SANDIA REPORT

SAND98–1699
Unlimited Release
Printed August 1998

Probabilistic Fusion of ATR Results

Katherine M. Simonson

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000.

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Available to DOE and DOE contractors from
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Prices available from (615) 576-8401, FTS 626-8401

Available to the public from
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Rd
Springfield, VA 22161

NTIS price codes
Printed copy: A03
Microfiche copy: A01

SAND98-1699
Unlimited Release
Printed August 1998

Probabilistic Fusion of ATR Results

Katherine M. Simonson
Signal and Image Processing Systems Department
Sandia National Laboratories
P.O. Box 5800
Albuquerque, NM 87185-0844

Abstract

The problem of combining multi-source information in applications related to automatic target recognition (ATR) is addressed. A mathematical approach is proposed for fusing the (possibly dependent) outputs of multiple ATR systems or algorithms. The method is derived from statistical principles, and the fused decision takes the form of an hypothesis test. The distribution of the test statistic is approximated as gamma, with parameters estimated from available training data. In a brief simulation study, the proposed method outperforms several alternative fusion techniques.

This work performed under sponsorship of the Air Force
Aeronautical Systems Center and Rome Laboratory.

Contents

List of Tables

List of Figures

1 – INTRODUCTION AND BACKGROUND

In recent years, researchers in the Automatic Target Recognition (ATR) community have been interested in developing techniques for pulling together information from different sources to improve the performance of ATR systems in locating and identifying targets of interest. The process by which information is combined is frequently referred to as "data fusion". Fusion can take place at various stages of the processing sequence, and the data to be fused may lie anywhere in the range from raw to highly processed.

Different approaches to fusion are appropriate at different stages of the processing sequence. If raw imagery from multiple sensors is to be combined before input to an ATR system, accurate modeling of some complex spatial and temporal relationships between and within images may be required. By contrast, when the data to be merged are highly processed (e.g., ATR outputs), the modeling task is often relatively simple, and can yield fusion metrics that are powerful and straightforward to interpret.

The term "ATR fusion" will be used to refer to the merging of different ATR outputs. ATR fusion techniques can be applied to results arising from any of the following: 1) multiple sensors, 2) multiple images taken with a single sensor, and 3) multiple ATR algorithms applied to a single image. The most elementary approach to ATR fusion uses voting metrics, in which each ATR output is converted to a binary "yes/no" format and final classifications are based on decision rules like "majority wins". In many applications, more sophisticated and effective metrics can be developed.

In this report, a new technique for ATR fusion is introduced. The method is derived from statistical principles in section 2.1. The decision rule takes the form of an hypothesis test, and an approximation to the distribution of the test statistic is discussed in sections 2.2 and 2.3. The power of the method is illustrated in section 3, in which the performances of the proposed metric and other common fusion techniques are compared over a series of simulated data sets. Practical implementation issues and possible extensions of the work are outlined in Section 4.

2 – THE MATHEMATICS OF PROBABILISTIC FUSION

2.1 - Deriving the Test Statistic

The probabilistic fusion metric is derived in the context of a target identification problem: the ATR system is asked to determine whether or not a specific target (vehicle type and orientation) is present at a particular location. This system may apply multiple identification algorithms to the same image, or it may apply a single algorithm to multiple images. In either case, several different ATR outputs are generated, which must be combined in the final decision. Each output is assumed to be one-dimensional, and typically represents a match score like a sum of mean-squared errors, a correlation coefficient, or some more sophisticated measure of the similarity between the input image and the target model.

Let N represent the number of available scores. Denote by Z_i the random variable representing the i^{th} score, and let z_i be a value assumed by Z_i. The N random variables $\{Z_1, \ldots, Z_N\}$ need not be mutually independent. Without loss of generality, assume for each i that a small value of Z_i represents a good match with the target model, while a large value represents a poor match. Because the N scores may differ in terms of magnitude, range and interpretation, the first step in the fusion process is to transform them into a new domain in which they may be combined in a straightforward manner.

Consider the i^{th} match score, and let F_i be the probability distribution function of the random variable Z_i for targets, such that:

$$F_i(z_i) = Prob\ (Z_i \leq z_i) \tag{1}$$

That is, $F_i(z_i)$ represents the probability that the i^{th} ATR score would take on a value of z_i or smaller for a target vehicle.

The function F_i may be known parametrically from theoretical considerations (e.g., the Central Limit Theorem), or it may be estimated empirically from available training data. It is assumed that F_i is continuous (or well approximated by a continuous function) for all i. It follows from the definition of the probability distribution function that the random variable $F_i(Z_i)$ has a distribution that is uniform on the interval [0, 1] [Silvey, 1975] for target vehicles. Because a large Z_i represents a poor match with the target model, it is expected that non-targets will take on values of $F_i(Z_i)$ that are close to unity.

Now define the mapping:

$$\begin{aligned} Y_i &= Y_i(Z_i) \\ &= -\log(1 - F_i(Z_i)) \end{aligned} \tag{2}$$

Because the function F_i is bounded between 0 and 1, the random variable Y_i is non-negative. Its value approaches infinity as the value of $F_i(Z_i)$ approaches unity. For targets, the quantity $1 - F_i(Z_i)$ has a uniform distribution on [0, 1]. It follows that the distribution of Y_i for targets is the standard exponential [Johnson and Kotz, 1970]. It is expected that Y_i will be large for non-targets.

The classification decision takes the form of a one-sided statistical test, in which the null hypothesis is that each of the $\{Y_1, \ldots, Y_N\}$ has the standard (unit mean) exponential distribution, with a specified multivariate correlation structure. The alternative hypothesis is that at least one of the Y_i's has a mean exceeding unity. Phrased another way, the null hypothesis is that the unknown is the target, and the alternative is that it is not.

The proposed test statistic combines the N transformed scores in a straightforward manner. Define the fused measure:

$$S_f = \sum_{i=1}^{N} Y_i \,. \tag{3}$$

Because each Y_i is expected to be large for non-targets, it is sensible to make target identification decisions by setting a threshold on the value of S_f. Images with a fused score at or below the chosen threshold are classified as targets, while those with scores above the threshold are classified as non-targets. The use of (3) dates to the work of R. A. Fisher [Fisher, 1932; van Zwet and Oosterhoff, 1967] in combining evidence from multiple independent statistical tests. For the present application, the requirement of independence is relaxed. In order to choose an appropriate threshold for the test statistic S_f, an approximation to its distribution is required. In the next subsection, one is proposed.

2.2 - Approximating the Distribution of S_f

If the Y_i's are mutually independent, then the distribution of the sum S_f is gamma, with shape parameter $r = N$ and scale parameter $\lambda = 1$ [Johnson and Kotz, 1970]. More generally, let $\hat{\rho}_{ij}$ represent the estimated correlation between Y_i and Y_j, and define the scalar:

$$C = \sum_{i=1}^{N} \sum_{j \neq i} \hat{\rho}_{ij} \tag{4}$$

The distribution of S_f is approximated as a gamma, with parameters:

$$\hat{r} = \frac{N^2}{N + C} \tag{5a}$$

$$\hat{\lambda} = \frac{N}{N + C} \tag{5b}$$

estimated by the method of moments. A decision on the consistency of the test data with the designated target vehicle can be made based on the gamma distribution function evaluated at S_f. Let $G(s; \hat{r}, \hat{\lambda})$ represent the probability that a gamma random variable with parameters $\hat{r}$ and $\hat{\lambda}$ will be less than or equal to the value s. Then a threshold, S_f^*, is selected to correspond to the desired probability of detection (PD), such that:

$$G\left(S_f^* ; \hat{r}, \hat{\lambda}\right) = PD \ . \tag{6}$$

If the fused measure (3) takes on a value less than or equal to S_f^*, the null hypothesis (presence of the target) is accepted. Otherwise, the test image is classified as non-target.

Note that the gamma model for the distribution of S_f is exact only when the Y_i's are mutually independent. The degree to which this approximation is reasonable for other situations must be considered. Because there is no unique multivariate uniform distribution [Barnett, 1980], a general answer is not possible. However, two specific cases that have been studied are discussed in the next section.

2.3 - Assessing the Gamma Approximation

The applicability of the gamma approximation to the distribution of the fused measure S_f (3) is now examined. Because the true distribution depends on the underlying models from which the exponential variates Y_i (Equation 2) are derived, the goodness of the approximation varies with the models. In the case where the Y_i's are mutually independent, the gamma approximation is exact [Johnson and Kotz, 1970]. In the case of dependent exponentials, two specific models have been studied, with excellent agreement found between the exact distributions and the gamma approximations in each case.

2.3.1 - The Morganstern Model

One case in which the exact distribution of a sum of dependent exponentials can be solved analytically is the Morganstern model. Although this model is limited to the bivariate case, it can still be useful in testing the accuracy of the gamma approximation. In the Morganstern-type bivariate uniform distribution with parameter ψ [Barnett, 1980], the random variables U_1 and U_2 have a joint density given by:

$$f(u_1, u_2; \psi) = 1 - \psi(1-2u_1)(1-2u_2), \tag{7}$$

for $0 \le u_1, u_2 \le 1$ and $|\psi| \le 1$. Here, both U_1 and U_2 have marginal distributions that are uniform on [0, 1]. Making the change of variables:

$$T_1 = -\log(U_1) \qquad T_2 = -\log(U_2), \tag{8}$$

each T_i has a marginal distribution that is standard exponential, and the joint density is given by:

$$g(t_1, t_2; \psi) = e^{-t_1-t_2} - \psi e^{-t_1-t_2} + 2\psi e^{-2t_1-t_2} + 2\psi e^{-t_1-2t_2} - 4\psi e^{-2t_1-2t_2}. \tag{9}$$

Now, let $W = T_1 + T_2$. The exact density of this sum is given by:

$$h(w; \psi) = we^{-w} - \psi we^{-w} + 4\psi e^{-w} - 4\psi e^{-2w} - 4\psi we^{-2w}, \tag{10}$$

with corresponding distribution function:

$$H(w; \psi) = 1 - e^{-w}(1 + 3\psi - \psi w + w) + e^{-2w}(3\psi + 2\psi w). \tag{11}$$

To determine the goodness of the approximation in the Morganstern case, the gamma distribution function is compared to the exact distribution function (11). From (9), it can be shown that:

$$corr(T_1, T_2) = -\frac{\psi}{4}. \tag{12}$$

It follows from equations (5) that the method of moments estimates for the gamma parameters are given by:

$$\hat{r}_\psi = \frac{8}{4-\psi} \tag{13a}$$

$$\hat{\lambda}_\psi = \frac{4}{4-\psi} . \tag{13b}$$

The distributions $G(w; \hat{r}_\psi, \hat{\lambda}_\psi)$ and $H(w; \psi)$ are compared in Figure 1 and Table 1, for several different values of the parameter ψ. Agreement is excellent: the differences between the exact and approximate detection probabilities for various thresholds are small enough to be of little practical significance.

2.3.2 - The Multivariate Normal Model

A second type of multivariate exponential is obtained by sampling from the M-dimensional normal distribution [Anderson, 1984]. The bivariate case is discussed in Barnett (1980), but a more general multivariate extension is possible. Under the normal model, the exact distribution function of the sum of standard exponentials cannot be derived analytically, due to the form of the normal density function. However, empirical estimates obtained via simulation can be compared to the gamma approximations.

Let the vector $\{V_1, \ldots, V_M\}$ represent a single realization from the multivariate normal, with mean vector $\underline{0}$ and covariance matrix:

$$\Sigma = [\sigma_{ij}], \qquad i, j = 1, \ldots, M , \tag{14}$$

where σ_{ii} is equal to unity for all i. Each of the V_i has a marginal distribution that is standard normal. Make the change of variables:

$$T_i = -\log(1 - \Phi(V_i)), \tag{15}$$

where Φ represents the standard normal distribution function. Then each T_i is distributed as standard exponential. The T_i's are correlated unless the matrix Σ is equal to the identity. However, the correlations between the exponentials cannot be solved analytically and must be computed numerically or estimated via simulation. Likewise, the distribution of the sum:

$$W = \sum_{i=1}^{M} T_i \tag{16}$$

is obtained via simulation or numerical methods.

To test the adequacy of the gamma approximation for the distribution of W in the multivariate normal model, the following approach was taken.

1. Begin with an initial "training" sample, used only to estimate the gamma parameters: sample 200 realizations from an M-dimensional normal distribution, with mean vector $\underline{0}$, and covariance matrix Σ. Convert each element to the exponential form (15). Now compute the sample covariance matrix of the exponential random variables, and insert the off-diagonal elements into equation (4). The method of moments estimates for the gamma parameters are found as in (5).

2. Discard the initial sample, and generate a new collection of 10,000 M-dimensional normal random variables, with mean vector $\underline{0}$ and covariance matrix Σ. Convert each element to the exponential form (15), and sum the M dimensions for each realization. This provides a sample of 10,000 sums of dependent exponentials. The empirical distribution of these sums is compared to the approximate distribution computed in the previous step.

Results for several different dimensions and covariance matrices are shown in Figure 2 and Table 2. In each case, the gamma approximation closely matches the empirical distribution obtained via simulation.

The results obtained for multivariate exponentials arising from the Morganstern and normal models do not establish the accuracy of the gamma approximation to the distribution of S_f for all possible models. However, they do suggest that the two-parameter gamma family represents a reasonable and flexible choice in the general case. Detection probabilities obtained with decision thresholds based on the gamma approximation are unlikely to differ greatly from those that would be obtained with exact thresholds.

3 – SOME COMPARATIVE RESULTS

3.1 - Alternative Approaches to ATR Fusion

There are many feasible alternatives to the fusion metric proposed in Section 2. A brief discussion of some of the more common approaches follows. Note that each of the metrics discussed below operates on the raw Z scores, and thus implicitly assumes that the scores are measured on the same scale, or have been normalized. For many fusion applications (e.g., the same ATR algorithm applied to multiple images) this is a reasonable assumption. As before, a small value of Z represents a close match with the target, while a large value represents a poor match.

The first candidate metric takes the smallest of the N scores recorded for a particular location, and uses this as a fused measure:

$$S_{min} = \min \{ Z_1, \ldots, Z_N \}. \tag{17}$$

Another option uses the largest score:

$$S_{max} = \max \{ Z_1, \ldots, Z_N \}. \tag{18}$$

The mean of the N scores may also be a reasonable choice:

$$S_{mean} = \frac{1}{N} \sum_{i=1}^{N} Z_i \tag{19}$$

Of these, S_{min} seems to be the least promising, since one small Z-value can give a low fused score to a non-target. If each Z_i represents a different ATR algorithm, the inclusion of a single algorithm lacking in power (the ability to reject non-targets) can significantly degrade system performance. Such a low-power algorithm would have a somewhat lesser impact on S_{mean}, and very little impact on S_{max}.

Figure 3 shows the bivariate acceptance regions (at $PD = 0.95$) for the each of the fused measures discussed above, for independent standard normal target scores. The region for S_f is in some ways a compromise between that of S_{max} and S_{mean}. It omits the extensions along the axes seen for S_{mean} and also cuts out the upper right-hand corner (where neither score represents a particularly strong match) of the S_{max} region. Note that if S_{mean} were computed in the Y domain rather than the Z domain, it would be equivalent to S_f.

Follmann [1996] discusses several other fusion measures, but these are applicable only in the multivariate normal case. Further, the shapes of the acceptance regions for these tests with dependent scores make them undesirable for our application: in some instances, increasing a single Z score (indicating poorer fit) can change a decision from "non-target" to "target".

3.2 - Simulation Studies

To test the power of the proposed fusion metric, and to compare its performance with the alternative approaches discussed above, a series of simulation studies were conducted. In each of these studies, ATR scores were generated to represent "targets" and "non-targets". The number, N, of scores per test varied. For targets, the scores were all distributed as standard normal ($\mu=0$, $\sigma=1$), with non-zero correlations in some cases. For non-targets, the test scores were also normally distributed, with positive means. The parameters used for each study are listed in Table 3, and bivariate samples of target and non-target scores are plotted in Figure 4.

The studies were conducted as follows:

1. Generate 10,000 sets of "target" Z-scores and 10,000 sets of "non-target" Z-scores.
2. From each set of raw scores, compute the metrics S_{min}, S_{max}, and S_{mean}.
3. Convert the raw scores according to (2), using the standard normal distribution function. Evaluate the fused metric S_f for each set of scores.
4. Compute the receiver-operator characteristic (ROC) curve [Poor, 1988] for each metric. Record the false alarm rate at selected values of PD.

Results for simulation studies I-a and I-b are shown in Figures 5 and 6. These studies used two scores, one which is a poor discriminator of non-targets (μ=5, σ=3) and one which is fair (μ=4, σ=2). In both the correlated and uncorrelated cases, the metrics S_f, S_{max}, and S_{mean} each significantly improve on the performance seen with either of the individual scores. Of these three, S_{mean} is not as powerful as the others, and S_{max} is particularly effective in the correlated case, due to the increased percentage of targets falling into the upper right-hand corner of the acceptance region (see Figure 4).

In simulation studies II-a and II-b, two fairly good discriminators, with power for different ranges of PD, are combined. Results are summarized in Figures 7 and 8. In both the correlated and uncorrelated cases, S_f, S_{max}, and S_{mean} perform about equally well, while S_{min} represents little improvement over the best of the individual scores for a given PD.

Figures 9 and 10 summarize the results of simulation studies III-a and III-b, in which four ATR scores are fused. Because none of the four gives particularly large values for non-targets (μ=2, σ=1), the performance of S_{max} is not competitive with that of S_f or S_{mean}. Note the strong effect that correlation has on each of the fused metrics: discrimination is decidedly superior when the individual scores are independent. This is not surprising, because the additional information provided by each score is greatest when the scores are minimally correlated. The results demonstrate that with as many as four independent scores available for fusion, excellent performance can be achieved.

In each of these simulations, the fusion metric S_f performs about as well as the best of the other metrics. In studies I-a and I-b, S_{mean} is not competitive, while S_{max} performs relatively poorly in studies III-a and III-b. The metric S_{min} is uniformly less powerful than the others.

4 – DISCUSSION AND CONCLUSIONS

The simulation results demonstrate that S_f can be a powerful metric for ATR fusion. However, the approach has some limitations which merit consideration. First, the method requires the user to completely specify (either parametrically or empirically) the marginal distributions of each of the individual scores, for target vehicles. These distributions must be continuous, or nearly so. In addition, the pairwise correlations between the exponential Y-scores (Equation 2) for targets are needed to set a rejection threshold on the fused score.

If ample training data are available, these requirements may not represent a major challenge. However, modeling errors can have significant implications for classifier performance. For example, if the training data are less variable than the test data (as is likely in many practical applications), one might experience an upward shift in the mean Z-scores for targets in the test set. With the transformation to exponential space, the impact of such a shift can be substantial. In practice, one can correct for this problem by specifying anticipated shifts for each Z-score, and subtracting these values from the raw scores before transformation to the exponential. While this approach is theoretically appealing, it is not a trivial problem to anticipate the effect of future test conditions.

A related issue is that of robustness. With the metric S_f, an extreme Z-value for just one of the individual scores can cause the rejection of a legitimate target, due to the nature of the nonlinear transformation from Z-space to Y-space. Metrics based on mean (or median) Z-scores may be more robust when all raw scores are reported in the same scale, but are not an option in other situations.

The robustness of the probabilistic fusion approach can potentially be improved by the use of a weighted sum of the Y_i's in place of the unweighted sum (3). Relatively low weights

would be assigned to the scores thought to be non-robust, with higher weights assigned to the robust scores. Measures of robustness may be inferred from auxiliary information like sensor capability, image quality, or statistical considerations. The distribution of any weighted sum can be modeled as gamma, with simple adjustments to the moment estimates of equation (5). This procedure would have the effect of stretching the acceptance region (Figure 3) along the axes of the down-weighted Z-scores.

Finally, it should be noted that the fusion metric is used to test a single null hypothesis: a particular target type at a specific location and orientation. In practical applications, several different target hypotheses may be in competition. When this is the case, separate tests are conducted for each hypothesis. If multiple target hypotheses are accepted at the chosen significance level, the one with fused score giving the largest p-value (computed from the appropriate Gamma distribution function) is selected.

The ATR fusion metric S_f has been developed in this report, using straightforward probabilistic reasoning. Simulation results demonstrate the power of the technique. However, the user needs to be cognizant of its significant modeling requirements, and of its potential difficulties with robustness. Practical suggestions have been made for resolving these issues.

5 – REFERENCES

Anderson, T.W. (1984). An Introduction to Multivariate Statistical Analysis, Second Edition. New York: John Wiley & Sons, Inc.

Barnett, V. (1980). Some bivariate uniform distributions. *Communications in Statistics - Theory and Methods,* vol. A9, no. 4. p. 453 - 461.

Fisher, R.A. (1932). Statistical Methods for Research Workers, 4th Edition. Edinburgh, London: Oliver and Boyd.

Follmann, D. (1996). A simple multivariate test for one-sided alternatives. *JASA,* vol. 91, no. 434. p. 854 - 861.

Johnson, N.L., and Kotz, S. (1970). Continuous Univariate Distributions - 1. New York: John Wiley & Sons, Inc.

Poor, H.V. (1988). An Introduction to Signal Detection and Estimation. New York: John Wiley & Sons, Inc.

Silvey, S.D. (1975). Statistical Inference. London: Chapman and Hall.

Van Zwet, W.R., and Oosterhoff, J. (1967). On the combination of independent test statistics. *Annals of Mathematical Statistics,* vol. 38, p. 659 – 680.

Table 1 - Gamma approximation to the Morganstern model. The exact and approximate distribution functions are compared. For each value of the parameter ψ, gamma quantiles were computed for probability of detection equal to 90%, 95%, 99%, and 99.9%. The exact distribution function (Equation 11) was evaluated at each of these quantile values, with results as shown in the body of the table. Even in the worst case, for $\psi = -1$, good agreement is observed. The exact and approximate density functions for each value of ψ are shown in Figure 1.

Estimated PD (Gamma Model)	*Exact PD (Morganstern Model)* $\psi = -1.0$	$\psi = -0.5$	$\psi = 0.25$	$\psi = 1.0$
.900	.894	.897	.901	.902
.950	.950	.950	.950	.950
.990	.992	.991	.990	.989
.999	.999	.999	.999	.999

Table 2 - Gamma approximation to the multivariate normal model. The empirical and approximate distribution functions are compared. In each set of simulations, all off-diagonal elements of the normal covariance matrix were set equal to R. For each pair of R and M (the dimension of the multivariate normal variates), an initial sample of size 200 was generated for use in estimating the gamma parameters. Comparisons were then made between the gamma approximations and the empirical distributions observed for independent samples of size 10,000. Results are shown for the probability of detection set at 90%, 95%, 99%, and 99.9%. The observed agreement is strong. Histograms of the empirical distribution functions, along with the corresponding gamma approximations, are shown in Figure 2.

Estimated PD (Gamma Model)	*Empirical PD (Normal Model)*			
	M=2, R=-.25	*M=2, R=.75*	*M=5, R=.25*	*M=5, R=.75*
.900	.901	.902	.899	.902
.950	.949	.951	.949	.951
.990	.989	.991	.991	.992
.999	.999	.999	.999	.999

Table 3 - Simulation study parameters. The first four studies used two dimensions each, while the final two studies used four. In each study, every dimension was standard normal ($\mu = 0$, $\sigma = 1$) for targets, while non-target scores were also normally distributed, with means and standard deviations (μ_{non} and σ_{non}, respectively) as specified on the table. A simple dependence model was used, in which the correlation matrices for targets and non-targets both had off-diagonal elements all equal to ρ.

MODEL	*PARAMETERS*		
I-a	$\mu_{non} = \{4, 5\}$	$\sigma_{non} = \{2, 3\}$	$\rho = 0.0$
I-b	$\mu_{non} = \{4, 5\}$	$\sigma_{non} = \{2, 3\}$	$\rho = 0.5$
II-a	$\mu_{non} = \{3, 5\}$	$\sigma_{non} = \{1, 2\}$	$\rho = 0.0$
II-b	$\mu_{non} = \{3, 5\}$	$\sigma_{non} = \{1, 2\}$	$\rho = 0.5$
III-a	$\mu_{non} = \{2, 2, 2, 2\}$	$\sigma_{non} = \{1, 1, 1, 1\}$	$\rho = 0.0$
III-b	$\mu_{non} = \{2, 2, 2, 2\}$	$\sigma_{non} = \{1, 1, 1, 1\}$	$\rho = 0.5$

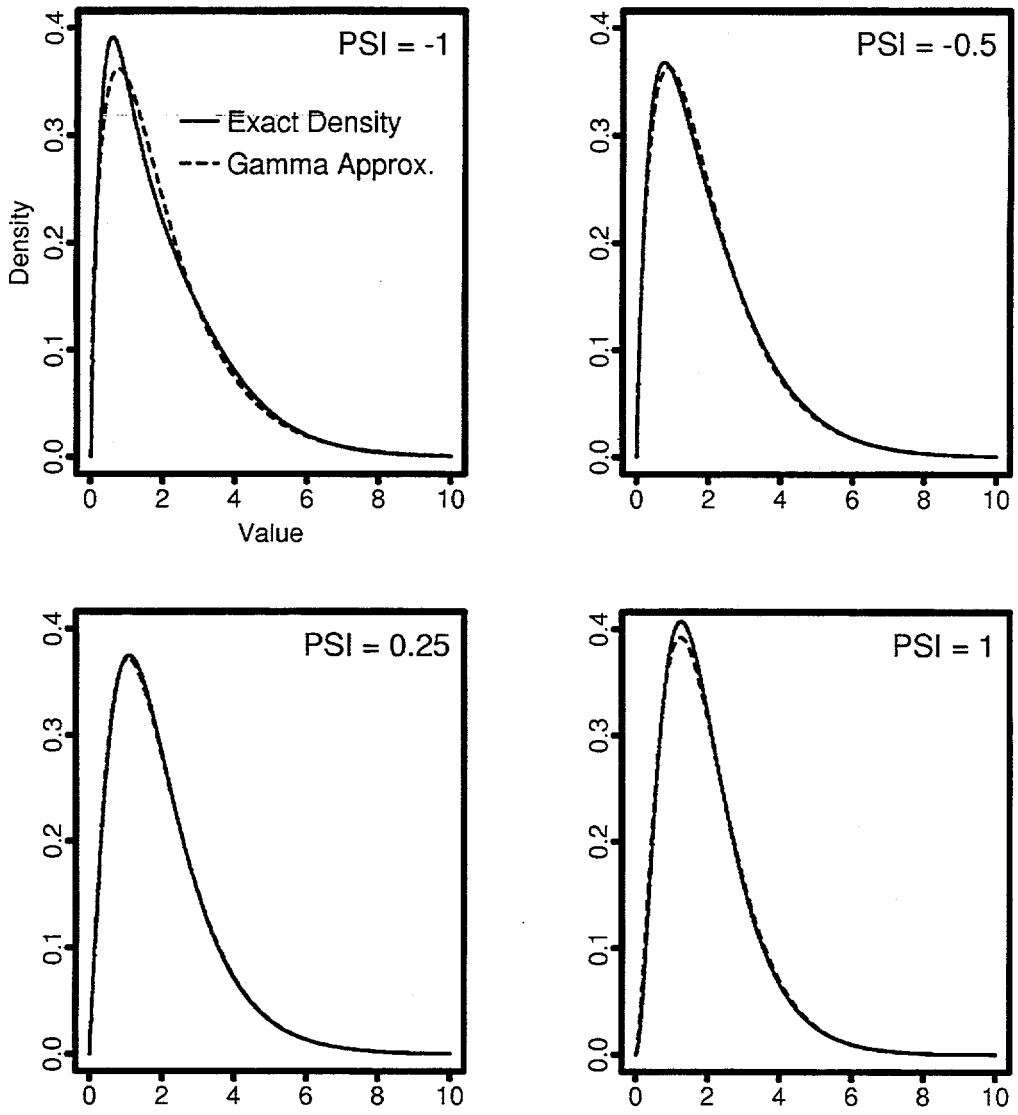

Figure 1 - Gamma approximation to the Morganstern model. The graphics compare the exact and approximate densities of a sum of dependent exponential random variates from the Morganstern-type bivariate uniform model. Results are shown for four different values of the parameter ψ, and the gamma approximation is excellent in all cases.

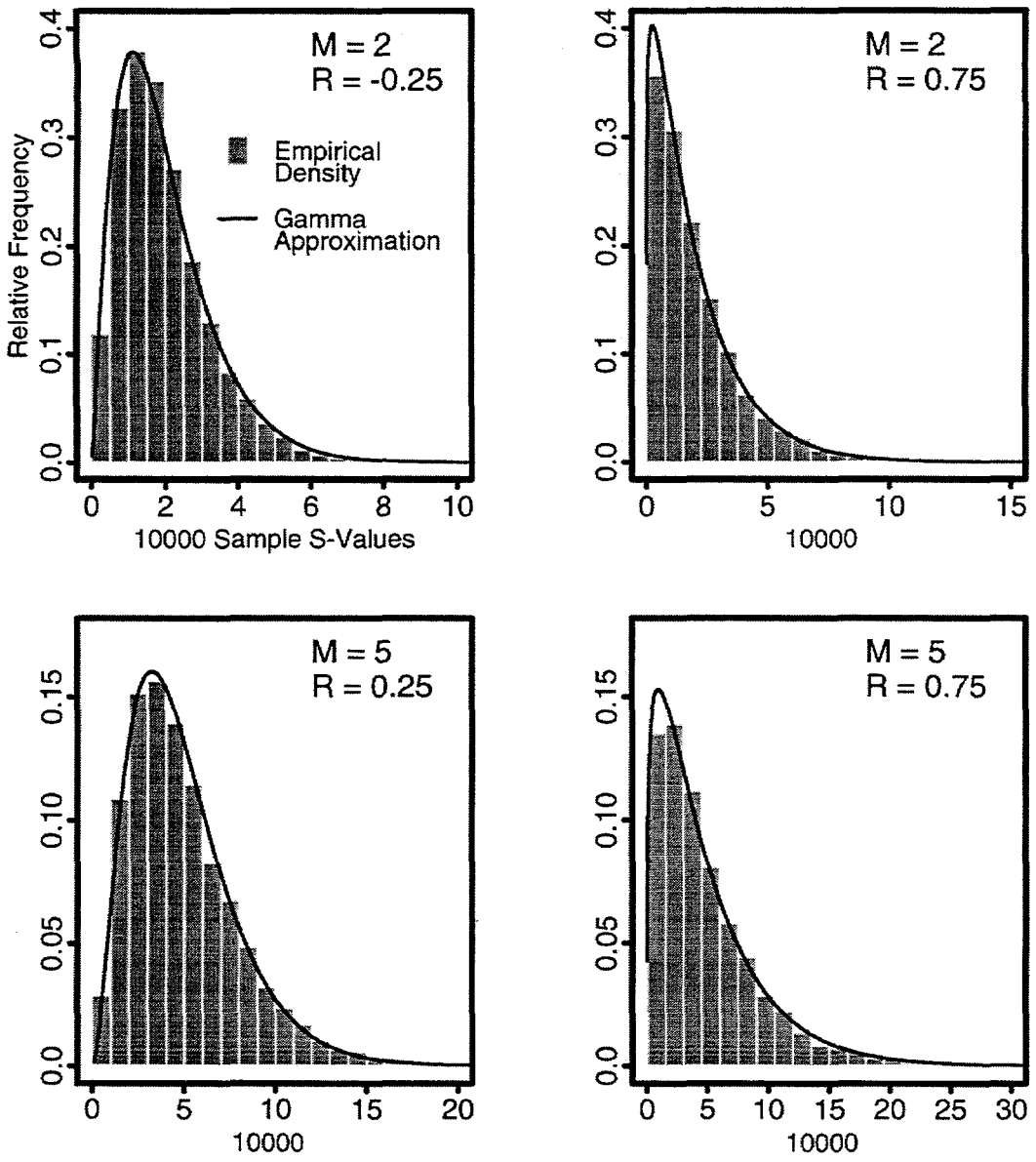

Figure 2 - Gamma approximation to the normal model. The histograms show empirical densities of sums of exponentials obtained from the multivariate normal model, while the lines represent the densities of the method-of-moments gamma approximations. Results are shown in two- and five-dimensional cases, with different levels of pairwise correlation between the underlying normal random variates.

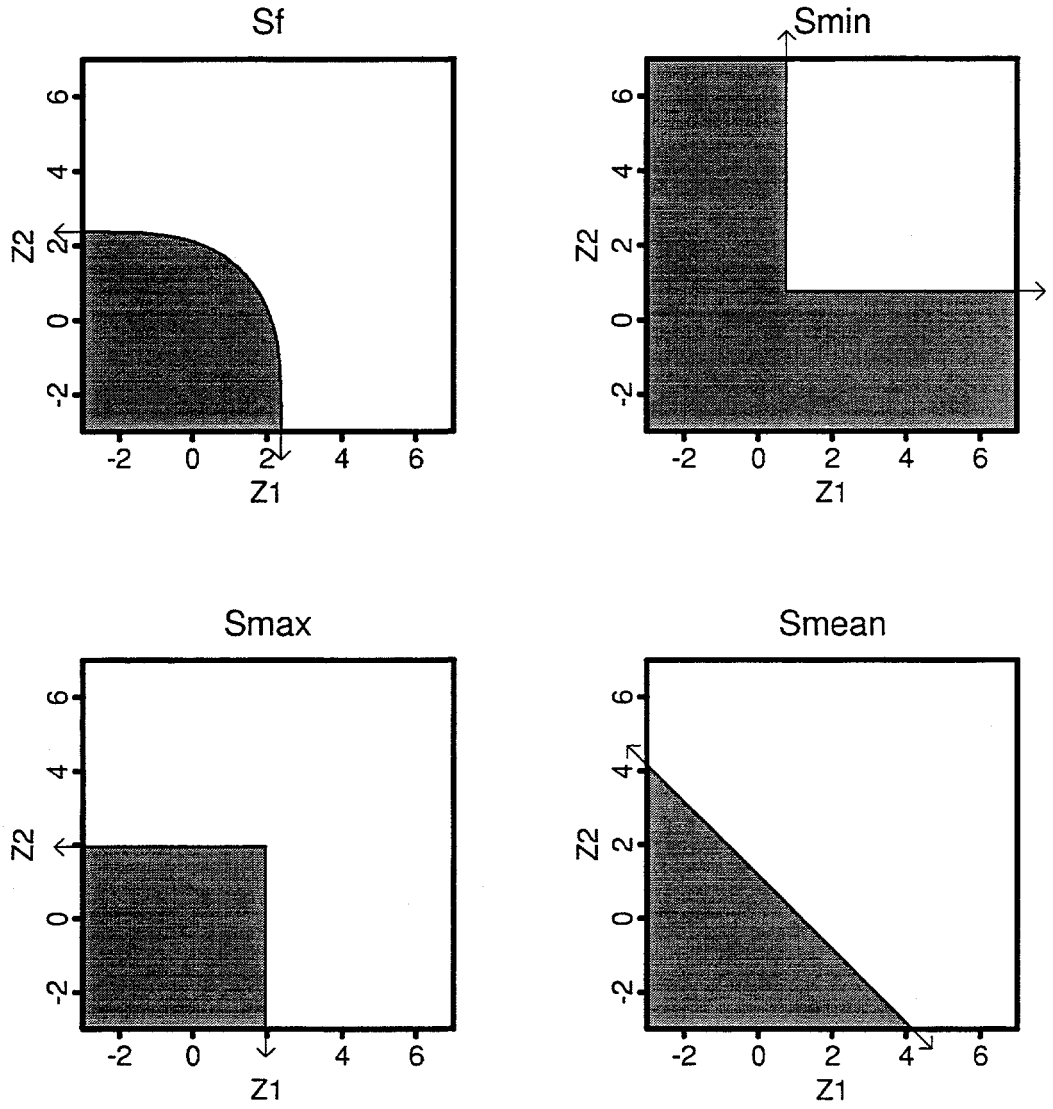

Figure 3 - Acceptance regions for PD=0.95. The shapes of the regions are shown for four different fusion metrics. These regions all correspond to the case where two independent standard normal *Z*-values are available.

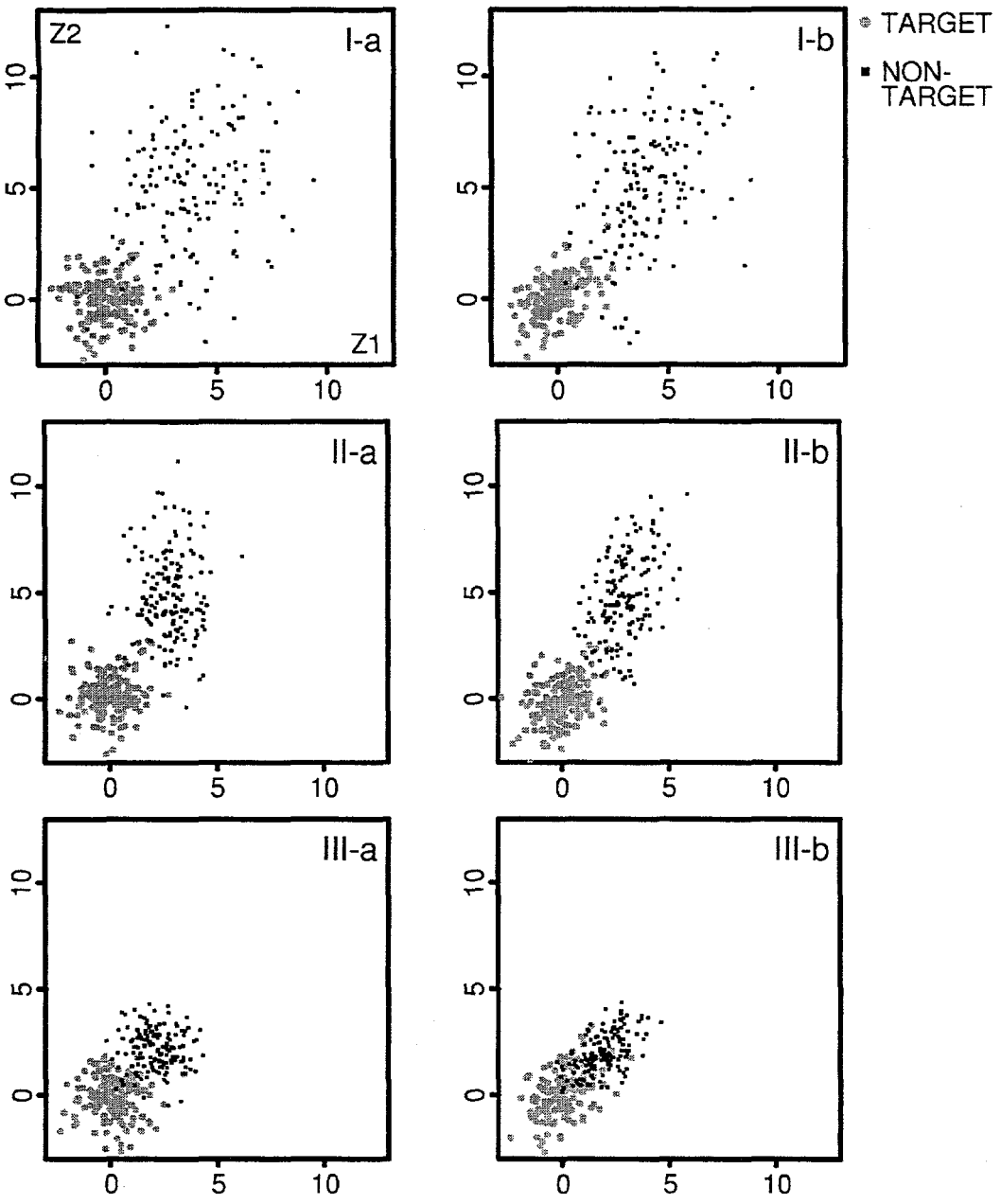

Figure 4 - Sample of bivariate scatterplots. The distributions of target and non-target *Z*-scores are plotted for each of the six simulation studies. For studies III-a and III-b, only the first two of four dimensions are shown. These samples represent a small subset of the 10,000 target and non-target scores generated in each simulation study.

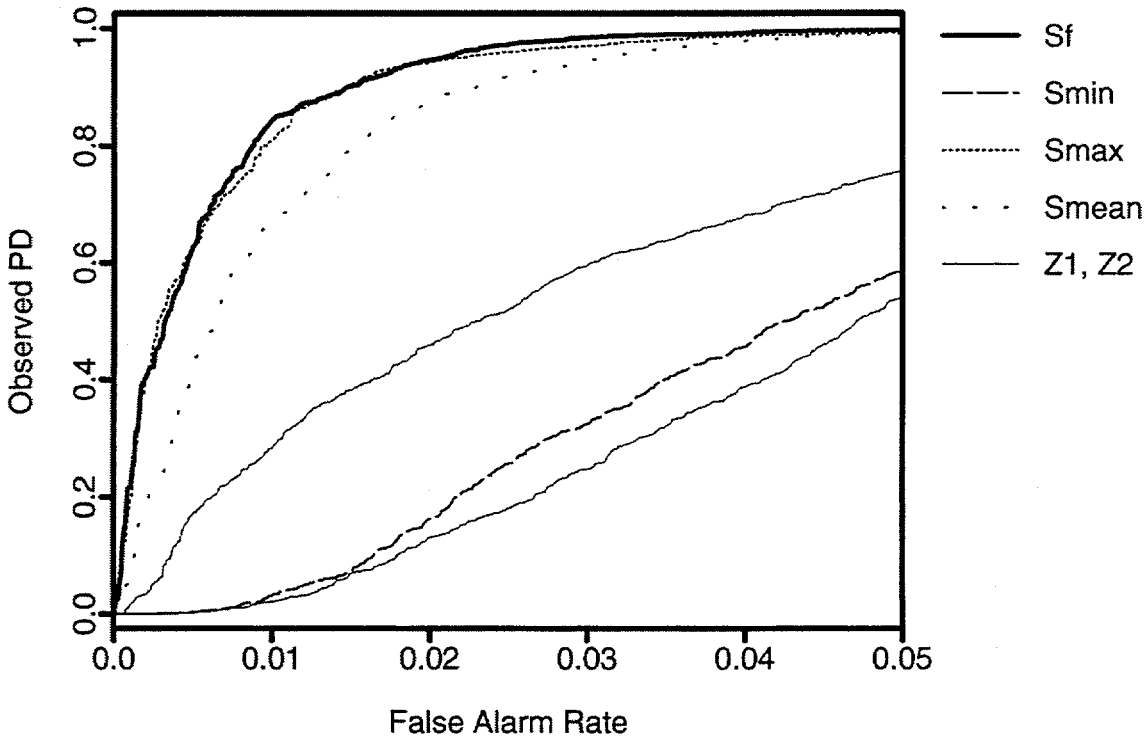

Figure 5 - ROC curves for study I-a. Two independent but fairly weak discriminators are fused. Each of the metrics S_f, S_{max}, and S_{mean} represents a substantial improvement over the Z-scores taken individually. False alarm rates are expressed in percentages of 10,000 simulated non-targets.

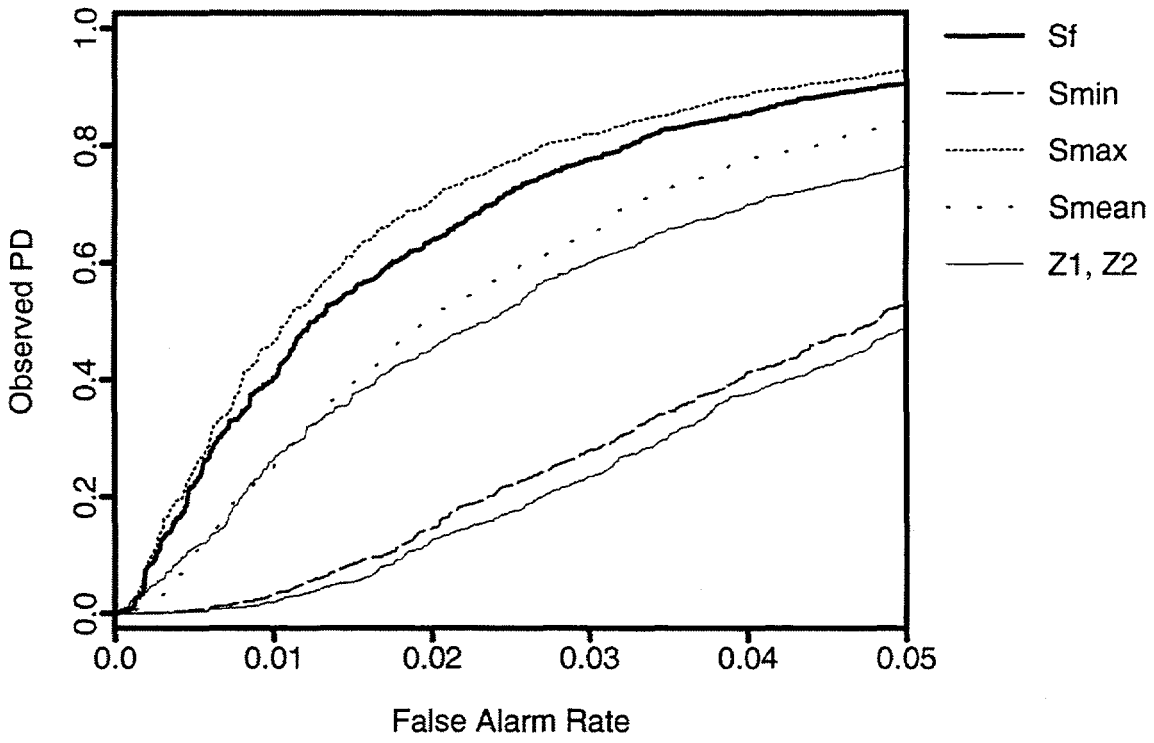

Figure 6 - ROC curves for study I-b. Two dependent and fairly weak discriminators are fused. While S_f, S_{max}, and S_{mean} continue to outperform the individual scores, the magnitude of the improvement decreases from that seen in the independence case (Figure 5). False alarm rates are expressed in percentages of 10,000 simulated non-targets.

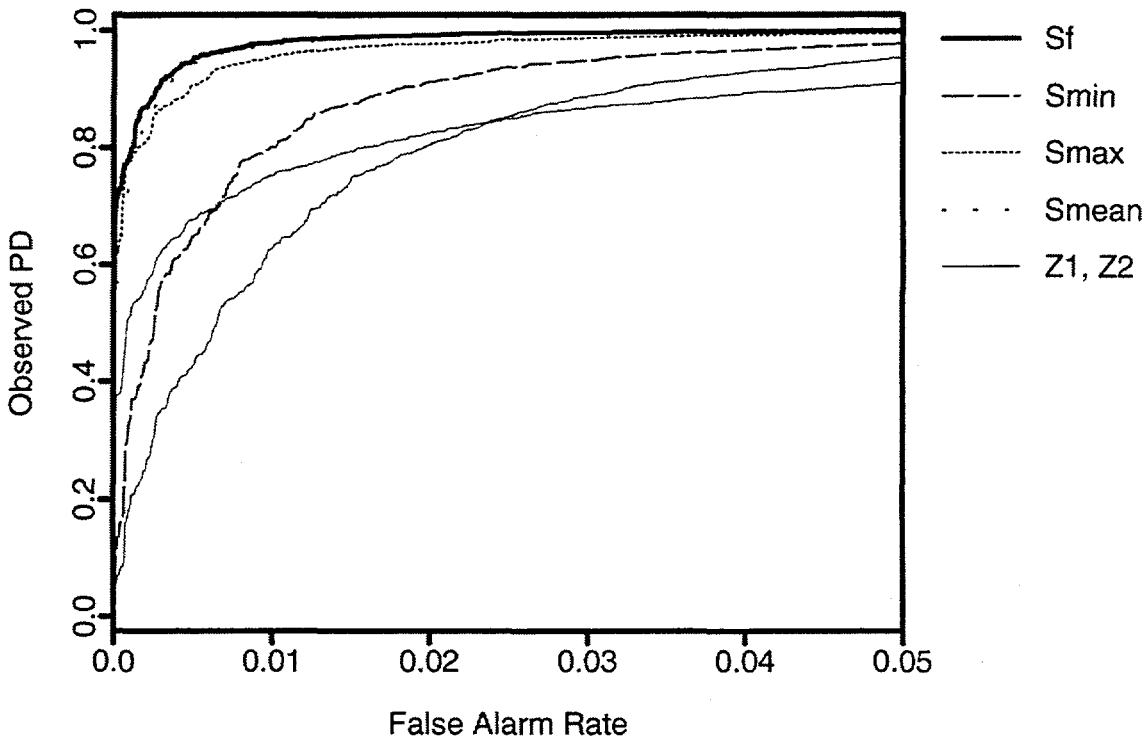

Figure 7 - ROC curves for study II-a. Here, two independent and moderately powerful discriminators are fused. The measures S_f, S_{max}, and S_{mean} all perform about equally, effectively discriminating between targets and non-targets. False alarm rates are expressed in percentages of 10,000 simulated non-targets.

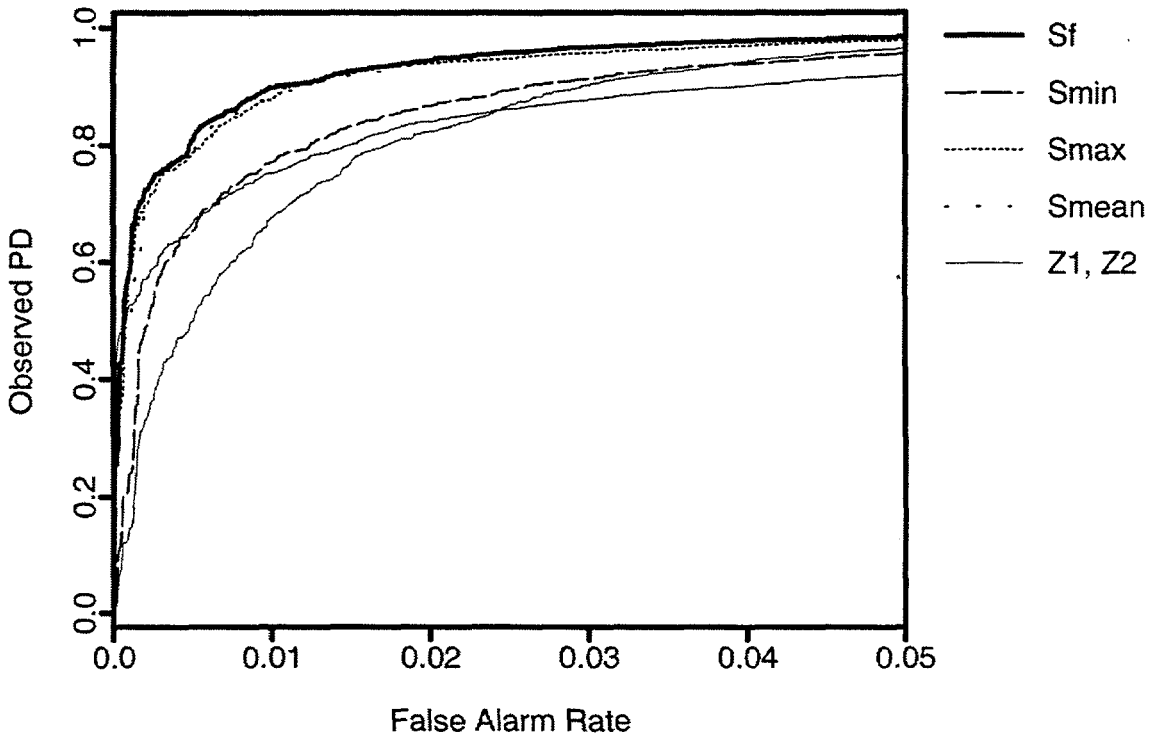

Figure 8 - ROC curves for study II-b. With correlated *Z*-scores, the effectiveness of the fusion metrics is reduced from the independence case (Figure 7). S_f, S_{max}, and S_{mean} perform about equally, while S_{min} represents little improvement over the best of the individual scores. False alarm rates are expressed in percentages of 10,000 simulated non-targets.

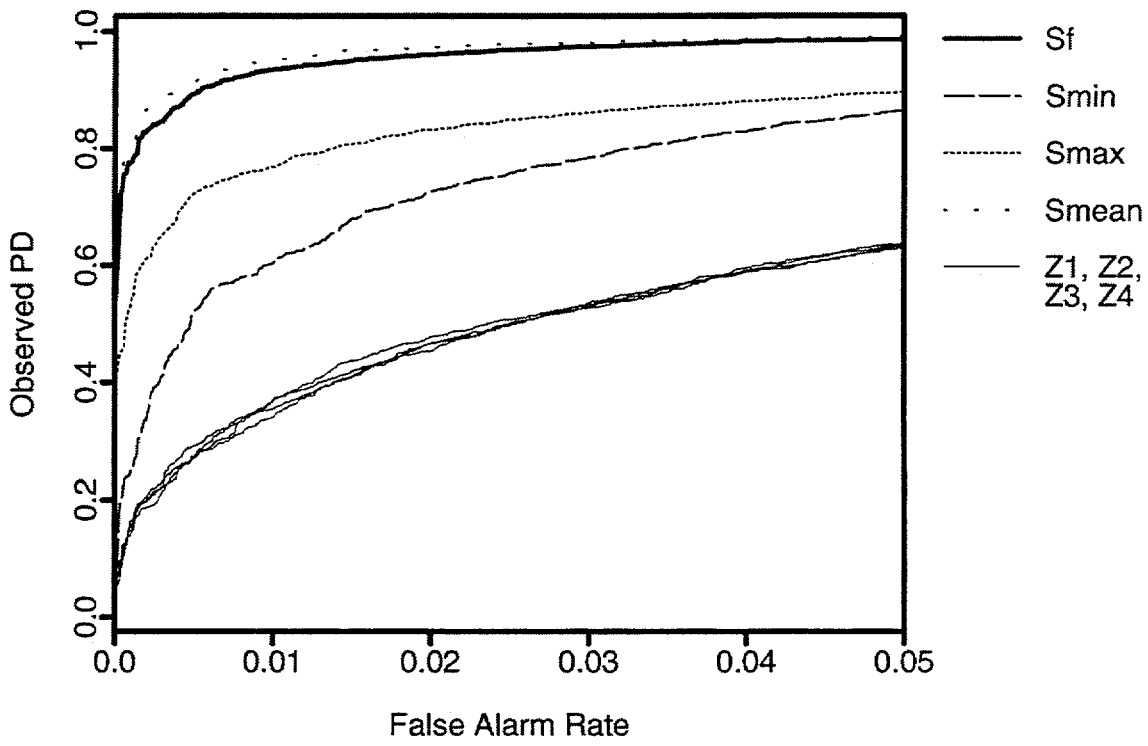

Figure 9 - ROC curves for study III-a. Four independent but fairly weak discriminants are fused. The measures S_f and S_{mean} perform well, while S_{max} and S_{min} are less effective. False alarm rates are expressed in percentages of 10,000 simulated non-targets.

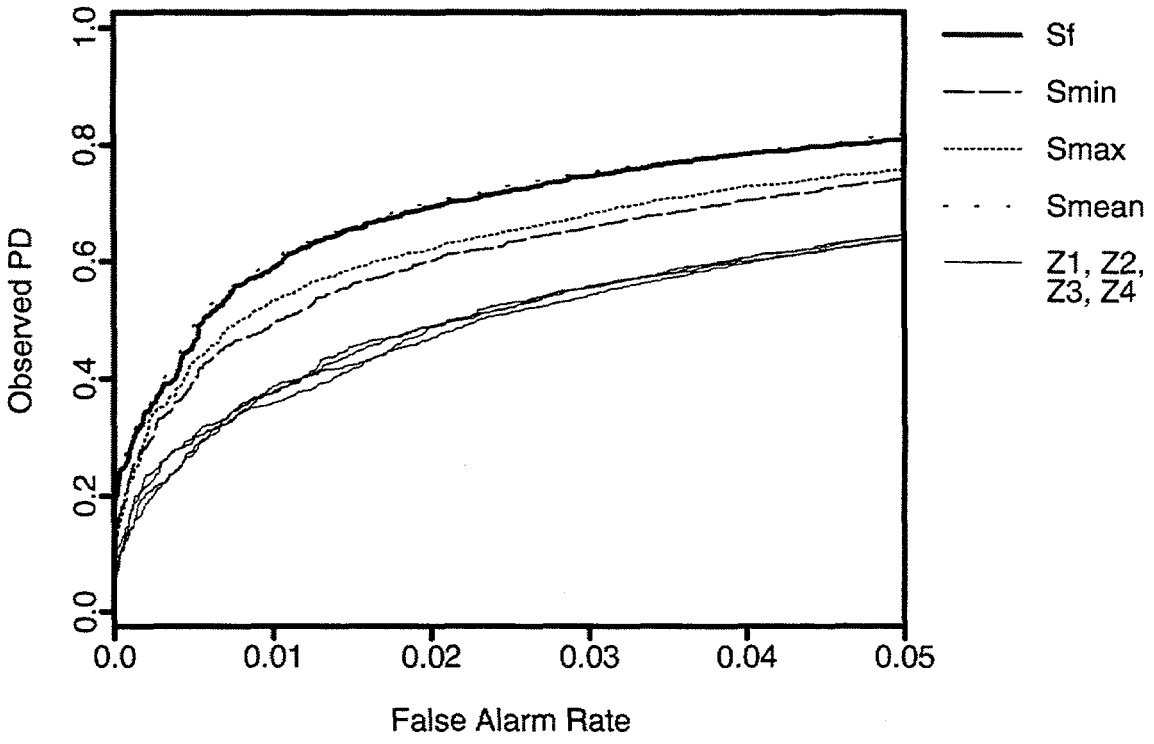

Figure 10 - ROC curves for study III-b. Here, four weak and correlated discriminants are fused. While the fused metrics still outperform the raw scores, the ROC curves show that the magnitude of the improvement is greatly reduced from that observed in the independence case (Figure 9). False alarm rates are expressed in percentages of 10,000 simulated non-targets.

DISTRIBUTION:

1	MS0844	Larry Hostetler, 2523
35	0844	Katherine Simonson, 2523
1	9018	Central Technical Files, 8940-2
2	0899	Technical Library, 4916
2	0129	Review & Approval Desk, 12690 For DOE/OSTI