, we have developed an adaptive fusion model for the independent case for both equiprobable and unequiprobable sources [13], [14]. In this paper, we derive another form of the maximum posterior probability (MAP)-based optimal fusion rule and extend our adaptive algorithm by considering dependent/correlated decisions. The paper is arranged as follows. In Section II, we develop and derive the optimal fusion rule for correlated decisions. The adaptive fusion rule and the proof of its convergence are discussed in Section III. The residue between error probabilities obtained using the optimal fusion rule and the adaptive fusion rule is analyzed in Section IV. Section V illustrates the effect of the number of sensors and correlation coefficients on the error probability in a Gaussian noise environment. Simulations are presented in Section VI. Conclusions are drawn in Section VII.

Manuscript received March 17, 1996; revised October 3, 1997.

J.-G. Chen was with the Center for Communications and Signal Processing, Electrical and Computer Engineering Department, New Jersey Institute of Technology, Newark, NJ 07102 USA. He is now with Lucent Technologies, Murray Hill, NJ 07974 USA (e-mail: jianguo11@lucent.com).

N. Ansari is with the Electrical and Computer Engineering Department, New Jersey Institute of Technology, Newark, NJ 07102 USA (e-mail: ang@njit.edu).

Publisher Item Identifier S 1094-6977(98)02518-8.

II. Optimal Fusion Rule for Correlated Decisions

Consider the binary hypothesis testing problem with N sensors, in which each sensor employs a predetermined decision rule. The two hypotheses have *a priori* probabilities $P(H_1)$ and $P(H_0)$, respectively. In binary detection theory, one of the most popular detection criterion is the likelihood ratio criterion. The likelihood ratio is expressed as

$$\begin{aligned}\Lambda(u_1, u_2, \cdots, u_N) &= \frac{P(u_1, u_2, \cdots, u_N|H_1)}{P(u_1, u_2, \cdots, u_N|H_0)} \\ &= \frac{P_1(u_1)P_1(u_2|u_1)\cdots P_1(u_N|u_1, u_2, \cdots, u_{N-1})}{P_0(u_1)P_0(u_2|u_1)\cdots P_0(u_N|u_1, u_2, \cdots, u_{N-1})}\end{aligned} \tag{1}$$

where $P_i(u_k|u_1, u_2, \cdots, u_{k-1}) = P(u_k|u_1, u_2, \cdots, u_{k-1}, H_i)$, $i = 0, 1$ are conditional probabilities and $u_1, u_2, \cdots, u_N$ are local decisions that are binary random variables. H_0 and H_1 represent the following two hypotheses: H_0 (signal is absent) and H_1 (signal is present). Here, u_k for $k = 1, 2, \cdots, N$ is defined by

$$u_k = \begin{cases} -1, & \text{if } H_0 \text{ is declared} \\ +1, & \text{if } H_1 \text{ is declared.}\end{cases}$$

Since

$$P_i(u_k|u_1, u_2, \cdots, u_{k-1}) = \frac{1}{\frac{P_i(u_1, u_2, \cdots, u_{k-1}, u_k=+1)}{P_i(u_1, u_2, \cdots, u_k)} + \frac{P_i(u_1, u_2, \cdots, u_{k-1}, u_k=-1)}{P_i(u_1, u_2, \cdots, u_k)}}$$

we have

$$P_i(u_k|u_1, u_2, \cdots, u_{k-1}) = \begin{cases} \dfrac{1}{1 + \dfrac{P_i(u_1, u_2, \cdots, u_{k-1}, u_k = -1)}{P_i(u_1, u_2, \cdots, u_{k-1}, u_k = +1)}}, & \text{if } u_k = +1 \\ \dfrac{1}{1 + \dfrac{P_i(u_1, u_2, \cdots, u_{k-1}, u_k = +1)}{P_i(u_1, u_2, \cdots, u_{k-1}, u_k = -1)}}, & \text{if } u_k = -1.\end{cases}$$

Let

$$\begin{aligned} p_k &= \frac{P_1(u_1, u_2, \cdots, u_{k-1}, u_k = -1)}{P_1(u_1, u_2, \cdots, u_{k-1}, u_k = +1)} \\ q_k &= \frac{P_0(u_1, u_2, \cdots, u_{k-1}, u_k = -1)}{P_0(u_1, u_2, \cdots, u_{k-1}, u_k = +1)}. \end{aligned} \tag{2}$$

So

$$P_1(u_k|u_1, u_2, \cdots, u_{k-1}) = \begin{cases} \dfrac{1}{1 + p_k}, & \text{if } u_k = +1 \\ \dfrac{p_k}{1 + p_k}, & \text{if } u_k = -1\end{cases} \tag{3}$$

$$P_0(u_k|u_1, u_2, \cdots, u_{k-1}) = \begin{cases} \dfrac{1}{1 + q_k}, & \text{if } u_k = +1 \\ \dfrac{q_k}{1 + q_k}, & \text{if } u_k = -1.\end{cases} \tag{4}$$

By defining the weight W_k for $k = 0, 1, \cdots, N$ as

$$W_k = \begin{cases} \log \dfrac{P(H_1)}{P(H_0)}, & \text{for } k = 0 \\ \begin{cases} \log \dfrac{P_1(u_1)}{P_0(u_1)}, & \text{if } u_k = +1 \\ \log \dfrac{P_0(u_1)}{P_1(u_1)}, & \text{if } u_k = -1 \end{cases} & \text{for } k = 1 \\ \begin{cases} \log \dfrac{P_1(u_k|u_1, u_2, \cdots, u_{k-1})}{P_0(u_k|u_1, u_2, \cdots, u_{k-1})}, & \text{if } u_k = +1 \\ \log \dfrac{P_0(u_k|u_1, u_2, \cdots, u_{k-1})}{P_1(u_k|u_1, u_2, \cdots, u_{k-1})}, & \text{if } u_k = -1 \end{cases} & \text{for } k > 1 \end{cases} \tag{5}$$

we have

$$W_k = \begin{cases} \log \dfrac{1+q_k}{1+p_k}, & \text{if } u_k = +1 \text{ and } k > 1 \\ \log \dfrac{q_k(1+p_k)}{p_k(1+q_k)}, & \text{if } u_k = -1 \text{ and } k > 1. \end{cases} \tag{6}$$

The MAP or minimum error probability detection rule is

$$\Lambda(u_1, u_2, \cdots, u_N) = \frac{P_1(u_1, u_2, \cdots, u_N)}{P_0(u_1, u_2, \cdots, u_N)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P(H_0)}{P(H_1)}. \tag{7}$$

According to the above equation, the optimal fusion rule, based on the MAP detection criterion for correlated local decisions, is equivalent to

$$\sum_{i=0}^{N} W_i \cdot u_i \underset{H_0}{\overset{H_1}{\gtrless}} 0 \tag{8}$$

where u_0 is always set to one.

It can be seen that, when the local decisions are independent, (5) is the same as that in [8]. Thus, (5) is a generalization of Chair and Varshney's result for the correlated case.

III. Adaptive Fusion Rule

The optimal fusion rule derived in Section II requires the knowledge of *a priori* probabilities and conditional probabilities that are either difficult to acquire or time varying. To realize optimal fusion, an adaptive algorithm is necessary to estimate these probabilities.

A. Adaptive Fusion Rule

Similar to the independent case [13], [14], denote the events of the fusion results being $+1$ and -1 by $\hat{H}_1$ and $\hat{H}_0$, respectively. In addition, let m be the number that $\hat{H}_1$ occurs, n the number that $\hat{H}_0$ occurs, and

$m_{k,1}$ the number $(u_1, u_2, \cdots, u_{k-1}, u_k = +1, \hat{H}_1)$ occurs
$m_{k,0}$ the number $(u_1, u_2, \cdots, u_{k-1}, u_k = -1, \hat{H}_1)$ occurs
$n_{k,1}$ the number $(u_1, u_2, \cdots, u_{k-1}, u_k = +1, \hat{H}_0)$ occurs
$n_{k,0}$ the number $(u_1, u_2, \cdots, u_{k-1}, u_k = -1, \hat{H}_0)$ occurs.

Similar to (5), define

$$\hat{W}_k = \begin{cases} \log \dfrac{\widehat{P(H_1)}}{\widehat{P(H_0)}}, & \text{for } k = 0 \\ \begin{cases} \log \dfrac{\widehat{P_1(u_1)}}{\widehat{P_0(u_1)}}, & \text{if } u_k = +1 \\ \log \dfrac{\widehat{P_0(u_1)}}{\widehat{P_1(u_1)}}, & \text{if } u_k = -1 \end{cases} & \text{for } k = 1 \\ \begin{cases} \log \dfrac{\hat{P}_1(u_k|u_1, u_2, \cdots, u_{k-1})}{\hat{P}_0(u_k|u_1, u_2, \cdots, u_{k-1})}, & \text{if } u_k = +1 \\ \log \dfrac{\hat{P}_0(u_k|u_1, u_2, \cdots, u_{k-1})}{\hat{P}_1(u_k|u_1, u_2, \cdots, u_{k-1})}, & \text{if } u_k = -1 \end{cases} & k > 1 \end{cases} \tag{9}$$

$$\hat{p}_k = \frac{P(u_1, u_2, \cdots, u_{k-1}, u_k = -1, \hat{H}_1)}{P(u_1, u_2, \cdots, u_{k-1}, u_k = +1, \hat{H}_1)}$$

$$\hat{q}_k = \frac{P(u_1, u_2, \cdots, u_{k-1}, u_k = -1, \hat{H}_0)}{P(u_1, u_2, \cdots, u_{k-1}, u_k = +1, \hat{H}_0)}$$

where the symbols with a "hat" are estimations of symbols without a "hat." $\hat{p}_k$ and $\hat{q}_k$ can be approximated by

$$\hat{p}_k \approx \frac{m_{k,0}}{m_{k,1}}, \qquad \hat{q}_k \approx \frac{n_{k,0}}{n_{k,1}}. \tag{10}$$

When $u_k = +1$

$$\hat{W}_k = \log \frac{1+\hat{q}_k}{1+\hat{p}_k} \approx \log \frac{m_{k,1}}{n_{k,1}} \frac{n_{k,1}+n_{k,0}}{m_{k,1}+m_{k,0}}.$$

Note that $n_{k,1} + n_{k,0} = n_{k-1,j}$, $m_{k,1} + m_{k,0} = m_{k-1,j}$, where j is the output of the $(k-1)$th local sensor, that is

$$j = \begin{cases} 1, & \text{if } u_{k-1} = +1 \\ 0, & \text{if } u_{k-1} = -1. \end{cases} \tag{11}$$

Thus

$$\hat{W}_k \approx \log \frac{m_{k,1}}{n_{k,1}} - \log \frac{m_{k-1,j}}{n_{k-1,j}}. \tag{12}$$

Following the same reasoning, the approximated weight for $u_k = -1$ is

$$\hat{W}_k \approx \log \frac{m_{k-1,j}}{n_{k-1,j}} - \log \frac{m_{k,0}}{n_{k,0}}. \tag{13}$$

$\hat{W}_k$ exhibits the following property:

$$\hat{W}_k|_{u_k=+1} + \hat{W}_k|_{u_k=-1} = \log \frac{m_{k,1}}{n_{k,1}} - \log \frac{m_{k,0}}{n_{k,0}} = \log \frac{\hat{q}_k}{\hat{p}_k}. \tag{14}$$

The partial derivatives of $\hat{W}_k$, with respect to $m_{k,1}$, $m_{k,0}$, $n_{k,1}$, and $n_{k,0}$, are

$$\frac{\partial \hat{W}_k}{\partial m_{k,1}} \approx \frac{1}{m_{k,1}}, \quad \frac{\partial \hat{W}_k}{\partial n_{k,1}} \approx -\frac{1}{m_{k,1}} \frac{m_{k-1,j}}{n_{k-1,j}} e^{\hat{W}_k}, \quad \text{if } u_k = +1$$

$$\frac{\partial \hat{W}_k}{\partial m_{k,0}} \approx -\frac{1}{m_{k,0}}, \quad \frac{\partial \hat{W}_k}{\partial n_{k,0}} \approx \frac{1}{m_{k,0}} \frac{m_{k-1,j}}{n_{k-1,j}} e^{-\hat{W}_k}, \quad \text{if } u_k = -1.$$

According to the concept of reinforcement learning [15], if the current local detector's decision conforms to that of the fusion center, its weight $\hat{W}_k$ should be reinforced. In this case

$$\Delta \hat{W}_k \approx \begin{cases} \dfrac{1}{m_{k,1}}, & \text{if } u_k = +1 \text{ and } \hat{H}_1 \\ \dfrac{1}{m_{k,0}} \dfrac{m_{k-1,j}}{n_{k-1,j}} e^{-\hat{W}_k}, & \text{if } u_k = -1 \text{ and } \hat{H}_0. \end{cases} \tag{15}$$

TABLE I
ADAPTIVE FUSION ALGORITHM

	$\widehat{H}_1$		$\widehat{H}_0$	
	u_k=+1	u_k=-1	u_k=+1	u_k=-1
$\Delta\widehat{W}_0$	$\frac{1}{m}$		$-\frac{1}{m}e^{\widehat{W}_0^-}$	
$\Delta\widehat{W}_k$	$\frac{1}{m_{k,1}}$	$-\frac{1}{m_{k,0}}$	$-\frac{1}{m_{k,1}}\frac{m_{k-1,j}}{n_{k-1,j}}e^{\widehat{W}_k^-}$	$\frac{1}{m_{k,0}}\frac{m_{k-1,j}}{n_{k-1,j}}e^{-\widehat{W}_k^-}$

On the other hand, if the current local decision contradicts that of the fusion, its weight $\hat{W}_i$ should be reduced. That is

$$\Delta\hat{W}_k \approx \begin{cases} -\dfrac{1}{m_{k,0}}, & \text{if } u_k = -1 \text{ and } \hat{H}_1 \\ -\dfrac{1}{m_{k,1}}\dfrac{m_{k-1,j}}{n_{k-1,j}}e^{\hat{W}_k}, & \text{if } u_k = +1 \text{ and } \hat{H}_0. \end{cases} \tag{16}$$

Hence, the adaptive fusion rule is

$$\hat{W}_k^+ = \hat{W}_k^- + \Delta\hat{W}_k, \qquad k = 0, 1, 2, \cdots, N \tag{17}$$

where $\hat{W}_k^+$ and $\hat{W}_k^-$ represent the weight after and before each updating. The change of weights ΔW_k is summarized in Table I.

B. Proof of Convergence

Since $\hat{W}_0$ is the same as that in the independent case, its convergence can be proved similarly to our previous work [13]. Here we only consider the convergence of $\hat{W}_k$ for $k > 1$. From (17), it is easily seen that convergence of $\hat{W}_k$ is equivalent to $\Delta\hat{W}_k \to 0$. According to (15) and (16), the expected value of $\Delta\hat{W}_k$ can be expressed as

$$\begin{aligned} E[\Delta\hat{W}_k] = &+\frac{1}{m_{k1}}P(u_k = +1, \hat{H}_1) + \frac{1}{n_{k0}}P(u_k = -1, \hat{H}_0) \\ &- \frac{1}{m_{k0}}P(u_k = -1, \hat{H}_1) - \frac{1}{n_{k1}}P(u_k + = 1, \hat{H}_0). \end{aligned}$$

When the number of iterations increases, m_{k1}, n_{k1}, m_{k0}, and n_{k0} will approach infinity, while $P(u_k = +1, \hat{H}_1)$, $P(u_k = -1, \hat{H}_1)$, $P(u_k = +1, \hat{H}_0)$, and $P(u_k = -1, \hat{H}_0)$ are always between zero and one. Thus

$$\lim_{n+m\to\infty} E[\Delta\hat{W}_k] = 0$$

where the number of iterations equals $n + m$. Following the same reasoning, it can be shown that the variance and higher moments of $\Delta\hat{W}_k$ approach zero when the number of iterations goes to infinity. According to the theory of probability [16], it can be concluded that $\lim_{n+m\to\infty}\Delta\hat{W}_k = 0$ with probability one. Thus, $\hat{W}_k$ converges asymptotically to a real number with probability one. This completes our proof. □

IV. ERROR ANALYSIS

Though the adaptive fusion rule converges asymptotically, it is not guaranteed to converge to the optimal weights. To compare the performance between the optimal and adaptive algorithms, error probabilities obtained by these two methods are investigated. Based on the previous analysis, the ideal optimum decision is

$$y = \log\frac{P(U, H_1)}{P(U, H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} 0. \tag{18}$$

Using the proposed adaptive algorithm, the decision becomes

$$\hat{y} = \log\frac{P(U, \hat{H}_1)}{P(U, \hat{H}_0)} \underset{H_0}{\overset{H_1}{\gtrless}} 0 \tag{19}$$

where y and $\hat{y}$ are linear combinations of the local decisions. Let $U = (u_1, u_2, \cdots, u_N)$ be the vector representation of the local decisions.

Since

$$P(\hat{H}_1, U) = P(H_0, U)P(\hat{H}_1|U, H_0) + P(H_1, U)P(\hat{H}_1|U, H_1) \tag{20}$$

$$P(\hat{H}_0, U) = P(H_0, U)P(\hat{H}_0|U, H_0) + P(H_1, U)P(\hat{H}_0|U, H_1) \tag{21}$$

the adaptive fusion algorithm can be written as

$$\hat{y} = y + \log\frac{P(\hat{H}_1|U, H_1) + e^{-y}P(\hat{H}_1|U, H_0)}{P(\hat{H}_0|U, H_0) + e^{+y}P(\hat{H}_0|U, H_1)}. \tag{22}$$

Thus, the decision rule derived in (19) becomes

$$y \underset{H_0}{\overset{H_1}{\gtrless}} T(U) \tag{23}$$

where

$$T(U) = \log\frac{P(\hat{H}_0|U, H_0) + e^{+y}P(\hat{H}_0|U, H_1)}{P(\hat{H}_1|U, H_1) + e^{-y}P(\hat{H}_1|U, H_0)}. \tag{24}$$

In comparing (18) and (23), the decision rule using the adaptive algorithm is equivalent to the optimal decision rule offset by $T(U)$.

The error probability using the optimum decision rule is defined by

$$P_e = P(H_0|H_1)P(H_1) + P(H_1|H_0)P(H_0) \tag{25}$$

where

$$\begin{aligned} P(H_0|H_1) &= \sum_{y<0} P(U|H_1) \\ P(H_1|H_0) &= \sum_{y>0} P(U|H_0). \end{aligned} \tag{26}$$

The error probability using the adaptive decision rule is

$$P_e' = P'(H_0|H_1)P(H_1) + P'(H_1|H_0)P(H_0) \tag{27}$$

where

$$\begin{aligned} P'(H_0|H_1) &= \sum_{y<T(U)} P(U|H_1) \\ P'(H_1|H_0) &= \sum_{y>T(U)} P(U|H_0). \end{aligned} \tag{28}$$

Since the optimum detection rule achieves the minimum error probability, P_e' is usually larger than P_e. The degradation in performance can be measured by the absolute difference between the two error probabilities, that is

$$\begin{aligned} \Delta P_e &= |P_e' - P_e| \\ &= \left| P(H_1)\left(\sum_{y<T(U)} P(U|H_1) - \sum_{y<0} P(U|H_1)\right) \right. \\ &\quad \left. + P(H_0)\left[\sum_{y>T(U)} P(U|H_0) - \sum_{y>0} P(U|H_0)\right]\right|. \end{aligned} \tag{29}$$

Comparing the decision rules using optimal and adaptive fusion, the following is desirable:

$$\begin{aligned} &\text{when } y > 0,\ T(U) - y < 0, \text{ and} \\ &\text{when } y < 0,\ T(U) - y > 0 \end{aligned}$$

such that $\Delta P_e = 0$. From (24), the offset increases with y monotonically. If $P(\hat{H}_0|U, H_0) > 0.5$ and $P(\hat{H}_1|U, H_1) > 0.5$ (which is the usual case), $T(U) - y$ decreases with y monotonically (see Fig. 1).

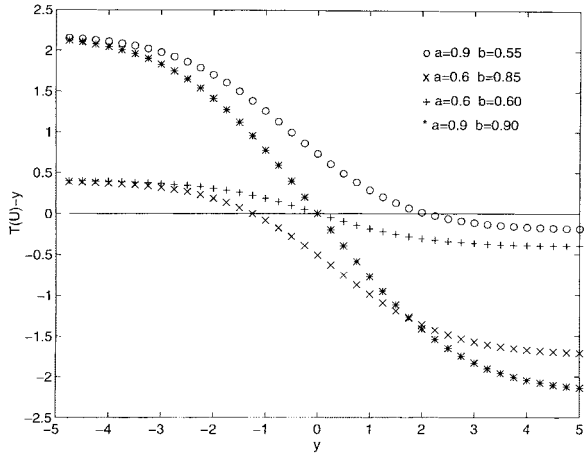

Fig. 1. $T(U) - y$ versus y, where $T(U)$ is the offset and y is the linear combination of local decisions.

Proposition: Let $a = P(\hat{H}_0|U, H_0)$, $b = P(\hat{H}_1|U, H_1)$, $1 - a = P(\hat{H}_1|U, H_0)$, $1 - b = P(\hat{H}_0|U, H_1)$, and y_T denote the zero-crossing point of $T(U) - y$ [i.e., $y_T = \{y: T(U) - y = 0\}$]. Since $T(U) - y$ decreases with y monotonically, y_T is unique

$$y_T = \log\left(\frac{a - 0.5}{b - 0.5}\right). \tag{30}$$

The closer the y_T is to zero, the smaller the degradation in error probability. The adaptive rule behaves as well as the optimal decision rule when $a = b$, in which case $y_T = 0$. Fig. 1 shows plots of $T(U) - y$ versus y for different values of a and b. By using the following relationship between y and $\Lambda(U)$:

$$y = \log\left(\Lambda(U)\,\frac{P(H_1)}{P(H_0)}\right) \tag{31}$$

many of the properties discussed in this section can also be expressed in term of the likelihood ratio function $\Lambda(U)$.

V. Performance in Gaussian Noise

To gain an insight into the proposed adaptive fusion algorithm, performance analysis in Gaussian noise that is both theoretically tractable and computationally feasible is examined in this section. Suppose all of the sensors are corrupted by Gaussian noise that has a zero mean and the same variance of σ^2. Let the correlation coefficient ρ between different sensors be the same. Thus, the observation vector X at the local sensors is Gaussian distributed. Let $\underline{U_1} = [1\ 1\ 1\ \cdots\ 1\ 1]$ and $\underline{U_0} = [-1\ -1\ -1\ \cdots\ -1\ -1]$ be its mean vectors for H_1 and H_0, respectively. The correlation matrix of X is

$$C = \sigma^2 \begin{bmatrix} 1 & \rho & \rho & \cdots & \rho \\ \rho & 1 & \rho & \cdots & \rho \\ \cdots & \cdots & \cdots & \cdots & \\ \rho & \rho & \cdots & \cdots & 1 \end{bmatrix}.$$

In addition, suppose all of the local sensors adopt the same decision threshold t, implying that the optimal fusion rule is the same as the k out of N rule. Furthermore, assume that $P(\hat{H}_0|U, H_0) = P(\hat{H}_1|U, H_1)$. Let $A_{N-k,\,k}(t, \rho)$ denote the joint probability of N random variables with the correlation coefficient ρ, of which k out of N random variables are greater than t and the other $N - k$ are less than t. If all of the random variables are identical, it can be shown that $A_{N-k,\,k}(t, \rho) = \binom{N}{k} P(x_1 < t, \cdots, x_{N-k} < t, x_{N-k+1} > t, \cdots, x_N > t)$ and

$$A_{N-k,\,k}(t, \rho) = \sum_{j=0}^{k} (-1)^j \binom{k}{j} A_{N-k+j,\,0}(t, \rho). \tag{32}$$

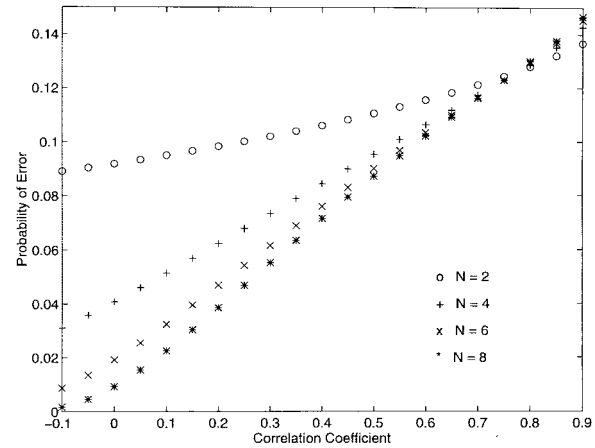

Fig. 2. P_e versus ρ, for $N = 2, 4, 6, 8$.

When x_i's are Gaussian random variables with a zero mean and correlation matrix C, $P(x_1 < t, \cdots, x_{N-k+j} < t)$ can be expressed as in [17]

$$P(x_1 < t, x_2 < t, \cdots, x_{N-k+j} < t) = \int_{-\infty}^{\infty} Q^{N-k+j}\left(\frac{t - \sqrt{\rho}y}{\sqrt{1-\rho}}\right) f(y)\, dy \tag{33}$$

where $f(\cdot)$ and $Q(\cdot)$ are the standard normal density and cumulative distribution functions. Equation (33) can be computed numerically. It follows from the above discussion that the two likelihood functions can be expressed as

$$\begin{aligned} P(U|H_1) &= A_{N-k,\,k}(-1, \rho) \\ P(U|H_0) &= A_{N-k,\,k}(+1, \rho). \end{aligned} \tag{34}$$

To study the effect of N and ρ on the error probability, consider the case where $P(H_0) = P(H_1) = 0.5$ and $\sigma^2 = 1$ (i.e., SNR = 0 dB). Thus, the fusion rule is simplified to

$$\Lambda(U) = \frac{P(U|H_1)}{P(U|H_0)} = \frac{A_{N-k,\,k}(-1, \rho)}{A_{N-k,\,k}(+1, \rho)} \underset{H_0}{\overset{H_1}{\gtrless}} 1 \tag{35}$$

where $A_{N-k,\,k}(-1, \rho)$ and $A_{N-k,\,k}(+1, \rho)$ can be computed numerically using (32) and (33). From this decision equation, there exists a K for given N and ρ, such that

$$\text{when } k \geq K, \qquad \frac{A_{N-k,\,k}(-1, \rho)}{A_{N-k,\,k}(+1, \rho)} \geq 1 \tag{36}$$

$$\text{when } k < K, \qquad \frac{A_{N-k,\,k}(-1, \rho)}{A_{N-k,\,k}(+1, \rho)} < 1. \tag{37}$$

In this case, the error probability can be expressed as

$$P_e = 0.5\left[\sum_{k=0}^{K-1} A_{N-k,\,k}(-1, \rho) + \sum_{k=K}^{N} A_{N-k,\,k}(+1, \rho)\right]. \tag{38}$$

For this special case, the error probability using the optimal decision and adaptive decision rules can be determined. Figs. 2 and 3 show the error probability versus ρ based on (38) when N is even and odd, respectively. When N is even, there exists a k, such that $\Lambda(U) = 1$, but when N is odd, no such k exists. $\Lambda(U) = 1$ corresponds to an undetermined case that can be considered as either H_0 or H_1. The contribution to the error probability for the undetermined case is considered as half of the probability it occurs.

It can be seen that better performance is achieved with a smaller correlation coefficient between sensors. This agrees with the conclusion of other fusion rules [10], [12]. Also, better performance can be achieved by increasing the number of sensors, but this advantage diminishes as the correlation coefficient ρ increases.

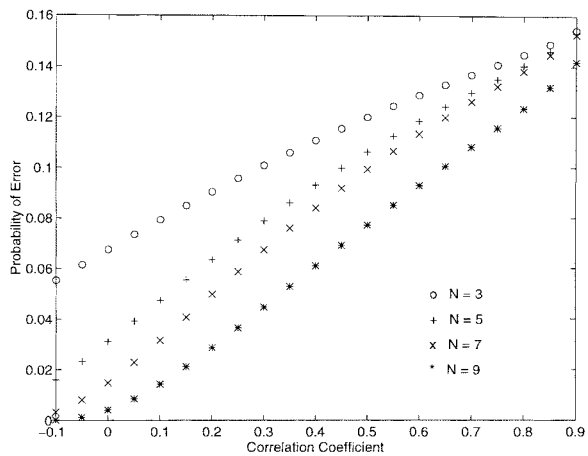

Fig. 3. P_e versus ρ, for $N = 3, 5, 7, 9$.

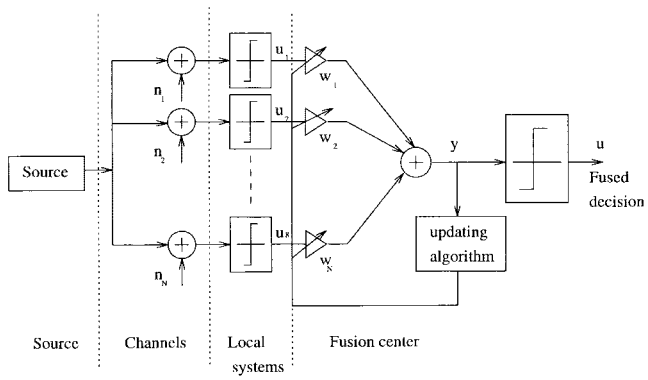

Fig. 4. Setup of computer simulations.

VI. Simulations

Fig. 4 shows the setup for our simulations. The source emits a sequence of $+1$ and -1. The probability of emitting $+1$ is $P(H_1)$ and that of emitting -1 is $P(H_0)$. The additive noise is zero-mean Gaussian with a variance of one. Each local sensor makes its decision u_i and transmits it to the fusion center. The fusion center computes the linear combination of local decisions to produce $\mathbf{y}$ and then compares it with a threshold (here, zero is used). If $\mathbf{y}$ is greater than zero, the final decision is $+1$, otherwise -1.

A. Generation of Correlated Gaussian Noise

In our simulations, we need to generate Gaussian noise with the specified correlation coefficient. The usual random number generator can only produce independent and identically distributed noise. Correlated noise can be obtained through some linear transformations. Let Z denote an N-dimensional correlated noise vector whose correlation matrix is C_z. Y is another N-dimensional noise vector defined by

$$Y = A^T Z. \tag{39}$$

If A is a square matrix whose column vectors are the eigenvectors of C_z, the correlation matrix C_y of Y becomes a diagonal matrix whose diagonal elements are the eigenvalue of C_z [18]. Denote C_y as

$$C_y = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \\ 0 & 0 & \cdots & \cdots & \lambda_N \end{bmatrix}$$

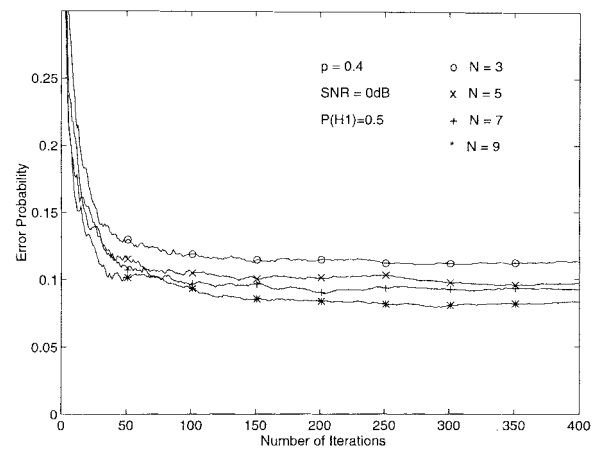

Fig. 5. Effect of the number of sensors on error probability.

where λ_i is an eigenvalue of C_z. Since λ_i are distinct even when each element in Z has the same variance and the same correlation coefficient, by introducing the next transformation

$$X = B^T Y \tag{40}$$

where

$$B = \begin{bmatrix} \frac{1}{\sqrt{\lambda_1}} & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{\sqrt{\lambda_2}} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \\ 0 & 0 & \cdots & \cdots & \frac{1}{\sqrt{\lambda_N}} \end{bmatrix}$$

the correlation matrix C_x of X becomes an identity matrix. Combining the above two transformations, we have

$$Z = (BA^T)^{-1} X. \tag{41}$$

According to the above definitions, X is an independent, zero-mean, unit-variance, Gaussian random vector that can be easily generated. Z becomes a zero-mean Gaussian random vector with correlation matrix C_z. B and A are determined by the eigenvalues and eigenvectors of C_z. In our experiment, C_z is specified to have the same correlation coefficient and variance.

B. Initialization

In the proposed adaptive algorithm, as specified by (12) and (13), the optimal weights are estimated by relative frequencies. Hence, the initialization of these relative frequencies is very crucial. If they are initialized randomly, the algorithm may converge to wrong weights that produce large error probability. To ensure that the algorithm converges appropriately, we employ the majority rule, which does not require *a priori* knowledge about the sensors and source to initialize the events described in Section VI-A. For every different local decision combination, the majority fusion rule is employed one time to form the fusion result. The relative frequencies are changed based on this fusion result. In this case, it will converge to the right direction.

C. Simulation Results

Consider the same situation as described in Section V. Theoretical analysis has shown that the number of sensors and correlation coefficients greatly affect the performance of fusion (see Figs. 2 and 3). These effects are also observed in the simulation results. Fig. 5 shows the plots of error probability versus the iteration for a different N with a fixed correlation coefficient. Fig. 6 shows the plot for a different correlation coefficient with a fixed N. Tables II and

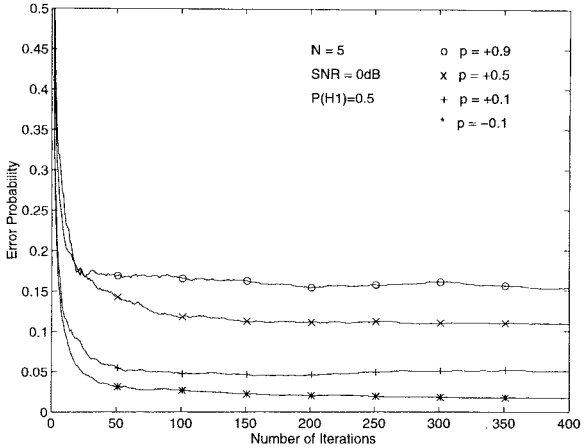

Fig. 6. Effect of the correlation coefficient on error probability.

TABLE II
THEORETICAL AND ESTIMATED VALUES OF ERROR PROBABILITY ($\rho = 0.4$)

N	3	5	7	9
Experimental Value	0.1130	0.0971	0.0923	0.0832
Theoretical Value	0.1107	0.0930	0.0840	0.0610

TABLE III
THEORETICAL AND ESTIMATED VALUES OF ERROR PROBABILITY ($N = 5$)

ρ	-0.1	0.1	0.5	0.9
Experimental Value	0.0173	0.0500	0.1092	0.1457
Theoretical Value	0.0160	0.0473	0.1062	0.1533

TABLE IV
ESTIMATED ERROR PROBABILITIES FOR DIFFERENT ALGORITHMS AND DIFFERENT SOURCES

	Independent Algorithm	Correlated Algorithm
Independent Source	0.0584	0.0358
Correlated Source	0.1100	0.1029

III summarize the corresponding theoretical and simulated values at the 400th iteration of the error probability. From these two figures and tables, it can be seen that the proposed algorithm converges and the steady-state error probabilities obtained from the simulations are very close to the theoretical values.

It is interesting to note, as illustrated by Fig. 7 and summarized in Table IV, that the adaptive algorithm developed for the correlated case always outperforms the one we previously developed for the independent case [13], [14], regardless of whether the local decisions are actually correlated. This may be attributed to the fact that the algorithm that considers correlated decisions includes more information in making its decision.

In this section, all error probabilities are calculated according to (25). The $P(H_0|H_1)$ and $P(H_1|H_0)$ are estimated by relative frequencies. All curves are results averaged over 20 times.

VII. CONCLUSION

In this paper, we have proposed an adaptive algorithm to solve the MAP-based optimal fusion problem when sensors are dependent from one another. The following main attributes of the algorithm can be concluded from the theoretical analysis and simulations.

1) It does not require *a priori* knowledge about the sensors and source and, thus, is more practical.
2) It adapts the weights from time to time and, thus, is more suitable for a time-varying environment.

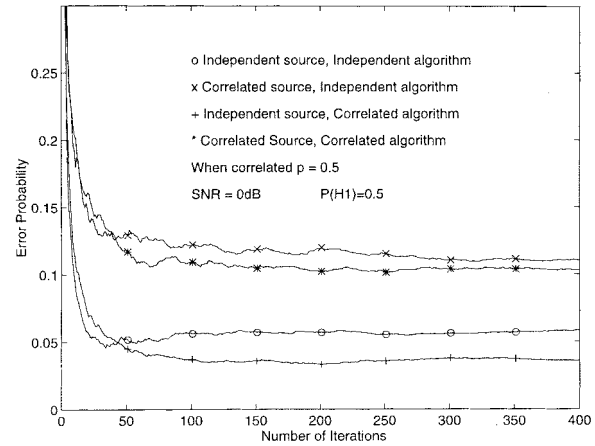

Fig. 7. Performance comparison between the correlated and independent algorithms.

3) Its computational complexity is low and, thus, is more implementable.

ACKNOWLEDGMENT

The authors would like to thank Z. Siveski for providing the code to calculate (33).

REFERENCES

[1] B. V. Dasarathy, *Decision Fusion.* Los Alamitos, CA: IEEE Comput. Soc., 1994.
[2] E. Waltz and J. Llinas, *Multisensor Data Fusion.* Norwood, MA: Artech House, 1990.
[3] R. R. Tenney and N. R. Sandell, Jr., "Detection with distributed sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-17, pp. 501–510, July 1981.
[4] F. A. Sadjadi, "Hypothesis testing in a distributed environment," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-22, pp. 134–137, Mar. 1986.
[5] S. C. A. Thomopoulos, R. Viswanathan, and D. C. Bougoolias, "Optimal decision fusion in multiple sensor systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-23, pp. 644–653, Sept. 1987.
[6] A. R. Reibman and L. W. Nolte, "Design and performance comparison of distributed detection networks," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-23, pp. 789–797, Nov. 1987.
[7] ——, "Optimal detection and performance of distributed sensor systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-23, pp. 24–30, Jan. 1987.
[8] Z. Chair and P. K. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-22, pp. 98–101, Jan. 1986.
[9] R. Srinivasan, "Distributed radar detection theory," *Proc. Inst. Elec. Eng.*, vol. 133, pt. F, pp. 55–60, Feb. 1986.
[10] V. Aalo and R. Viswanathan, "On distributed detection with correlated sensors: Two examples," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-25, pp. 414–421, May 1989.
[11] M. Kam, Q. Zhu, and W. Gray, "Optimal data fusion of correlated local decisions in multiple sensor detection systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-28, pp. 916–920, July 1992.
[12] E. Drakopoulos and C. Lee, "Optimum multi-sensor fusion of correlated local decision," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-27, pp. 593–605, July 1991.
[13] N. Ansari, E. Hou, B. Zhu, and J. Chen, "An adaptive fusion model for distributed detection systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, pp. 524–531, Apr. 1996.
[14] N. Ansari, J. Chen, and Y. Zhang, "Adaptive decision fusion for unequiprobable sources," *IEE Proc. Radar, Sonar Navig.*, vol. 144, pp. 105–111, June 1997.
[15] M. H. Hassoun, *Fundamentals of Artificial Neural Networks.* Cambridge, MA: MIT Press, 1995.
[16] W. Feller, *An Introduction to Probability Theory and its Applications*, 3rd ed. New York: Wiley, 1968, vol. 1.
[17] S. S. Gupta, "Probability integrals of multivariable normal and multivariate t," *Ann. Math. Stat.*, vol. 34, pp. 792–828, Sept. 1963.
[18] S. Haykin, *Adaptive Filter Theory*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1991.