AFOSR-TR-95

0658

REPORT DOCUMENTATION PAGE

1.AGENCY USE ONLY(leave blank)	2.REPORT DATE	3.REPORT TYPE AND DATES COVERED Sept. 1992 Final Progress Report 3'Aug. 1995

DTIC
SELECTED
OCT 1 8 1995
F

4. TITLE AND SUBTITLE Theoretical & Experimental Sutdies of Auditory Processing	5. FUNDING NUMBERS 61102F 2313 -AS
6. AUTHOR(S) S. Shamma & P.S. Krishnaprasad	

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Maryland Dept of Electrical Engineering College Park, MD 20742	8. PERFORMING ORGANIZATION REPORT NUMBER MD 910318-7388-360201

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) AFOSR/NL 110 Duncan Ave, Suite 0115 Bolling AFB, DC 20332-8080	10. SPONSORING/MONITORING AGENCY REPORT NUMBER F49620-92-J-0500

11. SUPPLEMENTARY NOTES
Papers & Manuscripts already submitted independently to Dr. Tangney

12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited.	12b. DISTRIBUTION CODE

13. ABSTRACT (maximum 200 words)

The research reported here has been conducted over the last three years under the AFOSR grant (F49620-92-J-0500). It is divided into four general categories of projects: (1) VLSI implementations of the early auditory stages. (2) Functional organization of the auditory cortex: Neurophysiology (3) Functional models of the auditory system: Psychoacoustics. (4) Analysis of neural network architectures with wavelet transforms. We shall review briefly the main results achieved in these four areas.

19951017 025

14. SUBJECT TERMS			15. NUMBER OF PAGES 19
			16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Not Classified	18. SECURITY CLASSIFICATION OF THIS PAGE U.	19. SECURITY CLASSIFICATION OF ABSTRACT U.	20. LIMITATION OF ABSTRACT
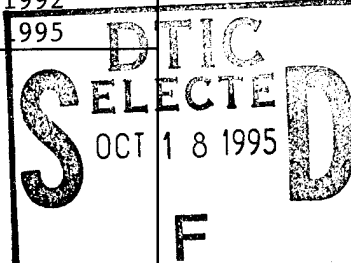
DTIC QUALITY INSPECTED 5

2 0 SEP 1995

SYSTEMS RESEARCH CENTER
and
DEPARTMENT of ELECTRICAL ENGINEERING
University of Maryland
College Park, MD 20742

Research Proposal

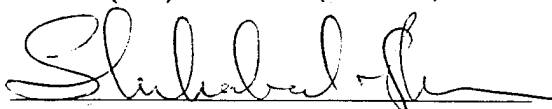## Theoretical and Experimental Studies of Auditory Processing

Prepared by

Shihab A. Shamma and P.S. Krishnaprasad

Theoretical and Experimental Studies of Auditory Processing

Shihab A. Shamma and P.S. Krishnaprasad

This proposal describes the theoretical and experimental research into the principles of sound processing in the auditory system. The proposal is divided into two parts. The first is a review of the work we accomplished in the last three years under the AFOSR grant (F49620-92-J-0500). The second part outlines our proposed research plans for the next three years.

Part I: Review of Research Results

Summary

The research reported here has been conducted over the last three years under the AFOSR grant (F49620-92-J-0500). It is divided into four general categories of projects: (1) VLSI implementations of the early auditory stages. (2) Functional organization of the auditory cortex: Neurophysiology (3) Functional models of the auditory system: Psychoacoustics. (4) Analysis of neural network architectures with wavelet transforms. We shall review briefly the main results achieved in these four areas.

Besides the two P.I.s' salaries, the grant supported several Ph.D. and M.S. students, a laboratory manager, and partially a post-doctoral fellow (see list of names and degrees at the end of this review).
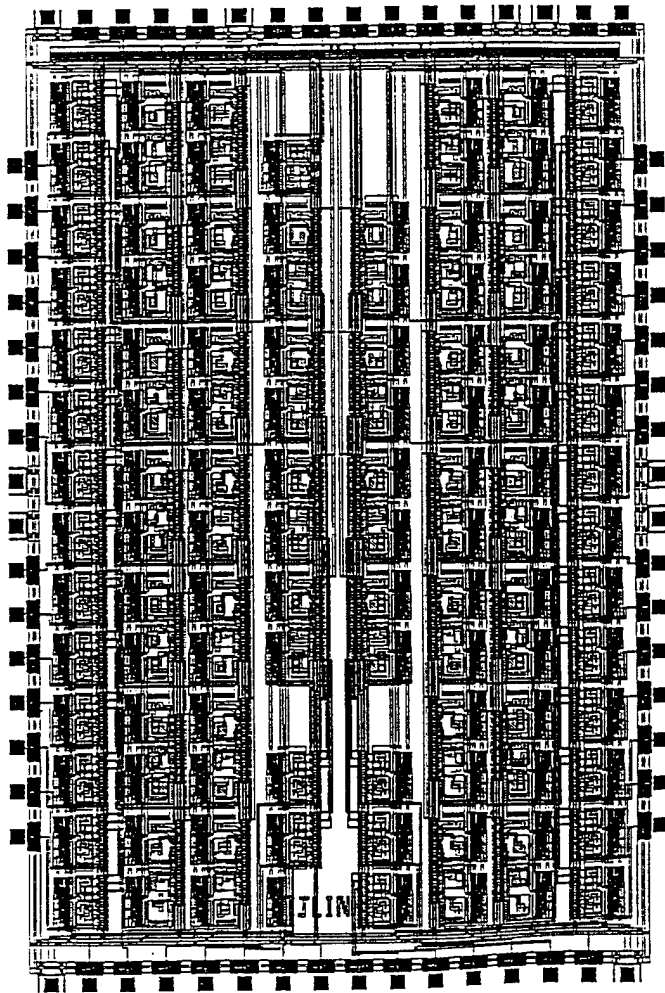
I. VLSI Design and Implementations of Early Auditory Processing

Over the last few years, we have been developing and analyzing detailed models of the early auditory stages. Our goal is to understand the underlying signal processing principles that endow such systems with their noise robustness and feature enhancement abilities. In order for these models to become useful components in such applications as automatic speech recognition, their computational cost has to be reduced drastically. One approach to accomplish this is to implement the algorithms in VLSI using (S)witched (C)apacitor (F)ilters (SCF) because they provide several advantages[10, 11, 15, 16].
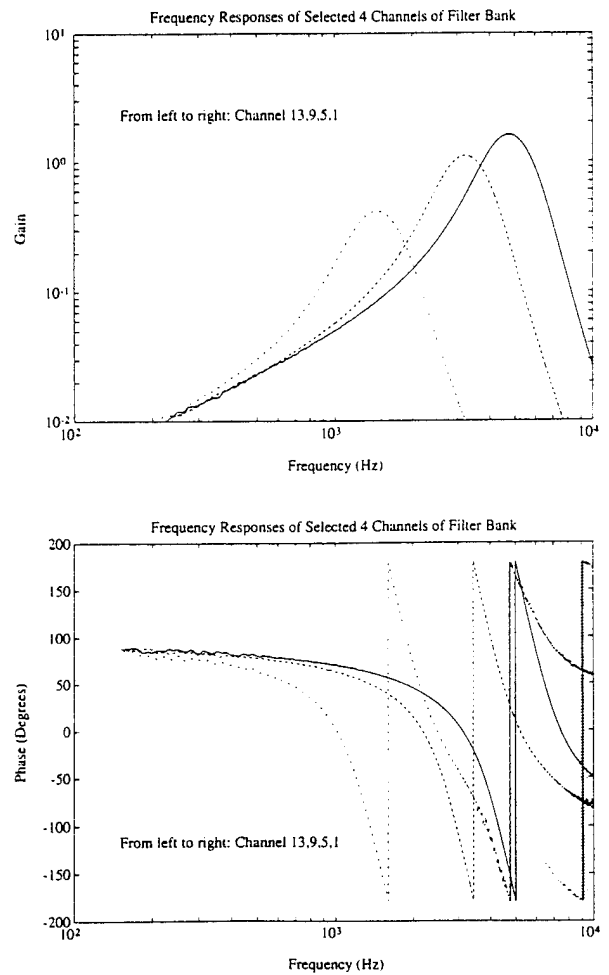
- Filter pole positions are determined not by the RC products, but by capacitor ratios.

- Capacitor ratios can be precisely controlled and are stable with temperature. Furthermore, accurate filter transfer functions can be implemented in a completely monolithic form.

- SCF's require very little silicon area to implement high-value resistors.

Over the last year, we have succeeded in developing such a system. Many obstacles were solved along the way, including the design of area efficient integrators working at the relatively low acoustic frequencies, and offset and parasatic insensitive Op-Amps for the channel adders (Fig.1). Details of these and many other design innovations are available in the paper by Lin et al. and J. Lin's Ph.D. thesis accompanying this review. VLSI chips of up to 32 channels/chip that can be combined in parallel to form a 64 channel system have been fabricated using MOSIS. A patent of the SCF design innovations was also granted (see Fig.2 for details).

(a)

(b)

Frequency Responses of Selected 4 Channels of Filter Bank

From left to right: Channel 13,9,5,1

Gain

Frequency (Hz)

Frequency Responses of Selected 4 Channels of Filter Bank

Phase (Degrees)

From left to right: Channel 13,9,5,1

Frequency (Hz)

(c)

Spatial-Temporal Responses of The Entire 32-Channel Filter Bank
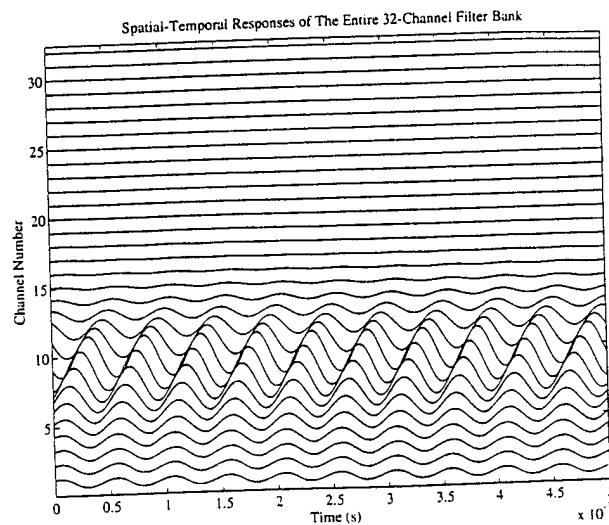
Channel Number

Time (s)

Figure 1: (a) IC-Chip layout of the 32-channel filter bank fabricated. (b) Transfer functions of selected 4 channels of the bank. (c) Spatiotemporal responses of the entire 32-channel filter bank to a single tone.

4

[54] **COCHLEAR FILTER BANK WITH SWITCHED-CAPACITOR CIRCUITS**

[75] Inventors: Jyhfong Lin, Hyattsville, Md.; Shihab A. Shamma, Washington, D.C.; Thomas G. Edwards, College Park, Md.

[73] Assignee: University of Maryland, College Park, Md.

[21] Appl. No.: 53,539

[22] Filed: Apr. 29, 1993

[51] Int. Cl.⁵ H03B 1/04; H04B 1/10
[52] U.S. Cl. 307/520; 307/490; 328/104; 328/127; 328/167
[58] Field of Search 307/490, 521, 520; 328/167, 165, 127, 104, 16, 17, 62

[56] **References Cited**

U.S. PATENT DOCUMENTS

3,639,842	4/1972	Zacone et al.	328/104
4,446,438	5/1984	Chang et al.	328/127
4,543,534	9/1985	Temes et al.	330/9
4,626,808	12/1986	Nossek	307/521
4,894,620	1/1990	Nagaraj	328/127
5,168,179	12/1992	Negahban-Hagh	328/167
5,182,521	1/1993	Chang et al.	328/167

OTHER PUBLICATIONS

Lin et al. "Realization of Cochlear Filters By VLT Switched-Capacitor Biquads", Mar. 1992, pp. 1-4, IEEE Intn'l Conf. on Acoustics.

Primary Examiner—William L. Sikes
Assistant Examiner—Dinh Le
Attorney, Agent, or Firm—Christopher N. Sears

[57] **ABSTRACT**

A parallel dilating-filters switched-capacitor filter bank is described in simulation of the cochlea. Area-saving is achieved by filter-sharing, effective sum-gain amplifier designs, and using area efficient nth-order filter designs, and in particular using a biquadratic filter design using charge-differencing. The structure is easily expandable to include more channels by extending with additional filters and output amplifiers, or by using several chips with different sampling frequencies in parallel connection. An offset-compensated area-efficient switched-capacitor sum-gain amplifier circuit design is described and can be used in the filter bank.

12 Claims, 6 Drawing Sheets

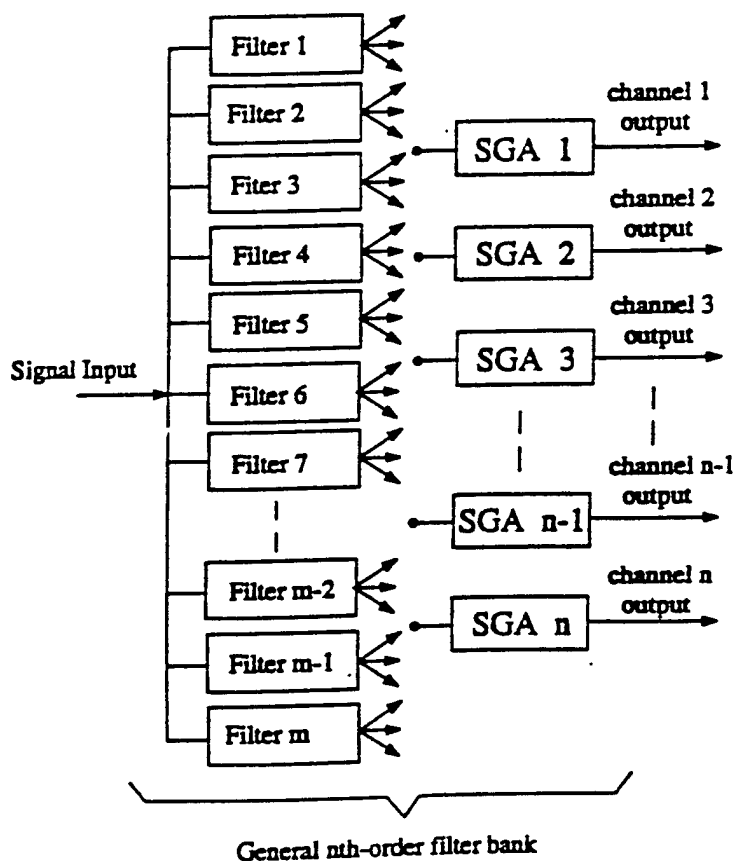
General nth-order filter bank
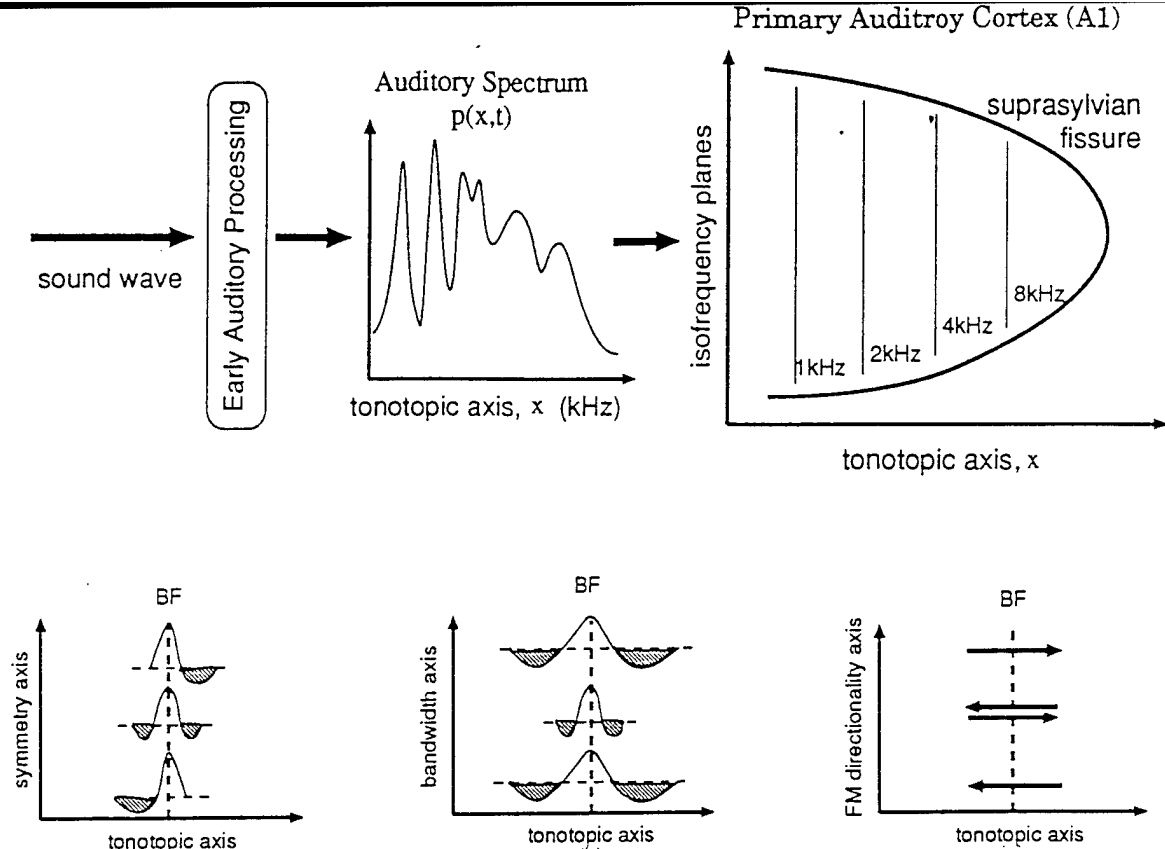
Figure 2: Details of patent awarded

Figure 3: (a) Schematic of the transformation of the acoustic stimulus into an auditory spectrum and then into a 2-dimensional pattern of activity in the auditory cortex. Responses of units along the isofrequency planes are differentiated by their binaural preferences, in addition to various monaural properties. (b) Three monaural response properties that are distinguishable along the isofrequency planes: Asymmetry, bandwidth, and FM directional selectivity of the response fields.

II. Functional Organization of the Auditory Cortex: Neurophysiology

There are several ongoing projects to explore the functional organization of the primary and anterior auditory cortex. They range from neurophysiological mappings of the responses of the various areas of the auditory cortex, to a detailed comparison of responses across the different fields, to the exploration of the linearity of cortical responses using complex broadband stimuli. Results from these projects are summarized below.

II.1 Neurophysiological mappings of the primary auditory cortex

The primary auditory cortex (AI) is essential for the perception and localization of sound. Its precise role in carrying out these functions, however, remains a mystery despite extensive knowledge gained from ablation experiments and from single and multi-unit recordings with various complex stimuli. Two general organizational features of AI have been previously firmly established: the spatially ordered tonotopic axis, and the alternating bands of binaural response properties that run perpendicularly to the isofrequency planes (Fig.3a). These axes relate to basic simple properties of the acoustic stimulus that are already established at much lower levels of the auditory pathway. With the exception of the more specialized auditory system of the bat, ordered responses to more complex stimulus features, analogous to the orientation columns and direction of motion selectivity in the visual cortex, have been more difficult to find in AI. At present, only a few reports hint at the existence of such maps in AI.

This issue was addressed in a series of experiments in the ferret AI, the results of which

6
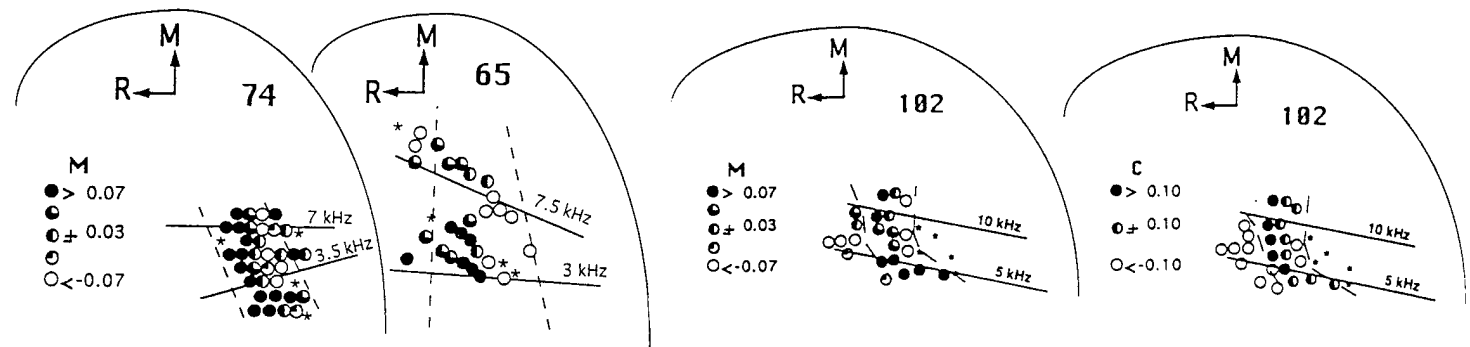
(a) Distribution of the 2-tone response symmetry measure, *M*, in the ferret primary auditory cortex from 2 animals (65 and 74). Circles: locations of the electrode penetrations along the isofrequency contours (——). Asterisks: penetrations with weak auditory responses. The arc in each map represents the location of the suprasylvian fissure; the dashed lines delineate the approximate borders of the band within which the *M* measure changes from extreme negative (○) to extreme positive values (●). A key for the classification scheme used is shown on the *left*. The medial (M) and rostral (R) directions are indicated by the arrows; the arrow lengths represent 0.5-mm distances on the surface of the cortex.

(b) comparison of the topographic distribution of 2-tone and frequency modulation (FM) responses in primary auditory cortex (AI). Map on the *left*: index *M* derived as in Fig. 11. Map on the *right*: corresponding distribution of the *C* index computed from FM responses of the same cells as in the *left* map. Map features and symbols are as in Fig. *a*. Filled circles: penetrations with selective responses to downward sweeps. Clear circles: penetrations with selective responses to upward sweeps. Partially shaded penetrations are less selective.

Figure 4:

were published in Shamma et al.(1992) appended to this review. The study explored the detailed organization of the excitatory and inhibitory responses of cortical cells, i.e., their so-called receptive fields. The aim was to establish whether any systematic changes in the balance of inhibitory and excitatory responses occur in cells along the isofrequency planes and, if so, to determine the implications of these changes to responses to frequency-modulated (FM) tones and spectrally shaped noise stimuli. These response features are more complex than the determination of a single best frequency BF (tonotopicity) or the (binary) nature of a binaural interaction (e.g., an Excitatory-Excitatory or Excitatory-Inhibitory response). The receptive field of a cell represents, to first order, its transfer function, i.e., the way it filters or processes the input spectrum. Similarly, FM tones reveal information about the dynamic interplay between of the inhibitory and excitatory responses of the cell.

The basic findings of the above experiments is the existence of a spatially ordered change in the symmetry of the receptive fields in any given isofrequency plane in AI (Fig.4). Considering the results of experiments from over 20 animals, the outline of the distribution is as follows : At the center of AI, units respond with a narrow excitatory tuning curve at BF, flanked by narrow symmetric inhibitory side-bands. The receptive fields become more asymmetric away from the center. In one direction (caudally in the ferret AI), the inhibitory side-bands above the BF become relatively stronger. The opposite occurs in the other direction. These response types tend to organize along one or more bands that parallel the tonotopic axis (i.e., orthogonal to the isofrequency planes).

Many more response properties were also examined. These include the relation between responses to spectrally shaped noise and the symmetry of the receptive fields, the selectivity of a cell's response to the direction of an FM tone in relation to its receptive field symmetry (Fig.4), and the dependence of these properties on stimulus parameters such as tone intensities and
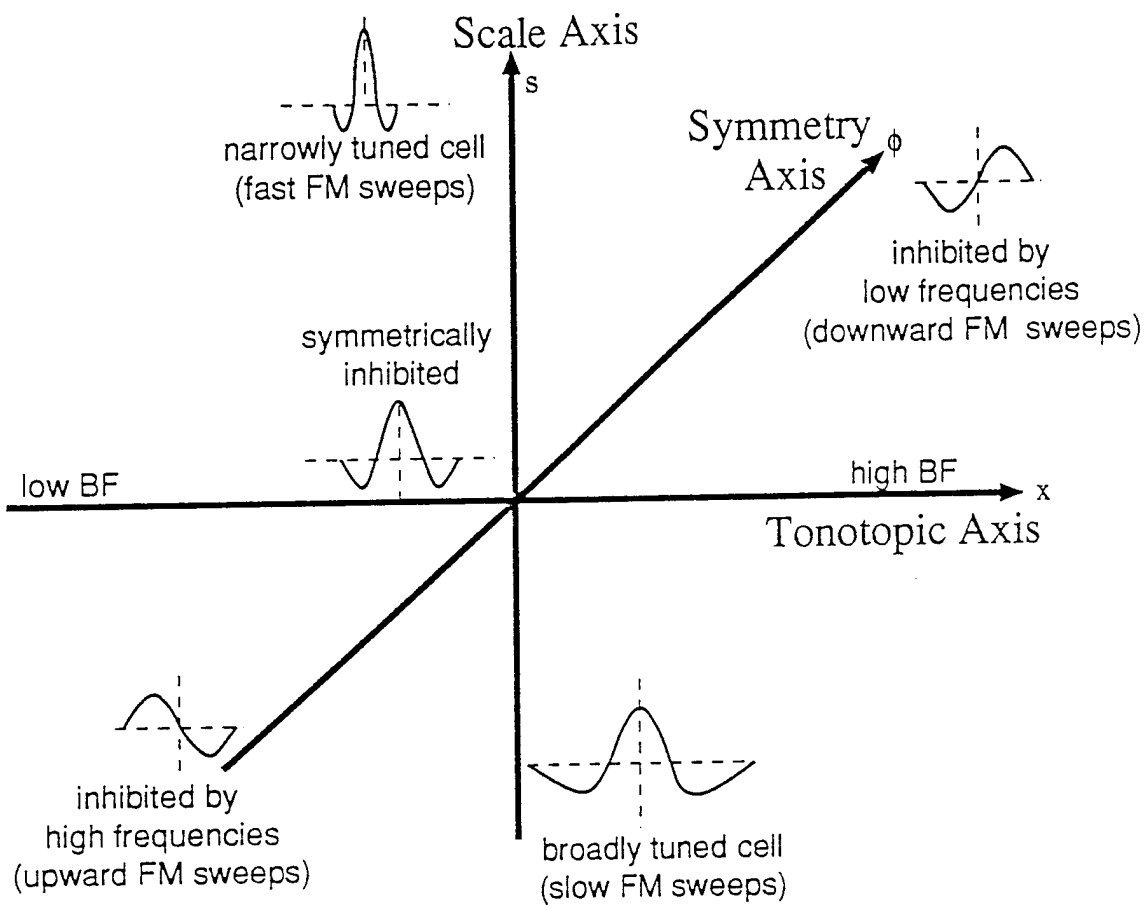
Figure 5: (a) Schematic of the three organizational axes of the response fields in AI

inter-tone delays. Another important finding of these mapping was the columnar organization of the responses, in which all units sampled in a given penetration were found to exhibit roughly similar receptive field symmetry and FM selectivity.

A fundamental conjecture suggested by these results was that units along the symmetry axis in fact encoded by their differential distribution along the isofrequency planes, a local measure of the shape of the acoustic spectrum - specifically, the locally averaged *gradient* of the spectrum. This conjecture follows from the schematics of Fig.3b where best responses to spectral peaks or edges of different symmetries are mapped systematically across the AI. The significance of such a map stems from its enhancement and explicit representation of such perceptually important features as the shape of spectral peaks, edges, and the spectral envelope. This gradient map can be viewed as a one dimensional analogue of the orientation columns of the visual cortex, since the orientation of a two-dimensional edge simply entails specifying its gradients in two directions. These physiological results in turn suggested a series of psychoacoustical experiments that are summarized later in this review.

II.2 Comparison of the responses in the anterior and primary auditory fields

The characteristics of an anterior auditory field (AAF) in the ferret auditory cortex were described in terms of its electrophysiological responses to tonal stimuli and compared to those of primary auditory cortex (AI). The AAF is located dorsal and rostral to AI on the ectosylvian gyrus and extends into the suprasylvian sulcus rostral to AI. The tonotopicity is organized with high frequencies at the top of the sulcus bordering the high- frequency area of AI, then reversing with lower BFs extending down into the sulcus. AAF contained single units that responded

8

to a frequency range of 0.3 - 30 kHz. Best frequency (BF) range, rate-level functions at BF, FM directional sensitivity, and variation in asymmetries of response areas were all comparable characteristics between AAF and AI. Responses in both areas were primarily phasic. The characteristics that were different between the two cortical areas were: latency to tone onset, excitatory bandwidth 20 dB above threshold (BW20) and preferred FM rate, as parameterized with the centroid (a weighted average of spike counts). The mean latency of AAF units was shorter than in AI (16.8 ms AAF, 19.4 ms AI). BW20 measurements in AAF were typically twice as large as those found in AI (2.5 oct AAF, 1.3 oct AI). The AI centroid population had a significantly larger standard deviation than the AAF centroid population. The relationship between centroid and BW20 was also examined to see if wider bandwidths were a factor in a unit's ability to detect fast sweeps. There was significant ($P < 0.05$) linear correlation in AAF but not in AI. In both fields, the variance of the centroid population decreased with increasing BW20. BW20 decreased as BF increased for units in both auditory fields.

II.3 Characterization of AI responses using broadband spectral ripples

As reviewed earlier in Sec.II.1, response areas of cells along the isofrequency planes of the mammalian primary auditory cortex (AI) are systematically organized with respect to two properties: their excitatory bandwidths and their asymmetry (Fig.5). To measure the response areas, these investigations employed simple tones which can be thought of as impulse-like stimuli along the tonotopic axis. If cortical cells were to respond linearly, the measured response areas would reflect the "impulse responses" of the system along the tonotopic axis, and hence could be used to predict the system's responses to arbitrary spectra. Furthermore, by Fourier transforming the impulse response, one would obtain the corresponding "transfer function", which represents the system's response to sinusoidally modulated spectra (Fig. 1B), more commonly known in the psychoacoustical literature as rippled spectra. Consequently, response properties measured by tonal stimuli might be equally evident from their ripple transfer function.

The suggestion that cortical cells are linear might appear farfetched given that their rate-level functions often exhibit threshold, saturation, and nonmonotonic behavior Nevertheless, just as measuring with tones a cell's bandwidth, tuning quality factor, or other linear systems response properties is considered meaningful, certain characteristics of the ripple responses may also prove useful, or possibly related to the properties measured with tones. It is possible as well that nonlinearities observed with tonal stimuli are less troublesome with broadband rippled spectra, or negligible over a certain range of stimulus parameters.

An analogous situation to the above has long existed in experimental studies of auditory-nerve responses. There, nonlinearities such as firing rate rectification, saturation, two-tone suppression, and adaptation are prevalent (see review in Pickles 1986). These nonlinearities, however, did not impede measurements of transfer characteristics of auditory-nerve fibers using single tones, noise stimuli, or acoustic clicks, all implying strong linear components in the responses.

Our primary goal in this work was to measure the responses of AI cells to rippled spectra at various ripple frequencies and phases, i.e., to measure their ripple transfer functions, and the dependence of this function on the amplitude of the ripples and the overall intensity of the sound. A second objective was to compare characteristic features of these transfer functions to
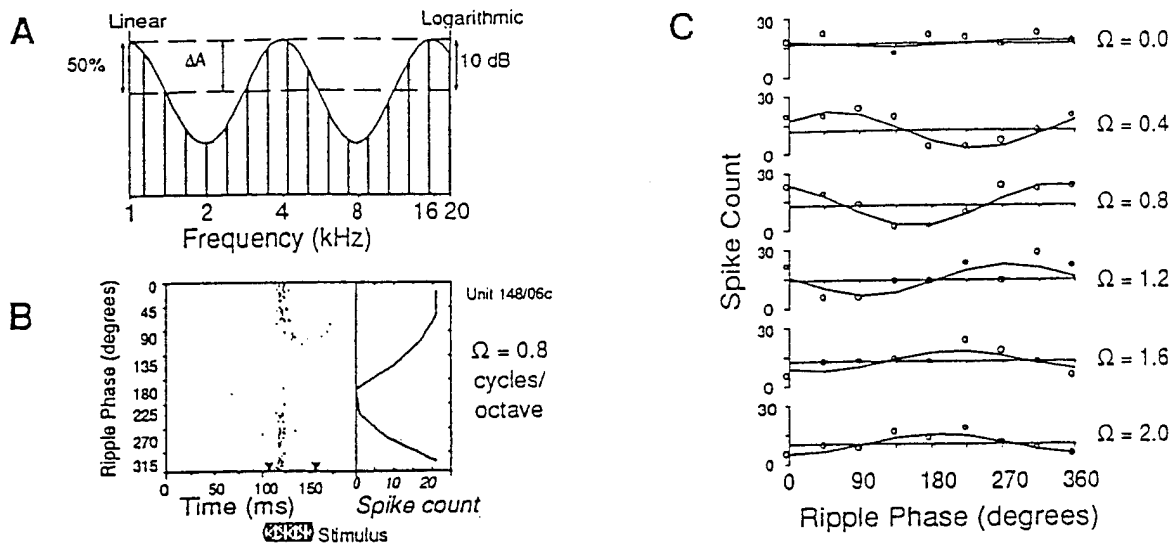
Figure 6a (A) Rippled complex stimulus composed of 101 tones equally spaced along the logarithmic frequency axis between 1 and 20 kHz. The envelope is sinusoidally modulated on a logarithmic frequency axis either on a linear (left ordinate) or a logarithmic (right ordinate) amplitude scale. The ripple phase is defined relative to the start of a ripple (i.e., low-frequency edge). The ripple stimuli were varied in ripple frequency Ω (0 - 4 cycle/octave in steps of 0.2 or 0.4) or ripple phase Φ (0-7π/4 in π/4 steps). The stimulus duration was 50 ms. (B) Dot display of responses of an AI cell (BF = 7.5 kHz) to a ripple of a fixed Ω at various ripple phases. The stimuli were started at 100 ms. Spike counts are computed over a 50 ms window as indicated by the bold arrows, and are displayed in the inset plot to the right. (C) Spike counts as a function of ripple phase for various ripple frequencies Ω. The spike counts are indicated by the circles, and for each Ω the abscissa is placed at the spike count averaged over all phases. The solid line is the best mean-square-error fit to a sinusoid.
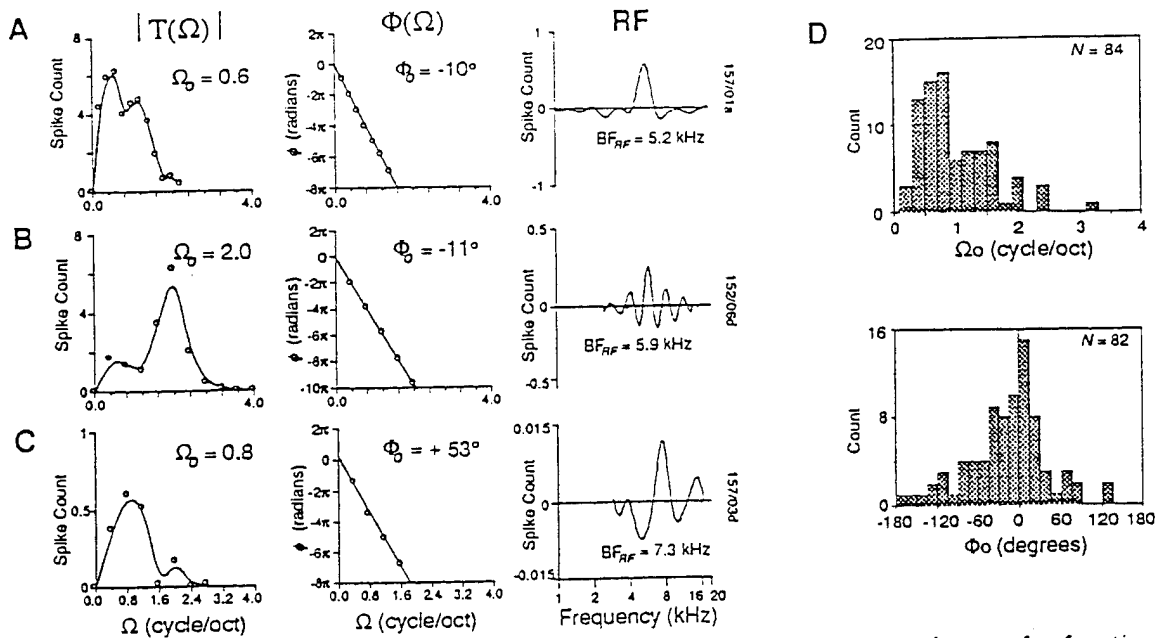

Figure 6b (A-C) Examples of ripple responses from three cells. Left: magnitude transfer function $|T(\Omega)|$; middle: phase transfer function $\Phi(\Omega)$; right: response field (inverse Fourier transform of transfer function). Note that width of $|T(\Omega)|$ is about 1 octave. (D) Distribution of Ω_o and phase Φ_o of single units in different experiments in ferret AI.
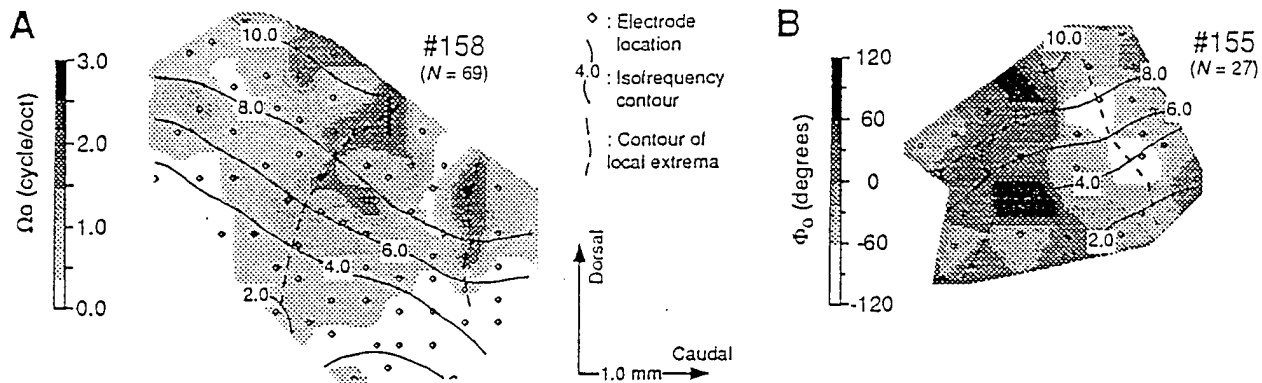
Figure 6:

10

Figure 6c *Cortical maps of Ω_o (A) and Φ_o (B) for multiunit recordings in ferrets #158 and #155. The scaling bar on the left indicates the values represented by the grey intensities. A Gaussian weighted filling was applied to obtain the values between penetration locations (marked by circles). Original values at the electrode locations were preserved. The map is made on the basis of responses where $|\Phi_o| \leq 100°$. The dashed lines connect local maxima of Ω_o, and maxima and minima of Φ_o, across isofrequency planes.*

response properties measurable using tonal stimuli, such as the bandwidth or the asymmetry of the response area.

Such an approach has proven fruitful in analogous studies of the primary visual cortex (De Valois and De Valois 1988). There, transfer functions measured using sinusoidally modulated gratings reveal much about the functional organization of the system, and its response to more complex stimuli such as oriented bars. In auditory physiology, such stimuli have only been reported by Calhoun and Schreiner (1993). Recently, several psychoacoustical studies (Hillier 1991; Summers and Leek 1994; Vranić-Sowers and Shamma 1994a, 1994b) have converged on the similar notion that measuring the perceptual thresholds of rippled spectra may help explain how spectral profiles are perceived.

The results we obtained in extensive recordings of single units in the primary auditory cortex of the ferret can be summarized as follows. Using broadband stimuli (1-20 kHz) with sinusoidally modulated spectral envelopes (ripples), the response magnitude of each cell was measured as a function of ripple frequency (Ω) and ripple phase (Φ), from which a "ripple transfer function" was constructed (Fig.6a-b). Most cells (approximately 90%) responded best around a specific (characteristic) ripple frequency, Ω_o. Values of Ω_o range from 0.2 to 3 cycles/octave, with the average of the distribution around 1.0 (Fig.6b). Most cells also exhibited a linear ripple phase as a function of Ω. The intercept of the phase function is interpreted as the best (characteristic) ripple phase to drive the cell, Φ_o; the slope of the line reflects the location of the response area of the cell along the tonotopic axis. Φ_o ranges over the full cycle in a Gaussian-like distribution around $0°$ (Fig.6b). By inverse Fourier transforming the transfer function, a "response field" (RF) of the cell was obtained, an analogue of the response area measured with tonal stimuli. Parameters of the RF were compared to parameters of the tonal

response area. The BF of the RF, BF_{RF}, was very similar to the tonal BF, and Ω_o and Φ_o were weakly but significantly correlated to the excitatory bandwidth and asymmetry index of the tonal response area, respectively. The RF was found to be a stable measure of a unit's response regardless of ripple amplitudes or overall stimulus levels. Responses to rippled spectra in AI closely resemble the response properties to sinusoidal gratings in the primary visual cortex (VI). This provides a unified framework within which to interpret the functional organization of both cortices. Details of this work are found in the manuscript entitled "Ripple Analysis in the Primary Auditory Cortex (Part I)", copies of which are appended to this proposal.

We have also examined the topographic distribution of response parameters using the ripple and tonal stimuli in the primary auditory cortex (AI) (Fig.6c). Both single-unit and multiunit recordings were used in these studies. As before, for each unit or cluster, responses to ripples were parametrized in terms of the characteristic ripple Ω_o and phase Φ_o (i.e., the best ripple frequency and phase, respectively). Two corresponding response area parameters (using tonal stimuli) were also measured: the excitatory bandwidth at 20 dB above threshold (BW20) which is roughly inversely proportional to Ω_o, and the asymmetry as reflected by the directional sensitivity index (C) to frequency-modulated (FM) tones (which is proportional to Φ_o). The response parameters measured from multiunit records corresponded well to those obtained from single units in the same cluster. The topographic distribution of the response parameters across the surface of AI was studied with multiunit recordings in four animals. In most maps, systematic patterns or clustering of response parameters could be discerned along the isofrequency planes. The distribution of the characteristic ripple Ω_o exhibited two trends. First, along the isofrequency planes, the largest values were grouped in one or two clusters near the middle of AI, with smaller values found towards the edges. The second trend occured along the tonotopic axis where the maximum Ω_o found in an isofrequency range increases with increasing BF. The distribution of the characteristic ripple phase, Φ_o, which reflects the asymmetry in the response field, also showed a clustering along the isofrequency axis. At the center of AI symmetric responses ($\Phi_o \approx 0$) predominated. Towards the edges, the RFs became more asymmetric with $\Phi_o < 0$ caudally, and $\Phi_o > 0$ rostrally. The asymmetric response types tended to cluster along repeated bands that paralleled the tonotopic axis. The distribution of the response area measures BW20 and C-index exhibited similar trends along the isofrequency planes as Ω_o and Φ_o .

Details of this work are found in the manuscripts entitled "Ripple Analysis in the Primary Auditory Cortex (Part II)", copies of which are appended to this proposal.

III. Functional Models of the Auditory System: Psychoacoustics

The experimental results described above suggested that specific features of the shape of the acoustic 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 symmetry 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. These are described in detail in the two manuscripts accompanying this proposal entitled "Representation of Spectral Profiles in the Auditory System: Parts I and II".

Over the last year, we have finished a series of experiments on the perception of spectral
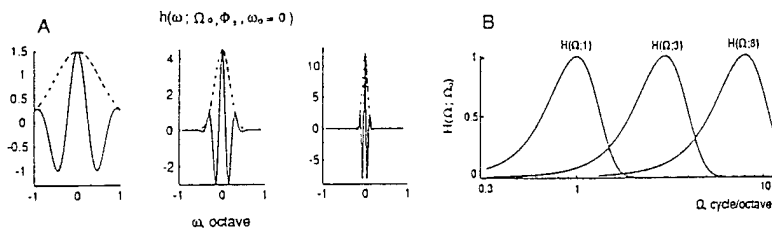
Figure 7a *(A) Impulse responses of three filters with characteristic ripple frequencies* $\Omega_o = 1, 3$, *and 8 cycle/octave and characteristic phase* $\Phi_o = 0$. *Filters are centered at* $\omega_o = 0$ *octave. The impulse response is computed for* $\sigma(\Omega_o) = 0.3\ \Omega_o$. *(B) Fourier transform of the three impulse responses of the filters in (a) plotted on a logarithmic* Ω *axis.*
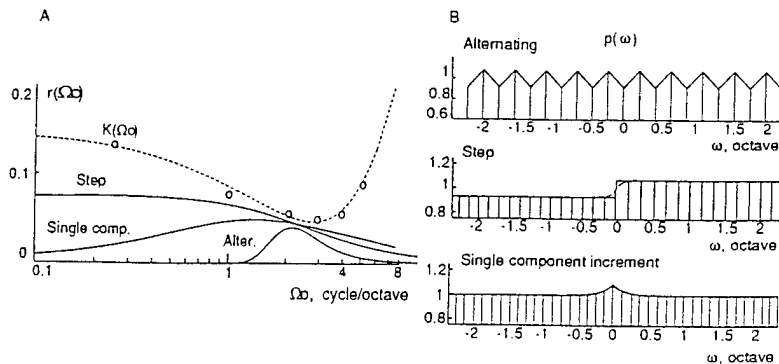

Figure 7b *Profile detection tests in the ripple analysis model. (A) The dashed line is a polynomial approximation to the perceptual thresholds measured with single ripples (reproduced from Fig. 3.27 in Hillier, 1991). Each data point (denoted by circles) represents the peak amplitude of* $r(\Omega_o)$ *due to a just-detectable ripple with frequency* Ω_o. *The solid lines represent* $r(\Omega_o)$ *computed for the profiles shown in (B). (B) The alternating, step, and single component increment profiles at their just-detectable amplitudes according to Bernstein and Green (1987).*

peak shapes, specifically their symmetry and bandwidth. The results we obtained could be explained elegantly by a model of auditory spectral profile perception that utilizes spectral ripples as the elementary representational features. In particular, the model assumes that the spectral profile is represented in the auditory system by a weighted sum of sinusoidally modulated spectra (ripples). The analysis is performed by a bank of bandpass filters, each tuned to a particular ripple frequency and ripple phase (Fig.7a). The parameters of the model are estimated using data from single ripple detection experiments. The model is then used to account for detection thresholds of more complex profiles such as the step, single component increment, and the alternating profiles (Fig.7b). Physiological and psychophysical evidences from the auditory and visual systems in support of this type of a model are reviewed in detail in the accompanying manuscripts.

Based on the ripple analysis model, detection thresholds of of shape changes in spectral peak profiles were then interpreted. Peak shape is uniquely described in terms of two parameters: bandwidth factor (BWF) which reflects the sharpness of a peak, and a symmetry factor (SF) which roughly measures the local evenness or oddness of a peak. Using profile analysis methods, thresholds to changes in these parameters (defined as δBWF/BWF and δSF) are measured together with the effects of several manipulations such as using different peak levels, varying spectral component densities, and randomizing the frequencies of the peaks. The ripple analysis model could account well for the measured thresholds (Fig.8). Predictions of three previously published models for the same profiles were also evaluated and discussed.
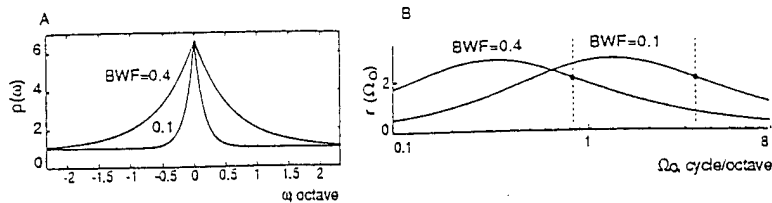
Figure 8a (A) Profiles of two symmetric peaks (described in detail in Vranić-Sowers and Shamma 1994a). The BWF measures the 3dB-bandwidth of a peak. (B) The ripple transform magnitudes, $r(\Omega_o)$, of the peaks in (A). The dashed lines denote the locations of the steepest low pass edges in $r(\Omega_o)$. The effect of a BWF change is a shift (and not a change in shape) of the ripple transform along the logarithmic Ω_o axis. For example, a four-fold increase in BWF (from 0.1 to 0.4 octaves), i.e., $\delta BWF/BWF=3$ or $\alpha = 0.25$ results in a $(\Delta = \log_2 \alpha =)$ 2 octave downward shift in $r(\Omega_o)$.
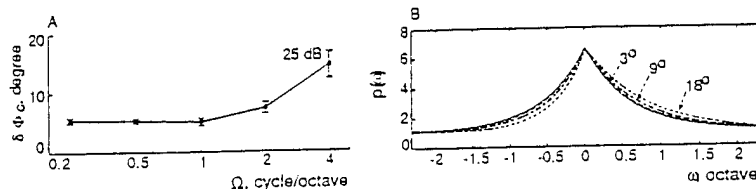

Figure 8b (A) Perceptual phase difference, $\delta\Phi_c$, at threshold (reproduced from Vranić-Sowers and Shamma, 1994b) for various ripple frequencies Ω, for 161 frequency components and at 25 dB peak-to-valley amplitude. The task is to detect a change in the phase of a ripple while keeping the ripple frequency constant. (B) Shifting the phase of the ripple transform phase of the peak profile by a constant angle simply changes the symmetry of the profile. Solid line is a symmetric peak (BWF = 0.4 octave) with ripple transform phase = 0. Dashed lines are the skewed peaks resulting from adding three different angles to the ripple transform phase (Vranić-Sowers and Shamma. 1994a).

Figure 8:

IV. 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 suggests that the auditory system analyzes an input spectral pattern along three independent dimensions: a logarithmic frequency axis, a local symmetry axis, and a local spectral bandwidth axis (Fig.5). It is shown that this representation is equivalent to performing an affine wavelet transform of the spectral pattern and preserving both the magnitude (a measure of the scale or local bandwidth of the spectrum) and phase (a measure of the local symmetry of the spectrum). Such an analysis is in the spirit of the cepstral analysis commonly used in speech recognition systems, the major difference being that the double Fourier-like transformation that the auditory system employs is carried out in a local fashion. Examples of such a representation for various speech and synthetic signals are discussed in the accompanying paper which has been accepted for publication in the IEEE Audio and Speech Processing.

V. Analysis of Neural Network Architectures with Wavelet Transforms

The goal of this work was to achieve a coherent theoretical foundation for a class of neural network architectures called *feedforward networks*. In the work of Y.C. Pati and P.S. Krishnaprasad, it was shown that it is possible to structure feedforward networks using the theory of discrete affine wavelet transforms. In particular, it was shown that for $L^2(R)$, it is possible to construct frames out of sigmoids, and hence represent elements of $L^2(R)$ as feedforward networks with one hidden layer. The dilations and shifts that appear in such a representation are

14

determined from prior knowledge of spatio-spectral concentration of the given function/map. A major advantage of this structuring is that the coefficients to be fitted from data, being simply the weights from the hidden layer to the output layer, enter linearly in the model, thereby leading to a convex/quadratic optimization problem. This result, being among the very first to place feedforward networks in the rigorous context of wavelet theory, inspired us to further investigate the use of wavelet transforms.

One of the difficulties in extending the above result to the multidimensional setting of $L^2(R^n)$ is that a naive approach based on tensor products of scalar frames can lead to computation-intensive formulations. Further it was not quite clear how to work out the frame theory in this context, –in particular certain key results of Daubechies needed to be extended. In the forthcoming M.S. thesis of T. Kugarajah, these problems are solved, and furthermore, the results have been applied to the problem of adaptive control in nonlinear systems. This goes quite a bit beyond the current literature based on the *ad hoc* use of radial basis functions to model nonlinear vector fields.

In continuation of our effort on feedforward networks, we started looking into the possibility of similar techniques for recurrent networks (in continuous time). More precisely, we started a careful study of the the problem of vector field approximations in nonlinear dynamics via basis vector fields. Prior efforts ignore the fundamental geometric differences between vector field approximations and function approximations (vector fields do not transform under coordinate change in the same way as scalar fields). We note for instance the use of radial basis functions in adaptive control. Our effort, based on an understanding of geometric approximation techniques for vector fields by polynomial vector fields, and vector wavelets (a subject that is undergoing rapid development due to the interests of researchers in fluid mechanics and quantum field theory) is likely to yield new insights into the structure of recurrent networks and more generally locally interacting dynamical systems. A graduate research assistant supported by AFOSR (Herbert Strumper) is involved in this study, as a part of his Ph.D. research.

To enhance the scope of our research in wavelet bases, we considered the problem of approximation of infinite-dimensional linear dynamical systems by finite dimensional linear systems. This fundamental problem, arising in many fields of applications of control theory ranging from control of vibrating structures, to control of heat-flow in furnaces, to cancellation of noise in real-time by destructive interference, is shown to have an elegant solution via rational wavelets in the Hardy space $H^2(\Pi^+)$. The construction of such bases has opened the way for a wide range of applications e.g. in the use of smart materials to carry out fast identification of structural dynamics.

A direct consequence of the work of on rational wavelets was the discovery of new ways to organize the design problem for switched capacitor filters for the auditory studies (Fig.9a). This led to useful collaboration between Daniel Lin (Ph.D. student of Shamma) and Y.C. Pati (Ph.D. student of Krishnaprasad). Further applications of rational wavelets are under way.

The work on rational wavelets for identfication would have remained a theoretical curiosity if it were not possible to do fast computation of such rational wavelet models. In recent work, new recursive algorithms have been devised for systematic approximation via basis function representations. The new algorithms known as orthogonal matching pursuit algorithms are applicable to a wide class of problems, ranging from fitting radial basis function approximations to wavelet-based models for transfer functions of linear systems. These new algorithms are

well-equipped to work with raw data as well as data subject to preliminary processing. It is of further interest that these algorithms are well-suited to the exploitation of certain forms of *a priori* knowledge (in the time-frequency plane). Several papers have resulted from this work, and software packages are now available for the use of such algorithms on Sun workstations. The packages include a commercial MATLAB based toolbox and a free package developed at Stanford and Maryland. These come with effective graphical user interfaces. One M.S. student has put the techniques and software to good use in the identification of the dynamics of flexible beams with surface mounted piezo-electric sensors and actuators (so-called smart structures), (Fig.9b). The algorithms are fast enough to merit consideration in real-time applications.

The recent discoveries in thalamo-cortical oscillations have led to the suggestion that oscillatory neural networks are playing an important part in the solution to the so-called "dynamic binding problem", where coherent oscillations encode the binding together of features of an image in a receptor field. Somewhat influenced by these exciting developments, we undertook a deep study of the properties of networks of oscillatory neurons (sometimes called rotor neurons). We have a better understanding of the mean-field theory of a class of such networks. We have proved new convergence theorems using arguments based on LaSalle's invariance principle. We have further insights into asymptotic behaviors. New implementations in analog networks are being considered. One M.S. student (Eric Justh) is involved in this project. He is however an NSF Graduate Fellow and hence does not need support from AFOSR. His thesis has been just completed and a preprint based on this work is available and is being readied for submission to a journal.

In addition to the papers listed below, several papers are under preparation including one based on a presentation at an April 1994 Symposium sponsored by the National Academy of Sciences Board on Mathematical Sciences. These and other papers also influenced by the current AFOSR project will be provided for review.

Publications acknowledging this grant

Papers that have appeared or submitted under this grant are:

X. Yang, K. Wang, and S. Shamma, " Auditory Representations of Acoustic Signals", *IEEE Trans. on Info. Theory*, 38, pp. 824-839, 1992.

M. Peckerar, S. Shamma, M. Robbert, J. Kosakowski, and P. Isaacson, " Passive Micro-electrode Arrays for Recording of Neural Signals", *Rev. Sci. Inst.*, 62(9), pp. 2276-2280, 1992.

S. Shamma, J. Fleshman, P. Wiser, H. Versnel "Response Area Organization in the Ferret Primary Auditory Cortex", *J. Neurophys*, 69(2), pp. 367-383, 1993.

K. Wang and S. Shamma "Self-Normalization and Noise Robustness in Early auditory processing", *IEEE Trans. Aud. and Speech*, 2(3), pp. 421-435, 1994
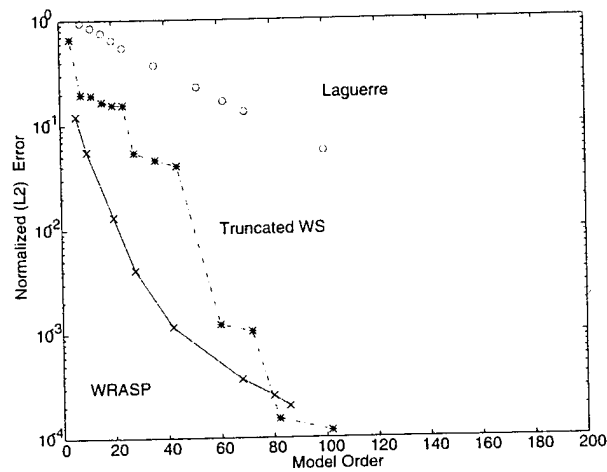
K. Wang and S. Shamma "Wavelet Representations of Sound in the Primary Auditory Cortex" *J. Optical Engineering*, 33(7), pp. 2143-2148, 1994.

J. Lin, T. Edwards, and S. Shamma, "Analog VLSI Implementations of Auditory Wavelet Transforms", *IEEE Trans Cir. Sys.*, 41(8), (17 pages) 1994.

Example I: Approximation of Cochlear Filter Response

Normalized ($H^2(\Pi^+)$) Approximation Error Versus Model-Order

Example II: Identification of Flexible Beam (Rezaiifar 1993)

Experimental Flexible Beam Setup.
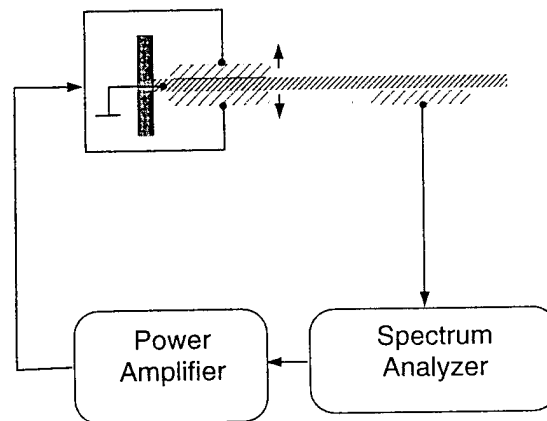

Figure 9:

S. Shamma, H. Versnel, and N. Kowalski, "Ripple Analysis in the Ferret Primary Auditory Cortex. I. Response Characteristics of Single Units to Sinusoidally Rippled Spectra," submitted to *J. Neurophys.*, approx. 28 pages, 1994.

H. Versnel, S. Shamma, and N. Kowalski, "Ripple Analysis in the Ferret Primary Auditory Cortex. II. Topographic and Columnar Distribution of Ripple Response Parameters," submitted to *J. Neurophys.*, approx. 20 pages, 1994.

S. A. Shamma, S. Vranic, and P. Wiser, "Spectral Gradient Columns in Primary Auditory Cortex: Physiological and Psychophysical Correlates", in *Auditory Physiology and Perception*, Ed. by Y. Cazals, K. Horner, and L. Demany, Pergamon Press, Oxford, pp. 397-406, 1992.

S. A. Shamma, " Common Principles in Auditory and Visual Processing", in *Neuroscience: From Neural Networks to Artificial Intelligence*, Edited by P. Rudomin, M. Arbib, F. Cervantes, and R. Romo. Springer Verlag, Heidelberg, Germany, pp. 189-205, 1994.

R. Lyon and S. Shamma, "Computational Strategies for Pitch and Timbre", n *Auditory Computations*, Ed. H. Hawkins, T. McMullen, A. Popper, and R. Fay, Springer Verlag, approx. 35 pages, 1994 (in press)

W. Byrne and S. Shamma, "Neurocontrol in Sequence Recognition", in *Neural Networks in Controls Applications*, Ed. by D. Elliott, approx. 15 pages, 1994 (in press).

S. Shamma, "Neural and Functional Models of the Auditory Cortex", in *Handbook of Brain Theory and Neural Networks*, Ed. by M. Arbib, Bradford Books, MIT Press, approx. 6 pages, 1994 (in press).

S. Shamma, S. Vranic, and H. Versnel "Representation of Spectral Profiles in the Auditory System: Theory, Physiology, and Psychoacoustics", in *Advances in Hearing Research*, Ed. G. Manley, G. Klump, C. Koppl, H. Fastl, and H. Oeckinghaus, World Scientific Publishers, Singapore, approx. 11 pages, 1994 (in press).

Pati, Y. C. and Krishnaprasad, P. S. and and Peckerar, M. and Marrian, C. R. K., "Neural Networks and Tactile Imaging", in *Proceedings INNS First Annual Conference, Boston MA*, Aug. 1988.

Pati, Y. C. and Krishnaprasad, P. S., "Feedforward Neural Networks: Analysis and Synthesis Using Discrete Affine Wavelet Transforms", *Proceedings IEEE Conference on Neural Information Processing Systems*, Denver, CO 1991.

Pati, Yagyensh C., "Neural Networks for Low-Level Processing of Tactile Sensory Data", University Of Maryland., College Park, MD. Systems Research Center Technical Report No. SRC-TR-89-8 1988.

Marrian, C. R. K. and Peckerar, M. C. and Mack, I. and Pati, Y. C., "Electronic 'Neural' Nets for Solving Ill-posed problems", in *Maximum Entropy and Bayesian Methods*, 1989, 371–376, Kluwer Academic Publishers.

Pati, Yagyensh C. and Krishnaprasad, P. S., "Analysis and Synthesis of Feedforward Neural Networks Using Discrete Affine Wavelet Transforms", University of Maryland, Systems Research Center, Technical Report, SRC TR 90-44, 1990.

Teolis, Anthony and Pati, Yagyensh C. and Peckerar, Martin C. and Sham ma, Shihab", "Cascaded Neural-Analog Networks For Real Time Decomposition of Superposed Radar Return Signals in the Presence of Noise", University of Maryland, Systems Research Center, Technical Report SRC TR 89-33, 1989.

Y.C. Pati and P.S. Krishnaprasad, "Decomposition of $H^2(\Pi^+)$ via Rational Wavelets" (1992), in *Proc. Conf. Information Sciences and Systems*, Princeton, pp 15-20.

Y.C. Pati and P.S. Krishnaprasad, "Approximations of Stable Linear Systems Via Rational Wavelets" (1992), in *Proceedings of 31st IEEE Conference on Decision and Control*, IEEE, New York, pp 1502-1506.

R. Rezaiifar, Y.C. Pati, P.S. Krishnaprasad and W.P. Dayawansa, "Wavelet Based Identification of Smart Structures with Surface Mounted Actuators and Sensors", (1993), in *Proceedings of 32nd IEEE Conference on Decision and Control*, IEEE, New York, pp 486-491.

Y.C. Pati and P. S. Krishnaprasad, "Analysis and Synthesis of Feedforward Neural Networks Using Discrete Affine Wavelet Transformations", (1993) in *IEEE Transactions on Neural Networks*, Vol. 4, No. 1, pp 73-85.

Y.C. Pati, R. Rezaiifar and P.S. Krishnaprasad, "Orthogonal Matching Pursuit: Recursive Function Approximation with Applications to Wavelet Decomposition", in *Proc. 27th Asilomar Conference on Signals, Systems and Computers*, Nov. 1-3, 1993, pp 40-44.

Y.C. Pati, R. Rezaiifar and P.S. Krishnaprasad, "A Fast Recursive Algorithm for System Identification and Model Reduction using Rational Wavelets", in *Proc. 27th Asilomar Conference on Signals, Systems and Computers*, Nov. 1-3, 1993, pp 35-39.

Y.C. Pati and P.S. Krishnaprasad, "Rational Wavelets in System Identification", Invited Paper for 33rd IEEE Conference on Decision and Control, December 1994.

Eric Justh (November 1994), Convergence Analysis of a Class of Networks of Nonlinear Coupled Oscillators, M.S. Thesis, Department of Electrical Engineering, University of Maryland, College Park.

Eric Justh and P.S. Krishnaprasad (October 1994), "Convergence Analysis of a Class of Networks of Coupled Nonlinear Oscillators" , Preprint, Institute for Systems Research, University of Maryland, College Park.

Tharmarajah Kugarajah (December 1994), Neural Networks for Control, M.S. Thesis, Department of Electrical Engineering ,University of Maryland, College Park. (to appear).

Herbert Strumper (1996) Vector field approximations, Ph.D. thesis work in progress. Department of Elect rical Engineering, University of Maryland, College Park.

Personnel

The research reported here is documented in the theses and reports of the students (with degrees awarded) listed below. Several are now supported by the AFOSR grant towards their Ph.D. degree.

Y. Pati (PhD), Geeth Chettiar (MS), Huib Versnel (Res. Faculty), Danniel Lin (PhD), Tony Tiolis (PhD), Preetham Gopalaswamy (Lab Manager), Svetlana Vranic (PhD), Kuansan Wang (PhD), Tom Edwards (MS), Yuan-Tai (MS), Po-wen Ru (MS).