Improved Type-Based Detection of Analog Signals

Don H. Johnson,[◇] *Paulo A. Gonçalvès,*[□] and *Richard G. Baraniuk*[◇]*

[◇] Department of Electrical and Computer Engineering
Rice University
Houston, TX 77005 USA

[□] INRIA Rocquencourt
Domaine de Voluceau — BP 105
78153 Le Chesnay Cedex, France

Abstract

When applied to continuous-time observations, type-based detection strategies are limited by the necessity to crudely quantize each sample. To alleviate this problem, we smooth the types for both the training and observation data with a linear filter. This post-processing improves detector performance significantly (error probabilities decrease by over a factor of three) without incurring a significant computational penalty. However, this improvement depends on the amplitude distribution and on the quantizer's characteristics.

1 Introduction

Type-based detectors provide totally adaptive detection by obtaining training data, then exploiting those data optimally to determine bit streams. A preliminary application to spread-spectrum communication [1] demonstrated the technique's effectiveness and adaptability. In both training and reception modes, the type-based detector requires received values to be members of a finite set. When continuous-valued signals occur, they must be quantized, and our preliminary results have demonstrated effective, but not optimal (i.e., that provided by likelihood ratio receivers), detection with 3-bit quantization. We describe here a post-quantization technique that allows type-based detectors to achieve performance levels closer to optimal.

2 Type-Based Detection

We consider the classical M-ary detection problem of deciding between M different hypotheses. We assume that for each hypothesis, we have available a training sequence $\left\{X_n^{(m)}\right\}_{n=0}^{N_X-1}$ that was measured when the m^{th} model was true. Given an observation sequence $\{R_n\}_{n=0}^{N_R-1}$ of unknown *a priori* origin, our task is to assign it to one of the M classes with a minimum probability of error.

"Type" is a word used in information theory for the histogram estimate of a discrete probability distribution [2, Chap. 12]. Given a stochastic sequence $\{Y_n\}_{n=0}^{N-1}$, each element of which is drawn from a finite alphabet $\mathcal{Y} = \{y_1, \ldots, y_L\}$, the *type* $\widehat{P}_Y(y)$ of our sequence equals

$$\sum_n I(Y_n = y)/N, \qquad y = y_1, \ldots, y_L. \tag{1}$$

Here $I(\cdot)$ is the indicator function, equaling one when its argument is true and zero otherwise.

A type-based detector has an extremely simple structure. Given the training data $X^{(m)}$ and the observation vector R, we form types of the $X^{(m)}$, of R, and of the *concatenated* sequences $Z^{(m)} = \{X_1^{(m)}, \ldots, X_{N_X}^{(m)}, R_1, \ldots, R_{N_R}\}$. The types of the concatenated sequences can be expressed in terms of the component sequences' types and need not be calculated independently:

$$\widehat{P}_{Z^{(m)}} = \left(N_X \widehat{P}_{X^{(m)}} + N_R \widehat{P}_R\right)/(N_X + N_R). \tag{2}$$

The decision rule calculates the test statistic [3]

$$S_m = \frac{N_X}{N_R} D(\widehat{P}_{X^{(m)}} || \widehat{P}_{Z^{(m)}}) + D(\widehat{P}_R || \widehat{P}_{Z^{(m)}}) \tag{3}$$

for each model m, where $D(P_1||P_2)$ denotes the Kullback-Leibler distance [4]. Qualitatively, the Kullback-Leibler distances will be small if the observations do coincide with the model, and large otherwise. Without going into details, the smaller of the test statistics indicates the decision [1].

This strikingly simple form of hypothesis test leads to an exponential error rate, defined to be $\lim_{N\to\infty} -\log P_e/N$,

*This work was supported by the National Science Foundation, grants MIP–9457438 and NCR–9628236, and the Office of Naval Research, grant no. N00014–95–1–0849.

Email: dhj@rice.edu, paulo.goncalves@inria.fr, richb@rice.edu
Web: http://www-dsp.rice.edu

equal to that of the optimal clairvoyant detector.[1] In this sense, the type-based detector has performance characteristics guaranteed to mimic those of the optimal. Our simulations indicate that many cases exist where the performance characteristics are virtually identical [1].

3 Type-Based Detection for Analog Signals

One important proviso in using type-based detectors is the *finite-alphabet requirement.* In many applications, the training data and observations are analog, with continuous-valued probability density functions (pdfs). Digital processing systems employ finite precision A/D converters, which do produce values drawn from a finite alphabet; however, as the alphabet size grows (more bits in the A/D converter), more training data is needed to provide optimal error rates [3]. Thus, type-based detectors usually employ relatively few quantization levels: about 8–16 in most applications. Such crude quantization, necessitated by training requirements, implies that type-based detector performance cannot equal that of an optimal detector that uses analog or full A/D converter values. At best, the ratio of error probabilities for the type-based detector and the optimal one based on full-precision observations equals a constant; at worst, the error rates do not agree.

When analog observations are quantized with an A/D converter, the type computed from them is really a *histogram estimate* of the pdf of the underlying analog random process. Histogram estimates are very simple, but provide only crude approximations of the true underlying pdfs. Histogram estimates are a special case of *kernel density estimates.* Kernel estimates typically outperform histogram estimates in the sense of integrated-mean-square error [5]. This realization might suggest a pre-processing approach to improved type-based detection of analog data: construct more accurate density estimates of the underlying pdfs — with a kernel smoother, for example — then sample the estimates to yield the types. A related approach would employ wavelet smoothing methods [6] for the density estimate. In either case, obtaining the pdf estimates is computationally expensive and can only be justified if the sampled estimates yield dramatic detector performance improvements.

A simpler option is post-processing. Can we perform processing on the quantized observations to improve performance? Our approach here is to *smooth the type with a digital lowpass filter,* essentially allowing data of surrounding quantization bins to influence the probability estimate in each bin. (See Figures 1 and 2.) While in the same spirit as a kernel density estimate, this approach is not in that class. Our approach quantizes then smoothes, whereas sampling a kernel estimate amounts to smoothing, then quantizing.

[1] A clairvoyant detector somehow knows the exact probabilistic model and uses it to implement the optimal likelihood ratio detector.

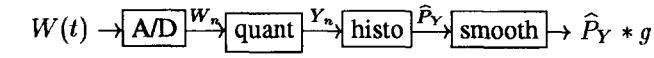


Figure 1. *Computation of a type $\widehat{P}_Y$ for the analog signal $W(t)$ involves conversion to discrete-time, quantization to L levels, and then computation of a histogram over N samples. In a smoothed type, we post-process $\widehat{P}_Y$ with a linear filter g.*

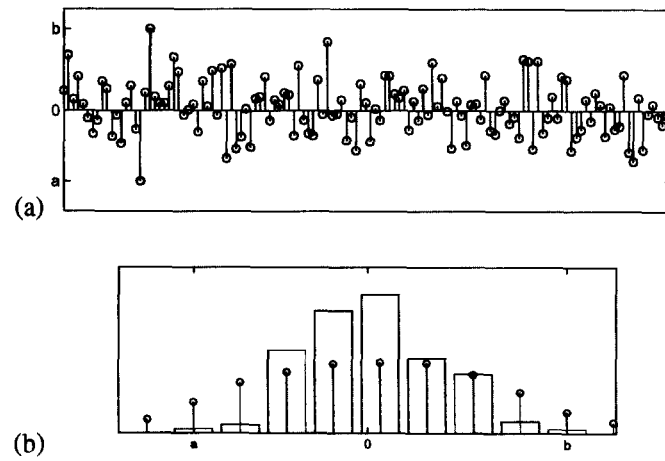


Figure 2. *Types and smoothed types. (a) An IID random sequence Y with Gaussian pdf. (b) Type $\widehat{P}(Y)$ (bars) and smoothed type $\widehat{P}(Y) * g$ (stems) of Y. The smoothed type was obtained by convolving the type with a 3-point boxcar lowpass filter g. Detection performance with the smoothed type is superior to that with the original type.*

In our approach, we automatically set the size of the quantization bins (an automatic gain control) using the "normal reference rule" of Scott [5] for optimal histogram density estimation

$$h^* = 3.5\,\widehat{\sigma}\,N^{-1/3}. \tag{4}$$

Here, $\widehat{\sigma}$ is the sample variance of the N observations. After constructing a type of the data, we regard the type as a discrete-time signal and smooth it through convolution. We have retained A/D converter quantization of only 2 to 6 bits (4–64 equally-spaced levels) and employed sampled Gaussian smoothing functions.

4 Computational Complexity

It is important to note that the smoothed type-based detector is only slightly more expensive (one small convolution) than a simple type-based detector, and therefore it is suited to real-time application at high data rates.

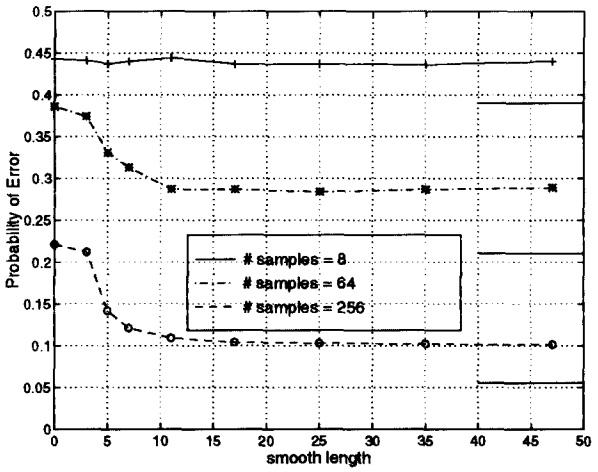


Figure 3. *Performance of a smoothed type-based detector versus the amount of training data. The detection task was to distinguish between two unit-variance Gaussian pdfs shifted in mean by 0.2.*

Vertical axis gives the probability of error P_e based on 1000 independent trials; horizontal axis gives the length of the smoothing window. In each experiment, we chose $N_X = N_R = \#\text{samples}$, which corresponds to the case of minimal training data. The types were smoothed with a sampled Gaussian window. The result for the zero-length window corresponds to that produced by the standard type with no smoothing. The performance of the optimal (matched filter) detector is indicated on the right for each case.

Assuming equal amounts of test and training data $N_X = N_R \equiv N$ and an alphabet of size L, the computational cost (in multiplies and additions) of building each type as in (1) is $O(N)$. The costs of concatenating the types as in (2) and computing the test statistic (3) are both $O(L)$. The cost of smoothing the types is $O(ML)$, with M the length of the smoothing filter. The total cost of the smoothed type-based detector is thus

$$O(N + 2L + ML) = O(N + ML)$$

compared with $O(N + L)$ for the simple type-based detector. Note that typically $M < L \ll N$.

5 Results

To illustrate the performance enhancement attainable by our proposed type post-processing, we consider the classical detection problem of distinguishing between two pdfs identical save for a shift in the mean. From Figure 3, we see that smoothing can improve the probability of error performance of a type-based detector by approximately a factor of 3.

In Figure 4, we demonstrate the effectiveness of smoothing with three different noise distributions: Gaussian, Laplacian, and a mixture composed of 99% Gaussian + 1% Cauchy. In each case, the probability of error decreases with smoothing, up to a maximum of about 3.5 times.

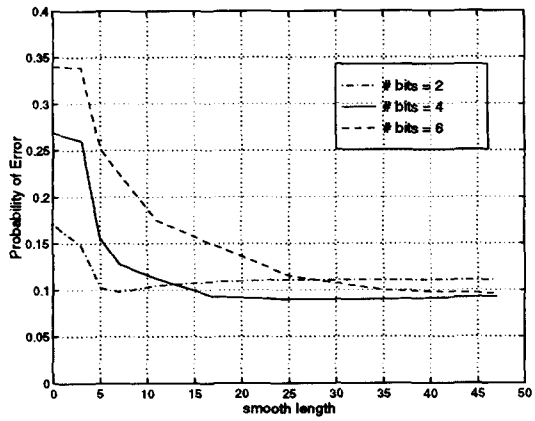
While simple and data-adaptive, the Scott rule (4) for setting the histogram bin widths is suboptimal with respect to the probability of error. In Figure 5, we show the performance of a smoothed type-based detector where the bin widths were chosen using a rule of Kelly [7] that requires *a priori* knowledge of the underlying pdfs. It is interesting to note that even in this case, we can improve the performance of this type-based detector by smoothing.

Caveat: We have found that smoothing does not improve performance in all detection scenarios. For example, smoothing seems to provide no performance gain in the case of two different amplitude distributions having equal means and variances. Fortunately, in communication problems we typically observe smooth pdfs of the same form but with differing means and/or variances. In our experience, some amount of smoothing has always provided a degree of performance gain in such scenarios.

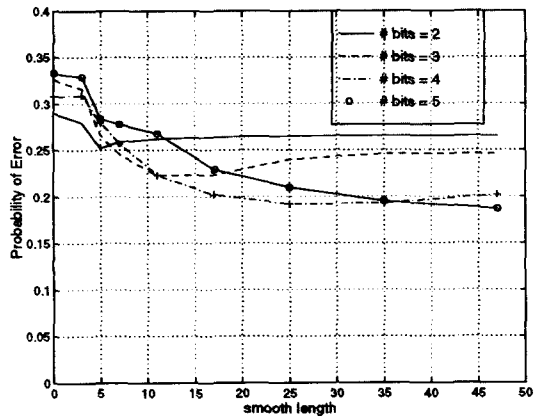
6 Conclusions

We have proposed a type-based adaptive detection algorithm tailored to analog data. The simplicity and robustness of the algorithm makes it suitable for real-time processing in applications. Its key ingredient is a post-processing smoothing filter that is applied to each type. Smoothing improves the performance of the detector in terms of probability of error. Equivalently, for the same error performance, smoothing allows us to reduce the amount of training data significantly. In a communications scenario, as in [1], this approach would allow for shorter preamble sequences.

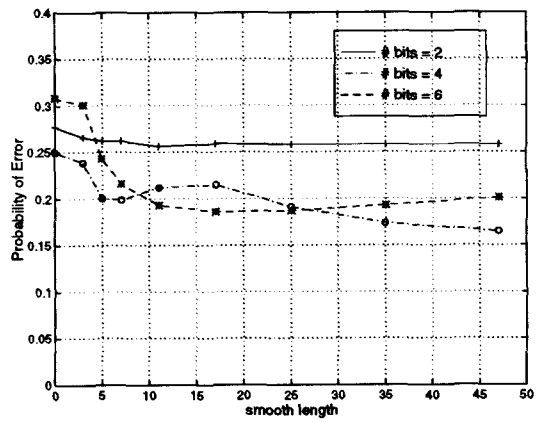
It is somewhat surprising that in many cases we cannot oversmooth the types; that is, we found no performance degradation as the length of the smoothing filter increases without bound. This indicates that the performance of the system will be relatively insensitive to the amount of smoothing applied to the types once a sufficient amount has been reached. This result differs from results in kernel-based density estimation, in which a particular binwidth minimizes mean-squared error. This behavior further emphasizes that improved pdf estimation does not necessarily lead to improved detector performance (in terms of probability of error) and *vice versa*.



(a)



(b)



(c)

Figure 4. *Performance of a smoothed type-based detector versus the length of the smoothing filter for different pdfs. The detection task was to distinguish between two pdfs identical save for a shift in the mean. (a) Gaussian pdf, (b) Laplacian pdf, (c) mixture pdf composed of 99% Gaussian + 1% Cauchy. The different curves correspond to different numbers $L = 2^{\#\text{bits}}$ of quantization levels. Each pdf had unit variance; the mean shift was 0.2 in each case. We used $N_X = N_R = 256$ samples of training and test data for each experiment.*

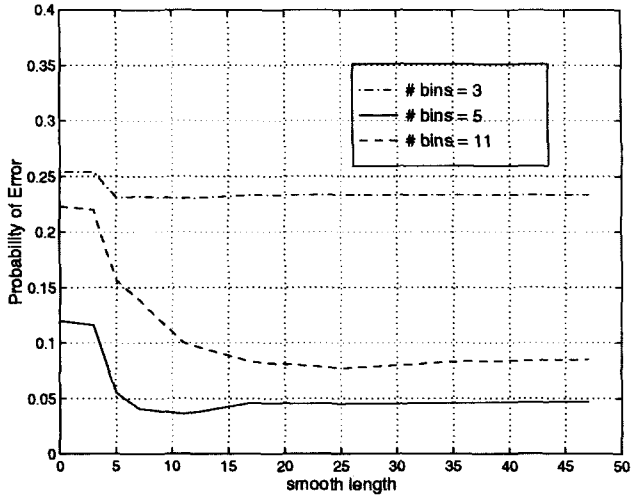


Figure 5. *Smoothed type-based detector performance for the same scenario as in Figure 4(a). Here we use a method of Kelly [7] for setting the bin widths of the histograms. This example demonstrates that the Scott rule (4) for setting the histogram bin widths is suboptimal with respect to the probability of error.*

References

[1] D. Johnson *et al.*, "Type-based detection for unknown channels," in *Proc. ICASSP*, (Atlanta, GA), 1996.

[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, Inc., 1991.

[3] M. Gutman, "Asymptotically optimal classification for multiple tests with empirically observed statistics," *IEEE Trans. Info. Th.*, vol. 35, pp. 401–408, 1989.

[4] S. Kullback, *Information Theory and Statistics*. New York: Wiley, 1959.

[5] D. W. Scott, *Multivariate Density Estimation: Theory, Practice, and Visualization*. New York: Wiley, 1992.

[6] D. L. Donoho, I. M. Johnstone, G. Kerkyacharian, and D. Picard, "Density estimation by wavelet thresholding," Tech. Rep. 426, Statistics, Stanford University, 1995.

[7] O. E. Kelly. Personal communication.