Neuroscience and Biobehavioral Reviews 35 (2011) 2064–2072

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

journal homepage: www.elsevier.com/locate/neubiorev

Review

Intensity-invariant coding in the auditory system

Dennis L. Barbour*

Laboratory of Sensory Neuroscience and Neuroengineering, Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, MO 63130, USA

ARTICLE INFO

Article history:
Received 21 September 2010
Received in revised form 9 April 2011
Accepted 11 April 2011

Keywords:
Efficient coding
Robust sensory processing
Perception
Intensity invariance
Auditory system
Sound processing
Auditory cortex
Adaptation
Forward masking

ABSTRACT

The auditory system faithfully represents sufficient details from sound sources such that downstream cognitive processes are capable of acting upon this information effectively even in the face of signal uncertainty, degradation or interference. This robust sound source representation leads to an invariance in perception vital for animals to interact effectively with their environment. Due to unique nonlinearities in the cochlea, sound representations early in the auditory system exhibit a large amount of variability as a function of stimulus intensity. In other words, changes in stimulus intensity, such as for sound sources at differing distances, create a unique challenge for the auditory system to encode sounds invariantly across the intensity dimension. This challenge and some strategies available to sensory systems to eliminate intensity as an encoding variable are discussed, with a special emphasis upon sound encoding.

Contents

1. Introduction

An important feature of biological sensory systems is their ability to extract meaningful environmental signals under a wide variety of conditions. This ability is necessary for successful prey acquisition, predator avoidance and mate localization, among other crucial behavioral tasks. These biological systems perform remarkably well even in the face of incomplete information, signal degradation or competing signals. Systems capable of extracting relevant information consistently under extremely variable environmental conditions are termed robust, and robust sensory pattern recognition is extremely useful for the survival of many species, including humans.

One stimulus dimension across which sounds are relatively consistently perceived is intensity or sound level. In other words, as the total power of a target sound is varied over many orders of magnitude, listeners are able either to correctly identify it or to process it correctly relative to other similar sounds (Buus and Florentine, 1991; Hanna et al., 1986; Viemeister and Bacon, 1988). This type of intensity invariance typically represents a straightforward achievement for artificial pattern recognizers because the overall stimulus pattern (e.g., the spectrotemporal distribution of sound power) remains relatively constant as overall power is added to the signal in a linear medium such as air. A simple signal normalization in the pattern recognizer can therefore enable robust identification of the signal relative to changes in its intensity. Intrinsic nonlinearities in biological sensory systems, on the other hand, often mean that even an operation as simple as adding power to a signal could alter the neural representation of that signal. These nonlinearities are partic-

---

* Correspondence address: Department of Biomedical Engineering, Washington University, One Brookings Dr., Campus Box 1097, Uncas Whitaker Hall Room 200E, St. Louis, MO 63130, USA. Tel.: +1 314 935 7548; fax: +1 314 935 7448.

E-mail address: dbarbour@biomed.wustl.edu

0149-7634/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neubiorev.2011.04.009

ularly strong in the auditory periphery, leaving open the question of how the central auditory system is able to create a consistent perception of a given sound as it changes in intensity.

Under environmental conditions, differences in sound intensity are often associated with differences in sound source distances. In fact, overall sound source intensity is one of the key stimulus features used to estimate target distance in adult humans (Ashmead et al., 1990; Litovsky and Clifton, 1992; Mershon and Bowers, 1979; Strybel and Perrott, 1984; Zahorik and Kelly, 2007). Manipulation of sound intensity (and therefore perceived loudness) leads to systematic errors in distance judgment for virtual sound sources (Mershon et al., 1981). Loudness perception itself has received considerable attention (Glasberg and Moore, 2006; Plack and Carlyon, 1995; Zhang and Zeng, 1997), while leaving relatively unexplored the mechanism of perceptual invariance across intensity in the auditory system. The goal of this review, therefore, will be to focus upon some of the potential strategies available to the nervous system for encoding sensory signals over a wide dynamic range while still preserving a representation of the signal that can be exploited for invariant or nearly invariant perception of the corresponding object. Strategies that appear to be used by the auditory system will be emphasized.

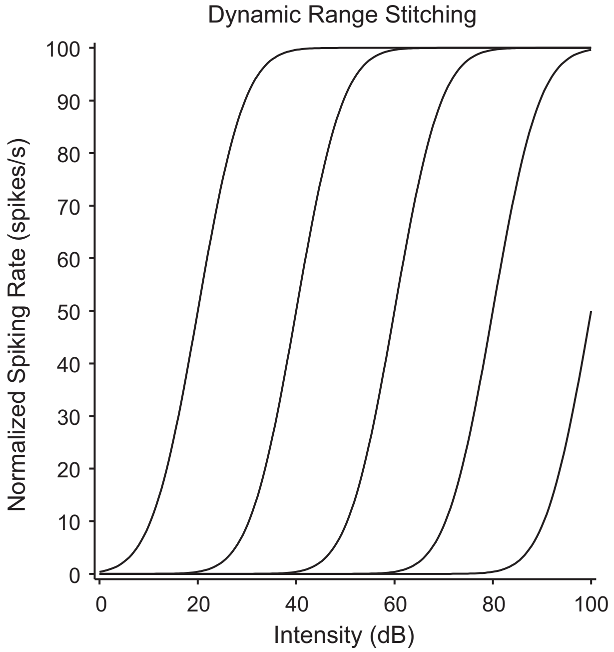

Fig. 1. Dynamic range stitching. One strategy to cover a wide overall intensity with a sensory system is to stitch individual coding elements together such that their more limited individual dynamic ranges combine to cover the total range of interest. This particular collection of 5 sigmoidal input/output functions is capable of collectively and equivalently encoding intensities from near 0 dB up to 100 dB, although each individual function would only be able to encode a relatively narrow range of intensities.

2. Dynamic range stitching

No individual coding element (e.g., receptor or neuron) in the sensory systems of higher animals is capable of encoding the entire intensity range to which the organism is sensitive. Perhaps the most obvious means of building an invariant representation across a wide intensity range using discrete elements of much narrower intensity ranges is to construct these elements such that their individual input/output functions combine to collectively span the total range of interest. In such a case, one would expect a range of neuronal thresholds such that combining or “stitching” together individual neuronal responses would allow the sensory system to represent the full range of intensity normally available to the organism. To some degree this strategy appears to be used by the auditory system and is depicted graphically in Fig. 1. Threshold measurements made in auditory nerve support the notion that individual neuronal dynamic ranges are dispersed somewhat across the total intensity range of hearing (Evans, 1972; Liberman, 1978; Liberman and Kiang, 1978; Sachs and Abbas, 1974). Thresholds of auditory nerve fibers have classically been evaluated as absolute spiking rate measures evoked by stimuli versus spontaneous rates, but similar trends hold true when statistical properties of rate responses are taken into account (Geisler et al., 1985; Young and Barta, 1986) and are logically extended when temporal information in the spike trains is considered (Carney, 1994; Colburn et al., 2003).

The strategy of dynamic range stitching can be also seen in neuronal responses from primary auditory cortex (A1). Fig. 2 depicts the distribution of relative tone thresholds measured from over 500 neurons located in awake marmoset monkey A1. The pattern that emerges reflects essentially the same pattern seen in the auditory nerve, whereby a subset of thresholds is indeed distributed across a wide intensity range, but the bulk of thresholds trend toward lower values. Because the average dynamic range of these neurons is around 15 dB, the overall machinery of auditory encoding appears to be biased largely toward intensities within 30 dB or so of hearing threshold (Watkins and Barbour, 2010b). This finding is problematic for two reasons. First, for the “stitching” process to be most effective, the neuronal dynamic ranges would be expected to span more or less uniformly the complete intensity range of hearing, which is not the case here. In fact, *prima facie* evidence for the auditory system’s ability to encode loud or moderately loud sounds is surprisingly lacking from these data. The issue

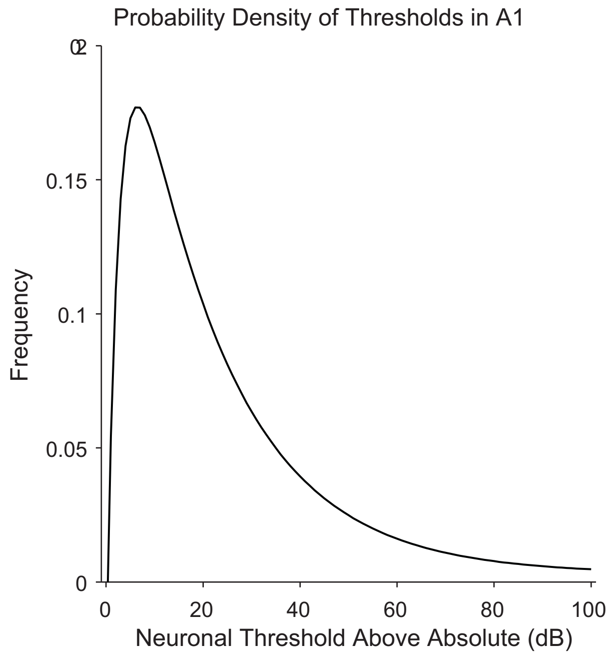

Fig. 2. Probability density function of thresholds measured from 544 neurons in awake marmoset primary auditory cortex. These measures demonstrate some neuronal coverage over a wide intensity range but a heavy skew toward the lowest intensities. The relative thresholds plotted are computed by subtracting out absolute thresholds of hearing at each frequency as determined by the collective neuronal responses. The curve plotted is the best difference of single exponentials that fits the actual data. Details of this data set can be found in (Watkins and Barbour, 2010b).