A Bayesian active-learning approach for obtaining notched-noise data

Josef Schlittenlacher[1], Thomas Baer[1], Richard E. Turner[2], Brian C. J. Moore[1]
[1] *University of Cambridge, Department of Psychology, Cambridge CB2 3EB, UK*
[2] *University of Cambridge, Department of Engineering, Cambridge CB2 1PZ, UK*
js2251, tb107, ret26, bcjm@cam.ac.uk

Introduction

Active-learning methods provide a promising way of running perceptual tests quickly while achieving high accuracy, at the cost of being computationally more expensive. Kontsevich and Tyler [1] used an active-learning approach to determine a psychometric function. For each trial, they chose the stimulus magnitude that was expected to be most informative about the mean and slope of the psychometric function. By using Gaussian Processes [2] and exploiting the commutativity of mutual information [3], active learning can be extended for use with more than two parameters, despite tough time constraints. In perceptual tests, the parameters for the next trial usually need to be chosen in an inter-trial interval of not more than 2 s.

A shorter test duration is particularly beneficial when determining auditory filter shapes using notched noise [4] in large scale research studies or clinical practice. Typically, several thresholds need to be obtained for each centre frequency in order to determine the filter parameters. Shen and Richards [5] suggested an active-learning test to overcome this problem. They directly learned the parameters of a simple model of the auditory-filter [6], assuming the filter to be symmetric. In the present work, the goal was to estimate the masker level at threshold as a function of the deviation of the lower edge of the notch from the signal frequency (Δf_l) and the upper edge of the notch from the signal frequency (Δf_u). This allowed the data to be fitted using an asymmetric model of the auditory filter.

Method

Ten normal-hearing subjects (18-44 yr, mean 27 yr, 5 female) were tested using their better ear. Each completed eight runs of a notched-noise active-learning test.

During a run, the pure-tone signal frequency (f_{sig}) and level (L_{sig}) were fixed, while the noise masker spectrum level (L_{mask}), Δf_l and Δf_u were varied. The signal parameters were f_{sig} = 250, 500, 1000, 2000 and 4000 Hz for L_{sig} = 40 dB SPL, and L_{sig} = 30 and 50 dB SPL for f_{sig} = 1000 Hz. The first and eighth runs used L_{sig} = 40 dB SPL and f_{sig} = 1000 Hz and the other conditions were run in random order. Each run consisted of 100 trials plus 20 catch trials with no signal. Twenty practice trials with L_{sig} = 40 dB SPL and f_{sig} = 1000 Hz were performed before the first run.

A trial contained three intervals. The first, second and third intervals contained, respectively, the signal alone, the noise alone, and the signal plus the noise. The task was to indicate whether the signal was audible in the third interval.

The signal consisted of three pulses, each with a duration of 150 ms, including raised-cosine rise and fall times of 20 ms. The pulses were separated by 100 ms. The noise had a duration of 850 ms, with rise and fall times of 20 ms. The signal started 100 ms after the start of the noise. The noise contained two bands, one above and one below L_{sig}. Each band had a width of $0.4f_{sig}$, except that the lower edge frequency was limited to 50 Hz. L_{mask} was limited to values between −20 and 55 dB/Hz. Δf_l and Δf_u varied from 0 to $0.8f_{sig}$ under the constraint $\Delta f_l + \Delta f_u \leq 1.2f_{sig}$.

For the first five trials the notch width was zero, and L_{mask} was increased to cover the range from the signal certainly being audible to certainly being inaudible. The next six trials used extreme values of Δf_l and Δf_u. After these, the parameters for the next trial were chosen to maximise the mutual information between the response and the parameters of the statistical model, as described in [3]. The underlying Gaussian Process [2] used to model the response probabilities as a function of the masker parameters had a constant mean that was optimized after each trial. Its covariance function was the sum of linear kernels in L_{mask}, Δf_l and Δf_u, and a squared-exponential kernel in Δf_l and Δf_u. The linear kernels captured the monotonicity in each dimension, and the additional squared-exponential kernel captured nonlinear perturbation effects. The likelihood function was a Gaussian cumulative density function, scaled to have a lapse rate of 0.02. This meant that the response probability as a function of L_{mask} resembled a typical psychometric function. The lapse rate allowed the model to recover from the effect of 'wrong' button presses.

Stimuli were presented through Sennheiser HD 580 headphones. Its transfer function, measured with a KEMAR, was taken into account when fitting auditory filters to the obtained threshold estimates.

Results

The values of L_{mask} for which the response probability was estimated to be 50% were taken as the thresholds. Auditory filters were fitted to the individual thresholds, and to the thresholds averaged across subjects. The model filter had a rounded-exponential shape [7] with a single exponential on the upper side (parameter p_u) and a sum of two exponential on the lower side to model the fiter tail (p_l for the main passband, t for the tail and w defining the transition between tail and the tip). Only p_l and p_u are discussed in this paper.

The symbols in Figure 1 show values of p_l and p_u for each condition, based on the average thresholds. The lines point to values predicted by a hearing model [8] for zero hearing loss. The actual p values were slightly smaller than predicted (except at 250 Hz), indicating somewhat broader auditory filters. All values corresponded to an outer hair cell loss less than 20 dB [8], i.e. within the normal range.

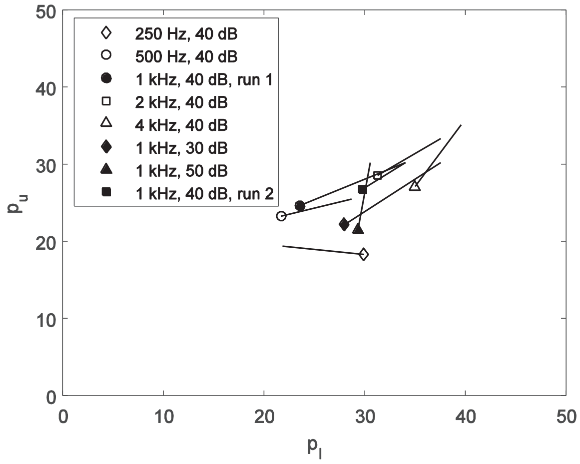

Figure 1: p_l and p_u based on average thresholds.

Figure 2 shows the parameter estimates for each subject for f_{sig} = 500 Hz (circles). The scatter is typical of that obtained for runs 1 to 7. The individual data were considerably more concentrated around the model estimate (triangle) for the last run. There was one outlier with p_l bigger than 50 for f_{sig} = 500 Hz, and another for a different subject for f_{sig} = 4000 Hz. All points were within or close to the area bounded by an outer hair cell loss of 20 dB (dotted lines). This was also typical of runs 1 to 7, except for f_{sig} = 250 Hz, for which the w and t estimates were probably less accurate. The goodness of fit of the parameter estimates, expressed as the root mean square difference between the thresholds predicted by the filter model and the obtained thresholds, was 3 to 4 dB for individual fits, and 1 to 2 dB for the mean data.

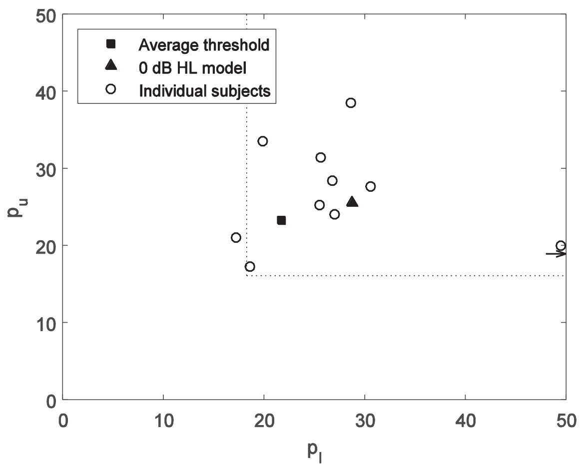

Figure 2: p_l and p_u for each subject for f_{sig} = 500 Hz (circles). The triangle shows the model prediction for zero hearing loss. The dotted lines enclose the area bounded by outer hair cell losses less than 20 dB. The arrow indicates that the value was bigger than 50.

The auditory filter becomes broader towards the low-frequency side with increasing level, and thus the ratio of p_l/p_u, representing the asymmetry, is expected to become smaller. The geometric means across subjects of p_l/p_u for f_{sig} = 1000 Hz were 1.32 for L_{sig} = 30 dB SPL, 1.05 for 40 dB SPL and 0.86 for 50 dB SPL (note, however, the difference to the fit to the average thresholds in Figure 1). A within-subject ANOVA on the logarithm of p_l/p_u confirmed the main effect of level, $F(2,18) = 9.2$, $p = 0.002$.

Discussion

Overall, the active-learning test yielded auditory-filter parameters for the average data across subjects that were close to expected values [4, 8]. The individual parameter values were also within or close to the expected range, with only two outliers out of 80 runs. However, a single run may not be accurate enough to quantify the asymmetry of the auditory filter. The scatter in Figure 2 is more likely to come from uncertainty in the filter estimates due to noisy responses than to true individual differences in the asymmetry of the auditory filter. The time taken for one run was about 10 to 15 minutes (120 trials). Extending this may not be practical for clinical applications. Solutions could be to use different parameters for the test, e.g. different ranges of Δf_l and Δf_u or other settings of the Gaussian Process, or a test that directly learns the parameters of the auditory filters, similar to [5]. Nonetheless the present method may be useful in large scale research studies, making it possible to sample informatively across multiple dimensions while keeping prior assumptions to a minimum.

Acknowledgments

The work was supported by the Engineering and Physical Sciences Research Council (UK, grant number RG78536).

References

[1] Kontsevich, L. L. and Tyler, C. W.: Bayesian adaptive estimation of psychometric slope and threshold, *Vision Research* 39 (1999), 2729-2737.

[2] Rasmussen, C. E. and Williams, C. K. Gaussian Processes for Machine Learning. Cambridge, MA, MIT press, 2006.

[3] Houlsby, N., Huszar, F., Ghahramani, Z., and Lengyel, M.: Bayesian active learning for classication and preference learning, *arXiv preprint* (2011).

[4] Glasberg, B. R. and Moore, B. C. J.: Derivation of auditory filter shapes from notched-noise data, *Hearing Research* 47 (1990), 103-138.

[5] Shen, Y. and Richards, V. M.: Bayesian adaptive estimation of the auditory filter, *Journal of the Acoustical Society of America* 134 (2013), 1134-1145.

[6] Patterson, R. D., Nimmo-Smith, I., Weber, D. L., and Milroy, R.: The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold, *Journal of the Acoustical Society of America* 72 (1982), 1788-1803.

[7] Glasberg, B. R., Moore, B. C. J., Patterson, R. D., and Nimmo-Smith, I.: Dynamic range and asymmetry of the auditory filter, *The Journal of the Acoustical Society of America* 76 (1984), 419-427.

[8] Moore, B. C. J. and Glasberg, B. R.: A revised model of loudness perception applied to cochlear hearing loss, *Hearing Research* 180 (2004), 70-88.