CONSIDERATIONS REGARDING ECHOLOCATION OF BATS

ANA-MARIA CHISELEV, LUMINITA MORARU

Dunarea de Jos University of Galati, Faculty of Science, 800201, Galati, Romania.

Abstract. The sonar system allows echolocating bats to navigate with high skill through a complex, threedimensional environment at high speed and low light. The auditory analysis of the echoes of their ultrasonic sounds requires a detailed comparison of the emission and echoes. This provides reasonably good predictions for both temporal and spectral behavioural sonar processing in terms of sonar delay, roughness, and phase sensitivity and in terms of sensitivity to the temporal separations in two-front targets and the classification of spectrally divergent phantom targets.

Keywords: echolocation, bats, impulse response.

1. INTRODUCTION

The auditory analysis of the echoes of self-generated sounds underlies the bats' extraordinary capability to manoeuvre quickly and at lowlight through their complex natural habitat. Bats shape the spectral-temporal structure of their ultrasonic emissions depending on the specific echo-acoustic task. Consequently, the auditory system must cope both with the variability of the emitted sounds and the variability of the echoes generated by these sounds. Some studies have shown that the auditory comparison of emission and echo appears to be implemented with such precision that only a fully coherent receiver, based on a cross-correlation of the call and echo waveform can explain sonar performance [1]. Both the spectrogram correlation and transformation (SCAT) model [2, 3] and the echolocation model of [4] focus on the time-domain reconstruction of the acoustic image of an object using a purely functional mathematic approach. Both call and echo, however, undergo complex, non-linear processing in the bat inner ear which renders a coherent receiver physiologically implausible.

A physiologically plausible model of bat sonar must thus start out with the neural representations of call and echo as they are represented after the basilar-membrane frequency-to-place transformation and the non-linear encoding of basilar-membrane motion by inner hair cells. Some current models of sonar neural processing are focus on the neural mechanisms extracting so-called glints from echoes [2, 3, 5]. A glint is defined as a conspicuously strong reflection of the call by a target. However, the echo-acoustic representation of a target is its impulse response (IR).

The impulse response of a target is defined as the sum of all reflections from a target when it is ensonified with a Dirac impulse. Complex natural targets produce many reflections with a large range of amplitudes and thus, the perceptually plausible definition of a glint is difficult. Specifically, compressive non-linear processing in the bat peripheral auditory system results in a distribution of reflection amplitudes which is quite different from the distribution of amplitudes in the IR waveform and thus, a reflection which appears relative prominent in the IR may be encoded in the peripheral auditory system at an only slightly higher level of excitation. Consequently, in the current paper, the definition and extraction of glints is avoided and instead, the auditory representation of call and echo is used to extract target information with the highest possible and plausible fidelity.

2. PERIPHERAL AUDITORY PROCESSING

The peripheral processing in the bats' outer, middle and inner ear has been described before [6]. Specifically, incoming sound is first convolved with the head-related transfer function [7] to imprint the directional information of echoes onto both the call and echo waveform. Second, a relatively broad band-pass filter is used to simulate the transfer characteristics of the bat middle ear. Inner ear, frequency-to-place conversion is implemented with a series of gammatone filters [8]. The spectral transmission characteristics of these filters have been adapted to fit the filtering characteristics of the *P. discolor* distortion product, otoacoustic emissions by [9]. The gammatone filters supplement these spectral characteristics with a gamma envelope which has been shown to be a reasonable approximation to the temporal filter characteristics of the basilar membrane. While already basilar membrane motion is non-linear through the effect of the outer hair cells, the inner hair cell mechano-electrical transduction is also highly non-linear with asymmetric compression best described with a Boltzman function. For simplicity, these non-linear transformations of the organ of corti are summarised by a simple half-wave rectification and exponential compression.

The temporal resolution of peripheral processing is critical for the design of a functionally plausible neural processing. Each filtering stage is accompanied by mandatory temporal integration and this integration increases with decreasing filter bandwidth. The strongest temporal integration is likely to be accompanied with the generation of the inner hair cell receptor potential [10, 11]. In mammals, this integration is characterised by a lowpass filter cut-off of 1–2 kHz and a filter slope of 12–24 dB per octave. Here, a 1 kHz low-pass filter with a slope of 12 dB per octave is used. The cumulative effect of the different peripheral–auditory temporal integration stages is illustrated in Fig. 1. Note that despite the relatively conservative choice of this low-pass filter, temporal integration in the current model is somewhat less than in the SCAT model which assumes a fixed, frequency-independent integration time of 350 μs. At this stage, a random noise with a fixed variance is added to the simulated cochlea output to limit overall encoding accuracy of the model.
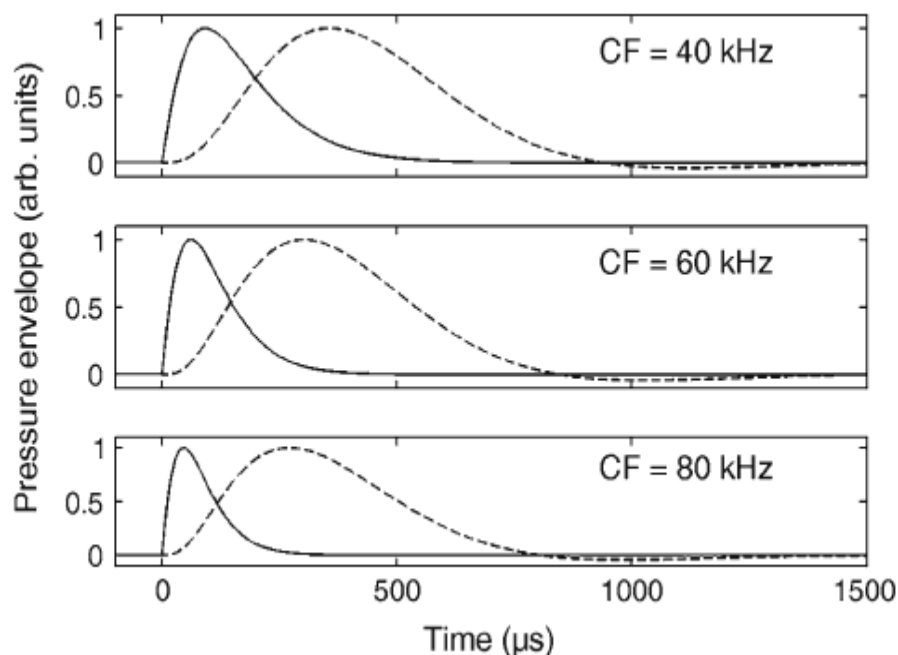

Fig. 1. Temporal integration in a peripheral auditory model of the bat, Phyllostomus discolor.

The data show the pressure envelope in response to an impulse at three different centre frequencies, 40, 60, and 80 kHz. The solid lines show the integration window for the inner ear bandpass filters; the dashed lines show the cumulative integration window for basilar-membrane and inner hair cell temporal integration. Note that the inner hair cell integration produces a stronger integration than basilar-membrane ringing.

3. NEURAL SONAR ANALYSIS

The current functional model of sonar analysis is inspired by human-auditory models of pitch extraction, and specifically, by the auditory-image model. In that model, the auditory spectrogram is subjected to strobed temporal integration. Here, the strobed temporal integration is replaced by a strobed, normalised autocorrelation.

The strobing requires a trigger signal. This trigger signal is derived from the auditory spectrogram by summing it across the tonotopic frequency axis. A trigger signal is generated when this summed activity exceeds a threshold. It is highly likely that the emission of the echolocation call, and not an echo, will issue a trigger. Here the trigger is extracted exclusively from auditory information. The neural activity following the trigger signal in an interval of about 20 ms is correlated with itself whereby negative correlation lags are ignored.

The normalisation ensures that, at a correlation lag of zero, the correlation in each channel is unity. The occurrence of echoes, i.e., of delayed reflections of the call, is encoded at the corresponding correlation lags. It is important to keep in mind that not call and echo itself but their temporal pattern of excitation in each frequency channel is correlated with itself. The data show that the normalised autocorrelation has two major effects: first, the frequency modulation of the call is not apparent in the autocorrelation output. Thus, the normalised autocorrelation can replace de-chirping mechanisms proposed by e.g. [5].

Second, due to the normalisation, also the spectral distribution of call energy is not preserved: even in frequency channels where call and echo have no acoustic power, the correlation coefficient at a correlation lag of 0 ms is unity. Instead, the estimate of echo delay appears to be defined better in frequency channels where the call was loud.

The autocorrelation of the frequency channel has a linear decay. In the 80 kHz channel, call and echo are loud and thus, the autocorrelation produces a good indicator of echo delay. There are spectral changes introduced through the head-related transfer function, changes introduced through atmospheric and geometric attenuation of the call travelling forth and back from the target, and changes introduced through spectral interference caused by the ensonified target, itself.

The final step of functional sonar processing is the preservation and integration of echo-acoustic information across a sequence of echolocation calls. To this end, the result of the channel-wise, normalised autocorrelation is stored in a sonar image buffer very similar to the auditory-image buffer of. In this image buffer, the autocorrelation output decays with a certain time constant which, as long as a psychophysical and electrophysiological quantification of this processing stage are missing, is set to 30 ms.

The system is retriggered with every sonar emission, and the autocorrelation output generated by subsequent trigger events is added to the current content of the image buffer. Thus, the auditory sonar representation of the ensonified target is stabilized, unlike in the auditory spectrogram, which is flowing in real time and is thus very difficult to assess. When an echolocating bat ensonifies a target from a fixed distance and angle, the sonar image buffer stabilizes the auditory sonar representation over time. Any changes in the geometric

arrangement, including distance, angle, and outer-ear movement, of course lead to a dynamic change of the auditory sonar representation in the image buffer.

Unless otherwise stated, the call used for the subsequent simulations is a synthesized, multi-harmonic echolocation call. Over duration of 1 ms, the fundamental frequency is modulated from 23 to 19 kHz. Five harmonics are generated with attenuations of 20, 15, 10, 0, and 5 dB, respectively. The call is ramped with 400 μs raised-cosine ramps. The input to the model always consists of the emitted call, convolved with a frontal head-related IR, the echo delay, and the echo. The latter consists of first the convolution of the call with the virtual-target IR and second, the convolution of the echo with the frontal head-related IR. Depending on the target-IR duration, zeros are appended to provide minimum sequence duration of 20 ms because this is the duration of the sequence subjected to the autocorrelation.

4. SPECTRAL MODEL PERFORMANCE

It is possible to investigate both spectral and temporal sonar encoding in the exact same model and decision device. Spectral interference in echoes occurs when the temporal separation of reflective surfaces of the ensonified target is shorter than the integration time of the inner-ear band-pass filters. In this case, the ringing in response to the first reflection interferes with the ringing in response to the second reflection and, dependent on the temporal separation and the band-pass centre frequency, this interference is constructive (resulting in a spectral peak) or destructive (resulting in a spectral notch). Based on the measured band width of the *P. discolor* inner-ear band-pass filters, these integration times vary in a frequency-dependent manner between about 100 and 200 μs. Note that, unlike in the SCAT model [2, 3] no fixed integration time is assumed.

The first spectral simulation is based on the capability of the echolocating bat, *Megaderma lyra*, to discriminate distances between two equally strong reflections of their own call [12]. The bats were trained to discriminate a two-front target with an internal temporal separation of 7.77 μs from two front targets with variable internal temporal separations between 8 and 26 μs. This temporal separation is encoded in the auditory frequency domain through a constructive and destructive interference pattern in the echo. In good agreement with the experimental data, the model threshold lies at an internal temporal separation between the two fronts of 8.3 μs. Thus, in the model, the temporal separations between the rewarded and unrewarded IRs was 0.6 μs compared to 1 μs measured experimentally.

The second spectrally challenging sonar simulation is based on the psychophysical phantom-target experiments of [13]. In this experiment, echolocating *M. lyra* were trained to discriminate two sets of IRs where one set had a spectral pass band around 41 kHz and the other set had a pass band around 82 kHz. Within the passband, the transmission across each set of 5 IRs was randomized to a limited extent. When the bats had learned this task, presentations of other IRs were randomly interspersed. These tests IRs were either spectrally expanded versions of one of the low-frequency IRs, or spectrally compressed versions of one of the high-frequency IRs. The bats' spontaneous association of these tests IRs to one of the trained groups was assessed.

For simulation purposes, image buffers were created for each of the trained and test IRs. The image buffers were averaged across each set of low or high frequency IRs to create a template as it may be used by the bats to evaluate unknown IRs.

The image buffer in response to each test IR was compared in terms of Euclidean distance to both the low- and the high-frequency template. As in [13], the classification in

'percent judged low' was quantified as the Euclidean distance to the high template divided by the sum of the distances to the low and high template times 100. This decision device is different from that used for the previous experiments because here IR classification was assessed, not a threshold for an IR change. The decision device for the current simulation is identical to that used in [13].

It must be pointed out that in several of the simulations shown above, although thresholds predicted from the model are similar to the experimental data, often the psychometric functions, and specifically their slopes around threshold, are not. Ideally, a model could predict the whole psychometric function and not just the 75% correct value. It is possible that further changes in the amount of internal noise and the soft thersholding in the decision device may yield even more plausible predictions of the experimental psychometric functions.

5. CONCLUSIONS

In summary, this paper presents a model of how echolocating bats may analysis the echoes of their sonar emissions to obtain information about their distance to an object and the three-dimensional shape of the object. We can predict the results of a variety of formal psychophysical experiments on bat sonar. It may serve as a roadmap to identify critical echoacoustic processing steps in the bat auditory brainstem and midbrain.

REFERENCES

[1]   Simmons, J.A., Ferragamo, M., Moss, C.F., Stevenson, S.B., Alters, R.A., *Journal of Comparative Physiology A*, **167**, 589, 1990.

[2]   Saillant, P.A., Simmons, J.A., Dear, S.P., McMullen, T.A., *Journal of the Acoustical Society of America*, **94**, 2691, 1993.

[3]   Peremans, H., Hallam, J., *Journal of the Acoustical Society of America,* **104**, 1101, 1998.

[4]   Matsuo, I., Kunugiyama, K., Yano, M., *Journal of the Acoustical Society of America,* **115**, 920, 2004.

[5]   Boonman, A., Ostwald, J., *Biological Cybernetics*, **97**, 159, 2007.

[6]   Grunwald, J.E., Schörnich, S., Wiegrebe, L., *The Proceedings of the National Academy of Sciences of United States of America*, **101**, 5670, 2004.

[7]   Firzlaff, U., Schuller, G, *Hearing Research*, **181**, 27, 2003.

[8]   Patterson, R.D., *Journal of the Acoustical Society of America,* **96**, 1409, 1994.

[9]   Wittekindt, A., Drexl, M., Koessl, M., *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, **191**, 31, 2005.

[10] Palmer, A.R., Russell, I.J., *Hearing Research*, **24**, 1, 1986.

[11] Chiselev, A.M., Moraru, L., Gogu, A., *Romanian Journal of Biophysics*, **19**(4), 251, 2009.

[12] Schmidt, S., *Journal of the Acoustical Society of America,* **91**, 2203, 1992.

[13] Weißenbacher, P., Wiegrebe, L., *Behavioral Neuroscience*, **117**, 833, 2003.