MAXIMUM LIKELIHOOD SIGNAL CLASSIFICATION USING SECOND-ORDER BLIND DECONVOLUTION PROBABILITY MODEL

Maya R. Gupta and Hyrum S. Anderson

University of Washington
Dept. of Electrical Engineering
Seattle, WA 98195

ABSTRACT

We address the problem of classifying a signal that has been corrupted by an unknown linear time-invariant filter. This problem is common in remote-sensing and non-destructive evaluation applications where multipath and spreading are prevalent. A traditional approach is blind deconvolution to estimate the original signal, followed by classification of the estimated signal. Blind deconvolution is an ill-posed estimation problem, and if only a classification is needed, then we hypothesize it is an unnecessary step. We present an alternative maximum likelihood classifier that uses second-order probability models for the original signal and the unknown corrupting filter. The resulting quadratic discriminant analysis classifier is shown to perform well over a range of signal-to-noise ratios for two different models of multipath, and in all cases performs consistently better than a standard blind deconvolution method followed by a quadratic discriminant analysis classifier.

***Index Terms*—** multipath, deconvolution, classification, Gaussian process

1. INTRODUCTION

Consider the problem of classifying a signal z as one of a finite number of classes $y = \{1, 2, \ldots, G\}$ where z is defined by

$$z = h * x + n, \tag{1}$$

where $*$ is the convolution operator, h has length M_h and is the impulse response of a linear time-invariant filter, and x is a signal of length M_x. The signal z is known and h, x and n are assumed to be unknown, but some information will be assumed about their distributions. This problem of classifying a signal that has undergone unknown filtering arises in many remote-sensing and non-destructive evaluation scenarios, where the signal-of-interest is x, but one receives the corrupted signal z, and h characterizes the effect of multipath or other spreading characteristics of the media through which the signal x has traveled.

Many researchers have approached the problem by suggesting that one first form an estimate of the deconvolved signal $\hat{x}$ using a blind deconvolution technique, and then classify $\hat{x}$ [1–3]. In previous work, we have demonstrated that simple methods to jointly address deconvolution and classification can outperform deconvolution followed by classification [4]. That approach treats signals deterministically, and the performance depends on how closely test signals match training signals in a deterministic sense.

In this paper we propose a maximum likelihood classifier that stochastically models x, h and n and avoids the ill-posed problem of estimating a deconvolved signal $\hat{x}$. Instead, the deconvolution is treated probabilistically as part of the maximum likelihood optimization that produces the class label. In Section 2 we present the classification method based on second order statistics of the unknown signals. Experimental setup is described in Section 3 and results are presented in Section 4. We close with a discussion in Section 5.

2. JOINT QDA CLASSIFIER

Let the class-conditional distribution of a random signal X belonging to class y have finite mean μ_{xy} and finite covariance Σ_{xy}. We model the noise signal n as a realization of the random noise signal N, and assume N is drawn independently and identically from $\mathcal{N}(0, \sigma_n^2 I)$ and is independent of X. Assume the filter's impulse response h is a realization of a random impulse response H that is independent of X and N and that the distribution of H has finite mean μ_h and finite covariance Σ_h. We model the received signal z as a realization of a random signal Z where

$$Z = X * H + N. \tag{2}$$

A standard approach to classification is quadratic discriminant analysis (QDA), which models each class-conditional distribution as a Gaussian distribution [5, 6]. We apply QDA to the $(M_x + M_h - 1)$-dimensional feature space formed by the coefficients of the corrupted signal z, such that $p(z|y)$ is

This work was supported by a United States Office of Naval Research Young Investigator Program Award.

modeled as a Gaussian distribution $\mathcal{N}(\mu_{zy}, \Sigma_{zy})$. One motivation for QDA is that the Gaussian distribution is the maximum entropy (i.e. least assumptive) distribution given fixed second-order statistics.

Given QDA's class-conditional Gaussian models for Z, the maximum likelihood class of the corrupted signal z is the class y^* that maximizes the class-conditional Gaussian probability $p(z|y)$, that is y^* solves

$$\arg\min_{y=1,2,\ldots,G} \left(\log|\Sigma_{zy}| + (z-\mu_{zy})^T \Sigma_{zy}^{-1}(z-\mu_{zy})\right). \quad (3)$$

Next, we give closed-form solutions for μ_{zy} and Σ_{zy} given μ_{xy}, Σ_{xy}, μ_h, Σ_h, and σ_n. Specifically, the mth component of the mean signal μ_{zy} is derived from (2) to be

$$\begin{aligned}
\mu_{zy}[m] &= E_{p(z|y)}[Z[m]] \quad (4)\\
&= E_{p(x|y)p(h)p(n)}\left[(X*H)[m] + N[m]\right]\\
&= E_{p(x|y)p(h)}\left[\sum_k X[k]H[m-k]\right]\\
&= \sum_k E_{p(x|y)}\left[X[k]\right] E_{p(h)}\left[H[m-k]\right]\\
&= \sum_k \mu_{xy}[k]\mu_h[m-k],
\end{aligned}$$

and thus $\mu_{zy} = \mu_{xy} * \mu_h$. Using $**$ to denote two-dimensional convolution,

$$\begin{aligned}
\Sigma_{zy} &= E_{p(z|y)}\left[ZZ^T\right] - E_{p(z|y)}\left[Z\right] E_{p(z|y)}\left[Z\right]^T\\
&= E\left[(X*H+N)(X*H+N)^T\right] - \mu_{zy}\mu_{zy}^T\\
&\overset{a}{=} E\left[(X*H)(X*H)^T\right] + E\left[NN^T\right] - \mu_{zy}\mu_{zy}^T\\
&\overset{b}{=} E\left[(XX^T)**(HH^T)\right] + \sigma_n^2 I - \mu_{zy}\mu_{zy}^T\\
&\overset{c}{=} E\left[(XX^T)\right]**E\left[(HH^T)\right] + \sigma_n^2 I - \mu_{zy}\mu_{zy}^T\\
&= (\Sigma_{xy} + \mu_{xy}\mu_{xy}^T)**(\Sigma_h + \mu_h\mu_h^T)\\
&\quad + \sigma_n^2 I - \mu_{zy}\mu_{zy}^T,
\end{aligned}$$

where the expectations are taken with respect to the appropriate distributions, recalling that $p(z|y) = p(x|y)p(h)p(n)$. Note that (c) follows from (b) due to the linearity of both expectation and convolution, and that (a) can be re-written as (b) because the nth-mth component of the outer product $(X*H)(X*H)^T$ can be expressed as

$$\begin{aligned}
&\left(\sum_k X[k]H[n-k]\right)\left(\sum_i X[i]H[m-i]\right)\\
&= \sum_k\sum_i X[k]H[n-k]X[i]H[m-i]\\
&= \sum_k\sum_i \left(X[k]X[i]\right)\left(H[n-k]H[m-i]\right),
\end{aligned}$$

and thus $(X*H)(X*H)^T = (XX^T)**(HH^T)$.

This classification approach requires the second-order statistics for the random processes generating h and x; most blind deconvolution algorithms also require some information or assumptions about either x or h or both [3, 7].

3. EXPERIMENTS

Two-class classification experiments were conducted with two different models of multipath filtering for the filter h. For all of the experiments the signals X were drawn from class-conditional Gaussian random processes, with known mean-signals and covariance matrices. Each signal X had length $M_x = 500$ and was drawn iid from one of two classes with equal probability, and then $X \sim \mathcal{N}(\mu_{xy}, \Sigma_{xy})$, where the μ_{xy} signal was constant over time and the Σ_{xy} was diagonal (these values are given in Tables 1 and 2). Each signal $Z = X*H+N$ was created by randomly drawing a signal X and H, and then different noise signals N were added for different SNRs. Results averaged over 1000 test signals are given in Tables 1 and 2 for different class distributions.

The first model for the multipath was that each coefficient of the random filter H was drawn independently from a Laplacian random process with parameters $\mu[k]$, $b[k]$:

$$p(h[k]|\mu[k], b[k]) = \frac{1}{2b[k]} e^{-(|x-\mu[k]|/b[k])}, \quad (5)$$

for all $k = 1, \ldots, 100$, where $\mu_h[1] = 1$ and $\mu_h[k] = 0$ for all $k > 1$, and the scale parameter b decays as k grows: $b[k] = 0.2e^{-0.024(k-1)}$. The decay parameter coefficients for this experiment were chosen to model oceanic multipath filtering of sonar signals. An example realization of a filter h drawn from (5) is shown in Fig. 1. Laplacian random processes create sparse realizations which are an appropriate model for some multipath filtering phenomena. Results with this multipath model are given in Table 1.

A second multipath model was formed such that the stochastic filter H had independent components where the nth component was

$$H[n] = \begin{cases} \exp(-\beta n) & \text{with probability } p/2\\ -\exp(-\beta n) & \text{with probability } p/2\\ 0 & \text{with probability } 1-p. \end{cases} \quad (6)$$

for $n > 1$, $H[1] = 1$, and to model oceanic-sonar multipath, we set $p = 0.15$ and $\beta = 0.024$. This stochastic model is similar to the standard K-impulse deterministic model of multipath where $h = \sum_{k=1}^{K} \alpha_k \delta[n-n_k]$ where $\alpha_k = \pm\exp(-\beta n_k)$ to model propagation loss at rate β [4, 8], but rather than specify the number of impulses K, in our stochastic model the probability of an impulse at a given time is specified. An example realization of the sparse multipath defined above is shown in Fig. 2. With multipath H specified in (6), the mean is $\mu_h[n] = 0$ for all $n > 1$ and $\mu_h[1] = 1$, and the covariance Σ_h is a diagonal matrix, where the first component of

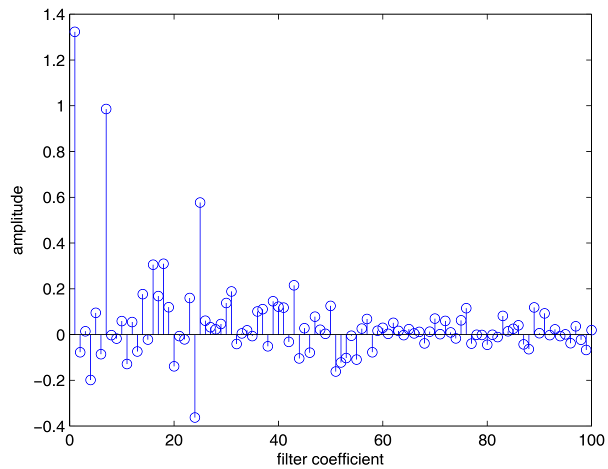

Fig. 1. Example realization of the random filter H drawn from a multivariate Laplace distribution to model multipath.

the diagonal is zero, and the nth component of the diagonal is $p\exp(-2\beta n)$ for $n > 1$.

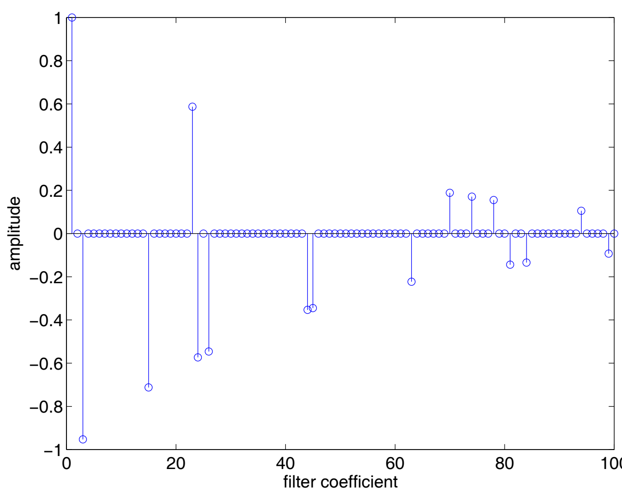

Fig. 2. Example realization of the random filter H drawn from the mulitpath model given in (6).

The proposed joint QDA classifier given in (3) was compared to a two-step classifier that first uses blind deconvolution to estimate the signal $\hat{x}$, then classifies $\hat{x}$ using quadratic discriminant analysis. That is, the signal z is classified based on the maximum-likelihood classification of the estimated deconvolved signal $\hat{x}$ such that the estimated class $\hat{y}$ maximizes $p(\hat{x}|y)$, that is $\hat{y}$ solves

$$\arg\min_{y=1,2,\ldots,G} \left(\log|\Sigma_{xy}| + (\hat{x} - \mu_{xy})^T \Sigma_{xy}^{-1} (\hat{x} - \mu_{xy}) \right). \tag{7}$$

The blind deconvolution method used to estimate $\hat{x}$ was proposed by Cabrelli for signals that have undergone unknown multipath filtering [9]. Cabrelli's method was shown to produce deconvolved signal estimates $\hat{x}$ that were significantly better correlated with the true x than the deconvolution estimates produced by other blind deconvolution methods, particularly in the presence of additive noise [3]. The approach uses the D-norm sparseness criterion to estimate the multipath filter h as well as an estimate of $\hat{x}^{-1}$ which must be inverted to form an estimate of $\hat{x}$ [3]. Given Cabrelli's $\hat{x}^{-1}$, we estimated $\hat{x}$ using a regularized inversion in the Fourier domain.

As a baseline comparison, the two methods were also compared to a QDA classifier that ignores multipath. That is, each test signal z was truncated to be of length $M_z = M_x = 500$, and QDA was performed on the truncated test signal using the true mean and covariance class-conditional statistics for random signals X. Each classification algorithm used the true mean and covariance statistics of X, and the joint QDA classifier also used the true mean and covariance statistics of H.

4. RESULTS

The results are given in Tables 1 and 2 and show that the proposed joint QDA classifier can use both differentiating information in the signal means and signal covariances to achieve significantly better classification accuracy than the compared methods. In particular, in the fourth case the means are close, but the covariances are very well-separated, and thus the two signal classes can be differentiated. The proposed Joint QDA method captures the differentiating covariancek information, and uses it to improve its performance over the 3rd case, where the class covariances were the same. In cases where the class covariances differed, the blind-deconvolution-then-QDA method classified almost all signals as the higher-variance class.

5. DISCUSSION AND OPEN QUESTIONS

We proposed a maximum likelihood architecture we termed Joint QDA to classify signals corrupted by unknown convolution noise and unknown additive noise that does not attempt ill-posed blind deconvolution as an intermediate step. The experiments showed that this joint-classification-and-deconvolution approach can outperform the two-step approach of Cabrelli's blind deconvolution followed by classification. Cabrelli's blind deconvolution method was designed to deconvolve out sparse filters h, and has been previously proposed for use with multipath [3, 9]. Joint QDA did not use the fact that h is expected to be sparse, but did use second-order statistics of the random process generating h.

Gaussian mixture model classifiers have less model bias than QDA, and it would increase the applicability of the proposed classifier if the class-conditional probability model for the corrupted signals z was extended to be a mixture of Gaussians. To improve robustness, the maximum likelihood architecture could be replaced by a Bayesian integration [6]. Also,

Table 1. Average classification accuracy for Laplacian multipath for 1000 iid test signals for each case.

Well-separated Means, Same Covariances $\mu_{x1} = -2, \Sigma_{x1} = I$ and $\mu_{x2} = 2, \Sigma_{x2} = I$					
SNR	**-10**	**-5**	**0**	**5**	**10**
Joint QDA	84.8	94.9	98.5	98.7	99.2
Blind Deconv+QDA	64.2	71.0	76.0	75.9	75.4
QDA w/o Deconv	67.2	77.1	79.5	80.0	79.7

Well-separated Means, Different Covariances $\mu_{x1} = -2, \Sigma_{x1} = I$ and $\mu_{x2} = 2, \Sigma_{x2} = 3I$					
SNR	**-10**	**-5**	**0**	**5**	**10**
Joint QDA	95.1	99.1	100	100	99.9
Blind Deconv+QDA	51.5	51.5	51.5	51.6	51.6
QDA w/o Deconv	54.0	72.3	79.2	82.4	83.2

Close Means, Same Covariances $\mu_{x1} = 2, \Sigma_{x1} = I$ and $\mu_{x2} = 2.5, \Sigma_{x2} = I$					
SNR	**-10**	**-5**	**0**	**5**	**10**
Joint QDA	53.2	58.7	58.8	64.8	66.6
Blind Deconv+QDA	49.8	53.0	53.8	52.6	55.2
QDA w/o Deconv	51.3	52.1	51.4	52.4	52.5

Close Means, Different Covariances $\mu_{x1} = 2, \Sigma_{x1} = I$ and $\mu_{x2} = 2.5, \Sigma_{x2} = 10I$					
SNR	**-10**	**-5**	**0**	**5**	**10**
Joint QDA	100	100	100	100	100
Blind Deconv+QDA	51.5	51.5	51.5	51.6	51.8
QDA w/o Deconv	49.2	50.0	51.5	52.8	53.2

Table 2. Average classification accuracy for sparse multipath model given in (6) for 1000 iid test signals for each case.

Well-separated Means, Same Covariances $\mu_{x1} = -2, \Sigma_{x1} = I$ and $\mu_{x2} = 2, \Sigma_{x2} = I$					
SNR	**-10**	**-5**	**0**	**5**	**10**
Joint QDA	74.4	89.4	96.2	99.1	99.9
Blind Deconv+QDA	60.8	67.3	69.8	71.9	71.7
QDA w/o Deconv	62.7	69.4	72.1	72.8	72.7

Well-separated Means, Different Covariances $\mu_{x1} = -2, \Sigma_{x1} = I$ and $\mu_{x2} = 2, \Sigma_{x2} = 3I$					
SNR	**-10**	**-5**	**0**	**5**	**10**
Joint QDA	98.9	100	100	100	100
Blind Deconv+QDA	49.1	49.1	49.1	49.1	49.1
QDA w/o Deconv	51.0	67.5	74.3	76.5	77.9

Close Means, Same Covariances $\mu_{x1} = 2, \Sigma_{x1} = I$ and $\mu_{x2} = 2.5, \Sigma_{x2} = I$					
SNR	**-10**	**-5**	**0**	**5**	**10**
Joint QDA	53.1	56.2	61.0	65.5	67.2
Blind Deconv+QDA	50.7	52.6	52.0	53.6	54.6
QDA w/o Deconv	48.9	50.8	52.5	52.7	52.9

Close Means, Different Covariances $\mu_{x1} = 2, \Sigma_{x1} = I$ and $\mu_{x2} = 2.5, \Sigma_{x2} = 10I$					
SNR	**-10**	**-5**	**0**	**5**	**10**
Joint QDA	100	100	100	100	100
Blind Deconv+QDA	49.1	49.1	49.1	49.1	49.1
QDA w/o Deconv	49.2	49.2	49.5	49.8	49.8

it is an open question how sensitive the classifier is to the accuracy of the second-order statistics, which in practice might have to be estimated from data.

6. REFERENCES

[1] D. Casasent and N. Kuljanyavivat, “Mine detection from multiple acoustic backscatter data,” *Proc. SPIE Conf. on Detection and Remediation Technologies for Mines and Minelike Targets III*, vol. 3392, pp. 370–381, 1998.

[2] D. H. Kil and F. B. Shin, *Pattern Recognition and Prediction with Applications to Signal Characterization*, AIP Press, Woodbury, New York, 1996.

[3] M. K. Broadhead and L. A. Pflug, “Performance of some sparseness criterion blind deconvolution methods in the presence of noise,” *J. Acoust. Soc. Am.*, vol. 107, no. 2, pp. 885–893, Februrary 2000.

[4] M. R. Gupta, H. S. Anderson, and Y. Chen, “Joint deconvolution and classification for signals with multipath,” *Proc. of the IEEE ICASSP Conf.*, pp. 1069–1072, 2007.

[5] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, Springer-Verlag, New York, 2001.

[6] S. Srivastava, M. R. Gupta, and B. A. Frigyik, “Bayesian quadratic discriminant analysis,” *Journal of Machine Learning Research*, vol. 8, pp. 1277–1305, 2007.

[7] E. Y. Lam and J. W. Goodman, “Iterative statistical approach to blind image deconvolution,” *J. Opt. Soc. Am.*, vol. 17, no. 7, pp. 1177–1184, July 2000.

[8] L. Cros, C. Gervaise, and I. Quidu, “A signal synthesis procedure designed for discreet ocean acoustic tomography,” *Proc. IEEE ICASSP*, vol. 3, pp. 1121–1124, 2007.

[9] C. A. Cabrelli, “Minimum entropy deconvolution and simplicity: A noniterative algorithm,” *Geophysics*, vol. 50, pp. 394–413, 1984.