International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:1, No:1, 2007

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, neuro-spike 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. Non-linearities 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.

Manuscript received November 20, 2004. This work is funded by the EU Framework Programme 5 under the contract IST-2001-35144.

T. Temel is with School of Informatics, IPAB, University of Edinburgh, Kings' Buildings, Edinburgh, EH9 3JZ, United Kingdom, (tel: +44 131 6513437; fax: +44 131 6513435;e-mail: ttemel@inf.ed.ac.uk).

J. Hallam is with the Maersk Mc Kinney Moller Institute for Production Technology, Campusvej 55, DK-5230, Odense M, University of Southern Denmark, Denmark (e-mail: john@mip.sdu.dk).

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.

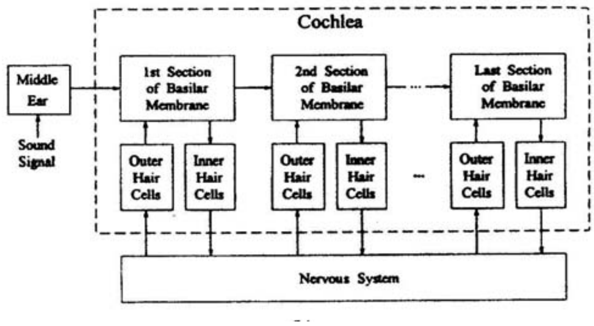

Fig.1. General auditory system model.

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 τ 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 searched for first-crossings of a number of thresholds to model spike generation. Müller [8] showed that successive inter-spike time intervals at threshold levels α_m and α_{m+1} with the condition $\Delta(\alpha_m,\alpha_{m+1})$ constitute a sufficient statistics for classifying a given echo source. He defined echo feature vectors as 3-tuples comprising

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

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

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

where $I_{[.]}$ is an indicator function giving 1 when the condition is met otherwise 0.

A cery common low-pass structure [9] for representing neural activities, and its invariant discrete-time form are given by

$$h_{LPF}(t) = \frac{1}{\tau} e^{\frac{-t}{\tau}} \Rightarrow H_{LPF}(z) = \frac{1}{\tau . f_s}\frac{1}{1 - e^{-1/(\tau . f_s)} z^{-1}} \quad (2)$$

However, in discrete-time realization, phase information should also be considered. It is desirable to have a smooth, constant phase characteristics in order to avoid variable group delays affecting computed firing timings. Moreover, first-order all-pole AR discrete-time systems tend to operate near oscillation for given $\tau > f_s$. In order to avoid these shortcomings, a low-pass ARMA filter model can be proposed as

$$H_{LPF}(z) = \frac{1}{\tau . f_s}\frac{1 + \alpha z^{-1}}{1 + \beta z^{-1}} \quad (3)$$

where parameters $-1<\alpha,\beta<1$ to be determined for proper low-pass magnitude and phase responses. From $z = e^{j\omega}$, stability and smooth phase response requirements, it can be shown that $-1<\beta<0$, $0<\alpha<1$ and $|\alpha| \approx |\beta| \approx 1$, respectively. By using the bilinear transform and eliminating discontinuous generalized functions from the LPF in (2), it can be shown that

$$\beta = -\frac{2\tau - 1/f_s}{2\tau + 1/f_s} \quad (4)$$

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, $C_o = \bigcup_{j\neq k} C_j$, $C = \bigcup_{\forall j} C_j$ with covariance matrices Σ_k, Σ_o, and Σ, respectively, optimum separation between C_k and C_o can be achieved with following descriptions

$$\theta_k = \max_{\theta}\frac{\left|\theta \Sigma_k \theta^T\right|}{\left|\theta \Sigma \theta^T\right|}, \quad \vartheta_k = \max_{\vartheta}\frac{\left|\vartheta \Sigma \vartheta^T\right|}{\left|\vartheta \Sigma_o \vartheta^T\right|} \quad (5)$$

where θ and ϑ 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 $\Sigma\Sigma_o^{-1}$, respectively. Then, a hard performance indicator for the chosen channel over class C_k with parameters τ and f_c can be defined as

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

where $\lambda_{\min}$ and $\eta_{\max}$ are the minimum and maximum eigenvalues of $\Sigma_k^{-1}\Sigma$ and $\Sigma\Sigma_o^{-1}$, 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=1MHz. 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_{-3dB}=10. LPF time constant τ 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,τ) while Fig.4 and 5 illustrate the preprocessing stage class-separation performance of LPFs given by (2) and (3).

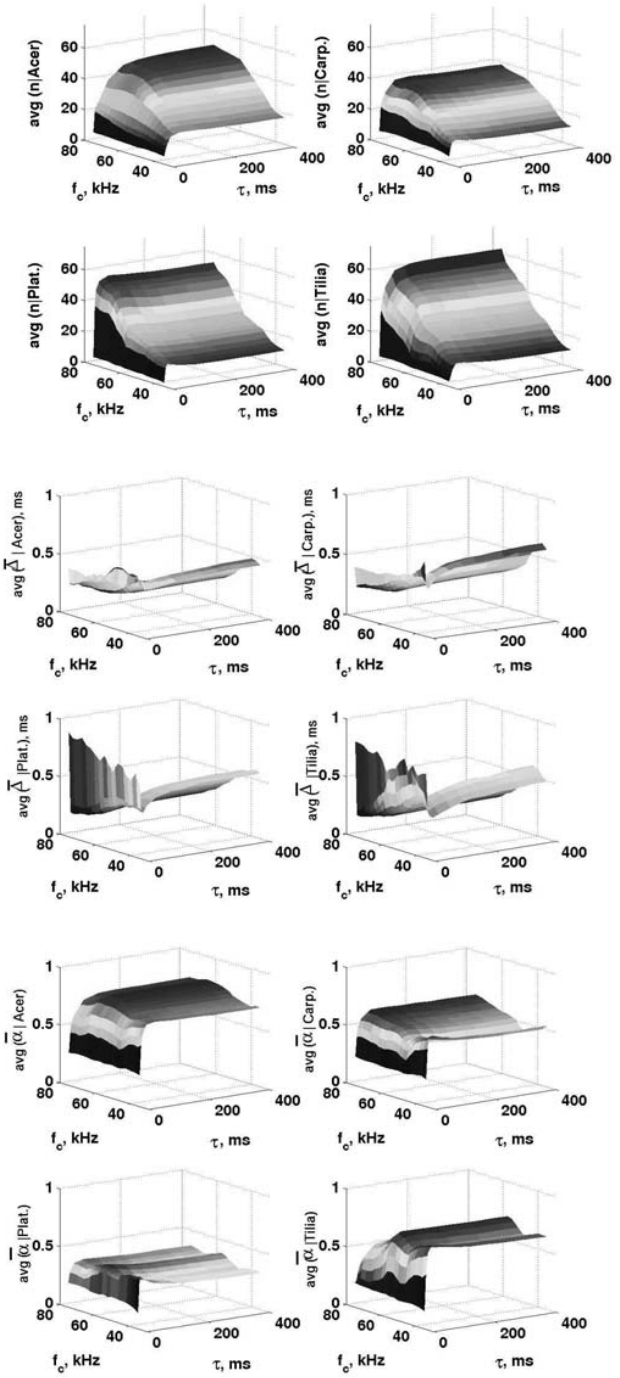

Fig.2. Sample-mean of features with LPF structure (2).

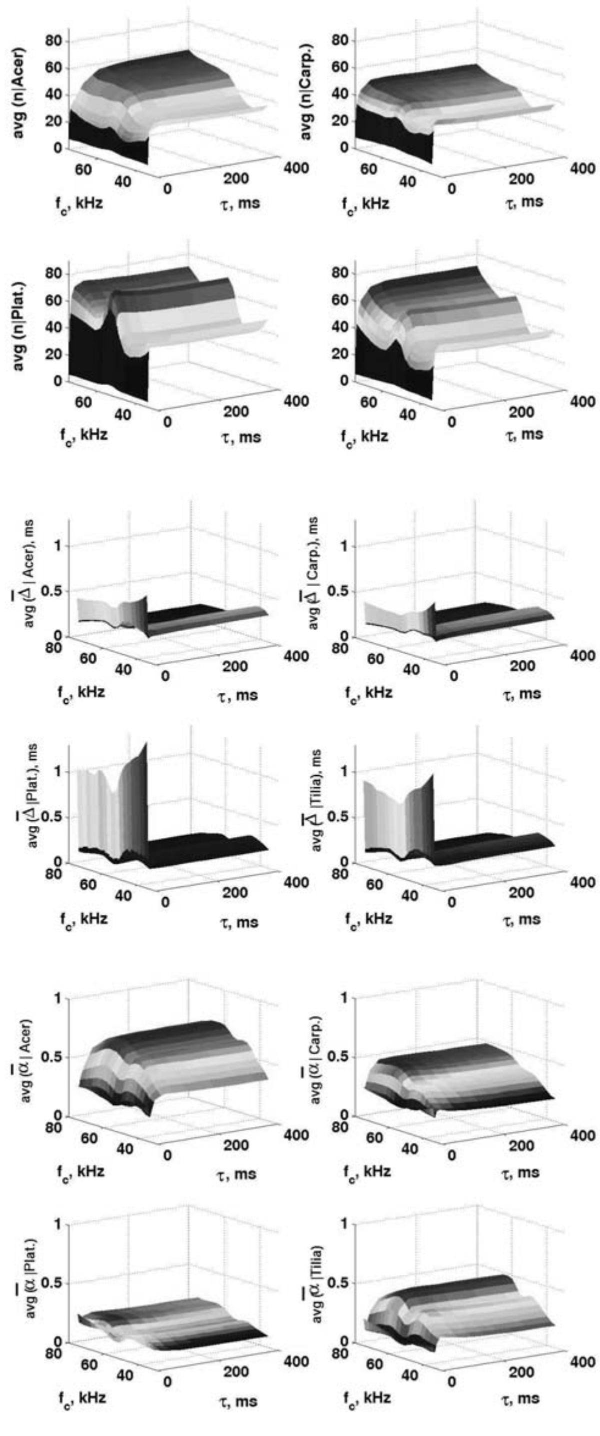

Fig.3. Sample-mean of features with LPF structure (3).

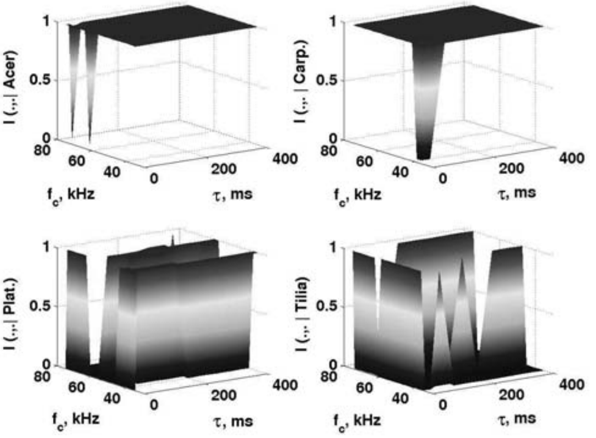

Fig. 4. Class separation performance with LPF structure (2).

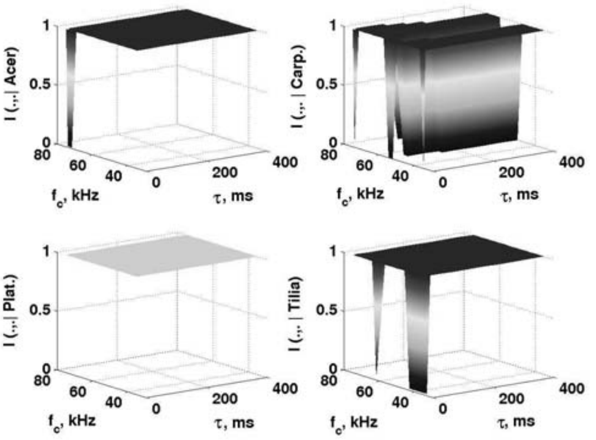

Fig. 5. Class separation performance with LPF structure (3).

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 $\tau >> 1/f_s$, 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

[1] R. F. Lyon, "A computational model of filtering, detection, and compression in cochlea," in *Proc. IEEE Intl. Conf. On Acoust.Speech, Signal Processing*, pp. 1282-1285, Paris, France, May, 1982.

[2] R.D. Patterson, and J. Holdsworth, "A functional model of neural activity patterns and auditory images," William A. Ainsworth (ed.), vol. 3, *Advances in Speech, Hearing and Language Processing*, JAI Press, London, 1996.

[3] M. Slaney, "An efficient implementation of the Patterson-Holdsworth auditory filterbank," Apple Computer Tech. Rep. 35, Apple Computer, 1988.

[4] R. Meddis, M. J. Hewitt, and T. M. Shackleton, "Implementation details of a computation model of the inner-hair cell/auditory –nerve synapse," *J. Acoust. Soc. Am.*, vol. 87, no. 4, pp. 1813-1816, Oct. 1990.

[5] S. B. Davis, and P. Mermelstein, "Comparison of parametric representations for monosyllabic word recognition in continuous spoken sentences," *IEEE Trans.Accoust. Speech and Signal Processing*,vol. 28, no. 4, pp. 357-366, 1980.

[6] R. F. Lyon, and C. Mead, "An analog electronic cochlea," *IEEE Trans. Accoust. Speech and Signal Processing*,vol. 36, no. 7, pp. 1119-1134, July 1988.

[7] E. Givelberg, and J. Brunn, "A comprehensive three-dimensional model of the cochlea," J. Comp. Physics, vol. 191, no. 2, pp. 377-391, Nov. 2003.

[8] R. Müller, "A computational theory for the classification of natural biosonar targets based on a spike code," *Network: Comput. Neural Syst.*, vol. 14, pp. 595-612, May 2003.

[9] J. M. Kates, "A time-domain digital cochlear model," *IEEE Trans. Signal Proc.*, vol. 39, no. 12, pp. 2573-2591, Dec. 1991.

[10] M. S. Srivastava, and E.M. Carter, *Applied Multivariate Statistics*, North Holland Amsterdam, 1983.