An Evaluation of Algorithms for Single-Echo Biosonar Target Classification

Turgay Temel and John Hallam

Abstract—A recent neuro-spiking coding scheme for feature extraction from biosonar echoes of various plants is examined with a variety of stochastic classifiers. Feature vectors derived are employed in well-known stochastic classifiers, including nearest-neighborhood, single Gaussian and a Gaussian mixture with EM optimization. Classifiers’ performances are evaluated by using cross-validation and bootstrapping techniques. It is shown that the various classifiers perform equivalently and that the modified preprocessing configuration yields considerably improved results.

Keywords—Classification, neuro-spike coding, non-parametric model, parametric model, Gaussian mixture, EM algorithm.

I. INTRODUCTION

When navigating in their natural habitat, the landmarks available to most bats are trees. The echolocation performed by bats is well established in theory, however, it remains a problematic area how to encode the received echoes for landmark recognition. McKerrow showed that echoes generated from continuous-time frequency modulated (CTFM) signals contain relevant cues for classifier design [1]. Frequency components are extracted by demodulating the echo with the same transmitted signal and indicate the distribution of echo scatterers with depth. The outcome is bundled into a database profile, which can be used for classification purposes. Kuc [2] presented a study with biomimetic sonar for classifying various objects where the outcome is bundled into a database profile, which can be used for classification purposes. Kuc [2] presented a study with biomimetic sonar for classifying various objects where the object identity information is embedded in the echo envelope, further, he presented a transformation of plant echoes into pseudo-action potentials where temporal differences are represented by a spatio-temporal field [3]. In [4], Kuc and Müller presented a neuro-spike representation of random echo ensembles and developed a hypothetical neural architecture.

In this study, we examine a more comprehensive yet easy-to-implement neuro-spike code representation of plant echoes as a random process, given by Müller [5]. The feature representation is based on summary statistics of inter-spike time differences at a number of threshold levels.

II. BIOLOGICAL SIGNAL PROCESSING FOR ECHOES, FEATURE EXTRACTION AND PROPOSED PREPROCESSING SCHEME

The echo preprocessing operation consists of two stages: cochlear filtering and coding, [5]. In the first stage, the waveform is passed through a fourth order gammatone filter with center frequency $f_c$ and $-3$ dB quality factor $Q_{-3dB}$ for modeling cochlear filters. This stage is followed by envelope extraction performed as half-wave rectification and low-pass filtering (LPF) with

$$h_{LPF}(t) = \frac{1}{\tau} e^{-t/\tau} \Rightarrow H_{LPF}(z) = \frac{1}{\tau} \frac{1}{1 - e^{-1/(\tau f_c) \Delta^{-1}}} \tag{1}$$

where $f_c$ is the sampling rate.

The LPF output is then normalized and searched for first-crossings of a number of thresholds to model spike generation. Müller [5] showed that successive inter-spike time intervals at threshold levels $\alpha_m$ and $\alpha_{m+1}$ with the condition $\Delta(\alpha_m, \alpha_{m+1}) \geq 1.5 f_c$ constitute sufficient statistics for classifying a given echo source. He defined echo feature vectors as $3$-tuples comprising

$$n = \sum_{\forall m} \mathbb{I}[\Delta(\alpha_m, \alpha_{m+1}) \geq 1.5 f_c]$$

$$\bar{\alpha} = \frac{1}{2n} \sum_{\forall m} (\alpha_m + \alpha_{m+1}) \mathbb{I}[\Delta(\alpha_m, \alpha_{m+1}) \geq 1.5 f_c] \tag{2}$$

$$\bar{\Delta} = \frac{1}{n} \sum_{\forall m} \Delta(\alpha_m, \alpha_{m+1}) \mathbb{I}[\Delta(\alpha_m, \alpha_{m+1}) \geq 1.5 f_c]$$

where $\mathbb{I}$ is an indicator function giving $1$ when the condition is met otherwise $0$.

In this study, we tested two alternative LPF structures. Since the LPF given by (1) introduces varying nonlinear phase response and hence group delay across the frequency range used, features extracted will deviate from the real quantities. A single-pole architecture is also prone to unstable operation i.e., unbounded impulse response. In order to remedy these shortcomings, we propose a discrete-time LPF given by...
The main characteristic of this filter is that it has an almost constant phase response and hence identical group delay. Impulse response will be more bounded compared to previous design.

III. DENSITY ESTIMATION MODELS

In classification, the objective is to decide the class label which best represents the data, \( x \), hence a minimum error probability, \( P_e \) amongst \( M \) different classes, \( C_k \), which results in the Bayesian decision rule [6]

\[
k = \arg\max_i P(C_i \mid x)
\]

(4)

From Bayes’ rule, the above posterior probability can be expressed in terms of likelihood densities, or conditional probability density functions (pdf) \( p(x \mid C_i) \) and \textit{a priori} probabilities \( P(C_i) \) as

\[
P(C_i \mid x) = \frac{P(x \mid C_i)P(C_i)}{\sum_{k=1}^{M} P(x \mid C_k)P(C_k)}
\]

(5)

Therefore, it is essential to develop the appropriate representations of the likelihood functions \( p(x \mid C_i) \). This can be done non-parametrically or parametrically.

A. Non-parametric Methods

Data is assigned to the class containing the closest training sample. In this study, the nearest-neighborhood algorithm [7] is adopted, which can be described as

\[
\hat{k} = \arg\min_i \|x - x_{j \in C_i}\|^2
\]

(6)

where the distance metric is the Mahalanobis normalization which, for column vector data, is given by

\[
\|x - x_{j \in C_i}\|^2 = (x - x_{j \in C_i})^T \Sigma^{-1}(x - x_{j \in C_i})
\]

(7)

The model parameter \( \Sigma \) is the within-class covariance matrix and its estimation is similar to parametric models, to be discussed next.

B. Parametric Methods (1): Single-Gaussian

A single (multi-variate) Gaussian density profile is given by

\[
p(x \mid C_k) = \frac{1}{(2\pi)^{d/2} |\Sigma_k|^{1/2}} \exp\left(-\frac{1}{2}(x - \mu_k)^T \Sigma_k^{-1}(x - \mu_k)\right)
\]

(8)

where \( d \) is the feature vector dimension. Model parameters mean, \( \mu \), and covariance matrix, \( \Sigma \), are estimated by using training set as

\[
\hat{\mu}_k = \frac{1}{N_k} \sum_{i \in C_k} x_{j \in C_k}
\]

\[
\hat{\Sigma}_k = \frac{1}{N_k - 1} \sum_{i \in C_k} (x_{j \in C_k} - \hat{\mu}_k)(x_{j \in C_k} - \hat{\mu}_k)^T
\]

(9)

C. Parametric Methods (2): Gaussian Mixture

Each class conditional pdf is expressed as a linear composition of \( M_k \) component Gaussian pdfs as

\[
p(x \mid C_k) = \sum_{i=1}^{M_k} P(c_i \mid C_k)p(x \mid c_i, C_k)
\]

(10)

with the constraint \( \sum_{i=1}^{M_k} P(c_i \mid C_k) = 1 \). A suitable number of subclasses, \( M_k \), can be determined by using Akaike’s Information Criterion [8], or Rissanen’s Minimum Description Length, [9], etc., an ongoing topic of clustering research.

Parameters can be optimized in the maximum likelihood sense by using the expectation-maximization (EM) algorithm [10] iteratively until the likelihood function reaches a local minimum or a predefined number of iterations have been used. EM description of the \( i \)-th component conditional model parameters at the \( (j+1) \)-th iteration with \( \beta = P(c_i \mid C_k) \), as follows:

\[
\beta_i^{(j+1)} = \frac{1}{N_k} \sum_{i \in C_k} P(c_i \mid x, C_k) = \frac{\sum_{i \in C_k} x_{j \in C_k} P(c_i \mid x, C_k)}{N_k \beta_i^{(j+1)}}
\]

\[
\mu_i^{(j+1)} = \frac{\sum_{i \in C_k} x_{j \in C_k} P(c_i \mid x, C_k)}{N_k \beta_i^{(j+1)}}
\]

\[
\Sigma_i^{(j+1)} = \frac{\sum_{i \in C_k} P(c_i \mid x, C_k)(x_{j \in C_k} - \mu_i^{(j+1)})(x_{j \in C_k} - \mu_i^{(j+1)})^T}{N_k \beta_i^{(j+1)}}
\]

IV. CLASSIFIER DESIGN AND PERFORMANCE MEASUREMENT

In experiments, we employed 2100 echoes for each of four tree types (acer, carpinus, platanus and tilia) from Müller’s database of 85000 echoes. The transmitted signal was a frequency-modulated chirp sweeping linearly from 120 kHz down to 20 kHz in 3 ms. Tree hedges were scanned by two receivers in three dimensions at an almost perpendicular angle and echoes were sampled at \( f_s = 1MHz \).

The filterbank consists of a single-channel with a band-pass filter of parameters \( f_s = 50 \) kHz and \( Q_{\text{filter}} = 10 \). Classifier performances are assessed versus the LPF time constant.
parameter (and the number of chosen components for the mixture models).

Fig. 1 illustrates first order, sample mean, statistics over a randomly chosen 500 echoes for each tree with the proposed discrete LPF characteristic. Reasonable separation of the classes can be seen.

Based on the above feature definitions, nearest neighborhood, single and mixture of Gaussian models were constructed. The mixture Gaussian model is implemented with EM and carried out for (maximum) 1000 iterations. Each classifier’s performance is evaluated with the leave-one-out cross-validation technique [11] with 2100 features by using 10 subgroups for each tree. Those 10 distinct sample average results of estimated correct classification probabilities are processed with boot-strapping to obtain a final classification performance within a given 95% confidence interval. The results are shown in Fig. 2-4 where the average proportion of each class correctly classified is plotted against the LPF time-constant. From the results, it is found that optimum component numbers with the EM mixture model for acer, carpinus, platanus and tilia are 2, 4, 4 and 2 respectively.

For the Gaussian mixture classifier with EM algorithm, the components are initialized by using K-Means algorithm [12]. Component model parameters, such as the mean and covariance matrix can be computed by using maximum likelihood method similar to a single-Gaussian model. The initial component probabilities, \(P(c_j|C_k)\), are given by relative distribution of subclass members.

For comparison purposes, a single-Gaussian classifier is designed with the LPF structure used by Müller [5] and its performance is shown in Fig. 5.

All the classifiers tested show similar (good) classification performance except for low LPF time-constants, and all perform equally. The new LPF characteristic results in an average correct decision proportion better than 90% for \(\tau > 50\) ms with labelled samples. This improves significantly on the classification performance achieved by Müller [5] using the LPF structure of (1) and kernel estimates of the likelihood pdfs computed from the full set of echoes for each tree species in the 85000 echo database.

However, each classifier has pros and cons. For example, nearest neighborhood classifier performance will be subject to the number of training samples, i.e., the larger the number, the longer to train the model and the greater the memory requirements. The single Gaussian model has the problem that the full likelihood pdfs shown in [5] are clearly non-Gaussian. The Gaussian mixture model with EM introduces a convergence problem with small number of training samples. None of these models is straightforward to adapt for on-line class learning.
The results demonstrate that the LPF employed with this study brings out better separation of feature statistics, hence, considerably improved classification performance. The effect can be accounted for by the almost constant phase characteristics of the filter, as explained before, while magnitudes are similar for both structures. It should be noted that, since $\tau \gg 1/f_s$, both filters operate in the asymptotical region and filter constants should be high precision.

V. CONCLUSION

Three probabilistic classifier models are examined for classifying various plant echoes with labelled data in a single-shot mode. For improving the feature first-order statistics’ separation further, a new LPF structure is proposed. Classifier performances are presented and compared to a previous auditory model [5] that employed kernel-estimated likelihood pdfs and the classical single-pole LPF structure of (1).

A variety of parametric and non-parametric classifiers are tested and found to perform adequately, showing the robustness of the chosen feature set.

ACKNOWLEDGEMENT

The authors thank Rolf Müller for his generous sharing of echo data and sample code.

REFERENCES