A New Model for Binaural Signal Detection

Jeroen Breebaart, Steven van de Par, and Armin Kohlrausch

IPO, Center for Research on User-System Interaction, P.O. Box 513, NL-5600 MB Eindhoven, The Netherlands

Abstract: Most present-day binaural models make use of coincidence detectors following an internal delay - so-called cross-correlation models. We present a new modeling approach based on a subtractive mechanism as found in excitatory-inhibitory type neurons in the auditory pathway. This approach transforms, for each auditory filter, time-domain waveforms into a time-varying two-dimensional activity pattern, where each unit within that pattern represents a characteristic interaural time delay and intensity difference. Limits of resolution are modeled by adding the same amount of internal noise to each unit. From this activity pattern, several features of the stimulus can be extracted, such as the apparent lateralization and the presence of a test signal within a masker. Depending on the condition, detection is limited by internal noise (e.g., NoSN) or by external fluctuations (e.g., NpSn, p < .95). This model can be used as an artificial observer in the same way as described by Dau et al. [J. Acoust. Soc. Am. 99, 3615-3622 (1996)] for monaural signals.

INTRODUCTION

In the past decades, several modeling efforts have been made to understand and quantitatively predict binaural masked thresholds. Also several schemes for the extraction of lateralization information have been proposed. Most models that have emerged rely on a cross-correlation approach (e.g., 2, 6, 7). However, the scope of these models separately is so far rather limited (cf. 1). Furthermore, the link between lateralization and detection for various combinations of spatial properties is still unclear. To resolve some of the limitations of the cross-correlation approach, we developed a model based on a different binaural interaction mechanism. A subclass of neurons in the auditory pathway is found to be excited by the signals from one ear and inhibited by the signals from the opposite ear (e.g., 5). The opposite influence of the two ears makes these cells sensitive to interaural differences. Based on this binaural interaction, our approach transforms, for each auditory filter, time-domain waveforms into an internal activity pattern from which several spatial features of the presented stimulus can be extracted.

MODEL DESCRIPTION

The model is divided into three stages. The first stage comprises the peripheral preprocessing of the outer, middle and inner ear. The combined outer and middle ear transfer function is modeled by a time-invariant bandpass filter with a roll-off of 6 dB/oct below 1 kHz and -6 dB/oct above 4 kHz. The spectral processing of the inner ear is simulated by a 2nd-order gammatone filterbank that correspond to the ERB (4) and a spectral spacing of 1 filter per ERB. The temporal processing properties are modeled by a half-wave rectifier, a 5th-order low-pass filter at 650 Hz (8) and a chain of 5 adaptation loops (3).

The second stage comprises interaction of the signals arriving at both ears. For each auditory filter (i), the difference (U_i) between the signals from both ears (L_i, R_i) is computed as a function of time (t), internal interaural delay (τ) and internal interaural gain factor (α):

\[ U_i(t, \alpha, \tau) = e^{\alpha L_i\left(t + \frac{\tau}{2}\right)} - e^{-\alpha R_i\left(t - \frac{\tau}{2}\right)} \]

(1)

For a range of delays and gain factors (τ,α), the above description results in a time-varying 2-D activity pattern which is calculated for each auditory filter. To each element in the 2-D pattern, an independent random variable is added to simulate limits of resolution. The task of the third stage is to extract spatial features form the activity patterns, such as the localization of a sound source or the presence of a binaural signal in a masker.
LATERALIZATION

When a diotic signal is presented to the model, for example a 100-Hz wide noise masker centered around 500 Hz, the activity pattern of an auditory filter centered on 500 Hz will take a form as depicted in the left panel of Fig. 1. Here, the activity $U_i$ is shown as a function of the internal variables $\alpha$ and $\tau$. For $\tau=0$, the internal representations of the signals arriving from both ears are exactly equal, hence resulting in a minimum of the internal activity. For other values of ($\tau, \alpha$), the internal representations will generally not be equal, resulting in an increase of the activity.

![Figure 1. Left panel: 2-D activity pattern for a diotic noise masker with a center frequency of 500 Hz as a function of the internal delay $\tau$ and the internal gain factor $\alpha$. Right panel: the activity pattern for NoS$\tau$ with S/N=-10 dB.](image)

If the signal is presented with an interaural time delay or interaural intensity difference, the point of minimal activity will shift across the pattern to a position at which the internal delay and intensity factor match the external spatial properties. Thus, by scanning the minimum in the activity pattern, the model can extract the externally presented IID and ITD. Hence information about the position of the sound source is present.

DETECTION

If an interaurally antiphase signal is added to a homophasic masker, the correlation between the signals arriving at both ears will differ from one. This results in an increase of activity at the minimum in the activity pattern. The activity pattern for an NoS$\tau$ condition for a signal-to-noise ratio of -10 dB is shown in the right panel of Fig. 1. A limit of resolution is obtained by adding an internal error variable to the activity pattern. This procedure enables to use the model as an ‘artificial observer’ using, e.g., a 3-interval, forced-choice procedure. In this procedure, the model’s task is to identify which interval contains the test signal. This is done in the following way. A template consisting of the mean internal representation of several masker-alone realizations is stored in memory. The computer model then determines which of the three intervals induces an internal representation which differs most from the template. In order to extract one decision variable from the internal representation, the differences that occur between the actual activity pattern and the template are weighted and summed across auditory filters. The weights are chosen in such a way that the weighting optimally reduces the internal error. The model then picks the interval which corresponds to the largest decision variable. Some examples of the predictive value of this model are given in Breebaart et al. (this issue).

REFERENCES