(U)
SECURITY CLASSIFICATION OF THIS PAGE

AD-A226 828

REPORT DOCUMENTATION PAGE

DTIC FILE COPY

1a.	1b. RESTRICTIVE MARKINGS
2a.	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE N/A	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 90 0964
6a. NAME OF PERFORMING ORGANIZATION University of Maryland / 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION AFOSR/NL
6c. ADDRESS (City, State and ZIP Code) College Park, MD 20742	7b. ADDRESS (City, State and ZIP Code) Bldg. 410 Bolling AFB, Washington, D.C. 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR / 8b. OFFICE SYMBOL (If applicable) NL	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR-88-0204
8c. ADDRESS (City, State and ZIP Code) Bldg. 410 Bolling AFB, Washington, D.C. 20332-6448	10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
61102F	2313	A8	

11. TITLE (Include Security Classification) **Signal Processing & Recognition in Adaptive Neural Networks**

12. PERSONAL AUTHOR(S)
Professors Shihab Shamma & P.S. Krishnaprasad

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Annual	FROM 08-01-89 TO 7-31-90	90 Aug 09	5

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	This research

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The research conducted under the AFOSR grant (AFOSR-88-0204) can be subdivided into four areas: (1) Models and neurophysiology of the auditory cortex: this includes mappings of physiological responses to sound, psychoacoustical studies, and mathematical models of the data. (2) Implementations of the cochlear and other auditory models both in DSP and VLSI forms. (3) Unsupervised learning algorithms applied to problems in sound segmentation, timbre characterization, and pitch extraction. (4) Applications of wavelet transforms to the analysis of neural networks.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS ☐	

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE NUMBER (Include Area Code)	22c. OFFICE SYMBOL
JOHN F TANGNEY	(202) 767-5021	NL

DD FORM 1473, 83 APR — EDITION OF 1 JAN 73 IS OBSOLETE.

(U)
SECURITY CLASSIFICATION OF THIS PAGE

90 09 25 038

Annual Technical Report (Second Year)

Shihab Shamma and P.S. Krishnaprasad

Summary

The research conducted under the AFOSR grant (AFOSR-88-0204) can be subdivided into four areas: (1) Models and neurophysiology of the auditory cortex: this includes mappings of physiological responses to sound, psychoacoustical studies, and mathematical models of the data. (2) Implementations of the cochlear and other auditory models both in DSP and VLSI forms. (3) Unsupervised learning algorithms applied to problems in sound segmentation, timbre characterization, and pitch extraction. (4) Applications of wavelet transforms to the analysis of neural networks.

In the following sections, we briefly discuss the progress accomplished over the last year, and the status of the work at present. Other supporting material (e.g., technical reports and papers) are appended at the end.

Models and neurophysiology of the auditory cortex

There are several ongoing projects in this area which range from neurophysiological mappings of the responses of the various areas of the auditory cortex, to the formulation of mathematical models of these data, to psychoacoustical testing of the functional hypotheses.

Neurophysiological mappings of the auditory cortex

We have been carrying out a systematic investigation into the functional organization of the various auditory cortical areas. Our experiments have focused on the determination of the receptive fields in the primary auditory cortex and in the anterior field. We have sought to determine if any systematic changes occur across the surface of the cortex, and if so, what functional significance such changes might imply. The initial phases of the work therefore focused on developing the experimental paradigms, and in carrying out an extensive series of experiments in which the auditory cortex was mapped. The the paradigms used and the results obtained can be summarized as follows:

1. We studied the topographic organization of the receptive fields obtained from single and multiunit recordings along the isofrequency planes of the primary auditory cortex (AI) in the barbiturate-anesthetized ferret.

2. Using a two-tone stimulus and bandlimited noise, the excitatory and inhibitory portions of the receptive field of each cell were determined and then parameterized in terms of a *symmetry index*. The index measures the net balance of excitatory and inhibitory influences around the best frequency (BF) of the cell.

3. The distribution of the symmetry index values along the isofrequency planes revealed systematic changes in the symmetry of the receptive fields. At the center, units with narrow and symmetric inhibitory sidebands predominated. These gave way gradually to asymmetric inhibition, with high frequencies (relative to the units' BFs) becoming more effective caudally, and low frequencies more effective rostrally. These response types tend to organize along one or more bands that parallel the tonotopic axis (i.e., orthogonal to the isofrequency planes).

4. Cell responses to spectrally shaped noise were consistent with the symmetry of their receptive fields. For instance, cells with strong inhibition from above the BF were most responsive

to stimuli that contain least spectral energy above the BF, i.e., stimuli with the opposite asymmetry. In six animals, it was demonstrated that the spectral gradient (or "orientation") of the most effective stimulus varies systematically along the isofrequency planes.

5. In five animals, the selectivity of unit responses to the direction of a frequency-modulated (FM) tone was tested and found to correlate strongly with the symmetry index of the receptive fields. Specifically, cells with strong inhibition from frequencies above (below) the BF prefer increasing (decreasing) sweeps. Thus, selectivity to FM direction is also mapped along the isofrequency planes of the AI.

6. One functional implication of the receptive field organization is that cortical responses encode the locally averaged *gradient* of the acoustic spectrum by their differential distribution along the isofrequency planes. This enhances the representation of such features as the symmetry (or "orientation") of spectral peaks, edges, and the spectral envelope. This scheme, together with FM selectivity, can be viewed as the one dimensional analogue of the orientation columns and direction of motion selectivity of the visual cortex.

Mathematical models of the auditory cortex

With these experimental data in hand, we then developed mathematical models of the receptive fields and analyzed the nature of the responses and potential features encoded by the cortex. The model approximates the systematic changes in the excitatory and inhibitory portions of the receptive fields along the isofrequency planes by a difference of gaussians function with spatially changing parameters. We considered intially only the response properties to *stationary* stimuli, i.e. those with non-varying spectra. The fundamental functional principle that emerged from the analysis of the model was that the primary auditory cortex encoded the shape of the acoustic spectrum in the distribution of its responses along the isofrequency planes. Specifically, it maps to each isofrequency plane a normalized measure of the locally averaged gradient of the input spectrum at that frequency. Physiological and psychoacoustical correlates and implications of these findings were elaborated, and parallels to the functional organization of the visual cortex were also considered.

Psychoacoustical studies of spectral shape discrimination

The experimental results and mathematical models discussed above suggested that specific features of the shape of the acoustuic spectrum are being extracted and mapped in the cortex. If so, then it is likely that important consequences must exist regarding the perception of such spectra. Very little direct evaluation of such features as the sensitivity of subjects to the symmetru of spectral peaks and local gradients exist. So we have developed experimental set-ups and paradigms with the help of Dr. David Green to carry out such experiments. We are now ready to start collecting the data, and we should be able to report the results in the next report.

Finally, these experimental results and mathematical models were submitted for publication early in this year, and copies of the SRC technical reports (manuscripts) of the two papers are enclosed with this progress report.

Implementations of the cochlear and other auditory models

In order for cochlear and other models of auditory processing to become useful computational structures, their computational cost has to be reduced drastically. This is important

not only for real-time applications, but also to facilitate experimentation with such models. At present, the fastest implementations of the cochlear models are about a 1000 times slower than real-time. Continued improvements in computational technology is expected to close this gap, but not sufficiently or rapidly enough. The two obvious alternative options are a radical reformulation of the algorithms to minimize their complexity without sacrificing the associated advantages, and the fabrication of specialized hardware.

The first option is only plausible once a clearer understanding of cochlear function is achieved. This is a goal of our work. Several groups have been pursuing the second option, that of fabricating specialized hardware to perform the cochlear operations. These efforts fall into two different approaches: The first aim to map the cochlear algorithms into digital algorithms that are then implemented using standard DSP chips. The second aim to fabricate various digital, analog, or mixed custom integrated circuits. In the first category, only moderate success has been achieved. For instance, we attempted two years ago to implement our cochlear model on the ODYSSEY Board, a TI product which contains four TMS32025 DSP Chips and associated circuitry, and sits on the backplane of the TI Explorer . Three processors worked in parallel to compute the cochlear outputs from 128 channels, while one coordinated all actions and memory transfers. The implementation worked well except for the massive I/O bottlenecks created by the need to pass data back and forth to the computer memory. This slow and unavoidable process offset any advantages gained from speeding up the computations. Overcoming this problem requires specialized parallel interfaces and software which we do not as yet possess.

Fabricating specialized integrated circuits has proven so far equally difficult. Several fundamentally different approaches have been tried, ranging from the sub-threshold analog circuits to digital filter banks Sub-threshold CMOS analog implementations have proven difficult because of the poor control over the gate thresholds, and hence the scatter in the center frequencies of the cochlear filters. Integrated digital filter banks are quite feasible, although the steep roll-offs of the cochlear filters and other time-domain cochlear stages have made the circuits excessively complex and unstable despite many simplifications.

An alternate approach to the fabrication of cochlear models is the direct implementation of its delay line structure without resort to equivalent frequency-domain filters. That is, to fabricate literally the equivalent circuit of the cochlea. A significant advantage of this approach is that any additional nonlinear processes can be incorporated directly and easily. One approach to implementing this circuit is through purely passive integrated components. We designed and simulated such a structure using capacitors, inductors and resistors. Active components were only used in the input and output stages. In order to minimize the component values, sound signals were translated to high frequencies (10 MHz), and total audio frequency range stretched to 2 MHz. This approach eventually proved unworkable because of the parasitic resistances that severely reduced the effective Q-values of the filters. The alternative technology that we settled on is the switched capacitor filter. This was arrived at after an exhaustive look at passive, active RC, digital, and continuous time filter approaches. Each of these had its advantages and disadvantages, and SCF's proved the best compromise. Our current strategy is to implement a direct analog of the tapped-delay-line of the cochlea using SCF's to replace the inductors. The overall structure will be fabricated in buffered segments of 10 stages each so as to minimize the absolute number of cascaded stages to ensure stability. Simulations of the necessary circuits

are underway, and layout of the masks will commence within a month.

In a different application, a network developed over the last two years under this research grant, called the *stereausis network*, was recently implemented both in hardware and fast software for real-time applications in speech recognition and pitch extraction. The hardware implementation came in the form of sub-threshold analog chip fabricated in Caltech by Dr. Carver Mead. The Software implementation on the Cray was carried out by Dr. Malcom Slany at Apple. Future experimentation with these implementations will be coordinated with the two above mentioned gruops.

Unsupervised learning algorithms

The goal of this project is to develop *unsupervised* learning algorithms to configure networks able to segment the sound stream based on either pitch or timber cues. The motivation behind this work is two-fold: (1) From a theoretical standpoint, such networks can explain how such functions emerge spontaneously in the auditory nervous system; and (2) they can in turn be applied as pitch or phoneme detector devices. Being unsupervised, the network connectivities reflect the structure of the input set, and not that of an imposed external "teacher" as is, for instance, the case in the backpropogation learning algorithm.

Such an adaptive pattern classifying algorithm was recently developed by T. Teolis (MS Thesis). The algorithm is based on a choice of a similarity function which is restricted to operate only on the network outputs. The learning rule is a gradient decent that aims to minimize an appropriate energy function with respect to an initial point that reflects the statistics of the input training set. The algorithm is computationally efficient and requires little storage capacity since it incorporates first order estimators of the input statistics rather than the actual input pattern sequence to compute its updates. The adaptive network has been demonstrated successfully as a segmentation algorithm of natural speech and of a sequence of musical notes from various instruments. The algorithm is currently being expanded to configure recursive and multilayer networks.

Application of wavelet transforms to the analysis of neural networks

Wavelet transforms have recently emerged as a means of decomposing functions in a manner which readily reveals local frequency content. Such time/space-frequency localization is beneficial in applications such as image analysis and speech processing. Affine wavelet transforms are transforms in which the "basis" functions used in the transformation are generated via translations and dilations of a single function. Discrete wavelet transforms in general are based upon the concept of *frames* in Hilbert spaces. *Frames* are essentially generalizations of orhonormal bases which do not require that the "basis" elements be orthogonal, normalized, or even form a true basis. Discrete affine wavelet transforms rely on the construction of frames which consist of elements generated via the action of a representation of the affine group (translations and dilations) on a single function.

Over the past year we have been studying the relationship of feedforward neural networks to discrete affine wavelet transforms. We noted that the operations of translation and dilation already exist in standard feedforward neural networks. For instance for a single hidden layer network, weights from the input layer to the hidden layer provide dilations and bias weights on the neurons of the hidden layer provide translations. Based upon this observation, we have shown that it is possible to construct a wavelet "basis" (frame) for $L^2(R^1)$ by appropriate

combinations of sigmoidal functions. The following are some of the benefits which can be derived from such a construction.

1. It is clear that any function in $L^2(R^1)$ can be learned by a feedforward neural network. This follows from the properties of frames.

2. Time-frequency localization properties of the frame elements help to provide a synthesis algorithm for feedforward networks which is based upon analysis of the training data. We have shown that the number of required hidden layer nodes, the weights from the input layer to the hidden layer, and the bias weights can all be determined based upon straightforward analysis of the training data.

3. Due to the synthesis algorithm mentioned above, the training problem is greatly simplified. The network can be trained via minimization of a *convex* function. Hence there will be global minima only. We have also shown that the training can be accomplished via the solution of a system of linear equations.

We are currently working on extensions of these constructions to higher dimensions ($L^2(R^n)$) and examining applications of the synthesis procedure which we have developed.

Publications

S. Shamma, J. Fleshman, and P. Wiser "Receptive Field Organization in the Primary Auditory Cortex", submitted to *J. Neurophys*, Mar. 1990.

S. Shamma and G. Chettiar, "Functional Model of the Primary Auditory Cortex", submitted to *J. Acoust. Soc. Am.*, Mar. 1990.

S. Shamma, N. Shen, and P. Gopalaswamy, "Stereausis: Binaural processing without neural delays", *J. Acoust. Soc. Am.*, 86: 989–1006, 1989.

T. Tiolis "Adaptive Pattern Recognition" to be submitted to the *J. Neural Networks*, 1990. Currently an SRC Tech. Report and MS Thesis. EE Department, 1989.

Personnel

The students (with degrees awarded) and research personnel involved in the research reported in this document are: Shihab Shamma (Assoc. Prof.), P. S. Krishnaprasad (Prof.), Y. Pati (MS), Geeth Chettiar (MS), James Fleshman (Res. Faculty, now at NIH), Philip Wiser (BS in Dec. 1990), Danniel Lin (MS), Tony Tiolis (MS), Preetham Gopalaswamy (MS), M. Niaming (MS), Svetlana Vranic (MS).