An Improved Preprocessing for Biosonar Target Classification

Turgay Temel and John Hallam

Abstract—An improved processing description to be employed in biosonar signal processing in a cochlea model is proposed and examined. It is compared to conventional models using a modified discrimination analysis and both are tested. Their performances are evaluated with echo data captured from natural targets (trees). Results indicate that the phase characteristics of low-pass filters employed in the echo processing have a significant effect on class separability for this data.

Keywords—Cochlea model, discriminant analysis, neurospike coding, classification.

I. INTRODUCTION

NEURAL representation of audio signals has been an attractive research field due to mainly lower bandwidth, adaptivity and versatility for the purpose of perception, and recognition. Since mammalian auditory systems share common features, most studies endeavor to model the human auditory system, particularly the inner ear structure called the cochlea.

The passive cochlea model for neural representation of audio signals described by Lyon [1], transforms sound signal into a probability of firing along the auditory nerve. It uses a pre-emphasis filter to simulate the frequency response of the middle and outer ear, a broadly tuned cascade of low-pass filters to model the traveling wave on the cochlea. Nonlinearities and envelope are detected with half-wave rectifiers, low-pass filters and inner hair cells (IHC) as four stages of automatic gain control with different time constants to simulate adaptation and masking.

A different model group consists of gammatone filterbank, described by Patterson et al. [2] and implemented in [3] and Meddis’ IHC model [4]. The filterbank is based on fourth order gammatone filters. Meddis’ IHC model simulates mechanical-to-neural transduction in each filter channel for modeling the transmitter release from hair cells into synaptic cleft.

A parametric sound representation is introduced by Davis et al., [5]: Mel-frequency cepstral coefficients, MFCC. The signal is passed through a mel-spaced filterbank based on FFTs, converted to a logarithmic scale, and then submitted to a cosine transform.

In this study, we present a new low-pass structure with cochlea model and its effect on the class separation of neural-representation statistics point of view for biosonar signals. It will be shown experimentally that discrete-time equivalence should be augmented with a zero term for better spatio-temporal firing separation.

II. AUDITORY SYSTEM MODELING

The general mammalian auditory system is composed of three parts: outer, middle and inner ear (cochlea). While the former two take on directivity and impedance matching functions, the cochlea performs waveform to neural transduction, Fig.1.

The cochlea is modeled as a rolled fluid-filled conical acoustic tube divided longitudinally by the basilar membrane. Wave parameters for the basilar membrane --- stiffness, mass and viscosity etc. --- are location dependent, giving rise to distributed resonant frequencies although mostly they are assumed to be constant for short discrete sections without reflection from a successive section. The solution of waveform of motion equation yields a band-pass filter system, [1]. The IHC with its attached neurons is an electromechanical transduction system that converts the sensed mechanical motion of the basilar membrane into neural firing activity. IHC modeling involves a single-pole low-pass filter structure with time constant $\tau$ ranging from 2.5 ms to 50 ms.

There is also another hair cell group, called outer hair cells...
OHC, that are believed to feedback neural signal from upper level neural system to cochlea for adaptation purposes. Little is known about their detailed function within the mammalian auditory system.

Depending on the modeling assumptions and complexity, different electro-mechanical models have been developed in 2-D [6] and in 3-D [7].

III. ALTERNATIVE SIGNAL PROCESSING SCHEMES AND FEATURE EXTRACTION FOR BIOSONAR ECHOES

The operation on echo consists of two stages: cochlear filtering (preprocessing) and coding, [8]. In the first stage, the waveform is passed through a fourth order gammatone filter with center frequency $f_c$ and -3 dB quality factor $Q_{3dB}$ to model the cochlear filters. This stage is followed by envelope extraction performed as half-wave rectification and low-pass filtering (LPF). The LPF output is then normalized and extracted as a low-pass ARMA filter model can be proposed as

$$\beta = -\frac{2r-1/f_s}{2r+1/f_s} \tag{4}$$

In a group of accompanying papers, low-pass structures (2) and (3) are tested for classification performances with a large range of time constant values. It is observed that a considerable improvement is achieved with the latter form. In the following sections, these alternative topologies will be studied in terms of class statistics separation, which is the main reason for classification performance improvement.

IV. CLASS SEPARATION: MODIFIED LINEAR DISCRIMINANT ANALYSIS

In order to assess the implication of the preprocessing over neural representation, separation of classes’ (feature) statistics is to be investigated. For this purpose, we will employ a modified version of the linear discriminant analysis [10] to give an overall performance measure. In the multivariate case, for class $C_k$, $C_o$, $C = \bigcup_{j \neq k} C_j$ with covariance matrices $\Sigma_k$, $\Sigma_o$, and $\Sigma$, respectively, optimum separation between $C_k$ and $C_o$ can be achieved with following descriptions

$$\theta_k = \max_{\theta} \frac{\theta^T \Sigma_k \theta^T}{\theta^T \Sigma_o \theta^T} \tag{5}$$

where $\theta$ and $\bar{\theta}$ are orthogonal transformation matrices for $C_k$ and $C_o$ with respect to $C$. The solution for above optimizations are given in terms of relative generalized right eigenvectors of $\Sigma_k^{-1} \Sigma$ and $\Sigma_o^{-1} \Sigma$, respectively. Then, a hard performance indicator for the chosen channel over class $C_k$ with parameters $\tau$ and $f_c$ can be defined as

$$I(f_c, \tau) = \begin{cases} 1 & \text{if } \lambda_{\min} > 1/\eta_{\max} \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

where $\lambda_{\min}$ and $\eta_{\max}$ are the minimum and maximum eigenvalues of $\Sigma_k^{-1} \Sigma$ and $\Sigma_o^{-1} \Sigma$, respectively.

V. EXPERIMENTS AND RESULTS

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 $f_s=1\text{MHz}$. The filterbank consists of a single-channel with a band-pass filter of parameters $f_c$ within the range 35 through 75 kHz and $Q_{3\text{db}}=10$. LPF time constant $\tau$ is varied in range 0 through 400ms. Fig. 2 and Fig. 3 illustrate the variation of sample mean of proposed features for each tree with respect to parameter set ($f_c, \tau$) while Fig.4 and 5 illustrate the preprocessing stage class-separation performance of LPFs given by (2) and (3).
From Fig. 5 we see that class separation capability of the single-channel cochlea model with LPF structure (3) is highly improved compared to that given by (2) for the class separation measure (6). The single-echo and sequential classifiers designed with LPF structure (2) are observed to have produced poor classification for the trees platanus and tilia, which effect is also visible in Fig. 4.

Performance improvement with LPF structure (3) 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 $r > 1/\tau$, both filters operate in the asymptotic region and filter constants should be high precision.

VI. CONCLUSION

An alternative low-pass structure is examined and tested with a general cochlea model in a single-channel for a neural representation of natural biosonar echoes. In order to measure the class separation capability offered by the cochlea model, a modified discrimination method is employed. Results indicate that LPF models to be used in cochlea models should take the phase characteristics of the low-pass structure into account, for better separation and hence improved classification.

ACKNOWLEDGEMENT

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

REFERENCES