Chapter 6

Discussion

The primary goal of this thesis was to identify computational principles that govern information processing and representation in the auditory system. For this purpose, we set to develop formal and quantitative methods that could identify such principles in a set of electrophysiological recordings from three brain regions.

We started the investigation by discussing methods to extract information from spike trains and reviewed in detail both the theoretical motivation and the practical considerations of reducing the dimensionality of spike trains and representing them by simpler statistics. We tested six different dimensionality reduction methods and compared the level of information they achieve. In all of these methods, IC neurons conveyed about twice more information about the identity of the presented stimulus than AI and MGB neurons.

We found that the maximal information can almost always be extracted by considering the distribution of temporal patterns of spikes. Surprisingly, the first spike latency carries almost the same level of information. In contrast, spike counts convey only half of the maximal information level. These results reveal a surprising observation regarding the nature of the neural code in the different stations. It shows that even though spikes in the cortex are not precisely locked in time to the stimulus (as in the IC), the occurrence of a single spike, the first spike since the onset, conveys the maximal information that could be achieved with any other methods tested.

We then advanced to investigate how small groups of neurons in different brain areas interact to represent the auditory stimuli. For this purpose we developed measures of informational redundancy in groups of cells, and described their properties. These measures can be reliably estimated in practice from empirical data using stimulus conditioned independence approximation. Since redundancy is biased by the baseline single-unit information level, we study this effect and show how it can be reduced with a proper normalization. Finally, we discussed redundancy biases due to a ceiling effect on the maximal information and the way to correct for these biases.

Applying these methods to our data we obtained three main results. First, we showed that small groups of IC cells are more redundant in the information they convey about stimulus identity than AI and MGB cells. In other words, cells of higher regions in the processing hierarchy tend to code features of the stimuli in a more independent manner. These findings put forward redundancy reduction as a possible generic organization principle of sensory systems. This principle was suggested 40 years ago by Barlow from theoretical considerations, and an empirical evidence for such a process is presented here for the first time.

Secondly, we showed that redundancy is significantly correlated with the best frequency (BF) of IC cells but not with the those of AI or MGB cells. This means that frequency characterization does not capture well the type of processing that AI and MGB cells perform. Finally, we found that AI and MGB cells convey an order of magnitude less information about the spectro-temporal structure of the stimuli as compared to IC neurons. This suggests that AI cells succeed to code well the identity of the stimuli without coding the precise acoustical structures in it.

The low redundancy in AI and MGB, and the lack of correlation of this redundancy with the best frequency of the cells has strong implications, which go far beyond the statement that BF responses are not predictive for complex sounds. The reason is that currently accepted methods of characterizing AI neurons, primarily the spectro temporal receptive field (STRF), imply redundancy between neurons that share spectro temporal characteristics. Although we find such redundant neurons in IC we do not find them in MGB or AI, showing that STRF characterization in AI misses crucial aspects of neuronal coding even for simple stimuli as used in our work.

Coding in an independent manner relates to the issue of *specialization* of neuronal responses. The standard meaning of specialization in the literature is that neurons respond only to a restricted set of stimuli. In this sense, for example, face neurons in the infero temporal cortex are specialized. In the context of primary auditory cortex, as discussed in [Nelken, 2002], most of the evidence today does not support the view that neurons in AI are more specialized than neurons in lower auditory stations. For example,[Middlebrooks *et al.*, 1994] showed that neurons in AI do not have specialized receptive fields for space,[Kowalski *et al.*, 1996a, Kowalski *et al.*, 1996b, Depireux *et al.*, 2001] used analysis of STRF in the ferret auditory cortex and demonstrated a rather simple structure in most of them (quadrant separability). The type of results quantified here suggests a more delicate type of specialization, since instead of mapping the parametric set of stimuli to which neurons respond, we measure the ability of neurons to discriminate between stimuli that are acoustically highly similar (see figure 1.9 in section 1.3.5). This type of informational specialization is manifested in the form of low redundancy. The fact that we do not find cortical neurons that are as redundant as IC neurons (even though they respond to stimuli and convey considerable amount of information about them) suggests that the informational redundancy measure is the more interesting measure of specialization.

The term *redundancy reduction* was originally coined by Barlow[Barlow, 1959b, Barlow, 1959a, Barlow, 1961]. As explained in Chapter 1, he suggested that a principal goal of sensory processing is to achieve an efficient code by compressing sensory inputs to obtain parsimonious representations. Barlow later suggested [Barlow, 2001] that the actual goal of the system is rather *redundancy exploitation*: During this process the statistical structures in the inputs are extracted and coded. This revised idea leads to inherently different consequences, and predicts that higher areas will contain many units, specialized to respond to complex structures of the stimuli, and largely independent. The results presented in our study are in agreement with this revised view. We find that neurons in higher processing stations are less informative but more independent than those at the lower levels, presumably because they convey information about more complex structures in the inputs.

Reducing redundancy during information processing while mapping stimuli to a higher dimensional feature space may provide better discrimination among complex stimuli, as in independent component analysis (ICA) [Bell and Sejnowski, 1995] and support vector machines[Vapnik, 1995]. Redundancy reduction may therefore be a generic organizational principle of sensory systems that allows for easier readout of stimulus aspects that are behaviorally relevant.

The last part of this dissertation addresses the problem of identifying the features that are relevant for neural responses. A major difficulty in this task is that the characterization of the processing performed in a single brain region requires filtering out the processing that takes place earlier in the processing hierarchy. This is an instance of a generic problem in unsupervised learning, of identifying relevant structures in data that contain many competing structures. We presented a formal definition of the problem in the framework of distributional clustering, as well as its analytical and algorithmic solutions. We showed its applicability in a variety of data domains as texts clustering and feature extraction for face recognition. Our results lay the groundwork for developing additional dimensionality reduction methods of data that use irrelevance data, and have already been extended to various problems and learning techniques such as the study of neural coding in evolving agents [Avraham *et al.*, 2003], data mining of the web [Gondek and Hofmann, 2003] linear projections of Gaussian variables and their relation to canonical correlation analysis [Chechik *et al.*, 2003, Chechik and Globerson, 2003], spectral based feature selection [Shashua and Wolf, 2003] and continuous embedding of categorical variables [Globerson *et al.*, 2003], as applied for example to feature extraction for face recognition.

Bibliography

[Avraham *et al.*, 2003] H. Avraham, G. Chechik, and E. Ruppin. Are there representations in embodied evolved agents? taking measures. In *Proceedings of the 7th European Conference on Artificial Life*, UK, 2003.

[Barlow, 1959a] H.B. Barlow. The coding of sensory messages, chapter XIII. In Thorpe and Zangwill, editors, *Current Problems in Animal Behavior*, pages 330–360. Cambridge University Press, 1959.

[Barlow, 1959b] H.B. Barlow. Sensory mechanisms, the reduction of redundancy, and intelligence. In *Mechanisation of thought processes*, pages 535–539. Her Majesty's stationary office, London, 1959.

[Barlow, 1961] H.B. Barlow. Possible principles underlying the transformation of sensory messages. In W.A. Rosenblith, editor, *Sensory communication*. MIT Press, Cambridge, MA, 1961.

[Barlow, 2001] H.B. Barlow. Redundancy reduction revisited. *Network: Computation in Neural Systems*, 12:241–253, 2001.

[Bell and Sejnowski, 1995] A.J. Bell and T.J. Sejnowski. An information maximization approach to blind seperation and blind deconvolution. *Neural Computation*, 7:1129–1159, 1995.

[Chechik and Globerson, 2003] G. Chechik and A. Globerson. Information bottleneck and linear projections of Gaussian processes. Technical Report 4, Leibniz Center for Research, School of Engineering and Computer Science, The Hebrew University of Jerusalem, May 2003.

[Chechik *et al.*, 2003] G. Chechik, A. Globerson, Naftali Tishby, and Yair Weiss. Information bottleneck for Gaussian variables. Submitted, 2003.

[Depireux *et al.*, 2001] D.A. Depireux, J.Z. Simon, D.J. Klein, and S.A. Shamma. Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex. *The Journal of Neurophysiology*, 85(3):1220–1234, 2001.

[Globerson *et al.*, 2003] A. Globerson, G. Chechik, and N. Tishby. Sufficient dimensionality reduction with irrelevant statistics. In *Uncertainty in Artificial Intelligence: Proceedings of the Nineteenth Conference (UAI-2003)*, San Francisco, CA, 2003. Morgan Kaufmann.

[Gondek and Hofmann, 2003] D. Gondek and T. Hofmann. Conditional information bottleneck clustering. In *3rd IEEE International Conference on Data Mining, Workshop on Clustering Large Data Sets*, 2003.

[Kowalski *et al.*, 1996a] N. Kowalski, D.A. Depireux, and S.A. Shamma. Analysis of dynamic spectra in ferret primary auditory cortex I. Characteristics of single unit responses to moving ripple spectra. *The Journal of Neurophysiology*, 76:3503–3523, 1996.

[Kowalski *et al.*, 1996b] N. Kowalski, D.A. Depireux, and S.A. Shamma. Analysis of dynamic spectra in ferret primary auditory cortex II. Prediction of unit responses to arbitrary spectra. *The Journal of Neurophysiology*, 76:3524–3534, 1996.

[Middlebrooks *et al.*, 1994] J.C. Middlebrooks, A.E. Clock, L. Xu, and D.M.A Green. A panoramic code for sound location by cortical neurons. *Science*, 264(5160):842–844, 1994.

[Nelken, 2002] I. Nelken. Feature detection by the auditory cortex. In Donata Oertel, Richard R. Fay, and A. N. Popper, editors, *Integrative Functions in the Mammalian Auditory Pathway*, volume 15 of *Handbook in Auditory Research*. Springer-Verlag, 2002.

[Shashua and Wolf, 2003] A. Shashua and L. Wolf. Sparse spectral-based feature selection with side information. Technical Report 57, Leibniz

Center for Research, School of Engineering and Computer Science, The Hebrew University of Jerusalem, June 2003.

[Vapnik, 1995] V.N. Vapnik. *The nature of statistical learning theory.* Springer-Verlag, Berlin, 1995.