UNDERGRADUATE REPORT

Stereausis: A Binaural Processing Model

by Samuel Jiawei Ng
Advisor: P.S. Krishnaprasad

UG 2001-6

STEREAUSIS: A BINAURAL PROCESSING MODEL

Samuel Jiawei Ng
University of California, Irvine
Advisor: Dr. Krishnaprasad

Research Experience for Undergraduates
Institute for Systems Research
University of Maryland, College Park

August 3, 2001

TABLE OF CONTENTS

STEREAUSIS: A BINAURAL PROCESSING MODEL

ABSTRACT

The stereausis network model measures interaural differences by detecting spatial disparities between instantaneous outputs from both ears. It can detect the location of a sound source at all frequencies as well as detect and enhance noisy signals by comparing the spatiotemporal responses of an auditory nerve fiber array at various horizontal shifts. Cross correlation of cochlear responses is found by combining an ipsilateral input at a characteristic frequency with contralateral inputs at other frequencies.

STEREAUSIS NETWORK MODEL

Binaural processing involves measuring similarities and differences between inputs to both ears. Shamma's two-dimensional stereausis network model measures interaural differences by detecting spatial disparities between instantaneous inputs from both ears. The network is able to determine the location of a sound source at all frequencies as well as detect and enhance noisy signals. This is done by comparing spatiotemporal responses of an array of auditory nerve fibers at various horizontal shifts. Since spatial disparity is proportional to temporal delay between both ears, binaural processing of interaural time differences (ITDs) can be seen as spatial operations. Spatial disparities due to interaural level differences (ILDs) that affect the amplitudes of the traveling waves can be seen in a similar manner.

Two axes exist in the network: 1) disparity or lateralization axis, where the perceived location of the sound is encoded and 2) spectral or characteristic frequency (CF) axis, where the spectrum of the stimulus is reflected. Each network cell has a maximum response to a characteristic spatial shift at a characteristic frequency. Stereausis operates by computing the cross-correlation between two cochlear responses. An input from one ear (also known as the ipsilateral input) at a given CF is combined (or cross-correlated) with inputs from the other ear (contralateral inputs) at other frequencies. Correlation is determined by using delays found in the traveling waves of the basilar membrane. Images are identical if the tone is centered. If one tone is delayed (phase shifted) in one ear compared to the other, the images are shifted in space.

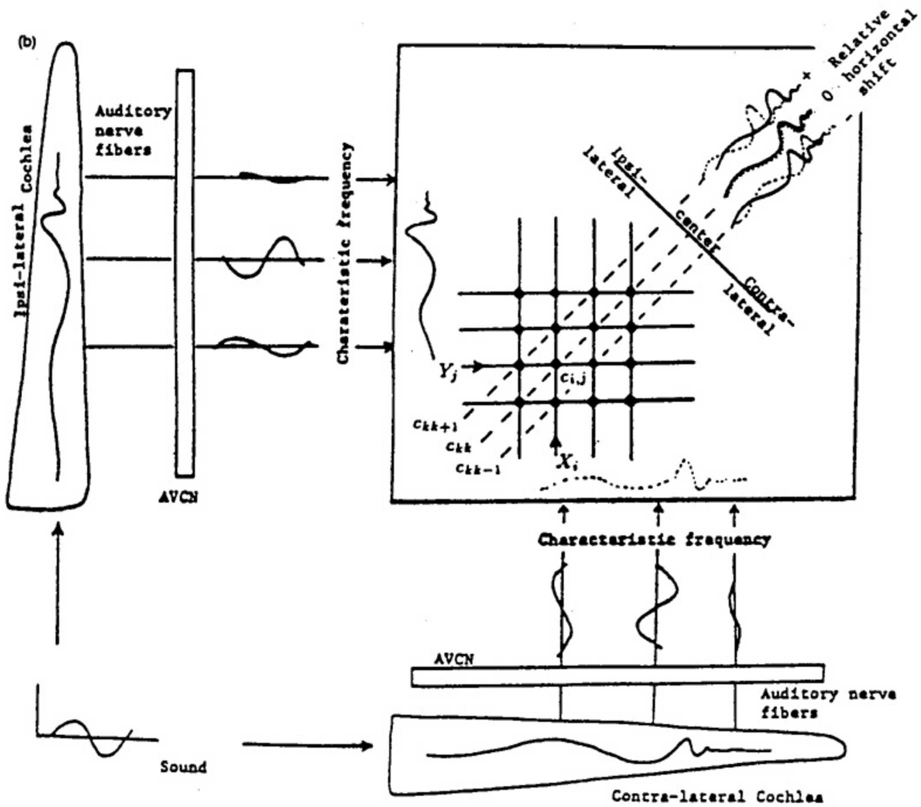

Figure 1. Stereausis binaural network

SIMULATIONS

Lateralization of Low-Frequency Tones

Low frequency tone lateralization can be represented by interaural time delays (ITDs) or phase shifts which are useful for tones below 1.5 kHz. Simulations for an 1100 Hz tone at 0, 60, 120, and 180 degrees phase shifts are shown in Figure 2. A dominant peak of activity appears along the diagonal for centered tones (zero disparity). The pattern shifts for a binaurally delayed tone where the relative height of the primary to the secondary peak gradually decreases with equally high peaks on either side of the midline at 180 degrees phase shift. Further shifts increase the height of the secondary peak making it the dominant peak.

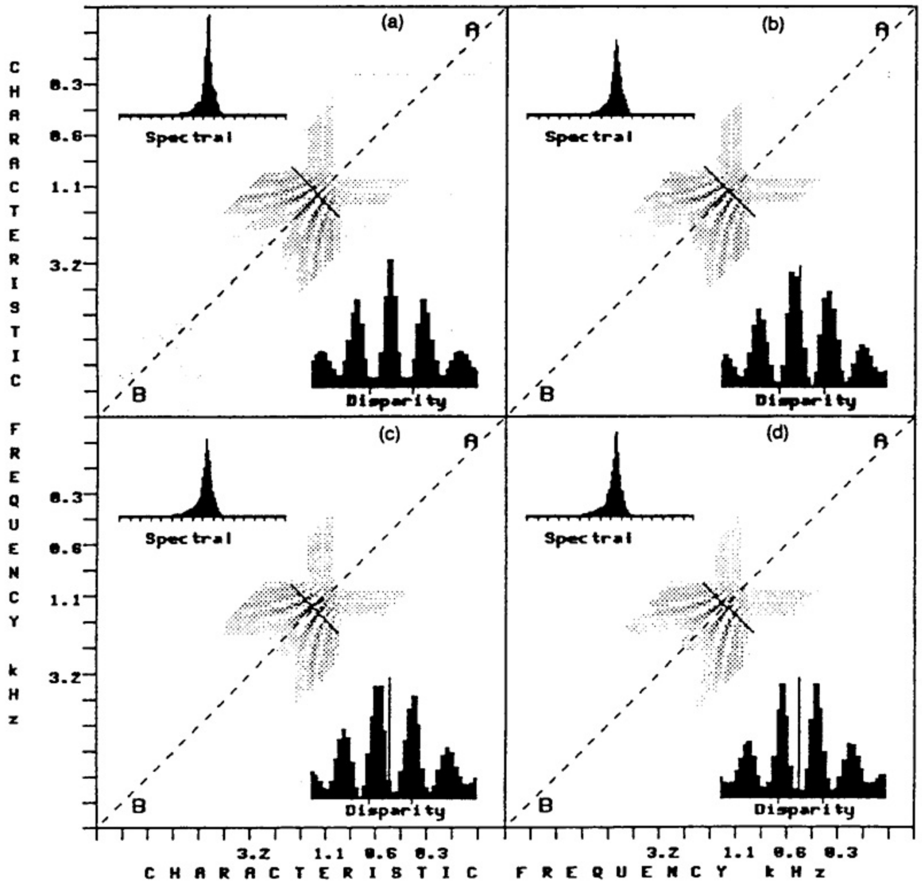

Figure 2. ITDs with low-frequency tones at 1100 Hz (a) Centered 1100 Hz tone
(b) Same as (a) but with 60-degree phase shift (c) Same as (a) but with 120-degree phase shift (d) Same as (a) but with 180-degree phase shift

Lateralization of High-Frequency Tones

Interaural level differences (ILDs) are used for frequencies greater than 1.5 kHz because the stereausis network is insensitive to ITDs at higher frequencies, especially above 3-4 kHz. Figure 3 shows simulated responses for a 3 kHz tone with ILDs of 0, 3, 6, and 12 dB. Disparity plots show the peak broadening and shifting for different ILDs. Increases in input ILDs are also shown in the relative levels of the two ridges of activity. Although the ILD plots are less accurate than those of ITDs, they are sufficient for human ILD detection.

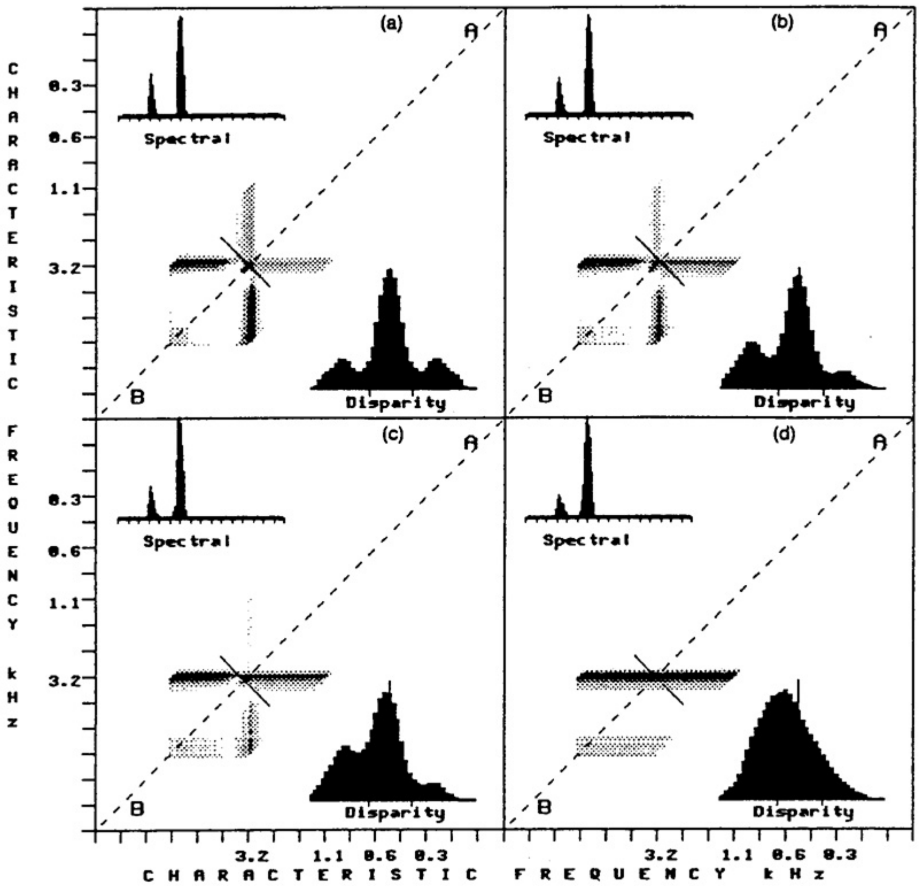

Figure 3. ILDs with high-frequency tones at 3 kHz (a) Centered tone, ILD = 0 dB (b) Same as (a) but with ILD = 3 dB (c) Same as (a) but with ILD = 6 dB (d) Same as (a) but with ILD = 12 dB

Time/Level Trading for Low-Frequency Tones

Amplitude and phase disparity information are preserved for low-frequency tones, meaning both ITDs and ILDs can influence lateralization in the stereausis network response. Figure 4 shows the effects of increasing ILD for a centered tone of 600 Hz. The primary peak is centered and fixed at 0 ITD, while increasing ILD causes the secondary peak to broaden and increase in height. According to Shamma, a particular ITD has a characteristic disparity and relative peak sizes in the patterns, while adding ILD leads to an imbalance in the outputs surrounding the primary peak.

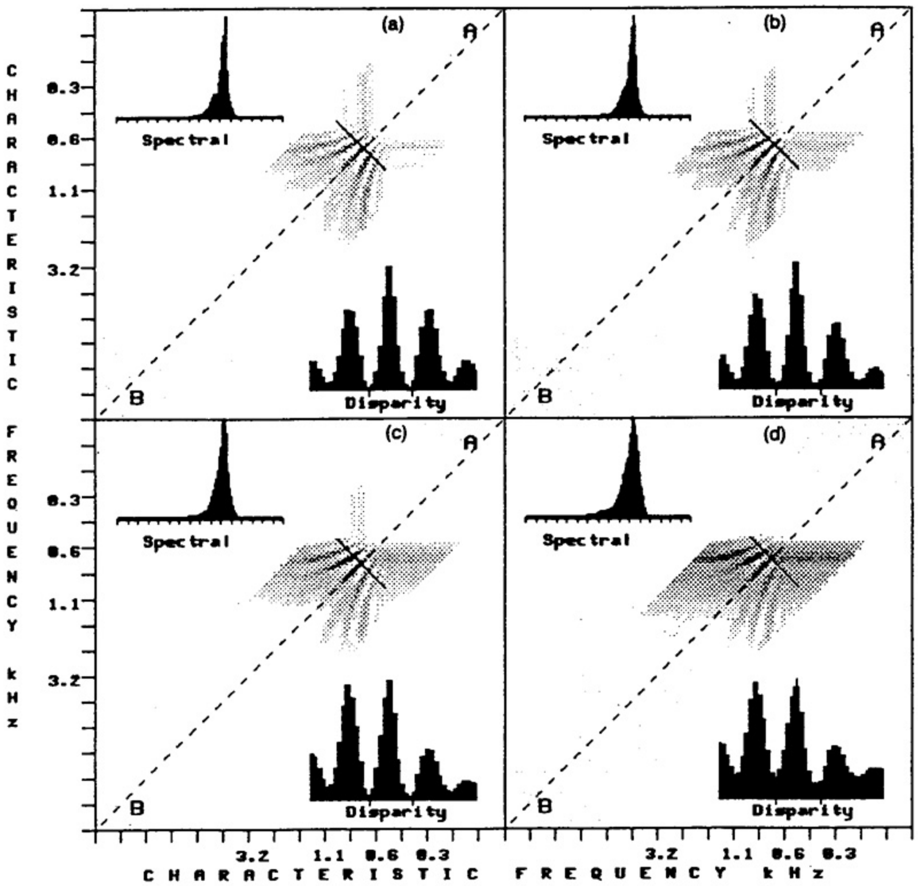

Figure 4. Effects of ILDs on low-frequency tones. (a) Centered 600-Hz tone, ILD = 0 dB (b) Same as (a) but ILD = 3 dB (c) Same as (a) but with ILD = 6 dB (d) Same as (a) but with ILD = 12 dB

Lateralization of Speech and Harmonic Complex Sounds

Both phase and amplitude disparities in complex signals are represented in the same way as single tones. The response diagonal has two spectral regions: 1) region of lower frequency resolved harmonics and 2) spectrally unresolved harmonics at higher frequencies. Spectral peaks lie along the diagonal for a centered harmonic complex. Patterns due to harmonics shift by an amount that increases with the harmonic number when an ITD is introduced. This is because of the higher velocities of the more basal traveling waves. Figure 5 shows some examples of complex stimuli.

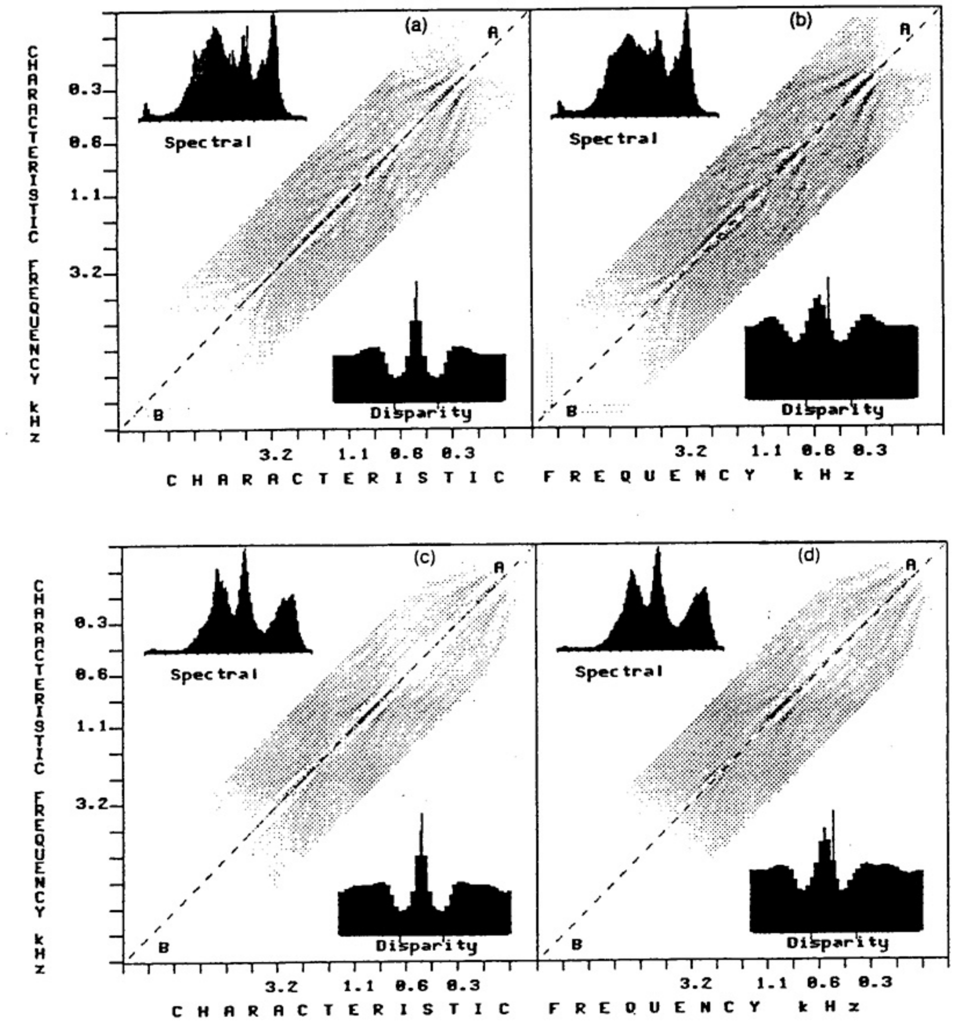

Figure 5. Lateralization of speech and complex harmonic series (a) Centered harmonic series with 300 Hz fundamental (b) Same as (a) with ITD = 250 μs (c) naturally spoken vowel /a/ (d) Same as (c) with ITD = 250 μs

Lateralization of Broadband Noise

Lateralization of noise operates on the same principles as complex sounds. Peak correlation for a centered noise occurs along the center diagonal, decreasing monotonically away from the diagonal. No secondary peaks exist because there isn't a pseudoperiodic spatial fine structure. The correlation peak shifts when an ITD is introduced with the amount of shift increasing for higher CFs, as shown in Figure 6.

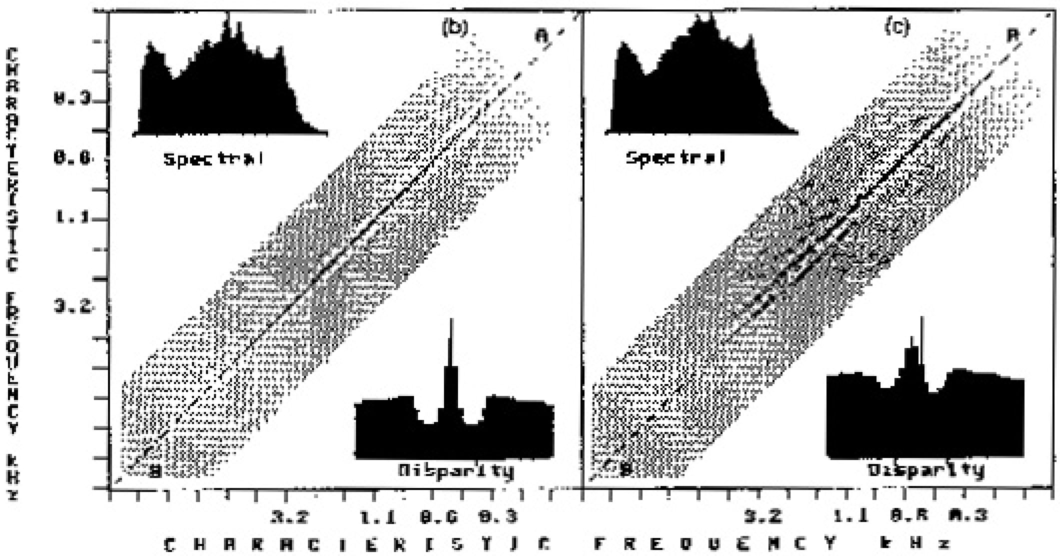

Figure 6. Lateralization of broadband noise (a) Network outputs with broadband noise centered (b) Same as (a) but with noise binaurally delayed by 250 μs

Detection and Enhancement of Tones in Noise

The stereausis network model is also able to enhance tone detection in the presence of competing noise. As shown in Figure 7, 1100 Hz centered tone in binaurally coherent noise has maximum response along the diagonal with a lot of noise around the peak. When either the tone or noise phase is reversed interaurally, there is improvement in signal detection since the phase reversal causes the output to shift off the noisy diagonal and stand as a separate component.

ACKNOWLEDGMENTS

This project would not have been possible without the help of the following individuals: Dr. Krishnaprasad, Tai Shi, Pamela White, Rebecca Copeland, and Diane Ihasz.

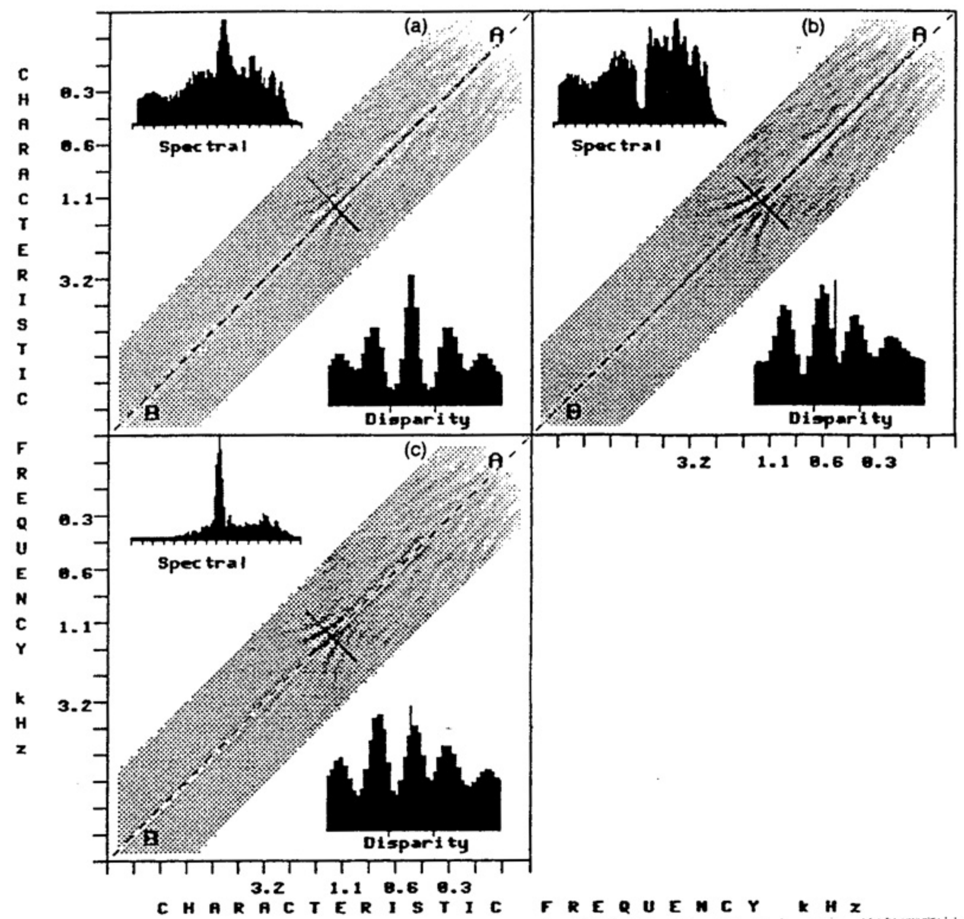

Figure 7. Detection and enhancement of tones in noise (a) Binaurally identical tone (1100 Hz) in noise background (0.1-10 kHz) and S/N = -10 dB (b) Same as (a) but with signal reversed in amplitude (c) Same as (a) but with noise reversed in amplitude

REFERENCES

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

Ru, Powen. (1997). Matlab files: loadload.m and stereaus.m